DISSERTATION

TRENDS AND TREE-RINGS: AN INVESTIGATION OF THE HISTORICAL AND PALEO PROXY HYDROCLIMATE RECORD OF THE KHANGAI MOUNTAIN REGION OF MONGOLIA

Submitted by

Niah B. H. Venable

Department of Geosciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2016

Doctoral Committee:

Advisor: Steven R. Fassnacht

Melinda J. Laituri
William E. Sanford
Peter M. Brown

ABSTRACT

TRENDS AND TREE-RINGS: AN INVESTIGATION OF THE HISTORICAL AND PALEO PROXY HYDROCLIMATE RECORD OF THE KHANGAI MOUNTAIN REGION OF MONGOLIA

The Khangai Mountain region of western central Mongolia is a diverse area of mountain, forest, steppe, and desert steppe landscapes reaching across and beyond the mountains. The tradition of nomadic pastoralism is strong in the region, with water for domestic and livestock needs supplied through lakes, springs, rivers, and wells. Herders of the region have felt impacts from the climatic extremes of the last few decades in terms of increasing temperatures and decreasing water supplies.

The main objective of this dissertation is to quantify the changing climate of Mongolia through analysis of key hydrometeorological variables over space and through time. The assessments of trends in the data and the paleo proxy analyses herein address interdisciplinary research questions using multidisciplinary approaches. In closing, this work also examines how the data and analyses presented are used as objects that cross disciplinary boundaries, and can facilitate communication and collaboration between different groups.

To provide context for this work, a countrywide view of changing maximum temperature, minimum temperature, and precipitation are examined using trend analyses of gridded datasets. Both minimum and maximum temperatures are significantly warming across the country. Significant decreases in precipitation are concentrated in the central and eastern parts of the country for the 50-year period of analysis.

Local knowledge of hydroclimatic change provides another source of climatic information with herders of the Khangai Mountain region observing temperature increases, though the exact time period over which change has occurred varies depending upon memory. Therefore, temperature data were analyzed from five meteorological stations with varying lengths of record from 15 to 50 years and varying start periods based on the available length of record. The most highly significant changes occurred for the longest time periods and for annual average minimum temperatures.

Issues of data availability, serial correlation, and homogeneity of climate records were explored using the Mann-Kendall test for trend significance and the Thiel-Sen method for determining trend slope or magnitude in precipitation and streamflow records. An additional step of prewhitening the data prior to testing was used to reduce the influence of autocorrelation on results. Homogeneity testing was also performed. Decreasing trends in annual, spring, and summer precipitation and/or streamflow were found at several Mongolian stations, particularly on the northern side of the mountains, with increasing winter precipitation trends at one site. Results were compared to analyses using Colorado data. Degradation of the Colorado hydroclimate records by shortening the time series and introducing gaps to simulate inconsistencies found in Mongolian datasets created significant trends where none previously existed.

Tree-ring reconstructions of Mongolian hydroclimate variables have provided insight on multidecadal and muticentennial trends in climate variability over many other parts of the country, but that work has not been extended to contextualize the recent sharply decreasing streamflows of the Khangai Mountain region. Cores from two new sites collected in the summer of 2012 and records from eight other moisture-sensitive sites in the region were used to reconstruct streamflow for four gages. Missing streamflow data were filled by multiple imputation/predictive mean matching methods with data from six nearby meteorological stations prior to use in multiple linear regression models developed for the reconstructions. A quantitative evaluation of reconstructed and historical extremes of wet and dry conditions in each basin and qualitative analyses of event synchrony are discussed. The drought events of the last decade and a half, while extreme are not beyond the range of natural variability found over the last 300+ years in the four Khangai Mountain region rivers and could be considered plausible flow conditions for the future, particularly under a warming and possibly drying climate.

Finally, this dissertation explores cross-boundary connections within each previous chapter and contributions of this work to selected goals of the Mongolian Rangelands and Resilience (MOR2) project, an interdisciplinary and cross-cultural collaboration investigating the resilience of Mongolian pastoral systems to climate change. Changes to the livelihoods of traditional nomadic pastoralists of Mongolia are not only attributable to climate, but also represent changes to socio-ecological, economic, and governmental/policy systems. The analyses of observational gridded, station-based, and paleo proxy data in this dissertation provide a quantitative foundation for continued investigations of the physical hydroclimate systems of the region and further themes developed in previous research from across Asia and within Mongolia. The results of this work will prove useful as a foundation for the development of water policy and infrastructure ideally favoring sustainable nomadic pastoral use of the region's finite water resources under a changing climate.

ACKNOWLEDGEMENTS

"Sometimes people come into your life for a moment, a day or a lifetime. It matters not the time they spent with you, but how they impacted your life at that time…"

- Anonymous

A friend that I only knew for a short time gave me this quote many years ago. It reminds me that regardless of the amount of time spent individually with mentors, colleagues, friends, and family while earning my doctoral degree at Colorado State University (CSU), I am grateful for the positive impacts that they have had on me, and the support they have given me during this challenging and wonderful period. Therefore, my gratitude first extends to the funding agencies that supported my work. While I may never know all the people working for and with these agencies, I believe they are largely composed of caring individuals committed to supporting quality student research. The National Science Foundation Dynamics of Coupled Natural and Human Systems Program contributed financially to much of my work. I received additional funding from the Colorado Water Institute through the National Institutes for Water Research, and the American Center for Mongolian Studies US-Mongolia Field Research Fellowship Program. I also benefitted from several CSU scholarships, fellowships, and travel grants. The completion of my degree was made possible by funding from the estate of Dr. William R. Venable, may he rest in peace.

Dr. Steven Fassnacht, my advisor, travel companion, and mentor has had a strong influence on my development as a researcher and scholar. He has been flexible when I needed it and has wholeheartedly supported my efforts to become a broad-based intellectual while requiring me to be scientifically rigorous. I would like to thank my committee members: Dr. Melinda Laituri, who enabled me to examine my work and the larger graduate school experience philosophically, empowering me to grow as a scholar and scientist. Dr. Peter Brown was instrumental in helping me realize my goals of applying the concepts and methods of dendrochronology to my dissertation research. Dr. Bill Sanford provided timely comments and insight, always seeming to ask the questions I was least prepared to answer in meetings and presentations, reminding me to think more deeply and present my results more concretely than I would by instinct alone.

Other colleagues and friends at CSU and elsewhere that deserve special thanks include: Dr. Maria Fernández-Giménez, the principal investigator on the Mongolian Rangelands and Resilience (MOR2) project that I came to regard as more than just the talented and demanding project leader but also as an insightful and inspiring colleague and friend. Dr. Arren Mendezona Allegretti (Class of 2016!) whose involvement in the MOR2 project, interests, expertise, and friendship provided a refreshing counterpoint to my studies and led to many long conversations about interdisciplinarity, teamwork, and the inherent struggles of trying to "walk the walk while talking the talk". Dr Jay Angerer, a colleague on the MOR2 project that in addition to his help on various data issues provided great conversation in his Southern accent that always reminded me of my Oklahoma and Texas roots. I am grateful to the Mongolian friends and colleagues I met and worked with during my time on the MOR2 project especially, Odgarav Jigjsuren, Dr. Khishigbayar Jamiyansharav, Dr. Tungalag Ulambayar, Dr. Chantsallkham Jamsranjav, Dr. Batkhishig Baival, Dr. Batbuyan Batjav, and Tumenjargal Sukh. Getting to know them, their families, and the beautiful country of Mongolia opened my eyes to the wonders and benefits of world travel and cultural experiences beyond my own. Similarly, I would like to thank Dr. Juan Ignacio López-Moreno for his insight and assistance with the climate data and facilitating my explorations of his home country of Spain. I also want to thank the two undergraduate students I mentored as part of the Skills for Undergraduate Participation in Ecological Research program, Katie Fitzsimmons and Marshall Wolf. It was a great experience working with them and using part of my dissertation research as a foundation for their investigations. Marshall in particular, was incredibly helpful by volunteering long hours at Dr. Jason Sibold's Biogeography Lab, assisting with the crossdating and measurement of my tree cores. I extend my thanks as well to faculty and staff of the CSU Geosciences Department, in particular Sharon Gale, whose aid in administrative matters proved invaluable on more than one occasion.

I would also like to thank my family. My parents for their encouragement through the years, enabling me to become the capable person I am, and my dear husband and children who have supported me with love, patience, and understanding throughout the "agonies and ecstasies" of this academic endeavor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1- INTRODUCTION

Climate change undermines our understanding of past water distribution, availability, and quality, affecting the technological, economical, political, and cultural choices we make about future water development, infrastructure, and policy (Gleick, 1989; Gleick, 2010). These choices must be informed by unbiased analysis of data collected from many scientific, technical, and socio-economic sources at varying spatial scales (IPCC, 2014a). The results of global climate change studies at coarse resolutions are popularly cited, but the effects of change are most felt on smaller spatial scales. On countrywide, regional, and/or local levels are where extremes of heat, cold, floods, and droughts have the biggest impact but these are also scales where climate change effects are difficult to separate from natural variability (Hoerling *et al*., 2010; Mote *et al*., 2011).

Each chapter in this work quantifies the changing climate of Mongolia over the last 50 years through analysis of key hydrometeorological variables over space and time. Most of the focus is placed on the Khangai Mountain region, a topographically diverse and ecologically important part of west central Mongolia (Figure 1.1). This region is distant enough from population centers in the north central part of the country to have had less urban and resource development than those areas. It is close enough however, that in the near future the limited natural resources of this important headwaters region will likely be placed under increasing development pressure. Impacts such as expanded mining (e.g., Regdel *et al*., 2012), irrigated agriculture (e.g., Pederson *et al*., 2013; Priess *et al*., 2015), and urbanization (e.g., Fan *et al*., 2016), when combined with the climatic extremes of the last decade and a half, will have unprecedented socio-ecological effects on the traditional herding livelihoods maintained in the Khangai Mountain region.

For initial spatial context, Chapter 2 takes a broad view of changing temperature and precipitation trends over the last 50 years for the entire country, using spatio-temporal analyses of coarse-scale gridded climate data products. The following third and fourth chapters of this work document trends in temperature, precipitation, and streamflow in the Khangai Mountain region, examining the length of record and investigating questions of data quality. In addition to these analyses using historical station-based data, Chapter 5 examines regional patterns of streamflow variability in the long-term context of the last several hundred years through the analysis of tree-ring paleo proxies. The sixth and final chapter examines how the data and analyses presented in the previous chapters are used interdisciplinarily. It explores cross-boundary connections within each chapter and contributions of this work to selected goals of the Mongolian Rangelands and Resilience (MOR2) project, an interdisciplinary and cross-cultural collaboration investigating the resilience of Mongolian pastoral systems to climate change.

This research does not explicitly identify the drivers of change or attribute them to strictly anthropogenic or natural causes. It does however support a critical need to recognize which variables exhibit the greatest amount of change as this gives credence to scenarios of change impacts. For example, sharply increased warming with no difference or even modest increases in precipitation, as generally seen in the combined results of Chapters 2, 3 and 4, could lead to “heat droughts” or periods of extreme heat that have been reported by Lesk *et al*. (2016) to be just as damaging to agricultural yields as drought alone in terms of annual production losses. The analyses in this dissertation provide a quantitative basis for further investigations of the physical hydroclimate systems of the Khangai Mountain region. The results of this work will prove useful as a foundation for scientists and practitioners pursuing holistic approaches to facilitating future development of water policy and infrastructure in the region, while ideally supporting the importance of sustainable nomadic pastoral use of finite water resources under a changing climate.

1.1 Continental to Countrywide Context

1.1.1 Climate Change in Asia

Focusing inward from the global scale to the continental scale, documented and projected climate changes in Asia vary due to the size and complexity of the landscapes represented. In general, changing climate patterns over Asia have been less studied than those in other regions of the world (IPCC, 2014b). This is in part due to a lack of long-term monitoring records in many regions and/or wide spacing between stations. Some authors are working to reverse engineer the station distribution problem through creation of more detailed models that are sensitive to terrain and climate features (e.g., inversions) using downscaling methods and reanalysis data (Gerelchuluun and Ahn, 2014). A further complication in data scarce and developing regions is the proper siting and long-term maintenance of stations, which can affect record quality and homogeneity (Peterson *et al*., 1998; Beaulieu *et al*., 2007; Jamiyansharav, 2010). Records from established stations may also be subject to long periods of missing data, possibly due to natural disaster, unstable political conditions, and/or changes in policy (Ojima and Chuluun, 2008; IPCC, 2014b). Such disturbances can also affect the water supply infrastructure of developing countries (Upton, 2010; IPCC, 2014b).

Findings summarized in the most recent Intergovernmental Panel on Climate Change study (IPCC, 2014b) show it is likely that mean annual temperatures across much of Asia have increased, with a drop in the number of cold days and nights. Previous assessments have stated that warming is expected to continue into the 21st century and beyond, which may lead to changes in ecozone boundaries, amounts of net primary production, frequency and duration of droughts, loss of soils to increased erosion, and invasion of grasslands by non-native species (Christensen *et al*., 2007; Angerer *et al*., 2008). In Northern Asia, which includes eastern Russia and Mongolia, there is evidence of increasing trends in heavy precipitation, but arid areas in Russia, Mongolia, and northern China could experience water shortages in the near and long term future exacerbating poverty and existing resource inequalities (Angerer *et al*., 2008; IPCC, 2014b). While large changes in moisture conditions are projected to occur in semi-arid regions, precipitation projections are uncertain resulting in lower confidence in the types (and direction) of changes that will occur in Asian drylands (IPCC, 2014b).

Precipitation in northern and central Asia has been studied with mixed results regarding the identification of common patterns of precipitation distribution and changes to those patterns through time. Li *et al*. (2009) identified ten regional moisture patterns that have varied over China and Mongolia during the period of 1951-2005. Increasing precipitation trends in mid to high-latitude Eurasia were reported by Dai *et al*. (1997) from the turn of the last century up to 1998, but Klein Tank *et al*. (2006) found no consistent pattern of change in precipitation extremes for indices calculated from 1961-2000. They did however see a significant increase in amounts of rainfall for very wet days. Several studies have resulted in findings of no significant trends, or divergent trends that may mask or possibly represent decadal-scale spatio-temporal variability across parts of northern and central Asia (e.g., Aizen *et al*, 2001; Batima *et al*., 2005; Becker *et al*., 2013; Poulter *et al*., 2013; Wang *et al*., 2014).

1.1.2 Mongolia

1.1.2.1 Geography

Mongolia is unique in both its culture and climate when compared to other parts of Asia such as neighboring Russia to the north and China to the south (Figure 1.1- inset). The population of Mongolia in 2013 was about 2.85 million people with over 1.35 million people living in the metropolitan area of the capital city, Ulaanbaatar. The country is politically divided into 21 *aimags* (provinces) and 351 *soums* (counties) with smaller divisions called *bags* (local municipalities with or without an administrative center). The capital is a federal municipality and is not considered part of the 21 *aimags*. The population is largely Khalkh Mongolian with several minority ethnic groups (UNSD, 2012).

Governmental agencies such as the National Statistical Office of Mongolia (www.en.nso.mn) and reports like the *Mongolia Second National Communication* (Ministry of Nature, Environment and Tourism, 2010) provide information on the status of the nation and give assessments of climate change and vulnerability. They also provide broad descriptions of the geography and observed climate of the country. For example, Mongolia covers 1,564,116 square kilometers with diverse landscapes and ecological features including mountain, forest, steppe, and desert regions. Elevations range from a high of 4653 meters above sea level at Khuiten Uul (mountain) in the Western Altai Mountains to a low of 566 meters above sea level at the Khokh Nuur (lake) depression. Ulaanbaatar is about 1310 meters above sea level.

Lakes are an important surface water resource in Mongolia. The three largest lakes in the country are Uvs-Nuur, Khuvsgol, and Khar Us-Nuur, and have areal coverages of 3,350, 2,620, and 1,852 square kilometers, respectively (Ministry of Nature, Environment and Tourism, 2010). There are many smaller lakes in the country, some of which are saline and significantly reduce their size or even completely dry up in extended periods of low rainfall, such as Orog Lake at the terminus of the Tuin River in the arid desert steppe of Bayankhongor *aimag* (Fassnacht *et al*., 2015).

1.1.2.2 Temperatures and Precipitation

The relatively high elevation and dry nature of this landlocked country results in an extreme continental climate with four distinctive seasons. Gradients of temperature and moisture occur across the landscape from north to south and east to west. In Chapter 2, mean annual maximum and minimum temperatures were calculated for the period of 1963-2012 from gridded climate data (Figures 2.1b and 2.1c) (see Harris *et al*., 2014 for data; Venable *et al*. 2015). Temperatures are higher in the southern desert regions and in the valleys than in the mountains or northernmost parts of the country. High temperatures in July can reach 20 to 25 degrees Celsius (C) on average in the southern part of the Eastern steppe and in the Gobi, with extremes of 40+ degrees C in the desert. In the mountains, summer average air temperatures may only reach 15 degrees C (Batima *et al*., 2005). In January, lows of -35 degrees C to -15 degrees C occur from the mountains down to the Gobi desert, respectively (Batima *et al*., 2005). Temperatures in the northern regions are cold enough to maintain permafrost though recent warming is affecting permafrost conditions in some areas (Sharkuu, 2003). These changes have implications for water resources, hydrological modeling, and understanding shifting vegetation phenology (Huelsmann *et al*., 2015; Sun *et al*., 2015).

Abrupt changes in temperature have been observed on a daily and seasonal basis (Miyazaki *et al*., 1999). Sharp differences in temperature and moisture conditions have been observed in areas of complex terrain where elevation and aspect play a distinct role in the distribution of soil moisture conditions and vegetation type (Dulamsuren and Hauck, 2008). Evapotranspiration processes have been studied as they affect the availability of soil moisture to vegetation and can effectively limit water yields to downslope river systems (Li *et al*., 2007; Minderlein and Menzel, 2014). Changes to observed soil moisture in relation to evapotranspiration and precipitation have been examined at several sites, especially regarding the sensitivity and resilience of vegetation to moisture conditions and drought (Shinoda *et al*., 2010; Nandintsetseg and Shinoda, 2011; Nandintsetseg and Shinoda, 2013; Shinoda *et al*., 2014). Understanding vegetation condition and response to climate extremes is critical in Asian countries such as Mongolia where a large proportion of the population relies on the agricultural sector/livestock industry for their livelihoods (Yu *et al*., 2003; Angerer *et al*., 2008; John *et al*., 2013).

A majority of the precipitation across Mongolia falls in the summer months of June, July, and August, with about 85% of the total precipitation received between April and September (Ministry of Nature, Environment and Tourism, 2010). Figure 2.1a in Chapter 2 shows gradients from higher mean total annual precipitation in the north to lower in the south with more in mountain regions than in the plains (Venable *et al*., 2015). Generally, annual mean precipitation is about 300-400 mm in the mountains, 250-300 mm in the Altai (western mountains) and western steppe zones, 150-250 mm in the central and eastern steppe, and about 50-100 mm in the Gobi desert regions (Ministry of Nature, Environment and Tourism, 2010). Extremes of precipitation can occur with record rainfall events as high as 130 mm per day, and high hourly precipitation values can occur in the range of 40 to 65 mm per hour (Ministry of Nature, Environment and Tourism, 2010). Flooding is not uncommon, especially seasonally, as a majority of the river systems of the country are completely unregulated or controls are poorly developed. As a result, there have been increasing economic losses to storms, floods, and drought over the last few decades of extreme weather (Daavasuren, 2008; Ministry of Nature, Environment and Tourism, 2010).

Unlike in countries to the south, the influence of the East Asian Monsoon generally reaches its limits south of Mongolia (e.g., Yang *et al*., 2014). Synoptic-scale disturbances provide summer moisture via westerly winds originating in central Asia, Siberia, and from within Mongolia itself (Aizen *et al*., 2001; Sato *et al*., 2007b). The spatial variability and influence on precipitation amounts of these Westerlies versus the East Asian Monsoon over time has been the topic of recent investigations (e.g., Cook *et al*., 2010; Lee *et al*., 2013; Li *et al*., 2014; Yang *et al*., 2014).

In the winter, high-pressure systems over Siberia and Eurasia (the Siberian High) move cold air masses south and west across Mongolia (Aizen *et al*., 2001; Morinaga *et al*., 2003). Contributions to the water budget from snow can be important, particularly in the northern and mountainous regions. Some correlations have been made to deeper snows and colder temperatures from teleconnections with the North American Oscillation and to changes in the influence of the El Niño-Southern Oscillation (Morinaga *et al*., 2003).

1.1.2.3 Natural Hazards

Deep snow conditions and cold temperatures can combine to make white *dzud*, or winter disaster with concomitant losses of livestock. The term white indicates snowy conditions and is one of the most common types of *dzud*, with other types attributed to extreme cold and/or icy conditions (Tachiiri *et al*., 2008). Droughts and *dzuds* are some of the worst natural disasters facing the country economically with large losses of livestock occurring particularly after summer droughts are followed by winter *dzuds* (Siurua and Swift, 2002; Begzsuren *et al*., 2004; Angerer *et al*., 2008; Sternberg *et al*., 2011; Middleton *et al*., 2015). These losses have severe impacts on the socio-ecological and economic systems of the country, particularly affecting the nomadic pastoralists of the nation, as nearly half of the population is rural (Angerer *et al*., 2008; Fernandez-Gimenez *et al*., 2012; Fernandez-Gimenez *et al*., 2015a). Models are under development to understand the potential re-occurrence and distribution of *dzud* to help mitigate losses (Tachiiri *et al*., 2008).

Soil moisture deficits due to drought combined with windy conditions can lead to severe dust storms. According to one study, the number of dusty days has increased between 1960 and 1999 with relative humidity averaging only 20 to 40% during storm events (Natsagdorj, *et al*., 2003). High winds and dry conditions increase the chances for other human and natural-caused disasters such as wildfire, which impacts both steppe and forest resources (Hessl *et al*., 2012; Salydyga *et al*., 2013).

Losses of lakes and springs, and the lowering of river flows are a concern as they provide critical water supplies for the livestock of local nomadic pastoralists and provide valuable riparian and wetland habitats for threatened and endangered species and migratory waterfowl. The losses are attributed to several reasons, including warming and drying of the climate, increasing usage of shallow groundwater supplies for mining and irrigation, and other changing land uses (Brutsaert and Sugita, 2008; Ministry of Nature, Environment and Tourism, 2010, Tao *et al*., 2015).

1.1.2.4 Hydroclimate Trends Across the Country

Nomadic pastoralists of Mongolia live in close connection with the land. They employ observations of seasonal and annual changes to their landscape in management of the resources needed for their survival (Fernandez-Gimenez, 2000; Green and Raygorodetsky, 2010). Abrupt changes and the perceived (or actual) increasing frequency of extremes strain their capacity to manage and/or adapt to a changing landscape. These changes are not only attributed to climate, but also represent changes to socio-ecological, economic, and governmental/policy systems (Fernandez-Gimenez, 2000; Fernandez-Gimenez and Batbuyan, 2004; Marin, 2010; Regdel *et al*., 2012; Lkhagvadorj *et al*. 2013; Bruegger *et al*., 2014; Fernandez-Gimenez *et al*., 2015a; Fernandez-Gimenez *et al*., 2015b; Jigjisuren *et al*., 2015).

Historical changes to countrywide precipitation and temperature patterns differ, with more coherence in temperature trends than precipitation trends. This finding is similar to the results from other parts of Asia. Decreasing precipitation trends are found for some parts of Mongolia and increasing trends in other parts of the country depending on the time interval examined and methods employed (e.g., Batima *et al*., 2005; Endo *et al*., 2006; Venable *et al*., 2012b; Venable *et al*., 2012c; Hilker *et al*., 2104; Nandintsetseg and Shinoda, 2014; Eckert *et al*., 2015). In a study focusing on metrics used to assess change in Mongolian observational records, Jamiyansharav (2010) examined trends in precipitation, temperature, and climate change indices at 17 stations across the country. Her results indicate increasing temperature-related trends but no unified temporal or spatial patterns in precipitation trends.

While many authors use linear or other basic statistical methods to infer changes to observed precipitation, some authors are interested in predicting precipitation occurrence and change. Kim *et al*. (2011) analyzed changing precipitation across the country and attempted to predict precipitation variability using a time series analysis model. To determine the effects of a warming climate, Sato *et al*. (2007a) used a predictive regional climate model of July precipitation to show future amounts of rainfall decreasing over northern Mongolia and increasing over southern areas, with decreases in modeled soil moisture due to rising temperatures.

As a comparison to countrywide station-based results by other authors, Chapter 2 uses the Mann-Kendall test for trend significance (Mann, 1945; Kendall and Gibbons, 1990) and Thiel-Sen slope (rate of change) estimator (Thiel, 1950; Sen, 1968) in an analysis of monthly gridded data at a 0.5-degree latitude by 0.5-degree longitude (~55 km) spatial resolution. Aggregate annual and seasonal values over the 50-year period from 1963-2012 were tested for significance and magnitude of trend with the spatio-temporal results on an annual timestep displayed in Figures 2.2a, 2.2b, and 2.2c. (see Harris *et al*., 2014 for data; Venable *et al*., 2015).

An advantage of using gridded data is that it provides a smoothed spatial surface with any missing data filled or not included in interpolation, providing continuous time series for analysis. Disadvantages to the use of gridded datasets include widely differing methods used for data filling and interpolation making comparison between datasets potentially difficult (e.g., Hijmans *et al*., 2005; Willmott and Matsuura, 2012; Harris *et al*., 2014; Schneider *et al.*, 2014). Also, due to the smoothing algorithms of the selected interpolation process, gridded datasets may not be able to capture the extremes of station-based point data, or may change the statistical characteristics of meteorological data by highly over or underestimating station values at certain times of the year, or in certain years (Figures 1.2a and 1.2b) (Ensor and Robeson, 2008; Venable *et al*., 2014a). Additionally, outputs from hydrological models can differ significantly when driven with inputs from station-based and/or gridded datasets from the same region, depending on choice of gridded product or station used (Venable *et al*., 2014b; Malsy *et al*., 2015).

In the results presented in Chapter 2, annual maximum temperatures were increasing for most areas, but annual minimum temperatures were significantly increasing across the entire nation (Figures 2.2b and 2.2c). Significant changes in annual precipitation however, only occurred in limited regions, primarily across the north central portion of the country (Figure 2.2a). An example of seasonal changes in precipitation calculated using the same methods is shown in Figure 1.3. Significantly decreasing trends are seen across the central and eastern parts of the country similar to the annual trends (Figure 2.2a) emphasizing the strong summer-dominated precipitation patterns in Mongolia. The magnitude of the significant summer trends is slightly less than that for the annual trends (Venable *et al*., 2015).

1.2 Regional Focus

The Khangai Mountain region of Mongolia is a diverse area of mountain, forest, steppe, and desert steppe landscapes reaching across and beyond the mountains. The region encompasses portions of three *aimags* and covers approximately 150,000 square kilometers (Figure 1.1). It is home to over 265,000 people (as of 2012), though the three largest administrative centers of Tsetserleg in Arkhangai *aimag*, Bayankhongor in Bayankhongor *aimag*, and Arvaikheer in Ovorkhangai *aimag* each have populations of less than 30,000 people (as of 2012). The tradition of nomadic pastoralism is strong in the region, with a reliance on water for domestic and livestock needs supplied naturally through lakes, springs and rivers, and through shallow hand-dug, and deeper drilled wells (located primarily at administrative centers).

Several investigations by researchers from Colorado State University have been conducted in the region over the last 20+ years. Surveys of nomadic pastoralists have been performed with questions focusing on themes such as indigenous knowledge, social and economic welfare (adaptive capacity), and environmental, ecological, and climate changes (e.g., Fernandez-Gimenez, 1993; Fernandez-Gimenez, 2000; Fassnacht *et al*., 2011; Fernandez-Gimenez *et al*., 2012; Sukh, 2012; Bruegger *et al*., 2014; Fernandez-Gimenez *et al*., 2015a). In an effort to better understand herder's observations of their changing ecological environment from a physically quantitative standpoint, researchers have paired survey responses with analyses of hydroclimatic variables (e.g., Marin, 2010), with varying levels of agreement and disagreement between responses and the station-based analysis results. General agreement between the two methods of inquiry suggests that herders are observant of changes to longer-term climate conditions and that their observations may provide a level of spatial and temporal detail not possible in coarse-scale analyses of gridded hydroclimatic data (e.g., Fernandez-Gimenez *et al*., 2015b) or for finer-scale, but still widely spaced, station-based data (e.g., Fassnacht *et al*., 2011; Sukh, 2012). Their observations, regardless of level of agreement with statistical hydroclimatic change evidence provide insight into those hydroclimatological variables or climate influenced changes that are not commonly measured or typically supported by data collected from Mongolian networks, such as rainfall intensity or the impact energy of rain events on soil (Marin, 2010).

1.2.1 Temperature Trends Through Time and Herder's Observations

As an extension of previous studies, Venable *et al*. (2012a) as Chapter 3, explores herder's perceptions [observations] of temperature change through time for five meteorological stations in the Khangai Mountain region using the Mann-Kendall test for trend significance and Thiel-Sen's slope analysis methods. The central questions of investigation in Chapter 3 are: "*What time period of reference is used when recalling observations of extremes and/or changes in temperature?"* and, "*How do these memories compare to temperature trend analyses of differing lengths (testing different periods through time)?"* These observations are recorded in terms of changes in the occurrence of warmer and cooler days and nights, or seasonal changes such as the onset or loss of snowcover. One extremely cold winter (2009-2010) was also examined in more detail to illustrate the effects of using differing record lengths on the significance of trends.

The most highly significant trends were in daily minimum temperatures for longer periods of record. At four out of the five stations tested, the 30-year period from 1966-1995 exhibited change significant at the $p<0.001$ level. Including the extreme winter of 2009-2010 in the analyses resulted in a decrease in rates of change (a decrease in warming) and a decrease in the significance level assigned to the different periods tested. This effect was particularly true for shorter periods of analysis. Even though the Mann-Kendall test for trend significance and Thiel-Sen slope estimators are considered less affected by outliers than other methods like regression analysis (e.g., Gilbert, 1987), these results suggest that even one extreme year can have an effect on determination of significant change in the observational temperature record.

1.2.2 Testing Trend Detection and Exploring Data Uncertainties

Station selection is critical for the determination of possible trends as station location can influence trend variability across a region (Burn and Hag Elnur, 2002; Pielke *et al*., 2002; Fassnacht *et al*., 2016). In developing countries like Mongolia however, there are only a few stations available with limited climatic records over large areas. The data are also highly variable and large quantities of data are often missing. Non-parametric statistical tests like the Mann-Kendall are robust to outliers and allow for missing values (Gilbert, 1987; Helsel and Hirsch, 2002). It is known that hydroclimatic variables can exhibit large degrees of persistence (autocorrelation) depending on the time step chosen for analysis such as daily, monthly, or annual (Helsel and Hirsch, 2002). High levels of data autocorrelation affect the results of many statistical analyses including the Mann-Kendall test, calling into question the significance of results of past trend analyses using these methods (Yue *et al*., 2002). Additionally, the datasets may contain inhomogeneities from station relocations or other causes which can influence the results of trend testing (Kundzewics and Robson, 2004; Gallagher *et al*., 2013)

Chapter 4 conducts trend analyses of precipitation and streamflow variables using data from six meteorological stations and four streamflow gaging sites across the Khangai Mountain region. Standard versions of the Mann-Kendall and Thiel-Sen approaches were used as in previous investigations (e.g., Fassnacht *et al*., 2011; Sukh, 2012; Venable *et al*., 2012b; Venable *et al*., 2012c). A variation developed by Yue *et al.* (2002) was also employed to remove autocorrelation in the records via prewhitening (results in observations that are statistically independent from one another in a given time series) prior to applying the Mann-Kendall test for trend significance and Thiel-Sen slope estimator. While no bias corrections were performed on the precipitation data, quality control measures were employed including tests for station change points (Pettitt, 1979; Zhang *et al*., 2004; Harris *et al*., 2014).

Seasonal and annual precipitation and annual mean and median streamflow were tested for trend and trend significance using the original data and re-tested again after prewhitening with similar results. Only two stations/gages had results that differed at the $p<0.05$ level for the two methods applied. These locations were on the northern side of the Khangai Mountains at Erdenemandal (a meterological station with decreasing summer precipitation) and at Ikhtamir (location of the Khoid Tamir River gage with decreasing median flow). When preparing the data prior to the trend analyses, certain assumptions were made concerning missing values. When these assumptions are not used, nearly 23% of the data are missing for the Khoid Tamir gage, and potentially 74% are missing for the Erdenemandal station. Missing data occurs randomly and in distinct blocks, and is likely related to a combination of factors such as equipment failure, lack of monitoring at certain times, differing rules applied to missing/trace measurements, or possibly data loss when stored in repositories. Inhomogeneities were found in the Khoid Tamir River gage time series, but no station movements were recorded for either the Erdenemandal or Khoid Tamir climate records.

1.2.3 Fieldwork in the Khangai Mountain Region

While analyses of existing datasets can give an idea of the statistical status and nature of hydroclimate systems, the opportunity to collect new data can add a dimension of relevance to investigations that is often missing when time is only spent in the laboratory. In the summer of 2012, funding was provided by the Mongolian Rangelands and Resilience project at Colorado State University to conduct hydrologic and hydraulic fieldwork in the Khangai Mountain region. Part of that work included visiting several of the meteorological stations and streamflow gaging sites of the region used in previous analyses to further understand the types of equipment used for measurements and the possible uncertainties associated with data collected from these sites (e.g., reasons for missing data, station relocations, reliability of stage measurements in varying flow conditions, etc.). Additional funding was secured from the American Center for Mongolian Studies to collect paleo proxy tree-ring data from two sites in the region. This provided new core material for the research conducted in Chapter 5. Another benefit of fieldwork was direct collaboration between the author and her advisor, and two Mongolian colleagues that participated in hydrologic/hydraulic and paleo proxy data collection (Venable and Fassnacht, 2012).

1.2.3.1 Challenges for Future Hydrological Modeling Projects

The hydrologic/hydraulic portion of the field expedition collected streamflow measurements at several locations along the Khoid Tamir and Tuin Rivers, though most of the measurements taken were of the Tuin River from near its source in the Khangai Mountains to its terminus at Orog Lake in the Gobi Desert. Measurements were collected using standard hydrologic field methods with hydraulic measurements of floodplain extents and geomorphic condition also made at several locations (Fassnacht *et al.*, 2015).

One goal of the fieldwork was to determine the feasibility of future hydrological modeling of the Khoid Tamir and Tuin River Basins, and possibly the Khanui and Baidrag Basins. The results show that a better understanding of flow conditions and more data are needed for successful hydrological modeling of bankfull, flood flow, and low flow conditions, such as those that were sampled in the summer of 2012 prior to the onset of the summer rains (Fassnacht *et al*., 2015).

A lack of data for modeling critical aspects of hydrological systems is common in developing regions (Hrachowitz *et al*., 2013). Several authors and research groups have attempted hydrological modeling of a number of basins across Mongolia, particularly those adjacent to larger population centers and in more developed parts of the country. In an early study, Ma *et al*. (2003) examined the hydrological regimes of the Selenge River basin using a soil, vegetation, and atmospheric transfer model. Only five years were modeled (1988-1992), and one of the main determinations of the study was that evaporation dominates the hydrological cycle in the basin. The authors admit that sparsely distributed and missing data provided difficulties for modeling, even within the short time period of interest.

More recently, concerns over land use change and an increase in policies aimed at improving food security in the country have generated new reasons to model hydrological systems. An interdisciplinary project to study Integrated Water Resources Management in Central Asia is being conducted in the Khaara River basin which lies north and to the east of the Khangai Mountain region, north of the capital of Ulaanbaatar. The project uses an integrative framework and both land and hydrological models to simulate land use and land cover dynamics and to study the rainfall-runoff characteristics of the catchment to examine the effects of sustainable and unsustainable irrigation practices (Menzel *et al*., 2011; Priess *et al*., 2011). Increasingly, the potential benefits of using Integrated Water Resources Management principles, such as increased social equity, gains in economic efficiency, and promotion of ecological sustainability (GWP, 2000), are being brought to the attention of land managers and policy makers in developing countries like Mongolia (e.g., Karthe *et al*., 2015; Priess *et al*., 2015). Researchers acknowledge mismatches in fit and interplay between institutional arrangements and hydro/biogeophysical systems but recognize that holistic and sustainable management practices are necessary to increase food security, reduce poverty, and assess and mitigate impacts from non-renewable resource extraction, such as mining and use of paleo waters for irrigation, while accommodating or even facilitating traditional uses of the landscape (Priess *et al*., 2011; Horlemann and Dombrowsky, 2012; Currell *et al*., 2012; Priess *et al*., 2015).

With the continued rise in irrigation and industrial water needs, some researchers are investigating the possibility of further developing paleo water resources based on preliminary paleo ecological work. These water supplies are sourced from aquifers that developed during previously wetter climate regimes in the desert regions of Mongolia. Current water exploration work focuses on the Ulaan Nuur basin to the south and east of the Khangai Mountain region, nearer to major mining developments, and in a region with herders and communities already utilizing shallow groundwater resources via hand-dug wells (Sternberg and Paillou, 2015). Past investigations of paleo climate have also been conducted in basins of the Khangai Mountain region including at Buuntsagaan Lake (terminus of the Baidrag River) and Orog Lake (terminus of the Tuin River) (e.g., Komatsu *et al*., 2001).

1.2.4 Paleo Proxies and Streamflow Variability

From a water resources management perspective, it has been shown that investigations of changes to the modern hydroclimate may benefit from the context gained by examining these changes over time periods longer than those of the historical record. This is accomplished through the use of moisture sensitive paleo proxies like tree-rings (e.g., Meko and Woodhouse, 2011). It is unlikely that the full range of hydroclimatic variability through time has been captured in the short observational records of streamflow, hindering the ability of water managers and researchers to develop plausible scenarios of drought extremes and magnitudes that differ from those seen in the observational record (e.g., Stockton and Jacoby, 1976; Woodhouse and Lukas, 2006; Woodhouse *et al*., 2006).

In Mongolia, work has been done in selected regions to quantify and extend flow regimes with long-term flow scenarios derived from tree-ring analyses (e.g., Pederson *et al*., 2001; Davi *et al*. 2006; Davi *et al*., 2013; Pederson *et al*., 2013). Additionally, researchers have examined unique spatial patterns of long-term hydroclimatic variability that provide context to existing investigations of precipitation and drought extremes in the historical climate record (e.g., Davi *et al*., 2010; Leland, 2011; Leland *et al*, 2013). Recent tree-ring reconstructions over central Mongolia of the Palmer Drought Severity Index, a widely used index of meteorological drought conditions (e.g., Dai *et al*., 2004), show that the droughts of the last decade and a half are unusually extreme if not unprecedented over the last millennium (Hessl *et al*., 2015). Work has also been completed examining seasonal variations in moisture conditions as derived from interannual tree-ring growth properties (e.g., De Grandpré *et al*., 2011; Wolf and Venable, 2015).

The Khangai Mountain region may exhibit a unique hydroclimatic signal through time as compared to other regions, which could influence the long-term variability of streamflow and the timing and magnitude of drought events in this area as compared to other basins in Mongolia (Leland *et al*., 2013). According to Fowell *et al*. (2013), 6000-year old paleo lake sediments from Lake Telmen located north of the Khangai Mountains indicate arid conditions at that time, which contrasts with the more humid conditions found for other paleo proxy site across Mongolia during the Holocene climatic optimum (9000 to 5000 years BCE). They assert that the Khangai Mountains may have acted as a limit line to expanded Asian Monsoon flow from the south. In modern times, the elevated terrain of the mountains still act as a barrier to the flow of moisture from north to south, and from west to east, resulting in different precipitation patterns in this area than in other regions like the Khentii Mountains to the east (Kim *et al*., 2011). The general concept of differing long-term climate conditions on different sides of the Khangai Mountains is tested in Chapter 5 through the development and analysis of streamflow reconstructions for the Khanui River at Erdenemandal, the Khoid Tamir River at Ikhtamir, the Baidrag River at Bayanburd, and the Tuin River at Bayankhongor (see Figure 1.1 for locations).

Additionally, a key question explored in Chapter 5 is: "*Are the steep decreases observed in the modern streamflow records of the region since the mid to late 1990's within the range of natural variability of these systems, or are they extraordinary over the 300+ year period of record available for analysis?*" The two new tree-ring core sets collected in the fieldwork completed in 2012 are used in part to answer this question.

While the reconstructions results are not the same for all four basins modeled, they are broadly similar and synchronous for certain extreme dry and wet events over a 300+

year period. More heterogeneity in event timing is seen on an east-west basis than north-south, which may be due to differing climate signals affecting the predictors used for modeling, or because different predictors were selected for each reconstruction based on correlations with the streamflow data. In a drought and pluvial (interval of relatively high precipitation) analysis, the top five driest and wettest non-overlapping intervals over the periods reconstructed were calculated for each basin.

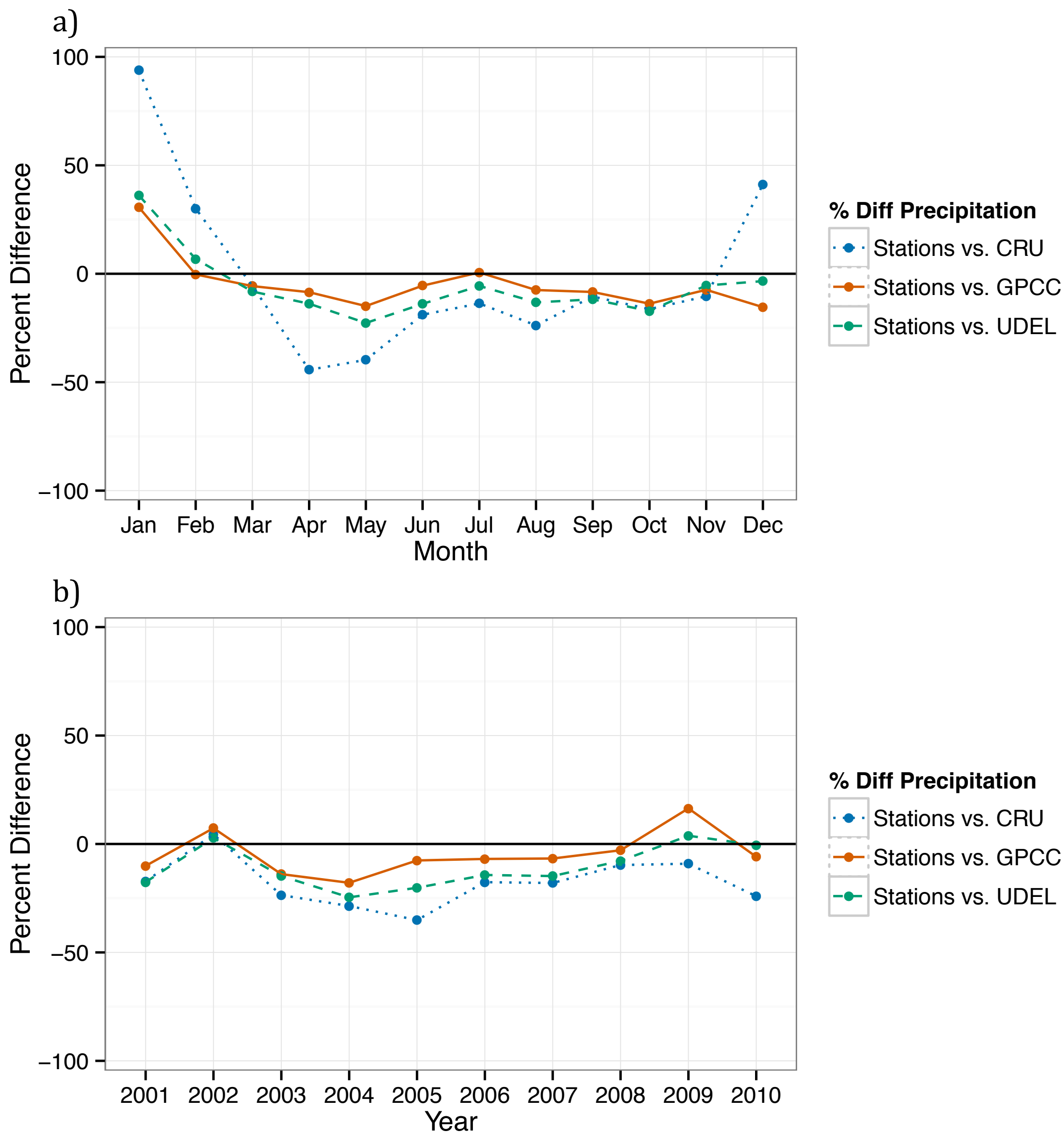
The new 2012 chronologies were used to create reconstructions that extended through the extreme low flows of the last decade and a half, since most of the other chronologies used in Chapter 5 extended to 2002 at the latest. Despite the shorter length and less sensitive nature of the new chronologies compared to the other series used for reconstruction, the alternate models including them (up to 2009) show mean flows during the decade of the 2000's that are at least as low as the top five most extreme dry periods reconstructed over longer periods for each basin. This result suggests that the steep decreases in modern streamflow while not common in the long records are not unprecedented. In some rivers like the Khanui, much drier events and those spanning longer durations than the modern have likely occurred in the past.

1.3 Hydrometeorological Analyses as Cross-Disciplinary Objects

The last chapter examines how the data and analyses presented throughout this dissertation are used in a cross-boundary research manner through explorations of interdisciplinary connections within each chapter and to selected goals of the larger MOR2 project. The boundary object concept developed by Star and Griesemer (1989) allows research results to act as objects or constructs that span multiple social (disciplinary) worlds enabling bridging between those worlds and fostering the collaboration and cooperation between groups/disciplines necessary to the success of this type of work (Star, 2010; Akkerman and Bakker, 2011).

The interdisciplinary approach undertaken in this dissertation merges methods and results that are multidisciplinary in nature into one document. Instead of delving deeply into any single element of the hydroclimatologic system of Mongolia, this work takes a broad look at the changing hydroclimate through the use of trend analyses and paleo climate investigations. It furthers themes developed in previous research from across Asia and within Mongolia, using the context and perspectives developed in analyses of data from the Khangai Mountain region to provide new insights. This work documents statistically significant increases in temperature, paticularly minimum temperatures across the country. In the Khangai Mountain region, herders have observed warming conditions which may affect their ability to sustain their traditional livelihoods. Statistical significance and rates of change in precipitation and streamflow are not uniform across the country and vary even within a region. Regional spatio-temporal heterogeneity is also found in reconstructed river flows for periods hundreds of years prior to the instrumented record. A common theme discussed throughout this dissertation is the effect of data type, length, quality, and chosen method of data analysis on research results. This work closes with an examination of how the analyses and syntheses presented within, when objectified or used as discussion points, can bridge disciplinary divides and provide a foundation for the scientific understanding needed to facilitate future water development and policy decisions in this dry landscape (e.g., Gleick, 1989; Gleick, 2010; Priess *et al*., 2011). Such policy and management decisions will require strong consideration of the climate change-related issues faced by nomadic pastoralists of the country.

FIGURE 1.1- Mongolia and surrounding countries (inset) with main map focus on portions of the three *aimags* of the Khangai Mountain region. Includes outlined boundaries of the four river basins of interest and locations of meteorological stations (black squares) and stream gages (black circles) highlighted in subsequent chapters.



FIGURES 1.2a- (top) and 1.2b- (bottom) Plots of the percent difference on a monthly (top) and annual basis (bottom) between values extracted from three different gridded precipitation products and station-based data from 40 sites across the Khangai Mountain region of Mongolia. CRU- Climatic Research Unit (Harris *et al*., 2014), GPCC- Global Precipitation Climatology Centre (Schneider *et al*., 2014), and UDEL- University of Delaware (Willmott and Matsuura, 2012)

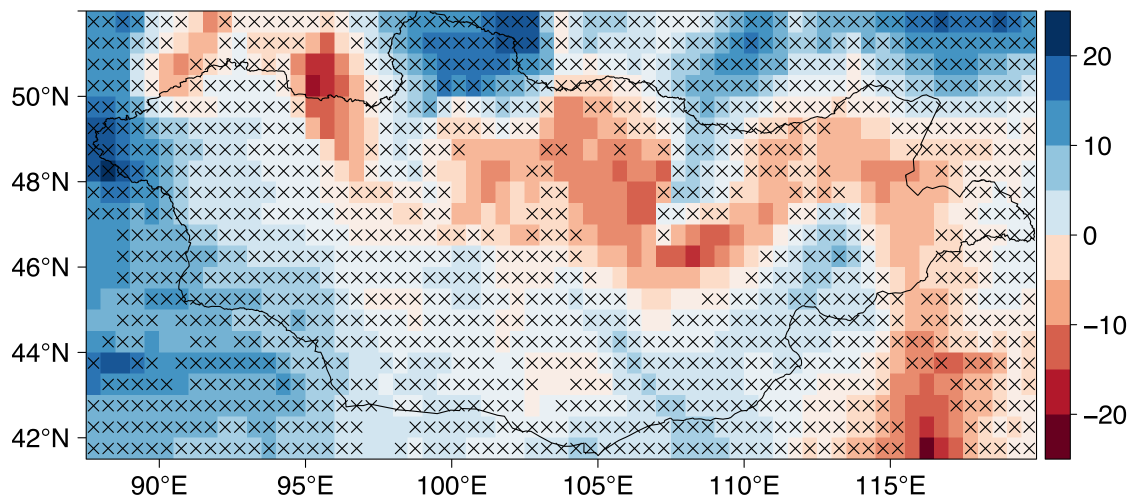


FIGURE 1.3- Trend of total summer (June, July, and August) precipitation from 1963-2012 in millimeters per decade for Mongolia. Areas without X's are significant at the $p<0.05$ level.

REFERENCES

Aizen EM, Aizen VB, Melack JM, Nakamura T, and Ohta T, 2001. Precipitation and atmospheric circulation patterns at mid-latitudes of Asia. *International Journal of Climatology*, 21, 535–556, doi:10.1002/joc.626.

Akkerman SF, and Bakker A, 2011. Boundary crossing and boundary objects. *Review of Educational Research*, 81, 132–169, doi:10.3102/0034654311404435.

Angerer J, Han G, Fujisaki I, and Havstad K, 2008. Climate change and ecosystems of Asia with emphasis on Inner Mongolia and Mongolia. *Rangelands,* 30(3), 46-51.

Batima P, Natsagdorj L, Gombluudev P, Erdenetsetseg B, 2005. Observed climate change in Mongolia. *Assessments and Adaptations to Climate Change* (AIACC), *www.aiaccproject.org*. Accessed 2011-11-21.

Beaulieu C, Ouarda T, Seidou O, 2007. A review of homogenization techniques for climate data and their applicability to precipitation series. *Hydrological Sciences Journal-Journal des Sciences Hydrologiques* 52, 18–37, doi:10.1623/hysj.52.1.18.

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, and Ziese M, 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5, 71–99, doi:10.5194/essd-5-71-2013.

Begzsuren S, Ellis JE, Ojima DS, Coughenour MB, and Chuluun T, 2004. Livestock responses to droughts and severe winter weather in the Gobi Three Beauty National Park, Mongolia. *Journal of Arid Environments*, 59, 785-796.

Bruegger RA, Jigjsuren O, and Fernandez-Gimenez ME, 2014. Herder observations of rangeland change in Mongolia: Indicators, causes, and application to community-based management. *Rangeland Ecology & Management*, 67, 119–131, doi:10.2111/REM-D-13-00124.1.

Brutsaert W, and Sugita M, 2008. Is Mongolia's groundwater increasing or decreasing? The case of the Kherlen River basin / Les eaux souterraines de Mongolie s'accroissent ou décroissent-elles? Cas du bassin versant la Rivière Kherlen. *Hydrological Sciences Journal*, 53, 1221–1229, doi:10.1623/hysj.53.6.1221.

Burn DH, and Hag Elnur MA, 2002. Detection of hydrologic trends and variability. *Journal of Hydrology* 255, 107–122, doi:10.1016/S0022-1694(01)00514-5.

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon WT, Laprise R, Rueda VM, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, and Whetton P, 2007. Regional climate projections. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the *Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL [Eds]. Cambridge University Press, Cambridge, UK and New York, NY, USA, 847-940.

Cook ER, Anchukaitis KJ, Buckley BM, D'Arrigo RD, Jacoby GC, and Wright WE, 2010. Asian Monsoon failure and megadrought during the last millennium. *Science*, 328, 486–489, doi:10.1126/science.1185188.

Currell MJ, Han D, Chen Z, and Cartwright I, 2012. Sustainability of groundwater usage in northern China: Dependence on palaeowaters and effects on water quality, quantity and ecosystem health. *Hydrological Processes*, 26, 4050–4066, doi:10.1002/hyp.9208.

Daavasuren J, 2008. Status and challenges of water infrastructure development in Mongolia. *First Regional Workshop on Development of the Eco-Efficient Water Infrastructure*, Ulaanbaatar, MN, 8 pp.

Dai A, Fung IY, and Del Genio AD, 1997. Surface observed global land precipitation variations during 1900–88. *Journal of Climate*, 10, 2943–2962, doi:10.1175/1520-0442(1997)010<2943:SOGLPV>2.0.CO;2.

Dai A, Trenberth KE, and Qian T, 2004. A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming. *Journal of Hydrometeorology*, 5, 1117–1130, doi:10.1175/JHM-386.1.

Davi NK, Jacoby GC, Curtis, AE, and Baatarbileg N, 2006. Extension of drought records for Central Asia using tree rings: West-Central Mongolia. *Journal of Climate*, 19, 288–299, doi:10.1175/JCLI3621.1.

Davi N, Jacoby G, Fang K, Li J, D'Arrigo R, Baatarbileg N, and Robinson D, 2010. Reconstructing drought variability for Mongolia based on a large-scale tree ring network: 1520–1993. *Journal of Geophysical Research: Atmospheres*, 115, D22103, doi:10.1029/2010JD013907.

Davi NK, Pederson N, Leland C, Nachin B, Suran B, and Jacoby GC, 2013. Is eastern Mongolia drying? A long-term perspective of a multidecadal trend. *Water Resources Research*, 49, 151–158, doi:10.1029/2012WR011834.

De Grandpré L, Tardif JC, Hessl A, Pederson N, Conciatori F, Green TR, Oyunsanaa B, and Baatarbileg N, 2011. Seasonal shift in the climate responses of Pinus sibirica, Pinus sylvestris, and Larix sibirica trees from semi-arid, north-central Mongolia. *Canadian Journal of Forest Resources*, 41, 1242–1255, doi:10.1139/x11-051.

Dulamsuren C, and Hauck M, 2008. Spatial and seasonal variation of climate on steppe slopes of the northern Mongolian mountain taiga. *Grassland Science*, 54, 217–230, doi:10.1111/j.1744-697X.2008.00128.x.

Eckert S, Huesler F, Liniger H, and Hodel E, 2015. Trend analysis of MODIS NDVI time series for detecting land degradation and regeneration in Mongolia. *Journal of Arid Environments*, 113, 16-28, doi:10.1016/j.jaridenv.2014.09.001.

Endo N, Kadota T, Matsumoto J, Ailikun B, & Yasunari T, 2006. Notes and correspondence: Climatology and trends in summer precipitation characteristics in Mongolia for the period 1960–98. *Journal of the Meteorological Society of Japan*, 84(3), 543–551.

Ensor LA, and Robeson SM, 2008. Statistical characteristics of daily precipitation: Comparisons of gridded and point datasets. *Journal of Applied Meteorology and Climatology*, 47, 2468–2476, doi:10.1175/2008JAMC1757.1.

Fan P, Chen J, and John R, 2016. Urbanization and environmental change during the economic transition on the Mongolian Plateau: Hohhot and Ulaanbaatar. *Environmental Research*, 144, Part B, 96–112, doi:10.1016/j.envres.2015.09.020.

Fassnacht SR, Sukh T, Fernandez-Gimenez ME, B Batbuyan, Venable NBH, Laituri M, and Adyabadam G, 2011. Local understanding of hydro-climatic changes in Mongolia. Proceedings of *H02 Symposium Cold Region Hydrology in a Changing Climate* held during IUGG2011 in Melbourne, Australia, July 2011, *IAHS Publication* 346, 120-129.

Fassnacht SR, Venable NBH, Odgarav J and Adyabadam G, 2015. A journey down the Tuin: The hydraulics of an internal draining river from the Khangai Mountains to the Gobi Desert. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 99-103.

Fassnacht SR, Cherry ML, Venable NBH, and Saavedra F, 2016. Snow and albedo climate change impacts across the United States Northern Great Plains. *The Cryosphere*, 10, 329-339.

Fernandez-Gimenez M, 1993. The role of ecological perception in indigenous resource management: A case study from the Mongolian forest-steppe. *Nomadic Peoples*, 33, 31–46.

Fernandez-Gimenez ME, 2000. The role of Mongolian nomadic pastoralists' ecological knowledge in rangeland management. *Ecological Applications* 10, 1318–1326, doi:10.1890/1051-0761(2000)010[1318:TROMNP]2.0.CO;2.

Fernandez-Gimenez ME, and B Batbuyan, 2004. Law and disorder: Local implementation of Mongolia's Land Law. *Development & Change*, 35, 141–165, doi:10.1111/j.1467-7660.2004.00346.x.

Fernandez-Gimenez ME, B Batkhishig, and B Batbuyan, 2012. Cross-boundary and cross-level dynamics increase vulnerability to severe winter disasters (dzud) in Mongolia. Global Environmental Change, 22, 836–851, doi:10.1016/j.gloenvcha.2012.07.001.

Fernandez-Gimenez ME, B Batkhishig, B Batbuyan, and Ulambayar T, 2015a. Lessons from the dzud: Community-based rangeland management increases the adaptive capacity of Mongolian herders to winter disasters. World Development 68, 48-65, doi:10.1016/j.worlddev.2014.11.015.

Fernandez-Gimenez ME, Angerer JP, Allegretti AM, Fassnacht SR, Byamba A, J Chantsallkham, Reid RS and Venable NBH, 2015b. Integrating herder observations, meteorological data and remote sensing to understand climate change patterns and impacts across an eco-climatic gradient in Mongolia. Proceedings of Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 228-234.

Fowell SJ, Hansen BCS, Peck JA, Khosbayar P, and Ganbold E, 2003. Mid to late Holocene climate evolution of the Lake Telmen basin, north central Mongolia, based on palynological data. Quaternary Research, 59, 353–363, doi:10.1016/S0033-5894(02)00020-0.

Gallagher C, Lund R, Robbins M, 2013. Changepoint detection in climate time series with long-term trends. Journal of Climate 26, 4994–5006, doi:10.1175/JCLI-D-12-00704.1.

Gerelchuluun B, and Ahn JB, 2014. Air temperature distribution over Mongolia using dynamical downscaling and statistical correction. International Journal of Climatology, 34, 2464–2476, doi:10.1002/joc.3853.

Gilbert RO, 1987. Statistical Methods for Environmental Pollution Monitoring. John Wiley & Sons, New York, 320 pp.

Gleick PH, 1989. Climate change, hydrology, and water resources. Reviews of Geophysics, 27, 329–344, doi:10.1029/RG027i003p00329.

Gleick PH, 2010. Forward. Written in Water: Messages of hope for Earth’s most precious resource. Salina I, [Ed]. National Geographic Society, Washington, DC, 7-12.

Green, D., and G. Raygorodetsky, 2010. Indigenous knowledge of a changing climate. Climatic Change, 100, 239–242, doi:10.1007/s10584-010-9804-y.

GWP, 2000. Global Water Partnership Technical Advisory Committee (GWP-TAC). Integrated Water Resources Management, GWP-TAC, Stockholm, Sweden, 71 pp.

Harris I, Jones PD, Osborn TJ, and Lister DH, 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. International Journal of Climatology, 34, 623–642, doi:10.1002/joc.3711.

Helsel DR, and Hirsch RM, 2002. Statistical methods in water resources. Techniques of Water-Resources Investigations of the United States Geological Survey, Book 4, Hydrologic Analysis and Interpretation, Chapter A3. http://water.usgs.gov/pubs/twri/twri4a3/. Accessed 2014-11-02.

Hessl AE, Ariya U, Brown P, Byambasuren O, Green T, Jacoby G, Sutherland EK, Nachin B, Maxwell RS, Pederson N, De Grandepré L, Saladyga T, Tardif JC, 2012. Reconstructing fire history in central Mongolia from tree-rings. International Journal of Wildland Fire, 21, 86–92, doi:10.1071/WF10108.

Hessl AE, Pederson N, Byambasuran O, Anchukaitis K, Leland C, 2015. How unusual was the 21st century drought in Mongolia? Placing recent extremes in an 1100-year context. Proceedings of Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 80-86.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, and Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25, 1965–1978, doi:10.1002/joc.1276.

Hilker T, Natsagdorj E, Waring RH, Lyapustin A, and Wang Y, 2014. Satellite observed widespread decline in Mongolian grasslands largely due to overgrazing. Global Change Biology, 20, 418–428, doi:10.1111/gcb.12365.

Hoerling M, Eischeid J, and Perlwitz J, 2010. Regional precipitation trends: Distinguishing natural variability from anthropogenic forcing. Journal of Climate, 23, 2131–2145, doi:10.1175/2009JCLI3420.1.

Horlemann L, and Dombrowsky I, 2012. Institutionalising IWRM in developing and transition countries: The case of Mongolia. Environmental Earth Sciences, 65, 1547-1559, doi:2048/10.1007/s12665-011-1213-7.

Hrachowitz M, Savenije HHG, Bloschl G, McDonnell JJ, Sivapalan M, Pomeroy JW, Arheimer B, Blume T, Clark MP, Ehret U, Fenicia F, Freer JE, Gelfan A, Gupta HV, Hut RW, Montanari A, Pande S, Tetzlaff D, Troch PA, Uhlenbrook S, Wagener T, Winsemius HC, Woods RA, Zehe E, and Cuennec C, 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. Hydrological Sciences Journal, 58, 1198–1255, doi:10.1080/02626667.2013.803183.

Huelsmann L, Geyer T, Schweitzer C, Priess J, and Karthe D, 2015. The effect of subarctic conditions on water resources: Initial results and limitations of the SWAT model applied to the Kharaa River Basin in Northern Mongolia. Environmental Earth Sciences, 73, 581–592, doi:10.1007/s12665-014-3173-1.

IPCC, 2014a. Summary for policymakers. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, and White LL [Eds]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1-32.

IPCC, 2014b. Chapter 24: Asia. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Barros VR, Field CB, Dokken DJ, Mastrandrea MD, Mach KJ, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, and White LL [Eds]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 688 pp.

Jamiyansharav K, 2010. Long-term analysis and appropriate metrics of climate change in Mongolia. Unpublished PhD Dissertation, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA.

Jigjisuren O, Baival B, Nayanaa K, Jargalsaikhan A, Dash K, Badamkhand B, Bud A, 2015. Evaluating the impact of climate change based on herder's observations and comparing it with hydroclimatic and remote sensing data. Proceedings of Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 235-243.

John R, Chen J, Ou-Yang, ZT, Xiao J, Becker R, Samanta A, Ganguly S, Yuan W, and Batkhishig O, 2013. Vegetation response to extreme climate events on the Mongolian Plateau from 2000 to 2010. Environmental Research Letters, 8, 035033, doi:10.1088/1748-9326/8/3/035033.

Karthe, D, Hofmann, J, Ibisch, R, Heldt, S, Westphal, K, Menzel, L, Avlyush, S and Malsy, M, 2015. Science-based IWRM implementation in a data-ccarce Central Asian region: Experiences from a research and development project in the Kharaa River Basin, Mongolia. Water, 7, 3486–3514, doi:2048/10.3390/w7073486.

Kendall MG and Gibbons JD, 1990. Rank Correlation Methods. 5th edition, Edward Arnold, Hodder and Stoughton, London, UK.

Kim B, Hossein S, and Choi G, 2011. Evaluation of temporal-spatial precipitation variability and prediction using seasonal ARIMA model in Mongolia. KSCE Journal of Civil Engineering, 15, 917–925, doi:10.1007/s12205-011-1097-9.

Klein Tank AMG, Peterson TC, Quadir DA, Dorji S, Zou X, Tang H, Santhosh K, Joshi UR, Jaswal AK, Kolli RK, Sikder AB, Deshpande NR, Revadekar JV, Yeleuova K, Vandasheva S, Faleyeva F, Gomboluudev P, Budhathoki KP, Hussain A, Afzaal M, Chandrapala L, Anvar H, Amanmurad D, Asanova VS, Jones PD, New MG, Spektorman T, 2006. Changes in daily temperature and precipitation extremes in central and south Asia. Journal of Geophysical Research, 111, D16105.

Komatsu G, Brantingham PJ, Olsen JW, and Baker VR, 2001. Paleoshoreline geomorphology of Böön Tsagaan Nuur, Tsagaan Nuur and Orog Nuur: The Valley of Lakes, Mongolia. Geomorphology, 39(3-4), 83–98.

Kundzewicz ZW and Robson AJ, 2004. Change detection in hydrological records—a review of the methodology. Hydrological Sciences Journal 49, 7–19, doi:10.1623/hysj.49.1.7.53993.

Lee MK, Lee YI, Lim HS, Lee JI, and Yoon HI, 2013. Late Pleistocene-Holocene records from Lake Ulaan, southern Mongolia: Implications for east Asian palaeomonsoonal climate changes. Journal of Quaternary Science, 28, 370–378, doi:10.1002/jqs.2626.

Leland C, 2011. Hydroclimatic variability in north-central Mongolia as inferred from tree rings. Unpublished MS Thesis, West Virginia University, Morgantown, WV, USA.

Leland C, Pederson N, Hessl A, Nachin B, Davi N, D'Arrigo R, and Jacoby G, 2013. A hydroclimatic regionalization of central Mongolia as inferred from tree rings. Dendrochronologia, 31, 205–215, doi:10.1016/j.dendro.2012.11.003.

Lesk C, Rowhani P, and Ramankutty N, 2016. Influence of extreme weather disasters on global crop production. Nature, 529, 84–87, doi:10.1038/nature16467.

Li J, Cook ER, D'Arrigo R, Chen F, and Gou X, 2009. Moisture variability across China and Mongolia: 1951–2005. Climate Dynamics, 32, 1173–1186, doi:10.1007/s00382-008-0436-0.

Li SG, Asanuma J, Kotani A, Davaa G, and Oyunbaatar D, 2007. Evapotranspiration from a Mongolian steppe under grazing and its environmental constraints. Journal of Hydrology, 333, 133–143, doi:10.1016/j.jhydrol.2006.07.021.

Li Y, Wang N, and Zhang C, 2014. An abrupt centennial-scale drought event and Mid-Holocene climate change patterns in monsoon marginal zones of East Asia. PLoS ONE, 9, e90241, doi:10.1371/journal.pone.0090241.

Lkhagvadorj D, Hauck M, Dulamsuren CH, and Tsogtbaatar J, 2013. Pastoral nomadism in the forest-steppe of the Mongolian Altai under a changing economy and warming climate. Journal of Arid Environments 88, 82-89.

Ma X, Yasunari T, Ohata T, Natsagdorj L, Davaa G, Oyunbaatar D, 2003. Hydrological regime analysis of the Selenge River basin, Mongolia. Hydrological Processes, 17, 2929-2945, doi:10.1002/hyp.1442.

Malsy M, Beek TAD, and Floerke M, 2015. Evaluation of large-scale precipitation data sets for water resources modelling in Central Asia. Environmental Earth Sciences, 73, 787–799, doi:10.1007/s12665-014-3107-y.

Mann HB, 1945. Nonparametric tests against trend. Econometrica, 13, 245–259.
Marin A, 2010. Riders under storms: Contributions of nomadic herder's observations to analyzing climate change in Mongolia. *Global Environmental Change,* 20, 162-176, doi:10.1016/j.gloenvcha.2009.10.004.

Meko DM, and Woodhouse CA, 2011. Chapter 8: Application of streamflow reconstruction to water resources management. *Dendroclimatology- Progress and prospects.* Hughes MK, Swetnam TW, and Diaz HF [Eds]. Springer Netherlands, Dordrecht, 231-261, ISBN:978-1-4020-5725-0.

Menzel L, Hofmann J, and Ibisch R, 2011. Studies of water and mass fluxes to provide a basis for an Integrated Water Resources Management (IWRM) in the catchment of the River Kharaa in Mongolia [primarily in German]. *Hydrologie und Wasserbewirtschaftung*, 55, 88–103.

Middleton N, Rueff H, Sternberg T, B Batbuyan, and Thomas D, 2015. Explaining spatial variations in climate hazard impacts in western Mongolia. *Landscape Ecology*, 30, 91–107, doi:10.1007/s10980-014-0091-2.

Minderlein S, and Menzel L, 2015. Evapotranspiration and energy balance dynamics of a semi-arid mountainous steppe and shrubland site in Northern Mongolia. *Environmental Earth Sciences*, 73, 593–609, doi:10.1007/s12665-014-3335-1.

Ministry of Nature, Environment, and Tourism, 2010. Mongolia Second National Communication, Dagvadorj D [Ed]. Ulaanbaatar, Mongolia, 160 pp.

Miyazaki S, Yasunari T, and Adyasuren T, 1999. Abrupt seasonal changes of surface climate observed in Northern Mongolia. *Journal of the Meteorological Society of Japan*, 77, 583–593.

Mote P, Brekke L, Duffy PB, and Maurer E, 2011. Guidelines for constructing climate scenarios. *EOS, Transactions, American Geophysical Union*, 92, 257–258.

Morinaga Y, Tian SF, and Shinoda M, 2003. Winter snow anomaly and atmospheric circulation in Mongolia. *International Journal of Climatology*, 23, 1627–1636, doi:10.1002/joc.961.

Nandintsetseg B, and Shinoda M, 2011. Seasonal change of soil moisture in Mongolia: Its climatology and modelling. *International Journal of Climatology*, 31, 1143–1152, doi:10.1002/joc.2134.

Nandintsetseg B, and Shinoda M, 2013. Assessment of drought frequency, duration, and severity and its impact on pasture production in Mongolia. *Natural Hazards*, 66, 995–1008, doi:10.1007/s11069-012-0527-4.

Nandintsetseg B, and Shinoda M, 2014. Multi-decadal soil moisture trends in Mongolia and their relationships to precipitation and evapotranspiration. *Arid Land Research and Management*, 28, 247–260, doi:10.1080/15324982.2013.861882.
Natsagdorj L, Jugder D, and Chung Y, 2003. Analysis of dust storms observed in Mongolia during 1937–1999. *Atmospheric Environment*, 37, 1401–1411, doi:10.1016/S1352-2310(02)01023-3.

Ojima D, and Chuluun T, 2008. Chapter 8: Policy changes in Mongolia: Implications for land use and landscapes. *Fragmentation in semi-arid and arid landscapes: Consequences for human and natural systems,* Galvin K, Reid R, Behnke Jr. R, Thompson Hobbs N [Eds]. Springer Publishers, Dordrecht, The Netherlands, 179-194.

Pederson N, Jacoby GC, D'Arrigo RD, Cook ER, Buckley BM, Dugarjav C, and Mijiddorj R, 2001. Hydrometeorological reconstructions for northeastern Mongolia derived from tree rings: 1651-1995. *Journal of Climate*, 14(5), 872-881, doi: 10.1175/1520-0442(2001)014.

Pederson N, Leland C, Nachin B, Hessl AE, Bell AR, Martin-Benito D, Davi NK, 2013. Three centuries of shifting hydroclimatic regimes across the Mongolian Breadbasket. *Agricultural and Forest Meteorology*, 178, 10–20.

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Førland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D, 1998. Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology* 18, 1493–1517, doi:10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T.

Pettitt AN, 1979. A non-parametric approach to the change point problem. *Applied Statistics* 28(2), 126–135.

Pielke Sr. RA, Stohlgren T, Schell L, Parton W, Doesken N, Redmond K, Moeny J, Mckee T, Kittel TGF, 2002. Problems in evaluating regional and local trends in temperature: An example from eastern Colorado, USA. *International Journal of Climatology*, 22(4), 421-434.

Poulter B, Pederson N, Liu H, Zhu Z, D'Arrigo R, Ciais P, Davi N, Frank D, Leland C, Myneni R, Piao S, and Wang T, 2013. Recent trends in Inner Asian forest dynamics to temperature and precipitation indicate high sensitivity to climate change. *Agricultural and Forest Meteorology*, 178, 31–45, doi:10.1016/j.agrformet.2012.12.006.

Priess JA, Schweitzer C, Wimmer F, Batkhishig O, Mimler M, 2011. The consequences of land-use change and water demands in Central Mongolia. *Land Use Policy* 28, 4-10.

Priess JA, Schweitzer C, Batkhishig O, Koschitzki T, and Wurbs D, 2015. Impacts of agricultural land-use dynamics on erosion risks and options for land and water management in Northern Mongolia. *Environmental Earth Sciences*, 73, 697-708, doi:10.1007/s12665-014-3380-9.

Regdel D, Dugarzhav C, and Gunin PD, 2012. Ecological demands on socioeconomic development of Mongolia under climate aridization. *Arid Ecosystems*, 2, 1–10, doi:10.1134/S2079096112010076.

Saladyga T, Hessl A, Nachin B, and Pederson N, 2013. Privatization, drought, and fire exclusion in the Tuul River Watershed, Mongolia. *Ecosystems*, 16, 1139–1151, doi:10.1007/s10021-013-9673-0.

Sato T, Kimura F., and Kitoh A, 2007a. Projection of global warming onto regional precipitation over Mongolia using a regional climate model. *Journal of Hydrology*, 333, 144–154, doi:10.1016/j.jhydrol.2006.07.023.

Sato T, Tsujimura M, Yamanaka T, Iwasaki H, Sugimoto A, Sugita M, Kimura F, Bavaa G, and Oyunbaatar D, 2007b. Water sources in semiarid northeast Asia as revealed by field observations and isotope transport model. *Journal of Geophysical Research*, 112, D17112, doi:10.1029/2006JD008321.

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, and Rudolf B, 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115, 15–40, doi:10.1007/s00704-013-0860-x.

Sen PK, 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association,* 63, 1379-1389.

Sharkuu N, 2003. Recent changes in the permafrost of Mongolia. Proceedings of the 2003 International Conference on Permafrost, Phillips, Springman and Arenson [Eds]. Swets & Zeitlinger, Lisse, 1029-1034, ISBN 90-5809-582-7.

Shinoda M, Nachinshonhor UG, and Nemoto M, 2010. Impact of drought on vegetation dynamics of the Mongolian steppe: A field experiment. *Journal of Arid Environments*, 74, 63–69, doi:10.1016/j.jaridenv.2009.07.004.

Shinoda M, Nandintsetseg B, Nachinshonhor UG, and Komiyama H, 2014. Hotspots of recent drought in Asian steppes. *Regional Environmental Change*, 14, 103–117, doi:10.1007/s10113-013-0464-0.

Siurua H, and Swift J, 2002. Drought and zud but no famine (yet) in the Mongolian herding economy. *IDS Bulletin*, 33, 88–97, doi:10.1111/j.1759-5436.2002.tb00048.x.

Star SL, and Griesemer JR, 1989. Institutional ecology, "translations" and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19, 387–420.

Star, SL, 2010. This is not a boundary object: Reflections on the origin of a concept. *Science Technology Human Values*, 35, 601–617, doi:10.1177/0162243910377624.

Sternberg T, Thomas D, and Middleton N, 2011. Drought dynamics on the Mongolian steppe, 1970–2006. *International Journal of Climatology*, 31, 1823–1830, doi:10.1002/joc.2195.

Sternberg T, and Paillou P, 2015. Mapping potential shallow groundwater in the Gobi Desert using remote sensing: Lake Ulaan Nuur. *Journal of Arid Environments*, 118, 21–27, doi:10.1016/j.jaridenv.2015.02.020.

Stockton CW, and Jacoby GC, 1976. Long-term surface-water supply and streamflow trends in the upper Colorado River Basin based on tree-ring analyses. *Lake Powell Research Project Bulletin Number 18*, March 1976. Collaborative Research on Assessment of Man's Activities in the Lake Powell Region. National Science Foundation, Washington, D.C.

Sukh T, 2012. Local understanding of hydro-climate changes in Mongolia. Unpublished MS Thesis, Colorado State University, Fort Collins, Colorado, USA.

Sun Z, Wang Q, Xiao Q, Batkhishig O, and Watanabe M, 2015. Diverse responses of remotely sensed grassland phenology to interannual climate variability over frozen ground regions in Mongolia. *Remote Sensing*, 7, 360–377, doi:10.3390/rs70100360.

Tachiiri K, Shinoda M, Klinkenberg B, and Morinaga Y, 2008. Assessing Mongolian snow disaster risk using livestock and satellite data. *Journal of Arid Environments*, 72, 2251–2263, doi:10.1016/j.jaridenv.2008.06.015.

Tao S, Fang J, Zhao X, Zhao S, Shen H, Hu H, Tang Z, Wang Z, Guo Q, 2015. Rapid loss of lakes on the Mongolian Plateau. *Proceedings of the National Academy of Sciences*, 112, 2281–2286, doi:10.1073/pnas.1411748112.

Thiel H, 1950. A rank-invariant method of linear and polynomial regression analysis, I, II, III. *Proceedings of the Royal Netherlands Academy of Sciences* 53, 386-392, 521-525, 1397-1412.

UNSD, 2012. United Nations Statistics Divison, *UNSD Demographic Statistics*, *http://data.un.org/Data.aspx?d=POP&f=tableCode:26#POP*. Accessed February 08, 2016.

Upton C, 2010. Wells, water rights and sustainability: new policies and practices on Mongolia's pastoral commons. *Water, Environmental Security and Sustainable Rural Development: Conflict and Cooperation in Central Eurasia*. Arsel M, and Spoor M [Eds]. 1st edition, Routledge, London, New York, 304 pp.

Venable NBH and Fassnacht SR, 2012. *Characterizing two semi-arid river basins near the Khangai Mountains of Mongolia through streamflow sampling*. American Center for Mongolian Studies, US-Mongolia Field Research Fellowship Program Final Report, Ulaanbaatar, Mongolia, 12 pp.

Venable NBH, Fassnacht SR, Adyabadam G, S Tumenjargal, Fernandez-Gimenez M, and B Batbuyan, 2012a. Does the length of station record influence the warming trend that is perceived by Mongolian herders near the Khangai Mountains? *Pirineos*, 167, 69–86, doi:10.3989/Pirineos.2012.167004.

Venable NBH, Fassnacht SR, Adyabadam G, and S Tumenjargal, 2012b. Perceived climatic warming and drying near the Khangai Mountains, Mongolia, explored through station record length analysis. Poster session, *32nd Annual American Geophysical Union Hydrology Days*, Colorado State University, Fort Collins, CO, March 23, 2012.

Venable NBH, Fassnacht SR, S Tumenjargal, B Batbuyan, J Odgarav, J Sukhbataar, Fernandez- Gimenez ME, and Adyabadam G, 2012c. Perspectives on hydro-climatic change in rivers sourced from the Khangai Mountains, Mongolia. Oral presentation, *2012 American Geophysical Union Fall Meeting*, San Francisco, CA, December 3, 2012.

Venable NBH, Fassnacht SR, and Adyabadam G, 2014a. Comparison of point and gridded data sources for use in hydroclimatic modeling of the Khangai Mountain region, Mongolia. Poster session, *34th Annual American Geophysical Union Hydrology Days*, Colorado State University, Fort Collins, CO, March 26th, 2014.

Venable NBH, Fassnacht SR, and Adyabadam G, 2014b. To Grid or not to grid… Precipitation data and hydrological modeling in the Khangai Mountain region of Mongolia. Oral presentation, *2014 American Geophysical Union Fall Meeting*, San Francisco, CA, December 15-19, 2014.

Venable NBH, Fassnacht SR, and Hendricks AS, 2015. Spatial changes in climate across Mongolia. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig , Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 73-77.

Wang X, Yang X, Liu T, Li F, Gao R, Duan L, and Luo Y, 2014. Trend and extreme occurrence of precipitation in a mid-latitude Eurasian steppe watershed at various time scales. *Hydrological Processes*, 28, 5547–5560, doi:10.1002/hyp.10054.

Willmott CJ, and Matsuura K, 2012. Terrestrial air temperature and precipitation: Gridded monthly time series (1900 - 2010), *http://climate.geog.udel.edu/~climate/html_pages/Global2011/README.GlobalTsT2011.html and GlobalTsP2011*. Accessed 2011-12-20.

Wolf JM, and Venable NBH, 2015. Earlywood, latewood, and adjusted latewood correlations to precipitation: A test case from the Khangai Mountains of Mongolia. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 87-93.

Woodhouse CA, and Lukas JJ, 2006. Multi-century tree-ring reconstructions of Colorado streamflow for water resource planning. *Climatic Change*, 78, 293–315, doi:10.1007/s10584-006-9055-0.

Woodhouse CA, Gray ST, and Meko DM, 2006. Updated streamflow reconstructions for the Upper Colorado River Basin. *Water Resources Research*, 42, W05415, doi:10.1029/2005WR004455.

Yang B, Kang S, Ljungqvist FC, He M, Zhao Y, and Qin C, 2014. Drought variability at the northern fringe of the Asian summer monsoon region over the past millennia. *Climate Dynamics*, 43, 845–859, doi:10.1007/s00382-013-1962-y.

Yu F, Price KP, Ellis J, and Shi P, 2003. Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sensing of Environment*, 87, 42–54, doi:10.1016/S0034-4257(03)00144-5.

Yue S, Pilon P, Phinney B, Cavadias G, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16, 1807–1829, doi:10.1002/hyp.1095.

Zhang Y, Ohata T, Yang D, and Davaa G, 2004. Bias correction of daily precipitation measurements for Mongolia. *Hydrological Processes*, 18, 2991–3005, doi:10.1002/hyp.5745.

CHAPTER 2- SPATIAL CHANGES IN CLIMATE ACROSS MONGOLIA[1]

2.1 Summary

Previous research using meteorological station data suggests that temperatures and precipitation have been changing more across the semi-arid and arid country of Mongolia than in many other locations across the globe. We used gridded monthly data to determine the annual and seasonal rate of change in total precipitation (P), maximum temperature (Tmax), and minimum temperature (Tmin), as computed from the non-parametric Thiel-Sen slope estimator method. The significance of those changes were computed from the Mann-Kendall test. The University of East Anglia Climatic Research Unit (CRU) dataset was used for the 50-year time period from 1963 through 2012 at a 0.5 degree (~55 km) resolution. For the first 30 years, 30 to 35 meteorological stations from across Mongolia were used to create the spatially distributed "High Resolution Gridded Data of Month-by-Month Variation in Climate" CRU product; 20 to 30 stations were used for the last 20 years due to a decrease in the number of operational stations. Results are presented as maps of 1) mean total annual P, and mean annual Tmax and Tmin, and ii) annual trends over the length of record (1963-2012) with significance overlain, for the three variables. Rates of change at annual and seasonal time scales varied spatially with more consistent increases in temperature; significant precipitation trends were observed over smaller areas than significant temperature trends.

[1] Chapter 2 was published in the Proceedings of the Trans-disciplinary Research Conference: Building Resilience of Mongolian Rangelands, Ulaanbaatar, Mongolia , June 9-10, 2015. Additional authors were S.R. Fassnacht and A. D. Hendricks.

2.2 Background

The climate of Mongolia is characterized as semi-arid with cold winters, warm summers and strongly seasonal precipitation patterns. Over the last few decades, a myriad of sources of change have affected traditional nomadic pastoralist lifestyles. These include, but are not limited to, major socio-economic and political changes and an increasingly warming and potentially drying landscape with changes in the frequency and severity of extreme climatic events, such as winter disasters or *dzud* (Batima *et al*., 2005; Fernandez-Gimenez *et al*., 2015). Simultaneously, changes are occurring such as the promotion of more intensive agricultural land uses with trends toward less nomadic practices, and increasing urban, industrial, and mining development (Ojima and Chuluun, 2008; Yamamura *et al*., 2013). Herders have observed changes to their environment and many of these observations correlate well with available climatic data (Marin, 2010; Fassnacht *et al*., 2011; Venable *et al*., 2012; Lkhagvadorj *et al*., 2013).

Climate research in Mongolia as reported in the international literature is limited, but there is a general consensus regarding the occurrence of increasing trends in mean, warm, and cool season temperatures but a relative lack of countrywide patterns in precipitation change (e.g. Batima *et al*., 2005; Ministry of Nature, Environment, and Tourism, 2010; Jamiyansharav, 2010). Longer-term climate research using climate proxies suggests that the most recent decade may be one of warming and drought unlike that seen over the last millennia (Pederson *et al*., 2014). Given the uncertainties inherent in station-based data due to sparse station availability and the amount of missing data, we chose to examine Mongolian climate trends using spatially and temporally coherent gridded datasets. Our work extends existing research by investigating climate variability seasonally and annually using data from across the country at spatial resolutions beyond the point-based station level.

2.3 Methods

Monthly maximum and minimum temperature (Tmax and Tmin) and precipitation (P) grids from the Climatic Research Unit (CRU) Time series 3.21 were acquired from the British Atmospheric Data Centre (Harris *et al.*, 2014). The 0.5 longitude by 0.5 latitude grids are interpolated from anomalies of station data (1961-1990 means) and then combined with existing climatologies to give absolute monthly grid values. The data were provided to CRU primarily by the World Meteorological Organization, via the Mongolian Institute for Meteorology, Hydrology and the Environment. While the number of contributing stations varies depending on the presence or absence of recorded data, generally 30 to 35 stations contributed from 1963 up to 1990 with a slight decrease in stations (20-30) recording to 2012.

The gridded files of monthly P, Tmax, and Tmin were compiled annually and seasonally (winter is December through February, spring is March through May, summer is June through August, and fall is September through November) for the length of record (1963-2012). The Thiel-Sen estimator for slope and the Mann-Kendall test for significance of trend (Gilbert, 1987) were calculated for the aggregate (total for P, and mean for Tmax and Tmin) time series at each grid cell using R statistical software (R Core Team, 2014).

2.4 Results

2.4.1 Climate Patterns

Mean annual total P, Tmax and Tmin are shown in Figure 2.1 for the 50-year length of record. Precipitation gradients are noticeable from north to south and in mountainous versus valley regions of the country (Figure 2.1a). Similar patterns are noticeable for Tmax and Tmin (Figures 2.1b and 2.1c), with cooler temperatures on average in the mountainous and more northerly portions of the country.

2.4.2 Annual Trends

On an annual basis, significantly decreasing precipitation trends occur in the eastern and central parts of the country (ranging from about -8 to -21 mm/decade), with slight decreases in parts of the far south-central Gobi region (from -7 and -9 mm/decade) and the northwestern part of the country near Zavkhan *aimag* (province), (between -12 and -14 mm/decade) (Figure 2.2a). Changes in annual mean maximum temperature through time were found to be significant across most of the country (ranging from 0.2 to 0.6 deg C/decade) (Figure 2.2b). Trends in annual mean minimum temperature were significant across the entire country (ranging from 0.2 to 0.7 deg C/decade) with the greatest rates of change occurring in the north central part of the nation near Lake Khovsgol (0.7 deg C/decade), across parts of the western Altai Mountains and Great Lakes region (up to 0.6 deg C/decade), and in the far eastern steppe (from 0.5 to 0.6 deg C/decade). Increasing mean minimum temperature trends were also greater in the central Gobi region (up to 0.5 deg C/decade) (Figure 2.2c).

2.4.3 Seasonal Trends

Seasonal trends (not shown) are more spatially diverse than the annual trends. Changes in precipitation are generally not significant across a majority of the country particularly in the winter and spring months. In summer however, significant decreases in precipitation are seen from the central northwest across the central forest steppe to the eastern steppe, similar in magnitude and location to and of somewhat greater extent than those patterns illustrated in the annual trend (Figure 2.2a). Decreasing trends in fall precipitation are centered in a region extending east from near the eastern edge of Khovsgol *aimag* across the central forest steppe and steppe to the area west of Ulaanbaatar (from -2 to -7 mm/decade).

Trends in mean maximum winter temperatures are not significant over a majority of Mongolia. Fall mean maximum temperature changes over the period of record are also generally not significant with the exception of the northwestern part of the country (up to 0.6 deg C/decade), the south central Gobi region (up to 0.3 deg C/decade) and areas south of the Khangai Mountains (from 0.3 to 0.4 deg C/decade). Increases in mean maximum fall temperatures are also seen north and west of Ulaanbaatar (up to 0.4 deg C/decade). Significant increases in spring and summer mean maximum temperatures are observed across most areas of the country (up to 0.4 deg C/decade in the summer), except for in the central Gobi (not significant in springtime) and parts of the eastern steppe.

Minimum mean temperature has been warming the most and has the largest extent of significant change throughout the seasons. While minimum winter and fall temperatures have increased significantly mainly over the western, central-southern and eastern portions of the country (overall ranges from 0.3 to 0.8 deg C/ decade in winter, and 0.2 to 0.7 deg C/decade increase in fall), minimum spring and summer temperatures have increased significantly across the entire nation. Only one small area of the far western Khangai Mountain region has not seen a significant warming of minimum spring temperatures over the last 50 years. Similar rates of change are seen in both spring and summer seasons (from 0.3 to 0.7 deg C/decade spring, 0.2 to 0.5 deg C/decade summer).

2.5 Discussion and Conclusions

The results of previous climate trend analyses by other authors often parallel the gridded results presented here. There are however, key differences. For example, when studying mean seasonal temperatures, Batima *et al*., (2005) concluded that primarily winter temperatures were increasing with increases also in spring and fall. They did not find clear increasing or decreasing trends in summer, though they found evidence of longer durations of periods with hot days. They did not find significant changes in seasonal precipitation, though they did acknowledge strong spatial variability in the precipitation results. Our analyses of these gridded datasets reveal clear temperature increases in summer, particularly for minimum temperatures and significant changes (decreasing) in precipitation in the summer and fall for nearly a quarter to a third of the country, depending on the season.

Other analyses of climate are more difficult to compare to our results due to the use of climate indices rather than explicit values of P, Tmax, and Tmin. In the *Mongolia Second National Communication* document (Ministry of Nature, Environment, and Tourism, 2010), results are presented as an increase in the frequency of extreme high temperatures and a drop in the occurrence of extreme low temperatures. Increases in winter precipitation are also mentioned with decreases in summer precipitation across the country. These results are somewhat correlative to the increasing mean maximum and minimum temperatures, as well as the significant decreases in summer precipitation shown herein.

Inherent in the use of climatic datasets are uncertainties introduced due to data collection, processing, and in the case of gridded datasets, the interpolation of climate data. Jamiyasharav (2010) documented biases that may be present in the Mongolian climate records due to station siting, movements of station locations, and changes to instrumentation. Whether the differences between existing studies and our results are artifacts of the original station data, the interpolation processes used in gridding the climate variables, or differences in trend analyses methodology, or all of these, it is clear that the historical climate record exhibits significant change over a 50-year period from 1963 through 2012.

Spatial trend analyses at annual and seasonal time steps using gridded datasets (e.g. Hendricks and Fassnacht, *in preparation*) provide a strong visual tool for examining significant climate change across Mongolia. These results suggest that significant warming trends and some drying trends are present in areas of the country that support much of the population. Mitigating adverse impacts from these changes will be particularly challenging under increasing agricultural and water-resource intensive mining development in these regions.

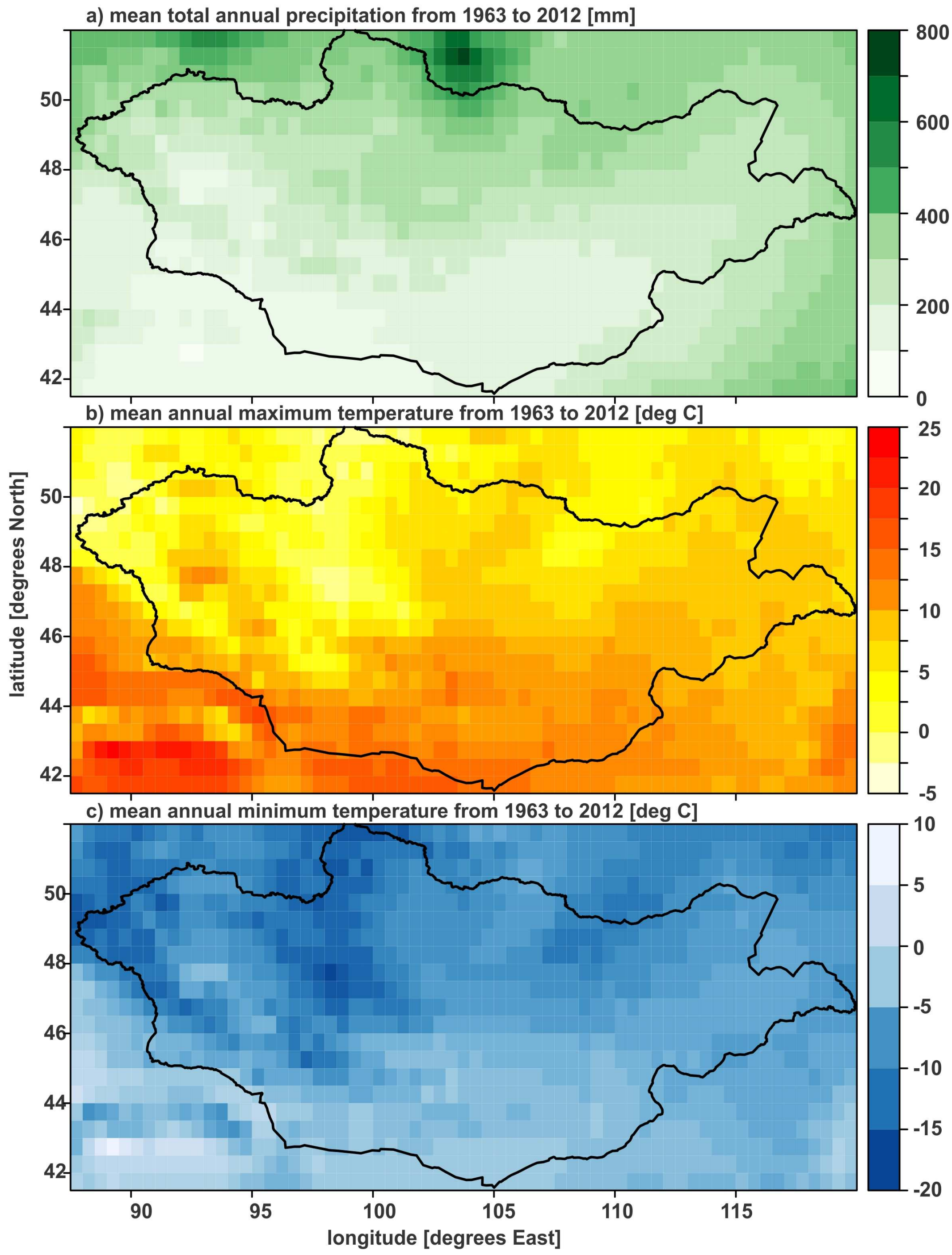


FIGURE 2.1- Mongolia mean annual (1963-2012) a) total precipitation in millimeters, b) maximum temperature in degrees Celsius, c) minimum temperature in degrees Celsius.

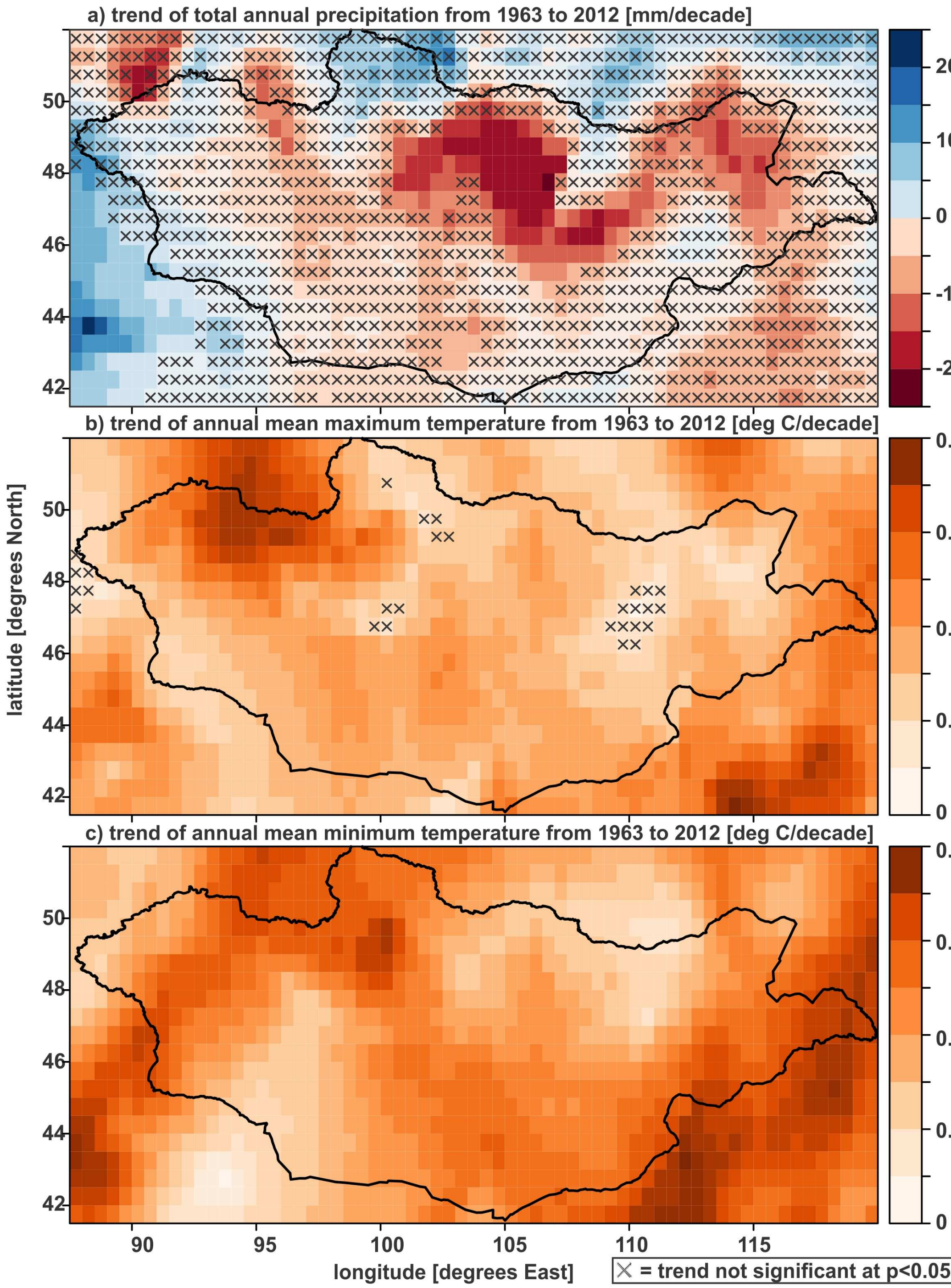


FIGURE 2.2- Trends per decade in the annual mean a) total precipitation in millimeters, b) maximum temperature in degrees Celsius, and c) minimum temperature in degrees Celsius. Note that the X's in the figure denote areas where the trend was not significant at the $p<0.05$ level.

REFERENCES

Batima P, Natsagdorj L, Gombluudev P, Erdenetsetseg B, 2005. Observed climate change in Mongolia. *Assessments and Adaptations to Climate Change* (AIACC), *www.aiaccproject.org*. Accessed 2011-11-21.

Fassnacht SR, Sukh T, Fernandez-Gimenez ME, B Batbuyan, Venable NBH, Laituri M, and Adyabadam G, 2011. Local understanding of hydro-climatic changes in Mongolia. Proceedings of *H02 Symposium Cold Region Hydrology in a Changing Climate* held during IUGG2011 in Melbourne, Australia, July 2011, *IAHS Publication* 346, 120-129.

Fernandez-Gimenez ME, B Batkhishig, B Batbuyan, and Ulambayar T, 2015. Lessons from the *dzud*: Community-based rangeland management increases the adaptive capacity of Mongolian herders to winter disasters. *World Development* 68, 48-65, doi:10.1016/j.worlddev.2014.11.015.

Gilbert RO, 1987. *Statistical Methods for Environmental Pollution Monitoring.* John Wiley & Sons, New York, 320 pp.

Harris I, Jones PD, Osborn TJ, and Lister DH, 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, 34, 623–642, doi:10.1002/joc.3711.

Jamiyansharav K, 2010. *Long-term analysis and appropriate metrics of climate change in Mongolia*. Unpublished PhD Dissertation, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA.

Lkhagvadorj D, Hauck M, Dulamsuren CH, and Tsogtbaatar J, 2013. Pastoral nomadism in the forest-steppe of the Mongolian Altai under a changing economy and warming climate. *Journal of Arid Environments* 88, 82-89.

Marin A, 2010. Riders under storms: Contributions of nomadic herder's observations to analyzing climate change in Mongolia. *Global Environmental Change,* 20, 162-176, doi:10.1016/j.gloenvcha.2009.10.004.

Ministry of Nature, Environment, and Tourism, 2010. Mongolia Second National Communication, Dagvadorj D [Ed]. Ulaanbaatar, Mongolia, 160 pp.

Ojima D, and Chuluun T, 2008. Chapter 8: Policy changes in Mongolia: Implications for land use and landscapes. *Fragmentation in semi-arid and arid landscapes: Consequences for human and natural systems,* Galvin K, Reid R, Behnke Jr. R, Thompson Hobbs N [Eds]. Springer Publishers, Dordrecht, The Netherlands, 179-194.

Pederson N, Hessl AE, Baatarbileg N, Anchukaitis KJ, and Di Cosmo N, 2014. Pluvials, droughts, the Mongol Empire, and modern Mongolia. *Proceedings of the National Academy of Sciences*, 111, 4375–4379, doi:10.1073/pnas.1318677111.

R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www.R-project.org/.*

Venable NBH, Fassnacht SR, Adyabadam G, S Tumenjargal, Fernandez-Gimenez M, and B Batbuyan, 2012. Does the length of station record influence the warming trend that is perceived by Mongolian herders near the Khangai Mountains? *Pirineos*, 167, 69–86, doi:10.3989/Pirineos.2012.167004.

Yamamura N, Fujita N, and Maekawa A, 2013. The Mongolian ecosystem network – Environmental issues under climate and social changes. *Ecological Research Monographs*, Iwasa Y [Ed]. Springer Publishers, Tokyo, Japan, 317 pp.

CHAPTER 3- DOES THE LENGTH OF STATION RECORD INFLUENCE THE WARMING TREND THAT IS PERCEIVED BY MONGOLIAN HERDERS NEAR THE KHANGAI MOUNTAINS?[2]

3.1 Summary

Temperatures changes can be difficult to infer from changes in vegetation patterns or other ecological changes, yet warming can be inferred through changes in the habits of people who live in close connection with their natural environment. Herders near the Khangai Mountains of central Mongolia have perceived a warming trend in recent years. Since it is difficult to determine the exact time period over which perceived warming has occurred, the statistical differences in temperature changes were examined based on the length of data and the specific period of record used in the analysis. Temperature data from five meteorological stations for up to 50 years (1961-2010) were used to examine trends in varying lengths of record from 15 to 50 years with varying start periods (1961 through 1986), based on the length of record. The most statistically significant temperature changes occurred for the longest time periods and for annual average minimum temperatures. One very cold winter, 2009-2010, decreased the warming trend and for shorter periods of record reduced the statistical significance.

3.2 Background

Mongolia is a country known for its vast rangelands. Over 83% of the country is dominated by this type of land cover/use (Angerer *et al.*, 2008). Recent threats to rangeland health include changes in the traditional nomadic use and management of pastureland, increasing pressure from changes in livestock numbers and composition of herds, and

[2] Chapter 3 was published in 2012 in *Pirineos* 167: 69-86. Additional authors were S.R. Fassnacht, G. Adyabadam, Tumenjargal S., M.E. Fernandez-Gimenez, and Batbuyan B.

desertification (Angerer *et al.*, 2008; Ministry of Nature, Environment, and Tourism, 2010). Other environmental threats include increasing pollution from urban expansion, mining, and climate change (Ministry of Nature, Environment, and Tourism, 2010).

Steppe vegetation condition is critical to herder livelihoods. Climate variability has altered vegetation (e.g. Yu *et al.*, 2003; Angerer *et al.*, 2008); specifically plant production has been limited by temperature and water supply changes (Yu *et al.*, 2003). These changes are particularly critical in mountain steppe areas, which provide highly productive forage for grazing livestock and a source of water supply for more arid desert steppe rangelands downstream (Angerer *et al.*, 2008; Fassnacht *et al.*, 2011).

The water supply for most herders and smaller communities comes from resources that are influenced by climate variability. Approximately 36 percent of the country's population relies on water from shallow groundwater wells and 10 percent rely on water from rivers, with the remainder of the population using municipal water supplies derived mainly from groundwater (Ministry of Nature, Environment, and Tourism, 2010). Access to these resources is critical for the daily needs of pastoralists, including their livestock. Herders have reported diminished volumes in lakes and rivers and lowering groundwater levels (Fassnacht *et al.*, 2011). These decreases may be linked to a variety of factors such as irrigation and agriculture needs, mining, and climate changes (Ministry of Nature, Environment, and Tourism, 2010).

Many of Mongolia's rivers flow from headwaters in the country's mountain ranges, including the Khangai Range in the west-central region (Figure 3.1). Between 56 and 75 percent of the country's annual runoff is derived from the mountains, mostly as rainfall, and the mountains contribute a significant portion of the groundwater baseflow that has a relatively short residence time (Ministry of Nature, Environment, and Tourism, 2010). Precipitation in Mongolia is spatially variable by elevation and latitude, with about 50 mm falling annually on Gobi desert regions to 400 mm in northern mountain regions (Ministry of Nature, Environment, and Tourism 2010). It is characterized by relatively intense rainfall events with most of the precipitation occurring between April and September (Ministry of Nature, Environment, and Tourism, 2010). Due to this spatial variability, trends in precipitation are not regionally consistent. Precipitation amounts and trends can vary significantly from station to station, especially in the more arid regions of the country (Batima *et al.*, 2005; Jamiyansharav, 2010).

Climate changes, such as variations in precipitation amounts and the occurrence of precipitation, may be inferred by changes in vegetation patterns or other ecological changes. It is more difficult to quantify changes in temperature. However, temperature changes, such as warming through an increase in minimum temperatures, can be inferred through changes in the habits of people who live in close connection with their natural environment (Alexander *et al.*, 2011).

Temperature trends play an indirect but critical role in the water budget, by affecting evapotranspiration, precipitation timing and distribution, and/or phase of precipitation. Potential evapotranspiration exceeds precipitation throughout Mongolia, and increasing temperatures may exacerbate this effect (Ministry of Nature, Environment, and Tourism 2010). In the drier steppe regions, early season warming raises evapotranspiration rates delaying the onset of spring green up, especially when rains are delayed, increasing the risk of rangeland degradation (Yu *et al.*, 2003; Jamiyansharav, 2010). For those areas with streamflow reliant on mountain snowmelt, changes in precipitation phase can affect the seasonal timing of streamflow (Barnett, *et al.*, 2005a). Therefore, the examination of temperature change trends is necessary to understand changes in hydrology and other resources.

3.2.1 Mongolian Temperature Regimes

Mongolia has a four-season climate with hot summers and long cold winters. Fluctuations in temperature from season to season, from day to day, and even over a single day can be great. Minimum temperatures can be colder than -52 degrees Celsius and maximums can be warmer than 43 degrees Celsius (Fassnacht *et al.*, 2011). Periods of extended heat and drought combined with subsequent long periods of cold and winter storms result in *dzud* (winter disaster) events that have caused large animal and even some herder mortality, loss of billions of dollars, and have severely impacted herder's livelihoods (Angerer *et al.*, 2008; Ministry of Nature, Environment, and Tourism, 2010; Marin, 2010).

Our planet's climate is changing, and changes in global temperatures are well documented (e.g. IPCC, 2007; Hansen *et al.*, 2010). Indices of temperature extremes for central and southern Asia suggest that there is a decrease in diurnal temperature ranges, and an increase in the number of warm nights/days for the period from 1961 to 2000 (Klein Tank *et al.*, 2006). Temperatures in Mongolia follow this trend, with warming recorded across the country (Batima *et al.*, 2005; Jamiyansharav, 2010). Central Mongolia has experienced long-term temperature increases of 2 to 4 degrees Celsius per century (Jamiyansharav, 2010). Potential problems for pastoralists include rapid warming then re-freezing periods in winter leading to ice-crusted rangeland, increases in drought conditions, and possible increasing *dzud* occurrence (Batima *et al.*, 2005; Ministry of Nature, Environment, and Tourism, 2010).

Global and large regional trends vary from local data gained from individual weather stations (Pielke *et al.*, 2002). In Mongolia, most long-term meteorological stations are located at an *aimag* (province) center or occasionally at a *soum* (county) center, so each station typically represents approximately 100,000 km^2. Temperature and the magnitude of trends are expected to vary from location to location due to spatial heterogeneity of landscape characteristics and seasonality. For example, in a study of data from 17 Mongolian meteorological stations, annual winter maximum temperatures increased at 14 stations, but decreased at three others (Jamiyansharav, 2010). The indigenous climate knowledge of pastoralists, especially nomadic people, may help bridge the spatial gap between local, point-collected weather data and the influence of large regional trends (Marin, 2010).

Since changes in resource availability are of utmost importance to pastoralists (Batima *et al.*, 2005), surveys have been conducted to collect indigenous knowledge covering a broad range of climatic change topics and the occurrence of extreme events (e.g., Marin, 2010; Fassnacht *et al.*, 2011). These surveys describe changes in Mongolian herder lifestyle as they have attempted to adapt to environmental changes (Marin, 2010), and they provide context for meteorological data analyses (Fassnacht *et al.*, 2011).

3.2.2 Herder Observations

Temperature changes over the past decades observed by herders were referenced to changes in seasonal extremes. These were exemplified by questions about warmer and/or cooler days and nights, and the onset of seasonal changes such as snowmelt. Specific statements from surveys in 2010 included that summer days seemed hotter yet summer nights were cooler, and winters seemed warmer than those in the past (Fassnacht *et al.*, 2011). Recent surveys in 2011 in Bayankhongor *aimag* suggested that spring temperatures were cooler than in the past but that snow amounts had decreased with earlier melting. One herder commented that “*there was a lot of snow this year*” (winter of 2010-2011), but another mentioned that snowmelt patterns have “*changed a lot*” compared to 20 years ago. Though some herders in the recent surveys perceived that winters now were cooler than in the past, one commented that, “*winter is like fall - no real winter*.” This observation is consistent with those from previously conducted surveys. The herders have observed temperature changes, but how far back in time is their period of reference and how does a variable period of record influence the station-based trends?

This paper aims to determine how the significance and rate of temperature trends are affected by the duration of the record, i.e., number of years, and the specific period of record, i.e., the start and end years. The available period of instrumented record will be used for each station. An individual extreme event is examined to determine if it will alter the statistical period of record analysis. Specifically, it is asked whether a single cold winter will reduce the trend and/or statistical significance of increasing temperature.

3.3 Study Areas

This study examines local scale temperature variability and focuses on stations in three *aimags* (provinces) in central Mongolia that represent both the mountain steppe and desert steppe regions. The Khangai Mountains divide the *aimags*; Arkhangai is to the north, Bayankhongor is south, and Uvurhangai (alternate spelling, Ovorkhangai) is located to the east. Five meteorological stations are used: Erdenemandal and Tsetserleg (Arkhangai *aimag*), Arvaikheer (alternate spelling, Arvaiheer) (Uvurkhangai *aimag*), and Bayankhongor and Horiult (alternate spelling, Khoriult) (Bayankhongor *aimag*) (Figure

3.1). The three northern-most stations are in the mountain steppe of the Khangai Mountains, Arvaikheer is in a more arid eastern edge of the mountain region while Horiult is in the desert steppe. The Erdenemandal and Tsetserleg stations are associated with Ikhtamir *soum* (county), while Bayankhongor and Horiult are associated with Jinst *soum*, as presented in the herder surveys by Fassnacht *et al.* (2011).

3.4 Methods

Meteorological data for this study were collated from two sources. Daily minimum and maximum temperature data for the Erdenemandal, Tsetserleg, Bayankhongor, and Horiult stations were obtained from the Mongolian Institute of Meteorology and Hydrology <http://www.icc.mn/Meteoins/index.html>. The longest length of record was 50 years (1961-2010) for Tsetserleg station, with a similar length of record for Erdenemandal (started in 1964) and Bayankhongor (started in 1963). Data collection at Horiult started in 1976. The Arvaikheer temperature data were obtained from the Global Historical Climatology Network of the National Climatic Data Center managed by the United States National Oceanic and Atmospheric Administration <http://www.ncdc.noaa.gov/ghcnm/>. There were substantial gaps in the daily maximum temperature data at Arvaikheer so only the minimum temperature data were used. Continuous data were available from 1969 through 1998 (Figure 3.2).

Annual average minimum and maximum temperatures were computed from daily data when there were less than 15 days of missing data. Typically periods of missing data were continuous and a month or more in duration. These temperature time series were analyzed using the non-parametric Mann-Kendall test for a monotonic increasing or decreasing trend with slope estimates or rates of change in degrees per century derived from the non-parametric Sen's method. The Mann-Kendall test determines the significance level of the change and is often reported at the 0.1, 1, 5 and 10% levels. For this study 1% ($p<0.01$), 5% ($p<0.05$) and 10% ($p<0.10$) significance levels are reported. As these methods are non-parametric, they allow for missing values, are robust with regards to not being biased by outliers, and the data do not need to conform to any particular type of distribution (Gilbert, 1987).

To assess the significance of trends of different record lengths and time periods, calculations were run for record lengths of 50 to 15 years at decreasing five year intervals. Time periods analyzed ranged from the longest record of 1961 to 2010 to the shortest period of 1996 to 2010. The length of record analyzed and time period were dependent upon the data available. These station-based analyses were performed in light of local traditional knowledge. Quantitative hydro-climate surveys and open-ended discussion were conducted with herders prior to and during the summer of 2010 (Fassnacht *et al.*, 2011). Herders with at least 14 years of experience in the field ranging in age from 30 to 78 years of age were asked to provide perceptions of climate change over the last 20 years. Temperature changes were assessed through examination of changes in seasonal conditions and length/timing of seasons (Fassnacht *et al.*, 2011).

Very cold winters in Mongolia are one type of extreme winter event called a *dzud* (Begszuren *et al.*, 2004). The winter of 2009-2010 was quite cold in the area of the Khangai Mountains in Mongolia (Fernandez-Gimenez *et al.*, 2011). Thus we compared the statistical analysis up to 2010 to the same length of record shifted two years earlier, i.e., up to 2008, to examine how the cold *dzud* of 2009-2010 changed the station analysis.

3.5 Results

The most significant trends were for the daily minimum temperatures (Table 3.1). These increasing trends were generally warmer (more) for the minimum than the maximum temperatures. The average temperatures (not shown) were statistically a combination of the minimum and maximum temperatures. Overall, the warming trend of longer time periods and including the more recent years (up to 2010) were the most statistically significant and often this warming trend was greater (Table 3.1).

Tsetserleg station has the longest records, and it exhibited highly significant changes, at the $p<0.001$ statistical significance level (not shown), for lengths of record from 30 to 50 years. Erdenemandal followed a similar pattern with highly significant trends from 25 to 45 years, with one period of 20 years being highly significant from 1981 to 2000. Trends at Bayankhongor were also highly significant from lengths of 30 to 40 years. For stations with shorter lengths of record, only the longest periods of 35 years at Horiult and 30 to 35 years at Arvaikheer were highly significant (Table 3.1).

The statistically significant rates of temperature change at the $p<0.001$ to $p<0.1$ levels for each station varied based on the period of record analyzed. For Erdenemandal the greatest rate of warming was for the 15 year period from 1971 through 1985, while the lowest rate of change was for a 25 year period from 1986 to 2010. Similarly at Tsetserleg, the highest rate of change was for a 15 year period from 1981 through 1995, while the lowest rate of change was a 35 year period from 1961 to 1995 (Figure 3.3). Several negative rates of change were noted for the maximum temperature at all stations and at Tsetserleg, Bayankhongor and Horiult for the minimum temperature. Most of these cooling trends were for shorter periods of record (15 to 20 years). Only the maximum temperature at

Horiult for the 15 year period from 1991 to 2005 was statistically significant at the $p<0.1$ level. The trend decreased during that time period at a rate of 15.1 degrees per century, while it increased at a rate of 11.3 degrees per century ($p<0.05$ significance level) for the same length of record 10 years earlier (1981 to 1995).

The rate of change for Bayankhongor was highest for a 15 year period (1966 to 1980) for both the maximum and minimum temperatures, yet the lowest rate of change for that station was a 25 year period (1971 to 1995). Similarly at Horiult, the greatest significant warming of minimum temperatures was a 15 year period (1986 to 2000) while the lowest rate of change was a 25 year period starting the same year (1986 to 2010). The same patterns exist for the minimum temperatures at Arvaikheer where the highest warming was a 15 year period (1966 to 1980) and the lowest was a 20 year period (1971 to 1990).

In Arkhangai *aimag*, the annual minimum temperatures for 2009 and 2010 were the coolest in the decade, except 2005 (Figure 3.2a and 3.2b). The average of the November through March daily maximum and minimum temperatures were the coldest on record at Tsetserleg and the fourth coldest at Erdenemandal. At the latter station 2005 was almost as cold and was the only other winter in the previous three decades that was similarly cold. At Tsetserleg only the winter of 1968-1969 had more individual cold days than 2009-2010. At Erdenemandal there were more cold days in 1968-1969, 1976-1977, and 2004-2005 than during the winter 2009-2010. The rate of temperature increase was higher for the periods ending in 2008 compared to those ending in 2010 (Table 3.2). When the increase was statistically significant it was greater at Tsetserleg than Erdenemandal. At shorter lengths of record, significance was greater for the periods up to 2008 (Table 3.2).

3.6 Discussion

The most significant temperature trends were for the daily minimum temperatures (Table 3.1). This is consistent with other studies in semi-arid regions of the world (e.g., Pielke *et al.*, 2002), including Mongolia, where research has documented increasing winter temperatures, warming annual minimum temperatures and decreasing numbers of cold days with minimum air temperatures below -5 degrees C (e.g. Batima *et al.*, 2005; Jamiyansharav, 2010; Ministry of Nature, Environment, and Tourism, 2010). Changes in minimum temperatures may have the most effect on natural resources such as increasing or delaying plant growth in spring and changing rain/snow regimes (Yu *et al.*, 2003).

The base period of temperature analysis used in global and regional studies varies. For the GISS (Goddard Institute for Space Studies) global analysis it is from 1951 to 1980, due to decent global data coverage and that it is a time remembered by most adults (Hansen *et al.*, 2010). Other studies, especially those examining sites in Asia, use periods such as 1961 to 2000 or later due to the length of record available (Klein Tank *et al.*, 2006; Marin, 2010). This analysis examined multiple time periods, due to the availability of data and the correlation to local knowledge. The data for one of the five stations (Tsetserleg) have been collected since 1961, while three started in the mid-1960s (Arvaikheer only had a good minimum temperature record and it ended in the late 1990s). The Horiult station only has data from 1976 to present. Clustering of relatively higher rates of change at a high 0.001 level of significance (not shown) occurs for the 30 year period of 1966 to 1995 for four out of the five selected stations. This period is similar to those chosen for baseline temperature analysis in other studies and could be a time of reference for comparison with greater amounts of statistically significant change occurring during that period.

Temperature changes are important to analyze as they influence many other natural processes. Warming (and cooling) temperatures drive processes such as evapotranspiration and the distribution and timing of precipitation events, which can negatively affect pastoralists due to their reliance on naturally pastured livestock (Ministry of Nature, Environment, and Tourism, 2010). Changes in minimums are more important than maximums or means in terms of biomass production, especially when warming temperatures limit moisture availability and delay green-up in the dry steppes (Yu *et al.*, 2003). Unfortunately, specific temperature changes are difficult to quantify when recalling climate phenomena of the past from memory.

Recent herder surveys (Fassnacht *et al.*, 2011) recorded a perception by some that the current winters are cooler than those in the past, yet one herder commented that, "*winter is like fall; there is no no real winter*." It is interesting to note that one of the herders who perceived cooler winters was younger than age 30. Recent winters in most areas have been much cooler than those in the 1990's, so how far back in time a person considers may influence their perception of change. For example, the herder that commented that winter temperatures now were more like fall (implying warmer winter temperatures) was nearly 50 years old. His memory back 20 years in time could have extended back to the mid-1980's when winter temperatures were on average cooler than winters in the 2000's.

Variability in temperature change trends may be related to larger-scale cyclicity. Periodicities are generally attributed to natural processes such as solar radiative forcing, large-scale patterns of circulation, and volcanism (Barnett *et al.*, 2005b). In an analysis of trend and periodicity, Chen and Grasby (2009) found that the start of an instrumented climate record relative to the phase of a particular cyclic phenomenon has a strong impact on estimation of trend. Their work shows that this impact lessens with lengthening record.

Utilizing the longest record lengths results in the highest statistical significance for temperature trends. Shorter periods of record support both higher and lower, and even negative trends with much variability depending upon period of record examined (Table 3.1). For a majority of the sites, the shortest periods that retained significance were those of the 15 year period from 1986 to 2000. The rates of change however for these short lengths of time were usually several degrees Celsius warmer per century than those for longer lengths of record.

Station location can also influence variability in the climate change trends. Many mountainous stations are higher than 1800 meters in elevation. Connections to climatic warming with elevation have been studied and while there is no strict relation it has been shown that the heterogeneous terrain of mountainous areas can lead to unequal climate trends especially in areas of incised valleys and flat locations (Pepin and Lundquist, 2008). The stations of Erdenemandal and Tsetserleg have the greatest number of highly statistically significant time periods. These two stations are located on the northern side of the Khangai Range, in more variable terrain than the other three stations. Bayankhongor is on the southern side of the mountain range and is also in somewhat rugged terrain. That station also has more highly significant periods than either Horiult or Arvaikheer.

Quantification of temperature trends based on a few point locations is problematic, especially related to issues of scale when interpolating or extrapolating meteorological data for use in modeling efforts (Daly, 2006). Regional averaging tends to smooth the effects of local temperature and other climate trends, and may hide important local trends (Pielke *et al.*, 2002). For example, the average rate of highly significant temperature change in

Erdenemandal is several degrees per century higher than most of the other station locations. Simple averaging of the highly significant trends of all five stations leads to a value that is several degrees lower per century than those at Erdenemandal. Perhaps the impacts of the ecological changes at Erdenemandal could be greater than those occurring at other locations with increasing minimum temperatures. This is where local knowledge of climate could be connected with regional models to test model applicability in an area (Marin, 2010).

Though not specifically examined in this paper, memory often favors extreme events. This is likely true of extremes of climatic conditions. Analysis of temperature minimums over varying time periods relates to extremes in that minimum temperature events such as the severely cold winter of 2009-2010 can affect longer-term climatic trendlines (Table 3.2). A comparison was made between periods excluding or before the 2009-2010 *dzud* years (up to 2008) and including (up to 2010) the anomalously cold winter. It illustrated that removing those years increased the significance of warming temperature trends and the rate of change. Temperature trend rates in most locations became lower for shorter periods including the cold winters. Interestingly, the effect of a cold winter on change in trend from up to 2008 to up to 2010 is minimized with an increase in length of record of 30 years but then increases for 35 years (Table 3.2).

In their analysis, Chen and Grasby (2009) used mainly synthetic data to illustrate the possible implication of the sampling period on the rate and significance of change, specifically sampling within a much longer cycle so that a highly significant upward or downward trend could actually just be part of a longer term oscillation. However, an individual extreme year could greatly influence pastoralists and others directly reliant on climate. In this case, the herders throughout Mongolia were devastated by the extreme cold of 2009-2010 with approximately 8.5 million livestock (20% of the livestock in Mongolia) perishing directly or indirectly from the *dzud*, affecting approximately 28% of Mongolia's population (UN Mongolia Country Team, 2010). The herders in Bayankhongor *aimag* were previously affected by a *dzud* in 1999-2002, and were thus more prepared for the 2009-2010 extreme cold (Fernandez-Gimenez *et al.*, 2011). Those in Arkhanagai were less prepared.

Statistically one extreme season decreased the rate of warming and for the moderate lengths of record, decreased the significance (Table 3.2). While different than what Chen and Grasby (2009) showed, this highlights the importance of careful analysis and interpretation of results. The implications of a changing climate, especially with an altered or possibly decreased preparedness for extremes, can be quite pronounced for those who live on the land and can provide local knowledge of change. This local knowledge can help inform and interpret environmental and climate change (see Alexander *et al.*, 2011 for examples).

Changing minimum temperatures can affect production of vegetation needed by livestock, changes to water supplies, and even a change in the amount and type of clothing needed to survive (Yu *et al.*, 2003; Batima *et al.*, 2005; Angerer *et al.*, 2008; Fassnacht *et al.*, 2011). Large-scale regional analyses and models are useful for national-level climate change mitigation and adaptability planning (Ministry of Nature, Environment, and Tourism, 2010), but local scale conditions will likely dictate changes in herder behavior as evidenced by responses to climate surveys (Marin, 2010; Fassnacht *et al.*, 2011).

3.7 Conclusion

In all cases, the most significant trends were those represented by the longest lengths of record. The most significant rates of change were not the highest or lowest rates of change at each station. The most highly significant rates of change in degrees per century were not uniform from station to station and covered a range of more than 5 degrees Celsius per century. Overall, the highest and most significant rates of change were approximately the 20 year period from 1981 to 2000, with a second clustering of significant values from 25 to 35 years in duration, starting in 1966 and extending until 1990 or 2000. This second period brackets the first cluster and may represent similar information on a longer time frame.

Increases in minimum temperatures are the most significant at a range of record lengths with a maximum of 50 years to about 30 years. A mid-range time period with higher rates of temperature change in degrees Celsius per century spanned the years from 1966 to 1995. Herder responses to climate surveys represent conditions experienced over a range of time periods. Observations however, often correlate with the documented changing climate trends. While the greatest statistical significance is noted with the longest records, shorter periods and extremes may more accurately represent the climate variability that is mentioned when herders are asked to recall natural phenomena of the past.

TABLE 3.1- Rates of temperature change for annual i) maximum and ii) minimum temperatures in degrees Celsius per century as a function of length of record and period (start-end years) of record. Significant rates are presented with three significant figures and are in italics for 10% significance (p<0.10), in italics and underlined for 5% significance (p<0.05), and in italics, underlined and bold for 1% significance (p<0.001). Non-significant rates are reported with one significant figure. Years with no data are represented as N/A.

ai) Erdenemandal maximum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
45	N/A	***4.40***						
40	N/A	***4.36***	***4.04***					
35	N/A	***4.44***	*3.84*	***4.77***				
30	N/A	3	4	*5.62*	***7.20***			
25	N/A	4	1	*6.71*	***8.86***	5		
20	N/A	1	2	2	***12.7***	6	4	
15	N/A	*11.4*	-7	4	9	9	6	-5

aii) Erdenemandal minimum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
45	N/A	***7.12***						
40	N/A	***7.28***	***6.61***					
35	N/A	***7.85***	***6.81***	***6.38***				
30	N/A	***7.85***	***7.55***	***6.41***	***5.98***			
25	N/A	***8.70***	***7.43***	***7.38***	***6.38***	*5.02*		
20	N/A	*6.53*	***8.64***	***6.46***	***9.55***	***6.58***	6	
15	N/A	7	*10.2*	*6.46*	***7.98***	*10.7*	6	8

bi) Tsetserleg maximum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
50	***3.69***							
45	***3.46***	***5.01***						
40	*2.81*	***5.06***	***4.59***					
35	1	***4.96***	***4.57***	***5.42***				
30	0	*3.93*	*4.35*	***5.92***	***6.98***			
25	-1	4	2	*5.55*	***7.53***	*6.28*		
20	0	3	-1	4	***8.67***	*7.53*	5	
15	-5	7	-4	-1	*6.59*	*11.0*	8	-4

bii) Tsetserleg minimum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
50	***3.57***							
45	***3.91***	***4.45***						
40	***3.56***	***5.04***	***3.84***					
35	***3.07***	***5.15***	***4.44***	***4.16***				
30	2	***5.38***	***4.14***	***5.13***	*3.96*			
25	0	*4.88*	***4.00***	***5.26***	***5.20***	3		
20	-1	4	2	***5.58***	*5.66*	4	2	
15	-2	6	0	4	*6.11*	6	4	-3

LEGEND: period of record (start-end years)

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
50	61-10							
45	61-05	66-10						
40	61-00	66-05	71-10					
35	61-95	66-00	71-05	76-10				
30	61-90	66-95	71-00	76-05	81-10			
25	61-85	66-90	71-95	76-00	81-05	86-10		
20	61-80	66-85	71-90	76-95	81-00	86-05	91-10	
15	61-75	66-80	71-85	76-90	81-95	86-00	91-05	96-10

significance level
0.01
0.05
0.10
NOT

cii) Arvaikheer minimum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
50	N/A							
45	N/A	N/A						
40	N/A	N/A	N/A					
35	N/A	***7.72***	N/A	N/A				
30	N/A	***6.98***	***6.11***	N/A	N/A			
25	N/A	***6.88***	***5.17***	***7.15***	N/A	N/A		
20	N/A	*9.59*	*2.81*	*5.34*	***9.69***	N/A	N/A	
15	N/A	***20.0***	1	2	*7.84*	*12.7*	N/A	N/A

di) Bayankhongor maximum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
45	N/A	***4.18***						
40	N/A	*3.50*	***3.87***					
35	N/A	*3.80*	*2.98*	***4.75***				
30	N/A	3	3	*4.00*	***6.38***			
25	N/A	2	1	5	*6.08*	*5.42*		
20	N/A	1	-1	3	*8.25*	4	*5.34*	
15	N/A	*10.3*	-6	2	8	7	3	3

dii) Bayankhongor minimum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
45	N/A	***5.10***						
40	N/A	***5.62***	***3.91***					
35	N/A	***6.38***	***4.19***	***4.63***				
30	N/A	***5.96***	***4.63***	***5.30***	***5.68***			
25	N/A	*5.08*	*3.25*	***6.88***	***7.61***	*3.78*		
20	N/A	5	1	*4.98*	***10.9***	*4.96*	1	
15	N/A	***12.5***	-4	3	*11.7*	*9.84*	1	-5

ei) Horiult maxiumum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
35	N/A	N/A	N/A	2				
30	N/A	N/A	N/A	1	2			
25	N/A	N/A	N/A	*4.00*	2	-1		
20	N/A	N/A	N/A	4	*4.81*	-4	-3	
15	N/A	N/A	N/A	4	*11.3*	3	*-15.1*	-3

eii) Horiult minimum

length of record \ start year	1961	1966	1971	1976	1981	1986	1991	1996
35	N/A	N/A	N/A	**5.36**				
30	N/A	N/A	N/A	4.69	***5.43***			
25	N/A	N/A	N/A	**8.05**	*4.68*	*4.03*		
20	N/A	N/A	N/A	5	*8.53*	2	3	
15	N/A	N/A	N/A	6	5	*8.85*	-4	0

TABLE 3.2- Rates of temperature change for annual i) maximum and ii) minimum temperatures in degrees Celsius per century as a function of length of record (15 to 35 years) for period ended at 2010 and at 2008. Significant rates are presented with three significant figures and are in italics for 10% significance (p<0.10), in italics and underlined for 5% significance (p<0.05), and in italics, underlined and bold for 1% significance (p<0.01). Non-significant rates are reported with one significant figure.

	a) Erdenemandal					
	i) maximum			ii) minimum		
	up to	**2008**	**2010**	**up to**	**2008**	**2010**
length of record	**35**	***5.83***	***4.77***	**35**	***7.13***	***6.38***
	30	***7.62***	***7.20***	**30**	***6.49***	***5.98***
	25	***9.50***	5	**25**	*7.46*	*5.02*
	20	6	4	**20**	*6.53*	6
	15	5	-5	**15**	10	8
	b) Tsetserleg					
	i) maximum			ii) minimum		
	up to	**2008**	**2010**	**up to**	**2008**	**2010**
length of record	**35**	***6.45***	***5.42***	**35**	***5.44***	***4.16***
	30	***7.47***	***6.98***	**30**	***5.20***	*3.96*
	25	***11.4***	*6.28*	**25**	***7.04***	3
	20	*9.17*	5	**20**	4	2
	15	7	-4	**15**	3	-3

significance level
0.01
0.05
0.10
NOT

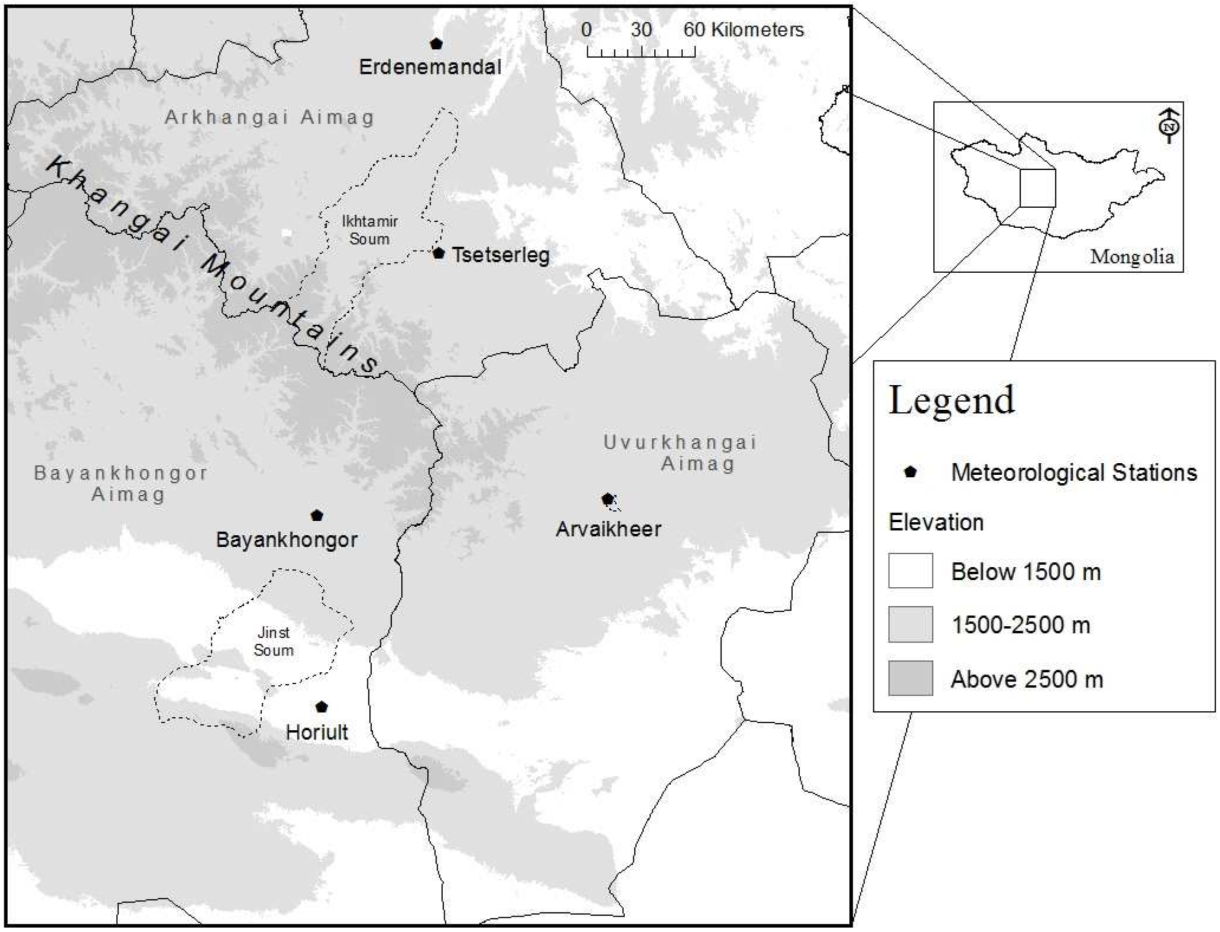


FIGURE 3.1- Map of study locations

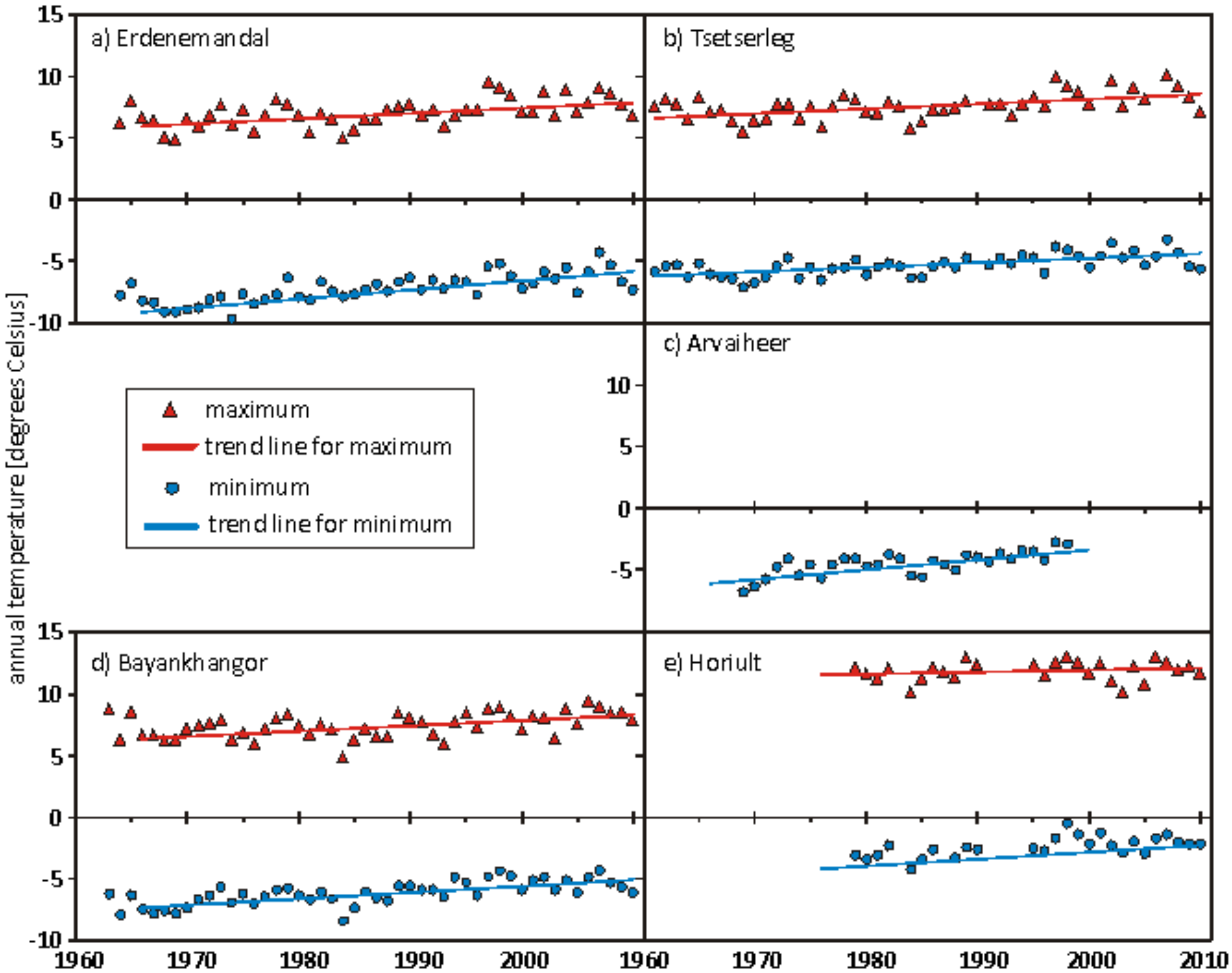


FIGURE 3.2- Annual maximum and minimum temperatures for the five meteorological stations. Stations are a) Erdenemandal, b) Tsetserleg, c) Arvaiheer, d) Bayankhongor, and e) Horiult

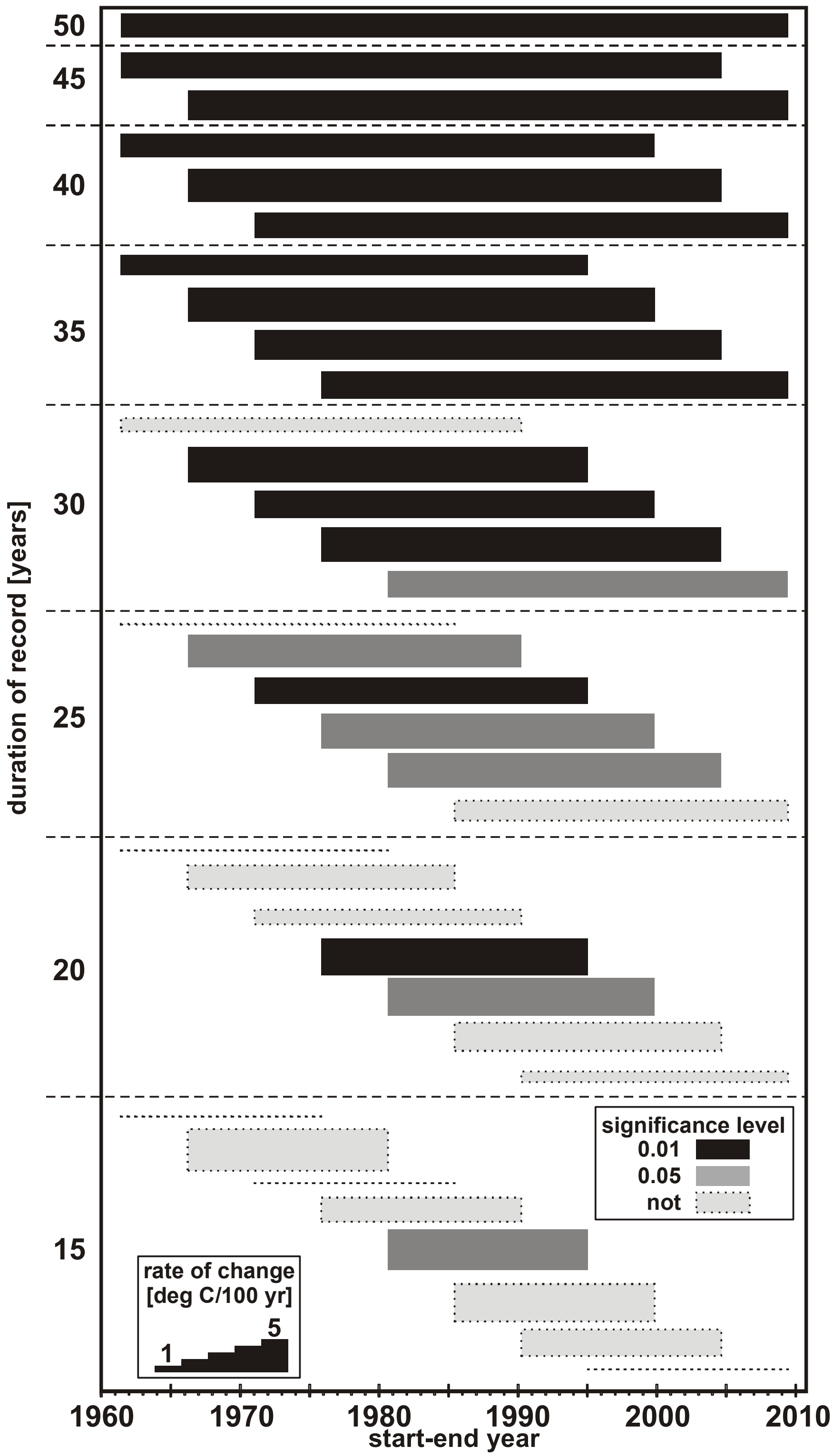


FIGURE 3.3- Graphical representation of the period and length of record calculations for annual minimum temperature at Tsetserleg.

REFERENCES

Alexander C, Bynum N, Johnson E, King U, Mustonen T, Neofotis P, Oettle N, Rosenzweig C, Sakakibara C, Shadrin V, Vicarelli M, Waterhouse J, Weeks B, 2011. Linking indigenous and scientific knowledge of climate change. *BioScience*, 61(6), 477-484.

Angerer J, Han G, Fujisaki I, and Havstad K, 2008. Climate change and ecosystems of Asia with emphasis on Inner Mongolia and Mongolia. *Rangelands,* 30(3), 46-51.

Barnett TP, Adam JC, and Lettenmaier DP, 2005a. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066), 303-309.

Barnett T, Zwiers F, Hegerl G, Allen M, Crowly T, Gillett N, Hasselmann K, Jones P, Santer B, Schnur R, Stott P, Taylor K, Tett S, 2005b. Detecting and attributing external influences on the climate system: A review of recent advances. *Journal of Climate*, 18(9), 1291-1314.

Batima P, Natsagdorj L, Gombluudev P, Erdenetsetseg B, 2005. Observed climate change in Mongolia. *Assessments and Adaptations to Climate Change* (AIACC), *www.aiaccproject.org*. Accessed 2011-11-21.

Begzsuren S, Ellis JE, Ojima DS, Coughenour MB, and Chuluun T, 2004. Livestock responses to droughts and severe winter weather in the Gobi Three Beauty National Park, Mongolia. *Journal of Arid Environments*, 59, 785-796.

Chen Z, and Grasby S, 2009. Impact of decadal and century-scale oscillations on hydroclimate trend analyses. *Journal of Hydrology* 365(1-2), 122-133, doi:10.1016/j.jhydrol.2008.11.031.

Daly C, 2006. Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, 26, 707-721.

Fassnacht SR, Sukh T, Fernandez-Gimenez ME, B Batbuyan, Venable NBH, Laituri M, and Adyabadam G, 2011. Local understanding of hydro-climatic changes in Mongolia. Proceedings of *H02 Symposium Cold Region Hydrology in a Changing Climate* held during IUGG2011 in Melbourne, Australia, July 2011, *IAHS Publication* 346, 120-129.

Fernandez-Gimenez M, B Batbuyan, B Batkhishig, 2011. Understanding Resilience in Mongolian Pastoral Social-ecological Systems*. Adapting to Disaster, Before, During, and After the 2010 Dzud- Year 1 Report*. Colorado State University & The Center For Nomadic Pastoralism Studies, Fort Collins, Colorado, USA, 119 pp.

Gilbert RO, 1987. *Statistical Methods for Environmental Pollution Monitoring.* John Wiley & Sons, New York, 320 pp.

Hansen J, Ruedy R, Sato M, & Lo K, 2010. Global surface temperature change. *Reviews of Geophysics*, 48(4), RG4004.

IPCC, 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the *Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL [Eds]. Cambridge University Press, Cambridge, United Kingdom, 237-383.

Jamiyansharav K, 2010. *Long-term analysis and appropriate metrics of climate change in Mongolia*. Unpublished PhD Dissertation, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA.

Klein Tank AMG, Peterson TC, Quadir DA, Dorji S, Zou X, Tang H, Santhosh K, Joshi UR, Jaswal AK, Kolli RK, Sikder AB, Deshpande NR, Revadekar JV, Yeleuova K, Vandasheva S, Faleyeva F, Gomboluudev P, Budhathoki KP, Hussain A, Afzaal M, Chandrapala L, Anvar H, Amanmurad D, Asanova VS, Jones PD, New MG, Spektorman T, 2006. Changes in daily temperature and precipitation extremes in central and south Asia. *Journal of Geophysical Research*, 111, D16105.

Marin A, 2010. Riders under storms: Contributions of nomadic herder's observations to analyzing climate change in Mongolia. *Global Environmental Change,* 20, 162-176, doi:10.1016/j.gloenvcha.2009.10.004.

Ministry of Nature, Environment, and Tourism, 2010. Mongolia Second National Communication, Dagvadorj D [Ed]. Ulaanbaatar, Mongolia, 160 pp.

Pepin NC, and Lundquist JD, 2008. Temperature trends at high elevations: Patterns across the globe. *Geophysical Research Letters*, 35, L14701.

Pielke Sr. RA, Stohlgren T, Schell L, Parton W, Doesken N, Redmond K, Moeny J, Mckee T, Kittel TGF, 2002. Problems in evaluating regional and local trends in temperature: An example from eastern Colorado, USA. *International Journal of Climatology*, 22(4), 421-434.

UN Mongolia Country Team, 2010. *Mongolian Dzud Appeal*. United National Development Programme, Office for the Coordination of Humanitarian Affairs, Consolidated Appeal, 53 pp.

Yu F, Price KP, Ellis J, and Shi P, 2003. Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sensing of Environment*, 87, 42–54, doi:10.1016/S0034-4257(03)00144-5.

CHAPTER 4- EVALUATING TREND DETECTION OF PRECIPITATION AND STREAMFLOW FROM SPARSE MONGOLIAN DATASETS

4.1 Summary

Climatic warming and potential drying may negatively impact the limited water supplies of semi-arid Mongolia. Precipitation over the Khangai Mountains supports streamflow in several major river systems of the region. Trends in these hydroclimatic variables are generally attributed to climate changes; however, the detection and attribution of trends may be influenced by several information-based factors such as data source, length/quality of records, step changes from anthropogenic sources, amount of serial correlation, and method of analysis. After inhomogeneity testing, we applied two variations of the Mann-Kendall test and the Thiel-Sen approach to precipitation and steamflow records from six meteorological stations and six streamflow gages in the Khangai Mountain region to determine the existence and magnitude of significant trends. Step changes were noted in some Mongolian datasets, possibly influencing trend detection. Little difference was found between the two Mann-Kendall and Thiel-Sen methods used. For comparison, two long period hydroclimate records from the San Luis Valley of Colorado, USA were analyzed using the same methods. Results suggest decreasing trends in annual, spring, and summer precipitation at several Mongolian stations, with increasing winter precipitation trends at one site. Decreases in streamflow were of greatest magnitude for rivers on the northern side of the Khangai Mountains. The Colorado datasets do not exhibit step changes and have no significant trends over the total length of record. Degradation of the Colorado hydroclimate records by shortening the time series and introducing gaps to simulate inconsistencies found in Mongolian datasets yielded

significant trends. Interpretation of the results of trend analyses on datasets of shorter periods and of uncertain quality should be made in light of the influences of data structure on the detection and significance and/or magnitude of trends.

4.2 Background

Water resources are limited across the vast semi-arid regions of the world. This is also true in Mongolia where climatic warming and potential drying are a concern for scientists, policy makers, and the nomadic pastoralists of the country (e.g., Batima *et al*., 2005; Ministry of Nature, Environment, and Tourism, 2010; Marin, 2010; Fassnacht *et al*., 2011). While the case for climatic warming has been well established in many mountainous and semi-arid regions, including in Mongolia and in the Rocky Mountain region of Colorado, USA, trends in precipitation and streamflow are less clear (e.g., Batima *et al*., 2005; Ray *et al*., 2008; Ministry of Nature, Environment, and Tourism, 2010; Jamiyansharav, 2010; Venable *et al.*, 2012).

Precipitation is spatially variable across the Mongolian landscape with more occurring annually in the higher latitude and higher elevation regions of the country. Precipitation events are often intense with most of the rainfall recorded between April and September (Ministry of Nature, Environment, and Tourism, 2010). The seasonal nature of precipitation affects the availability of rangeland vegetation for grazing livestock and surface and shallow subsurface water resources. Trend analyses of hydroclimatic variables are needed for assessing negative impacts of possible climatic changes that affect herders' livelihoods and influence developing agricultural and mining interests and burgeoning population centers (Fernandez-Gimenez, 2000; Ministry of Nature, Environment, and Tourism, 2010; Konagaya, 2013; Suzuki, 2013).

Several issues affect the determination of trends in Mongolian precipitation and streamflow as compared to those of temperature. These are related to the spatial and temporal variability of hydroclimate data in Mongolia and the quality of the data. High inter-annual variability, serial correlation (autocorrelation), and relatively short climatic records (less than 50 years, and often less than 30 years) with many missing measurements affect trend analyses and make significant trends attributed to a changing climate difficult to separate from natural variability and/or artificially introduced change points such as station relocations (Askew, 1987; Yue *et al*., 2002; Chen and Grasby, 2009; Venable *et al*., 2012; Gallagher *et al*., 2013).

In this work, trend analyses from several meteorological stations and streamflow gaging stations in the topographically and ecologically diverse Khangai Mountain region of Mongolia (Figure 4.1a) are presented, along with an analysis of data problems that affect the significance of trends found in precipitation and streamflow data of the region. Specifically, we assess the structure and quality of several Mongolian climate datasets, including record length and missing observations, possible step-changes to the time series, and presence of autocorrelation. We test these data for trend using two different methods, one of which accounts for the effects of positive autocorrelation on the detection of a significant trend. The results of the Khangai Mountain region station analyses are compared to trend analyses of higher quality, long period, hydroclimatic records from Colorado, USA (Figure 4.1b) using the same methods. The Colorado records were subsequently degraded to represent the general data quality of several of the Khangai Mountain region time series through shortening of the record and the inclusion of gaps in the data from one to five years in length, similar to those gaps noted in the Mongolian records. The results regarding changes to trend detection and significance were analyzed.

4.3 Methods

4.3.1 Data Sources

Meteorological and hydrological datasets used in this work are from two sources. The Khangai Mountain region datasets are from the Mongolian Institute of Meteorology, Hydrology, and the Environment (IMHE). Daily precipitation and streamflow data were analyzed for six meteorological stations and six stream gages (Tables 4.1 and 4.2). Two of the meteorological stations (Erdenemandal and Tsetserleg) and two of the stream gage sites (the Khanui River at Erdenemandal and the Khoid Tamir at Iktamir) are located on the northern side of the Khangai Mountains (Figure 4.1a). On the south side of the mountains, the meteorological stations of Baidrag and Galuut are near the gage sites of the Baidrag River at Baidrag and the Baidrag River at Bayanburd (Figure 4.1a). The meteorological stations at Baynkhongor and Khoriult are near the Tuin River at Bayankhongor and the Tuin River at Bogd, respectively (Figure 4.1a).

The Colorado data analyzed in comparison were downloaded from the United States National Oceanic and Atmospheric Administration National Climatic Data Center (http://ncdc.noaa.gov/), the United States Geological Survey National Water Information Center (http://waterdata.usgs.gov/co/nwis/), and the Colorado Division of Water Resources (http://www.dwr.state.co.us/). Daily precipitation data were from Del Norte and daily streamflow data were from [North] Crestone Creek (Tables 4.1 and 4.2). The station at Del Norte is part of the Global Historical Climatology Network of long-record stations that are subject to standardized quality assurance reviews. Crestone Creek is relatively pristine with no diversions above the gage and has headwaters in the Sangre de Cristo Wilderness Area. Both sites are located in the semi-arid San Luis Valley of south-central Colorado in the western United States (Figure 4.1b).

4.3.2 Data Quality and Length of Record

The average lengths of records for the Mongolian stations are 39 years, with the longest record of 63 years at Tsetserleg and the shortest record of 19 years at Khoriult. There are many observational gaps in the individual time series. The precipitation data has many potential missing values with several months of any given year only having a few observations recorded per month. If the precipitation data were considered "regular" as in regularly sampled every day of the year, the amount of missing data per station is potentially quite large. From 67% to 90% of the precipitation data could be considered missing (Table 4.1). For streamflow, the sampling schedule is more regular with less daily data missing, however almost 2% up to 23% are unavailable (Table 4.2). The Colorado data were specifically selected for their length of record and completeness, with an 82-year record at Del Norte and a 67-year record for Crestone Creek. Only 1.2% of the daily data are missing for Del Norte and no daily data are missing for Crestone Creek (Tables 4.1 and 4.2).

Precipitation and streamflow data were processed using R statistical software (R Core Team, 2014). The data were quality controlled for obvious erroneous values through visual examination of the datasets and determination of outliers. Outliers were considered as values that fall more than 3.0 standard deviations from the mean (4.0 for precipitation) (e.g., Harris *et al*., 2014). Limited obvious errors were found in the Mongolian datasets; the few values of precipitation and streamflow beyond the standard deviation limit tests were not considered to be outside the range of high natural variability inherent in these systems. The most common errors detected were related to decimal point conversions from commas in the original datasets. After the quality control checks were completed, daily precipitation data were aggregated into total monthly, seasonal, and annual time series, and daily streamflow data were aggregated into mean and median monthly, and mean and median annual time series.

When at least one day or more of precipitation observations were recorded per month, the month was not considered missing. This decision was made to preserve as much precipitation data as possible for analysis under the assumption that only days with precipitation were recorded in the original datasets. As the streamflow data were recorded on a more regular daily time interval, when aggregated to monthly values, any months with less than 23 days (~75%) of data were considered missing (e.g., Harris *et al*., 2014). When aggregating the precipitation and streamflow data to annual values, any years with less than 12 months of data were considered missing. As much of the precipitation data was considered zero values and not missing, the distribution of missing data in the annual record for this work is generally related to the length of record (Figure 4.2a). The original streamflow data had several missing months or years, such as for the Khoid Tamir River at Ikhtamir between 2005 and 2008 (Figure 4.2b). Also, the more stringent requirement of 75% of daily values per month and 12 months per year resulted in missing data dispersed throughout annual record.

4.3.3 Change Point Analyses

A common problem in climate and hydrological data series is the presence of inhomogeneities. A majority of these appear as abrupt changes in the average values, but also appear as changes in the trend of the series (Alexandersson and Moberg, 1997). Inhomogeneities in temporal series can result in substantial misinterpretation of the behavior and evolution of climate trends. Inhomogeneities generally arise from human causes such as changes in the location of the observation station, alteration of the surrounding environment, observer changes and instrument replacement (Karl and Williams, 1987).

A variety of methods have been developed to identify inhomogeneities in hydroclimatic data series (see the reviews in Beaulieu *et al*., 2007 and Peterson *et al*., 1998). There are two general types of homogenization procedures: 1) absolute, which considers only the information in the time series being tested, and 2) relative, in which data from other stations are also used. In this study, the lack of other reasonably close reliable precipitation and runoff series with similar length of the study period forced us to use absolute methods. We employed the Mann-Whitney Pettitt test (Pettitt, 1979) to identify breaks in both precipitation and runoff series. This is a widely used test for detecting inhomogeneities in hydroclimatic series (Kundzewicz and Robson, 2004; Lorenzo-Lacruz *et al.*, 2012). We assumed that anthropogenic induced inhomogeneities should affect a considerable number of individual months. When metadata was available (as in the case of the US data, but not for Mongolia), years noted with several breaks in the monthly series should coincide with years flagged in the metadata as potential causes of inhomogeneity.

4.3.4 Trend Analyses

All of the daily and monthly data were highly autocorrelated, with each observation having a degree of dependence on previous observations. Since this dependence could influence analysis results, we chose to focus the testing on seasonal and annual timesteps where autocorrelation is reduced between subsequent values in the time series of interest. Based on the short record lengths, amounts of missing data, and previous work with Mongolian hydroclimatic datasets, full periods of record for each location were used (i.e. Fassnacht *et al*., 2011; Venable *et al*., 2012) (Tables 4.1 and 4.2). Data were first checked for autocorrelation above a 5% significance threshold of the series length. The seasonal and annual precipitation and annual streamflow time series were analyzed using the Mann-Kendall (MK) test for trend (Mann, 1945; Kendall, 1990). If the presence of a trend was detected, the Thiel-Sen (TS) approach was applied to estimate the magnitude (slope) of the trend (Thiel, 1950; Sen, 1968). If no trend was detected with the MK method no further analysis of that data was completed. Median values were also tested for trend due to the resistance of the median to the influence of extreme observation values, common in precipitation and streamflow datasets (Helsel and Hirsch, 2002).

Simulations by Yue *et al.* (2002), demonstrate the presence of autocorrelation in data affects the detection of significant trends when using the MK test. Additionally, the presence of trend in a time series was shown to influence the magnitude of the estimate of autocorrelation present in the series. Therefore, in our study a multi-step process was employed.

If significant trend was detected using the initial MK method, the trend-free pre-whitening procedure (TFPW) prescribed by Yue *et al.* (2002) was applied to the time series by first computing the slope of the trend using TS. The trend was then assumed to be linear and was removed from the data. The autocorrelation coefficient at a lag of one timestep (lag-1) was computed from the resulting detrended residual time series and the AR(1) process (autoregressive model of order 1) was removed. This ideally resulted in the creation of an independent time series. The modified residual series was then combined with the identified trend and the MK test was reapplied to assess the significance of the trend.

4.3.5 Effects of Time series Length and Continuity

To test the possibility of station record length and completeness affecting the detection of a trend, the Colorado meteorological station and streamflow gage records for annual and seasonal precipitation and annual mean streamflow were degraded using two methods similar to those used in Loew (2014). The first involved testing the step-wise reduction of the time series length to simulate the lengths of the shorter Mongolian time series. The second incorporated a moving window of missing years across the length of record. Gaps from one up to five years were used, as those gap lengths were similar to those in the original Mongolian records. The TFPW MK and TS methods initially used on the Mongolian and Colorado datasets were re-applied to the degraded Colorado datasets to detect the presence of trend induced by changes to the structure of the original datasets.

4.4 Results

4.4.1 Potential Change Points

The results of testing for potential change points revealed the likelihood of a step-change in streamflow of the Khanui at Erdenemandal. Runoff is noticeably lower after 1996 than in preceding years. A change was also noted for the Khoid Tamir at Ikhtamir around the same period of 1996-1997. No inhomogeneities were detected in the precipitation data. No changes attributable to anthropogenic influences can be inferred from the Colorado data.

4.4.2 Initial Testing

At an annual timestep, most precipitation values are not significantly autocorrelated at the 5% significance threshold of series length, but seasonally nearly half the stations analyzed had significant summer autocorrelation. Significant autocorrelation exists at this level for annual mean streamflow for the Khanui at Erdenemandal and the Tuin at Bogd, and also for annual median flow for the Khanui at Erdenemandal and for Crestone Creek. Initial testing using MK and TS on the datasets revealed that a rejection of the null hypothesis of no trend was possible for both annual and summer precipitation at Erdenemandal, for winter precipitation at Baidrag, and for summer precipitation at Galuut (Table 4.3). For annual mean streamflow only the Khanui at Erdenemandal, the Khoid Tamir at Ikhtamir , and the Tuin at Bayankhongor were found to have trends different from zero. Annual median flows had similar results, with the Khanui at Erdenemandal, and the Khoid Tamir at Ikhtamir having significant trend (Table 4.3).

All significant trends were decreasing with the exception of winter precipitation at Baidrag (2.9 millimeters per decade). The greatest losses in precipitation were noted at Erdenemandal (-16.3 and -11.0 millimeters per decade, annually and summer respectively). Summer precipitation at Galuut decreased similarly (-10.3 millimeters per decade). Decreasing trends in annual mean flow for the Khanui at Erdenemandal, and the Khoid Tamir at Ikhtamir were similar in magnitude (-2.3 and -3.4 cubic meters per second per decade, respectively), but the Tuin at Bayankongor recorded a minimal decrease of -0.6 cubic meters per second per decade. Median annual flow decreases on the Khanui River at Erdenemandal and the Khoid Tamir at Ikhtamir were -1.6 and -1.3 cubic meters per second per decade respectively.

4.4.3 Testing After Pre-Whitening

The MK TFPW and TS results are the generally the same as those of testing using the unmodified MK and TS. The only differences were that trends in summer precipitation at Erdenemandal and annual median flow for the Khoid Tamir River at Iktamir were no longer significant at the $p<0.05$ level (Table 4.3). Autocorrelation at each of those sites were calculated as 0.21 and 0.16 before applying the unmodified MK test and were 0.14 and 0.13 after applying the MK TFPW procedure. The 5% length of series significance cutoff was for 0.25 for seasonal precipitation and 0.32 for annual median flow.

Performing the same tests at a higher level of significance (the $p<0.01$ level) resulted in no significant trends in precipitation for the stations analyzed on an annual or seasonal basis (results not shown). The only significant changes in mean/median streamflow were for the Khanui River at Erdenemandal and for mean flows on the Khoid Tamir River at Ikhtamir (results not shown).

4.4.4 Altered Time series Length and Continuity of the CO Datasets

Shortening the lengths of record of the Colorado datasets by moving the trend analysis start dates resulted in the appearance of significant trends (at the $p<0.05$ level) in both the annual precipitation (Figures 4.3a and 4.3b) and spring and summer seasonal precipitation (results not shown) and annual mean streamflow data (Figures 4.4a and 4.4b). Significant trends appeared primarily between 20 and 40 years of record for all datasets, with some significant trends occurring at very short lengths of record, <15 years in the seasonal data. An example of a period of significant decreasing trend for annual precipitation is given in Figure 4.3a. This length of record is likely significant due to the unique nature of the climate history of the San Luis Valley; the trends during that time reflect a bridging of above average wet periods at the start of the shortened record and below average dry periods at the end. This result confirms that changing record length can affect trend occurrence and magnitude, particularly over shorter lengths of record in this region.

Gap creation also results in the occurrence of significant trends where none were previously detectable using the entire length of record. For example, if even 1 year were removed from the 82-year annual precipitation record at Del Norte, it would result in a significant decreasing trend in precipitation of -5.1 mm per decade (results not shown). Fall seasonal precipitation was sensitive to the addition of gaps in the data, with periods of missing data from one to five years creating significant increasing trends of up to 3.2 mm per decade at three to nine different points across the time series. Figure 4.5b shows an example of the effects of five-year gaps on the detection of significant trend over the 82-year length of record for fall precipitation. The addition of gaps did not result in the detection of any significant trends in seasonal precipitation for winter, spring, or summer, or for streamflow.

4.5 Discussion

Station selection is a critical procedure for trend analyses (Burn and Hag Elnur, 2002). In developing countries such as Mongolia, however, there are only a few stations with limited climatic records available over large areas. Therefore, it is important to understand the nature of the data and possible sources of error associated with these records. Confounding factors in the detection of significant trends include the existence of abrupt change points both human-caused and those of an unidentified nature, the presence of autocorrelation and possible long-terms trends, the type of trend analyses applied, and the quality of the data used for analysis.

The time series used in these analyses have not been homogenized, though tests for step changes in trend have been performed. Inhomogeneities discovered in the Khanui at Erdenemandal and the Khoid Tamir at Ikhtamir records could be the source of the significant trends documented in streamflow, though only the Khanui at Erdenemandal had a known change in location. Difficulties arise as well when testing for step changes without a reliable reference network (Easterling and Pederson, 1995). For example, the great distances between stations may not make the use of other stations as references reasonable, particularly for precipitation as it can have high spatial variability.

Little information exists on potential relocations of Mongolian hydroclimatic instruments. Jamiyansharv's work (2010) examined the site characteristics of several meteorological stations in Mongolia, including the stations at Tsetserleg and Bayankhongor. It is unlikely that either of these stations moved since their establishment, and no information was available for the other meteorological stations used in this study. Some metadata regarding the movement of the stream gage sites was provided by the IMHE, suggesting that the gage for the Khanui River at Erdenemandal was moved at least once to a location 9 km upstream from a previous location (date of move unknown). No other movements were recorded for the other river gages under study. Field visits to the Khoid Tamir River at Ikhtamir and the Tuin River from its source in the Khangai Mountains to its terminus at Orog Lake including both the gage sites at Bayankhongor and Bogd, suggest no movement of the gage sites and no significant man-made structures or diversions of flow on either river (Fassnacht *et al.*, 2015). In comparison, the station instruments at Del Norte have been moved several times in the past with the largest moves occurring between 1997 and 1998 (into the town of Del Norte, about 2 kilometers east of the original location) and to the most recent location from 2004 to the present on the north side of the Rio Grande floodplain (approximately 2 kilometers north and east of the original location). The Crestone Creek gage was moved once (in 2006) to a location 48.8 meters downstream of the original location.

Autocorrelation is commonly present in hydroclimatic data due to the persistent nature of the system (i.e., high streamflow values tend to follow high streamflow values until there is a change in system inputs) (Helsel and Hirsch, 2002), and is evident in the Mongolian datasets even at larger aggregate timesteps of the seasonal and annual scale. It is known to affect the results of the MK test (e.g., von Storch, 1995) and may have interactive effects when a trend and autocorrelation are present in a time series (Yue *et al.*, 2002). Of interest is that the results from only two stations/gages were affected by a change in the trend testing methodology over the initial application of the MK and TS tests. These were summer precipitation at Erdenemandal and median annual flow on the Khoid Tamir at Ikhtamir. Of these two, inhomogeneities were only found for the Khoid Tamir at Ikhtamir, but some positive autocorrelation (lag-1) was detected in both records. The amounts however, were below the 5% significance limits.

Khangai climate records are short and have much missing data. The assumption that all of the precipitation data is acceptable as received, aside from minor quality control checks, is likely a poor choice for use in trend analysis. It is unclear how much data is missing or can be recorded as zero. While some months may truly have only a few precipitation events occurring, with the strong seasonal precipitation patterns of the

Khangai region, it is unlikely that certain months, particularly those occurring in the summer, would only have a small number of events recorded. What appear to be low periods of precipitation may actually be measurement recording error.

The results support our hypotheses that changing hydroclimatic record lengths of long datasets with no apparent trend and inserting missing periods affects the detection of significant trends. For example, when the start date of the analysis was moved for Crestone Creek, significant trends ($p<0.05$) in mean streamflow occurred at shorter lengths of record of 43 years, 41 years, and in a block from 39 down to 21 years (Figure 4.4b). This effect will be different and unique to each climate time series as a comparison of Figures 4.3b and 4.4b shows. Trends attributed to a changing climate may reflect shorter-term climate variability or oscillations in climate (Kundzewicz and Robson, 2004; Chen and Grasby, 2009).

In the gap analysis tests using the Colorado data, precipitation at an annual timestep was more sensitive to gaps from one to five years in length than streamflow (results not shown). Gaps at one and two year lengths were found to be significant at a few intervals within the record (i.e. one year missing in 2002 changed the trend from non-significant to significant) but for gap lengths of three to five years, significant trends were clustered at the ends of the records indicating a shortening of the record was more likely the cause of the significant trend than a gap within the time series tested. Significant trends from gaps within the time series were more often detected in the fall precipitation data than in any of the other time series. While a similar clustering of significant trends was seen when the 5-year gaps were placed toward the end of the record, a period of several years around 1960 were also determined significant at the $p<0.05$ level when gaps were inserted around that time (Figure 4.5b). This is likely due to the specific data structure and seasonality of the Del Norte record as compared to that of the Mongolian records (Figure 4.5a).

In other published research, trend analyses of precipitation and streamflow in Mongolia use station-based and gridded datasets. While most employ the MK test, several use other methods than those used here. Results vary but many suggest decreasing precipitation trends in the region of interest or in neighboring areas. For example, one of the longest periods of analysis (1940-2001) of Mongolian data was by Batima *et al*. (2005), but their work averaged data from stations throughout Mongolia, used normalized anomalies of the data, and applied linear regression to estimate trends in precipitation. Changes of -30+ millimeters of precipitation over the last 30 years were calculated for central parts of the Khangai region (Batima *et al*., 2005). In a desert-steppe region east of the Khangai, Marin (2010) finds decreasing annual and summer precipitation amounts using the MK test over a period from 1961-2007, however most trends were not statistically significant. Only one station had significant decreasing August and September and increasing February precipitation (Marin, 2010). Our work found significant decreasing annual and summer trends (includes August) at some stations in the study area and increasing winter precipitation (includes February) at one station. Another study by Endo *et al*. (2006), over the period of 1960 to 1998 examined daily and monthly summer total, type, and intensity of precipitation for trends and extremes across Mongolia and into China and Russia. Their work acknowledges possible biases in the data from instrument measurement error but they did not make corrections due to a lack of additional data. Several quality control checks with data archived in multiple databases and with surrounding stations were made (Endo *et al*., 2006). The authors note the difficulty in distinguishing between trace and/or zero millimeters of precipitation, and the determination of no observation made. They conclude decreasing total summer precipitation trends are occurring across the central part of the country using the MK test, but none are significant (Endo *et al*., 2006)

One of the few studies employing gridded data averaged several different gridded datasets to reduce the bias that may exist in any single product and then applied the MK test for trend (Liu *et al.*, 2013). Their study of data from 1998 to 2008 found significantly decreasing trends in the northeastern and western parts of our study region, including the areas around the stations of Erdenemandal, Baidrag, and Galuut. Our station-based analysis saw significantly decreasing annual precipitation only at Erdenemandal ($p<0.05$). Similarly, another study by Venable *et al*. (2015), using monthly gridded data and the MK test over the period of 1963 to 2012 found significant decreases in annual total precipitation north of the Khangai Mountains, but no significant changes on the southern side of the mountains. No significant changes were seen in either winter or spring precipitation, but decreases were noted in the same area for summer and fall (Venable *et al*., 2015).

Aside from studies examining herders' observations of climate including precipitation and/or streamflow such as Marin (2010) and Fassnacht *et al*. (2011), most studies of changes in Mongolian streamflow near the Khangai are focused on hydrological modeling (Ma *et al*., 2003), or streamflow reconstruction (Davi *et al*., 2006; Pederson *et al*., 2013) using proxy climate records such as tree rings. These works assess the Selenge River, of which the Khanui is an upper tributary. Ma *et al*. (2003) discusses difficulties in the application and interpretation of model results due to the low density of stations and the high interannual variability of Mongolian climate data. Interestingly, the works by Davi *et* *al*. (2006) and Pederson *et al*. (2013) suggest that trends in flow during the 20th century for the Selenge River were generally increasing compared to those of the previous century. Further supporting the idea that length of trend analyzed can have an effect on perception of increasing or decreasing climatologic trends. While Batima *et al*. (2005), mentions a drying of rivers and streams through time, trend analyses of streamflow in Mongolia by other authors are limited.

4.6 Conclusions

Analysis of trends in precipitation and streamflow data are complicated by many factors such as data source, length and quality of record, possible step changes from climate or human influences, amount of autocorrelation present and chosen method of analysis. Our work explores these problems in the context of datasets from the Khangai Mountain region of Mongolia, where nomadic pastoralists rely on natural water supplies for sustaining their flocks and livelihoods. We assessed the structure and quality of these datasets in comparison to longer period hydroclimatic records from Colorado, USA and tested the annual and seasonal precipitation and mean and median streamflow records for trend using the standard Mann-Kendall and Thiel-Sen tests and the Mann-Kendall trend-free pre-whitening procedure and Thiel-Sen test proposed by Yue *et al*., (2002). The results suggest overall decreasing trends in annual, spring, and summer precipitation at several Mongolian stations, with only one site with increasing winter precipitation trends. Decreases were also found in mean and median streamflow for several river gage locations, of greatest magnitude for the rivers on the northern side of the Khangai Mountains. Step changes were noted in some of the Mongolian streamflow datasets however, which may influence trend detection. Little difference was noted in trend detection between the two

Mann-Kendall and Thiel-Sen methods used. When the higher quality, longer record, Colorado datasets were degraded by record shortening and the insertion of gaps in the record, significant trends were detected where none were observed using the complete length of record. These results highlight a need to carefully consider the significance and magnitude of trend analyses conducted on shorter period hydroclimate records with known inhomogeneities and potentially large amounts of missing data as those analyzed from the Khangai Mountain region in Mongolia.

TABLE 4.1- Meteorological stations including amount of observational record selected for analysis from the Khangai Mountain region of Mongolia and Colorado, USA.

Precipitation Station Name	Latitude (°)- Longitude (°)	Elev (m)	Observational Record	Days Missing[1]	% of Record Missing[2]
Erdenemandal	48.53 N - 101.38 E	1509	1964-2012	13263	74.1
Tsetserleg	47.45 N - 101.47 E	1691	1950-2012	15581	67.7
Baidrag	47.20 N - 99.60 E	2199	1993-2012	5931	81.2
Galuut	46.70 N - 100.13 E	2126	1956-2012	16712	80.3
Bayankhongor	46.13 N - 100.68 E	1859	1963-2012	15079	82.6
Khoriult	45.19 N - 100.57 E	1276	1994-2012	6269	90.3
Del Norte	37.67 N - 106.35 W	2403	1933-2014	343	1.2

TABLE 4.2- Streamflow stations including amount of observational record selected for analysis from the Khangai Mountain region of Mongolia and Colorado, USA.

Streamflow Gage Name	**Latitude (°)-Longitude (°)**	**Elev (m)**	**Observational Record**	**Days Missing[1]**	**% of Record Missing[2]**
Khanui R. at Erdenemandal	48.61 N- 101.38 E	1487	1976-2010	2955	23.1
Khoid Tamir R. at Ikhtamir	47.48 N - 100.89 E	1740	1976-2010	2918	22.8
Baidrag R. at Baidrag	47.13 N - 99.67 E	2148	1985-2010	662	7.0
Baidrag R. at Bayanburd	46.67 N - 99.27 E	1880	1976-2010	195	1.5
Tuin R. at Bayankhongor	46.14 N - 100.72 E	1863	1976-2010	1918	15.0
Tuin R. at Bogd	45.19 N - 100.78 E	1271	1971-2010	2202	15.1
Crestone Ck., N. near Crestone	38.01 N -105.69 W	2553	1948-2014	0	0

TABLE 4.3- Significant annual and seasonal trends and rates of change (slopes) in millimeters/decade (precipitation) or cubic meters per second/decade (streamflow) for the Khangai region and Colorado data before and after trend-free pre-whitening (TFPW) procedures.

Station/Gage Name	Variable	Slope No TFPW	Slope TFPW Applied
Erdenemandal	Annual Precipitation	-16.3	-16.3
Erdenemandal	Summer Precipitation	-11.0	NS[1]
Baidrag	Winter Precipitation	2.9	2.9
Galuut	Summer Precipitation	-10.3	-10.3
Khanui River at Erdenemandal	Mean Flow	-2.3	-2.3
Khanui River at Erdenemandal	Median Flow	-1.6	-1.6
Khoid Tamir River at Ikhtamir	Mean Flow	-3.4	-3.4
Khoid Tamir River at Ikhtamir	Median Flow	-1.3	NS[1]
Tuin River at Bayankhongor	Mean Flow	-0.6	-0.6

[1] Results not significant at $p<0.05$.

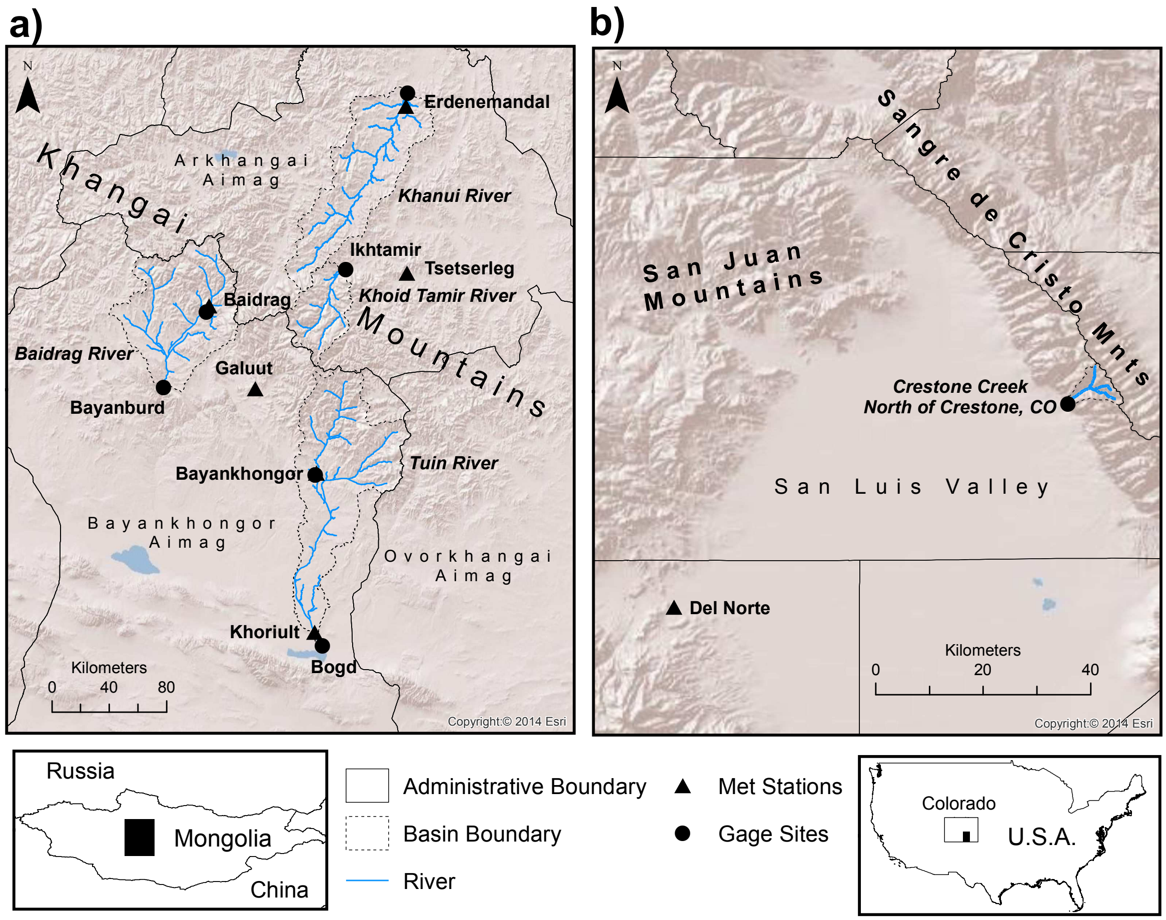


FIGURE 4.1a- (left) Study area in the Khangai Mountain region, Mongolia. Meteorological station locations used in the analyses are triangles and stream gage locations are circles.

FIGURE 4.1b- (right) Study area in the San Luis Valley, Colorado, USA. Meteorological station locations used in the analyses are triangles and stream gage locations are circles.

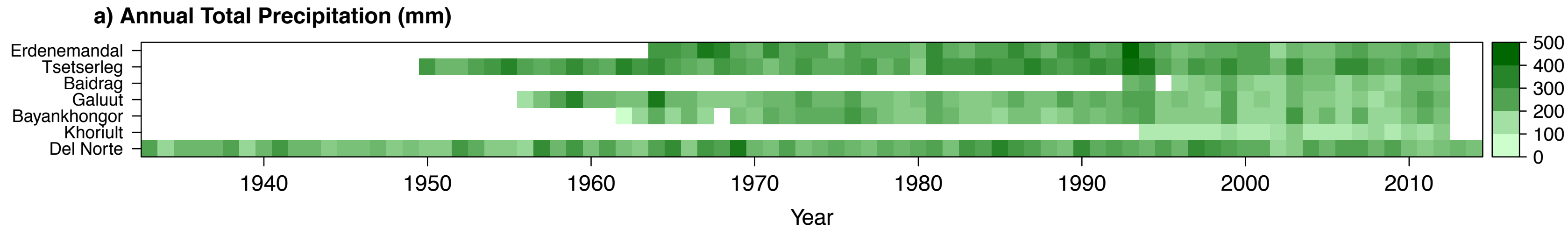


FIGURE 4.2a- Annual total precipitation in millimeters over the period of record for each meteorological station of analysis. The upper six locations on each plot are from the Khangai Mountain region of Mongolia and the last site in each plot is from Colorado, USA. White spaces denote missing data.

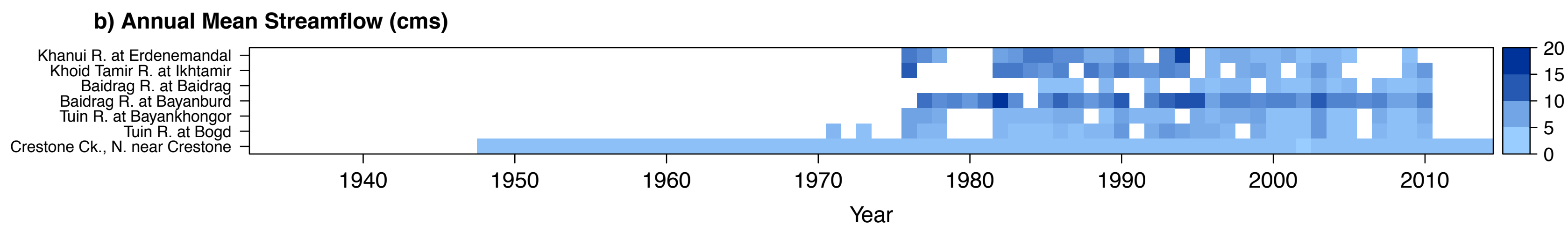


FIGURE 4.2b- Annual mean streamflow over the period of record for each gaging site of analysis. The upper six locations on each plot are from the Khangai Mountain region of Mongolia and the last site in each plot is from Colorado, USA. White spaces denote missing data.

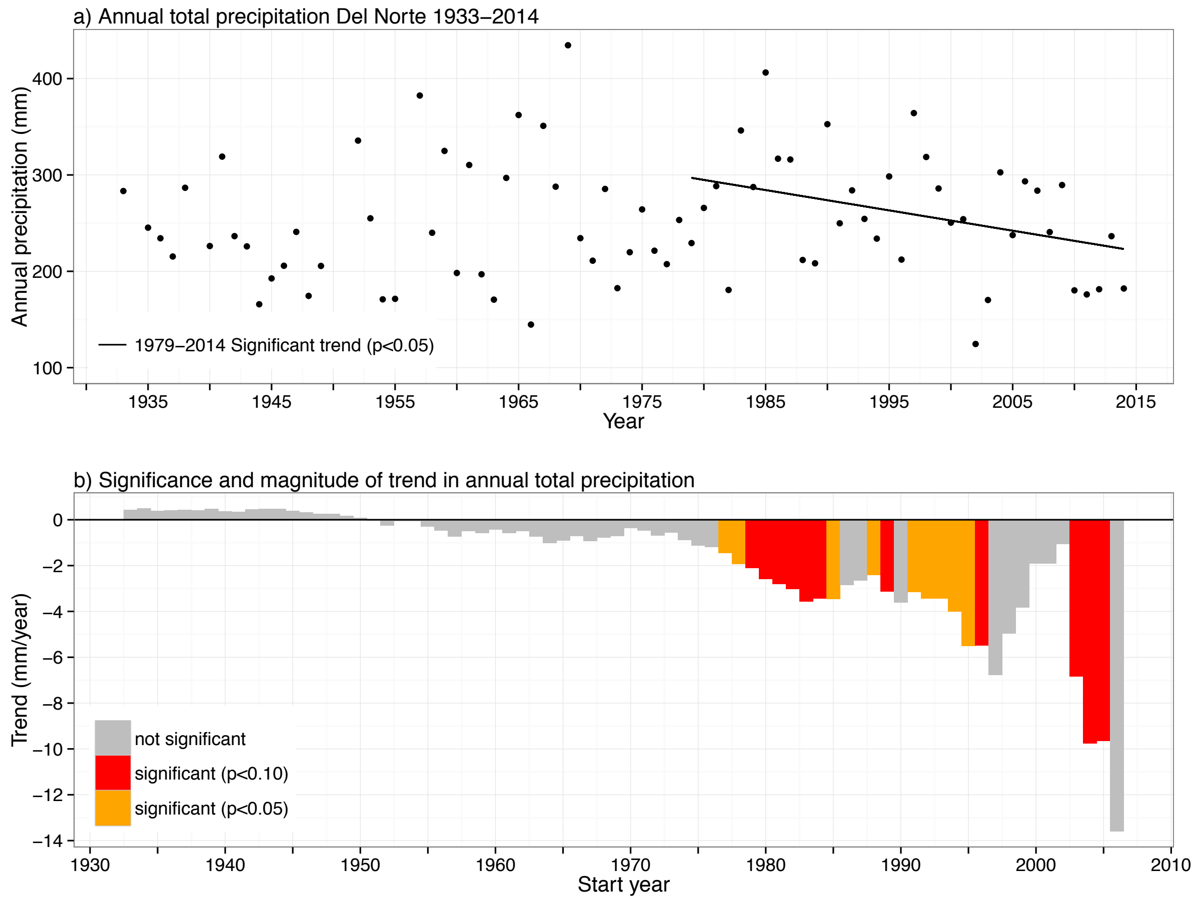


FIGURE 4.3a-(top) Annual total precipitation record for Del Norte in millimeters. The trendline is an example of a significant trend (p<0.05 level) resulting from the shortening of the record from 1933-2014 to 1979-2014 (slope of -21 mm/decade).

FIGURE 4.3b- (bottom) Plot displaying the significance and magnitude of trend in the Del Norte total annual precipitation record with increasingly shorter periods of record starting at the complete record of 1933-2014 and ending with the last testable period of 2006-2014. Of the testable record lengths, 15% were significant at the p<0.05 level.

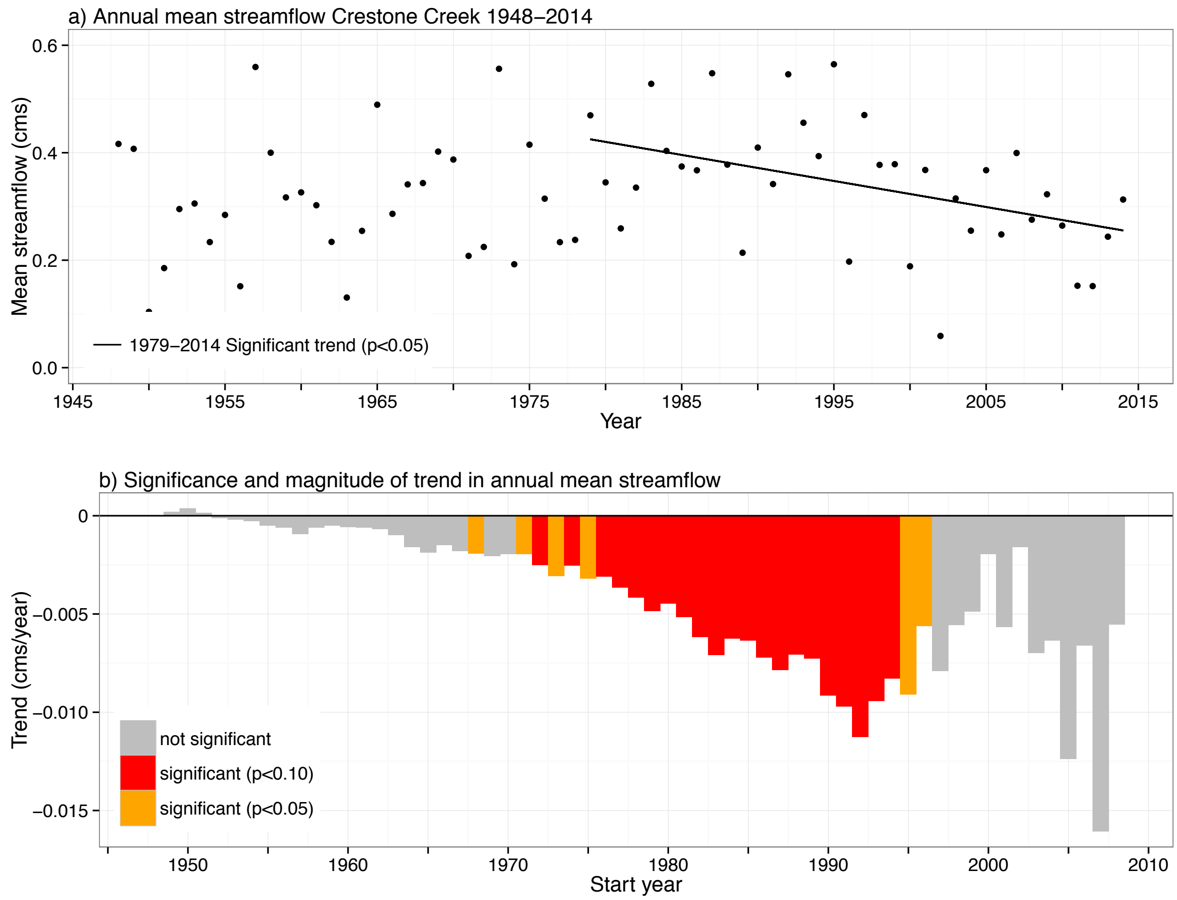


FIGURE 4.4a- (top) Annual mean streamflow record for Crestone Creek in cubic meters per second. The trendline is an example of a significant trend (p<0.05 level) resulting from the shortening of the record from 1933-2014 to 1979-2014 (slope of -0.04 cms/decade).

FIGURE 4.4b- (bottom) Plot displaying the significance and magnitude of trend in the Crestone Creek annual mean streamflow record with increasingly shorter periods of record starting at the complete record of 1948-2014 and ending with the last testable period of 2008-2014. Of the testable record lengths, 34% were significant at the p<0.05 level.

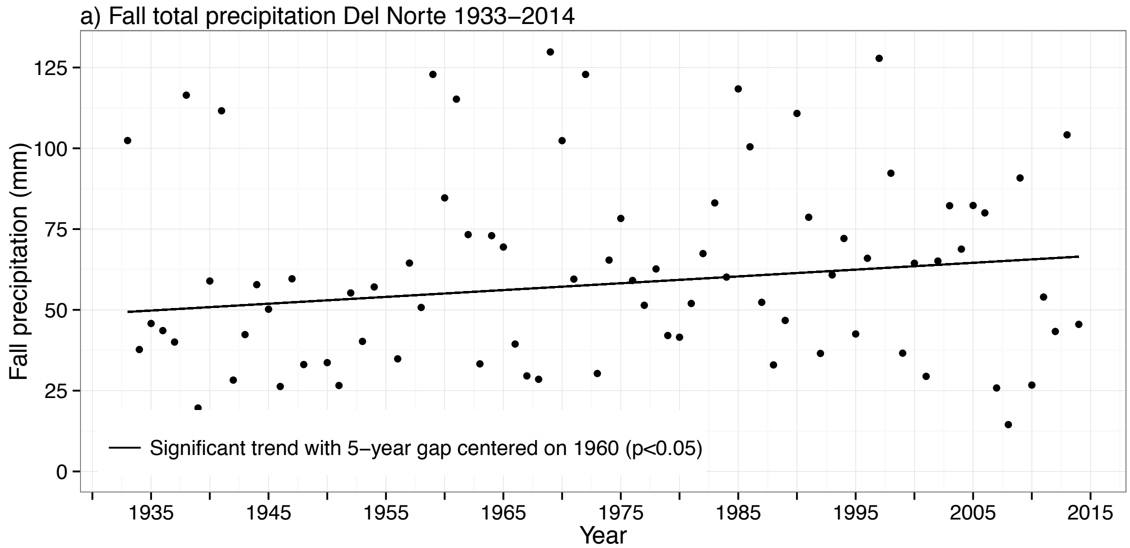


FIGURE 4.5a- Fall total precipitation record for Del Norte in millimeters. The trendline is an example of a significant trend ($p<0.05$ level) resulting from the addition of 5-year gaps in record centered at 1960 (slope of -2.1 mm/decade).

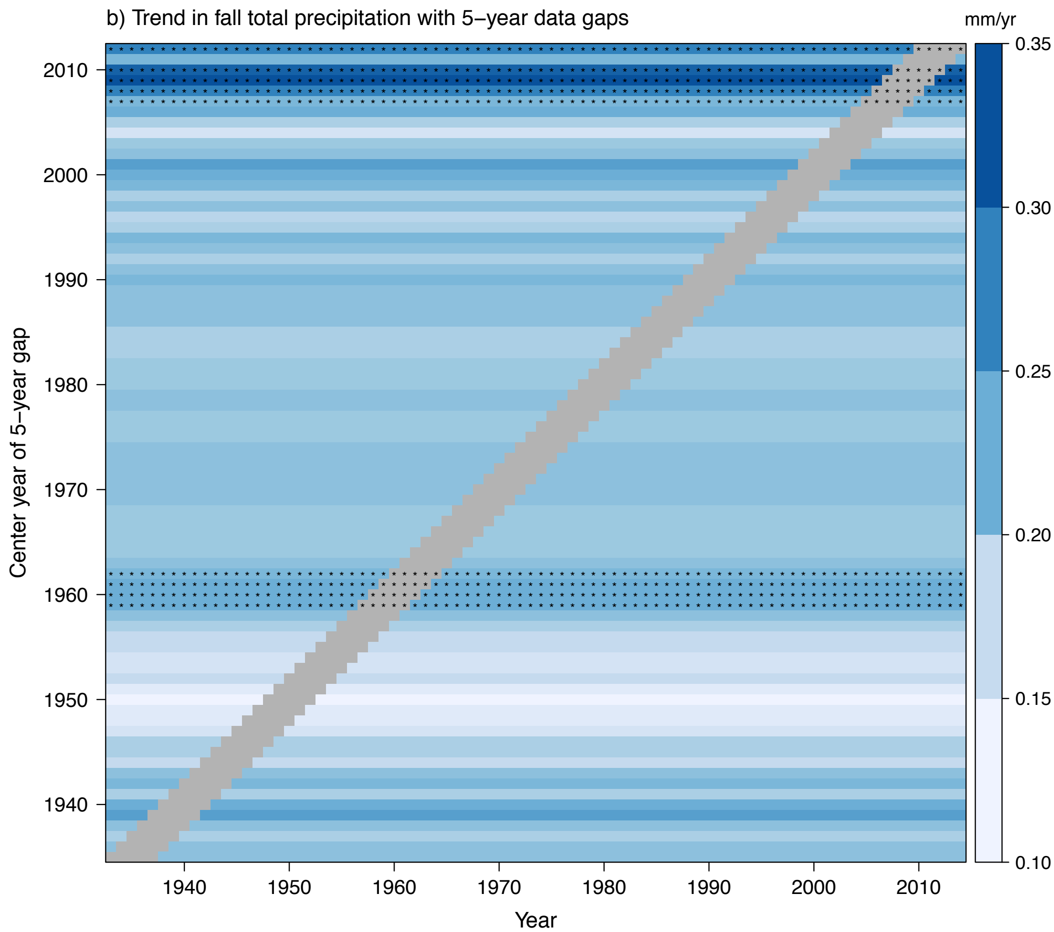


FIGURE 4.5b- Plot displaying the significance and magnitude of trend in the Del Norte fall annual precipitation record with gaps of 5 years added sequentially (grey areas) to the total record from 1933-2014. Of all the record lengths tested, 12% were significant at the p<0.05 level (denoted by stars on the plot).

REFERENCES

Alexandersson H, Moberg A, 1997. Homogenization of Swedish temperature data. Part I: Homogeneity test for lineal trends. *International Journal of Climatology* 17(1), 25–34, doi:10.1002/(SICI)1097-0088(199701)17:1<25::AID-JOC103>3.0.CO;2-J.

Askew AJ, 1987. Climate change and water resources. *The influence of climate change and climatic variability on the hydrologic regime and water resources*, Solomon SI, Beran M., Hogg W [Eds]. *IAHS Publication* 168, 421-430.

Batima P, Natsagdorj L, Gombluudev P, Erdenetsetseg B, 2005. Observed climate change in Mongolia. *Assessments and Adaptations to Climate Change* (AIACC), *www.aiaccproject.org*. Accessed 2011-11-21.

Beaulieu C, Ouarda T, Seidou O, 2007. A review of homogenization techniques for climate data and their applicability to precipitation series. *Hydrological Sciences Journal-Journal des Sciences Hydrologiques* 52, 18–37, doi:10.1623/hysj.52.1.18.

Burn DH, and Hag Elnur MA, 2002. Detection of hydrologic trends and variability. *Journal of Hydrology* 255, 107–122, doi:10.1016/S0022-1694(01)00514-5.

Chen Z, and Grasby S, 2009. Impact of decadal and century-scale oscillations on hydroclimate trend analyses. *Journal of Hydrology* 365(1-2), 122-133, doi:10.1016/j.jhydrol.2008.11.031.

Davi NK, Jacoby GC, Curtis, AE, and Baatarbileg N, 2006. Extension of drought records for Central Asia using tree rings: West-Central Mongolia. *Journal of Climate*, 19, 288–299, doi:10.1175/JCLI3621.1.

Easterling DR, Peterson TC, 1995. A new method for detecting undocumented discontinuities in climatological time-series. *International Journal of Climatology* 15, 369–377, doi:10.1002/joc.3370150403.

Endo N, Kadota T, Matsumoto J, Ailikun B, & Yasunari T, 2006. Notes and correspondence: Climatology and trends in summer precipitation characteristics in Mongolia for the period 1960–98. *Journal of the Meteorological Society of Japan*, 84(3), 543–551.

Fassnacht SR, Sukh T, Fernandez-Gimenez ME, B Batbuyan, Venable NBH, Laituri M, and Adyabadam G, 2011. Local understanding of hydro-climatic changes in Mongolia. Proceedings of *H02 Symposium Cold Region Hydrology in a Changing Climate* held during IUGG2011 in Melbourne, Australia, July 2011, *IAHS Publication* 346, 120-129.

Fassnacht SR, Venable NBH, Odgarav J and Adyabadam G, 2015. A journey down the Tuin: The hydraulics of an internal draining river from the Khangai Mountains to the Gobi Desert. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 99-103.

Fernandez-Gimenez ME, 2000. The role of Mongolian nomadic pastoralists' ecological knowledge in rangeland management. *Ecological Applications* 10, 1318–1326, doi:10.1890/1051-0761(2000)010[1318:TROMNP]2.0.CO;2.

Gallagher C, Lund R, Robbins M, 2013. Changepoint detection in climate time series with long-term trends. *Journal of Climate* 26, 4994–5006, doi:10.1175/JCLI-D-12-00704.1.

Harris I, Jones PD, Osborn TJ, and Lister DH, 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, 34, 623–642, doi:10.1002/joc.3711.

Helsel DR, and Hirsch RM, 2002. Statistical methods in water resources. *Techniques of Water-Resources Investigations of the United States Geological Survey, Book 4, Hydrologic Analysis and Interpretation, Chapter A3. http://water.usgs.gov/pubs/twri/twri4a3/*. Accessed 2014-11-02.

Jamiyansharav K, 2010. *Long-term analysis and appropriate metrics of climate change in Mongolia*. Unpublished PhD Dissertation, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA.

Karl TR, Williams CN, 1987. An approach to adjusting climatological time series for discontinuous inhomogeneities. *Journal of Climate and Applied Meteorology* 26(12), 1744–1763, doi:10.1175/15200450(1987)026<1744:AATACT>2.0.CO;2.

Kendall MG and Gibbons JD, 1990. *Rank Correlation Methods*. 5th edition, Edward Arnold, Hodder and Stoughton, London, UK.

Konagaya Y, 2013. The impact of agricultural development on nomadic pastoralism in Mongolia. *The Mongolian ecosystem network environmental issues under climate and social changes*, Fujita N, Maekawa A, Yamamura N [Eds]. Ecological Research Monographs. Springer, Tokyo, 255-268.

Kundzewicz ZW and Robson AJ, 2004. Change detection in hydrological records—a review of the methodology. *Hydrological Sciences Journal* 49, 7–19, doi:10.1623/hysj.49.1.7.53993.

Liu YY, Evans JP, McCabe MF, de Jeu RAM, van Dijk AIJM, Dolman AJ, Saizen I, 2013. Changing climate and overgrazing are decimating Mongolian steppes. *PLoS ONE* 8(2) e57599, doi:10.1371/journal.pone.0057599.

Loew A, 2014. Terrestrial satellite records for climate studies: how long is long enough? A test case for the Sahel. *Theoretical and Applied Climatology* 115, 427–440, doi:10.1007/s00704-013-0880-6.

Lorenzo-Lacruz J, Vicente-Serrano SM, Lopez-Moreno JI, Gonzalez-Hidalgo JC, Moran-Tejeda E, Zabalza, J, 2012. Recent trends in Iberian river flows (1945-2005). *Journal of Hydrology* 414-415: 463-475, doi:10.1016/j.jhydrol.2011.11.023.

Ma X, Yasunari T, Ohata T, Natsagdorj L, Davaa G, Oyunbaatar D, 2003. Hydrological regime analysis of the Selenge River basin, Mongolia. *Hydrological Processes*, 17, 2929-2945, doi:10.1002/hyp.1442.

Mann HB, 1945. Nonparametric tests against trend. *Econometrica,* 13, 245–259.

Marin A, 2010. Riders under storms: Contributions of nomadic herder's observations to analyzing climate change in Mongolia. *Global Environmental Change,* 20, 162-176, doi:10.1016/j.gloenvcha.2009.10.004.

Ministry of Nature, Environment, and Tourism, 2010. Mongolia Second National Communication, Dagvadorj D [Ed]. Ulaanbaatar, Mongolia, 160 pp.

Pederson N, Leland C, Nachin B, Hessl AE, Bell AR, Martin-Benito D, Davi NK, 2013. Three centuries of shifting hydroclimatic regimes across the Mongolian Breadbasket. *Agricultural and Forest Meteorology*, 178, 10–20.

Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S, Auer I, Boehm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey T, Salinger J, Førland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker D, 1998. Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology* 18, 1493–1517, doi:10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T.

Pettitt AN, 1979. A non-parametric approach to the change point problem. *Applied Statistics* 28(2), 126–135.

R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, *http://www.R-project.org/*.

Ray AJ, Barsugli JJ, Averyt KB, Wolter K, Hoerling M, Doesken N, Udall B, and Webb RS, 2008. *Climate Change in Colorado: A synthesis to support water resources management and adaptation*. Report for the Colorado Water Conservation Board. University of Colorado, Boulder, 58 pp.

Sen PK, 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association,* 63, 1379-1389.

Suzuki Y, 2013. Conflict between mining development and nomadism in Mongolia. *The Mongolian ecosystem network environmental issues under climate and social changes*, Fujita N, Maekawa A, Yamamura N [Eds]. Ecological Research Monographs. Springer, Tokyo, 269-294.
Thiel H, 1950. A rank-invariant method of linear and polynomial regression analysis, I, II, III. *Proceedings of the Royal Netherlands Academy of Sciences* 53, 386-392, 521-525, 1397-1412.

Venable NBH, Fassnacht SR, Adyabadam G, S Tumenjargal, Fernandez-Gimenez M, and B Batbuyan, 2012. Does the length of station record influence the warming trend that is perceived by Mongolian herders near the Khangai Mountains? *Pirineos*, 167, 69–86, doi:10.3989/Pirineos.2012.167004.

Venable NBH, Fassnacht SR, and Hendricks AS, 2015. Spatial changes in climate across Mongolia. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig , Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 73-77.

Von Storch H, 1995. Misuses of statistical analysis in climate research. In *Analysis of Climate Variability: Applications of Statistical Techniques*, Von Storch H, Navarra A [Eds]. Springer-Verlag: Berlin, 11–26.

Yue S, Pilon P, Phinney B, Cavadias G, 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16, 1807–1829, doi:10.1002/hyp.1095.

CHAPTER 5- SYNCHRONY AND HETEROGENEITY OF 300+ YEARS OF RECONSTRUCTED STREAMFLOW IN THE KHANGAI MOUNTAIN REGION OF MONGOLIA

5.1 Summary

Sharply decreasing flows in rivers of the Khangai Mountain region over the last decade and a half raise concerns about long-term water availability in this part of Mongolia. The threat of continued climatic warming and increased resource development pressures, combined with the needs of local nomadic pastoralists provide reasons for investigating the long-term hydrologic variability of the limited surface water supplies of the region. Previous analyses of gridded and station-based hydroclimatic time series show greater decreases in moisture over the northern portions of the region than southern. Flows from four river basins that originate in the Khangai Mountains are analyzed along with tree-rings from ten sites in the region, two of which are new sites with cores collected in 2012. The Khanui and Khoid Tamir rivers flow north from the mountains into the forest steppe and steppe of central Mongolia and the Baidrag and Tuin Rivers flow to the south from the steppe into the desert steppe of the Gobi region. Cores from the two new tree-ring sites were processed with total ring-width measurements from the other eight moisture-sensitive sites using standard laboratory and statistical methods. Correlation analyses identified the strongest predictors for each basin. The short observational records (1976-2012) of the four streamflow gaging sites include missing data which were filled by multiple imputation/predictive mean matching methods using data from six meteorological stations in the region. Filled flow volumes are used with chronologies of total ring widths from selected tree-ring sites in multiple linear regression models to

reconstruct scenarios of past flow in each river basin. Estimates of model performance are given and the best-fitting models were additionally tested using cross validation methods due to the short observational record lengths. For comparison, results of model fits using differing correlated predictors are also shown. The recent low flow events of the observational record are placed in context with the 300+ year scenarios of past streamflow through a quantitative evaluation of extremes of wet and dry conditions in each basin and a qualitative analysis of synchrony of events. More heterogeneity in event timing is seen on an east-west basis than north-south, which may be due to differing climate signals affecting the predictors used for reconstruction. Other possible influences on the nature and synchrony of the reconstructed records include the effects of choice of chronology processing method, tree-growth morphology (stripbark effects), and repeated insect defoliation events common in the region through time. Synchrony of flow in the Khoid Tamir and Tuin Rivers are greatest of the models considered and are roughly correlative to many events noted in the Selenge River reconstructions from other dendroclimatic investigations in Mongolia. As the best predictors in the original reconstructions only extend the scenarios to 1998, prior to the largest drops in observational streamflow, extended models were developed using only those sites with tree-ring data extending to at least 2009. The Khanui, Khoid Tamir, and Baidrag River models created to 2008 have mean flows during the droughts of the 2000's that are at least as extreme in terms of dryness as the top 5 most extreme dry periods originally reconstructed for each basin. No extended model could be developed for the Tuin River due to poor model fits, but the observational streamflow record for the Tuin is not as dry in the 2000's as the most extreme events modeled for the basin over the last 300+ years up to 1998. The low flows in this basin during the most recent droughts are not as severe a drop from earlier conditions as seen in other basins in the region. In the case of the Khanui River, much more extreme drought conditions occurred in the past than those seen in the last decade and a half. The recent low flows were about a third of the volumes seen in wetter periods. The results of reconstruction of flows in the four Khangai Mountain region rivers studied here show that the drought events of the last decade and a half, while extreme, are not beyond the range of natural variability over the last 300+ years in these systems. Also, patterns of climate variability are more similar in the easternmost basins than the westernmost with slight variations in event timing from basin to basin, particularly between the Khanui and Baidrag, and together in comparison of those two rivers with the more similarly synchronous Khoid Tamir and Tuin Rivers. These findings have implications for development and management of water resources in these basins. Conditions like the most recent extreme dry periods and low flows have likely occurred in the past, and while not common, should be considered as plausible flow conditions for the future, particularly under a warming and possibly drying climate.

5.2 Background

Warming temperatures and increasing resource development pressures have the potential to reduce the already limited surface water supplies of central and western Mongolia (Ministry of Nature, Environment and Tourism, 2010; Priess *et al*., 2011). Significant increases in temperatures and decreases in summer rains have been observed by nomadic pastoralists of the Khangai Mountain region and their observations are supported in part by the results of trend analyses of station-based climate records from the area (Fassnacht *et al*., 2011; Venable *et al*., 2012) (also see Chapter 4). Record low flows in rivers over the last decade and a half and the drying of local lakes and springs are also a concern for herders in this region (Ministry of Nature, Environment, and Tourism, 2010; Fassnacht *et al.*, 2011; Tao *et al.*, 2015).

The Khangai Mountain region as described here encompasses roughly 100,000 square kilometers in three different *aimags*; Arkangai to the north, Bayankhongor to the south and west, and Ovorkhangai to the south and east (Figure 5.1). Four rivers flow to the north and south from the divide of the Khangai Mountains. The Khanui and the Khoid Tamir Rivers flow to the north and east through forest steppe and steppe landscapes. The Khanui River (R.) gage at Erdenemandal is at 1487 meters above sea level, and the Khoid Tamir R. gage at Ikhtamir is at 1740 meters above sea level (Table 5.1, Figure 5.1). Annual precipitation averaged over the length of the available records from three meteorological stations near the northern basins is 275 millimeters. The Khanui River joins the Selenge River and the Khoid Tamir River joins the Orkhon River, downstream. The confluence of both rivers is near the northern border of Mongolia and flow continues into Lake Baikal in Russia. The Baidrag and Tuin Rivers flow from the Khangai Mountains south into the drier steppe and desert-steppe terminating at Buuntsagaan and Orog Lakes, respectively. The elevation of the Baidrag R. gage at Bayanburd is 1880 meters above sea level, and the Tuin R. gage at Bayankhongor is at 1863 meters above sea level (Table 5.1, Figure 5.1). Average annual precipitation from stations near these basins is 200 millimeters per year. The locations of six meteorological stations used in filling missing streamflow data and the four streamflow gaging sites used for reconstruction are plotted with the river basins in Figure 5.1.

Analyses of grid-based temperature and precipitation products from the period of 1963-2012 show statistically significant warming across all of Mongolia and distinct patterns of summer drying in the northern parts of the Khangai Mountain region but no significant summer drying in the southern parts of the region (Venable *et al*. 2015). A majority of the rainfall in Mongolia occurs in the summer months of June, July, and August and the growth of vegetation to support herders' livestock in the region is closely tied to seasonal, annual, and extreme moisture conditions (Yu *et al*., 2003; Batima *et al*., 2005; John *et al.*, 2013).

Trees in this semi-arid region also respond strongly to summer moisture conditions and the width of tree-rings from selected sites reflect hydroclimatic variability over the last several hundred years (e.g., Jacoby et al., 1999). In this work the recent low streamflow events seen in the observational record are placed in context with scenarios of past streamflow for the four river basins of interest. Patterns and timing of wet and dry conditions are explored with a focus on the variability between models using different tree-ring predictors for reconstructions over the last 300+ years. Of particular interest is determining if the north to south differences in historical trends are observed in the reconstructed records of streamflow.

5.3 Methods

5.3.1 Hydroclimate Data

Hydroclimatic records in Mongolia are generally short with periods of missing data and are from stations that are sparsely distributed across the country. This is true for records in the Khangai Mountain region with days, months, and even entire years of data missing. Existing streamflow data from the four basins of interest and six regional precipitation gauges were provided by the Mongolian Research and Information Institute of Meteorology, Hydrology, and Environment (RIIMHE) (Table 5.1). Data were examined to evaluate the distributions of the time series, the amount of autocorrelation or persistence in the data, and were tested for correlations between the variables of interest. Data were visually checked for errors and outliers, and most found were associated with decimal point conversions from commas. The threshold for streamflow data outliers were values that fell more than 3.0 standard deviations from the mean (e.g., Harris, *et al*., 2014). For precipitation, a threshold of 4.0 standard deviations from the mean was considered appropriate (e.g., Harris *et al*., 2014). The few values of streamflow or precipitation that occurred beyond these bounds were not considered to be outside the range of values possible in these highly variable natural systems.

Daily values were aggregated to a monthly timestep. As no overall distinction was made between zeros or missing values for precipitation, when at least one day of precipitation was recorded the month was not considered missing. For streamflow, daily values were more regularly recorded. For aggregation purposes, any month with less than 23 days (~75%) of data was considered missing (e.g., Harris *et al*., 2014). A maximum amount of aggregated streamflow and precipitation data were available for the basins of interest between the years of 1976 and 2012. Records were truncated to this period of analysis to aid in the comparison of results between basins. Mean monthly streamflow values given in meters per second (m^3/s) were converted to volumes of million cubic meters (MCM) per month. Due to the high levels of skewness and outlying values inherent in hydroclimatic datasets, particularly in those from semi-arid regions, all aggregate streamflow and precipitation data were transformed using the square root function to better meet the assumptions needed for use in the regression-based reconstruction analyses (Helsel and Hirsch, 2002). Estimated streamflow values from reconstruction modeling were back transformed to original units (MCM) for interpretation of results.

The four basins of interest had differing amounts of missing data (Table 5.1). Entire years of data could be missing. For example, no streamflow data were reported for the Khoid Tamir R. at Ikhtamir for the years of 1979-1981, 1995, 2006-2008, and 2011. Some of the same years (1979-1981, 2006) were missing for two of the other rivers under study. The transformed volumetric streamflow values were filled using the transformed monthly station-based precipitation data via multiple imputation or Predictive Mean Matching (PMM) methods (e.g., Rubin, 1987, van Buuren and Groothuis-Oudshoorn, 2011). The hydroclimate variables were not lagged prior to imputation as three of the four basins had streamflow and precipitation correlations that were highest in the current month. In the fourth basin (Tuin R. at Bayankhongor), correlations were only slightly higher with a lag of one month between occurrence of precipitation and resultant streamflow. A similar lag between precipitation and streamflow was found for the Kherlen River in eastern Mongolia (Pederson *et al.*, 2001). The PMM method imputes values in a column of incomplete data through iteration given the other columns of variables in the dataset. The order of imputation is controlled through correlation of the predictor variables. Though regression is used as a metric for selecting possible values for missing variables that are close to those predicted for non-missing variables, the process of imputation is not regression-based. It draws values from a pool of plausible predictors generated from donors with a similar predictive mean. When close cases are identified, a random draw from those cases results in filling the missing value with the observed value of the chosen case (Rubin, 1987; van Buuren, 2012).

Resulting filled streamflow values follow a similar distribution and range as other values in the dataset. No core mathematical theory defines PMM and assessment of method skill relies on Monte Carlo methods, but it is thought to be robust to model misspecification from non-normality, heteroscedasticity of residuals, and non-linear relations (van Buuren, 2012; Morris *et al*., 2014). Five hundred imputations of streamflow were conducted and the mean of those imputation results were used as plausible values to fill missing cases in the final streamflow time series. Due to the amount of missing station-based precipitation data the non-parametric, distribution-free Kolmogorov-Smirnov test for was run to compare the robustness of the imputation method using predictors with missing values (the station-based data) to predictors with no missing values (results not shown). The non-missing predictors were interpolated monthly gridded precipitation values from the Global Precipitation Climatology Centre (GPCC) (Schneider *et al*., 2014). The GPCC values from 1976-2010 were extracted over each basin of interest using a weighted mean function. Differences in the imputation process using transformed and untransformed values were also tested. No significant differences in imputation results were found. Precipitation values used in the imputation process were highly correlated from basin to basin (Pearson's r=>0.92). Final filled streamflow values were also well correlated with values ranging from r=0.72 to r=0.83.

5.3.2 Tree-Ring Data

Tree-ring data from Siberian larch, *Larix sibirica*, were collected from two new sites at Jargalant Bag (JGB) and Khuush Uul (KHU) in the Khangai Mountains of central Mongolia in the summer of 2012 to complement existing tree-ring datasets in the region (Figure 5.1, Table 5.2). Seven other chronologies used in this study were created from raw ring width records archived in the International Tree-Ring Data Bank (ITRDB). All sites are Siberian larch, *Larix sibirica* with the exception of the Khorgo Lava Pine (KLP), which are Siberian pine, *Pinus sibirica.* Many were collected as part of the Mongolian-American Tree Ring Project (D'Arrigo *et al*., 2000; Davi *et al*., 2006; Cook *et al*., 2010; Pederson, *et al*., 2014). Of the seven sites, two were from the Khorgo Lava site, Pine (KLP) and Larch (KLL). The others were Mandal Hill Monastery (MHM), Namyin Davaa Bayankhongor (NDB), Suuleen Bagtraa (SLB), Zuun Salaa Mod (ZSM), and Zurkh Togol (ZTG). One additional set of ring widths, Orkhon Gol Hushree (OGH), were contributed by another researcher (Leland, pers. comm., 2014) (Figure 5.1).

The existing core sets from the ITRDB were collected from moisture-sensitive sites ranging from those on xeric lava flow substrates (KLL- Davi *et al*., 2006; KLP- Pederson *et al*., 2014), to open canopy forests with sparse or grassy understories (ZSM-Jacoby *et al*., 1999; Davi *et al*., 2006), and one high elevation site where temperature may have a greater influence on growth than moisture (SLB- D'Arrigo *et al*., 2000). The higher elevation (2500 m) site was included here as it was shown in previous research to correlate moderately well with precipitation from stations in the southern Khangai region (Fitzsimmons and Venable, 2015; Venable and Fassnacht, 2015b). The JGB site has an average elevation of 2556 m (ranging from 2414 m to 2706 m) as a result the JGB site may be more temperature sensitive than moisture sensitive, though that relation is not investigated here.

The sites selected for analysis display differing levels of significant correlation of ring widths ($p>0.10$ level) with one another. As expected, some sites geographically near each another were highly correlated (KLL and KLP, r=0.89), while other sites that were geographically distant from one another were still significantly correlated but at a much lower level (OGH and ZSM, r=0.58). Other sites closer to one another were not significantly correlated, which may be due to differing climatic influences such as temperature relations or regional moisture patterns (e.g., SLB and KLP; OGH and NDB/ZTG) (D'Arrigo *et al*., 2000, Leland *et al*., 2013).

5.3.2.1 Core Processing

At the new sites, at least two or more cores were collected from each tree using a 4.3mm or 5.15 mm diameter increment corer. Over 40 cores were collected per site. The sites were generally steep and rocky with primarily open canopy conditions. On north-facing and lower angle slopes at one site however, canopy conditions were generally closed, particularly where the oldest trees were found. Trees exhibiting old age characteristics such as larger tree diameters and lower branches were sampled in both sites (Swetnam and Brown, 1992). Most of the older trees exhibited heart rot limiting the length of the series collected. Tree cores of a range of ages were collected when possible to aid cross dating in the last few decades of dry conditions in the region. One stand was found to be essentially two-aged upon analysis and the cores were divided into two sets, retaining the older core series in the final chronology. Disturbances at the new JGB and KHU sites were limited to a few fire scars at higher elevations at one site (JGB) and to the presence of small pockets of dead trees at the other site (KHU), with mortality possibly caused by recent insect outbreaks. Sampling was focused on healthy-appearing trees. Insect infestations have affected many stands throughout Mongolia, with outbreak intensities increasing during the recent droughts (Ghent and Ohnken, 2004). Few stumps were noted in either stand.

Cores were air-dried then mounted and sanded for inspection following standard methods (Stokes and Smiley, 1996). Wood anatomical anomalies such as collapsed cells, narrow latewood, and false rings were observed in several of the individual cores from the new sites as observed in samples from other sites in Mongolia (De Grandpré et al., 2011; Khishigjargal et al., 2014). Measurements of cores were completed on a Velmex measuring stage to the nearest 0.001 mm of precision. Cores from part of one of the existing sets however, were processed using 0.01 mm precision (Leland, pers. comm., 2015). Visual cross dating (Stokes and Smiley, 1996) was used for all cores and COFECHA software (Holmes, 1983; Grissino-Mayer, 2001) was used to statistically check the cross dating of the archived cores. Similar functions in the R software package, dplR were used to check accuracy of the cross dating of the two new core sites and to perform all other tree-ring analyses in this study (Bunn, 2008; Bunn, 2010; R Core Team, 2015). The characteristics of the other eight chronologies were also visually and statistically assessed, as the original wood cores were not available for examination.

As the longest resolvable climate fluctuation information in a series is related to the length of the series, cores with about 200 years or more of record were preferentially selected for analysis (Cook *et al*., 1995). This rule resulted in median series lengths of 300 to nearly 400 years per site, with the exception of the KHU site where only about 100 years of information were available (Table 5.2). A few series were truncated earlier in the time series due to poor correlation between series, or in some cases when very long chronologies included remnant wood measurements not used for these reconstructions.

5.3.2.2 Tree-Ring Indices

Time curve standardization methods were employed for generating dimensionless indices from the ring-width series by fitting a curve, a negative exponential function for example, to each series then dividing the ring-width values by the corresponding curve value to remove growth effects and generate growth indices (Fritts, 2001). To reduce non-climatic effects on tree growth from endogenous or exogenous disturbances a cubic smoothing spline was applied to all series (Cook and Peters 1981; Cook *et al*., 1990). For reproducibility, a uniform spline length of 220 years with a frequency cutoff of 50% was applied to every series based on 67% of the average length of all the series used in the study. Due to the flexibility of the spline fit and based on testing with a subset of the data, additional tree-ring series processing methods such as power transformation of residuals or double-detrending of series as used in other studies were not employed despite the generally small size of the tree-rings and disturbance features present in some of the series (Cook and Peters 1997; Pederson *et al*., 2013). The uniform application of the spline to all series may result in some loss of long-term low frequency climate signals or insufficient removal of disturbance signals.

Standard and prewhitened (residual) chronologies were calculated using a bi-weight robust mean function with a constant value of nine, following the methods of Cook *et al*. (1990). The latter chronologies were created through application of an autoregressive model selected by Akaike's Information Criterion (AIC) and removing the low-order persistence in each series prior to averaging (Kutner *et al*., 2005; Bunn, 2008). ARSTAN chronologies were not used in this analysis, though they can minimize the effects of endogenous disturbance through application of an autoregressive model to pool and retain the common autoregressive signals in the collection (Cook, 1985). Variance stabilization was applied to the chronologies to reduce changes in variance resulting from the number of series contributing to the mean chronology and their temporal dependence through time (Osborn *et al*., 1997, Frank *et al*., 2006). Windows of 50-years without overlap were used for correlation computation (rbar, e.g., Wigley *et al*., 1984) and the chronologies were adjusted by multiplying by the square root of the effective sample size (Frank *et al*., 2006). The adjusted chronologies were then rescaled using the reciprocal of the rbar used at that time step (Osborn *et al*., 1997). To retain a high amount of common signal in the chronologies, the adjusted time series were truncated when the between-tree expressed population signal (EPS) fell below 0.85 (Wigley *et al*., 1984). The median series lengths used to create most of the chronologies were over 300 years. Excluding the shorter KHU chronology, the 350-year common period of overlapping chronologies extends from 1650 to 1999 (Table 5.2).

5.3.3 Modeling

A screening method with a level of $p<0.05$ was used for detecting significant correlations between the standard chronologies and streamflow. Standard chronologies were used as the persistence (autocorrelation) better matched that observed in the transformed hydroclimate records. To maximize the length of the reconstructions, the shortest period chronology of KHU was not included in these analyses, despite strong correlations of ring-with indices to regional monthly (not shown) and seasonal hydroclimate variables (Figures 5.2a and 5.2b). Correlations between each of the remaining chronology indices and transformed hydroclimate variables for the months of May through September were used to inform the selection of predictors for multiple linear regression model development. As tree growth responds to previous year's growing conditions, lagged variables were also compared (Fritts, 2001).

Due to the very short length of the observational streamflow records, training models were calibrated on the entire common period of the tree-ring predictors and streamflow records (22-32 years with lagged predictors). Several methods were used to develop parsimonious models including forward stepwise searches using Aikake's Information Criterion (AIC); LEAPS, an exhaustive search best subsets technique using branch and bound algorithms; and the use of the best-correlated chronologies as previously mentioned. Training models were evaluated using multiple and adjusted R^2 values; SE, the standard error of the estimate; Mallow's Cp (a measure of variable multicollinearity and bias); the PRESS criterion (a gauge of prediction error); and variance inflation factors (VIF) to assess multicollinearity of variables (Kutner *et al*., 2005). Residuals were checked graphically for normality and constancy of variance and using the Kolmogorov-Smirnov test (Kutner *et al.,* 2005). The final eleven test models were developed from the most robust training models identified using these methods. The best models identified had either two or three predictors, eight contained the OGH predictor and seven contained the JGB predictor.

The best-fitting models for the common period were tested using leave-one-out cross validation methods. Additional bias and variance inflation may be introduced when using cross validation procedures for model validation, as compared to that from longer time series using conventional split calibration/validation methods (Kutner *et al*., 2005; Arlot and Celisse, 2010). Estimates of performance including the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), and adjusted R^2 values (adjR^2) were calculated for the initial calibration models and the cross-validated calibration models. Synchrony of wet (pluvials) and dry (droughts) conditions through time in the four basins was studied using ranked non-overlapping 5-year periods.

5.4 Results

5.4.1 Correlation and Regression Analyses

On a monthly basis, the results of the simple correlations (not shown) between the tree-ring indices at each tree-ring site and streamflow for each of the four basins was broadly similar, with the highest correlations for the greatest number of sites expectedly occurring in the current year summer months of June and July, when precipitation is highest, and to a lesser extent in August. Peak streamflows in the four basins generally occur around August due to summer rains, but soil moisture in these headwater systems is also replenished by snowpack, with one site (Baidrag R. at Bayanburd) exhibiting dual peak flows occurring in May and August. The significant positive correlations also extended on both sides of the summer months to the current year spring months of April and May and into the fall month of September. Significant positive correlations were also found in the fall months of October, November, and the winter month of December of the current year, particularly at the northern tree-ring and streamflow sites. The growth relation of these values to current year tree-ring width is not known as the growing season at most sites in central Mongolia is approximately June to late August (Davi *et al.*, 2006).

Seasonal correlations (significant positive previous Summer: June, July, August; current Spring: March, April, May; and current Summer: June, July, and August correlations shown) generalize monthly relations revealing differing patterns depending on site. For example, precipitation extracted from the GPCC grids for the area over the north-flowing

Khoid Tamir R. at Ikhtamir correlates positively during the growing season for several tree-ring sites (Figure 5.2a). Correlations are slightly higher over a longer period for streamflow than for precipitation. In contrast, precipitation values for the Tuin R. at Bayankhongor correlate more highly with growth at several tree-ring sites in the previous year than the current year and only one site has significant positive correlations for streamflow (Figure 5.2b). Based on these results and previous reconstructions of flow in other Mongolian basins, the dendroclimatic year (previous August through current July) (e.g., Pederson *et al.*, 2001) was used as the timestep for reconstruction.

5.4.2 Streamflow Reconstructions

Following the correlation results, unique models were constructed for each basin to strengthen model fits and minimize errors. Regression models contained between two and three predictors and also included year t and year t+1 lagged predictors. Regression coefficients were used to create reconstructed time series of streamflow for each basin. These scenarios of flow were different lengths depending upon the overall length of the predictors used in the model, but they cover the nine chronology common period for the unlagged and lagged predictors of 1650-1998.

The statistical model results for the four basins of interest are summarized in Table 5.3. Several measures of model fit are presented, including the adjusted R^2, which is the coefficient of multiple determination for measuring reduction in the variation of the dependent variable given the set of independent variables used for prediction adjusted by the number of predictors used in the model (Kutner *et al.*, 2005). The Root Mean Squared (RMSE) and Mean Absolute Error (MAE), both of which describe model performance error, are also included. The RMSE and MAE express the average model prediction error in terms of million cubic meters (MCM) of streamflow. The magnitude of RMSE varies depending upon the time domain of the model and the distribution of the error magnitudes (Willmott and Matsuura, 2005). It is however, a widely reported and useful statistic when comparing models from similar domains. The MAE is a measure of average model error magnitude and is more comparable from model to model over differing domains (Willmott and Matsuura, 2005). Including both statistics provides a better estimate of model error, especially when considering the distribution of errors, or the effects of outliers on the magnitude of RMSE (Chai and Draxler, 2014).

Due to the lack of large differences between the calibration and verification statistics, the cross validation results for the adjR^2, RMSE, and MAE are given in Table 5.3. The model explaining the greatest amount of variance was for the Khoid Tamir R. at Ikhtamir. The adjR^2 was 0.69, with RMSE and MAE values of 60.7 and 47.7 MCM, respectively. Models for the Khanui R. at Erdenemandal and for the Baidrag R. at Bayanburd had similar errors and slightly lower adjusted R^2 values (Table 5.3). The basin model with the lowest errors (RMSE=23.4 MCM, MAE=19.9 MCM) was for the Tuin R. at Bayankhongor, but it explained substantially less of the variance averaged over the multiple leave-one-out periods of cross validation (adjR^2=0.26).

5.4.3 Drought and Pluvial Timing Between Basins

Hydroclimatic state is approximated by conditions above (pluvials/high flows from increased rainfall) or below (droughts/low flows from decreased rainfall) the long-term mean of the reconstruction. The results of calculating ranked 5-year non-overlapping means (e.g., Pederson *et al.*, 2001) for wet and dry conditions in each basin are given in

Table 5.4. Agreement in scenario state for the top ranked extremes from basin to basin varies through time.

In the reconstructions, the late 1600's were a period of extreme drier conditions, with low flows particularly in the Khanui River basin, but much wetter conditions and higher flows prevailing by the early 1700's in the Khanui, Khoid Tamir, and Tuin River basins (Figure 5.3). The mid to late 1700's were very dry in the Khoid Tamir and Tuin basins, with wetter conditions by the very end of that period in those two basins. Dry, low flow conditions persisted at that time in the Baidrag River Basin. The early to mid 1800's were drier in most basins with extremes of drought in all river basins by the end of the 19th century. The last 20 years of that century saw a swing from the top ranked wet interval to the top ranked dry interval for the Baidrag River basin. The early 20th century was dry in all of the basins, with a transition to wet, high flow conditions by the mid-1900's particularly in the Baidrag River basin. At that time however, dry conditions were found in the Khanui River basin. By the late 20th century, most sites experience wet, record high flow conditions. These above long-term mean flows are seen in the observational record, with a downturn in flow regimes at all sites by the very end of the 20th century (Figures 5.3-5.6).

5.5 Discussion

5.5.1 Spatial Variability and Model Predictors

Data quality and modeling choices in terms of predictor selected (tree-ring site indices) affect the resultant reconstructions. The sensitivity of a particular site to antecedent and current soil moisture conditions and differences in response between sites necessitates the use of multivariate methods (Fritts, 2001). Principal Components Analysis regression (PCA) is widely used in dendroclimatological analyses to identify common modes of variation and to reduce the effects of multicollinearity among predictors. In this work, careful selection of predictors for multiple regression through application of model diagnostics results in robust model fits for most basins where there are strong correlations between moisture conditions and growth while reducing the effects of multicollinearity on model results (Figures 5.2a and 5.2b, Table 5.3).

Some tree-ring sites in the chosen network are well correlated with each other. Often the well-correlated sites are geographically near one another, as with the JGB and ZTG sites (r=0.69). These sites are within roughly 20 kilometers of each other in the upper reaches of the Tuin River basin and are each a predictor in two different robust streamflow models for the Khoid Tamir R. at Ikhtamir. Figure 5.4 provides evidence of the small effect the choice of two well-correlated predictors have on reconstruction results. While the time series are different, both display similar trends, fits to the observed data, and mean values of 223.0 and 215.5 MCM respectively, for Model 1 (OGH, SLB, and ZTG predictors) and Model 2 (OGH, SLB, and JGB predictors) (Figure 5.4). Statistics for Model 1 are given in Table 5.3. Model 2 has an adjusted R^2 of 0.61, a cross validation RMSE of 60.7 MCM, and a cross-validation MAE of 45.8 MCM. Either predictor would be suitable in the final model chosen for reconstruction; the final choice of OGH, SLB, and ZTG was based on slightly improved model statistics and a moderately longer chronology length.

It is expected that robust, yet alternative models with similar and/or well-correlated predictors would provide similar results. However, testing results with alternatives of less robust models with much less correlated differing predictors yielded similar results (Figure 5.5). Two models for the Tuin River at Bayankhonger had two of the same predictors (JGBp and SLBp, where p is previous year), with a third predictor that differed.

In Model 1, (see Table 5.3 and Figure 5.3) the differing predictor was OGH. In the alternative model it was replaced with MHM; correlation between OGH and MHM was only r=0.38. The two sites OGH and MHM are 140 kilometers apart east to west on opposite sides of the hydrologic and geographic divide between the northern side of the Khangai Mountains and the southern side. The adjusted R^2 values for each model were 0.26 and 0.28 and mean flow values over the length of each record (1650-1998) were 85.4 MCM and 85.0 MCM, for Model 1 and Model 2 respectively. Despite the low correlations and poor model fits, the signal between the two models is strong.

5.5.2 Spatio-Temporal Reconstruction Synchrony

The results of the alternate model comparisons show a very high level of synchrony between the different models for each basin, and between the Khoid Tamir and Tuin River basins. Less synchrony is found for the Baidrag and Khanui Rivers though predictors used for these models are some of the same as for the other models (OGH, JGB, and SLB). They also include predictors from sites much further east and north of the other sites (ZSM and KLL) (Table 5.3).

Comparisons of hydroclimate patterns suggest distinct heterogeneity in drought and pluvial event timing across the common reconstruction period (1650-1998) for the Khanui and Baidrag Rivers versus the Khoid Tamir and Tuin Rivers (see Figure 5.3). Times when high or low flow events in the Khanui and Baidrag River basins were very different than the Khoid Tamir and Tuin Basins occurred in the mid 1700’s, and the early to mid-1900’s. There are some minor distinctions between event timing for the Baidrag and Khanui River reconstructions particularly in the mid-1800’s with slightly above mean flows in the Khanui River and below mean flows in the Baidrag River basin (Figure 5.3).

These results suggest similar tree growth and resulting reconstructed hydroclimate patterns for the two easternmost basins (with more persistent periods of wetter or drier conditions). They also show a level of coherence in reconstructed patterns with greater variability between dry and wet conditions seen in the westernmost basins. There are also slight differences from north to south between event timing, especially in the Khanui and Baidrag River basins. Explicit differences between high and low flow event timings in the basins may also be a result of exogenous effects on tree growth aside from climate effects. Differences in the late 20th century could also be related to end of record effects from detrending when creating the chronologies (Cook *et al*., 1990).

Insect defoliation events are present in the SLB site chronology and exist in some of the other chronologies. Tree growth is suppressed in the years following defoliation mimicking low moisture conditions. Events tend to be cyclic, occurring historically in Mongolia every 10-12 years (Ghent and Onken, 2004). The time between outbreaks has been decreasing however, with more trees succumbing to defoliation and eventual mortality with increasingly warmer temperatures, disturbance from wildfires, logging, and livestock grazing (Ghent and Onken, 2004; Hessl *et al*., 2012).

Three of the four original models use the SLB predictor (Table 5.3, Figure 5.3). More conservative detrending processes are inadequate in removing the influence of these disturbances and even spline fit detrending methods as used here may not be adequate without unduly affecting climate signals. Evidence of recent insect defoliation was noted in the mountain forests near where the JGB and KHU cores were collected. The KHU site in particular had dead trees within sight of the collection area and despite the young age of most of the trees sampled at that site, many missing and narrow rings were noted in the last decade of growth. In the field, it was uncertain whether these patterns were primarily related to the effects of insect defoliation, but they are also similar to patterns of low growth observed during this time period in other core sets that extend into the last decade such as OGH and KLP (Figure 5.6). The reduced growth patterns observed in the last decade or two of the KHU core, are likely a result of the combination of insect attack and low moisture conditions in the region, as ring-width patterns seen in the cores for 1999-2011 still relate well to trends in observed streamflow for that period.

Insect defoliators are a natural part of the Mongolian forested landscape. It is often assumed that insect defoliations of the past were constrained to small areas and would not affect multiple core sites synchronously, allowing identification of a defoliation signal at an individual site separate from the climate signals that would ideally be reflected in all sites in a region. Increasing temperatures and/or increasing droughts, will cause greater tree stress resulting in increased damage to larger areas of larch and to a lesser extent, pine trees from insects than previously known (Mattson and Haack, 1987; Ghent and Ohken, 2004). It is challenging to identify and remove the effects of these exogenous disturbances from the chronologies used in Mongolian streamflow reconstruction and more research and publication needs to be done regarding the extent and synchrony of insect defoliation events in the Khangai Mountain region and other areas of central Mongolia.

5.5.3 Comparisons to Other Reconstructions of Mongolian Streamflow

The simple drought and pluvial analysis given here corresponds to reconstructions in other basins ranging from the Kherlen River in the Eastern Steppe, to rivers in the central and northern portions of the country including the Yeruu and Selenge Rivers. Reconstructions of flow for the Kerlen River basin did not generally correspond to the patterns seen in the four basins studied here. A few extreme dry events in the late 1600's, the late 1700's and the early to mid-1900's did however occur synchronously with the northern Khanui and Khoid Tamir River Basins. Pluvial timings also matched up in the early to mid-1700's in the easternmost basins of the Khoid Tamir and Tuin Rivers and in the late 1700's in Khanui River basin. Wet periods in the early to mid-1900's were found for the Kherlen River and for all four Khangai Mountain region rivers (Pederson *et al*., 2001; Davi *et al*., 2013). Based on the great distances between the Kherlen River and rivers in the Khangai Mountain region, it is likely that any common events represent countrywide conditions of drought or wetness, especially the above average precipitation periods observed in the 1900's, some of which are captured in historical observational records.

Synchrony with flow in the central part of Mongolia is more common with some similarly timed dry and wet events in the Khoid Tamir and Tuin Rivers. In particular, these correspond to extreme dry and wet conditions in the Selenge River basin in the mid to late 1700's, the mid-1800's and the early 1900's (Davi *et al*., 2006). One or two extreme pluvial events also occurred across similar periods in the Khanui and Baidrag River basins and the Selenge River basin in the mid 1700's and early to mid 1900's. The Khanui River is a headwaters tributary of the Selenge River, with origins to the south and slightly east of the main stem of the Selenge River studied in Davi *et al*. (2006) and Pederson *et al*. (2013). The Hutag gaging station used in those reconstructions is located west of where the Khanui River joins the Selenge, and is almost directly north of Erdenemandal by almost 100 km. The portion of the Selenge previously studied receives most of its moisture in runoff from mountains to the south of Lake Khovsgol and mountains to the south and west of the river that border the drainage divide with the Khanui River basin (Davi *et al*., 2006).

Though geographically not that separate, the modeled results for the Selenge at Hutag are not completely similar to the results of modeling for the Khanui River, though there is broad agreement at selected periods, such as below mean flows in the late 19th century and above mean flows in the early part of the 20th century. This may be due to the choice of predictors, though the PCA used by Davi *et al*. (2006) did include the ZSM site, as did the Khanui River regression model used here. It is also possible that the more northerly and westerly sites used for the Selenge reconstruction are responding to different climatic signals than the overall more southern and eastern-located predictors used for the Khanui (and Baidrag) River models. However, when studying the Selenge and the Yeruu Rivers of western and eastern central Mongolia, respectively, Pederson *et al*. (2013) discovered that the western region (Selenge River) has a greater tendency toward bimodal, more persistent wet and dry conditions, similar to the reconstruction results of the Khoid Tamir and Tuin Rivers. The Yeruu River to the east has wet and dry extremes that exhibit less persistence through time (Pederson *et al*., 2013). While likely in a different hydroclimatic region (e.g., Leland *et al*., 2013) than the Yeruu River, the Khanui and Baidrag Rivers located much further to the west and south also exhibit more rapidly changing flow conditions through time.

5.5.4 10 More Years of Data

Some reconstructions of Mongolian streamflow use nested PCA regression to extend the reconstruction period further back in time or to step them forward closer to the present using differing predictors (i.e. Davi *et al*, 2010; Davi *et al*., 2013). Due to the availability of four chronologies that extended to the 2009-2011 period (Table 5.2), several models were constructed to examine relations between streamflow and tree-growth over the decade of 1998-2008. Statistically significant declines in streamflow have occurred in the Khanui and Khoid Tamir River basins over the 1976-2010 historical period of record (see Chapter 4). Increasingly intense and more prolonged droughts occurred in 1999-2002 and 2006-2009 than in the preceding few decades, and are attributed to increasing temperatures and reduced precipitation (Nandintsetseg and Shinoda, 2013). These droughts are reflected in the observational streamflow record for all basins under study.

An example model for the Khoid Tamir R. at Iktamir extends the reconstruction record from 1998 to 2008 and shows the precipitous decline in streamflow that occurred during that period in this basin (Figure 5.6). Similar steep declines occur in the observational records and flow reconstructions, (results not shown) for the Khanui River, using OGH and KHU predictors, and to a lesser extent in the Baidrag River, using OGH, JGB, and KLP predictors. The observational record for the Tuin River also shows the effects of the droughts in the 2000's but the drops in observed streamflow are not as steep as for the other three basins. No extended reconstruction could be built for this basin due to poor model fits. One caveat with the extended model for the Khoid Tamir River is that while there is good agreement between the low observed streamflow conditions in the last decade of the model and the reconstructed values (overall model adjR^2=0.67), the extended time series does not completely agree with the trends seen in the other models of flow for the basin (Figures 5.3 and 5.4). This is likely due to growth issues with the KLP predictor.

About 44% of the trees at the KLP site exhibit strip-bark conditions, where cambial dieback results in a positive ring-width response with increasing age resulting in decadal to centennial variations in growth that differ from reconstructions created from trees without this issue (Pederson *et al.*, 2014). The expression of these effects is dependent upon standardization method used in core processing, and it is unlikely these were completely removed with the spline detrending method applied here.

Despite complications with the KLP predictor, the extended model suggests very different hydroclimatic conditions occurred in the last decade of the reconstructions than in the past decades to centuries prior to that time. These drought conditions are especially harsh given the previous wetter conditions of the 20th century pluvial seen in the reconstructions. For comparison, non-overlapping 5-year mean values of streamflow in million cubic meters (MCM) were computed for each basin modeled with the extended predictors. For the Baidrag and Khoid Tamir Rivers, the reconstructed means from 1999-2003 and 2004-2008 (229.6 and 244.4 MCM for the Baidrag R. and 102.0 and 107.1 MCM for the Khoid Tamir R., respectively) would fall into the ranges of the top five non-overlapping means from the original reconstructions (Table 5.4). This implies that the low flows of the last modeled decade were at least as extreme as the most extreme dry periods seen throughout the 300+ year periods of reconstruction in those basins.

Interestingly, the modeled flows of the extended Khanui River reconstruction are wetter than those given as the top five driest (91.1 and 100.9 MCM, 1999-2003, and 2004-2008 respectively), but are only a third the volumes of the top wettest mean values. Much drier conditions have existed in past scenarios of flow for the Khanui River. For example, the top two dry periods in the basin occurred from 1682-1691, a decade of dry conditions exceeding the current drought in magnitude (Table 5.4). Conditions were also dry in the Tuin River basin at this time (Table 5.4 and expanded ranking results not shown).

Though no extended models demonstrate the 21st century droughts in the Tuin River basin, the observed record is not as dry as the most extreme events modeled for this basin (89.5 and 82.8 MCM, 1999-2003 and 2004-2008, respectively) suggesting that the drought is not as extreme in this area as in other basins to the north and west. Additionally, no significant decreasing trends in annual precipitation were found for this basin using gridded precipitation data over the period of 1963-2012 (Venable *et al*., 2015). An alternative interpretation of the last decade being wetter than extreme dry conditions from past periods in the Tuin River basin is that the poorer model fits may not capture the true range of hydroclimatic variation in this basin as well as those constructed for the other basins.

5.6 Conclusion

The differing flow conditions over the last 300+ years seen in the Khoid Tamir and Tuin River basins versus the Khanui and Baidrag River basins suggest different climatic influences on the predictor sites used for these models. In particular, the northern and westernmost sites (KLL, KLP, and ZSM) may have different influences than those located to the south in the eastern and central Khangai Mountains (i.e. OGH, NDB, ZTG). These results are complicated however by exogenous disturbances affecting the chronologies used from some sites. These issues at a minimum may include increases in ring width with age from strip-bark growth conditions and/or chronology end effects from detrending methods. Some expression of variability of growth with multiple suppression and release episodes from pervasive insect defoliation events is also possible in these chronologies. Additional analyses should be conducted taking both a more conservative and less conservative approach to detrending to determine possible effects on chronology development.

The two new sites provide extended records of tree-growth that give long-term context to the steep declines in streamflow seen across the region. Also, the high-elevation

Jargalant Bag (JGB) site may have strong relations to temperature. The exploration of these relations could corroborate the significant increases in temperature observed by local herders and in the meteorological station records of the region (Venable *et al*., 2012). Additional tree-ring site development to the south and west of the sites used in this study is recommended, particularly for the western portions of the Baidrag River basin to ascertain the possible expression of different regional hydroclimatic signals in that area.

An understanding of past hydroclimatic conditions across the country is needed to better manage future development of scarce hydrologic resources. As larger-scale agricultural and industrial (mining) development and growth continues in central Mongolia and spreads further into the Khangai Mountain region, these natural river systems will be increasingly used beyond the current domestic and livestock watering needs of the smaller population centers and nomadic pastoralists of the area. The recent decreases in streamflow in these four rivers while unusual in the context of the pluvial events of the mid-20th century, do not appear to be uncommon over the last 300+ years. More extreme low flows and longer periods of low flows exist in all of the new reconstructions, providing water managers with long-term context for the recent drought events observed across the country.

.

TABLE 5.1- Meteorological stations and streamflow gaging sites (1976-2012).

Name	Latitude (°), Longitude (°)	Elev. (m)	% Months Missing[1]	Mean Annual Flow (MCM)[2]
Erdenemandal Station	48.53 N, 101.38 E	1509	3.8	--
Tsetserleg Station	47.45 N, 101.47 E	1691	0.2	--
Tariat Station	48.16 N, 99.88 E	2040	2.3	--
Baidrag Station	47.20 N, 99.60 E	2199	50.5	--
Galuut Station	46.70 N, 100.13 E	2126	3.2	--
Bayankhongor Station	46.13 N, 100.68 E	1859	5.4	--
Khanui R. at Erdenemandal	48.61 N, 101.38 E	1487	24.5	155
Khoid Tamir R. at Ikhtamir	47.48 N, 100.89 E	1740	24.3	192
Baidrag R. at Bayanburd	46.67 N, 99.27 E	1880	1.6	317
Tuin R. at Bayankhongor	46.14 N, 100.72 E	1863	14.1	95

[1]Percentage of months missing is determined using the data quality protocols for all months in the 1976-2012 observational record.

[2]Mean annual flow calculated over 1976-2012 after data imputation, given in million cubic meters (cubic meters x10^6).

TABLE 5.2- Tree-ring sites used for reconstruction of basins. Codes are referenced on site map and in text. All sites are *Larix sibirica* with the exception of the Khorgo Lava Pine (KLP), which are *Pinus sibirica*.

Site	Latitude (°), Longitude (°)	Code	No. of Series	Chronology Span	EPS[a]	Yr. EPS >0.85[b]	Series Length (Yrs)[c]
Jargalant Bag	46.68 N, 100.93 E	JGB	21	1650-2011	0.88	1650	361
Khuush Uul	47.54 N, 101.08 E	KHU	26	1829-2011	0.96	1829	113
Khorgo Lava Larch	48.17 N, 99.87 E	KLL	55	1405-2000	0.97	1405	301
Khorgo Lava Pine	48.17 N, 99.87 E	KLP	35	1448-2011	0.98	1448	382
Mandal Hill Monastery	46.82 N, 100.12 E	MHM	16	1576-2002	0.91	1576	305
N. D. Bayankhongor	46.32 N, 101.32 E	NDB	18	1630-2001	0.91	1630	362
Orkhon Gol Hushree	46.79 N, 101.95 E	OGH	25	1582-2009	0.91	1582	284
Suleen Bagtraa	47.27 N, 100.03 E	SLB	32	1455-1999	0.94	1455	405
Zuun Salaa Mod	48.15 N, 100.28 E	ZSM	36	1564-2001	0.98	1564	340
Zurkh Togol	46.52 N, 100.95 E	ZTG	33	1639-2002	0.87	1639	303

[a]Overall between-tree EPS of series comprising the chronology
[b]Year when EPS>0.85 (from running 50-yr windows)
[c]Median series length after EPS truncation

TABLE 5.3- Model results. adjR^2=adjusted R^2, RMSE=Root Mean Squared Error (MCM[1]), MAE=Mean Absolute Error (MCM[1]), all statistics are from cross-validation results. The years reconstructed are based on the record lengths of the selected predictors.

Model	Predictors[2]	adjR^2	RMSE[1]	MAE[1]	Years of Reconst.
Khanui R. at Erdenemandal	OGH, ZSM, SLBp	0.61	60.9	43.0	1582-1998
Khoid Tamir R. at Ikhtamir	OGH, SLB, ZTG	0.69	60.7	47.7	1639-1998
Baidrag R. at Bayanburd	JGB, NDB, KLLp	0.67	56.4	45.9	1650-1998
Tuin R. at Bayankhongor	OGH, JGBp, SLBp	0.26	23.4	19.9	1650-1998

[1]Million cubic meters

[2]p denotes previous year lagged predictor

TABLE 5.4- Top five dry and wet 5-year intervals for each basin reconstruction. Mean flow conditions for the intervals in million cubic meters (MCM) follow the dates.

	Rank	Khanui River	Khoid Tamir River	Baidrag River	Tuin River
Dry	1	1687-1691 (26.1)	1757-1761 (78.0)	1897-1901 (201.6)	1757-1761 (53.0)
	2	1682-1686 (30.0)	1747-1751 (90.5)	1912-1916 (203.4)	1747-1751 (54.0)
	3	1807-1811 (39.2)	1752-1756 (104.2)	1797-1801 (207.6)	1752-1756 (58.0)
	4	1927-1931 (47.3)	1772-1776 (113.0)	1957-1961 (236.6)	1882-1886 (60.2)
	5	1667-1671 (67.5)	1872-1876 (132.5)	1717-1721 (244.4)	1692-1696 (61.7)
Wet	1	1672-1676 (311.4)	1712-1716 (404.0)	1882-1886 (731.3)	1712-1716 (132.1)
	2	1762-1766 (306.8)	1792-1796 (386.4)	1922-1926 (708.2)	1717-1721 (120.0)
	3	1917-1921 (278.8)	1702-1706 (341.9)	1927-1931 (698.4)	1707-1711 (115.8)
	4	1732-1736 (270.5)	1727-1731 (335.8)	1942-1946 (696.1)	1792-1796 (114.8)
	5	1727-1731 (266.8)	1707-1711 (329.4)	1652-1656 (558.9)	1972-1976 (112.7)

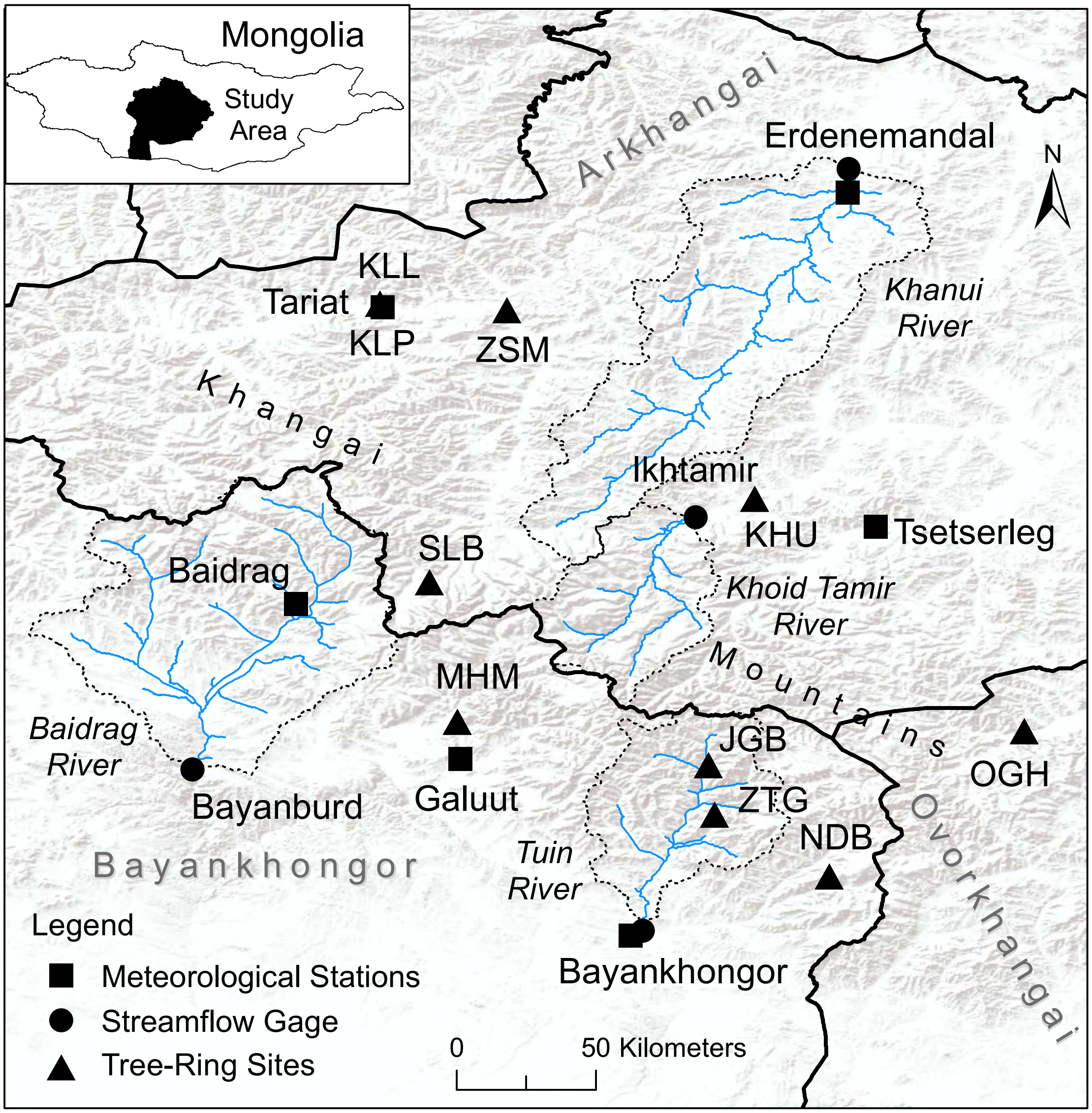


FIGURE 5.1- Map of the study region in Mongolia, with meteorological stations as black squares, streamflow gages as black circles, and tree-ring sites as black triangles. Study aimags and river basins of interest are outlined.

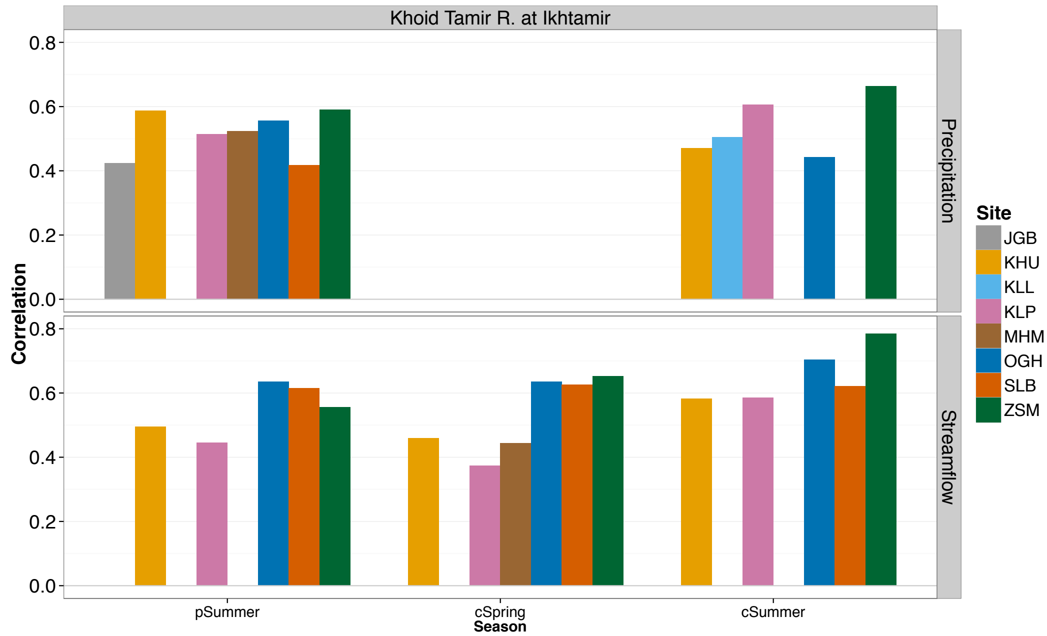


FIGURE 5.2a- Significant correlations (p<0.05) between tree-ring sites and seasonal precipitation and streamflow for the Khoid Tamir R. at Ikhtamir for previous Summer (pSummer: June, July, August), current Spring (cSpring: March, April, May), and current Summer (cSummer: June, July, August).

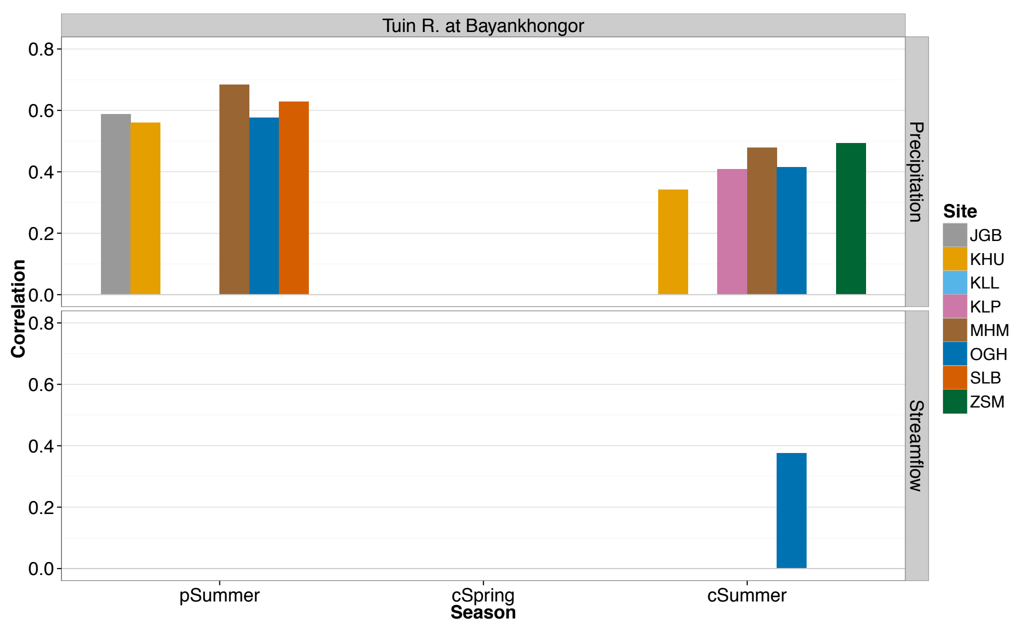


FIGURE 5.2b- Significant correlations between tree-ring sites and seasonal precipitation and streamflow for the Tuin R. at Bayankhongor for previous Summer (pSummer: June, July, August), current Spring (cSpring: March, April, May), and current Summer (cSummer: June, July, August) .

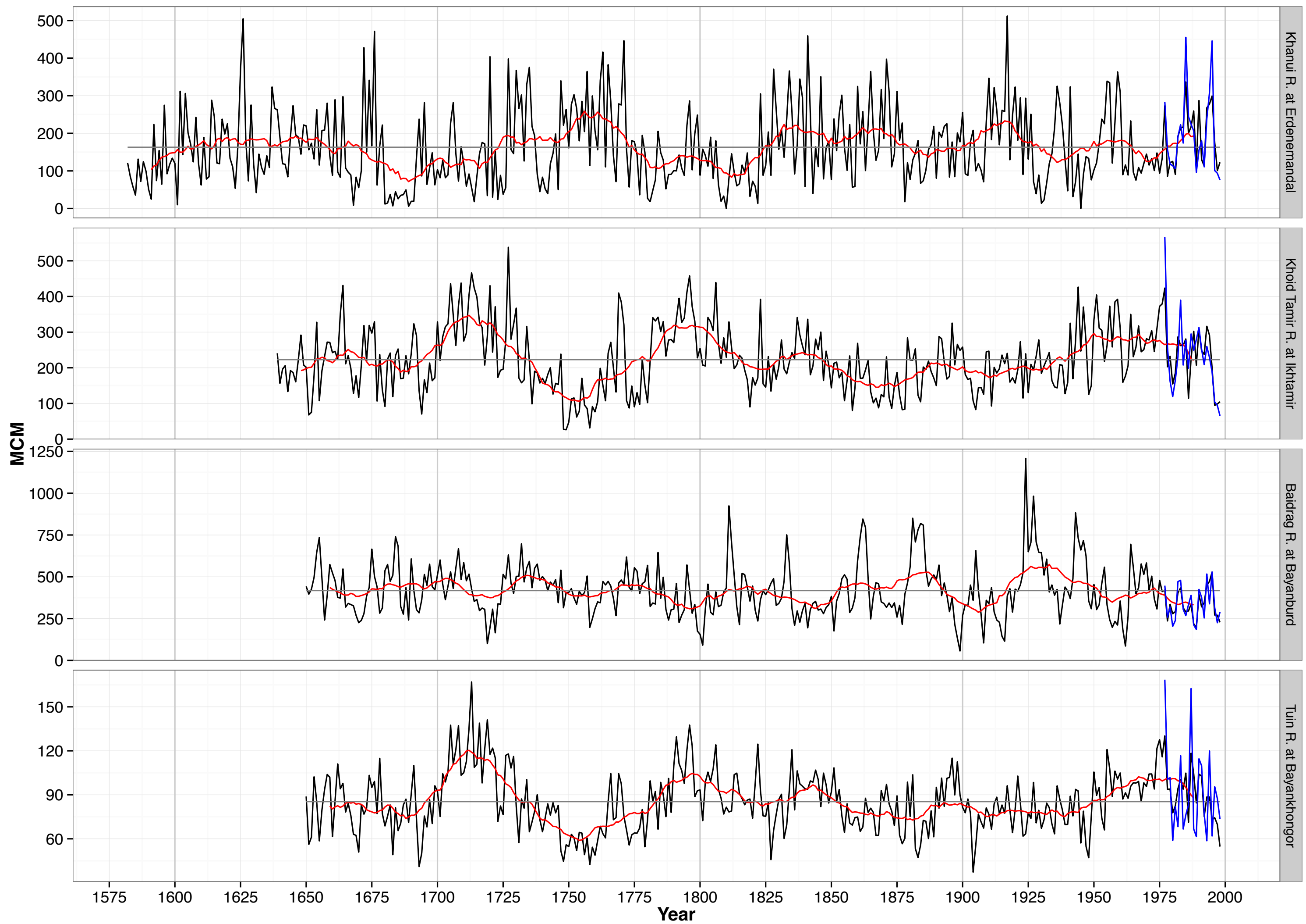


FIGURE 5.3- Model results for four basins of interest. Black thin lines are reconstructed streamflow in millions of cubic meters (MCM) of annual flow, note the differing y-axes. Differing y-axis scales were used to facilitate comparison of flow variability between reconstructions. Horizontal grey lines through each time series are the reconstruction mean, red lines are a 20-year windowed mean to aid interpretation, and blue lines are observed streamflow for the 1977-1998 period.

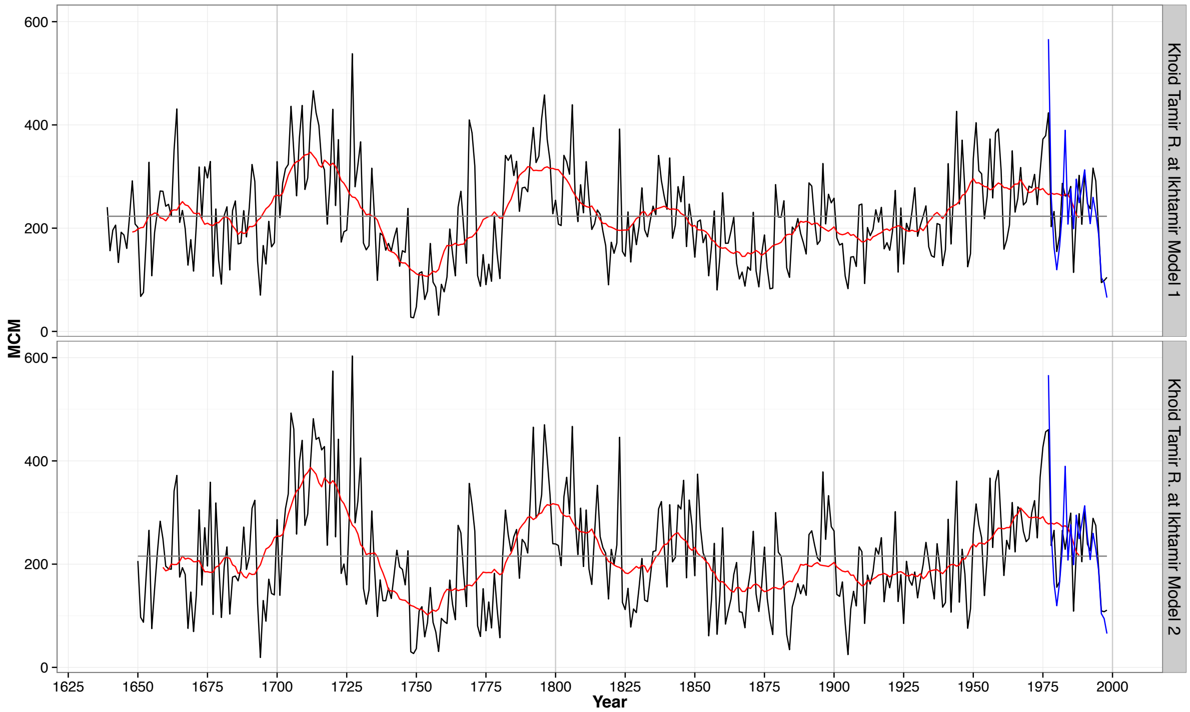


FIGURE 5.4- Alternate models for the Khoid Tamir R. at Ikhtamir using sites OGH, SLB, and ZTG (Model 1) or JGB (Model 2).

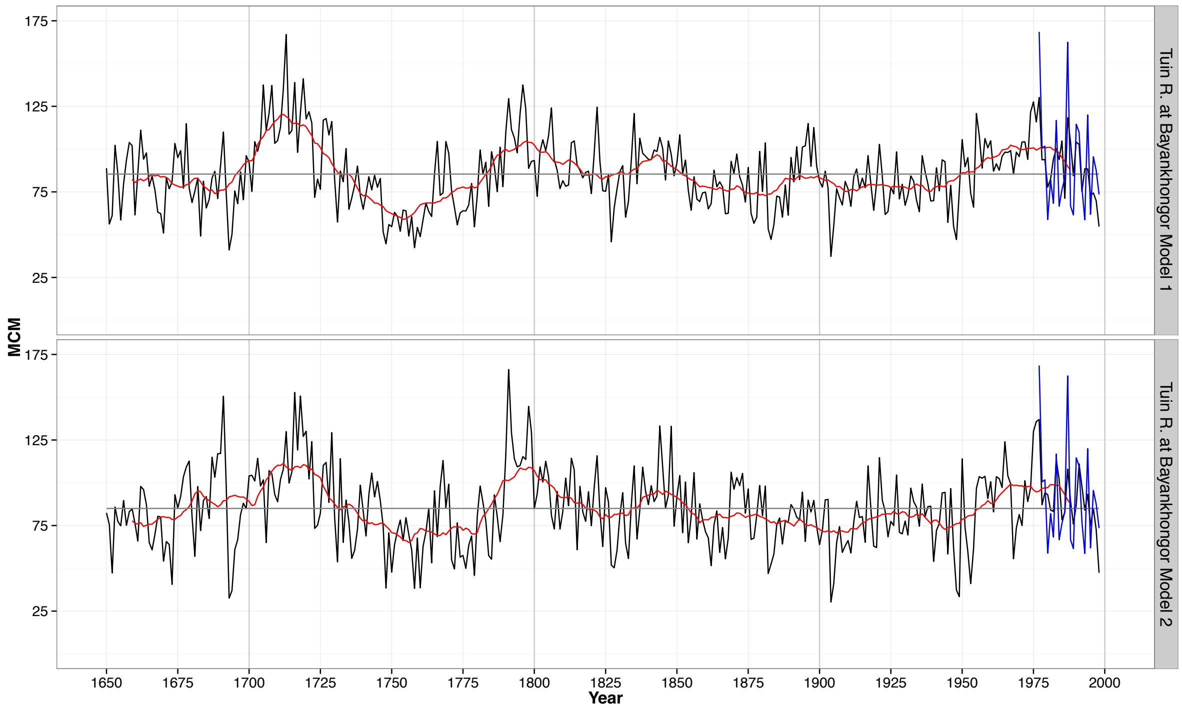


FIGURE 5.5- Alternate models for the Tuin R. at Bayankhongor using sites JGBp and SLBp, and OGH (Model1) or MHM (Model 2), where p indicates previous year.

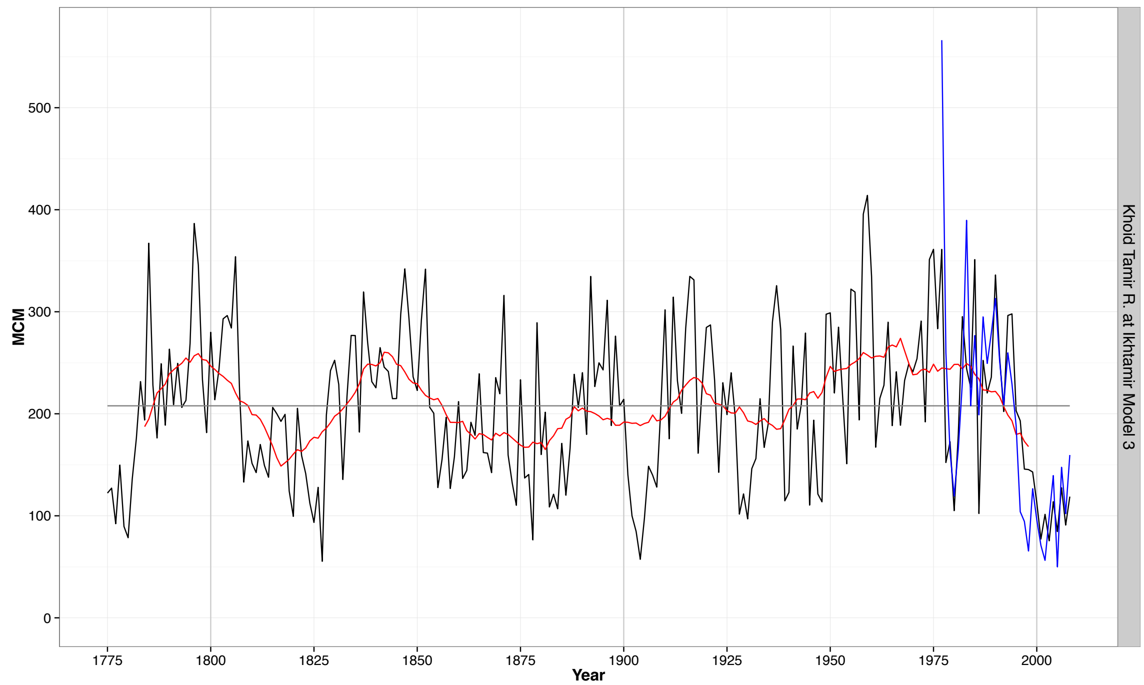


FIGURE 5.6- Extended model for the Khoid Tamir R. at Ikhtamir using sites OGH and KLPp (previous year), truncated to 1775-2008 to show detail. Note the severe drop in modeled and observed streamflow in the last decade of the plot (1998-2008).

REFERENCES

Arlot S, and Celisse A, 2010. A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40-79, doi:1214/09-SS054.

Batima P, Natsagdorj L, Gombluudev P, Erdenetsetseg B, 2005. Observed climate change in Mongolia. *Assessments and Adaptations to Climate Change* (AIACC), *www.aiaccproject.org*. Accessed 2011-11-21.
Bunn AG, 2008. A dendrochronology program library in R (dplR). *Dendrochronologia*, 26, 115–124, doi:10.1016/j.dendro.2008.01.002.

Bunn AG, 2010. Statistical and visual crossdating in R using the dplR library. *Dendrochronologia*, 28, 251–258, doi:10.1016/j.dendro.2009.12.001.

Chai T, and Draxler RR, 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7, 1247-1250, doi: 10.5194/gmd-7-1247-2014.

Cook ER, 1985. *A time series analysis approach to tree ring standardization*. Unpublished PhD Dissertation, University of Arizona Press, Tucson, AZ, USA, *http://arizona.openrepository.com/arizona/handle/10150/188110*. Accessed 2013-07-13.

Cook ER, and Peters K, 1997. Calculating unbiased tree-ring indices for the study of climatic and environmental change. *The Holocene,* 7, 361-370, doi: 10.1177/095968369700700314.
Cook ER, Briffa K, Shiyatov S, Mazepa A, Jones, PD, 1990. Chapter 3: Data analysis. In: Cook, ER, Kairiukstis, LA [Eds]. *Methods of Dendrochronology: Applications in the Environmental Sciences*. Kluwer Academic Publishers, Dordrecht, pp. 97–162.

Cook ER, Briffa KR, Meko DM, Graybill DA, and Funkhouser G, 1995. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. *The Holocene*, 5, 229–237.

Cook ER, Anchukaitis KJ, Buckley BM, D'Arrigo RD, Jacoby GC, and Wright WE, 2010. Asian Monsoon failure and megadrought during the last millennium. *Science*, 328, 486–489, doi:10.1126/science.1185188.

D'Arrigo R, Jacoby G, Pederson N, Frank D, Buckley B, Nachin B, Mijiddorj R, and Dugarjav C, 2000. Mongolian tree-rings, temperature sensitivity and reconstructions of Northern Hemisphere temperature. *The Holocene*, 10, 669–672, doi:10.1191/09596830094926.

Davi NK, Jacoby GC, Curtis, AE, and Baatarbileg N, 2006. Extension of drought records for Central Asia using tree rings: West-Central Mongolia. *Journal of Climate*, 19, 288–299, doi:10.1175/JCLI3621.1.

Davi N, Jacoby G, Fang K, Li J, D'Arrigo R, Baatarbileg N, and Robinson D, 2010. Reconstructing drought variability for Mongolia based on a large-scale tree ring network: 1520–1993. *Journal of Geophysical Research: Atmospheres*, 115, D22103, doi:10.1029/2010JD013907.

Davi NK, Pederson N, Leland C, Nachin B, Suran B, and Jacoby GC, 2013. Is eastern Mongolia drying? A long-term perspective of a multidecadal trend. *Water Resources Research*, 49, 151–158, doi:10.1029/2012WR011834.

De Grandpré L, Tardif JC, Hessl A, Pederson N, Conciatori F, Green TR, Oyunsanaa B, and Baatarbileg N, 2011. Seasonal shift in the climate responses of Pinus sibirica, Pinus sylvestris, and Larix sibirica trees from semi-arid, north-central Mongolia. *Canadian Journal of Forest Resources*, 41, 1242–1255, doi:10.1139/x11-051.

Fassnacht SR, Sukh T, Fernandez-Gimenez ME, B Batbuyan, Venable NBH, Laituri M, and Adyabadam G, 2011. Local understanding of hydro-climatic changes in Mongolia. Proceedings of *H02 Symposium Cold Region Hydrology in a Changing Climate* held during IUGG2011 in Melbourne, Australia, July 2011, *IAHS Publication* 346, 120-129.

Fritts HC, 2001. *Tree Rings and Climate*. Blackburn Press: Caldwell, 567 p., ISBN: 1-930665-39-3.

Fitzsimmons K, and Venable NBH, 2015. Correlation of tree-rings and precipitation stations in the Khangai Mountain region of Mongolia*. Department of Ecosystem Science and Sustainability Spring Research Fair: Skills for Undergraduate Participation in Research Program*, Colorado State University, April 23rd, 2015.

Frank D, Esper J, and Cook E, 2006. On variance adjustments in tree-ring chronology development. *TRACE - Tree Rings in Archaeology, Climatology and Ecology, Vol. 4: Proceedings of the DENDROSYMPOSIUM 2005*, April 21st – 23rd 2005, Fribourg, Switzerland. Heinrich I, Gärtner H, Monbaron M, Schleser G, [Eds]. *Schriften des Forschungszentrums Jülich, Reihe Umwelt* Vol. 61, p. 56 – 66, *http://treering.de/sites/default/files/TRACE_pdf/Volume_3/Frank_et_al_TraceVol_4.pdf*. Accessed 2015-10-06.

Ghent JH, and Onken AH, 2004. Trip report on assistance to Mongolian Ministry for Nature and Environment for the control of forest defoliators. USDA Forest Service, Forest Health Protection Unit, *http://www.fao.org/forestry/892808215bc1f05de14f2f8d1556ab3ddee74.pdf*. Accessed 2015-01-20.

Grissino-Mayer HD, 2001. Evaluating cross dating accuracy: A manual tutorial for the computer program COFECHA. *Tree-Ring Research* 57, 205–221.
Harris I, Jones PD, Osborn TJ, and Lister DH, 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, 34, 623–642, doi:10.1002/joc.3711.

Helsel DR, and Hirsch RM, 2002. Statistical methods in water resources. *Techniques of Water-Resources Investigations of the United States Geological Survey, Book 4, Hydrologic Analysis and Interpretation, Chapter A3. http://water.usgs.gov/pubs/twri/twri4a3/*. Accessed 2014-11-02.

Hessl AE, Ariya U, Brown P, Byambasuren O, Green T, Jacoby G, Sutherland EK, Nachin B, Maxwell RS, Pederson N, De Grandepré L, Saladyga T, Tardif JC, 2012. Reconstructing fire history in central Mongolia from tree-rings. *International Journal of Wildland Fire*, 21, 86–92, doi:10.1071/WF10108.

Holmes RL, 1983. Computer-assisted quality control in tree-ring dating and measurement. *Tree-Ring Bulletin*, 43, 69–78.

Jacoby G, Arrigo RD, Pederson N, Buckley B, Dugarjav C, and Mijiddorj R, 1999. Temperature and precipitation in Mongolia Based on dendroclimatic investigations. *IAWA Journal*, 20, 339–354, doi:10.1163/22941932-90000696.

John R, Chen J, Ou-Yang ZT, Xiao J, Becker R, Samanta A, Ganguly S, Yuan W, and Batkhishig O, 2013. Vegetation response to extreme climate events on the Mongolian Plateau from 2000 to 2010. *Environmental Research Letters*, 8, 035033, doi:10.1088/1748-9326/8/3/035033.

Khishigjargal M, Dulamsuren C, Leuschner HH, Leuschner C, and Hauck M, 2014. Climate effects on inter- and intra-annual larch stemwood anomalies in the Mongolian forest-steppe. *Acta Oecologia-International Journal of Ecology*, 55, 113–121, doi:10.1016/j.actao.2013.12.003.

Kutner MH, Nachtscheim C, Neter J, and Li W, 2005. *Applied Linear Statistical Models*, 5th Edition, McGraw-Hill Irwin, ISBN:0072386886.

Leland C, Pederson N, Hessl A, Nachin B, Davi N, D'Arrigo R, and Jacoby G, 2013. A hydroclimatic regionalization of central Mongolia as inferred from tree rings. *Dendrochronologia*, 31, 205–215, doi:10.1016/j.dendro.2012.11.003.

Mattson WJ, Haack RA, 1987. The role of drought in outbreaks of plant-eating insects. *BioScience*, 37, 110–118.
Ministry of Nature, Environment, and Tourism, 2010. Mongolia Second National Communication, Dagvadorj D [Ed]. Ulaanbaatar, Mongolia, 160 pp.

Morris TP, White IR, Royston P, 2014. Tuning multiple imputation by predictive mean matching and local residual draws. *BMC Medical Research Methodology*, 14(75), 1-13, doi:10.1186/1471-2288-14-75.

Nandintsetseg B, and Shinoda M, 2013. Assessment of drought frequency, duration, and severity and its impact on pasture production in Mongolia. *Natural Hazards*, 66, 995–1008, doi:10.1007/s11069-012-0527-4.

Osborn TJ, Briffa KR, Jones PD, 1997. Adjusting variance for sample-size in tree-ring chronologies and other regional mean time series. *Dendrochronologia*, 15, 89-99.

Pederson N, Jacoby GC, D'Arrigo RD, Cook ER, Buckley BM, Dugarjav C, and Mijiddorj R, 2001. Hydrometeorological reconstructions for northeastern Mongolia derived from tree rings: 1651-1995. *Journal of Climate*, 14(5), 872-881, doi: 10.1175/1520-0442(2001)014.

Pederson N, Leland C, Nachin B, Hessl AE, Bell AR, Martin-Benito D, Davi NK, 2013. Three centuries of shifting hydroclimatic regimes across the Mongolian Breadbasket. *Agricultural and Forest Meteorology*, 178, 10–20.

Pederson N, Hessl AE, Baatarbileg N, Anchukaitis KJ, and Di Cosmo N, 2014. Pluvials, droughts, the Mongol Empire, and modern Mongolia. *Proceedings of the National Academy of Sciences*, 111, 4375–4379, doi:10.1073/pnas.1318677111.

Priess JA, Schweitzer C, Wimmer F, Batkhishig O, Mimler M, 2011. The consequences of land-use change and water demands in Central Mongolia. *Land Use Policy* 28, 4–10.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www.R-project.org/.*

Rubin DB, 1987. *Multiple imputation for nonresponse in surveys*. Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, New York, 258 p. ISBN 0-471-08705-X.

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, and Rudolf B, 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115, 15–40, doi:10.1007/s00704-013-0860-x.

Stokes MA, and Smiley TL, 1996: *An Introduction to Tree-Ring Dating*. University of Arizona Press,Tucson, AZ, 73 pp.

Swetnam TW, and Brown PM, 1992. Oldest known conifers in the southwestern United States: temporal and spatial patterns of maximum age. *Old Growth Forests in the Southwest and Rocky Mountain Regions. USDA Forest Service General Technical Report RM-213*, 24–38.

Tao S, Fang J, Zhao X, Zhao S, Shen H, Hu H, Tang Z, Wang Z, Guo Q, 2015. Rapid loss of lakes on the Mongolian Plateau. *Proceedings of the National Academy of Sciences*, 112, 2281–2286, doi:10.1073/pnas.1411748112.

van Buuren S, and Groothuis-Oudshoorn K, 2011. mice: Multivariate imputation by chained equations in R. *Journal of statistical software*, 45, *http://doc.utwente.nl/78938/*. Accessed 2015-07-24.

van Buuren S, 2012 *Flexible imputation of missing data*. Chapman & Hall/CRC Press, Boca Raton, FL, 316 p, ISBN: 9781439868249.

Venable NBH, Fassnacht SR, Adyabadam G, S Tumenjargal, Fernandez-Gimenez M, and B Batbuyan, 2012. Does the length of station record influence the warming trend that is perceived by Mongolian herders near the Khangai Mountains? *Pirineos*, 167, 69–86, doi:10.3989/Pirineos.2012.167004.

Venable NBH, and Fassnacht SR, 2015. Patterns of hydroclimatic variability in proxy: Tree-rings from the Khangai Mountains of Mongolia. Poster presentation, *2015 American Geophysical Union Fall Meeting*, San Francisco, CA, December 14-18, 2015.

Venable NBH, Fassnacht SR, and Hendricks AS, 2015. Spatial changes in climate across Mongolia. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig , Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 73-77.

Wigley T, Briffa K, and Jones P, 1984. On the average value of correlated time-series, with applications in dendroclimatology and hydrometeorology. *Journal of Climate and Applied Meteorology*, 23, 201–213, doi:10.1175/1520-0450(1984)023<0201:OTAVOC>2.0.CO;2.

Willmott CJ, and Matsuura K, 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79-82.

Yu F, Price KP, Ellis J, and Shi P, 2003. Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sensing of Environment*, 87, 42–54, doi:10.1016/S0034-4257(03)00144-5.

CHAPTER 6- HYDROMETEOROLOGICAL DATA ANALYSES AS BOUNDARY OBJECTS IN INTERDISCIPLINARY CLIMATE CHANGE RESEARCH

6.1 Introduction

Study of the Earth's changing climate is a complex and challenging endeavor employing a diverse and ever-broadening assemblage of disciplines to attain knowledge. Humans have impacted most Earth systems, with significant and long lasting consequences (Zalsiewicz *et al*., 2011). Innovative and integrative solutions are needed to manage the challenges faced by the rapidly changing socio-ecological systems of the Anthropocene (Rockstrom *et al*., 2009). As such, scientists are routinely performing collaborative work that crosses discipline boundaries to address larger and more complex problems, or are at minimum, employing the language of multi-, inter-, and transdisciplinarity in climate change related research to further their individual research goals (Rhoten and Parker, 2004; Sundberg, 2007).

In this regard, the main objective of my dissertation is to quantify the changing climate of Mongolia through analysis of key hydrometeorological variables over space and through time. The assessments of trends in the data conducted in Chapters 2-4, plus the paleo proxy analyses conducted in Chapter 5, address interdisciplinary research questions using multidisciplinary approaches. The primary purpose of this chapter is to examine how the data and analyses previously presented are used in a cross-boundary research sense by exploring specific connections within each chapter, and then concluding with a brief discussion of the contributions of this work to selected goals of the Mongolian Rangelands and Resilience (MOR2) project lead by researchers at Colorado State University in collaboration with national and international colleagues.

6.2 Data and Analyses as Boundary Objects

Star and Griesemer (1989) developed the concept of boundary objects as a way to describe how objects [or results] of scientific inquiry can inhabit different intersecting social worlds (e.g., the social spheres of influence and interaction of professional scientists in their own community versus those of another distinct group such as the administrators at a university that interact with the scientists), or disciplinary realms (e.g., ecologists, or hydrologists) while adapting interpretively across these boundaries or social/disciplinary divides to meet the needs of each of these groups. These objects maintain a common identity, but individually, or within a particular group they may have a strong structure or tailored use, a specific identity or definition. In common use by different groups, they may have a weaker interpretation or structural function allowing interpretive flexibility (Star and Griesemer, 1989; Star, 2010). These objects act as temporary bridges to joint and future endeavors, but as objects cannot displace actual collaboration or communications by actors in differing groups (Star, 2010; Akkerman and Bakker, 2011).

The conceptualization of climate change data and/or scientific research results as boundary objects has precedent (e.g., Lynch *et al*., 2008; Star, 2010; Whetton *et al*., 2012; van Pelt *et al*., 2015; Blades *et al*., 2016). Here, hydrometeorological analyses become boundary objects when interpreted through *efforts of translation* [finding shared meaning and/or reconciliation of a diversity of definitions or understandings] between different groups or worlds (Akkerman and Bakker, 2011). As more than just static research products, they are capable of stimulating communication between disciplinary communities through the creation of compatible conceptual frameworks (Lynch *et al*., 2008). Ideally, they may be presented as actionable items that become catalysts for new research ideas through collaboration (Kimble *et al*., 2010). The idea of trend analyses, a subset of broader climate change research results, as boundary objects informing further interdisciplinary work, is a means of scaling the concept that has limited coverage in the climate literature. For a conceptually related example see Sundberg (2007), where the parameterization of climate models is discussed as a boundary object that both climate modelers and experimentalists connect to across occupational and epistemic boundaries.

6.3 Multidisciplinary Approaches Facilitate Objectification of Research Results

The boundary object model can initially be established at the data level, as climate variables form the foundation of the hydroclimatic change investigations performed in previous chapters. To illustrate this idea, an example from Chapter 4 is used as that chapter focuses on effects of data quality on analysis results.

6.3.1 Data Quality Issues

Within the meteorological/climatological research community, expertise has been developed to process and analyze climate data using specialized procedures (e.g., Willmott and Robeson, 1995; Easterling *et al*., 1996; Mitchell and Jones, 2005; Beaulieu *et al*., 2007). Outside of that community however, data are frequently used "as-is" after basic quality control pre-processing, or derived products are sought that minimize discrepancies in data. These are generally in the form of continuous gridded climate surfaces that fill and smooth spatiotemporal gaps between data points (e.g., Becker *et al*., 2013; Harris *et al*., 2014). Interpolated and time gap-filled products may or may not adequately represent the underlying point-based records, which could significantly affect research results depending on the application (Willmott *et al*., 1985; Ensor and Robeson, 2008; Venable *et al*., 2014a).

In the case of station-based Mongolian climate data there is an underlying problem of the uncertainty associated with the *delegated work* (Star, 2010) of the data collection and archiving process that cannot be controlled by end-users. One way this manifests is in the question of if missing [specifically precipitation] data represents a zero measurement or was actually missing, as in not recorded. The designation of data points as missing rather than zero values has strong effects on the results of trend analyses, likely more so than the other issues addressed in the chapter such as the presence of a trend in the data prior to statistical analysis, or due to the presence of autocorrelation or persistence between observations.

An anonymous external review of a previous version of Chapter 4 reinforced the importance of this issue. Specifically, a scientist in the climatological community, which is known for its rigorous treatment of meteorological data collection and processing methods, raised concerns that the methods employed in the chapter may not be optimal due to data quality. The Mann-Kendall (Mann, 1945; Kendall and Gibbons, 1990) and Thiel-Sen (Thiel, 1950; Sen, 1968) approaches used in the chapter however, are considered robust non-parametric methods for analyzing datasets with missing values and are widely accepted and employed by the hydrological community (e.g., Gilbert, 1987; Helsel and Hirsch, 2002). The reviewer's comments transformed the data from a scientific product, a *thing*, to an object capable of initiating a *reflective* process highlighting differences between treatments of the data by what are perceived to be closely related disciplinary communities (Akkerman and Bakker, 2011). This example serves to reinforce disciplinary disparities, while also allowing the boundary between those disciplines to work as a means to "*look at oneself through the eyes of other worlds*" (Akkerman and Bakker, 2011, p.145) when considering future versions of the work. Ironically, the reviewer's comments further reinforced one of the original intents of the chapter, specifically, to inform groups outside of the climatological community of the caveats of using data with many missing values (Venable and Fassnacht, 2015a).

6.3.2 Crossing Disciplines with Hydrometeorological Analyses

Connections can also be drawn between the scientific results from one discipline and how those results relate and integrate into analyses from other disciplines. In Chapters 3 and 5 multidisciplinary approaches are used to answer different research questions that are rooted in similar socio-ecological contexts of nomadic pastoralist use of natural resources, and impacts to those resources under a changing climate. Boundary objects discussed here include socio-ecological surveys of nomadic pastoralists, hydroclimatic trend analyses, and reconstructions of long-term streamflow variability.

6.3.2.1 Observations of Change

The research questions explored in Chapter 3 (i.e., Venable *et al*., 2012b) include: *"Are the warming trends perceived [observed] by nomadic pastoralists of the Khangai Mountain region found in the results of trend analyses of observational climate records, and if so, what period of record is most statistically significant?"* This work has a clear multidisciplinary focus, connecting the socio-ecological world and assessments of indigenous knowledge, and the hydroclimatological realm concerned with analyzing and understanding spatiotemporal variability in station and gage-based hydrometeorological data.

The nomadic pastoralists of Mongolia live and work in a land of weather and climate extremes and must perceive and respond to ecological change to succeed economically and culturally (Fernandez-Gimenez, 2000; Brugger *et al*., 2014). Historically, anthropologists and ethnographers primarily conducted studies of culture, but broader disciplinary groups like ecologists and even hydrologists are now conducting these surveys to understand herder lifestyles and their adaptation strategies to a variety of internal and external influences. Interdisciplinary scientists from Colorado State University and other institutions have been studying facets of Mongolian traditional and ecological knowledge for decades (e.g., Fernandez-Gimenez, 1993; Lkhagvadorj et al., 2013; Brugger *et al*., 2014). Recent surveys incorporate explicit questions of herder's observations of the hydroclimate (e.g., Marin, 2010; Fassnacht *et al*., 2011). Some surveys, such as those conducted by Sukh (2012), are composed entirely of questions about hydroclimatic change.

Herder surveys can be seen as two different types of boundary objects. First, as objects sharing *coincident* boundaries (Star and Griesmer, 1989). These objects, while designed or implemented by different groups (i.e., surveys designed by ecologists or surveys designed by hydroclimatologists) are in this case, collecting information coincidently about the same place, time, or people. Each captures social, ecological, and climatic knowledge differently depending upon who designed the survey instrument and who is asking the questions. Different goals [answers to research questions] are resolved through these objects, and diverse researchers may understand and use the internal contents of each object differently (Star and Griesemer, 1989; McGreavy *et al*., 2013).

Hydrologists/hydroclimatologists find value in existing surveys of ecological change. Based on their expertise, they may relate changes in ecology to the influence of changing hydrometeorological processes, but are not necessarily concerned with the fine details of ecological change as defined and/or understood by ecologists. As part of an iterative organic process that occurs between groups (Star, 2010), objects can provide insight to one group from another, resulting in work that bridges disciplines. This is the case of the temperature analyses of Chapter 3. In this example, one object [socio-ecological surveys] is used to inform the creation of another [hydrometeorological analyses]. Sociologists and ecologists then use the latter object synergistically in further research recognizing that an understanding of trends in hydroclimatic variables through time can contribute to the process of identifying impacts of climate change on herders (e.g., Fernandez-Gimenez *et al*., 2015).

A second way of using the surveys as boundary objects is to consider them *repositories* (Star and Griesemer, 1989). In this case, portions of objects originally designed and implemented specifically for one group are extracted and used directly, or refined by a different group for their own purposes (Star and Griesemer, 1989; Star, 2010). This is functionally what the original MOR2 herder surveys became when used cross-disciplinarily. Chapter 3 employed this model, extracting the results of herder surveys pertaining to temperature change and making comparisons to the historical climate record.

Ideally, survey instruments used in the manner of Chapter 3 should be constructed to capture key relations between observations made by herders and measured climate observations to facilitate comparisons to hydrometeorological analyses. This integrative and collaborative approach was implemented with later herder surveys conducted by the MOR2 team and used in subsequent research (e.g., Brugger *et al*., 2014; Fernandez-Gimenez *et al*., 2015). This synthesis was a challenging task both from the standpoint of a need for a certain level of consensus between groups on survey design and/or implementation, and as the personal observations collected in the surveys provide a level of spatiotemporal detail that may not be compatible with coarser-scale climate products or may provide information that is difficult to support or refute as it is not measured using standard instrumented methods (e.g., Marin, 2010; Fassnacht *et al*., 2011; Fernandez-Gimenez *et al*., 2015). The uncertainty inherent in the pairing of survey and hydroclimatic data is a key point of the chapter. The idea that herder's observations may extend over variable time periods was represented by the testing of various lengths of record from the full period of 50 years for some stations, down to 15-year periods for all stations, and by shifting the start and end years for each length of record tested.

6.3.2.2 Dendrohydrology

Broadly defined, dendroclimatology is concerned with the study of past climatic conditions using tree-rings (Fritts, 2001). Dendrohydrology is a subset of that discipline, focused on understanding long-term hydrologic phenomena of the past (Loaiciga *et al*., 1993). As an interdisciplinary science, it incorporates multi-disciplinary approaches using statistical relations between analyses of hydrometeorological variables from historical periods and data obtained from the specialized processing and analysis of tree growth rings to reconstruct past scenarios of hydrological variability over the length of the selected paleo proxy record (e.g., Fritts, 2001). Inherent in this approach is the crossing of disciplinary boundaries, not only driven by the desires of the hydrological and dendroclimatological scientific communities to pursue *basic research* (Parker and Crona, 2012), but also extending into the public and practitioner spheres to help water managers understand how or why this work can assist them in meeting long-term supply needs for their constituencies (Meko and Woodhouse, 2007).

The practical application of this interdisciplinary field has been shaped by the intensely water-managed landscapes of the American West. Early researchers like Stockton (1971) made explicit connections between moisture-limited tree growth and variations in streamflow through time. The results of these and later investigations became boundary objects, or bridges for knowledge about the long-term variability and availability of water supplies for human use crossing from one disciplinary community (dendrohydrologists) to another (water resource managers). In essentially unmanaged river systems like those found across most of Mongolia, knowledge of long-term hydrologic regimes can help government officials, pastoralists, and farmers understand patterns of hydroclimatic variability over the last few centuries to place recent extreme climate conditions into a longer term context (e.g., Davi *et al*., 2013; Pederson *et al*., 2013). This is particularly important in those basins facing shifts in land use through intensification of agriculture or increases in water-intensive industries like mining (e.g., Regdel *et al*., 2012; Pederson *et al*., 2013). The work performed in Chapter 5 fits into this framework, filling in gaps in knowledge about long-term streamflow variability in the Khangai Mountain region.

While the efforts of other researchers in the dendroclimatological community studying Mongolia (i.e., Leland, 2011) certainly inspired my ideas for Chapter 5, in large part the investigations in that chapter were a result of previous explorations of historical hydroclimatic data performed in preparatory work for Chapters 3-5 and in the analysis efforts of the chapters themselves (e.g., Venable *et al*., 2012a; Venable *et al*., 2012b; Venable *et al*., 2012c, Venable and Fassnacht, 2013; Venable *et al*., 2013; Venable *et al*., 2014b; Venable and Fassnacht 2015b; Wolf and Venable, 2015). As an interdisciplinary scientist, my own research results morphed into a type of *ideal* boundary object that was internally useful (Star and Greisemer, 1989; Star, 2010). The details of my hydroclimatic analysis results became less critical when thinking about dendrohydrological research questions like: "*How does the 300+ year paleo record for the Khangai Mountain region compare to the last 50 to 35 years of station and gage-based data I had been studying previously?*" yet provided context to the new questions nonetheless. I made a transition from tailored and concrete use of the hydroclimatologic data in the analyses of my previous work, to relatively vague use of the hydroclimatic results as I crossed the disciplinary threshold into the tree-ring community. I then returned to a sharper hydroclimatological focus when connecting the results of my long-term streamflow reconstructions back to the historical trends. This iterative behavior is characteristic of the working use of boundary objects (Star, 2010). It is important to note that certain points of the process are not necessarily interdisciplinary or consensual when working on either side of a disciplinary line (Star, 2010). It can be argued that since my works must adhere to the standards of the individual disciplines pursued they do not possess true interdisciplinary characteristics, but these ways of working are often seen as dynamic and productive tensions needed to create new knowledge (Klein, 1996). Regardless of structural terminology, these chapters generally recognize a *shared problem space* (Akkerman and Bakker, 2011), where the approaches of the differing disciplines can co-mingle to emphasize the combined disciplinary context of the work (Parker and Crona, 2012; Wyborn, 2015).

6.4 Research Results as Objects for Further Interdisciplinary Work

The idea that exact interpretation (and full comprehension) of research results is needed for cross-boundary conversation is not supported by the boundary object construct (Star and Greisemer, 1989). It is common for a group [discipline] to maintain a fuzzy understanding of the explicit work of another group and yet be able to use the object for building their own work and contributing to a larger interdisciplinary undertaking (Star, 2010). As such, maps are a prime means of exemplifying the abstractions that can occur when utilizing ideal types of boundary objects to stimulate conversation and cooperation (Star and Greisemer, 1989).

Chapter 2 establishes a broad countrywide view of trends in the key meteorological variables of precipitation and maximum and minimum temperatures over selected periods of record via several maps. Trends are identified by magnitude, with the depiction of significant trends focusing the user's eye to certain parts of the map, or possibly the whole country when examining minimum temperature trends (e.g., Figure 2.2) (i.e., Venable *et al.*, 2015). Where change occurs is critical to understand when extending the work to other disciplinary spheres, such as for ecologists planning vegetation surveys, or regional and local resource managers developing long-term plans for their communities. There are questions of scale however, that may prove barriers to use of this work on interdisciplinary problems at finer resolutions (Fernandez-Gimenez *et al.*, 2015). The seasonal maps bridge other cross-boundary uses by displaying trends across the landscape in periods significant to ecological work or relevant to remote sensing analyses (summer precipitation trends map, Figure 1.3). For example, the spring seasonal map (not shown) highlights spatiotemporal changes in spring precipitation averages that could be linked to changes in emergence of spring vegetation (e.g., Yu *et al.*, 2003).

The social significance of the hydroclimatic maps is interesting and important. Diverse approaches are needed to support the adaptation of governments and communities to climate change (Lynch *et al.*, 2008). The maps are in an accessible format, easily interpreted, and much more likely to reach a broader disciplinary audience than a paper explaining the same research results. These maps [as objects] highlight the ongoing need for academics to think about how others perceive, interpret, and can use their work (Parker and Crona, 2012). The Chapter 2 map products are more powerful than point-based analysis results (even when mapped) as there is a sense of spatial continuity across the landscape. The smoothed, black box nature of the depictions however, masks uncertainty in the analysis results, which is a common problem when communicating scientific and/or disciplinary results to others (van Pelt *et al*., 2015).

6.5 Conclusion

Each chapter of this dissertation crosses disciplinary boundaries and the analysis results and data lend themselves to the boundary object role, both within the creation and interpretation of the research results of each chapter, and when making connections to the broader research questions of the MOR2 and other Mongolian climate change investigations. The Mongolian Rangelands and Resilience project is a type of *boundary organization* (Parker and Crona, 2012), fulfilling competing missions of creating knowledge, producing boundary objects, and collaborating with actors from universities and other places such as herders, teachers, research institutions, non-governmental organizations, and the Mongolian government. My individual research was made possible by my participation in this project, because as a team member I had access to proprietary hydroclimate data and other essential team information and research products. Therefore my results address selected climate change research project goals as set forth in the original National Science Foundation MOR2 proposal.

Specifically, the aim of the MOR2 project to: "*Seek to understand patterns of climate change across Mongolia and their effects on hydrological and ecological systems and their dynamics…"* was considered in the formulation of all of my research questions. This was particularly reflected in the spatiotemporal trend analyses of Chapter 2, in which "…*meteorological trends and indices [were] mapped for Mongolia.*" In Chapter 3 I used herder surveys and climate data to "…*investigate the relationship between these [climatic] changes and the human system.*" In Chapter 4, I more critically examined underlying problems of data quality that could affect trend analysis results. The long-term paleo proxy investigations of Chapter 5 used statistical modeling to create scenarios of long-term streamflow variability establishing context for future hydrological "…*modeling [that] will allow us to examine hydrological processes across each basin.*"

While the hydometeorological trend analyses and paleo proxy streamflow reconstructions presented in this dissertation meet selected goals of the broader MOR2 project, more importantly, this chapter shows how my work is useful in a cross-boundary research sense. This is true both internally as an interdisciplinary scientist evaluating my own research, and externally when making connections to existing Mongolian socio-ecological and climate change knowledge, and the disciplinary communities creating and using that knowledge. The specifics and uncertainties of my results may or may not be critical to some groups when used as *ideal* boundary objects as defined by Star and Greisemer (1989), or groups may criticize or question my results allowing internal re-interpretation of the work in a reflective manner. The analysis results may be used in whole or in part by other disciplinary groups as a type of repository of information, by focusing on an individual variable, or looking at a suite of variables or analysis results to paint a different and/or more cohesive picture depending on the research questions they are posing. Finally, the results may be used by the broader scientific community in the form of this dissertation document and the publications that are connected to it as a foundation or structure to support other investigations, or simply as talking points when translating the work to related problems of nomadic pastoralist reliance on natural resources, and impacts to those resources under a changing climate.

REFERENCES

Akkerman SF, and Bakker A, 2011. Boundary crossing and boundary objects. *Review of Educational Research*, 81, 132–169, doi:10.3102/0034654311404435.

Beaulieu C, Ouarda T, Seidou O, 2007. A review of homogenization techniques for climate data and their applicability to precipitation series. *Hydrological Sciences Journal-Journal des Sciences Hydrologiques* 52, 18–37, doi:10.1623/hysj.52.1.18.

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, and Ziese M, 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5, 71–99, doi:10.5194/essd-5-71-2013.

Blades JJ, Klos PZ, Kemp KB, Hall TE, Force JE, Morgan P, and Tinkham WT, 2016. Forest managers' response to climate change science: Evaluating the constructs of boundary objects and organizations. *Forest Ecology and Management*, 360, 376–387, doi:10.1016/j.foreco.2015.07.020.

Bruegger RA, Jigjsuren O, and Fernandez-Gimenez ME, 2014. Herder observations of rangeland change in Mongolia: Indicators, causes, and application to community-based management. *Rangeland Ecology & Management*, 67, 119–131, doi:10.2111/REM-D-13-00124.1.

Davi NK, Pederson N, Leland C, Nachin B, Suran B, and Jacoby GC, 2013. Is eastern Mongolia drying? A long-term perspective of a multidecadal trend. *Water Resources Research*, 49, 151–158, doi:10.1029/2012WR011834.

Easterling DR, Peterson TC, and Karl TR, 1996. On the development and use of homogenized climate datasets. *Journal of Climate*, 9, 1429–1434, doi:10.1175/1520-0442(1996)009<1429:OTDAUO>2.0.CO;2.

Ensor LA, and Robeson SM, 2008. Statistical characteristics of daily precipitation: Comparisons of gridded and point datasets. *Journal of Applied Meteorology and Climatology*, 47, 2468–2476, doi:10.1175/2008JAMC1757.1.

Fassnacht SR, Sukh T, Fernandez-Gimenez ME, B Batbuyan, Venable NBH, Laituri M, and Adyabadam G, 2011. Local understanding of hydro-climatic changes in Mongolia. Proceedings of *H02 Symposium Cold Region Hydrology in a Changing Climate* held during IUGG2011 in Melbourne, Australia, July 2011, *IAHS Publication* 346, 120-129.

Fernandez-Gimenez M, 1993. The role of ecological perception in indigenous resource management: A case study from the Mongolian forest-steppe. *Nomadic Peoples*, 33, 31–46.

Fernandez-Gimenez ME, 2000. The role of Mongolian nomadic pastoralists' ecological knowledge in rangeland management. *Ecological Applications* 10, 1318–1326, doi:10.1890/1051-0761(2000)010[1318:TROMNP]2.0.CO;2.

Fernandez-Gimenez ME, Angerer JP, Allegretti AM, Fassnacht SR, Byamba A, J Chantsallkham, Reid RS and Venable NBH, 2015. Integrating herder observations, meteorological data and remote sensing to understand climate change patterns and impacts across an eco-climatic gradient in Mongolia. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 228-234.

Fritts HC, 2001. *Tree Rings and Climate*. Blackburn Press: Caldwell, 567 p., ISBN: 1-930665-39-3.

Gilbert RO, 1987. *Statistical Methods for Environmental Pollution Monitoring.* John Wiley & Sons, New York, 320 pp.

Harris I, Jones PD, Osborn TJ, and Lister DH, 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, 34, 623–642, doi:10.1002/joc.3711.

Helsel DR, and Hirsch RM, 2002. Statistical methods in water resources. *Techniques of Water-Resources Investigations of the United States Geological Survey, Book 4, Hydrologic Analysis and Interpretation, Chapter A3. http://water.usgs.gov/pubs/twri/twri4a3/*. Accessed 2014-11-02.

Kendall MG and Gibbons JD, 1990. *Rank correlation methods*. 5th edition, Edward Arnold, Hodder and Stoughton, London, UK.

Kimble C, Grenier C, and Goglio-Primard K, 2010. Innovation and knowledge sharing across professional boundaries: Political interplay between boundary objects and brokers. *International Journal of Information Management*, 30, 437–444, doi:10.1016/j.ijinfomgt.2010.02.002.

Klein JT, 1996. *Crossing Boundaries: Knowledge, Disciplinarities, and Interdisciplinarities*. University of Virginia Press, 300 pp.

Leland C, 2011. *Hydroclimatic variability in north-central Mongolia as inferred from tree rings*. Unpublished MS Thesis, West Virginia University, Morgantown, WV, USA.

Loaiciga HA, Haston L, and Michaelsen J, 1993. Dendrohydrology and long-term hydrologic phenomena. *Reviews of Geophysics*, 31, 151–171.

Lkhagvadorj D, Hauck M, Dulamsuren CH, and Tsogtbaatar J, 2013. Pastoral nomadism in the forest-steppe of the Mongolian Altai under a changing economy and warming climate. *Journal of Arid Environments* 88, 82-89.

Lynch AH, Tryhorn L and Abramson R, 2008. Working at the boundary - Facilitating interdisciplinarity in climate change adaptation research. *Bulletin of the American Meteorological Society*, 89, 169-179, doi:10.1175/BAMS-89-2-169.

Mann HB, 1945. Nonparametric tests against trend. *Econometrica,* 13, 245–259.

Marin A, 2010. Riders under storms: Contributions of nomadic herder's observations to analyzing climate change in Mongolia. *Global Environmental Change,* 20, 162-176, doi:10.1016/j.gloenvcha.2009.10.004.

McGreavy B, Hutchins K, Smith H, Lindenfeld L, and Silka L, 2013. Addressing the complexities of boundary work in sustainability science through communication. *Sustainability*, 5, 4195–4221, doi:10.3390/su5104195.

Meko DM, and Woodhouse CA, 2007. Chapter 8: Application of streamflow reconstruction to water resources management. *Dendroclimatology*, Developments in Paleoenvironmental Research 11, Hughes MK, Swetnam TW, and Diaz HF [Eds.], Springer, New York, 231-261, doi:10.1007/978-1-4020-5725-0_8

Mitchell TD, and Jones PD, 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, 25, 693–712, doi:10.1002/joc.1181.

Parker J, and Crona B, 2012. On being all things to all people: Boundary organizations and the contemporary research university. *Social Studies of Science*, 42, 262–289, doi:10.1177/0306312711435833.

Pederson N, Leland C, Nachin B, Hessl AE, Bell AR, Martin-Benito D, Davi NK, 2013. Three centuries of shifting hydroclimatic regimes across the Mongolian Breadbasket. *Agricultural and Forest Meteorology*, 178, 10–20.

Regdel D, Dugarzhav C, and Gunin PD, 2012. Ecological demands on socioeconomic development of Mongolia under climate aridization. *Arid Ecosystems*, 2, 1–10, doi:10.1134/S2079096112010076.

Rhoten D, and Parker A, 2004. Education: Risks and rewards of an interdisciplinary research path. *Science*, 306, 2046–2046, doi:10.1126/science.1103628.

Rockstrom J, Steffen W, Noone K, Persson A, Chapin III FS, Lambin EF, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, Nykvist B, de Wit CA, Hughes T, van der Leeuw S, Rodhe H, Sorlen S, Snyder PK, Costanza R, Svedin U, Falkenmark M, Karlberg L, Corell RW, Fabry VJ, Hansen J, Walker B, Liverman D, Richardson K, Crutzen P, and Foley JA, 2009. A safe operating space for humanity. *Nature*, 461, 472–475, doi:10.1038/461472a.

Sen PK, 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association,* 63, 1379-1389.

Star SL, and Griesemer JR, 1989. Institutional ecology, "translations" and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19, 387–420.

Star SL, 2010. This is Not a Boundary Object: Reflections on the origin of a concept. *Science Technology Human Values*, 35, 601–617, doi:10.1177/0162243910377624.

Stockton CW, 1971. The feasibility of augmenting hydrologic records using tree-ring data. Unpublished PhD Dissertation, The University of Arizona, Tucson, AZ, USA, 185 pp.

Sukh T, 2012. Local understanding of hydro-climate changes in Mongolia. Unpublished MS Thesis, Colorado State University, Fort Collins, Colorado, USA.

Sundberg M, 2007. Parameterizations as boundary objects on the climate arena. *Social Studies of Science*, 37, 473–488, doi:10.1177/0306312706075330.

Thiel H, 1950. A rank-invariant method of linear and polynomial regression analysis, I, II, III. *Proceedings of the Royal Netherlands Academy of Sciences* 53, 386-392, 521-525, 1397-1412.

van Pelt SC, Haasnoot M, Arts B, Ludwig F, Swart R, and Biesbroek R, 2015. Communicating climate (change) uncertainties: Simulation games as boundary objects. *Environmental Science and Policy*, 45, 41–52, doi:10.1016/j.envsci.2014.09.004.

Venable NBH, Fassnacht SR, G Adyabadam, and S Tumenjargal, 2012a. Perceived climatic warming and drying near the Khangai Mountains, Mongolia, explored through station record length analysis. Poster session, *32nd Annual American Geophysical Union Hydrology Days*, Colorado State University, Fort Collins, CO, March 23, 2012.

Venable NBH, Fassnacht SR, Adyabadam G, S Tumenjargal, Fernandez-Gimenez M, and B Batbuyan, 2012b. Does the length of station record influence the warming trend that is perceived by Mongolian herders near the Khangai Mountains? *Pirineos*, 167, 69–86, doi:10.3989/Pirineos.2012.167004.

Venable NBH, Fassnacht SR, S Tumenjargal, B Batbuyan, J Odgarav, J Sukhbataar, Fernandez- Gimenez ME, and Adyabadam G, 2012c. Perspectives on hydro-climatic change in rivers sourced from the Khangai Mountains, Mongolia. Oral presentation, *2012 American Geophysical Union Fall Meeting*, San Francisco, CA, December 3, 2012.

Venable NBH, and Fassnacht SR, 2013. Reconstructing a water balance for North Crestone Creek: Streamflow variability and extremes in a snowmelt-dominated internal drainage basin. *Colorado Water: Newsletter of the Water Center of Colorado State University*, 30(5), 10-14.

Venable NBH, Brown PM and Fassnacht SR, 2013. Reconstructing streamflow to nowhere: Long-term variability of flow into the San Luis Valley, Colorado, USA. Poster session, *33rd Annual American Geophysical Union Hydrology Days*, Colorado State University, Fort Collins, CO, March 27th, 2013.

Venable NBH, Fassnacht SR, and Adyabadam G, 2014a. Comparison of point and gridded data sources for use in hydroclimatic modeling of the Khangai Mountain region, Mongolia. Poster session, *34th Annual American Geophysical Union Hydrology Days*, Colorado State University, Fort Collins, CO, March 26th, 2014.

Venable NBH, Fassnacht SR, and Adyabadam G, 2014b. To Grid or not to grid… Precipitation data and hydrological modeling in the Khangai Mountain region of Mongolia. Oral presentation, *2014 American Geophysical Union Fall Meeting*, San Francisco, CA, December 15-19, 2014.

Venable NBH, and Fassnacht SR, 2015a. Climate change assessment issues and implications for hydro-ecology in Mongolia. Oral presentation, *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference*, Ulaanbataar, Mongolia, June 9, 2015.

Venable NBH, and Fassnacht SR, 2015b. Patterns of hydroclimatic variability in proxy: Tree-rings from the Khangai Mountains of Mongolia. Poster presentation, *2015 American Geophysical Union Fall Meeting*, San Francisco, CA, December 14-18, 2015.

Venable NBH, Fassnacht SR, and Hendricks AS, 2015. Spatial changes in climate across Mongolia. Proceedings of *Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference* held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig , Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 73-77.

Whetton P, Hennessy K, Clarke J, McInnes K, and Kent D, 2012. Use of representative climate futures in impact and adaptation assessment. *Climatic Change*, 115, 433–442, doi:10.1007/s10584-012-0471-z.

Willmott CJ, Rowe CM, and Philpot WD, 1985. Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *The American Cartographer*, 12, 5–16.

Willmott, CJ, and Robeson, SM, 1995. Climatologically aided interpolation (CAI) of terrestrial air temperature. *International Journal of Climatology*, 15, 221–229.

Wolf JM, and Venable NBH, 2015. Earlywood, latewood, and adjusted latewood correlations to precipitation: A test case from the Khangai Mountains of Mongolia. Proceedings of

Building Resilience of Mongolia's Rangelands: A Transdisciplinary Conference held June 9-10, 2015. Fernandez-Gimenez ME, B Batkhishig, Fassnacht SR, and Wilson D [Eds]. Nutag Action and Research Institute, Ulaanbaatar, Mongolia, 87-93.

Wyborn C, 2015. Connectivity conservation: Boundary objects, science narratives and the co-production of science and practice. *Environmental Science and Policy*, 51, 292–303, doi:10.1016/j.envsci.2015.04.019.

Yu F, Price KP, Ellis J, and Shi P, 2003. Response of seasonal vegetation development to climatic variations in eastern central Asia. *Remote Sensing of Environment*, 87, 42–54, doi:10.1016/S0034-4257(03)00144-5.

Zalasiewicz J, Williams M, Haywood A, and Ellis M, 2011. The Anthropocene: a new epoch of geological time? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369, 835–841, doi:10.1098/rsta.2010.0339.

APPENDIX A

A.1 Chapter 2 R Code

The following code was used for creating the climate maps for Chapter 2 and works with R software, version 3.1.3 (2015-03-09), "Smooth Sidewalk" and previous versions (R Team, 2015). While any errors or omissions are mine, please be aware that no warranty of any kind is offered with this code and users need to carefully proof and/or modify the code to suit their needs.

A.1.1 Extracting Climate Data and Trend Analyses

```
#-------------------------------------------------------------------------
#  TITLE: Extracting CRU Climate Grids Script for all Mongolia 1963-2012
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2014-02-12
#  LAST REVISION: 2016-01-08
#  DESCRIPTION: This script provides code for extracting climate data from CRU grids.
#  PACKAGES REQUIRED:
#  VARIABLES/DATA USED:
#       NAME:
# TYPE:
# COMMENT:NOTE that this code is not meant to be run without careful comment and uncomment as
necessary to avoid duplication errors in products.
# Also, a change in the way rasters are converted to data frames (new R version and packages??)
resulted in the appending of x,y spatial data to the end of the variable values rather than the
beginning changing the hard-coded subsetting of the dataframes for analysis- so check your work!
#-------------------------------------------------------------------------

#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/")

#libraries
library(raster)
library(rgdal)
library(ncdf4)
library(RColorBrewer)
library(plyr)
library(rasterVis)
library(rkt)

#---------------------------------------------------------------------------------------------
```

```
#import shapefile
mong <-
shapefile("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Shapefiles/Mongolia_bou
ndary.shp")
#plot MN outline (lat long)
plot(mong)

#three <-
shapefile("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Shapefiles/all3aimag.shp
")
three <-
shapefile("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Shapefiles/the3aimag.sh
p")

region <-
shapefile("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Shapefiles/MOR2_AIMAG
_Regions2.shp")

#convert region to lat long, not utm
region2 <- spTransform(region, CRS("+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84
+towgs84=0,0,0"))
plot(region2)

#boundbox
#min      max
#x 87.73567 119.93303
#y 41.58105  52.14925

#-----------------------------------------------------------------------------------
#examine how many stations per gridcell

#temp
statfile1 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
1951.1960.tmp.st0.nc"
statfile2 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
1961.1970.tmp.st0.nc"
statfile3 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
1971.1980.tmp.st0.nc"
statfile4 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
1981.1990.tmp.st0.nc"
statfile5 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
1991.2000.tmp.st0.nc"
statfile6 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
2001.2010.tmp.st0.nc"
statfile7 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/temp/cru_ts3.21.
2011.2012.tmp.st0.nc"
```

```
#precip
statfile1 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.1951.1960.pre.st0.nc"
statfile2 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.1961.1970.pre.st0.nc"
statfile3 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.1971.1980.pre.st0.nc"
statfile4 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.1981.1990.pre.st0.nc"
statfile5 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.1991.2000.pre.st0.nc"
statfile6 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.2001.2010.pre.st0.nc"
statfile7 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/stn/precip/cru_ts3.2
1.2011.2012.pre.st0.nc"

statfile <-list(statfile1, statfile2, statfile3, statfile4, statfile5, statfile6, statfile7)

#loop to open each file, crop to the MN bounding box extent, and save as a gridfile

stat.stack<-stack()

for (i in 1:length(statfile)) {

#open file and make brick
varibname <-"st0"
varst <-brick(statfile[[i]], varname=varibname)

#set extent for analysis
varst.extent <- extent(87.73567, 119.93303, 41.58105, 52.14925)
varst.aoi <- crop(varst, varst.extent)

#export file to raster stack
stat.stack <-addLayer(stat.stack,varst.aoi)
}

#test plots
plot(stat.stack, 744)
plot(mong, add=TRUE)

#extract the total number of stations in Mongolia for each layer in the raster stack
stat.stackmn <-extract(stat.stack, mong, fun=sum)

#transpose vector and add year column for aggregating
stat.stackmnt <-as.data.frame(t(stat.stackmn))
colnames(stat.stackmnt) <-"no_stations"
stat.stackmnt$year <-substr(rownames(stat.stackmnt), 2, 5)
```

```
stat.stackyr <-ddply(stat.stackmnt, .(year), summarize, no_stats=mean(no_stations))

#export as csv
write.csv(stat.stackyr,"CRU_Temp_Station_Counts_By_Year")
write.csv(stat.stackyr,"CRU_Precip_Station_Counts_By_Year")

bp <-barplot(stat.stackyr$no_stats, xaxt="n", main="Average Number of Stations per Year (Precip)")
labs <-paste(stat.stackyr$year)
text(cex=0.5, x=bp-.25, y=-1.25, labs, xpd=TRUE, srt=45)

#-------------------------------------------------------------------------------------
#Step 1:open each file, crop and stack to one time series

#tmax
#varfile1 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.195
1.1960.tmx.dat.nc"
#varfile2 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.196
1.1970.tmx.dat.nc"
#varfile3 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.197
1.1980.tmx.dat.nc"
#varfile4 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.198
1.1990.tmx.dat.nc"
#varfile5 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.199
1.2000.tmx.dat.nc"
#varfile6 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.200
1.2010.tmx.dat.nc"
#varfile7 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmax/cru_ts3.21.201
1.2012.tmx.dat.nc"

#tmin
varfile1 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.1951
.1960.tmn.dat.nc"
varfile2 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.1961
.1970.tmn.dat.nc"
varfile3 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.1971
.1980.tmn.dat.nc"
varfile4 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.1981
.1990.tmn.dat.nc"
varfile5 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.1991
.2000.tmn.dat.nc"


varfile6 <-
```

```
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.2001
.2010.tmn.dat.nc"
varfile7 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/tmin/cru_ts3.21.2011
.2012.tmn.dat.nc"

#precip
varfile1 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.19
51.1960.pre.dat.nc"
varfile2 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.19
61.1970.pre.dat.nc"
varfile3 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.19
71.1980.pre.dat.nc"
varfile4 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.19
81.1990.pre.dat.nc"
varfile5 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.19
91.2000.pre.dat.nc"
varfile6 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.20
01.2010.pre.dat.nc"
varfile7 <-
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/CRU_TS3_21/precip/cru_ts3.21.20
11.2012.pre.dat.nc"

varfile <-list(varfile1, varfile2, varfile3, varfile4, varfile5, varfile6, varfile7)

#loop to open each file, crop to the MN bounding box extent, and save as a gridfile
#remember to change variable names!

var.stack<-stack()

for (i in 1:length(varfile)) {

 #open file and make brick
 #varibname <-"tmx"
 varibname <-"tmn"
 #varibname <-"pre"
 varst <-brick(varfile[[i]], varname=varibname)

 #set extent for analysis
 varst.extent <- extent(87.73567, 119.93303, 41.58105, 52.14925)
 varst.aoi <- crop(varst, varst.extent)

 #export file to raster stack
 var.stack <-addLayer(var.stack,varst.aoi)
}

#Move to seasonal analysis step fom here

#extract LOR of interest (1963-2012)
```

```
var.stacklor <-subset(var.stack, 145:744)

#tmax.stack <-var.stacklor
tmin.stack <-var.stacklor
#pre.stack <-var.stacklor

#-----------------------------------------------------------------------------------------------
#Annual precipitation rasters- sum each 12 month period and export each raster to a stack
precip.dates <-seq(as.Date("1963-01-01"), as.Date("2012-12-01"), "months")
preann.stack<-stack()

for (i in seq(0,588, by=12)) {

 #open file and make brick
 mon.br <-subset(var.stacklor, (i+1):(i+12))

 #sum precipitation in that subset
 rast.montime <-setZ(mon.br, precip.dates[(i+1):(i+12)], "months")
 rast.sum <- zApply(rast.montime, by="months", fun=sum)

 #export file to raster stack
 preann.stack <-addLayer(preann.stack,rast.sum)
}

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(preann.stack) <-rast.year

#test plots
plot(pre.stack,6)
plot(mong, add=TRUE)

plot(preann.stack,6)
plot(mong, add=TRUE)

#---------------------------------------------------------------------------------
#Step 2 calculate mean for each stack over LOR
#Note: Alyssa's work extracted raster values (as points) then performed time series analysis on a
matrix of those values
#she applied the functions needed to the matrix, then converted back to a raster.

#use zApply() from raster package for mean to apply to rasters directly
#for tmax, tmin
rast.date <-seq(as.Date("1963-01-01"), as.Date("2012-12-01"), "months")
#for precip
#rast.date <-seq(as.Date("1963-01-01"), as.Date("2012-12-01"), "years")

#For tmax, tmin, precip
rast.time <-setZ(var.stacklor, rast.date, "months")
#rast.time <-setZ(preann.stack, rast.date, "years")

rast.mean <- zApply(rast.time, by="months", fun=mean)
#rast.mean <- zApply(rast.time, by="years", fun=mean)

#rast.ext <-getValues(rast.mean)
```

```
#for ESRI use, but only 1 layer at a time
#writeRaster(rast.mean, "MN_Mean_Annual_Tmax_1963-2012.asc", format="ascii")
writeRaster(rast.mean, "MN_Mean_Annual_Tmin_1963-2012.asc", format="ascii")
#writeRaster(rast.mean, "MN_Mean_Annual_Precip_1963-2012.asc", format="ascii")

#Files for import
#writeRaster(rast.mean, "MN_Mean_Annual_Tmax_1963-2012.grd")
writeRaster(rast.mean, "MN_Mean_Annual_Tmin_1963-2012.grd")
#writeRaster(rast.mean, "MN_Mean_Annual_Precip_1963-2012.grd")

#compare to extraction of values (as SpatialPointsDataFrame) and get mean then back to a raster
#rast.xyz <-rasterToPoints (tmax.stack, spatial=TRUE)
#rast.df <-as.data.frame(rast.xyz)
#rast.dfnc <-rast.df[,3:602]
#rast.xyzm <-apply(rast.dfnc, MARGIN=1, FUN=mean)

#test plot the result
plot(rast.mean)

#---------------------------------------------------------------------------
#Step/Loop 2: Trends and Significance
#annual trends- avg for tmin/tmax, sum for P
#use raster stacks
#tmax.stack
#tmin.stack
#preann.stack

#for tmin and tmax want to get mean annual values first for each year
temp.dates <-seq(as.Date("1963-01-01"), as.Date("2012-12-01"), "months")
tmpann.stack<-stack()

for (i in seq(0,588, by=12)) {

 #open file and make brick
 #mon.br <-subset(tmax.stack, (i+1):(i+12))
 mon.br <-subset(tmin.stack, (i+1):(i+12))

 #avg temp in that subset
 rast.montime <-setZ(mon.br, temp.dates[(i+1):(i+12)], "months")
 rast.avg <- zApply(rast.montime, by="months", fun=mean)

 #export file to raster stack
 tmpann.stack <-addLayer(tmpann.stack,rast.avg)
}

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(tmpann.stack) <-rast.year

#make unique stack
#tmaxann.stack <-tmpann.stack
tminann.stack <-tmpann.stack

#annual mean temp use tmpann.stack values (for double check of plots and method)
#rast.date <-seq(as.Date("1963-01-01"), as.Date("2012-12-01"), "years")
```

```
#rast.time <-setZ(tmaxann.stack, rast.date, "years")
#rast.time <-setZ(tminann.stack, rast.date, "years")
#rast.mean <- zApply(rast.time, by="years", fun=mean)
#rast.ext <-getValues(rast.mean)

#use unique 50 year stacks for trend analyses (cannot use Zapply, so make into df)
Years <-c(1963:2012)
#rast.xyz <-rasterToPoints (tmaxann.stack, spatial=TRUE)
#rast.coord <-coordinates(tmaxann.stack)
#rast.extent <-extent(tmaxann.stack)
rast.xyz <-rasterToPoints (tminann.stack, spatial=TRUE)
rast.coord <-coordinates(tminann.stack)
rast.extent <-extent(tminann.stack)
#rast.xyz <-rasterToPoints (preann.stack, spatial=TRUE)
#rast.coord <-coordinates(preann.stack)
#rast.extent <-extent(preann.stack)
rast.df <-as.data.frame(rast.xyz)
rast.dfnc <-rast.df[,1:50]
rast.xyztrend <-apply(rast.dfnc, MARGIN=1, FUN=function(x) rkt(Years, x))

#export slope and pval to make trend raster
trend.slope <-sapply(rast.xyztrend, "[[", 3)
trend.pval <-sapply(rast.xyztrend, "[[", 1)

#fill a raster with these value and export
rast.pval <-raster(rast.extent, nrows=21 , ncols=65, crs="+proj=longlat +datum=WGS84
+ellps=WGS84 +towgs84=0,0,0")
values(rast.pval) <-trend.pval

#writeRaster(rast.pval, "Annual_Tmax_Pval.asc", format="ascii")
writeRaster(rast.pval, "Annual_Tmin_Pval.asc", format="ascii")
#writeRaster(rast.pval, "Annual_P_Pval.asc", format="ascii")

#make all pvals less than 0.05= NA i.e. to get points of non-significance
trend.pval[trend.pval<0.05 ] <-NA
trend.pvaldf <-as.data.frame(trend.pval)
trend.pvaldfc <-data.frame(rast.coord, trend.pvaldf)
trend.pvaldfs <-subset(trend.pvaldfc, !is.na(trend.pvaldfc$trend.pval))

#write point locations of nonsignificance to file for use in plotting
#write.csv(trend.pvaldfs, "Tmax_annual_non_significant_pval_locs_ANN.csv")
write.csv(trend.pvaldfs, "Tmin_annual_non_significant_pval_locs_ANN.csv")
#write.csv(trend.pvaldfs, "Precip_annual_non_significant_pval_locs_ANN.csv")

#locations of non-significant pvals to use as spatial points for plotting
trend.pvalxy <-trend.pvaldfs[,1:2]

#convert to raster
rast.slope <-raster(rast.extent, nrows=21 , ncols=65, crs="+proj=longlat +datum=WGS84
+ellps=WGS84 +towgs84=0,0,0")
values(rast.slope) <-trend.slope

#for ESRI use, but only 1 layer at a time
#writeRaster(rast.slope, "Trend_Slope_Tmax_ANN.asc", format="ascii")
writeRaster(rast.slope, "Trend_Slope_Tmin_ANN.asc", format="ascii")
```

```
#writeRaster(rast.slope, "Trend_Slope_Pre_ANN.asc", format="ascii")

#writeRaster(rast.slope, "Trend_Slope_Tmax_ANN.grd")
writeRaster(rast.slope, "Trend_Slope_Tmin_ANN.grd")
#writeRaster(rast.slope, "Trend_Slope_Pre_ANN.grd")

#---------------------------------------------------------------------------
#Step 3: Seasonal Trends and Significance  (DJF, MAM, JJA, SON)
#For data stack run step 1 for each variable then through seasonal processing to gridded and csv
outputs

#extract LOR of interest (1962-2012, seasonal)
#First, make a sequence for those three months for LOR
#For temp need seasonal mean, as mean of 3 months each "year" and then trend of those means over
the time series
#For precip need seasonal sum, as sum of 3 months each "year" and then trend of those sums over
the time series

#Winter-DJF
dec <- seq(from=144, to=732, by=12)
jan <-seq(from=145, to=733, by=12)
feb <-seq(from=146, to=734, by=12)
djf <-c(dec, jan, feb)
djf.o <-djf[order(djf) ]

var.stacklordjf <-subset(var.stack, djf.o)

dectemp.dates <-seq(as.Date("1962-12-01"), as.Date("2011-12-01"), "years")
jantemp.dates <-seq(as.Date("1963-01-01"), as.Date("2012-01-01"), "years")
febtemp.dates <-seq(as.Date("1963-02-01"), as.Date("2012-02-01"), "years")

djftemp.dates <-c(dectemp.dates, jantemp.dates, febtemp.dates)
djftemp.dateso <-djftemp.dates[order(djftemp.dates)]

#Loop gives mean of each 3 month period per year for temperature
tmpseas.stackdjf<-stack()

for (i in seq(0,150, by=3)) {

 #open file and make brick
 three.br <-subset(var.stacklordjf, (i+1):(i+3))

 #avg temp in that subset
 rast.threetime <-setZ(three.br, djftemp.dateso[(i+1):(i+3)], "months")
 rast.avg <- zApply(rast.threetime, by="months", fun=mean)

 #export file to raster stack
 tmpseas.stackdjf <-addLayer(tmpseas.stackdjf,rast.avg)
}
#gives error:Error in .local(x, ...) : not a valid subset but values appear valid?!

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(tmpseas.stackdjf) <-rast.year
```

```
#Loop gives sum of each 3 period per year for precipitation
preseas.stackdjf<-stack()

for (i in seq(0,150, by=3)) {

  #open file and make brick
  three.br <-subset(var.stacklordjf, (i+1):(i+3))

  #sum precip in that subset
  rast.threetime <-setZ(three.br, djftemp.dateso[(i+1):(i+3)], "months")
  rast.avg <- zApply(rast.threetime, by="months", fun=sum)

  #export file to raster stack
  preseas.stackdjf <-addLayer(preseas.stackdjf,rast.avg)
}
#gives error:Error in .local(x, ...) : not a valid subset but values appear valid?!

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(preseas.stackdjf) <-rast.year

#Spring-MAM
mar <- seq(from=147, to=735, by=12)
apr <-seq(from=148, to=736, by=12)
may <-seq(from=149, to=737, by=12)
mam <-c(mar, apr, may)
mam.o <-mam[order(mam) ]

var.stacklormam <-subset(var.stack, mam.o)

martemp.dates <-seq(as.Date("1963-03-01"), as.Date("2012-03-01"), "years")
aprtemp.dates <-seq(as.Date("1963-04-01"), as.Date("2012-04-01"), "years")
maytemp.dates <-seq(as.Date("1963-05-01"), as.Date("2012-05-01"), "years")

mamtemp.dates <-c(martemp.dates, aprtemp.dates, maytemp.dates)
mamtemp.dateso <-mamtemp.dates[order(mamtemp.dates)]

#Loop gives mean of each 3 month period per year for temperature
tmpseas.stackmam<-stack()

for (i in seq(0,150, by=3)) {

  #open file and make brick
  three.br <-subset(var.stacklormam,(i+1):(i+3))

  #avg temp in that subset
  rast.threetime <-setZ(three.br, mamtemp.dateso[(i+1):(i+3)], "months")
  rast.avg <- zApply(rast.threetime, by="months", fun=mean)

  #export file to raster stack
  tmpseas.stackmam <-addLayer(tmpseas.stackmam,rast.avg)
}

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
```

```
names(tmpseas.stackmam) <-rast.year

#Loop gives sum of each 3 period per year for precipitation
preseas.stackmam<-stack()

for (i in seq(0,150, by=3)) {

 #open file and make brick
 three.br <-subset(var.stacklormam, (i+1):(i+3))

 #sum precip in that subset
 rast.threetime <-setZ(three.br, mamtemp.dateso[(i+1):(i+3)], "months")
 rast.avg <- zApply(rast.threetime, by="months", fun=sum)

 #export file to raster stack
 preseas.stackmam <-addLayer(preseas.stackmam,rast.avg)
}
#gives error:Error in .local(x, ...) : not a valid subset but values appear valid.

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(preseas.stackmam) <-rast.year

#Summer-JJA
jun <- seq(from=150, to=738, by=12)
jul <-seq(from=151, to=739, by=12)
aug <-seq(from=152, to=740, by=12)
jja <-c(jun, jul, aug)
jja.o <-jja[order(jja) ]

var.stacklorjja <-subset(var.stack, jja.o)

juntemp.dates <-seq(as.Date("1963-06-01"), as.Date("2012-06-01"), "years")
jultemp.dates <-seq(as.Date("1963-07-01"), as.Date("2012-07-01"), "years")
augtemp.dates <-seq(as.Date("1963-08-01"), as.Date("2012-08-01"), "years")

jjatemp.dates <-c(juntemp.dates, jultemp.dates, augtemp.dates)
jjatemp.dateso <-jjatemp.dates[order(jjatemp.dates)]

#Loop gives mean of each 3 month period per year for temperature
tmpseas.stackjja<-stack()

for (i in seq(0,150, by=3)) {

 #open file and make brick
 three.br <-subset(var.stacklorjja, (i+1):(i+3))

 #avg temp in that subset
 rast.threetime <-setZ(three.br, jjatemp.dateso[(i+1):(i+3)], "months")
 rast.avg <- zApply(rast.threetime, by="months", fun=mean)


 #export file to raster stack
 tmpseas.stackjja <-addLayer(tmpseas.stackjja,rast.avg)
}
```

```
#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(tmpseas.stackjja) <-rast.year

#Loop gives sum of each 3 period per year for precipitation
preseas.stackjja<-stack()

for (i in seq(0,150, by=3)) {

  #open file and make brick
  three.br <-subset(var.stacklorjja, (i+1):(i+3))

  #sum precip in that subset
  rast.threetime <-setZ(three.br, jjatemp.dateso[(i+1):(i+3)], "months")
  rast.avg <- zApply(rast.threetime, by="months", fun=sum)

  #export file to raster stack
  preseas.stackjja <-addLayer(preseas.stackjja,rast.avg)
}
#gives error:Error in .local(x, ...) : not a valid subset but values appear valid.

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(preseas.stackjja) <-rast.year

#Fall-SON
sep <- seq(from=153, to=741, by=12)
oct <-seq(from=154, to=742, by=12)
nov <-seq(from=155, to=743, by=12)
son <-c(sep, oct, nov)
son.o <-son[order(son) ]

var.stacklorson <-subset(var.stack, son.o)

septemp.dates <-seq(as.Date("1963-09-01"), as.Date("2012-09-01"), "years")
octtemp.dates <-seq(as.Date("1963-10-01"), as.Date("2012-10-01"), "years")
novtemp.dates <-seq(as.Date("1963-11-01"), as.Date("2012-11-01"), "years")

sontemp.dates <-c(septemp.dates, octtemp.dates, novtemp.dates)
sontemp.dateso <-sontemp.dates[order(sontemp.dates)]

#Loop gives mean of each 3 month period per year for temperature
tmpseas.stackson<-stack()

for (i in seq(0,150, by=3)) {

  #open file and make brick
  three.br <-subset(var.stacklorson, (i+1):(i+3))

    #avg temp in that subset
    rast.threetime <-setZ(three.br, sontemp.dateso[(i+1):(i+3)], "months")
    rast.avg <- zApply(rast.threetime, by="months", fun=mean)

    #export file to raster stack
```

```
tmpseas.stackson <-addLayer(tmpseas.stackson,rast.avg)
}

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(tmpseas.stackson) <-rast.year

#Loop gives sum of each 3 month period per year for precipitation
preseas.stackson<-stack()

for (i in seq(0,150, by=3)) {

  #open file and make brick
  three.br <-subset(var.stacklorson, (i+1):(i+3))

  #sum precip in that subset
  rast.threetime <-setZ(three.br, sontemp.dateso[(i+1):(i+3)], "months")
  rast.avg <- zApply(rast.threetime, by="months", fun=sum)

  #export file to raster stack
  preseas.stackson <-addLayer(preseas.stackson,rast.avg)
}
#gives error:Error in .local(x, ...) : not a valid subset but values appear valid?!

#for clarity name each raster in the stack by year
rast.year <-seq(1963, 2012, 1)
names(preseas.stackson) <-rast.year


#For trend analyses
#make unique stack seasonally
#tmaxseas.stack <-tmpseas.stackdjf
#tmaxseas.stack <-tmpseas.stackmam
#tmaxseas.stack <-tmpseas.stackjja
#tmaxseas.stack <-tmpseas.stackson

#tminseas.stack <-tmpseas.stackdjf
#tminseas.stack <-tmpseas.stackmam
#tminseas.stack <-tmpseas.stackjja
#tminseas.stack <-tmpseas.stackson

preseas.stack <-preseas.stackdjf
#preseas.stack <-preseas.stackmam
#preseas.stack <-preseas.stackjja
#preseas.stack <-preseas.stackson

#run trend analysis on stack
#use unique 50 year stacks for trend analyses (cannot use Zapply, so make into df)
Years <-c(1963:2012)

#rast.xyz <-rasterToPoints (tmaxseas.stack, spatial=TRUE)
#rast.coord <-coordinates(tmaxseas.stack)
#rast.extent <-extent(tmaxseas.stack)

#rast.xyz <-rasterToPoints (tminseas.stack, spatial=TRUE)
```

```
#rast.coord <-coordinates(tminseas.stack)
#rast.extent <-extent(tminseas.stack)

rast.xyz <-rasterToPoints(preseas.stack, spatial=TRUE)
rast.coord <-coordinates(preseas.stack)
rast.extent <-extent(preseas.stack)

rast.df <-as.data.frame(rast.xyz)
rast.dfnc <-rast.df[,1:50]
rast.xyztrend <-apply(rast.dfnc, MARGIN=1, FUN=function(x) rkt(Years, x))

#export slope and pval to make trend raster
trend.slope <-sapply(rast.xyztrend, "[[", 3)
trend.pval <-sapply(rast.xyztrend, "[[", 1)

#fill a raster with these value and export
rast.pval <-raster(rast.extent, nrows=21 , ncols=65, crs="+proj=longlat +datum=WGS84
+ellps=WGS84 +towgs84=0,0,0")
values(rast.pval) <-trend.pval

#writeRaster(rast.pval, "Annual_Tmax_Pval_DJF.asc", format="ascii")
#writeRaster(rast.pval, "Annual_Tmax_Pval_MAM.asc", format="ascii")
#writeRaster(rast.pval, "Annual_Tmax_Pval_JJA.asc", format="ascii")
#writeRaster(rast.pval, "Annual_Tmax_Pval_SON.asc", format="ascii")

#writeRaster(rast.pval, "Annual_Tmin_Pval_DJF.asc", format="ascii")
#writeRaster(rast.pval, "Annual_Tmin_Pval_MAM.asc", format="ascii")
#writeRaster(rast.pval, "Annual_Tmin_Pval_JJA.asc", format="ascii")
#writeRaster(rast.pval, "Annual_Tmin_Pval_SON.asc", format="ascii")

writeRaster(rast.pval, "Annual_P_Pval_DJF.asc", format="ascii")
#writeRaster(rast.pval, "Annual_P_Pval_MAM.asc", format="ascii")
#writeRaster(rast.pval, "Annual_P_Pval_JJA.asc", format="ascii")
#writeRaster(rast.pval, "Annual_P_Pval_SON.asc", format="ascii")

#make all pvals less than 0.05= NA i.e. to get points of non-significance
trend.pval[trend.pval<0.05 ] <-NA
trend.pvaldf <-as.data.frame(trend.pval)
trend.pvaldfc <-data.frame(rast.coord, trend.pvaldf)
trend.pvaldfs <-subset(trend.pvaldfc, !is.na(trend.pvaldfc$trend.pval))

#Write point locations of nonsignificance to file for use in plotting
#write.csv(trend.pvaldfs, "Tmax_annual_non_significant_pval_locs_djf.csv")
#write.csv(trend.pvaldfs, "Tmax_annual_non_significant_pval_locs_mam.csv")
#write.csv(trend.pvaldfs, "Tmax_annual_non_significant_pval_locs_jja.csv")
#write.csv(trend.pvaldfs, "Tmax_annual_non_significant_pval_locs_son.csv")

#write.csv(trend.pvaldfs, "Tmin_annual_non_significant_pval_locs_djf.csv")
#write.csv(trend.pvaldfs, "Tmin_annual_non_significant_pval_locs_mam.csv")
#write.csv(trend.pvaldfs, "Tmin_annual_non_significant_pval_locs_jja.csv")
#write.csv(trend.pvaldfs, "Tmin_annual_non_significant_pval_locs_son.csv")

write.csv(trend.pvaldfs, "Precip_annual_non_significant_pval_locs_djf.csv")
#write.csv(trend.pvaldfs, "Precip_annual_non_significant_pval_locs_mam.csv")
#write.csv(trend.pvaldfs, "Precip_annual_non_significant_pval_locs_jja.csv")
```

```
#write.csv(trend.pvaldfs, "Precip_annual_non_significant_pval_locs_son.csv")

#locations of non-significant pvals to use as spatial points for plotting
trend.pvalxy <-trend.pvaldfs[,1:2]

#convert to raster
rast.slope <-raster(rast.extent, nrows=21 , ncols=65, crs="+proj=longlat +datum=WGS84
+ellps=WGS84 +towgs84=0,0,0")
values(rast.slope) <-trend.slope

#writeRaster(rast.slope, "Trend_Slope_Tmax_DJF.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Tmax_MAM.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Tmax_JJA.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Tmax_SON.asc", format="ascii")

#writeRaster(rast.slope, "Trend_Slope_Tmin_DJF.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Tmin_MAM.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Tmin_JJA.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Tmin_SON.asc", format="ascii")

writeRaster(rast.slope, "Trend_Slope_Pre_DJF.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Pre_MAM.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Pre_JJA.asc", format="ascii")
#writeRaster(rast.slope, "Trend_Slope_Pre_SON.asc", format="ascii")

#As gridfiles for plotting
#writeRaster(rast.slope, "Trend_Slope_Tmax_DJF.grd")
#writeRaster(rast.slope, "Trend_Slope_Tmax_MAM.grd")
#writeRaster(rast.slope, "Trend_Slope_Tmax_JJA.grd")
#writeRaster(rast.slope, "Trend_Slope_Tmax_SON.grd")

#writeRaster(rast.slope, "Trend_Slope_Tmin_DJF.grd")
#writeRaster(rast.slope, "Trend_Slope_Tmin_MAM.grd")
#writeRaster(rast.slope, "Trend_Slope_Tmin_JJA.grd")
#writeRaster(rast.slope, "Trend_Slope_Tmin_SON.grd")

writeRaster(rast.slope, "Trend_Slope_Pre_DJF.grd")
#writeRaster(rast.slope, "Trend_Slope_Pre_MAM.grd")
#writeRaster(rast.slope, "Trend_Slope_Pre_JJA.grd")
#writeRaster(rast.slope, "Trend_Slope_Pre_SON.grd")

#--------------------------------------------------------------------------------
#Extract grid values for data archiving

#ncol 65
#nrow 21
#ncell
rastcell <-seq(1:1365)

#LOR annual values
rasttmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/MN_Mean_Annual_
Tmax_1963-2012.grd")
#rasttmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/MN_Mean_Annual_
```

```
Tmin_1963-2012.grd")
#rastpre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/MN_Mean_Annual_
Precip_1963-2012.grd")

rastlocs <-xyFromCell(rasttmax, rastcell)
#rastlocs <-xyFromCell(rasttmin, rastcell)
#rastlocs <-xyFromCell(rastpre, rastcell)

rastval <-getValues (rasttmax)
#rastval <-getValues (rasttmin)
#rastval <-getValues (rastpre)

rastlocval <-data.frame(rastlocs, rastval)

write.csv(rastlocval, "MN_Mean_Annual_Tmax_GridVals.csv")
#write.csv(rastlocval, "MN_Mean_Annual_Tmin_GridVals.csv")
#write.csv(rastlocval, "MN_Mean_Annual_Precip_GridVals.csv")

#for seasonal values
#raster of interest
filesuf <-list(c("ANN", "DJF", "MAM", "JJA", "SON"))
filetmax <-
paste("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/Trend_Slope_Tmax
_", filesuf[[1]], ".grd", sep="")
filetmin <-
paste("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/Trend_Slope_Tmin_
", filesuf[[1]], ".grd", sep="")
filepre <-
paste("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/Trend_Slope_Pre_",
filesuf[[1]], ".grd", sep="")

filetmaxpval <-
paste("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Export_Data/Annual_Tmax_P
val_", filesuf[[1]], ".asc", sep="")
filetminpval <-
paste("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Export_Data/Annual_Tmin_P
val_", filesuf[[1]], ".asc", sep="")
fileprepval <-
paste("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Export_Data/Annual_P_Pval_
", filesuf[[1]], ".asc", sep="")

seaslist <-c("ANN", "DJF", "MAM", "JJA", "SON")

#for (i in 1:length(filetmax)) {
#for (i in 1:length(filetmin)) {
for (i in 1:length(filepre)) {

 #open raster, extract location and cell value
 #rasterfile <-raster(filetmax[i])
 #rasterfile <-raster(filetmin[i])
 rasterfile <-raster(filepre[i])

 rastlocs <-xyFromCell(rasterfile, rastcell)
 rastval <-getValues (rasterfile)
```

```
 rastlocval <-data.frame(rastlocs, rastval)

 #open  pval grid
 #pval <-raster(filetmaxpval[i])
 #pval <-raster(filetminpval[i])
 pval <-raster(fileprepval[i])

 rastlocspv <-xyFromCell(pval, rastcell)
 rastvalpv <-getValues (pval)
 rastlocvalpv <-data.frame(rastlocspv, rastvalpv)

 #append signif to raster vals
 rastlocvalsig <-data.frame(rastlocval, rastlocvalpv[,3])

 #create column denoting significance
 rastlocvalsig$sig <-ifelse(rastlocvalsig[,4]<0.05, "Y", NA)

 #truncate and write output file
 rastpvalf <-rastlocvalsig[ , c(1:3,5)]
 #write.csv(rastpvalf, paste("filetmax_",seaslist[i] ,".csv",sep=""))
 #write.csv(rastpvalf, paste("filetmin_",seaslist[i] ,".csv",sep=""))
 write.csv(rastpvalf, paste("filepre_",seaslist[i] ,".csv",sep=""))
}

#testing results using get cell values from cell number and get value from csv
#test cell
#col=30, row=1
#cell No. 30
xymat <-matrix(c(102.25,  51.75), ncol=2, nrow=1)

#testval <-matrix(ncol=4, nrow=length(filetmax), NA)
#testval <-matrix(ncol=4, nrow=length(filetmin), NA)
testval <-matrix(ncol=4, nrow=length(filepre), NA)

#for (i in 1:length(filetmax)) {
#for (i in 1:length(filetmin)) {
for (i in 1:length(filepre)) {

 #open raster, check raster value given xy location
 #rasterfile <-raster(filetmax[i])
 #rasterfile <-raster(filetmin[i])
 rasterfile <-raster(filepre[i])

 rastcell <-cellFromXY(rasterfile, xymat)
 rastval <-getValues(rasterfile, row=1)
 #extract #30
 rastval30 <-rastval[30]

 #open pval locations raster
 #pval <-raster(filetmaxpval[i])
 #pval <-raster(filetminpval[i])
 pval <-raster(fileprepval[i])

 #extract value of interest
 pvalval <-getValues(pval, row=1)
```

```
 #extract #30
 pval30 <-pvalval[30]
 pvalsig <-ifelse(pval30<0.05, "Y", NA)

 #put results in a matrix (by seas)
 testval[i,1] <-rastcell
 testval[i,2] <-rastval30
 testval[i,3] <-pvalsig
 testval[i,4] <-seaslist[i]
}

#For annual LOR means
#rasttmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/MN_Mean_Annual_
Tmax_1963-2012.grd")
#rasttmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/MN_Mean_Annual_
Tmin_1963-2012.grd")
rastpre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Gridfiles/MN_Mean_Annual_
Precip_1963-2012.grd")

#rasterfile <-rasttmax
#rasterfile <-rasttmin
rasterfile <-rastpre

rastcell <-cellFromXY(rasterfile, xymat)
rastval <-getValues(rasterfile, row=1)
#extract #30
rastval30 <-rastval[30]

c(rastcell, rastval30)

#---------------------------------------------------------------------------------
#LOR plots for pub
#levelplot of mean of tmax, tmin, and precip

#stacked levelplots
#note: output with no lat long text and no heading text in case of rescaling.
#One legend for temp range and one for precip

#import rasters to plot
rast.meantmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/MN_M
ean_Annual_Tmax_1963-2012.grd")
rast.meantmax2 <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/MN_M
ean_Annual_Tmax_1963-2012.grd")
rast.meantmin <-raster(
"/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/MN_Mean_An
nual_Tmin_1963-2012.grd")
rast.meanpre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/MN_M
ean_Annual_Precip_1963-2012.grd")
```

```
#find max and min values each grid for setting legend
rastmaxmin.tmax <-getValues(rast.meantmax)
rastmaxmin.tmin <-getValues(rast.meantmin)
rastmaxmin.pre <-getValues(rast.meanpre)

max(rastmaxmin.tmax)
[1] 21.7075
min(rastmaxmin.tmax)
[1] -2.263833

max(rastmaxmin.tmin)
[1] 8.167167
min(rastmaxmin.tmin)
[1] -16.66133

max(rastmaxmin.pre)
[1] 891.898
min(rastmaxmin.pre)
[1] 32.874

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
heat.theme <-rasterTheme(region=rev(heat.colors(n=7)))
tmax.plot <-levelplot(rast.meantmax,colorkey=list(at=c(-5,0,5,10,15,20,25),labels=list(c('-5',' 0','
5','10', '15', '20', '25'), cex=1.5), col=rev(heat.colors(n=7))), scales=list(x=x.scale, y=y.scale),
ylab=NULL, xlab=NULL, margin=FALSE, par.settings=heat.theme, main=list(expression(paste("b)
Mean Annual Maxiumum Temperature 1963-2012 (",degree,"C)")), x=0.07, cex=1.5, just="left"))
tmax.plot+ layer(sp.polygons(mong))

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
cool.theme <-rasterTheme(region=rev(brewer.pal(7, "Blues")))
tmin.plot <-levelplot(rast.meantmin, colorkey=list(at=c(-20,-15, -10, -5, 0, 5, 10),labels=list(c('-20','-
15', '-10',' -5.0',' 0',' 5', ' 10'), cex=1.5), col=rev(brewer.pal(7, "Blues"))),scales=list(x=x.scale,
y=y.scale),
ylab=NULL, xlab=NULL, margin=FALSE, par.settings=cool.theme,main=list(expression(paste("c)
Mean Annual Minimum Temperature 1963-2012 (",degree,"C)")), x=0.07, cex=1.5, just="left"))
tmin.plot+ layer(sp.polygons(mong))

#sig.themep <-rasterTheme(region=(brewer.pal(11, "RdBu")))
#pre.plot <-levelplot(rast.meanpre, margin=FALSE,  par.settings=sig.themep, colorkey=list(at=c(30,
120, 210,300,390,480,570,660,750,840,930),labels=c(as.character(c(30, 120,
210,300,390,480,570,660,750,840,930))), col=brewer.pal(11, "RdBu")), main="Average Annual
Precipitation 1963-2012")
#pre.plot+ layer(sp.polygons(mong))

#sig.themep <-rasterTheme(region=(brewer.pal(9, "Greens")))
#pre.plot <-levelplot(rast.meanpre, margin=FALSE,  par.settings=sig.themep, colorkey=list(at=c(30,
130, 230,330,430,530,630,730,830),labels=c(as.character(c(30, 130,
230,330,430,530,630,730,830))), col=brewer.pal(9, "Greens")), main="Average Annual Precipitation
1963-2012")
#pre.plot+ layer(sp.polygons(mong))

#sig.themep <-rasterTheme(region=(brewer.pal(10, "RdBu")))
#pre.plot <-levelplot(rast.meanpre, margin=FALSE,  par.settings=sig.themep, colorkey=list(at=c(0,
```

```
100, 200,300,400,500,600,700,800,900),labels=c(as.character(c(0, 100,
200,300,400,500,600,700,800,900))), col=brewer.pal(10, "RdBu")), main="Average Annual
Precipitation 1963-2012")
#pre.plot+ layer(sp.polygons(mong))

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
sig.themep <-rasterTheme(region=(brewer.pal(9, "Greens")))
pre.plot <-levelplot(rast.meanpre, margin=FALSE,  par.settings=sig.themep, colorkey=list(at=c(0,
100, 200,300,400,500,600,700,800),labels=list(c('   0', '100', '200','300','400','500','600','700','800'),
cex=1.5), col=brewer.pal(9, "Greens")),
ylab=NULL, xlab=NULL, scales=list(x=x.scale, y=y.scale), main=list(expression(paste("a) Mean Total
Annual Precipitation 1963-2012 (mm)")), x=0.07, cex=1.5, just="left"))
pre.plot+ layer(sp.polygons(mong))

#thot.breaks <- c(-5, 0,5,10,15,20,25)
#heat.theme <-rasterTheme(region=rev(heat.colors(n=15)))
#tmax.plot <-levelplot(rast.meantmax, margin=FALSE, par.settings=heat.theme, xlab=NULL,
               #ylab=list(expression(paste("Latitude (",degree, " North)")), x=0, cex=1.8),
               #scales=list(alternating=2), colorkey=list(labels=list(cex=1.5), at=thot.breaks))
#tmax.plot+ layer(sp.polygons(mong))

#tcool.breaks <- c(-20, -15, -10, -5, 0,5,10)
#cool.theme <-rasterTheme(region=rev(brewer.pal(9, "Blues")))
#tmin.plot <-levelplot(rast.meantmin, margin=FALSE, par.settings=cool.theme, ylab=NULL,
               #xlab=list(expression(paste("Longitude (",degree, " East)")), cex=1.8),
               #scales=list(alternating=2), colorkey=list(labels=list(cex=1.5), at=tcool.breaks))
#tmin.plot+ layer(sp.polygons(mong))

#p.breaks <-c(0, 200, 400, 600, 800, 1000)
#pre.plot <-levelplot(rast.meanpre, margin=FALSE, par.settings=RdBuTheme, ylab=NULL,
              #xlab=NULL,scales=list(alternating=2), colorkey=list(labels=list(cex=1.5),at=p.breaks))
#pre.plot+ layer(sp.polygons(mong))

#example plot with main header text
#pre.plot <-levelplot(rast.meanpre, margin=FALSE, par.settings=RdBuTheme, ylab=NULL,
# xlab=NULL,main=list("a) Average annual total precipitation from 1963 to 2012", x=0.05, cex=1.7,
just="left"),
#scales=list(alternating=2), colorkey=list(labels=list(cex=1.5)))
#  pre.plot+ layer(sp.polygons(mong))

#Trend and Significance Plots Annual Data
#import non-signif pval csv and use locations (1st 2 cols)
nonsig.tmax <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmax_annual_non_significant_pval_locs_ANN.csv")
nonsig.tmax2 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Tmax_annual
_non_significant_pval_locs2.csv")
#note: there are no non-significant pvalue locastions
#nonsig.tmin <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Values_csvs/Tmin_annual
_non_significant_pval_locs.csv")
nonsig.pre <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
```

```
Precip_annual_non_significant_pval_locs_ANN.csv")

trend.pvalxytmax <-nonsig.tmax[,1:2]
coordinates(trend.pvalxytmax) <-~x+y
trend.pvalxytmax2 <-nonsig.tmax2[,1:2]
coordinates(trend.pvalxytmax2) <-~x+y
trend.pvalxytmin <-nonsig.tmin[,1:2]
trend.pvalxypre <-nonsig.pre[,1:2]
coordinates(trend.pvalxypre) <-~x+y

rast.slopetmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmax_ANN.grd")
rast.slopetmax2 <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/Trend_Slope_Tmax2.grd")

rast.slopetmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmin_ANN.grd")
rast.slopepre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Pre_ANN.grd")

#use this method to adjust to trends per decade or century
rast.slopetmax10 <-rast.slopetmax*10
rast.slopetmax102 <-rast.slopetmax2*10
rast.slopetmin10 <-rast.slopetmin*10
rast.slopepre10 <-rast.slopepre*10

#find max and min values each grid for setting legend
rastmaxmin.tmaxslope <-getValues(rast.slopetmax)
rastmaxmin.tminslope <-getValues(rast.slopetmin)
rastmaxmin.preslope <-getValues(rast.slopepre)

#filter values by significance precip- orginally written to look at all signif values in one raster, rather
than an overlay of non-sig points.
#extract by non sig location values
rast.slopeprenon <-extract(rast.slopepre,trend.pvalxypre, method="simple")
rast.slopeprenondf <-as.data.frame(cbind(rast.slopeprenon, rast.slopeprenon))
colnames(rast.slopeprenondf) <-c("slope","non")
rastmaxmin.preslopedf <-as.data.frame(rastmaxmin.preslope)
names(rastmaxmin.preslopedf) <-"slope"
#match to original slopes raster
rast.slopeprej <-join(rastmaxmin.preslopedf, rast.slopeprenondf, type="left", by="slope")
#select non not signif values and get range of slopes
rast.slopeprej$non[!is.na(rast.slopeprej$non)] <- -9999
rast.slopeprejnum <-rast.slopeprej$non
rast.slopeprenotna <-rast.slopepre
setValues(rast.slopeprenotna, rast.slopeprejnum)
#export to asc
writeRaster(rast.slopeprenotna, "Signif_Precip_Annual.asc", format="ascii")

#filter values by significance tmin (all are significant)
#rastmaxmin.tminslope
writeRaster(rast.slopetmin, "Signif_Tmin_Annual.asc", format="ascii")
```

```
#filter values by significance tmax
#extract by non sig location values
rast.slopetmaxnon <-extract(rast.slopetmax,trend.pvalxytmax, method="simple")
rast.slopetmaxnondf <-as.data.frame(cbind(rast.slopetmaxnon, rast.slopetmaxnon))
colnames(rast.slopetmaxnondf) <-c("slope","non")
rastmaxmin.tmaxslopedf <-as.data.frame(rastmaxmin.tmaxslope)
names(rastmaxmin.tmaxslopedf) <-"slope"
#match to original slopes raster
rast.slopetmaxj <-join(rastmaxmin.tmaxslopedf, rast.slopetmaxnondf, type="left", by="slope")
#select non not signif values and get range of slopes and make into raster
rast.slopetmaxj$non[!is.na(rast.slopetmaxj$non)] <- -9999
rast.slopetmaxjnum <-rast.slopetmaxj$non
rast.slopetmaxnotna <-rast.slopetmax
setValues(rast.slopetmaxnotna, rast.slopetmaxjnum)
#export as asc
writeRaster(rast.slopetmaxnotna, "Signif_Tmax_Annual.asc", format="ascii")

#For legend/raster plots
max(rastmaxmin.tmaxslope)
[1] 0.0547619
min(rastmaxmin.tmaxslope)
[1] 0.01287879

max(rastmaxmin.tminslope)
[1] 0.07083333
min(rastmaxmin.tminslope)
[1] 0.01458333

max(rastmaxmin.preslope)
[1]
min(rastmaxmin.preslope)
[1]

#plotting non-significance using x's
#sig.theme <-rasterTheme(region=(brewer.pal(9, "Oranges")))
#tmax.plot <-levelplot(rast.slopetmax, margin=FALSE, par.settings=sig.theme, main="Trends in
Maximum Temperature 1963-2012")
#tmax.plotm <-tmax.plot+ layer(sp.polygons(mong))
#tmax.plotm +layer(sp.points(trend.pvalxytmax, cex=1.2, pch=4, col=1))

#note: no pvals of min temp are insignificant so don't need last line for plot
#sig.theme <-rasterTheme(region=(brewer.pal(9, "Oranges")))
#tmin.plot <-levelplot(rast.slopetmin, margin=FALSE, par.settings=sig.theme, main="Trends in
Minimum Temperature 1963-2012")
#tmin.plotm <-tmin.plot + layer(sp.polygons(mong))
#tmin.plotm + layer(sp.points(trend.pvalxytmin, cex=1.2, pch=4, col=1))

#pre.plot <-levelplot(rast.slopepre, margin=FALSE, par.settings=RdBuTheme, main="Trends in
Annual Precipitation 1963-2012")
#pre.plotm <-pre.plot+ layer(sp.polygons(mong))
#pre.plotm +layer(sp.points(trend.pvalxypre, cex=1.2, pch=4, col=1))

#plotting using adjusted legend/colors
x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
```

```
sig.theme <-rasterTheme(region=(brewer.pal(9, "Oranges")))
tmax.plot <-levelplot(rast.slopetmax10, par.settings=sig.theme, colorkey=list(at=c(0.0,0.1, 0.2, 0.3,
0.4, 0.5, 0.6, 0.7, 0.8),labels=list(c('0.0','0.1', '0.2', '0.3', '0.4', '0.5', '0.6', '0.7','0.8'), cex=1.5),
col=brewer.pal(8, "Oranges")),
scales=list(x=x.scale, y=y.scale),xlab=NULL, ylab=NULL,margin=FALSE,
main=list(expression(paste("b) Trend of Annual Mean Maximum Temperature 1963-2012
(",degree,"C/decade)")), x=0.07, cex=1.5, just="left"))
#tmax.plotm <-tmax.plot+ layer(sp.polygons(mong))
#tmax.plotm <-tmax.plot+ layer(sp.polygons(region2))
#tmax.plotm +layer(sp.points(trend.pvalxytmax, cex=1.2, pch=4, col=1))
#tmax.plotm +layer(sp.points(trend.pvalxytmax, cex=1, pch=15, col="grey50"))
tmax.plotm <-tmax.plot+ layer(sp.points(trend.pvalxytmax, cex=1, pch=15, col="grey50"))
tmax.plotm +layer(sp.polygons(region2))

#note: no pvals of min temp are insignificant so don't need last line for plot
x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
tmin.plot <-levelplot(rast.slopetmin10, par.settings=sig.theme, colorkey=list(at=c(0.0,0.1, 0.2, 0.3,
0.4, 0.5, 0.6, 0.7, 0.8),labels=list(c('0.0','0.1', '0.2', '0.3', '0.4', '0.5', '0.6', '0.7','0.8'), cex=1.5),
col=brewer.pal(8, "Oranges")),
scales=list(x=x.scale, y=y.scale),xlab=NULL, ylab=NULL,margin=FALSE,
main=list(expression(paste("c) Trend of Annual Mean Minimum Temperature 1963-2012
(",degree,"C/decade)")), x=0.07, cex=1.5, just="left"))
#tmin.plotm <-tmin.plot + layer(sp.polygons(mong))
#tmin.plotm + layer(sp.points(trend.pvalxytmin, cex=1.2, pch=4, col=1))
#tmin.plotm <-tmin.plot+ layer(sp.polygons(region2))
#tmin.plotm +layer(sp.points(trend.pvalxytmin, cex=1, pch=15, col="grey50"))
#tmin.plotm <-tmin.plot+ layer(sp.points(trend.pvalxytmin, cex=1, pch=15, col="grey50"))
tmin.plot +layer(sp.polygons(region2))

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
sig.themep <-rasterTheme(region=(brewer.pal(11, "RdBu")))
pre.plot <-levelplot(rast.slopepre10, margin=FALSE, par.settings=sig.themep, colorkey=list(at=c(-25,
-20, -15, -10, -5, 0, 5, 10,  15, 20, 25),labels=list(c('-25', '-20', '-15', '-10', ' -5', '  0', '  5', ' 10',  ' 15', ' 20',
' 25'), cex=1.5), col=brewer.pal(11, "RdBu")),
scales=list(x=x.scale, y=y.scale),xlab=NULL, ylab=NULL, main=list(expression(paste("a) Trend of
Annual Total Precipitation 1963-2012 (mm/decade)")), x=0.07, cex=1.5, just="left"))
#pre.plotm <-pre.plot+ layer(sp.polygons(mong))
#pre.plotm +layer(sp.points(trend.pvalxypre, cex=1.2, pch=4, col=1))
#pre.plotm <-pre.plot+ layer(sp.polygons(region2))
#pre.plotm +layer(sp.points(trend.pvalxypre, cex=1, pch=15, col="grey50"))
pre.plotm <-pre.plot+layer(sp.points(trend.pvalxypre, cex=1, pch=15, col="grey50"))
pre.plotm +layer(sp.polygons(region2))

#Plot Seasonal trends
#check titles before plotting!
#read rasters for plotting
#rast.slopetmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmax_DJF.grd")
#rast.slopetmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmax_MAM.grd")
#rast.slopetmax <-
```

```
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmax_JJA.grd")
rast.slopetmax <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmax_SON.grd")

#rast.slopetmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmin_DJF.grd")
#rast.slopetmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmin_MAM.grd")
#rast.slopetmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmin_JJA.grd")
rast.slopetmin <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Tmin_SON.grd")

#rast.slopepre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Pre_DJF.grd")
#rast.slopepre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Pre_MAM.grd")
#rast.slopepre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Pre_JJA.grd")
rast.slopepre <-
raster("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Gridfiles/Trend
_Slope_Pre_SON.grd")

#multiply by 10 for decadal trends
rast.slopetmax10=rast.slopetmax*10
rast.slopetmin10=rast.slopetmin*10
rast.slopepre10=rast.slopepre*10

#find max and min values each grid for setting legend
#rastmaxmin.tmaxslope <-getValues(rast.slopetmax)
#rastmaxmin.tminslope <-getValues(rast.slopetmin)
#rastmaxmin.preslope <-getValues(rast.slopepre)

#For legend/raster plots
#DJF
max(rastmaxmin.tmaxslope)
[1] 0.08148148
min(rastmaxmin.tmaxslope)
[1] -0.025
max(rastmaxmin.tminslope)
[1] 0.1074074
min(rastmaxmin.tminslope)
[1] -0.005982906
max(rastmaxmin.preslope)
[1] 0.65625
```

```
min(rastmaxmin.preslope)
[1] -0.3235294

#MAM
max(rastmaxmin.tmaxslope)
[1] 0.06875
min(rastmaxmin.tmaxslope)
[1] 0.008333333
max(rastmaxmin.tminslope)
[1] 0.08137255
min(rastmaxmin.tminslope)
[1] 0.02037037
max(rastmaxmin.preslope)
[1] 0.9588235
min(rastmaxmin.preslope)
[1] -0.3

#JJA
max(rastmaxmin.tmaxslope)
[1] 0.0462963
min(rastmaxmin.tmaxslope)
[1] -0.003174603
max(rastmaxmin.tminslope)
[1] 0.05632184
min(rastmaxmin.tminslope)
[1] -0.003571429
max(rastmaxmin.preslope)
[1] 1.035294
min(rastmaxmin.preslope)
[1] -2.245455

#SON
max(rastmaxmin.tmaxslope)
[1] 0.05833333
min(rastmaxmin.tmaxslope)
[1] 0.007017544
max(rastmaxmin.tminslope)
[1] 0.08039216
min(rastmaxmin.tminslope)
[1] 0.003030303
max(rastmaxmin.preslope)
[1] 0.6974359
min(rastmaxmin.preslope)
[1] -0.88

#import non-signif pval csv and use locations (1st 2 cols)
#nonsig.tmax <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmax_annual_non_significant_pval_locs_djf.csv")
#nonsig.tmax <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmax_annual_non_significant_pval_locs_mam.csv")
#nonsig.tmax <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmax_annual_non_significant_pval_locs_jja.csv")
```

```
nonsig.tmax <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmax_annual_non_significant_pval_locs_son.csv")

#nonsig.tmin <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmin_annual_non_significant_pval_locs_djf.csv")
#nonsig.tmin <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmin_annual_non_significant_pval_locs_mam.csv")
#nonsig.tmin <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmin_annual_non_significant_pval_locs_jja.csv")
nonsig.tmin <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Tmin_annual_non_significant_pval_locs_son.csv")

#nonsig.pre <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Precip_annual_non_significant_pval_locs_djf.csv")
#nonsig.pre <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Precip_annual_non_significant_pval_locs_mam.csv")
#nonsig.pre <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Precip_annual_non_significant_pval_locs_jja.csv")
nonsig.pre <-
read.csv("/Users/niah/Documents/CSU/Mongolia/MN_2015/Spatial_MN/1963_2012/Values_csvs/
Precip_annual_non_significant_pval_locs_son.csv")

trend.pvalxytmax <-nonsig.tmax[,1:2]
coordinates(trend.pvalxytmax) <-~x+y
trend.pvalxytmin <-nonsig.tmin[,1:2]
coordinates(trend.pvalxytmin) <-~x+y
trend.pvalxypre <-nonsig.pre[,1:2]
coordinates(trend.pvalxypre) <-~x+y

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
sig.theme <-rasterTheme(region=(brewer.pal(9, "Oranges")))
tmax.plot <-levelplot(rast.slopetmax10, par.settings=sig.theme, colorkey=list(at=c(-0.50, -0.25, 0.0,
0.25, 0.50, 0.75, 1.0, 1.25, 1.50),labels=list(c('-0.50', '-0.25', '0.0', '0.25', '0.50', '0.75', '1.0', '1.25',
'1.50'), cex=1.5), col=brewer.pal(8, "Oranges")),
scales=list(x=x.scale, y=y.scale),xlab=NULL, ylab=NULL,margin=FALSE,
main=list(expression(paste("b) Trend of Fall Mean Maximum Temperature 1963-2012
(",degree,"C/decade)")), x=0.07, cex=1.5, just="left"))
tmax.plotm <-tmax.plot+ layer(sp.polygons(mong))
tmax.plotm +layer(sp.points(trend.pvalxytmax, cex=1.2, pch=4, col=1))

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
tmin.plot <-levelplot(rast.slopetmin10, par.settings=sig.theme, colorkey=list(at=c(-0.50, -0.25, 0.0,
0.25, 0.50, 0.75, 1.0, 1.25, 1.50),labels=list(c('-0.50', '-0.25', '0.0', '0.25', '0.50', '0.75', '1.0', '1.25',
'1.50'), cex=1.5), col=brewer.pal(8, "Oranges")),
scales=list(x=x.scale, y=y.scale),xlab=NULL, ylab=NULL,margin=FALSE,
```

```
main=list(expression(paste("c) Trend of Fall Mean Minimum Temperature 1963-2012
(",degree,"C/decade)")), x=0.07, cex=1.5, just="left"))
tmin.plotm <-tmin.plot + layer(sp.polygons(mong))
tmin.plotm + layer(sp.points(trend.pvalxytmin, cex=1.2, pch=4, col=1))

x.scale <- list(cex=1.5, alternating=1)
y.scale <- list(cex=1.5, alternating=1)
sig.themep <-rasterTheme(region=(brewer.pal(11, "RdBu")))
pre.plot <-levelplot(rast.slopepre10, margin=FALSE, par.settings=sig.themep, colorkey=list(at=c(-25,
-20, -15, -10, -5, 0, 5, 10,  15, 20, 25),labels=list(c('-25', '-20', '-15', '-10', ' -5', '  0', '  5', ' 10',  ' 15', ' 20',
' 25'), cex=1.5), col=brewer.pal(11, "RdBu")),
scales=list(x=x.scale, y=y.scale),xlab=NULL, ylab=NULL, main=list(expression(paste("a) Trend of Fall
Total Precipitation 1963-2012 (mm/decade)")), x=0.07, cex=1.5, just="left"))
pre.plotm <-pre.plot+ layer(sp.polygons(mong))
pre.plotm +layer(sp.points(trend.pvalxypre, cex=1.2, pch=4, col=1))

#With aimag boundary added
#pre.plot <-levelplot(rast.slopepre, margin=FALSE, par.settings=RdBuTheme, main="Trends in
Summer Precipitation 1963-2012")
pre.plotm <-pre.plot+ layer(sp.polygons(mong))
pre.plotnon <-pre.plotm +layer(sp.points(trend.pvalxypre, cex=1.2, pch=4, col=1))
pre.plotnon +layer(sp.polygons(three))
```

REFERENCES

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www.R-project.org/*.

APPENDIX B

B.1 Chapter 3 Code

The MAKESENS Microsoft Excel spreadsheet template from the Finnish Meteorological Institute was used for the Chapter 3 trend analyses (Salmi *et al.*, 2002). The spreadsheet is available from:

www.ilmanlaatu.fi/ilmansaasteet/julkaisu/pdf/MAKESENS_1_0.xls

REFERENCES

Salmi T, Maatta A, Anttila P, Ruoho-Airola T, and Amnell T, 2002. Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's Slope estimates- The Excel template application MAKESENS. Publications on Air Quality, No. 31, Finnish Meteorological Institute, Helsinki, Finland, 35 pp., *http://www.ilmanlaatu.fi/ilmansaasteet/julkaisu/pdf/MAKESENS-Manual_2002.pdf*.

APPENDIX C

C.1 Chapter 4 R Code

The following code was used for creating the climate maps for Chapter 4 and works with R software, version 3.1.3 (2015-03-09), "Smooth Sidewalk" and previous versions (R Team, 2015). While any errors or omissions are mine, please be aware that no warranty of any kind is offered with this code and users need to carefully proof and/or modify the code to suit their needs.

C.1.1 Creating Regular Time series and Calculating Percentage of Missing Values

```
#------------------------------------------------------------------------
#  TITLE: Creating Regular Time series from Irregular Data and Calculating Percent Data Missing
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2014-03-03
#  LAST REVISION: 2014-12-08
#  DESCRIPTION: This script provides code for creating regular time series from irregular series
#  PACKAGES REQUIRED: chron
#  VARIABLES/DATA USED:
#        NAME:
# TYPE:
# COMMENT:
#------------------------------------------------------------------------

#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Regular_TS_R/")

#libraries
library(chron)
library(plyr)
library(doBy)
library(reshape2)

#--------------------------------------------------------------------------------------------
#import precipitation and streamflow data

pvar <-read.csv("Khangai_Daily_P.csv")
qvar <-read.csv("Khangai_Daily_Q.csv")

#make precip data in wide format for easier formatting
day.precipw <-dcast(pvar, yrmonday~name, value.var="dayP")
pvar2 <-day.precipw

qvar.date <-paste(qvar$Day,qvar$Month,qvar$Year, sep="" )
```

```
qvar.dateyrmonday <-as.Date(qvar.date, "%d%b%Y")

qvar$yrmonday <-qvar.dateyrmonday
qvar2 <-data.frame(qvar[,11], qvar[, 5:10])
colnames(qvar2) <-c("yrmonday", "Baidrag_Baidrag", "ikhkturd_Baidrag", "Bayankhongor_Tuin",
"Bogd_Tuin", "Erdenemandal_Khanui", "Ikhtamir_Khoid_Tamir")

#making regular time series from entire irregular series then matching up and parsing out by station
#create regular ts using chron
var.regdayp <- seq.dates("01/01/1950", "12/31/2012", by = "days")
var.regsp <-as.data.frame(as.Date(var.regdayp))
colnames(var.regsp) <-"yrmonday"

var.regdayq <- seq.dates("01/01/1971", "12/31/2010", by = "days")
var.regsq <-as.data.frame(as.Date(var.regdayq))
colnames(var.regsq) <-"yrmonday"

#join datasets and sort
pvar.j <-join(pvar2, var.regsp, by="yrmonday", type="full")
pvar.js <-pvar.j[order(pvar.j$yrmonday),]

qvar.j <-join(qvar2, var.regsq, by="yrmonday", type="right")

#divide to stations and truncate to record starts
baidb <-qvar.j[5115:14610,1:2]
ikhkt <-qvar.j[1827:14610,c(1, 3)]
bayant <-qvar.j[1827:14610, c(1,4)]
bogdt <-qvar.j[, c(1,5)]
erdk <-qvar.j[1827:14610, c(1,6)]
ikhkt <-qvar.j[1827:14610, c(1,7)]

baid <-pvar.js[15707:23011, 1:2]
bayan <-pvar.js[4749:23011,c(1, 3)]
erd <-pvar.js[5114:23011, c(1,4)]
gal <-pvar.js[2192:23011, c(1,5)]
khor <-pvar.js[16072:23011, c(1,6)]
tset <-pvar.js[, c(1,7)]

#for precip sum missing by month
#Baidrag
#add month column for sum
baid$yrmonth <-substr(baid$yrmonday, 1,7)
baidmiss <-ddply(baid, .(yrmonth), summarize, baidmon=sum(is.na(Baidrag)))

#if more than 23 days missing then month is NA
#summarize all months
baidmon <-ddply(baid, .(yrmonth), summarize, baidmon=sum(Baidrag, na.rm=TRUE))
baidmissna <-ifelse(baidmiss$baidmon>=23, NA, baidmon$baidmon)
baidmissmon <-data.frame(baidmon[,1], baidmissna)
colnames(baidmissmon) <-c("yrmon","monthp")

#summarize to annual, if less than 6 months data per year is na, then year is na
baidmissmon$yr <-substr(baidmissmon$yrmon, 1,4)
baidmissann <-ddply(baidmissmon,.(yr), summarize, Baidp=sum(monthp, na.rm=TRUE))
baidmissyr <-ddply(baidmissmon, .(yr), summarize, baidmon=sum(is.na(monthp)))
```

```
baidmissnann <-ifelse(baidmissyr$baidmon>6, NA, baidmissann$Baidp)
baidmissingyr <-data.frame(baidmissyr[,1], baidmissnann)
colnames(baidmissingyr) <-c("yr","Baidp")
#Bayankhongor
bayan$yrmonth <-substr(bayan$yrmonday, 1,7)
bayanmiss <-ddply(bayan, .(yrmonth), summarize, bayanmon=sum(is.na(Bayankhongor)))

#if more than 23 days missing then month is NA
#summarize all months
bayanmon <-ddply(bayan, .(yrmonth), summarize, bayanmon=sum(Bayankhongor, na.rm=TRUE))
bayanmissna <-ifelse(bayanmiss$bayanmon>=23, NA, bayanmon$bayanmon)
bayanmissmon <-data.frame(bayanmon[,1], bayanmissna)
colnames(bayanmissmon) <-c("yrmon","monthp")

#summarize to annual, if less than 6 months data per year is na, then year is na
bayanmissmon$yr <-substr(bayanmissmon$yrmon, 1,4)
bayanmissann <-ddply(bayanmissmon,.(yr), summarize, bayanp=sum(monthp, na.rm=TRUE))
bayanmissyr <-ddply(bayanmissmon, .(yr), summarize, bayanmon=sum(is.na(monthp)))
bayanmissnann <-ifelse(bayanmissyr$bayanmon>6, NA, bayanmissann$bayanp)
bayanmissingyr <-data.frame(bayanmissyr[,1], bayanmissnann)
colnames(bayanmissingyr) <-c("yr","Bayanp")

#Erdenemandal
#add month column for sum
erd$yrmonth <-substr(erd$yrmonday, 1,7)
erdmiss <-ddply(erd, .(yrmonth), summarize, erdmon=sum(is.na(Erdenemandal)))

#if more than 23 days missing then month is NA
#summarize all months
erdmon <-ddply(erd, .(yrmonth), summarize, erdmon=sum(Erdenemandal, na.rm=TRUE))
erdmissna <-ifelse(erdmiss$erdmon>=23, NA, erdmon$erdmon)
erdmissmon <-data.frame(erdmon[,1], erdmissna)
colnames(erdmissmon) <-c("yrmon","monthp")

#summarize to annual, if less than 6 months data per year is na, then year is na
erdmissmon$yr <-substr(erdmissmon$yrmon, 1,4)
erdmissann <-ddply(erdmissmon,.(yr), summarize, erdp=sum(monthp, na.rm=TRUE))
erdmissyr <-ddply(erdmissmon, .(yr), summarize, erdmon=sum(is.na(monthp)))
erdmissnann <-ifelse(erdmissyr$erdmon>6, NA, erdmissann$erdp)
erdmissingyr <-data.frame(erdmissyr[,1], erdmissnann)
colnames(erdmissingyr) <-c("yr","Erdp")

#Galuut
#add month column for sum
gal$yrmonth <-substr(gal$yrmonday, 1,7)
galmiss <-ddply(gal, .(yrmonth), summarize, galmon=sum(is.na(Galuut)))

#if more than 23 days missing then month is NA
#summarize all months
galmon <-ddply(gal, .(yrmonth), summarize, galmon=sum(Galuut, na.rm=TRUE))
galmissna <-ifelse(galmiss$galmon>=23, NA, galmon$galmon)
galmissmon <-data.frame(galmon[,1], galmissna)
colnames(galmissmon) <-c("yrmon","monthp")

#summarize to annual, if less than 6 months data per year is na, then year is na
```

```
galmissmon$yr <-substr(galmissmon$yrmon, 1,4)
galmissann <-ddply(galmissmon,.(yr), summarize, galp=sum(monthp, na.rm=TRUE))
galmissyr <-ddply(galmissmon, .(yr), summarize, galmon=sum(is.na(monthp)))
galmissnann <-ifelse(galmissyr$galmon>6, NA, galmissann$galp)
galmissingyr <-data.frame(galmissyr[,1], galmissnann)
colnames(galmissingyr) <-c("yr","Galp")

#Khoriult
#add month column for sum
khor$yrmonth <-substr(khor$yrmonday, 1,7)
khormiss <-ddply(khor, .(yrmonth), summarize, khormon=sum(is.na(Khoriult)))

#if more than 23 days missing then month is NA
#summarize all months
khormon <-ddply(khor, .(yrmonth), summarize, khormon=sum(Khoriult, na.rm=TRUE))
khormissna <-ifelse(khormiss$khormon>=23, NA, khormon$khormon)
khormissmon <-data.frame(khormon[,1], khormissna)
colnames(khormissmon) <-c("yrmon","monthp")

#summarize to annual, if less than 6 months data per year is na, then year is na
khormissmon$yr <-substr(khormissmon$yrmon, 1,4)
khormissann <-ddply(khormissmon,.(yr), summarize, khorp=sum(monthp, na.rm=TRUE))
khormissyr <-ddply(khormissmon, .(yr), summarize, khormon=sum(is.na(monthp)))
khormissnann <-ifelse(khormissyr$khormon>6, NA, khormissann$khorp)
khormissingyr <-data.frame(khormissyr[,1], khormissnann)
colnames(khormissingyr) <-c("yr","khorp")

#Tsetserleg
#add month column for sum
tset$yrmonth <-substr(tset$yrmonday, 1,7)
tsetmiss <-ddply(tset, .(yrmonth), summarize, tsetmon=sum(is.na(Tsetserleg)))

#if more than 23 days missing then month is NA
#summarize all months
tsetmon <-ddply(tset, .(yrmonth), summarize, tsetmon=sum(Tsetserleg, na.rm=TRUE))
tsetmissna <-ifelse(tsetmiss$tsetmon>=23, NA, tsetmon$tsetmon)
tsetmissmon <-data.frame(tsetmon[,1], tsetmissna)
colnames(tsetmissmon) <-c("yrmon","monthp")

#summarize to annual, if less than 6 months data per year is na, then year is na
tsetmissmon$yr <-substr(tsetmissmon$yrmon, 1,4)
tsetmissann <-ddply(tsetmissmon,.(yr), summarize, tsetp=sum(monthp, na.rm=TRUE))
tsetmissyr <-ddply(tsetmissmon, .(yr), summarize, tsetmon=sum(is.na(monthp)))
tsetmissnann <-ifelse(tsetmissyr$tsetmon>6, NA, tsetmissann$tsetp)
tsetmissingyr <-data.frame(tsetmissyr[,1], tsetmissnann)
colnames(tsetmissingyr) <-c("yr","Tsetp")

#collate results
#no precipitation stations have a 12 month period without missing data
#collate if 6 months or more of data
#add name column each
baidname <-rep("Baidrag", length(baidmissingyr[,1]))
baidmissingyr$Name <-baidname
colnames(baidmissingyr) <-c("Year", "AnnP", "Name")
```

```
bayanname <-rep("Bayankhongor", length(bayanmissingyr[,1]))
bayanmissingyr$Name <-bayanname
colnames(bayanmissingyr) <-c("Year", "AnnP", "Name")
erdname <-rep("Erdenemandal", length(erdmissingyr[,1]))
erdmissingyr$Name <-erdname
colnames(erdmissingyr) <-c("Year", "AnnP", "Name")

galname <-rep("Galuut", length(galmissingyr[,1]))
galmissingyr$Name <-galname
colnames(galmissingyr) <-c("Year", "AnnP", "Name")

khorname <-rep("Khoriult", length(khormissingyr[,1]))
khormissingyr$Name <-khorname
colnames(khormissingyr) <-c("Year", "AnnP", "Name")

tsetname <-rep("Tsetserleg", length(tsetmissingyr[,1]))
tsetmissingyr$Name <-tsetname
colnames(tsetmissingyr) <-c("Year", "AnnP", "Name")

#stack
annp.mnmiss <-rbind(khormissingyr,bayanmissingyr,galmissingyr, baidmissingyr, tsetmissingyr,
erdmissingyr)
write.csv(annp.mnmiss, "Khangai_Annual_P_Miss.csv")

#for each station count na values
#calculate percent missing as total record/ missing days

baidb.na <-sum(is.na(baidb))
baidb.p <-baidb.na/length(baidb[,2])
bayanb.na <-sum(is.na(bayanb))
bayanb.p <-bayanb.na/length(bayanb[,2])
bayant.na <-sum(is.na(bayant))
bayant.p <-bayant.na/length(bayant[,2])
bogdt.na <-sum(is.na(bogdt))
bogdt.p <-bogdt.na/length(bogdt[,2])
erdk.na <-sum(is.na(erdk))
erdk.p <-erdk.na/length(erdk[,2])
ikhkt.na <-sum(is.na(ikhkt))
ikhkt.p <-ikhkt.na/length(ikhkt[,2])

baid.na <-sum(is.na(baid))
baid.p <-baid.na/length(baid[,2])
bayan.na <-sum(is.na(bayan))
bayan.p <-bayan.na/length(bayan[,2])
erd.na <-sum(is.na(erd))
erd.p <-erd.na/length(erd[,2])
gal.na <-sum(is.na(gal))
gal.p <-gal.na/length(gal[,2])
khor.na <-sum(is.na(khor))
khor.p <-khor.na/length(khor[,2])
tset.na <-sum(is.na(tset))
tset.p <-tset.na/length(tset[,2])
```

C.1.2 Trend Analyses for Khangai Mountain Region Hydroclimate Data

```
#-----------------------------------------------------------------------
#  TITLE: Trend Analysis Script for LOR PQ
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2014-11-04
#  LAST REVISION: 2015-05-12
#  DESCRIPTION: This script provides code for analyzing Khangai met and streamflow data for trends
#  PACKAGES REQUIRED: plyr, reshape2, ggplot2
#  VARIABLES/DATA USED: text files of Khangai precipitation and streamflow
#  NAME:
# TYPE:
# COMMENT:
#-----------------------------------------------------------------------------
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/")

library(plyr)
library(reshape2)
library(ggplot2)
library(zyp)
library(Kendall)
library(mblm)

#note: original data from IMHE assumes that zero values are not NA.
# NA columns added for same period of record for all stations

#Extracting met data from excel files for Khangai stations (from proposal code)
##For txt import
varfilep <- "Ark_Bay_Met_Precip.txt"

precip.t <- read.table(varfilep, colClasses=c(rep("character", 5), "numeric"), header=TRUE,
blank.lines.skip=TRUE, na.strings="NA")

#remove any blank lines
precip.t <-precip.t[complete.cases(precip.t), ]

#create new categories for aggregating
#yearmon column
precip.t$yrmon <- paste(precip.t$year, precip.t$month, sep='-')

#yearmonday column
precip.t$yrmonday <- paste(precip.t$year, precip.t$month, precip.t$day, sep='-')

#monthday column
precip.t$monday <- paste(precip.t$month, precip.t$day, sep='-')

#replace -999 with NA
#find which are -999 use tmaxna <-tmax.t[ which(tmax.t$tmax =='-999'),]
precip.t$precip[precip.t$precip=='-999'] <- NA

#aggregate to daily precip
day.precip <- ddply(precip.t, .(ID, yrmonday), summarize, dayP=sum(precip))
#add monday column to daily precip
```

```
day.precip$monday <-paste(substr(day.precip$yrmonday, 6,10))

#check max and min values
max(day.precip$dayP, na.rm=TRUE)
min(day.precip$dayP, na.rm=TRUE)

#check values for extreme outliers (see Harris et al 2013, sec 3.1, precip within 4 SD)
precip.m <- ddply(day.precip, .(ID, monday), summarize, Mean=mean(dayP, na.rm=TRUE))
precip.sd <- ddply(day.precip, .(ID, monday), summarize, StdDev=sd(dayP, na.rm=TRUE))
precip.sdcol <- cbind(precip.sd[ , 3], rep(4,length(precip.sd)))
precip.sdc <- precip.sdcol[ , 1]*precip.sdcol[,2]
precip.sdfcol <-cbind(precip.m, precip.sdc)
precip.sdfcols <-precip.sdfcol[ , 3]+precip.sdfcol[, 4]
precip.sdf <-cbind(precip.m, precip.sdfcols)
names(precip.sdf)[names(precip.sdf)=="precip.sdfcols"] <- "StdDev4"

#merge in 4sd column to precip data for comparison
precip.sdm <- join(precip.sdf,day.precip, by=c("ID", "monday"))

#find which timesteps exceed 4 sd
precip.sdex <- subset(precip.sdm, subset = precip.sdm$dayP > precip.sdm$StdDev4)
#results in 58 values (out of 52,319) outside this range, but none seem erroneous

#join in station names using ID
metnames <- read.csv("Met_Names.csv", colClasses="character")
day.precipn <- join(day.precip, metnames, by="ID")

#export daily data truncated to stations of interest to csv
day.precipaoi <-subset(day.precipn, c(name=="Tsetserleg" | name=="Erdenemandal" |
name=="Baidrag" | name=="Galuut" | name=="Bayankhongor" | name=="Khoriult"))
#write.csv(day.precipaoi, file= "Khangai_Daily_P.csv")

#counting values for QC (must count after aggregated to daily data to get days/month)
#note that for precip, missing days are considered to be "0" precipitation, not NA, see Bayankhongor
data from IMHE with zero fills.
precip.col <-substr(day.precipaoi$yrmonday, 1, 7)
day.precipcol <-cbind(day.precipaoi, precip.col)
names(day.precipcol)[names(day.precipcol)=="precip.col"] <- "yrmon"
precip.ct <- count(day.precipcol, c("ID","yrmon"))

# not using count for QC due to data not recorded is assumed to be 0 not NA

#aggregate to monthly precip
mon.precip <- ddply(precip.t, .(ID, yrmon), summarize, monP=sum(precip))

#join in station names using ID
mon.precipn <- join(mon.precip, metnames, by="ID")

#subset to stations of interest and export monthly data to csv
mon.precipaoi <-subset(mon.precipn, c(name=="Tsetserleg" | name=="Erdenemandal" |
name=="Baidrag" | name=="Galuut" | name=="Bayankhongor" | name=="Khoriult"))
#write.csv(mon.precipaoi, file= "Khangai_Monthly_P.csv")


#seasonal precip
```

```
#add mon column
monaoi <-substr(mon.precipaoi$yrmon, 6, 7)
mon.precipaoi$month <-monaoi

#add season column
mon.precipaoi$season[mon.precipaoi$month=="12"|mon.precipaoi$month=="01"|mon.precipaoi$m
onth=="02"] <-"winter"
mon.precipaoi$season[mon.precipaoi$month=="03"|mon.precipaoi$month=="04"|mon.precipaoi$m
onth=="05"] <-"spring"
mon.precipaoi$season[mon.precipaoi$month=="06"|mon.precipaoi$month=="07"|mon.precipaoi$m
onth=="08"] <-"summer"
mon.precipaoi$season[mon.precipaoi$month=="09"|mon.precipaoi$month=="10"|mon.precipaoi$m
onth=="11"] <-"fall"

#add year column
yraoi <-as.numeric(substr(mon.precipaoi$yrmon, 1, 4))
mon.precipaoi$year <-yraoi
mon.precipaoi$yrwin<-yraoi+1

mon.precipaoi$yrseas <-ifelse(mon.precipaoi$month=="12", mon.precipaoi$yrwin,
mon.precipaoi$year)

#aggregate to seasonal values
seas.precipaoi <- ddply(mon.precipaoi, .(ID, name, yrseas, season), summarize, seasP=sum(monP,
na.rm=TRUE))

#no first winter values to make NA as theere are no vals for precious december of first year
write.csv(seas.precipaoi, file= "Khangai_Seasonal_P.csv")

#annual precipitation

#aggregate to annual values
yr.precipaoi <- ddply(mon.precipaoi, .(ID, name, year), summarize, yrP=sum(monP))
#write.csv(yr.precipaoi, file= "Khangai_Annual_P.csv")

#-----------------------------------------------------------------------------------
#streamflow data
#import data
baidrag <-read.csv("baidrag_Q.csv", header=TRUE, na.strings= "NA", colClasses= "numeric")
bayanb <-read.csv("bayanburd_Q.csv", header=TRUE, na.strings= "NA", colClasses= "numeric")
bayank <-read.csv("bayank_Q.csv", header=TRUE, na.strings= "NA", colClasses= "numeric")
bogd <-read.csv("bogd_Q.csv", header=TRUE, na.strings= "NA", colClasses= "numeric")
khanui <-read.csv("khanui_Q.csv", header=TRUE, na.strings= "NA", colClasses= "numeric")
ktamir <-read.csv("ktamir_Q.csv", header=TRUE, na.strings= "NA", colClasses= "numeric")

#add date columns
months <- c(rep("Jan",31), rep("Feb",29), rep("Mar",31), rep("Apr", 30),
rep("May",31),rep("Jun",30),rep("Jul",31),rep("Aug",31), rep("Sep", 30), rep("Oct",31),
rep("Nov",30), rep("Dec", 31))
month <- factor(months, levels=c("Jan","Feb","Mar", "Apr","May","Jun","Jul","Aug", "Sep", "Oct",
"Nov", "Dec"))

days <- c(seq(1,31, 1), seq(1,29, 1), seq(1,31,1), seq(1, 30,1),
seq(1,31,1),seq(1,30,1),seq(1,31,1),seq(1,31,1), seq(1, 30,1), seq(1,31,1), seq(1,30,1), seq(1, 31,1))
day <- formatC(days, width=2 ,format="d", flag="0")
```

```
years <- seq(1971,2010, 1)

#add name identifier columns
baidragn <- rep("Baidrag_Baidrag", 366)
bayanbn <- rep("Bayanburd_Baidrag", 366)
bayankn <- rep("Bayankhongor_Tuin", 366)
bogdn <- rep("Bogd_Tuin", 366)
khanuin <- rep("Erdenemandal_Khanui", 366)
ktamirn <- rep("Ikhtamir_Khoid_Tamir", 366)

#bind dates, names, and data
baidrag.d <-data.frame(month, day, baidragn, baidrag)
colnames(baidrag.d) <-c("Month", "Day", "Name", years)

bayanb.d <-data.frame(month, day, bayanbn, bayanb)
colnames(bayanb.d) <-c("Month", "Day", "Name", years)

bayank.d <-data.frame(month, day, bayankn, bayank)
colnames(bayank.d) <-c("Month", "Day", "Name", years)

bogd.d <-data.frame(month, day, bogdn, bogd)
colnames(bogd.d) <-c("Month", "Day", "Name", years)

khanui.d <-data.frame(month, day, khanuin, khanui)
colnames(khanui.d) <-c("Month", "Day", "Name", years)

ktamir.d <-data.frame(month, day, ktamirn, ktamir)
colnames(ktamir.d) <-c("Month", "Day", "Name", years)

#reshape frame to y-m-d format
baidrag.m <- melt(baidrag.d, id.vars=c("Month", "Day"), na.rm=FALSE,
value.name="Baidrag_Baidrag", variable.name="Year", measure.vars= c("1971", "1972", "1973",
"1974", "1975", "1976", "1977", "1978", "1979", "1980", "1981", "1982", "1983", "1984", "1985",
"1986", "1987", "1988", "1989", "1990", "1991", "1992", "1993", "1994", "1995", "1996", "1997",
"1998", "1999", "2000", "2001", "2002", "2003" , "2004",  "2005", "2006", "2007", "2008", "2009",
"2010"))
bayanb.m <- melt(bayanb.d, id.vars=c("Month", "Day"), na.rm=FALSE,
value.name="Bayanburd_Baidrag", variable.name="Year", measure.vars= c("1971", "1972", "1973",
"1974", "1975", "1976", "1977", "1978", "1979", "1980", "1981", "1982", "1983", "1984", "1985",
"1986", "1987", "1988", "1989", "1990", "1991", "1992", "1993", "1994", "1995", "1996", "1997",
"1998", "1999", "2000", "2001", "2002", "2003" , "2004",  "2005", "2006", "2007", "2008", "2009",
"2010"))
bayank.m <- melt(bayank.d, id.vars=c("Month", "Day"), na.rm=FALSE,
value.name="Bayankhongor_Tuin", variable.name="Year", measure.vars= c("1971", "1972", "1973",
"1974", "1975", "1976", "1977", "1978", "1979", "1980", "1981", "1982", "1983", "1984", "1985",
"1986", "1987", "1988", "1989", "1990", "1991", "1992", "1993", "1994", "1995", "1996", "1997",
"1998", "1999", "2000", "2001", "2002", "2003" , "2004",  "2005", "2006", "2007", "2008", "2009",
"2010"))
bogd.m <- melt(bogd.d, id.vars=c("Month", "Day"), na.rm=FALSE, value.name="Bogd_Tuin",
variable.name="Year", measure.vars= c("1971", "1972", "1973", "1974", "1975", "1976", "1977",
"1978", "1979", "1980", "1981", "1982", "1983", "1984", "1985", "1986", "1987", "1988", "1989",
"1990", "1991", "1992", "1993", "1994", "1995", "1996", "1997", "1998", "1999", "2000", "2001",
"2002", "2003" , "2004",  "2005", "2006", "2007", "2008", "2009", "2010"))
khanui.m <- melt(khanui.d, id.vars=c("Month", "Day"), na.rm=FALSE,
value.name="Erdenemandal_Khanui", variable.name="Year", measure.vars= c("1971", "1972",
```

```
"1973", "1974", "1975", "1976", "1977", "1978", "1979", "1980", "1981", "1982", "1983", "1984",
"1985", "1986", "1987", "1988", "1989", "1990", "1991", "1992", "1993", "1994", "1995", "1996",
"1997", "1998", "1999", "2000", "2001", "2002", "2003" , "2004",  "2005", "2006", "2007", "2008",
"2009", "2010"))
ktamir.m <- melt(ktamir.d, id.vars=c("Month", "Day"), na.rm=FALSE,
value.name="Ikhtamir_Khoid_Tamir", variable.name="Year", measure.vars= c("1971", "1972",
"1973", "1974", "1975", "1976", "1977", "1978", "1979", "1980", "1981", "1982", "1983", "1984",
"1985", "1986", "1987", "1988", "1989", "1990", "1991", "1992", "1993", "1994", "1995", "1996",
"1997", "1998", "1999", "2000", "2001", "2002", "2003" , "2004",  "2005", "2006", "2007", "2008",
"2009", "2010"))

#bind all locations together
gages <-data.frame(baidrag.m, bayanb.m[,4], bayank.m[,4], bogd.m[,4], khanui.m[,4], ktamir.m[,4])
names(gages) <-c("Month", "Day", "Year", "Baidrag_Baidrag",
"Bayanburd_Baidrag","Bayankhongor_Tuin","Bogd_Tuin","Erdenemandal_Khanui","Ikhtamir_Khoid_
Tamir")

#export data
#write.csv(gages, "Khangai_Daily_Q.csv")

#convert daily flows to mm for comparison across basins
#basin size conversion to mm
#(((xxm^3/sec*86400sec)/1,000,000m^2)*1000mm)/xxxkm^2
#basin sizes at gages:Baidrag 1622 km^2, Bayanburd 5887 km^2, Bayankhongor 2436 km^2, Bogd
7564 km^2, Erdenenmandal 5799 km^2, Ikhtamir 1963 km^2
baid.b <-1622
bayanb.b <-5887
bayank.b <- 2436
bogd.b <-7564
erd.b <-5799
ikh.b <- 1963

basin <-function(x,y) {
  ((((x*86400)/1000000)*1000)/y)
}

#daily flow in mm
gages.mm <-data.frame(gages[,1:3], basin(gages[,4],baid.b),
basin(gages[,5],bayanb.b),basin(gages[,6],bayank.b),basin(gages[,7],bogd.b),basin(gages[,8], erd.b),
basin(gages[,9],ikh.b))
names(gages.mm) <-c("Month","Day", "Year", "Baidrag_Baidrag",
"Bayanburd_Baidrag","Bayankhongor_Tuin","Bogd_Tuin","Erdenemandal_Khanui","Ikhtamir_Khoid_
Tamir")

#export data
#write.csv(gages.mm, "Khangai_Daily_Q_mm.csv")

#check each location for missing data (counting non-NA values)
count.q <-ddply(gages, .(Year, Month), summarize, BaidB = sum(!is.na(Baidrag_Baidrag)),BayanB =
sum(!is.na(Bayanburd_Baidrag)), BayanT=sum(!is.na(Bayankhongor_Tuin)),
BogdT=sum(!is.na(Bogd_Tuin)), ErdK=sum(!is.na(Erdenemandal_Khanui)),
IkhKT=sum(!is.na(Ikhtamir_Khoid_Tamir)))

#extract those months with less than 75% data (less than 23 days) and make NA
#use if statement to match those months with less than 23 days back to original df
```

```
gages.na <-gages

#join in na indicator by month and station
#if <23 then NA, otherwise 1
count.qbaidb <-ifelse(count.q[,3]<23, NA, 1)
count.qbayanb <-ifelse(count.q[,4]<23, NA, 1)
count.qbayant <-ifelse(count.q[,5]<23, NA, 1)
count.qbogdt <-ifelse(count.q[,6]<23, NA, 1)
count.qerdk <-ifelse(count.q[,7]<23, NA, 1)
count.qikhkt <-ifelse(count.q[,8]<23, NA, 1)

#rejoin to year mon
count.qna <-data.frame(count.q[,1:2], count.qbaidb,count.qbayanb,count.qbayant,
count.qbogdt,count.qerdk,count.qikhkt)
count.qna$yrmon<-paste(count.qna$Month, "-", count.qna$Year, sep="")

#join to original gages.na df
gages.na$yrmon<-paste(gages.na$Month, "-", gages.na$Year, sep="")

count.qnaday <-join(gages.na, count.qna, by="yrmon", type="left")

#make column values NA if count values are 1 keep original values if count values are 0
baidbna <-ifelse(is.na(count.qnaday[,13]),-9999, count.qnaday[,4] )
bayanbna <-ifelse(is.na(count.qnaday[,14]),-9999, count.qnaday[,5] )
bayantna <-ifelse(is.na(count.qnaday[,15]),-9999, count.qnaday[,6] )
bogdtna <-ifelse(is.na(count.qnaday[,16]),-9999, count.qnaday[,7] )
erdkna <-ifelse(is.na(count.qnaday[,17]),-9999, count.qnaday[,8] )
ikhtna <-ifelse(is.na(count.qnaday[,18]),-9999, count.qnaday[,9] )

#join columns back to df
gages.namon <-data.frame(gages.na[,1:3],baidbna,bayanbna,bayantna,bogdtna,erdkna, ikhtna)
colnames(gages.namon) <-c("Month", "Day", "Year", "Baidrag_Baidrag",
"Bayanburd_Baidrag","Bayankhongor_Tuin","Bogd_Tuin","Erdenemandal_Khanui","Ikhtamir_Khoid_
Tamir")

#replace -9999 values with NA
gages.namon[gages.namon == -9999] <- NA

#aggregate to monthly (mean and median flow)
month.qmean <- ddply(gages.namon, .(Year, Month), summarize, BaidB_meanQ =
mean(Baidrag_Baidrag, na.rm=TRUE),BayanB_meanQ = mean(Bayanburd_Baidrag, na.rm=TRUE),
BayanT_meanQ=mean(Bayankhongor_Tuin, na.rm=TRUE), BogdT_meanQ=mean(Bogd_Tuin,
na.rm=TRUE), ErdK_meanQ=mean(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_meanQ=mean(Ikhtamir_Khoid_Tamir, na.rm=TRUE))
month.qmedian <- ddply(gages.namon, .(Year, Month), summarize, BaidB_medianQ =
median(Baidrag_Baidrag, na.rm=TRUE),BayanB_medianQ = median(Bayanburd_Baidrag,
na.rm=TRUE), BayanT_medianQ=median(Bayankhongor_Tuin, na.rm=TRUE),
BogdT_medianQ=median(Bogd_Tuin, na.rm=TRUE),
ErdK_medianQ=median(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_medianQ=median(Ikhtamir_Khoid_Tamir, na.rm=TRUE))


month.qmedianlor <- ddply(gages.namon, .(Month), summarize, BaidB_medianQ =
median(Baidrag_Baidrag, na.rm=TRUE),BayanB_medianQ = median(Bayanburd_Baidrag,
na.rm=TRUE), BayanT_medianQ=median(Bayankhongor_Tuin, na.rm=TRUE),
```

```
BogdT_medianQ=median(Bogd_Tuin, na.rm=TRUE),
ErdK_medianQ=median(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_medianQ=median(Ikhtamir_Khoid_Tamir, na.rm=TRUE))

#write.csv(month.qmean, "Mean_Monthly_Q.csv")
#write.csv(month.qmedian, "Median_Monthly_Q.csv")

#convert to mm and sum
gagesnamon.mm <-data.frame(gages.namon[,1:3], basin(gages.namon[,4],baid.b),
basin(gages.namon[,5],bayanb.b),basin(gages.namon[,6],bayank.b),basin(gages.namon[,7],bogd.b),ba
sin(gages.namon[,8], erd.b), basin(gages.namon[,9],ikh.b))
names(gagesnamon.mm) <-c("Month","Day", "Year", "Baidrag_Baidrag",
"Bayanburd_Baidrag","Bayankhongor_Tuin","Bogd_Tuin","Erdenemandal_Khanui","Ikhtamir_Khoid_
Tamir")
month.qcum <- ddply(gagesnamon.mm, .(Year, Month), summarize, BaidB_cumQ =
sum(Baidrag_Baidrag, na.rm=TRUE),BayanB_cumQ = sum(Bayanburd_Baidrag, na.rm=TRUE),
BayanT_cumQ=sum(Bayankhongor_Tuin, na.rm=TRUE), BogdT_cumQ=sum(Bogd_Tuin,
na.rm=TRUE), ErdK_cumQ=sum(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_cumQ=sum(Ikhtamir_Khoid_Tamir, na.rm=TRUE))

#write.csv(month.qcum, "Cumulative_Monthly_Q_mm.csv")

#count months in a year to see if any months missing
#If less than 12 months data avaialble, make year NA
count.monq <-ddply(month.qmean, .(Year), summarize, BaidB_mon = sum(!is.na(BaidB_meanQ)),
BayanB_mon = sum(!is.na(BayanB_meanQ)), BayanT_mon=sum(!is.na(BayanT_meanQ)),
BogdT_mon=sum(!is.na(BogdT_meanQ)), ErdK_mon=sum(!is.na(ErdK_meanQ)),
IkhKT_mon=sum(!is.na(IkhKT_meanQ)))

#use monthly gages file
gages.nayr <-gages.namon

#join in na indicator by year and station
#if <12 then NA, otherwise 1
count.qbaidbyr <-ifelse(count.monq[,2]<12, NA, 1)
count.qbayanbyr <-ifelse(count.monq[,3]<12, NA, 1)
count.qbayantyr <-ifelse(count.monq[,4]<12, NA, 1)
count.qbogdtyr <-ifelse(count.monq[,5]<12, NA, 1)
count.qerdkyr <-ifelse(count.monq[,6]<12, NA, 1)
count.qikhktyr <-ifelse(count.monq[,7]<12, NA, 1)

#rejoin to year
count.qnayr <-data.frame(count.monq[,1], count.qbaidbyr,count.qbayanbyr,count.qbayantyr,
count.qbogdtyr,count.qerdkyr,count.qikhktyr)
colnames(count.qnayr) <-c("Year","count.qbaidbyr","count.qbayanbyr","count.qbayantyr",
"count.qbogdtyr","count.qerdkyr","count.qikhktyr" )

#join to original gages.na df
count.qnayrf <-join(gages.nayr, count.qnayr, by="Year", type="left")

#make column values NA if count values are 1 keep original values if count values are 0
baidbnayr <-ifelse(is.na(count.qnayrf[,10]),-9999, count.qnayrf[,4] )
bayanbnayr <-ifelse(is.na(count.qnayrf[,11]),-9999, count.qnayrf[,5] )
bayantnayr <-ifelse(is.na(count.qnayrf[,12]),-9999, count.qnayrf[,6] )
bogdtnayr <-ifelse(is.na(count.qnayrf[,13]),-9999, count.qnayrf[,7] )
```

```
erdknayr <-ifelse(is.na(count.qnayrf[,14]),-9999, count.qnayrf[,8] )
ikhtnayr <-ifelse(is.na(count.qnayrf[,15]),-9999, count.qnayrf[,9] )

#join columns back to df
gages.nayrf <-data.frame(gages.nayr[,1:3],baidbnayr,bayanbnayr,bayantnayr,bogdtnayr,erdknayr,
ikhtnayr)
colnames(gages.nayrf) <-c("Month", "Day", "Year", "Baidrag_Baidrag",
"Bayanburd_Baidrag","Bayankhongor_Tuin","Bogd_Tuin","Erdenemandal_Khanui","Ikhtamir_Khoid_
Tamir")

#replace -9999 values with NA
gages.nayrf[gages.nayrf == -9999] <- NA

#aggregate to annual data (cum flow, mean, and median flow)
annual.qmean <- ddply(gages.nayrf, .(Year), summarize, BaidB_meanQ = mean(Baidrag_Baidrag,
na.rm=TRUE),BayanB_meanQ = mean(Bayanburd_Baidrag, na.rm=TRUE),
BayanT_meanQ=mean(Bayankhongor_Tuin, na.rm=TRUE), BogdT_meanQ=mean(Bogd_Tuin,
na.rm=TRUE), ErdK_meanQ=mean(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_meanQ=mean(Ikhtamir_Khoid_Tamir, na.rm=TRUE))
annual.qmedian <- ddply(gages.nayrf, .(Year), summarize, BaidB_medianQ =
median(Baidrag_Baidrag, na.rm=TRUE),BayanB_medianQ = median(Bayanburd_Baidrag,
na.rm=TRUE), BayanT_medianQ=median(Bayankhongor_Tuin, na.rm=TRUE),
BogdT_medianQ=median(Bogd_Tuin, na.rm=TRUE),
ErdK_medianQ=median(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_medianQ=median(Ikhtamir_Khoid_Tamir, na.rm=TRUE))

#write.csv(annual.qmean, "Mean_Annual_Q.csv")
#write.csv(annual.qmedian, "Median_Annual_Q.csv")

#convert to mm and sum
gagesnayr.mm <-data.frame(gages.nayrf[,1:3], basin(gages.nayrf[,4],baid.b),
basin(gages.nayrf[,5],bayanb.b),basin(gages.nayrf[,6],bayank.b),basin(gages.nayrf[,7],bogd.b),basin(
gages.nayrf[,8], erd.b), basin(gages.nayrf[,9],ikh.b))
names(gagesnayr.mm) <-c("Month","Day", "Year", "Baidrag_Baidrag",
"Bayanburd_Baidrag","Bayankhongor_Tuin","Bogd_Tuin","Erdenemandal_Khanui","Ikhtamir_Khoid_
Tamir")

annual.qcum <- ddply(gages.nayrf, .(Year), summarize, BaidB_cumQ = sum(Baidrag_Baidrag,
na.rm=TRUE),BayanB_cumQ = sum(Bayanburd_Baidrag, na.rm=TRUE),
BayanT_cumQ=sum(Bayankhongor_Tuin, na.rm=TRUE), BogdT_cumQ=sum(Bogd_Tuin,
na.rm=TRUE), ErdK_cumQ=sum(Erdenemandal_Khanui,na.rm=TRUE),
IkhKT_cumQ=sum(Ikhtamir_Khoid_Tamir, na.rm=TRUE))

#write.csv(annual.qcum, "Cumulative_Annual_Q_mm.csv")

#-----------------------------------------------------------------
#Checking for autocorrelation in datasets for trend analysis
#precipitation: day.precipaoi, mon.precipaoi , yr.precipaoi

#make precip data like format of Q data
day.precipw <-dcast(day.precipaoi, yrmonday~name, value.var="dayP")
mon.precipw <-dcast(mon.precipaoi, yrmon~name, value.var="monP")
yr.precipw <-dcast(yr.precipaoi, year~name, value.var="yrP")
#segregate to season
seas.precipaoi$yearseas <-paste(seas.precipaoi$yrseas, seas.precipaoi$season, sep="_")
```

```
win.precipaoi <-subset(seas.precipaoi, season=="winter")
spr.precipaoi <-subset(seas.precipaoi, season=="spring")
sum.precipaoi <-subset(seas.precipaoi, season=="summer")
fal.precipaoi <-subset(seas.precipaoi, season=="fall")
seas.precipwwin <-dcast(win.precipaoi, yearseas~name, value.var="seasP")
seas.precipwspr <-dcast(spr.precipaoi, yearseas~name, value.var="seasP")
seas.precipwsum <-dcast(sum.precipaoi, yearseas~name, value.var="seasP")
seas.precipwfal <-dcast(fal.precipaoi, yearseas~name, value.var="seasP")

#daily autocorrelation
baidacfd <-acf(day.precipw[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfd <-acf(day.precipw[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfd <-acf(day.precipw[,4], lag.max=5, na.action=na.pass, plot=FALSE)
galacfd <-acf(day.precipw[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfd <-acf(day.precipw[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfd <-acf(day.precipw[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.dayp <-rbind(baidacfd$acf,bayancfd$acf,erdacfd$acf,galacfd$acf,khoracfd$acf,tsetacfd$acf)
rownames(acf.dayp) <-colnames(day.precipw[2:7])
colnames(acf.dayp) <-c("0", "1", "2", "3", "4", "5")

#monthly autocorrelation
baidacfm <-acf(mon.precipw[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfm <-acf(mon.precipw[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfm <-acf(mon.precipw[,4], lag.max=5, na.action=na.pass, plot=FALSE)
galacfm <-acf(mon.precipw[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfm <-acf(mon.precipw[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfm <-acf(mon.precipw[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.monp <-rbind(baidacfm$acf,bayancfm$acf,erdacfm$acf,galacfm$acf,khoracfm$acf,tsetacfm$acf)
rownames(acf.monp) <-colnames(mon.precipw[2:7])
colnames(acf.monp) <-c("0", "1", "2", "3", "4", "5")

#seasonal autocorrelation winter
baidacfswin <-acf(seas.precipwwin[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfswin <-acf(seas.precipwwin[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfswin <-acf(seas.precipwwin[,4], lag.max=5, na.action=na.pass, plot=FALSE)
galacfswin <-acf(seas.precipwwin[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfswin <-acf(seas.precipwwin[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfswin <-acf(seas.precipwwin[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.seaspwin <-
rbind(baidacfswin$acf,bayancfswin$acf,erdacfswin$acf,galacfswin$acf,khoracfswin$acf,tsetacfswin$
acf)
rownames(acf.seaspwin) <-colnames(seas.precipwwin[2:7])
colnames(acf.seaspwin) <-c("0", "1", "2", "3", "4", "5")


#seasonal autocorrelation spring
baidacfsspr <-acf(seas.precipwspr[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfsspr <-acf(seas.precipwspr[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfsspr <-acf(seas.precipwspr[,4], lag.max=5, na.action=na.pass, plot=FALSE)
```

```
galacfsspr <-acf(seas.precipwspr[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfsspr <-acf(seas.precipwspr[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfsspr <-acf(seas.precipwspr[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.seaspspr <-
rbind(baidacfsspr$acf,bayancfsspr$acf,erdacfsspr$acf,galacfsspr$acf,khoracfsspr$acf,tsetacfsspr$acf
)
rownames(acf.seaspspr) <-colnames(seas.precipwspr[2:7])
colnames(acf.seaspspr) <-c("0", "1", "2", "3", "4", "5")

#seasonal autocorrelation summer
baidacfssum <-acf(seas.precipwsum[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfssum <-acf(seas.precipwsum[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfssum <-acf(seas.precipwsum[,4], lag.max=5, na.action=na.pass, plot=FALSE)
galacfssum <-acf(seas.precipwsum[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfssum <-acf(seas.precipwsum[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfssum <-acf(seas.precipwsum[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.seaspsum <-
rbind(baidacfssum$acf,bayancfssum$acf,erdacfssum$acf,galacfssum$acf,khoracfssum$acf,tsetacfssu
m$acf)
rownames(acf.seaspsum) <-colnames(seas.precipwsum[2:7])
colnames(acf.seaspsum) <-c("0", "1", "2", "3", "4", "5")

#seasonal autocorrelation fall
baidacfsfal <-acf(seas.precipwfal[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfsfal <-acf(seas.precipwfal[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfsfal <-acf(seas.precipwfal[,4], lag.max=5, na.action=na.pass, plot=FALSE)
galacfsfal <-acf(seas.precipwfal[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfsfal <-acf(seas.precipwfal[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfsfal <-acf(seas.precipwfal[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.seaspfal <-
rbind(baidacfsfal$acf,bayancfsfal$acf,erdacfsfal$acf,galacfsfal$acf,khoracfsfal$acf,tsetacfsfal$acf)
rownames(acf.seaspfal) <-colnames(seas.precipwfal[2:7])
colnames(acf.seaspfal) <-c("0", "1", "2", "3", "4", "5")


#annual autocorrelation
baidacfy <-acf(yr.precipw[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayancfy <-acf(yr.precipw[,3], lag.max=5, na.action=na.pass, plot=FALSE)
erdacfy <-acf(yr.precipw[,4], lag.max=5, na.action=na.pass, plot=FALSE)
galacfy <-acf(yr.precipw[,5], lag.max=5, na.action=na.pass, plot=FALSE)
khoracfy <-acf(yr.precipw[,6], lag.max=5, na.action=na.pass, plot=FALSE)
tsetacfy <-acf(yr.precipw[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.yrp <-rbind(baidacfy$acf,bayancfy$acf,erdacfy$acf,galacfy$acf,khoracfy$acf,tsetacfy$acf)
rownames(acf.yrp) <-colnames(yr.precipw[2:7])
colnames(acf.yrp) <-c("0", "1", "2", "3", "4", "5")

#5% intervals
```

```
acf.dayp5 <-
rbind(baidacfd$n.used,bayancfd$n.used,erdacfd$n.used,galacfd$n.used,khoracfd$n.used,tsetacfd$n.u
sed)
acf.dayp5lev <-2/sqrt(acf.dayp5)

acf.monp5 <-
rbind(baidacfm$n.used,bayancfm$n.used,erdacfm$n.used,galacfm$n.used,khoracfm$n.used,tsetacfm
$n.used)
acf.monp5lev <-2/sqrt(acf.monp5)

acf.seaspwin <-
rbind(baidacfswin$n.used,bayancfswin$n.used,erdacfswin$n.used,galacfswin$n.used,khoracfswin$n.
used,tsetacfswin$n.used)
acf.seaspspr <-
rbind(baidacfsspr$n.used,bayancfsspr$n.used,erdacfsspr$n.used,galacfsspr$n.used,khoracfsspr$n.us
ed,tsetacfsspr$n.used)
acf.seaspsum <-
rbind(baidacfssum$n.used,bayancfssum$n.used,erdacfssum$n.used,galacfssum$n.used,khoracfssum
$n.used,tsetacfssum$n.used)
acf.seaspfal <-
rbind(baidacfsfal$n.used,bayancfsfal$n.used,erdacfsfal$n.used,galacfsfal$n.used,khoracfsfal$n.used,t
setacfsfal$n.used)

acf.monp5levwin <-2/sqrt(acf.seaspwin)
acf.monp5levspr <-2/sqrt(acf.seaspspr)
acf.monp5levsum <-2/sqrt(acf.seaspsum)
acf.monp5levfal <-2/sqrt(acf.seaspfal)

acf.yrp5 <-
rbind(baidacfy$n.used,bayancfy$n.used,erdacfy$n.used,galacfy$n.used,khoracfy$n.used,tsetacfy$n.us
ed)
acf.yrp5lev <-2/sqrt(acf.yrp5)

#-------------------------------------------------------------------
#Checking for autocorrelation in datasets for trend analysis
#streamflow: gages, gages.mm, month.qmean, month.qmedian, month.qcum (in mm), annual.qmean,
annual.qmedian, annual.qcum (in mm)

#daily autocorrelation cms
baidbacfd <-acf(gages[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfd <-acf(gages[,5], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfd <-acf(gages[,6], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfd <-acf(gages[,7], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfd <-acf(gages[,8], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfd <-acf(gages[,9], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.dayq <-
rbind(baidbacfd$acf,bayanbcfd$acf,bayantacfd$acf,bogdtacfd$acf,erdkacfd$acf,ikhktacfd$acf)
rownames(acf.dayq) <-colnames(gages[,4:9])
colnames(acf.dayq) <-c("0", "1", "2", "3", "4", "5")

#daily autocorrelation mm
baidbacfdmm <-acf(gages.mm[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfdmm <-acf(gages.mm[,5], lag.max=5, na.action=na.pass, plot=FALSE)
```

```
bayantacfdmm <-acf(gages.mm[,6], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfdmm <-acf(gages.mm[,7], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfdmm <-acf(gages.mm[,8], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfdmm <-acf(gages.mm[,9], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.dayqmm <-
rbind(baidbacfdmm$acf,bayanbcfdmm$acf,bayantacfdmm$acf,bogdtacfdmm$acf,erdkacfdmm$acf,ik
hktacfdmm$acf)
rownames(acf.dayqmm) <-colnames(gages.mm[,4:9])
colnames(acf.dayqmm) <-c("0", "1", "2", "3", "4", "5")

#monthly autocorrelation mean
baidbacfmonmean <-acf(month.qmean[,3], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfmonmean <-acf(month.qmean[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfmonmean <-acf(month.qmean[,5], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfmonmean <-acf(month.qmean[,6], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfmonmean <-acf(month.qmean[,7], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfmonmean <-acf(month.qmean[,8], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.monmean <-
rbind(baidbacfmonmean$acf,bayanbcfmonmean$acf,bayantacfmonmean$acf,bogdtacfmonmean$acf,
erdkacfmonmean$acf,ikhktacfmonmean$acf)
rownames(acf.monmean) <-colnames(month.qmean[,3:8])
colnames(acf.monmean) <-c("0", "1", "2", "3", "4", "5")

#monthly autocorrelation median
baidbacfmonmedian <-acf(month.qmedian[,3], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfmonmedian <-acf(month.qmedian[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfmonmedian <-acf(month.qmedian[,5], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfmonmedian <-acf(month.qmedian[,6], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfmonmedian <-acf(month.qmedian[,7], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfmonmedian <-acf(month.qmedian[,8], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.monmedian <-
rbind(baidbacfmonmedian$acf,bayanbcfmonmedian$acf,bayantacfmonmedian$acf,bogdtacfmonmed
ian$acf,erdkacfmonmedian$acf,ikhktacfmonmedian$acf)
rownames(acf.monmedian) <-colnames(month.qmedian[,3:8])
colnames(acf.monmedian) <-c("0", "1", "2", "3", "4", "5")

#monthly autocorrelation cumulative
baidbacfmonqcum <-acf(month.qcum[,3], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfmonqcum <-acf(month.qcum[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfmonqcum <-acf(month.qcum[,5], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfmonqcum <-acf(month.qcum[,6], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfmonqcum <-acf(month.qcum[,7], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfmonqcum <-acf(month.qcum[,8], lag.max=5, na.action=na.pass, plot=FALSE)


#collate results
acf.monqcum <-
rbind(baidbacfmonqcum$acf,bayanbcfmonqcum$acf,bayantacfmonqcum$acf,bogdtacfmonqcum$acf,
erdkacfmonqcum$acf,ikhktacfmonqcum$acf)
```

```
rownames(acf.monqcum) <-colnames(month.qcum[,3:8])
colnames(acf.monqcum) <-c("0", "1", "2", "3", "4", "5")

#annual autocorrelation mean
baidbacfannmean <-acf(annual.qmean[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfannmean <-acf(annual.qmean[,3], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfannmean <-acf(annual.qmean[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfannmean <-acf(annual.qmean[,5], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfannmean <-acf(annual.qmean[,6], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfannmean <-acf(annual.qmean[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.annmean <-
rbind(baidbacfannmean$acf,bayanbcfannmean$acf,bayantacfannmean$acf,bogdtacfannmean$acf,er
dkacfannmean$acf,ikhktacfannmean$acf)
rownames(acf.annmean) <-colnames(annual.qmean[,2:7])
colnames(acf.annmean) <-c("0", "1", "2", "3", "4", "5")

#annual autocorrelation median
baidbacfannmedian <-acf(annual.qmedian[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfannmedian <-acf(annual.qmedian[,3], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfannmedian <-acf(annual.qmedian[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfannmedian <-acf(annual.qmedian[,5], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfannmedian <-acf(annual.qmedian[,6], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfannmedian <-acf(annual.qmedian[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.annmedian <-
rbind(baidbacfannmedian$acf,bayanbcfannmedian$acf,bayantacfannmedian$acf,bogdtacfannmedian
$acf,erdkacfannmedian$acf,ikhktacfannmedian$acf)
rownames(acf.annmedian) <-colnames(annual.qmedian[,2:7])
colnames(acf.annmedian) <-c("0", "1", "2", "3", "4", "5")

#annual autocorrelation cumulative
baidbacfannqcum <-acf(annual.qcum[,2], lag.max=5, na.action=na.pass, plot=FALSE)
bayanbcfannqcum <-acf(annual.qcum[,3], lag.max=5, na.action=na.pass, plot=FALSE)
bayantacfannqcum <-acf(annual.qcum[,4], lag.max=5, na.action=na.pass, plot=FALSE)
bogdtacfannqcum <-acf(annual.qcum[,5], lag.max=5, na.action=na.pass, plot=FALSE)
erdkacfannqcum <-acf(annual.qcum[,6], lag.max=5, na.action=na.pass, plot=FALSE)
ikhktacfannqcum <-acf(annual.qcum[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.annqcum <-
rbind(baidbacfannqcum$acf,bayanbcfannqcum$acf,bayantacfannqcum$acf,bogdtacfannqcum$acf,er
dkacfannqcum$acf,ikhktacfannqcum$acf)
rownames(acf.annqcum) <-colnames(annual.qcum[,2:7])
colnames(acf.annqcum) <-c("0", "1", "2", "3", "4", "5")

#5% intervals
acf.dayq5 <-2/sqrt(baidbacfd$n.used)

acf.monp5mean <-2/sqrt(baidbacfmonmean$n.used)
acf.monp5median <-2/sqrt(baidbacfmonmedian$n.used)
acf.monp5cum <-2/sqrt(baidbacfmonqcum$n.used)
```

```
acf.yrp5mean <--2/sqrt(baidbacfannmean$n.used)
acf.yrp5median <--2/sqrt(baidbacfannmedian$n.used )
acf.yrp5cum <--2/sqrt(baidbacfannqcum$n.used)

#-----------------------------------------------------------------
#trend analyses
#first testing for trend in all time series no adjustments for autocorrelation
#precipitation: day.precipaoi, mon.precipaoi , yr.precipaoi, day.precipw, month.precipw, yr.precipw
#streamflow: gages, gages.mm, month.qmean, month.qmedian, month.qcum (in mm), annual.qmean,
annual.qmedian, annual.qcum (in mm)

#checking annual values first, no adjustment for autocorrelation
#precipitation
mk.yrprecip <-apply(yr.precipw[,2:7], 2, MannKendall)

#checking that NA's don't affect the results-seems ok
#baidragtest <-yr.precipw[,2]
#baidragtestn <-baidragtest[complete.cases(baidragtest)]
#baidragmk <-MannKendall(baidragtestn)

#baidragtest <-yr.precipw[,2]
#baidragtestn <-baidragtest[complete.cases(baidragtest)]
#baidragmk <-MannKendall(baidragtestn)

mk.winprecip <-apply(seas.precipwwin[,2:7], 2, MannKendall)
mk.sprprecip <-apply(seas.precipwspr[,2:7], 2, MannKendall)
mk.sumprecip <-apply(seas.precipwsum[,2:7], 2, MannKendall)
mk.falprecip <-apply(seas.precipwfal[,2:7], 2, MannKendall)

#loop for ts winter with y intercept only from years of data
seas.precipwdwin <-seas.precipwwin[,2:7]
ts.seaspwin <-matrix(nrow=6, ncol=2)
rownames(ts.seaspwin) <-colnames(seas.precipwwin[,2:7])
colnames(ts.seaspwin) <-c("intercept", "slope")

for (i in 1:6){
 year.p<-as.numeric(substr(seas.precipwwin[,1],1,4))
 station <-seas.precipwdwin[,i]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.ps <-year.p[stationc]
 ts.seas <-zyp.sen(stationcomp~year.ps)
 ts.seaspwin[i,] <-ts.seas$coefficients
}

#loop for ts spring
seas.precipwdspr <-seas.precipwspr[,2:7]
ts.seaspspr <-matrix(nrow=6, ncol=2)
rownames(ts.seaspspr) <-colnames(seas.precipwspr[,2:7])
colnames(ts.seaspspr) <-c("intercept", "slope")

for (i in 1:6){
 year.p <-as.numeric(substr(seas.precipwspr[,1],1,4))
 station <-seas.precipwdspr[,i]
 stationc<-which(!is.na(station))
```

```
  stationcomp <-station[stationc]
  year.ps <-year.p[stationc]
  ts.seas <-zyp.sen(stationcomp~year.ps)
  ts.seaspspr[i,] <-ts.seas$coefficients
}

#loop for ts summer
seas.precipwdsum <-seas.precipwsum[,2:7]
ts.seaspsum <-matrix(nrow=6, ncol=2)
rownames(ts.seaspsum) <-colnames(seas.precipwsum[,2:7])
colnames(ts.seaspsum) <-c("intercept", "slope")

for (i in 1:6){
  year.p <-as.numeric(substr(seas.precipwsum[,1],1,4))
  station <-seas.precipwdsum[,i]
  stationc<-which(!is.na(station))
  stationcomp <-station[stationc]
  year.ps <-year.p[stationc]
  ts.seas <-zyp.sen(stationcomp~year.ps)
  ts.seaspsum[i,] <-ts.seas$coefficients
}

#loop for ts fall
seas.precipwdfal <-seas.precipwfal[,2:7]
ts.seaspfal <-matrix(nrow=6, ncol=2)
rownames(ts.seaspfal) <-colnames(seas.precipwfal[,2:7])
colnames(ts.seaspfal) <-c("intercept", "slope")

for (i in 1:6){
  year.p <-as.numeric(substr(seas.precipwfal[,1],1,4))
  station <-seas.precipwdfal[,i]
  stationc<-which(!is.na(station))
  stationcomp <-station[stationc]
  year.ps <-year.p[stationc]
  ts.seas <-zyp.sen(stationcomp~year.ps)
  ts.seaspfal[i,] <-ts.seas$coefficients
}

#loop for ts
yr.precipwd <-yr.precipw[,2:7]
ts.yrp <-matrix(nrow=6, ncol=2)
rownames(ts.yrp) <-colnames(yr.precipw[,2:7])
colnames(ts.yrp) <-c("intercept", "slope")

for (i in 1:6){
  year.p <-as.numeric(yr.precipw[,1])
  station <-yr.precipwd[,i]
  stationc<-which(!is.na(station))
  stationcomp <-station[stationc]
  year.ps <-year.p[stationc]
  ts <-zyp.sen(stationcomp~year.ps)
  ts.yrp[i,] <-ts$coefficients
}

#streamflow
```

```
mk.yrqmean <-apply(annual.qmean[,2:7], 2, MannKendall)
mk.yrqmedian <-apply(annual.qmedian[,2:7], 2, MannKendall)
mk.yrqcum <-apply(annual.qcum[,2:7], 2, MannKendall)

#loop for ts

annual.qmeand <-annual.qmean[,2:7]
ts.yrqmean <-matrix(nrow=6, ncol=2)
rownames(ts.yrqmean) <-colnames(annual.qmean[,2:7])
colnames(ts.yrqmean) <-c("intercept", "slope")

for (i in 1:6){
 year.qf <-annual.qmean[,1]
 year.q <-as.numeric(levels(year.qf)[year.qf])
 station <-annual.qmeand[,i]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.qs <-year.q[stationc]
 ts <-zyp.sen(stationcomp~year.qs)
 ts.yrqmean[i,] <-ts$coefficients
}

annual.qmediand <-annual.qmedian[,2:7]
ts.yrqmedian <-matrix(nrow=6, ncol=2)
rownames(ts.yrqmedian) <-colnames(annual.qmedian[,2:7])
colnames(ts.yrqmedian) <-c("intercept", "slope")

for (i in 1:6){
 year.qf <-annual.qmean[,1]
 year.q <-as.numeric(levels(year.qf)[year.qf])
 station <-annual.qmediand[,i]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.qs <-year.q[stationc]
 ts <-zyp.sen(stationcomp~year.qs)
 ts.yrqmedian[i,] <-ts$coefficients
}

#make 0 cumulative flow NA
annual.qcum[annual.qcum == 0]<-NA
annual.qcumd <-annual.qcum[,2:7]
ts.yrqcum <-matrix(nrow=6, ncol=2)
rownames(ts.yrqcum) <-colnames(annual.qcum[,2:7])
colnames(ts.yrqcum) <-c("intercept", "slope")

for (i in 1:6){
 year.qf <-annual.qmean[,1]
 year.q <-as.numeric(levels(year.qf)[year.qf])
 station <-annual.qcumd[,i]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.qs <-year.q[stationc]
 ts <-zyp.sen(stationcomp~year.qs)
 ts.yrqcum[i,] <-ts$coefficients
}
```

```
#trend slope using mblm Thiel-Sen estimator for comparison to zyp.sen results
#truncate each to no NA based on non-NA values in station
stationc<-which(!is.na(station))
stationcomp <-station[complete.cases(station)]
year.ps <-year.p[stationc]
ts.test<-mblm(stationcomp~year.ps, repeated=FALSE)

#-----------------------------------------------
#using TFPW methods on annual data
#reformat data
#precipitation: yr.precipaoi,  yr.precipw, seas.precipwwin, seas.precipwspr, seas.precipwsum,
seas.precipwfal

#seasonal winter
seas.precipwtwin <-t(seas.precipwwin)
seas.precipdfwin <-data.frame(seas.precipwtwin[2:7,])
colnames(seas.precipdfwin) <-seas.precipwsum[,1]

#run ts analysis
#winter
seas.preciptspwwin <-zyp.trend.dataframe(seas.precipdfwin, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#seasonal spring
seas.precipwtspr <-t(seas.precipwspr)
seas.precipdfspr <-data.frame(seas.precipwtspr[2:7,])
colnames(seas.precipdfspr) <-seas.precipwspr[,1]

#run ts analysis
seas.preciptspwspr <-zyp.trend.dataframe(seas.precipdfspr, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#seasonal summer
seas.precipwtsum <-t(seas.precipwsum)
seas.precipdfsum <-data.frame(seas.precipwtsum[2:7,])
colnames(seas.precipdfsum) <-seas.precipwsum[,1]

#run ts analysis
seas.preciptspwsum <-zyp.trend.dataframe(seas.precipdfsum, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#seasonal fall
seas.precipwtfal <-t(seas.precipwfal)
seas.precipdffal <-data.frame(seas.precipwtfal[2:7,])
colnames(seas.precipdffal) <-seas.precipwfal[,1]

#run ts analysis
seas.preciptspwfal <-zyp.trend.dataframe(seas.precipdffal, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#annual
yr.precipwt <-t(yr.precipw)
yr.precipdf <-data.frame(yr.precipwt[2:7,])
colnames(yr.precipdf) <-yr.precipw[,1]
```

```
#run ts analysis
yr.preciptspw <-zyp.trend.dataframe(yr.precipdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#streamflow: annual.qmean, annual.qmedian, annual.qcum (in mm)
#mean annual
yr.qmeant <-t(annual.qmean)
yr.qmeandf <-data.frame(yr.qmeant[2:7,])
colnames(yr.qmeandf) <-annual.qmean[,1]

yr.qmeantspw <-zyp.trend.dataframe(yr.qmeandf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#median annual
yr.qmediant <-t(annual.qmedian)
yr.qmediandf <-data.frame(yr.qmediant[2:7,])
colnames(yr.qmediandf) <-annual.qmedian[,1]

yr.qmedianspw <-zyp.trend.dataframe(yr.qmediandf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#cumulative annual
yr.qcumt <-t(annual.qcum)
yr.qcumdf <-data.frame(yr.qcumt[2:7,])
colnames(yr.qcumdf) <-annual.qcum[,1]

yr.qcumtspw <-zyp.trend.dataframe(yr.qcumdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#-------------------------------------------------
#precip stats for comparison with CO data
#avg annual precip over LOR
yr.precipavg <-apply(yr.precipw[,2:7], 2, FUN=median, na.rm=TRUE)

#median monthly P over LOR
#add month col
mon.precipw$month <-substr(mon.precipw$yrmon, 6, 7)
mon.precipavg <-ddply(mon.precipw, .(month), summarize,
Med_P_Baidrag=median(Baidrag,na.rm=TRUE),Med_P_Bayankhongor=median(Bayankhongor,na.rm
=TRUE), Med_P_Erdenemandal=median(Erdenemandal,na.rm=TRUE),
Med_P_Galuut=median(Galuut,na.rm=TRUE), Med_P_Khoriult=median(Khoriult,na.rm=TRUE),
Med_P_Tsetserleg=median(Tsetserleg, na.rm=TRUE))

#max daily precip recorded
day.precipw$day <-substr(day.precipw$yrmonday, 9,10)
day.precipmax <-ddply(day.precipw, .(day), summarize,
Max_P_Baidrag=max(Baidrag,na.rm=TRUE),Max_P_Bayankhongor=max(Bayankhongor,na.rm=TRUE)
, Max_P_Erdenemandal=max(Erdenemandal,na.rm=TRUE), Max_P_Galuut=max(Galuut,na.rm=TRUE),
Max_P_Khoriult=max(Khoriult,na.rm=TRUE), Max_P_Tsetserleg=max(Tsetserleg, na.rm=TRUE))
day.precipmaxavg <-apply(day.precipmax[,2:7], 2, max)
#streamflow stats for comparison
#average annual flows and Median annual flows over LOR
annual.qmeanlor <-apply(annual.qmean[,2:7], 2, mean, na.rm=TRUE)
annual.qmedianlor <-apply(annual.qmedian[,2:7], 2, median, na.rm=TRUE)
```

```
#average peak flow over LOR
gages.pk <-ddply(gages, .(Month, Day), summarize, med_BaidB =median(Baidrag_Baidrag,
na.rm=TRUE), med_BayB= median(Bayanburd_Baidrag, na.rm=TRUE),
med_BayT=median(Bayankhongor_Tuin, na.rm=TRUE), med_BogdT= median(Bogd_Tuin,
na.rm=TRUE), med_ErdK= median(Erdenemandal_Khanui, na.rm=TRUE),
med_IkhKT=median(Ikhtamir_Khoid_Tamir, na.rm=TRUE))
gages.pkmax <-apply(gages.pk[,3:8], 2, max)
```

C.1.3 Trend Analyses for Colorado Hydroclimate Data

```
#----------------------------------------------------------------------
# TITLE: Trend Analysis Script for LOR PQ for CO data
# AUTHOR: Niah Venable
# DATE WRITTEN: 2014-11-04
# LAST REVISION: 2015-05-11
# DESCRIPTION: This script provides code for analyzing CO met and streamflow data for trends
# PACKAGES REQUIRED: plyr, reshape2, ggplot2
# VARIABLES/DATA USED: text files of Khangai precipitation and streamflow
# NAME:
# TYPE:
# COMMENT:
#---------------------------------------------------------------------------
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/CO_Data/")

library(plyr)
library(reshape2)
library(ggplot2)
library(zyp)
library(Kendall)
library(chron)
library(doBy)

#import data from USGS and NCDC
varfilep <- "CO_Met_P_mm.csv"

precip.t <- read.csv(varfilep, header=TRUE, blank.lines.skip=TRUE, na.strings="-999.9")


#create new categories for aggregating
#yearmon column
precip.t$yrmon <- substr(precip.t$Date, 1,7)
precip.t$year <-substr(precip.t$Date, 1,4)
precip.t$monday <-substr(precip.t$Date, 6,10)

#check max and min values
max(precip.t$P_mm, na.rm=TRUE)
min(precip.t$P_mm, na.rm=TRUE)

#check values for extreme outliers (see Harris et al 2013, sec 3.1, precip within 4 SD)
precip.m <- ddply(precip.t, .(Name, monday), summarize, Mean=mean(P_mm, na.rm=TRUE))
precip.sd <- ddply(precip.t, .(Name, monday), summarize, StdDev=sd(P_mm, na.rm=TRUE))
precip.sdcol <- cbind(precip.sd[ , 3], rep(4,length(precip.sd)))
precip.sdc <- precip.sdcol[ , 1]*precip.sdcol[,2]
```

```
precip.sdfcol <-cbind(precip.m, precip.sdc)
precip.sdfcols <-precip.sdfcol[ , 3]+precip.sdfcol[, 4]
precip.sdf <-cbind(precip.m, precip.sdfcols)
names(precip.sdf)[names(precip.sdf)=="precip.sdfcols"] <- "StdDev4"

#merge in 4sd column to precip data for comparison
precip.sdm <- join(precip.sdf,precip.t, by=c("Name", "monday"))

#find which timesteps exceed 4 sd
precip.sdex <- subset(precip.sdm, subset = precip.sdm$P_mm > precip.sdm$StdDev4)
#results in 6388 values (out of 350186) outside this range, and many are high, but QC'd by NOAA?

#want the most complete and longest records, check na totals each station
#convert to regular time series 1st

#convert to wide format
day.precipw <-dcast(precip.t, Date~Name, value.var="P_mm")

#make a regular time series (in case it wasn't)
#create regular ts using chron
var.regdayp <- seq.dates("01/01/1933", "12/31/2012", by = "days")
var.regsp <-as.data.frame(as.Date(var.regdayp))
colnames(var.regsp) <-"Date"

pvar.j <-join(day.precipw, var.regsp, by="Date", type="full")
#doesn't seem to be any difference with day.precipw

#to find percent missing for LOR need to know LOR for each station
#find location of first non NA value

firstnon <-matrix(ncol=16, nrow=1)

for (i in 1:16) {
  nonaindex <- which(!is.na(day.precipw[,i]))
  firstnon[i] <- min(nonaindex)
}

#of wide format reg time series, total # nas
day.na <-colSums(is.na(day.precipw))

#subtract pre-data start nas to get actual data nas
prena <-day.na-firstnon

#actual record lengths
lor <-29220-firstnon

#percent missing is prena/lor*100
per <-(prena/lor)*100
colnames(per) <-names(day.na)
per[,1] <-NA
per[order(per)]

#top 3 least missing stations
#Lama  1-1-1933 to 12-31-2012 0.72 percent missing
#Del 1-1-1933 to 12-31-2012 1.17 percent missing
```

```
#Karv 8-1-1941 to 12-31-2012 1.51 percent missing

#truncate datatset to these three stations
precip.aoi <-day.precipw[,c(1,12,4,11)]

#add yrmon and mon day cols
precip.aoi$yrmon <-substr(precip.aoi$Date, 1,7)
precip.aoi$monday <-substr(precip.aoi$Date, 6,10)
precip.aoi$year <-substr(precip.aoi$Date, 1,4)
precip.aoi$month <-substr(precip.aoi$Date, 6,7)

#counting values for QC
lama <-precip.aoi[,c(7,8,5,6,2)]
colnames(lama) <-c("year", "month", "yrmon", "monday", "dayP")
del <-precip.aoi[,c(7,8,5,6,3)]
colnames(del) <-c("year", "month", "yrmon", "monday", "dayP")
karv <-precip.aoi[3135:29220,c(7,8,5,6,4)]
colnames(karv) <-c("year", "month", "yrmon", "monday", "dayP")

lama.ct <-count(lama, vars="yrmon")
del.ct <-count(del, vars="yrmon")
karv.ct <-count(karv, vars="yrmon")

#any months with less than 23 days data? none!
which(lama.ct$freq<23)
which(del.ct$freq<23)
which(karv.ct$freq<23)

#aggregate to monthly precip
mon.precipl <- ddply(lama, .(yrmon), summarize, monP=sum(dayP, na.rm=TRUE))
mon.precipd <- ddply(del, .(yrmon), summarize, monP=sum(dayP, na.rm=TRUE))
mon.precipk <- ddply(karv, .(yrmon), summarize, monP=sum(dayP, na.rm=TRUE))

#add station names
mon.precipl$name <-rep("Lama", length(mon.precipl))
mon.precipd$name <-rep("Del", length(mon.precipd))
mon.precipk$name <-rep("Karv", 857)

mon.precip <-rbind(mon.precipl, mon.precipd, mon.precipk)

#export monthly data to csv
write.csv(mon.precip, file= "CO_Monthly_P.csv")

#aggregate to seasonal values winter=DJF, spring=MAM, summer=JJA, fall=SON for Del Norte
#note: truncate to winter 1934 to present due to no december 1932 data in this set.
#add season heading
del$season[del$month=="12"|del$month=="01"|del$month=="02"] <-"winter"
del$season[del$month=="03"|del$month=="04"|del$month=="05"] <-"spring"
del$season[del$month=="06"|del$month=="07"|del$month=="08"] <-"summer"
del$season[del$month=="09"|del$month=="10"|del$month=="11"] <-"fall"
yraoi <-as.numeric(del$year)
del$yrwin<-yraoi+1

del$yrseas <-ifelse(del$month=="12", del$yrwin, del$year)
```

```
#aggregate to seasonal values
seas.precipd <-ddply(del, .(yrseas, season), summarize, seasP=sum(dayP, na.rm=TRUE))

#winter 1933 NA due to no December 1932
seas.precipd[4,3] <-NA

#winter 2015 NA due to no Jan 2015
seas.precipd[329,3] <-NA
write.csv(seas.precipd, file= "DelNorte_Seasonal_P.csv")

#aggregate to annual values
yr.precipl <- ddply(lama, .(year), summarize, yrP=sum(dayP, na.rm=TRUE))
yr.precipd <- ddply(del, .(year), summarize, yrP=sum(dayP, na.rm=TRUE))
yr.precipk <- ddply(karv, .(year), summarize, yrP=sum(dayP, na.rm=TRUE))

yr.precipl$name <-rep("Lama", length(yr.precipl))
yr.precipd$name <-rep("Del", length(yr.precipd))
yr.precipk$name <-rep("Karv", length(yr.precipk))

yr.precip <-rbind(yr.precipl, yr.precipd, yr.precipk)

write.csv(yr.precip, file= "CO_Annual_P.csv")

#precip stats for comparison with CO data
#avg annual precip over LOR
yr.precipavgl <-median(yr.precipl$yrP, na.rm=TRUE)
yr.precipavgd <-median(yr.precipd$yrP, na.rm=TRUE)
yr.precipavgk <-median(yr.precipk$yrP, na.rm=TRUE)

#median monthly P over LOR
mon.precipl$month <-substr(mon.precipl$yrmon, 6,7)
mon.precipd$month <-substr(mon.precipd$yrmon, 6,7)
mon.precipk$month <-substr(mon.precipk$yrmon, 6,7)

mon.precipavgl <-ddply(mon.precipl, .(month), summarize, Med_P=median(monP,na.rm=TRUE))
mon.precipavgd <-ddply(mon.precipd, .(month), summarize, Med_P=median(monP,na.rm=TRUE))
mon.precipavgk <-ddply(mon.precipk, .(month), summarize, Med_P=median(monP,na.rm=TRUE))

#max daily precip recorded
day.precipw$day <-substr(day.precipw$Date, 9,10)
colnames(day.precipw) <- c("Date", "CENT", "CHER", "DEL",  "FORT", "HOLL", "HOLY", "IDAL", "JOES",
"JULE", "KARV", "LAMA", "MONT", "SEDG", "WALS", "YUMA", "day")
day.precipmax <-ddply(day.precipw, .(day), summarize, Max_P_Lama=max(LAMA,na.rm=TRUE),
Max_P_Del=max(DEL,na.rm=TRUE),Max_P_Karv=max(KARV,na.rm=TRUE))
day.precipmaxavg <-apply(day.precipmax[,2:4], 2, max)

#-------------------------------------------------------------------------------------
#streamflow data
#import data

#varfileq <-"Apishapa.csv"
varfileq <-
"/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/CO_Data/Crestone/Crest
one_daily_1948_2014.csv"
```

```
#use same object names and run thru with crestone data
apis.lor <-read.csv(varfileq, header=TRUE, na.strings= "-9999")

#truncate to 1940 start (and 2012-12-31 end)date due to missing data
#apis<-apis.lor[6485:33148,]
#use all as no missing data for Crestone Ck.
apis <-apis.lor

#add date columns
apis$yrmon <-substr(apis$Date, 1,7)
apis$yr <-substr(apis$Date, 1,4)
apis$month <-substr(apis$Date, 6,7)
apis$day <-substr(apis$Date, 9,10)

#convert cfs to cms
#1 cfs=0.028316847 cms

apis.cms <-apis$CFS*0.028316847
apis$CMS<-apis.cms

#annual mean and median
annual.qmedian <-ddply(apis, .(yr), summarize, med <-median(CMS, na.rm=TRUE))
annual.qmean <-ddply(apis, .(yr), summarize, mean <-mean(CMS, na.rm=TRUE))

#write.csv(annual.qmean, "Apishapa_Mean_Annual_Q.csv")
#write.csv(annual.qmedian, "Apishapa_Median_Annual_Q.csv")

write.csv(annual.qmean, "Crestone_Mean_Annual_Q.csv")
write.csv(annual.qmedian, "Crestone_Median_Annual_Q.csv")


#streamflow stats for comparison
#average annual flows and Median annual flows over LOR
annual.qmeanlor <-mean(annual.qmean[,2], na.rm=TRUE)
annual.qmedianlor <-mean(annual.qmedian[,2], na.rm=TRUE)

#average peak flow over LOR
gage.pk <-ddply(apis, .(month, day), summarize, med =median(CMS, na.rm=TRUE))
gage.pkmax <-max(gage.pk$med)
gage.pk$monthday <-paste(gage.pk$month, "-", gage.pk$day, sep="")
gage.pk$julian <-seq(1, 366)

plot(gage.pk$julian, gage.pk$med, type="l")

#median flows by month over LOR
month.qmedianLOR <-ddply(apis, .(month), summarize, med <-median(CMS, na.rm=TRUE))

#---------------------------------------------------
#convert daily flows to mm for comparison across basins
#basin size conversion to mm
#(((xxm^3/sec*86400sec)/1,000,000m^2)*1000mm)/xxxkm^2
#apish <- 1963
cres <-27.71

basin <-function(x,y) {
```

```
  ((((x*86400)/1000000)*1000)/y)
}

#daily flow in mm
#apish.mm <-data.frame(apis[,1:7], basin(apis[,7],apish))
apish.mm <-data.frame(apis[,1:7], basin(apis[,7],cres))
names(apish.mm) <-c(colnames(apis), "Q_mm")

#export data
#write.csv(apish.mm, "Apishapa_Daily_Q_mm.csv")
write.csv(apish.mm, "Crestone_Daily_Q_mm.csv")

#aggregate to monthly (mean and median flow)
month.qmean <- ddply(apish.mm, .(yr, month), summarize, Apish_meanQ = mean(CMS,
na.rm=TRUE))
month.qmedian <- ddply(apish.mm, .(yr, month), summarize, Apish_medianQ = median(CMS,
na.rm=TRUE))

#write.csv(month.qmean, "Apishapa_Mean_Monthly_Q.csv")
#write.csv(month.qmedian, "Apishapa_Median_Monthly_Q.csv")

write.csv(month.qmean, "Crestone_Mean_Monthly_Q.csv")
write.csv(month.qmedian, "Crestone_Median_Monthly_Q.csv")

#convert to mm and sum
month.qcum <- ddply(apish.mm, .(yr, month), summarize, Apishapa_cumQ = sum(Q_mm,
na.rm=TRUE))

#write.csv(month.qcum, "Apishapa_Cumulative_Monthly_Q_mm.csv")
write.csv(month.qcum, "Crestone_Cumulative_Monthly_Q_mm.csv")

annual.qcum <- ddply(apish.mm, .(yr), summarize, Apishapa_cumQ = sum(Q_mm, na.rm=TRUE))

#write.csv(annual.qcum, "Apishapa_Cumulative_Annual_Q_mm.csv")
write.csv(annual.qcum, "Crestone_Cumulative_Annual_Q_mm.csv")

#Dates of Q25, Q50, and Q75 by year
#join annual cumulative values to daily values, accumulated
apish.mmcum <-ddply(apish.mm, .(yr), summarize, Apish_cumsum=cumsum(Q_mm))

apish.mmcumyr <-join(apish.mmcum, annual.qcum, by="yr", type="left")
apish.mmcumyrday <-data.frame(apish.mmcumyr, apish.mm[,c(1,3,5,6)])

#percentile flow
apish.per <-apish.mmcumyrday[,2]/apish.mmcumyrday[,3]

apish.mmper <-data.frame(apish.mmcumyrday[,4],apish.per*100)
names(apish.mmper) <-c("Date", "Percentile")

#write.csv(apish.mmper, "Apishapa_PercentileFlowDate.csv")
write.csv(apish.mmper, "Crestone_PercentileFlowDate.csv")

#percentile flow plot (NOTE:not run for Crestone, as the streamflow regime is different)
#get data note that old file used julian date for y value, add this for plotting?
apishper25 <-
```

```r
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/CO_Data/Cump
ercent_Apishapa_25.csv", colClasses=c("character", "numeric", "factor", "factor", "factor", "numeric"))
apishper50 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/CO_Data/Cump
ercent_Apishapa_50.csv",colClasses=c("character", "numeric", "numeric", "numeric", "character"))
apishper75 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/CO_Data/Cump
ercent_Apishapa_75.csv",colClasses=c("character", "numeric", "numeric", "numeric", "character"))

kyears <-seq(1940,2012, 1)

plot(kyears,apishper25$Julian,main="Apishapa Date of Percentile of Streamflow",
col="blue",yaxt="n",ylab="Month", type="p",pch=19,xlab="Year", ylim=c(0,366))
points(kyears,apishper50$Julian, pch=19)
points(kyears,apishper75$Julian, pch=19, col="red")
axis(2,at=c(1,32,61,92,122,153,183,214,245,275,306,336), lab=c("Jan","", "Mar", "","May", "", "Jul", "",
"Sep", "", "Nov", ""))
legend("topleft",c("25% Total","50% Total", "75% Total"), bty="n",pch=19,
cex=0.85,col=c("blue",1,"red"), inset=0.06)

#-----------------------------------------------------------------
#Checking for autocorrelation in datasets for trend analysis
#precipitation: precip.aoi, mon.precip , yr.precip

#make precip data like format of Q data
mon.precipw <-dcast(mon.precip, yrmon~name, value.var="monP")
yr.precipw <-dcast(yr.precip, year~name, value.var="yrP")
seas.precipw<-dcast(seas.precipd, yrseas~season, value.var="seasP")

#daily autocorrelation
lamaacfd <-acf(precip.aoi[,2], lag.max=5, na.action=na.pass, plot=FALSE)
delacfd <-acf(precip.aoi[,3], lag.max=5, na.action=na.pass, plot=FALSE)
karvacfd <-acf(precip.aoi[,4], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.dayp <-rbind(lamaacfd$acf,delacfd$acf,karvacfd$acf)
rownames(acf.dayp) <-colnames(precip.aoi[2:4])
colnames(acf.dayp) <-c("0", "1", "2", "3", "4", "5")

#monthly autocorrelation
lamaacfm <-acf(mon.precipw[,4], lag.max=5, na.action=na.pass, plot=FALSE)
delacfm <-acf(mon.precipw[,2], lag.max=5, na.action=na.pass, plot=FALSE)
karvacfm <-acf(mon.precipw[,3], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.monp <-rbind(lamaacfm$acf,delacfm$acf,karvacfm$acf)
rownames(acf.monp) <-colnames(mon.precipw[c(4,2,3)])
colnames(acf.monp) <-c("0", "1", "2", "3", "4", "5")

#seasonal autocorrelation
delacfswin <-acf(seas.precipw[,5], lag.max=5, na.action=na.pass, plot=FALSE)
delacfsspr <-acf(seas.precipw[,3], lag.max=5, na.action=na.pass, plot=FALSE)
delacfssum <-acf(seas.precipw[,4], lag.max=5, na.action=na.pass, plot=FALSE)
delacfsfal <-acf(seas.precipw[,2], lag.max=5, na.action=na.pass, plot=FALSE)
```

```
#annual autocorrelation
lamaacfy <-acf(yr.precipw[,4], lag.max=5, na.action=na.pass, plot=FALSE)
delacfy <-acf(yr.precipw[,2], lag.max=5, na.action=na.pass, plot=FALSE)
karvacfy <-acf(yr.precipw[,3], lag.max=5, na.action=na.pass, plot=FALSE)

#collate results
acf.yrp <-rbind(lamaacfy$acf,delacfy$acf,karvacfy$acf)
rownames(acf.yrp) <-colnames(yr.precipw[c(4,2,3)])
colnames(acf.yrp) <-c("0", "1", "2", "3", "4", "5")

#5% intervals
acf.dayp5 <-rbind(lamaacfd$n.used,delacfd$n.used,karvacfd$n.used)
acf.dayp5lev <-2/sqrt(acf.dayp5)

acf.monp5 <-rbind(lamaacfm$n.used,delacfm$n.used,karvacfm$n.used)
acf.monp5lev <-2/sqrt(acf.monp5)

acf.seasp5 <-rbind(delacfsspr$n.used)
acf.seasp5lev <-2/sqrt(acf.seasp5)

acf.yrp5 <-rbind(lamaacfy$n.used,delacfy$n.used,karvacfy$n.used)
acf.yrp5lev <-2/sqrt(acf.yrp5)

#-------------------------------------------------------------------
#Checking for autocorrelation in datasets for trend analysis
#streamflow: apis, apish.mm, month.qmean, month.qmedian, month.qcum (in mm), annual.qmean,
annual.qmedian, annual.qcum (in mm)

#daily autocorrelation cms
apisacfd <-acf(apis[,7], lag.max=5, na.action=na.pass, plot=FALSE)

#daily autocorrelation mm
apisacfdmm <-acf(apish.mm[,8], lag.max=5, na.action=na.pass, plot=FALSE)

#monthly autocorrelation mean
apisacfmonmean <-acf(month.qmean[,3], lag.max=5, na.action=na.pass, plot=FALSE)

#monthly autocorrelation median
apisacfmonmedian <-acf(month.qmedian[,3], lag.max=5, na.action=na.pass, plot=FALSE)

#monthly autocorrelation cumulative
apisacfmonqcum <-acf(month.qcum[,3], lag.max=5, na.action=na.pass, plot=FALSE)

#annual autocorrelation mean
apisacfannmean <-acf(annual.qmean[,2], lag.max=5, na.action=na.pass, plot=FALSE)

#annual autocorrelation median
apisacfannmedian <-acf(annual.qmedian[,2], lag.max=5, na.action=na.pass, plot=FALSE)


#annual autocorrelation cumulative
apisacfannqcum <-acf(annual.qcum[,2], lag.max=5, na.action=na.pass, plot=FALSE)

#5% intervals
acf.dayq5 <-2/sqrt(apisacfd$n.used)
```

```
acf.monp5mean <-2/sqrt(apisacfmonmean$n.used)
acf.monp5median <-2/sqrt(apisacfmonmedian$n.used)
acf.monp5cum <-2/sqrt(apisacfmonqcum$n.used)

acf.yrp5mean <--2/sqrt(apisacfannmean$n.used)
acf.yrp5median <--2/sqrt(apisacfannmedian$n.used )
acf.yrp5cum <--2/sqrt(apisacfannqcum$n.used)

#--------------------------------------------------------------
#trend analyses
#first testing for trend in all time series no adjustments for autocorrelation
#precipitation: precip.aoi, mon.precip , yr.precip, month.precipw, yr.precipw
#streamflow: gages, gages.mm, month.qmean, month.qmedian, month.qcum (in mm), annual.qmean,
annual.qmedian, annual.qcum (in mm)

#checking annual values first, no adjustment for autocorrelation
#precipitation
mk.yrprecip <-apply(yr.precipw[,2:4], 2, MannKendall)

#loop for ts
yr.precipwd <-yr.precipw[,2:4]
ts.yrp <-matrix(nrow=3, ncol=2)
rownames(ts.yrp) <-colnames(yr.precipw[,2:4])
colnames(ts.yrp) <-c("intercept", "slope")

for (i in 1:3){
 year.p <-as.numeric(yr.precipw[,1])
 station <-yr.precipwd[,i]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.ps <-year.p[stationc]
 ts <-zyp.sen(stationcomp~year.ps)
 ts.yrp[i,] <-ts$coefficients
}

#checking seasonal values , no adjustment for autocorrelation
#precipitation
seas.precip <-apply(seas.precipw[,2:5], 2, MannKendall)

#loop for ts
seas.precipwd <-seas.precipw[,2:5]
ts.seasp <-matrix(nrow=4, ncol=2)
rownames(ts.seasp) <-colnames(seas.precipw[,2:5])
colnames(ts.seasp) <-c("intercept", "slope")

for (i in 1:4){
 year.p <-as.numeric(seas.precipw[,1])
 station <-seas.precipwd[,i]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.ps <-year.p[stationc]
 tsseas <-zyp.sen(stationcomp~year.ps)
 ts.seasp[i,] <-tsseas$coefficients
}
```

```
#streamflow
mk.yrqmean <-MannKendall(annual.qmean[,2])
mk.yrqmedian <-MannKendall(annual.qmedian[,2])
mk.yrqcum <-MannKendall(annual.qcum[,2])

#for ts

annual.qmeand <-annual.qmean[,2]
year.q <-as.numeric(annual.qmean[,1])
station <-annual.qmeand
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.qs <-year.q[stationc]
ts <-zyp.sen(stationcomp~year.qs)
ts.yrqmean <-ts$coefficients

annual.qmediand <-annual.qmedian[,2]
year.q <-as.numeric(annual.qmedian[,1])
station <-annual.qmediand
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.qs <-year.q[stationc]
ts <-zyp.sen(stationcomp~year.qs)
ts.yrqmedian <-ts$coefficients

annual.qcumd <-annual.qcum[,2]
year.q <-as.numeric(annual.qcum[,1])
station <-annual.qcumd
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.qs <-year.q[stationc]
ts <-zyp.sen(stationcomp~year.qs)
ts.yrqcum <-ts$coefficients

#-----------------------------------------------
#using TFPW methods on annual data
#reformat data
#precipitation: yr.precipaoi,  yr.precipw, seas.precipw

#seasonal
seas.precipwt <-t(seas.precipw)
seas.precipdf <-data.frame(seas.precipwt[2:5,])
colnames(seas.precipdf) <-seas.precipw[,1]

#run ts analysis
seas.preciptspw <-zyp.trend.dataframe(seas.precipdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)


#annual
yr.precipwt <-t(yr.precipw)
yr.precipdf <-data.frame(yr.precipwt[2:4,])
colnames(yr.precipdf) <-yr.precipw[,1]

#run ts analysis
```

```
yr.preciptspw <-zyp.trend.dataframe(yr.precipdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#check against vector form
yr.precipd <-as.numeric(yr.precipwt[2,])
yr.preciptspw <-zyp.trend.vector(yr.precipd, method="yuepilon", conf.intervals=TRUE)

#streamflow: annual.qmean, annual.qmedian, annual.qcum (in mm)
#mean annual use vector form
yr.qmeand<-as.numeric(annual.qmean[,2])
yr.qmeantspw <-zyp.trend.vector(yr.qmeand, method="yuepilon", conf.intervals=TRUE)

#median annual use vector form
yr.qmedian<-as.numeric(annual.qmedian[,2])
yr.qmediantspw <-zyp.trend.vector(yr.qmedian, method="yuepilon", conf.intervals=TRUE)

#cumulative annual use vector form
yr.qcumd <-as.numeric(annual.qcum[,2])
yr.qcumtspw <-zyp.trend.vector(yr.qcumd, method="yuepilon", conf.intervals=TRUE)
```

C.1.4 Creating Gaps and Testing for Trend with Missing Data

```
#------------------------------------------------------------------------
#  TITLE: Trend Analysis Script for LOR PQ for MN and CO data creating missing data
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2015-03-24
#  LAST REVISION: 2015-05-13
#  DESCRIPTION: This script provides code for analyzing CO met and streamflow data for trends and
making gaps in data
#  PACKAGES REQUIRED: plyr, reshape2, ggplot2
#  VARIABLES/DATA USED: text files of Khangai and CO precipitation and streamflow
#  NAME:
# TYPE:
# COMMENT:
#---------------------------------------------------------------------------
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Missing/")

library(plyr)
library(reshape2)
library(ggplot2)
library(zyp)
library(Kendall)
library(chron)
library(doBy)

#import data
#annual P
co.pannfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/CO_Annual_P.csv", header=TRUE, colClasses=c("numeric", "numeric", "character"))
#select del norte
dn.pannfile <-subset(co.pannfile, Name=="Del")
dn.pann <-dn.pannfile[,1:2]
```

```
names(dn.pann) <-c("Year", "Del_Norte")

mn.pannfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Khangai_Annual_P.csv", header=TRUE, colClasses=c("numeric", "numeric", "character",
"numeric"))

#make wide format
mn.pann <-dcast(mn.pannfile, Year~Name, value.var="AnnP")

#join
pann <-join(dn.pann, mn.pann, by="Year", type="full")

#annual Q mean
co.qannmeanfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Crestone_Mean_Annual_Q.csv", header=TRUE, colClasses="numeric")
mn.qannmeanfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Khangai_Mean_Annual_Q.csv", header=TRUE, colClasses="numeric")
qannmean <-join(co.qannmeanfile, mn.qannmeanfile, by="Year", type="full")

#annual Q median
co.qannmedfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Crestone_Median_Annual_Q.csv",header=TRUE, colClasses="numeric")
mn.qannmedfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Khangai_Median_Annual_Q.csv",header=TRUE, colClasses="numeric")
qannmed <-join(co.qannmedfile, mn.qannmedfile, by="Year", type="full")

#seasonal P
co.pseasfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Seasonal/DelNorte_Seasonal_P.csv", header=TRUE, colClasses=c("numeric", "factor","numeric"))
mn.pseasfile <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Seasonal/Khangai_Seasonal_P.csv", header=TRUE, colClasses= c("numeric", "factor","numeric",
"character", "numeric"))

#make wide format
mn.pseas <-dcast(mn.pseasfile[,1:4], Year+Season~Name, value.var="SeasP")

#join
pseas <-join(co.pseasfile, mn.pseas, by=c("Year", "Season"), type="full")

#normalize values for comparison (Co only)
normdel <-(pann[,2]-min(pann[,2]))/(max(pann[,2])-min(pann[,2]))
normmatdel <-as.matrix(normdel)
normcovdel<-cov(normmatdel, y=NULL, use="complete.obs")

normcres <-(qannmean[,2]-min(qannmean[,2]))/(max(qannmean[,2])-min(qannmean[,2]))
normmatcres <-as.matrix(normcres)
normcovcres<-cov(normmatcres, y=NULL, use="complete.obs")
```

```
#check covariance of p and q- NOTE this is UNSCALED
pmat <-as.matrix(pann)
pcov <-cov(pmat, y=NULL, use="complete.obs")
write.csv(pcov, "Covariance_AnnP.csv")

qmatmean <-as.matrix(qannmean)
qcov <-cov(qmatmean, y=NULL, use="complete.obs")
write.csv(qcov, "Covariance_MeanQ.csv")
#-----------------------------------------------------------------
#make Del Norte record like Erdenemandal record in terms of length and data missingness
pann.dnerd<-ifelse(is.na(pann$Erdenemandal),NA, pann$Del_Norte)
pann.dnerdyr <-data.frame(pann$Year, pann.dnerd)
names(pann.dnerdyr) <-c("Year", "AnnP")

#make Crestone like the Khoid Tamir
qmean.creskt <-ifelse(is.na(qannmean$IkhKT_meanQ),NA, qannmean$Crestone_meanQ)
qmean.cresktyr <-data.frame(qannmean$Year, qmean.creskt)
names(qmean.cresktyr) <-c("Year", "Ann_MeanQ")

#check for autocorr
pann.dnerdacf <-acf(pann.dnerdyr$AnnP,lag.max=5, na.action=na.pass, plot=FALSE)

qmean.cresktacf <-acf(qmean.cresktyr$Ann_MeanQ,lag.max=5, na.action=na.pass, plot=FALSE)

#5% levels
acf.yrp5 <-pann.dnerdacf$n.used
acf.yrp5lev <-2/sqrt(acf.yrp5)

acf.yrp5 <-qmean.cresktacf$n.used
acf.yrp5lev <-2/sqrt(acf.yrp5)

#check for trend
mk.dnerd <- MannKendall(pann.dnerdyr$AnnP)

ts.dnerd <-zyp.sen(AnnP~Year, pann.dnerdyr)
ts.yrpdnerd <-ts.dnerd$coefficients
names(ts.yrpdnerd) <-c("intercept", "slope")

mk.creskt <- MannKendall(qmean.cresktyr$Ann_MeanQ)

ts.creskt <-zyp.sen(Ann_MeanQ~Year, qmean.cresktyr)
ts.yrqcreskt <-ts.creskt$coefficients
names(ts.yrqcreskt) <-c("intercept", "slope")

#--------------------------------------------------------------------
#changing start date and length of record by moving start date forward one year each time then
testing for trend

#pann dim 82, 8
#pseas dim 329, 8
#qannmean dim 67, 8
#qannmed dim 67, 8

#precipitation MK-TS using TFPW changing lengths on Del Norte
pann.dn <-pann[,2]
```

```
names(pann.dn) <-pann[,1]

pann.dnl<-list()

for (i in 1:length(pann.dn)){
pann.dnl[[i]] <-pann.dn[i:length(pann.dn)]
}

pann.rn <-unique(unlist(sapply(pann.dnl, names)))
pann.mx <-matrix(nrow=length(pann.rn), ncol=length(pann.dnl), dimnames=list(pann.rn, NULL))

for (i in seq(length(pann.dnl))) {
  y <-pann.dnl[[i]]
  pann.mx[names(y), i] <-y
}

pann.t <-t(pann.mx)
pann.tdf <-data.frame(pann.t)
colnames(pann.tdf) <-rownames(pann.mx)
rownames(pann.tdf) <-sort(seq(1,length(pann.dnl)), decreasing=TRUE)

annp.tfpw <-zyp.trend.dataframe(pann.tdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annp.tfpw, "Shortening_LOR_AnnP_TRPW.csv")

#for seasonal precip analyze in same way, just divide to 4 different initial matrices by season and
year.
#extract each season
pseas.win <-subset(pseas, Season == "winter")
pseas.spr <-subset(pseas, Season == "spring")
pseas.sum <-subset(pseas, Season == "summer")
pseas.fal <-subset(pseas, Season == "fall")

#scaled
normmatdelw <-as.matrix(pseas.win[,3])
normmatdelw2 <-normmatdelw[2:82]
normdelw <-(normmatdelw2-min(normmatdelw2))/(max(normmatdelw2)-min(normmatdelw2))
normdelmatmatw <-as.matrix(normdelw)
normcovdelw<-cov(normdelmatmatw, y=NULL, use="complete.obs")

normmatdelsp <-as.matrix(pseas.spr[,3])
normdelsp <-(normmatdelsp-min(normmatdelsp))/(max(normmatdelsp)-min(normmatdelsp))
normdelmatmatsp <-as.matrix(normdelsp)
normcovdelsp<-cov(normdelmatmatsp, y=NULL, use="complete.obs")

normmatdelsu <-as.matrix(pseas.sum[,3])
normdelsu <-(normmatdelsu-min(normmatdelsu))/(max(normmatdelsu)-min(normmatdelsu))
normdelmatmatsu <-as.matrix(normdelsu)
normcovdelsu<-cov(normdelmatmatsu, y=NULL, use="complete.obs")

normmatdelf <-as.matrix(pseas.fal[,3])
normdelf <-(normmatdelf-min(normmatdelf))/(max(normmatdelf)-min(normmatdelf))
normdelmatmatf <-as.matrix(normdelf)
normcovdelf<-cov(normdelmatmatf, y=NULL, use="complete.obs")
```

```
#unscaled
pseas.winmat<-as.matrix(pseas.win[,3:9])
pseas.wincov <-cov(pseas.winmat, y=NULL, use="complete.obs")
write.csv(pseas.wincov, "Covariance_PWinter.csv")

pseas.sprmat<-as.matrix(pseas.spr[,3:9])
pseas.sprcov <-cov(pseas.sprmat, y=NULL, use="complete.obs")
write.csv(pseas.sprcov, "Covariance_PSpring.csv")

pseas.summat<-as.matrix(pseas.sum[,3:9])
pseas.sumcov <-cov(pseas.summat, y=NULL, use="complete.obs")
write.csv(pseas.sumcov, "Covariance_PSummer.csv")

pseas.falmat<-as.matrix(pseas.fal[,3:9])
pseas.falcov <-cov(pseas.falmat, y=NULL, use="complete.obs")
write.csv(pseas.falcov, "Covariance_PFall.csv")

#pseason <-pseas.win
#pseason <-pseas.spr
#pseason <-pseas.sum
pseason <-pseas.fal

#pseas.fal <-read.csv()

pseas.dn <-pseason[,3]
names(pseas.dn) <-pseason[,1]

pseas.dnl<-list()

for (i in 1:length(pseas.dn)){
  pseas.dnl[[i]] <-pseas.dn[i:length(pseas.dn)]
}

pseas.rn <-unique(unlist(sapply(pseas.dnl, names)))
pseas.mx <-matrix(nrow=length(pseas.rn), ncol=length(pseas.dnl), dimnames=list(pseas.rn, NULL))

for (i in seq(length(pseas.dnl))) {
  y <-pseas.dnl[[i]]
  pseas.mx[names(y), i] <-y
}

pseas.t <-t(pseas.mx)
pseas.tdf <-data.frame(pseas.t)
colnames(pseas.tdf) <-rownames(pseas.mx)
rownames(pseas.tdf) <-sort(seq(1,length(pseas.dnl)), decreasing=TRUE)

seasp.tfpw <-zyp.trend.dataframe(pseas.tdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#write.csv(seasp.tfpw, "Shortening_LOR_P_Winter_TRPW.csv")
#write.csv(seasp.tfpw, "Shortening_LOR_P_Spring_TRPW.csv")
#write.csv(seasp.tfpw, "Shortening_LOR_P_Summer_TRPW.csv")
write.csv(seasp.tfpw, "Shortening_LOR_P_Fall_TRPW.csv")
```

```
#mean Q MK-TS using TFPW changing lengths on Crestone
qann.cr <-qannmean[,2]
names(qann.cr) <-qannmean[,1]

qann.crl<-list()

for (i in 1:length(qann.cr)){
  qann.crl[[i]] <-qann.cr[i:length(qann.cr)]
}

qann.rn <-unique(unlist(sapply(qann.crl, names)))
qann.mx <-matrix(nrow=length(qann.rn), ncol=length(qann.crl), dimnames=list(qann.rn, NULL))

for (i in seq(length(qann.crl))) {
  y <-qann.crl[[i]]
  qann.mx[names(y), i] <-y
}

qann.t <-t(qann.mx)
qann.tdf <-data.frame(qann.t)
colnames(qann.tdf) <-rownames(qann.mx)
rownames(qann.tdf) <-sort(seq(1,length(qann.crl)), decreasing=TRUE)

annqmean.tfpw <-zyp.trend.dataframe(qann.tdf, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annqmean.tfpw, "Shortening_LOR_AnnQ_Mean_TRPW.csv")

#--------------------------------------------------------------------------
#note that MN Q has up to 5 years in a row missing (IkhKT), so test for what a gap of 1-5 years in a
row, in the data at different times can do
#precipitation MK-TS using TFPW on Del Norte for gaps of 1 year
pann.dn <-pann[,2]
names(pann.dn) <-pann[,1]

#make matrix length and width of record and na 1 year
pann.dnm1 <-matrix(rep(pann.dn, length(pann.dn)) , ncol=length(pann.dn), byrow=TRUE)
rownames(pann.dnm1) <-seq(1:length(pann.dn))
colnames(pann.dnm1) <-names(pann.dn)

for (i in 1:length(pann.dn)){
   pann.dnm1[i,i]<-NA
}

annpg1.tfpw <-zyp.trend.dataframe(pann.dnm1, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annpg1.tfpw, "Gap1_LOR_AnnP_TRPW.csv")

#make matrix length and width of record and na 2 years
pann.dnm2 <-matrix(rep(pann.dn, length(pann.dn)) , ncol=length(pann.dn), byrow=TRUE)
rownames(pann.dnm2) <-seq(1:length(pann.dn))
colnames(pann.dnm2) <-names(pann.dn)

for (i in 1:(length(pann.dn)-1)){
```

```
  pann.dnm2[i,c(i,i+1)]<-NA
}

#truncate last row
pann.dnm2t <-pann.dnm2[1:(nrow(pann.dnm2)-1),]

annpg2.tfpw <-zyp.trend.dataframe(pann.dnm2t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annpg2.tfpw, "Gap2_LOR_AnnP_TRPW.csv")

#make matrix length and width of record and na 3 years
pann.dnm3 <-matrix(rep(pann.dn, length(pann.dn)) , ncol=length(pann.dn), byrow=TRUE)
rownames(pann.dnm3) <-seq(1:length(pann.dn))
colnames(pann.dnm3) <-names(pann.dn)

for (i in 1:(length(pann.dn)-2)){
  pann.dnm3[i,c(i,i+1,i+2)]<-NA
}

#truncate last 2 rows
pann.dnm3t <-pann.dnm3[1:(nrow(pann.dnm3)-2),]

annpg3.tfpw <-zyp.trend.dataframe(pann.dnm3t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annpg3.tfpw, "Gap3_LOR_AnnP_TRPW.csv")

#make matrix length and width of record and na 4 years
pann.dnm4 <-matrix(rep(pann.dn, length(pann.dn)) , ncol=length(pann.dn), byrow=TRUE)
rownames(pann.dnm4) <-seq(1:length(pann.dn))
colnames(pann.dnm4) <-names(pann.dn)

for (i in 1:(length(pann.dn)-3)){
  pann.dnm4[i,c(i,i+1,i+2, i+3)]<-NA
}

#truncate last 3 rows
pann.dnm4t <-pann.dnm4[1:(nrow(pann.dnm4)-3),]

annpg4.tfpw <-zyp.trend.dataframe(pann.dnm4t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annpg4.tfpw, "Gap4_LOR_AnnP_TRPW.csv")

#make matrix length and width of record and na 5 years
pann.dnm5 <-matrix(rep(pann.dn, length(pann.dn)) , ncol=length(pann.dn), byrow=TRUE)
rownames(pann.dnm5) <-seq(1:length(pann.dn))
colnames(pann.dnm5) <-names(pann.dn)

for (i in 1:(length(pann.dn)-4)){
  pann.dnm5[i,c(i,i+1,i+2, i+3, i+4)]<-NA
}

#truncate last 4 rows
```

```
pann.dnm5t <-pann.dnm5[1:(nrow(pann.dnm5)-4),]

annpg5.tfpw <-zyp.trend.dataframe(pann.dnm5t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(annpg5.tfpw, "Gap5_LOR_AnnP_TRPW.csv")

#gap analysis for seasonal P
pseasonal <-pseas.win
#pseasonal <-pseas.spr
#pseasonal <-pseas.sum
pseasonal <-pseas.fal

pseas.fall <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Seasonal/DelNorte_Seasonal_P.csv")
pseas.fal <-subset(pseas.fall, Season=="fall")

#precipitation MK-TS using TFPW on Del Norte for gaps of 1 year
pseason.dn <-pseasonal[,3]
names(pseason.dn) <-pseasonal[,1]

#make matrix length and width of record and na 1 year
pseason.dnm1 <-matrix(rep(pseason.dn, length(pseason.dn)) , ncol=length(pseason.dn),
byrow=TRUE)
rownames(pseason.dnm1) <-seq(1:length(pseason.dn))
colnames(pseason.dnm1) <-names(pseason.dn)

for (i in 1:length(pseason.dn)){
 pseason.dnm1[i,i]<-NA
}

pseasg1.tfpw <-zyp.trend.dataframe(pseason.dnm1, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#write.csv(pseasg1.tfpw, "Gap1_LOR_SeasP_Winter_TRPW.csv")
#write.csv(pseasg1.tfpw, "Gap1_LOR_SeasP_Spring_TRPW.csv")
#write.csv(pseasg1.tfpw, "Gap1_LOR_SeasP_Summer_TRPW.csv")
write.csv(pseasg1.tfpw, "Gap1_LOR_SeasP_Fall_TRPW.csv")

#make matrix length and width of record and na 2 years
#pseasonal <-pseas.win
#pseasonal <-pseas.spr
#pseasonal <-pseas.sum
pseasonal <-pseas.fal

pseason.dn <-pseasonal[,3]
names(pseason.dn) <-pseasonal[,1]

pseason.dnm2 <-matrix(rep(pseason.dn, length(pseason.dn)) , ncol=length(pseason.dn),
byrow=TRUE)
rownames(pseason.dnm2) <-seq(1:length(pseason.dn))
colnames(pseason.dnm2) <-names(pseason.dn)

for (i in 1:(length(pseason.dn)-1)){
```

```
  pseason.dnm2[i,c(i,i+1)]<-NA
}

#truncate last row
pseason.dnm2t <-pseason.dnm2[1:(nrow(pseason.dnm2)-1),]

pseasg2.tfpw <-zyp.trend.dataframe(pseason.dnm2t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#write.csv(pseasg2.tfpw, "Gap2_LOR_SeasP_Winter_TRPW.csv")
#write.csv(pseasg2.tfpw, "Gap2_LOR_SeasP_Spring_TRPW.csv")
#write.csv(pseasg2.tfpw, "Gap2_LOR_SeasP_Summer_TRPW.csv")
write.csv(pseasg2.tfpw, "Gap2_LOR_SeasP_Fall_TRPW.csv")


#make matrix length and width of record and na 3 years
#pseasonal <-pseas.win
#pseasonal <-pseas.spr
#pseasonal <-pseas.sum
pseasonal <-pseas.fal

pseason.dn <-pseasonal[,3]
names(pseason.dn) <-pseasonal[,1]

pseason.dnm3 <-matrix(rep(pseason.dn, length(pseason.dn)) , ncol=length(pseason.dn),
byrow=TRUE)
rownames(pseason.dnm3) <-seq(1:length(pseason.dn))
colnames(pseason.dnm3) <-names(pseason.dn)

for (i in 1:(length(pseason.dn)-2)){
  pseason.dnm3[i,c(i,i+1,i+2)]<-NA
}

#truncate last 2 rows
pseason.dnm3t <-pseason.dnm3[1:(nrow(pseason.dnm3)-2),]

pseasg3.tfpw <-zyp.trend.dataframe(pseason.dnm3t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#write.csv(pseasg3.tfpw, "Gap3_LOR_SeasP_Winter_TRPW.csv")
#write.csv(pseasg3.tfpw, "Gap3_LOR_SeasP_Spring_TRPW.csv")
#write.csv(pseasg3.tfpw, "Gap3_LOR_SeasP_Summer_TRPW.csv")
write.csv(pseasg3.tfpw, "Gap3_LOR_SeasP_Fall_TRPW.csv")

#make matrix length and width of record and na 4 years
#pseasonal <-pseas.win
#pseasonal <-pseas.spr
#pseasonal <-pseas.sum
pseasonal <-pseas.fal
pseason.dn <-pseasonal[,3]
names(pseason.dn) <-pseasonal[,1]

pseason.dnm4 <-matrix(rep(pseason.dn, length(pseason.dn)) , ncol=length(pseason.dn),
byrow=TRUE)
rownames(pseason.dnm4) <-seq(1:length(pseason.dn))
```

```
colnames(pseason.dnm4) <-names(pseason.dn)

for (i in 1:(length(pseason.dn)-3)){
 pseason.dnm4[i,c(i,i+1,i+2, i+3)]<-NA
}

#truncate last 3 rows
pseason.dnm4t <-pseason.dnm4[1:(nrow(pseason.dnm4)-3),]

pseasg4.tfpw <-zyp.trend.dataframe(pseason.dnm4t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#write.csv(pseasg4.tfpw, "Gap4_LOR_SeasP_Winter_TRPW.csv")
#write.csv(pseasg4.tfpw, "Gap4_LOR_SeasP_Spring_TRPW.csv")
#write.csv(pseasg4.tfpw, "Gap4_LOR_SeasP_Summer_TRPW.csv")
write.csv(pseasg4.tfpw, "Gap4_LOR_SeasP_Fall_TRPW.csv")

#make matrix length and width of record and na 5 years
#pseasonal <-pseas.win
#pseasonal <-pseas.spr
#pseasonal <-pseas.sum
pseasonal <-pseas.fal

pseason.dn <-pseasonal[,3]
names(pseason.dn) <-pseasonal[,1]

pseason.dnm5 <-matrix(rep(pseason.dn, length(pseason.dn)) , ncol=length(pseason.dn),
byrow=TRUE)
rownames(pseason.dnm5) <-seq(1:length(pseason.dn))
colnames(pseason.dnm5) <-names(pseason.dn)

for (i in 1:(length(pseason.dn)-4)){
 pseason.dnm5[i,c(i,i+1,i+2, i+3, i+4)]<-NA
}

#truncate last 4 rows
pseason.dnm5t <-pseason.dnm5[1:(nrow(pseason.dnm5)-4),]

pseasg5.tfpw <-zyp.trend.dataframe(pseason.dnm5t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

#write.csv(pseasg5.tfpw, "Gap5_LOR_SeasP_Winter_TRPW.csv")
#write.csv(pseasg5.tfpw, "Gap5_LOR_SeasP_Spring_TRPW.csv")
#write.csv(pseasg5.tfpw, "Gap5_LOR_SeasP_Summer_TRPW.csv")
write.csv(pseasg5.tfpw, "Gap5_LOR_SeasP_Fall_TRPW.csv")

#streamflow MK-TS using TFPW on Crestone Creek for gaps of 1 year
qann.cr <-qannmean[,2]
names(qann.cr) <-qannmean[,1]
#make matrix length and width of record and na 1 year
qann.crm1 <-matrix(rep(qann.cr, length(qann.cr)) , ncol=length(qann.cr), byrow=TRUE)
rownames(qann.crm1) <-seq(1:length(qann.cr))
colnames(qann.crm1) <-names(qann.cr)

for (i in 1:length(qann.cr)){
```

```
  qann.crm1[i,i]<-NA
}

meanqg1.tfpw <-zyp.trend.dataframe(qann.crm1, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(meanqg1.tfpw, "Gap1_LOR_MeanQ_TRPW.csv")

#make matrix length and width of record and na 2 years
qann.crm2 <-matrix(rep(qann.cr, length(qann.cr)) , ncol=length(qann.cr), byrow=TRUE)
rownames(qann.crm2) <-seq(1:length(qann.cr))
colnames(qann.crm2) <-names(qann.cr)

for (i in 1:(length(qann.cr)-1)){
  qann.crm2[i,c(i,i+1)]<-NA
}

#truncate last row
qann.crm2t <-qann.crm2[1:(nrow(qann.crm2)-1),]

meanqg2.tfpw <-zyp.trend.dataframe(qann.crm2t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(meanqg2.tfpw, "Gap2_LOR_MeanQ_TRPW.csv")

#make matrix length and width of record and na 3 years
qann.crm3 <-matrix(rep(qann.cr, length(qann.cr)) , ncol=length(qann.cr), byrow=TRUE)
rownames(qann.crm3) <-seq(1:length(qann.cr))
colnames(qann.crm3) <-names(qann.cr)

for (i in 1:(length(qann.cr)-2)){
  qann.crm3[i,c(i,i+1,i+2)]<-NA
}

#truncate last 2 rows
qann.crm3t <-qann.crm3[1:(nrow(qann.crm3)-2),]

meanqg3.tfpw <-zyp.trend.dataframe(qann.crm3t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(meanqg3.tfpw, "Gap3_LOR_MeanQ_TRPW.csv")

#make matrix length and width of record and na 4 years
qann.crm4 <-matrix(rep(qann.cr, length(qann.cr)) , ncol=length(qann.cr), byrow=TRUE)
rownames(qann.crm4) <-seq(1:length(qann.cr))
colnames(qann.crm4) <-names(qann.cr)

for (i in 1:(length(qann.cr)-3)){
  qann.crm4[i,c(i,i+1,i+2, i+3)]<-NA
}

#truncate last 3 rows
qann.crm4t <-qann.crm4[1:(nrow(qann.crm4)-3),]

meanqg4.tfpw <-zyp.trend.dataframe(qann.crm4t, metadata.cols=0, method="yuepilon",
```

```
conf.intervals=TRUE)

write.csv(meanqg4.tfpw, "Gap4_LOR_MeanQ_TRPW.csv")

#make matrix length and width of record and na 5 years
qann.crm5 <-matrix(rep(qann.cr, length(qann.cr)) , ncol=length(qann.cr), byrow=TRUE)
rownames(qann.crm5) <-seq(1:length(qann.cr))
colnames(qann.crm5) <-names(qann.cr)

for (i in 1:(length(qann.cr)-4)){
 qann.crm5[i,c(i,i+1,i+2, i+3, i+4)]<-NA
}

#truncate last 4 rows
qann.crm5t <-qann.crm5[1:(nrow(qann.crm5)-4),]

meanqg5.tfpw <-zyp.trend.dataframe(qann.crm5t, metadata.cols=0, method="yuepilon",
conf.intervals=TRUE)

write.csv(meanqg5.tfpw, "Gap5_LOR_MeanQ_TRPW.csv")

#----------------------------------------------------------------------------------------------------------
#plots
library(raster)
library(rasterVis)
library(gridExtra)

#figure of "holes" in data
#annual precip
annp.mn <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Khangai_Annual_P.csv")
annp.co <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/CO_Annual_P.csv")
meanq.mn <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Khangai_Mean_Annual_Q.csv")
meanq.co <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Crestone_Mean_Annual_Q.csv")

#aggregate locations
meanq <-join(meanq.co, meanq.mn, by="Year", type="left")
colnames(meanq) <-c("Year", "Crestone", "BaidB", "BayanB", "BayanT", "BogdT", "ErdK", "IkhKT")
annp <-rbind(annp.mn[,1:3], annp.co)

#convert to matrix of data (wide format)
yr.precip <-dcast(annp, Year~Name, value.var="AnnP")
#join in q data
annual <-join(yr.precip, meanq, by="Year", type="left")

#reorder to two matrices for plotting
annual.op <-data.frame(annual$Erdenemandal, annual$Tsetserleg, annual$Baidrag, annual$Galuut,
annual$Bayankhongor, annual$Khoriult,annual$Del)
```

```
annual.opt <-t(annual.op)
colnames(annual.opt) <-annual[,1]
annual.mp <-as.matrix(annual.opt)

annual.oq <-data.frame(annual$ErdK, annual$IkhKT, annual$BaidB, annual$BayanB, annual$BayanT,
annual$BogdT, annual$Crestone)
annual.oqt <-t(annual.oq)
colnames(annual.oqt) <-annual[,1]
annual.mq <-as.matrix(annual.oqt)

#make into rasters
annp.y <-rownames(annual.mp)
annp.x <-colnames(annual.mp)
annp.ext <-extent(1932.5, 2014.5, 0.5, 7.5)

annp.rast <-raster(ncol=82, nrow=7, ext=annp.ext)
values(annp.rast) <-annual.mp

annq.y <-rownames(annual.mq)
annq.x <-colnames(annual.mq)
annq.ext <-extent(1932.5, 2014.5, 0.5, 7.5)

annq.rast <-raster(ncol=82, nrow=7, ext=annq.ext)
values(annq.rast) <-annual.mq

#plot it
pal1 <-colorRampPalette(c("#CCFFCC", "#006600"))
numshades<-6
colors6 <-pal1(numshades)
xlabel1 <-"Year"
mainlabel1 <-"a) Annual Total Precipitation (mm)"
p.theme <-rasterTheme(region=(colors6))
xp.scale <-list(at=c(1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010), labels=c("1940", "1950",
"1960", "1970", "1980", "1990", "2000", "2010"), cex=1.2)
yp.scale <-list(at=c(1,2,3,4,5,6,7),labels=c("Del Norte","Khoriult",
"Bayankhongor","Galuut","Baidrag","Tsetserleg","Erdenemandal"), cex=0.9)
pbreaks <-list(at=c(0,100, 200, 300, 400,500),labels=list(c('0', '100','200','300','400', '500'), cex=1.0),
col=colors6)
plot1 <-levelplot(annp.rast, scales=list(y=yp.scale, x=xp.scale), colorkey=pbreaks, xlab=list(xlabel1,
cex=1.2),
    xscale.components=xscale.raster.subticks, margin=FALSE, par.settings=p.theme,
main=list(mainlabel1, x=0.07,just="left",cex=1.2))

pal2 <-colorRampPalette(c("#99CCFF", "#003399"))
numshades2<-5
colors5 <-pal2(numshades2)
xlabel1 <-"Year"
mainlabel2 <-"b) Annual Mean Streamflow (cms)"
q.theme <-rasterTheme(region=(colors5))
xq.scale <-list(at=c(1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010), labels=c("1940","1950",
"1960", "1970", "1980", "1990", "2000", "2010"), cex=1.2)
yq.scale <-list(at=c(1,2,3,4,5,6,7),labels=c("Crestone Ck., N. near Crestone", "Tuin R. at Bogd", "Tuin R.
at Bayankhongor", "Baidrag R. at Bayanburd","Baidrag R. at Baidrag","Khoid Tamir R. at Ikhtamir",
"Khanui R. at Erdenemandal"), cex=0.85)
qbreaks <-list(at=c(0, 5, 10, 15,20),labels=list(c('0', '5','10','15','20'), cex=1.0), col=colors5)
```

```
plot2 <-levelplot(annq.rast,scales=list(y=yq.scale, x=xq.scale), colorkey=qbreaks, xlab=list(xlabel1,
cex=1.2),
        xscale.components=xscale.raster.subticks, margin=FALSE, par.settings=q.theme,
        main=list(mainlabel2, x=0.07,just="left",cex=1.2))

#grid.arrange(plot1, plot2, ncol=1 )

#--------------------------------
#barplot results of shortening
#for precip
#import csv
annp <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Missing/Shorte
ning/Shortening_LOR_AnnP_TRPW.csv")

#each trend significance level is a facet (non signif, p<0.10, and p<0.05)
#add column with determination
annp$sigf <-rep(100, length(annp$sig))
annp$sigf[annp$sig<0.10] <-10
annp$sigf[annp$sig<0.05] <-5
annp$sigff <-as.factor(annp$sigf)

#boxplot will be with trend values. Boxes are shaded based on trend significance.
annpbox <-ggplot(data=annp, aes(x=x, y=trend, fill=sigff))+
geom_bar(stat="identity", width=1)+ theme_bw()+
  ylab("Trend (mm/year)")+ xlab("Start year")+ ggtitle("b) Significance and magnitude of trend in
annual total precipitation")+
  theme(plot.title=element_text(hjust=0, size=14), axis.text=element_text(size=12),
axis.title=element_text(size=14), legend.position=c(0,0), legend.justification=c(0,0),
legend.title=element_blank(), legend.text=element_text(size=12))+
  scale_x_continuous(breaks=c(seq(1930, 2015, by=5)), labels=c("1930","", "1940","", "1950","",
"1960","", "1970","", "1980", "","1990","", "2000","", "2010",""))+
  scale_y_continuous(breaks=c(-16,-14, -12, -10, -8, -6, -4, -2, 0, 2),labels=c("-16","-14", "-12", "-10", "-
8", "-6", "-4", "-2", "0", "2"))+
  scale_fill_manual(values=c("red", "orange", "grey88"), labels=c("significant (p<0.05)","significant
(p<0.10)","not significant"))+
  geom_abline(slope=0, intercept=0)

#plotting time series points
#import timseries of annual precip
annpts <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/CO_Annual_P.csv")
annp.dn <-subset(annpts, Name=="Del")

#insert trend lines for a few examples of significant trends by calculation of trend and y intercept
using standard TS method
#1979-2014 -2.1 mm/yr
#subset to period of interest
annp.dnp <-subset(annp.dn, annp.dn$Year>1978)

year.p <-as.numeric(annp.dnp[,1])
station <-annp.dnp[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
```

```
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)
#Intercept   year.ps
#4468.690    -2.108

annp.g <-ggplot(annp.dn, aes(Year, AnnP))+ geom_point()+
  theme_bw()+theme(axis.text=element_text(size=12), axis.title.x=element_text(size=14),
axis.title.y=element_text(size=14), plot.title=element_text(hjust=0,
size=14),legend.title=element_blank(),legend.position=c(0,0), legend.justification=c(0,0),
legend.text=element_text(size=12), legend.key=element_blank())+
  ggtitle("a) Annual total precipitation Del Norte 1933-2014")+ ylab("Annual precipitation (mm)")+
xlab("Year")+
  scale_x_continuous(breaks=c(seq(1930, 2015, by=5)), labels=c("","1935", "","1945", "","1955",
"","1965", "","1975", "", "1985","","1995", "","2005", "","2015"))+
  expand_limits(y=c(100,400))+ geom_segment(aes(x=1979, y=296.958, xend=2014, yend=223.178,
color="1979-2014 Significant trend (p<0.05)"),linetype=1)+
  scale_colour_manual(values=c("black"))

grid.arrange(annp.g, annpbox)

#for streamflow
#import csv
annq <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Missing/Shorte
ning/Shortening_LOR_AnnQ_Mean_TRPW.csv")

#each trend significance level is a facet (non signif, p<0.10, and p<0.05)
#add column with determination
annq$sigf <-rep(100, length(annq$sig))
annq$sigf[annq$sig<0.10] <-10
annq$sigf[annq$sig<0.05] <-5
annq$sigff <-as.factor(annq$sigf)

#boxplot will be with trend values. Boxes are shaded based on trend significance.
annqbox <-ggplot(data=annq, aes(x=x, y=trend, fill=sigff))+
  geom_bar(stat="identity", width=1)+ theme_bw()+
  ylab("Trend (cms/year)")+ xlab("Start year")+ ggtitle("b) Significance and magnitude of trend in
annual mean streamflow")+
  theme(plot.title=element_text(hjust=0, size=14), axis.text=element_text(size=12),
axis.title=element_text(size=14), legend.position=c(0,0), legend.justification=c(0,0),
legend.title=element_blank(), legend.text=element_text(size=12))+
  scale_x_continuous(breaks=c(seq(1945, 2015, by=5)), labels=c("", "1950","", "1960","", "1970","",
"1980", "","1990","", "2000","", "2010",""))+
  scale_y_continuous(breaks=c(-0.020,-0.015, -0.010, -0.005, 0, 0.005),labels=c("-0.020","-0.015", "-
0.010", "-0.005", "0", "0.005"))+
  scale_fill_manual(values=c("red", "orange", "grey88"), labels=c("significant (p<0.05)", "significant
(p<0.10)", "not significant"))+
  geom_abline(slope=0, intercept=0)

#plotting time series points
#import timseries of annual precip
annqts <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/Crestone_Mean_Annual_Q.csv")
```

```
#insert trend lines for a few examples of significant trends by calculation of trend and y intercept
using standard TS method
#1979-2014 -2.1 mm/yr
#subset to period of interest
annqtss <-subset(annqts, annqts$Year>1978)

year.q <-as.numeric(annqtss[,1])
station <-annqtss[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)
#Intercept    year.ps
#10.017207    -0.004847

annq.g <-ggplot(annqts, aes(Year, Crestone_meanQ))+ geom_point()+
  theme_bw()+theme(axis.text=element_text(size=12), axis.title.x=element_text(size=14),
axis.title.y=element_text(size=14), plot.title=element_text(hjust=0,
size=14),legend.title=element_blank(),legend.position=c(0,0), legend.justification=c(0,0),
legend.text=element_text(size=12), legend.key=element_blank())+
  ggtitle("a) Annual mean streamflow Crestone Creek 1948-2014")+ ylab("Mean streamflow (cms)")+
xlab("Year")+
  scale_x_continuous(breaks=c(seq(1945, 2015, by=5)), labels=c("1945", "","1955", "","1965",
"","1975", "", "1985","","1995", "","2005", "","2015"))+
  expand_limits(y=c(0,0.6))+ geom_segment(aes(x=1979, y=0.424994, xend=2014, yend=0.255349,
color="1979-2014 Significant trend (p<0.05)"),linetype=1)+
  scale_colour_manual(values=c("black"))

grid.arrange(annq.g, annqbox)

#-----------------------------------------------------------------------------------------
#raster results of shortening
#import csv
annp <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Missing/Shorte
ning/Shortening_LOR_AnnP_TRPW.csv")

#set up for fill of raster, each row has slope value for LOR and NAs for rest
mat.dat <-matrix(nrow=73, ncol=82, NA)

for (i in 1:(nrow(mat.dat)-1)){
  mat.datfill <-c(rep(NA,((i-1)+1)), rep(annp$trend[i+1],ncol(mat.dat)-i))
  mat.dat[i,] <-mat.datfill
}

#trim last line and add first line to matrix
mat.dat1 <-mat.dat[1:72,]

mat.datr1 <-rep(annp$trend[1], 82)
mat.dat2 <-rbind(mat.datr1, mat.dat1)
colnames(mat.dat2) <-seq(1933, 2014, 1)
rownames(mat.dat2)<-seq(82,10, -1)

#make raster
annp.y <-rownames(mat.dat2)
```

```
annp.x <-colnames(mat.dat2)
annp.ext <-extent(1932.5, 2014.5, 9.5, 82.5)

annp.rast <-raster(ncol=82, nrow=73, ext=annp.ext)
values(annp.rast) <-mat.dat2

#plot it
x.scale <- list(cex=1.2, alternating=1)
y.scale <- list(cex=1.2, alternating=1)
sig.themep <-rasterTheme(region=rev(brewer.pal(8, "Reds")))
slope.plot <-levelplot(annp.rast,colorkey=list(at=c(-12,-10,-8,-6,-4,-2,0,2),labels=list(c('-12', '-10','-
8','-6','-4', '-2','0', '2')), cex=1.5, col=rev(brewer.pal(8, "Reds"))), scales=list(x=x.scale, y=y.scale),
 xlab=list("Start Year", cex=1.2), xscale.components=xscale.raster.subticks,
yscale.components=yscale.raster.subticks, ylab=list("Length of Record in Years", cex=1.2),
margin=FALSE, par.settings=sig.themep, main=list( "Trend in Annual Precipitation", cex=1.3))

#add overlay of points to denote which cells are significant
#extract rows from matrix that are significant based off sig value in original dataframe
annp.sigy <-ifelse(annp$sig<0.05, annp$y, NA)
annp.sigx <-ifelse(annp$sig<0.05, annp$x, NA)

annp.sigr <-data.frame(annp.sigx, annp.sigy)
colnames(annp.sigr) <-c("x", "y")

#remove na rows
annp.sigrd <-annp.sigr[complete.cases(annp.sigr),]
coordinates(annp.sigrd) <-~x+y

slope.plot +layer(sp.points(annp.sigrd, cex=1, pch=1, col=1))
#plots don't quite line up prob from using whole numbers in the extent

#make matrix of only those lines that are significant
#set up for fill of raster, only those with sig<0.05 for each row has slope value for LOR,  NAs for rest
datelist=list(seq(1979, 2014, 1), seq(1980,2014, 1), seq(1981, 2014, 1),seq(1982, 2014, 1),
seq(1983, 2014, 1), seq(1984, 2014, 1), seq(1989, 2014, 1), seq(1996, 2014, 1), seq(2003, 2014, 1),
(2004, 2014, 1), (2005, 2014, 1)))

mat.datna <-matrix(nrow=73, ncol=82, NA)

for (i in 0:nrow(mat.datna)){
 if(annp$sig[(i+1)]<0.05)
   mat.datfillna <-c(rep(NA,(i+1)), rep(annp$trend[i+1],ncol(mat.datna)-(i+1)))
 else
   mat.datfillna <-rep(NA,(ncol(mat.datna)))
 mat.datna[i,] <-mat.datfillna
}

#trim last line and add first line to matrix
mat.datna1 <-mat.datna[1:72,]
mat.datnar1 <-ifelse(annp$sig[1]<0.05, rep(annp$trend[1], 82), rep(NA,82))
mat.datna2 <-rbind(mat.datnar1, mat.datna1)
colnames(mat.datna2) <-seq(1933, 2014, 1)
rownames(mat.datna2)<-seq(82,10, -1)

#convert to raster
```

```
annp.y2 <-rownames(mat.datna2)
annp.x2 <-colnames(mat.datna2)
annp.ext2 <-extent(1932.5, 2014.5, 9.5, 82.5)

annp.rast2 <-raster(ncol=82, nrow=73, ext=annp.ext2)
values(annp.rast2) <-mat.datna2

#outline a raster cells by a polygon
#annp.sig1box <-rasterToPolygons(annp.rast2)
annp.sig1box <-rasterToPoints(annp.rast2)
#convert to matrix of points
annp.sigpts <-as.data.frame(annp.sig1box)
coordinates(annp.sigpts) <-~x+y
slope.plot+layer(sp.points(annp.sigpts,cex=1, pch="*", col=1))

#-------------------------------------------------------------------------------------------
#plotting time series points
#library(ggplot2)

#import timseries of annual precip
annp <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Annual/CO_Annual_P.csv")
annp.dn <-subset(annp, Name=="Del")

#insert trend lines for a few examples of significant trends by calculation of trend and y intercept
using standard TS method
#1979-2014 -2.1 mm/yr
#subset to period of interest
annp.dnp <-subset(annp.dn, annp.dn$Year>1978)

 year.p <-as.numeric(annp.dnp[,1])
 station <-annp.dnp[,2]
 stationc<-which(!is.na(station))
 stationcomp <-station[stationc]
 year.ps <-year.p[stationc]
 ts <-zyp.sen(stationcomp~year.ps)
#Intercept    year.ps
#4468.690     -2.108

#1989-2014 -3.1 mm/yr
annp.dnp <-subset(annp.dn, annp.dn$Year>1988)

year.p <-as.numeric(annp.dnp[,1])
station <-annp.dnp[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)
#Intercept    year.ps
#6494.18      -3.12

#1996-2014 -5.5 mm/yr
annp.dnp <-subset(annp.dn, annp.dn$Year>1995)
```

```
year.p <-as.numeric(annp.dnp[,1])
station <-annp.dnp[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)
#Intercept    year.ps
#11223.234    -5.479

annp.g <-ggplot(annp.dn, aes(Year, AnnP))
annp.g +geom_point() +theme_bw()+theme(axis.text=element_text(size=12),
axis.title.x=element_text(size=14), axis.title.y=element_text(size=14),
plot.title=element_text(hjust=0))+
  labs(title="a) Annual Total Precipitation Del Norte 1933-2014", y="Annual Precipitation (mm)",
x="Year")+
  scale_x_continuous(breaks=c(1930,1940, 1950,1960, 1970,1980, 1990,2000, 2010),
labels=c("1930","1940", "1950","1960", "1970", "1980", "1990","2000", "2010"))+
  expand_limits(y=c(100,400))+
  geom_segment(aes(x=1979, y=296.958, xend=2014, yend=223.178),linetype=2)+
  geom_segment(aes(x=1989, y=288.5, xend=2014, yend=210.5), linetype=3)+
  geom_segment(x=1996, y= 287.15, xend=2014, yend=188.528)

#trends for streamflow are lower slope, harder to see on graph?

#-----------------------------------------------------------------------------------------------------------
#gaps
#raster plot of gaps in data? 5 year as example
#import csv
fallp <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Missing/Gap/G
ap5_LOR_SeasP_Fall_Invert_TRPW.csv")

#column for center of 5 year gap will be y axis, x axis will be year in the series, color will be trend
level, significance will be overlay

#set up for fill of raster, each row has slope value for LOR
mat.dat <-matrix(nrow=78, ncol=82, NA)

for (i in 1:(nrow(mat.dat)-1)){
  mat.datfill <-c(rep(fallp$trend[i],ncol(mat.dat)))
  mat.dat[i,] <-mat.datfill
}

#add last line to matrix
mat.dat1 <-mat.dat[1:77,]

mat.datr1 <-rep(fallp$trend[78], 82)
mat.dat2 <-rbind(mat.dat1, mat.datr1)
colnames(mat.dat2) <-seq(1933, 2014, 1)
rownames(mat.dat2)<-seq(2012,1935, -1)

#insert gaps as with analysis
mat.datn <-matrix(nrow=nrow(mat.dat2), ncol=ncol(mat.dat2), NA)

for (i in 1:78){
```

```
 mat.datr <-mat.dat2[i,]
 mat.datr[c(79-i,80-i,81-i, 82-i, 83-i)]<-NA
 mat.datn[i,] <-mat.datr
}

#make raster
fallp.y <-rownames(mat.datn)
fallp.x <-colnames(mat.datn)
fallp.ext <-extent(1932.5, 2014.5, 1934.5, 2012.5)

fallp.rast <-raster(ncol=82, nrow=78, ext=fallp.ext)
values(fallp.rast) <-mat.datn

#plot it
x.scale <- list(cex=1.1, at=c(1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010), labels=c("1940",
"1950", "1960", "1970", "1980", "1990", "2000", "2010"))
y.scale <- list(cex=1.1, at=c(1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010), labels=c("1940",
"1950", "1960", "1970", "1980", "1990", "2000", "2010"))
sig.themep <-rasterTheme(region=(brewer.pal(6, "Blues")))
slope.plot <-levelplot(fallp.rast,colorkey=list(at=c(0.10,0.15, 0.20,0.25, 0.30,
0.35),labels=list(c("0.10", "0.15", "0.20", "0.25", "0.30", "0.35"), cex=1), col=(brewer.pal(6, "Blues"))),
scales=list(x=x.scale, y=y.scale),
     ylab=list("Center year of 5-year gap", cex=1.2), xscale.components=xscale.raster.subticks,
     yscale.components=yscale.raster.subticks, xlab=list("Year", cex=1.2), margin=FALSE,
par.settings=sig.themep,main=list( "b) Trend in fall total precipitation with 5-year data gaps",
cex=1.2, x=0.1, just="left", font=1),
     panel = function(...) {
      panel.fill(col = "black")
      panel.levelplot(...)
     })

     trellis.focus("legend", side="right", clip.off=TRUE, highlight=FALSE)
     grid.text("mm/yr", 0.2, 1.03) #hjust=0.5, vjust=1.3
     trellis.unfocus()

#add overlay of points to denote which cells are significant
#set up for fill of raster, only those with sig<0.05 for each row has slope value for LOR
datelist=list(seq(1979, 2014, 1), seq(1980,2014, 1), seq(1981, 2014, 1),seq(1982, 2014, 1),
seq(1983, 2014, 1), seq(1984, 2014, 1), seq(1989, 2014, 1), seq(1996, 2014, 1), seq(2003, 2014, 1),
seq(2004, 2014, 1), seq(2005, 2014, 1))

mat.datna <-matrix(nrow=78, ncol=82, NA)
rownames(mat.datna)<-rownames(mat.dat2)
colnames(mat.datna)<-colnames(mat.dat2)

for (i in 1:nrow(mat.datna)){
 if(fallp$sig[i]<0.05)
  mat.datfillna <-rep(fallp$trend[i],ncol(mat.datna))
 else
  mat.datfillna <-rep(NA,(ncol(mat.datna)))
 mat.datna[i,] <-mat.datfillna
}

#convert to raster
fallp.y2 <-rownames(mat.dat2)
```

```
fallp.x2 <-colnames(mat.dat2)
fallp.ext2 <-extent(1932.5, 2014.5, 1934.5, 2012.5)

fallp.rast2 <-raster(ncol=82, nrow=78, ext=fallp.ext2)
values(fallp.rast2) <-mat.datna

fallp.sig1box <-rasterToPoints(fallp.rast2)

#convert to matrix of points
fallp.sigpts <-as.data.frame(fallp.sig1box)
coordinates(fallp.sigpts) <-~x+y
slope.plot+layer(sp.points(fallp.sigpts,cex=1, pch="*", col=1))

trellis.focus("legend", side="right", clip.off=TRUE, highlight=FALSE)
grid.text("mm/yr", 0.2, 1.03) #hjust=0.5, vjust=1.3
trellis.unfocus()

#----------------------------------------------------------------

#looking at gaps of 1 year, three years, and 5 years, all centered on 1960.
#1 year gap:
#1960=NA then run trend analysis stripping this gap to get y intercept can prob plot using abline
#3 year gap:
#1959,60,61=NA
#5 year gap:
#1958,59, 60, 61, 62=NA

#import timseries of seasonal precip
fallp <-
read.csv("/Users/niah/Documents/CSU/Mongolia/LOR/LOR_PQ/Trend_Analyses_R/Aggregate_Files
/Seasonal/DelNorte_Seasonal_P.csv")
fallp.dn <-subset(fallp, Season=="fall")

#1960
fallp1 <-fallp.dn[,c(1,3)]
fallp1[28,2]<-NA

year.p <-as.numeric(fallp1[,1])
station <-fallp1[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)
#Intercept    year.ps
#-358.6893     0.2111

#1959-61
fallp3 <-fallp.dn[,c(1,3)]
fallp3[27:29,2]<-NA
year.p <-as.numeric(fallp3[,1])
station <-fallp3[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)
```

```
#Intercept    year.ps
#-420.3507    0.2419

#1958-62
fallp5 <-fallp.dn[,c(1,3)]
fallp5[26:30,2]<-NA

year.p <-as.numeric(fallp5[,1])
station <-fallp5[,2]
stationc<-which(!is.na(station))
stationcomp <-station[stationc]
year.ps <-year.p[stationc]
ts <-zyp.sen(stationcomp~year.ps)

#Intercept    year.ps
#-416.3507    0.2397

x=0.07,just="left"

#dot plot
fallp.xy <-(fallp.dn[,c(1,3)])
fallp.x <-as.numeric(fallp.dn[,1])
fallp.xy <-data.frame(fallp.x, fallp.dn[,3])
colnames(fallp.xy) <-c("Year", "Del_Norte")
fallp.g <-ggplot(fallp.xy, aes(Year, Del_Norte))+geom_point()+
  theme_bw()+theme(axis.text=element_text(size=12), axis.title.x=element_text(size=14),
axis.title.y=element_text(size=14), plot.title=element_text(hjust=0,
size=14),legend.title=element_blank(),legend.position=c(0,0), legend.justification=c(0,0),
legend.text=element_text(size=12), legend.key=element_blank())+
  ggtitle("a) Fall total precipitation Del Norte 1933-2014")+ ylab("Fall precipitation (mm)")+
xlab("Year")+
  scale_x_continuous(breaks=c(seq(1930, 2015, by=5)), labels=c("","1935", "","1945", "","1955",
"","1965", "","1975", "", "1985","","1995", "","2005", "","2015"))+
  expand_limits(y=c(0,130))+scale_y_continuous(breaks=c(0,25,50,75,100,125),
labels=c("0","25","50","75","100","125"))+
  geom_segment(aes(x=1933, y=49.367, xend=2014, yend=66.4661,color="Significant trend with 5-
year gap centered on 1960 (p<0.05)"),linetype=1)+
  scale_colour_manual(values=c("black"))
```

REFERENCES

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www.R-project.org/*.

APPENDIX D

D.1 Chapter 5 R Code

The following code was used for creating the climate maps for Chapter 5 and works with R software, version 3.1.3 (2015-03-09), "Smooth Sidewalk" and previous versions (R Team, 2015). While any errors or omissions are mine, please be aware that no warranty of any kind is offered with this code and users need to carefully proof and/or modify the code to suit their needs.

D.1.1 Extracting Precipitation Data from Global Precipitation Climatology Centre Grids

```
#-------------------------------------------------------------------------
#  TITLE: Extracting Climate Grids Script
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2014-02-12
#  LAST REVISION: 2015-10-14
#  DESCRIPTION: This script provides code for extracting climate data from GPCC grids.
#  PACKAGES REQUIRED: netcdf, raster
#  VARIABLES/DATA USED:
#NAME:
# TYPE:
# COMMENT:
#-----------------------------------------------------------------------------------------------
## For GPCC datafiles
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPCC_
Extract_R/")

#libraries
library(raster)
library(rgdal)
library(ncdf4)
library(RColorBrewer)
library(rgeos)
library(maptools)

# create raster brick directly from netCDF file
gpccfile <- "full_data_v6_precip_05.nc"

variname <-"p"
vartp <- brick(gpccfile, varname=variname)

#plot to visualize file
plot(vartp, 1)
```

```r
image(vartp,col=rev(brewer.pal(9,"RdBu")))

#set extent for analysis
vartp.extent <- extent(97.25, 104.75, 42.25, 49.75)
vartp.aoi <- crop(vartp, vartp.extent)

#subset brick to last 50 years of data, 1976-2010
vartp.aois <- subset(vartp.aoi, subset=901:1320)

#plot to visualize the new extent
plot(vartp.aois, 1)
image(vartp.aois,col=rev(brewer.pal(9,"RdBu")))

#write raster for other analysis in raster format
#vartp.aoir <-writeRaster(vartp.aois, filename="gpcc_pre_aoi.grd", bandorder='BIL', overwrite=TRUE)

#---------------------------------------------------------
#import basin locations
tuin <-
shapefile("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPCC_Extract_R/Basins_GIS/Tuin_Bayankhongor.shp")
baid <-
shapefile("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPCC_Extract_R/Basins_GIS/Baidrag_Bayanburd.shp")
khan <-
shapefile("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPCC_Extract_R/Basins_GIS/Khanui_Erdenemandal.shp")
khoid <-
shapefile("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPCC_Extract_R/Basins_GIS/KhoidTamir_Ikhtamir.shp")
bogd <-
shapefile("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPCC_Extract_R/Basins_GIS/Tuin_Bogd.shp")

#extract weighted mean precip each month from basin area
basins <-list(baid,tuin, khan, khoid)
basin.pre <-matrix(ncol=length(basins), nrow=420, NA)
for (i in 1:length(basins)) {
 basin.p <-extract(vartp.aois, basins[[i]], weights=TRUE, fun=mean)
 basin.pre[,i] <-basin.p
}
rownames(basin.pre) <-colnames(basin.p)
colnames(basin.pre) <-c("BayanB","BayanT", "ErdK", "IkhKT")

write.csv(basin.pre, "GPCC_Mean_P_1976_2010_Basins.csv")

#extract mean precip over basin for BogdT
basins <-list(bogd)
basin.pre <-matrix(ncol=length(basins), nrow=420, NA)
for (i in 1:length(basins)) {
 basin.p <-extract(vartp.aois, basins[[i]], weights=TRUE, fun=mean)
 basin.pre[,i] <-basin.p
}
rownames(basin.pre) <-colnames(basin.p)
colnames(basin.pre) <-c("Bogd")
```

```
write.csv(basin.pre, "GPCC_Mean_P_1976_2010_Bogd.csv")

#-------------------------------------------------------------------
#extract centroid coordinate of each basin
basins <-list(baid,tuin, khan, khoid)
basin.cent <-matrix(nrow=length(basins), ncol=2, NA)
basin.centpt <-list()
for (i in 1:length(basins)) {
 basin.c <-gCentroid(basins[[i]])
 basin.ext<-basin.c@coords
 basin.cent[i,] <-basin.ext
 basin.centpt[[i] ]<-basin.c
}
colnames(basin.cent) <-c("long","lat")
rownames(basin.cent) <-c("BayanB","BayanT", "ErdK", "IkhKT")
write.csv(basin.cent, "Basin_Centroids.csv")

#--------------------------------------------------------------------
#calculate distance from met stations to basin centroid
library(maptools)

#metlocs file has all stations used in previous work, will need to select out stations of interest
metlocs <-
shapefile("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPC
C_Extract_R/Basins_GIS/metloc3aimag.shp")

#rename Tsetserleg soum to avoid confusion
metlocs$Name[18]="Tsetserleg_s"

#testing
plot(vartp.aois, 1)
plot(metlocs, pch=19,add=TRUE)
plot(basin.centpt[[1]], pch=19,col="red", add=TRUE)

#calculate distance from met stations to centroid locations

stat.dist<-matrix(ncol=4, nrow=40, NA)
for (i in 1:4){
 dist <-pointDistance(metlocs, basin.centpt[[i]], lonlat=TRUE)
 distkm <-dist/1000
 stat.dist[,i] <-distkm
}
colnames(stat.dist) <-c("BayanB","BayanT", "ErdK", "IkhKT")
rownames(stat.dist) <-metlocs$Name

#select stations of interest
#Khoriult, Bayankhongor, Galuut, Baidrag, Tsetserleg, and Erdenemandal

stat.distaoi <-stat.dist[c(2,17,25,31,34,36),]


#export distances to CSV
write.csv(stat.distaoi, "Station_Distances_to_Basin_Centroid.csv")
```

D.1.2 Khangai Mountain Region Precipitation and Streamflow Analysis

```
#-----------------------------------------------------------------------
#  TITLE: PQ_Khangai_for_TR
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2015-08-30
#  LAST REVISION: 2015-11-03
#  DESCRIPTION: This script provides code for analyzing Khangai streamflow and precipitation, station and
grid based.
#  PACKAGES REQUIRED:
#  VARIABLES/DATA USED:
#  NAME:
# TYPE:
# COMMENT:
#------------------------------------------------------------------------
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/")

#libraries
library(zoo)
library(reshape2)
library(plyr)
library(mice)
library(ggplot2)
library(lmtest)
library(moments)
library(MASS)

#---------------------------------------------------------------------------------------------------------
#Station-based daily P and Q data distribution
#----------
#for P use Erdenemandal, Tsetserleg, Tariat, Baidrag, Galuut and Bayankhongor
#for Q use Erdenemandal, Ikhtamir, Bayanburd, and Bayankhongor
#streamflow
#----------------
#streamflow
khq.d <-read.csv("Khangai_Daily_Q_1976_2012.csv")

#add yrmonday col
khq.d$yrmonday <-paste(khq.d$Year,sprintf("%02d", khq.d$Mon), sprintf("%02d", khq.d$Day), sep="-")

khq.day<-khq.d[,c(9,5:8,2:4)]

#check distribution of data
qqnorm(khq.day[,2])
qqline(khq.day[,2])
qqnorm(khq.day[,3])
qqline(khq.day[,3])
qqnorm(khq.day[,4])
qqline(khq.day[,4])
qqnorm(khq.day[,5])
qqline(khq.day[,5])

#check autocorrelation of data
```

```
qacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
  qcorr <-acf(khq.day[,i+1], type="correlation", na.action=na.pass)
  qacf[i,] <-qcorr$acf[[2]]
}
colnames(qacf) <-"lag1"
rownames(qacf) <-colnames(khq.day[,2:5])

#High levels of autocorr for all basins

#check for skewness and kurtosis
qskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qsk <-skewness(khq.day[,i+1], na.rm=TRUE)
  qskew[i,]<-qsk
}
colnames(qskew) <-c("skewness")
rownames(qskew) <-colnames(khq.day[,2:5])

qkurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qku <-kurtosis(khq.day[,i+1], na.rm=TRUE)
  qkurt[i,]<-qku
}
colnames(qkurt) <-c("kurtosis")
rownames(qkurt) <-colnames(khq.day[,2:5])

#check distribution of residuals
time <-seq(1:length(khq.day[,2]))
bayanb <-khq.day[,2]
bayanb.lm <-lm(bayanb~time, na.action=na.omit)
time.r <-seq(1:length(bayanb.lm$residuals))
plot(time.r, bayanb.lm$residuals, pch=19)
abline(0,0)

bayant <-khq.day[,3]
bayant.lm <-lm(bayant~time, na.action=na.omit)
time.r2 <-seq(1:length(bayant.lm$residuals))
plot(time.r2, bayant.lm$residuals, pch=19)
abline(0,0)

erdk <-khq.day[,4]
erdk.lm <-lm(erdk~time, na.action=na.omit)
time.r4 <-seq(1:length(erdk.lm$residuals))
plot(time.r4, erdk.lm$residuals, pch=19)
abline(0,0)

ikhkt <-khq.day[,3]
ikhkt.lm <-lm(ikhkt~time, na.action=na.omit)
time.r6 <-seq(1:length(ikhkt.lm$residuals))
plot(time.r6, ikhkt.lm$residuals, pch=19)
abline(0,0)

#try transform using sqrt first
khq.daysqrt<- sqrt(khq.day[,2:5])
```

```
#autocorr in transformed variables
qacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
  qcorr <-acf(khq.daysqrt[,i], type="correlation", na.action=na.pass)
  qacf[i,] <-qcorr$acf[[2]]
}
colnames(qacf) <-"lag1"
rownames(qacf) <-colnames(khq.daysqrt[,1:4])

#transformed with square root
qskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qsk <-skewness(khq.daysqrt[,i], na.rm=TRUE)
  qskew[i,]<-qsk
}
colnames(qskew) <-c("skewness")
rownames(qskew) <-colnames(khq.daysqrt[,1:4])

qkurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qku <-kurtosis(khq.daysqrt[,i], na.rm=TRUE)
  qkurt[i,]<-qku
}
colnames(qkurt) <-c("kurtosis")
rownames(qkurt) <-colnames(khq.daysqrt[,1:4])

#check normalcy
qqnorm(khq.daysqrt[,1])
qqline(khq.daysqrt[,1])
qqnorm(khq.daysqrt[,2])
qqline(khq.daysqrt[,2])
qqnorm(khq.daysqrt[,3])
qqline(khq.daysqrt[,3])
qqnorm(khq.daysqrt[,4])
qqline(khq.daysqrt[,4])

#check residuals
bayanb.sqrt <-khq.daysqrt[,1]
bayanb.sqrtlm <-lm(bayanb.sqrt~time, na.action=na.omit)
time.r1 <-seq(1:length(bayanb.sqrtlm$residuals))
plot(time.r1, bayanb.sqrtlm$residuals, pch=19)
abline(0,0)

#plot of absolute residuals
plot(time.r1, abs(bayanb.sqrtlm$residuals), pch=19)

bayant.sqrt <-khq.daysqrt[,2]
bayant.sqrtlm <-lm(bayant.sqrt~time, na.action=na.omit)
time.r3 <-seq(1:length(bayant.sqrtlm$residuals))
plot(time.r3, bayant.sqrtlm$residuals, pch=19)
abline(0,0)

erdk.sqrt <-khq.daysqrt[,3]
erdk.sqrtlm <-lm(erdk.sqrt~time, na.action=na.omit)
time.r5 <-seq(1:length(erdk.sqrtlm$residuals))
```

```
plot(time.r5, erdk.sqrtlm$residuals, pch=19)
abline(0,0)

ikhkt.sqrt <-khq.daysqrt[,2]
ikhkt.sqrtlm <-lm(ikhkt.sqrt~time, na.action=na.omit)
time.r7 <-seq(1:length(ikhkt.sqrtlm$residuals))
plot(time.r7, ikhkt.sqrtlm$residuals, pch=19)
abline(0,0)

#transformed with log +0.1
khq.daylog <-log(khq.day[,2:5]+0.1)

#autocorr in transformed variables
qacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
  qcorr <-acf(khq.daylog[,i], type="correlation", na.action=na.pass)
  qacf[i,] <-qcorr$acf[[2]]
}
colnames(qacf) <-"lag1"
rownames(qacf) <-colnames(khq.daylog[,1:4])

#skewness and kurtosis
qskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qsk <-skewness(khq.daylog[,i], na.rm=TRUE)
  qskew[i,]<-qsk
}
colnames(qskew) <-c("skewness")
rownames(qskew) <-colnames(khq.daylog[,1:4])

qkurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qku <-kurtosis(khq.daylog[,i], na.rm=TRUE)
  qkurt[i,]<-qku
}
colnames(qkurt) <-c("kurtosis")
rownames(qkurt) <-colnames(khq.daylog[,1:4])

#check normalcy of distribution
qnorm(khq.daylog[,1])
qqline(khq.daylog[,1])
qqnorm(khq.daylog[,2])
qqline(khq.daylog[,2])
qqnorm(khq.daylog[,3])
qqline(khq.daylog[,3])
qqnorm(khq.daylog[,4])
qqline(khq.daylog[,4])


#check residuals transform as log (add a constant and transform to log)
bayanblogt <-log(bayanb+0.1)
bayanb.loglm <-lm(bayanblogt~time, na.action=na.omit)
time.r9 <-seq(1:length(bayanb.loglm$residuals))
plot(time.r9, bayanb.loglm$residuals, pch=19)
abline(0,0)
```

```
bayantlogt <-log(bayant+0.1)
bayant.loglm <-lm(bayantlogt~time, na.action=na.omit)
time.r10 <-seq(1:length(bayant.loglm$residuals))
plot(time.r10, bayant.loglm$residuals, pch=19)
abline(0,0)

erdklogt <-log(erdk+0.1)
erdk.loglm <-lm(erdklogt~time, na.action=na.omit)
time.r12 <-seq(1:length(erdk.loglm$residuals))
plot(time.r12, erdk.loglm$residuals, pch=19)
abline(0,0)

ikhktlogt <-log(ikhkt+0.1)
ikhkt.loglm <-lm(ikhktlogt~time, na.action=na.omit)
time.r8 <-seq(1:length(ikhkt.loglm$residuals))
plot(time.r8, ikhkt.loglm$residuals, pch=19)
abline(0,0)

#all daily values except bayanburd have better distribution of residuals when log transformed

#precip
#-------------------------------------------------------------
#precip
khp.d <-read.csv("Khangai_Daily_P.csv")
#make wide format
khp.dw <-dcast(khp.d, yrmonday~name, value.var="dayP")
khp.dw$Year <-substr(khp.dw$yrmonday, 1,4)
khp.dw$Month <-substr(khp.dw$yrmonday, 6,7)
khp.dw$Day <-substr(khp.dw$yrmonday, 9,10)

#truncate to period of interest
khp.day <-khp.dw[3732:9719,]

#qqplots
qqnorm(khp.day[,2])
qqline(khp.day[,2])
qqnorm(khp.day[,3])
qqline(khp.day[,3])
qqnorm(khp.day[,4])
qqline(khp.day[,4])
qqnorm(khp.day[,5])
qqline(khp.day[,5])
qqnorm(khp.day[,6])
qqline(khp.day[,6])
qqnorm(khp.day[,7])
qqline(khp.day[,7])


#skew and kurtosis
pskew <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  psk <-skewness(khp.day[,i+1], na.rm=TRUE)
  pskew[i,]<-psk
}
colnames(pskew) <-c("skewness")
```

```
rownames(pskew) <-colnames(khp.day[,2:7])

pkurt <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
 pku <-kurtosis(khp.day[,i+1], na.rm=TRUE)
 pkurt[i,]<-pku
}
colnames(pkurt) <-c("kurtosis")
rownames(pkurt) <-colnames(khp.day[,2:7])

#autocorrelation
pacf <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6) {
 pcorr <-acf(khp.day[,i+1], type="correlation", na.action=na.pass)
 pacf[i,] <-pcorr$acf[[2]]
}
colnames(pacf) <-"lag1"
rownames(pacf) <-colnames(khp.day[,2:7])

#check residuals
timep <-seq(1:length(khp.day[,2]))
baidrag.lm <-lm(khp.day[,2]~timep, na.action=na.omit)
time.pb <-seq(1:length(baidrag.lm$residuals))
plot(time.pb, baidrag.lm$residuals, pch=19)
abline(0,0)

bayankhongor.lm <-lm(khp.day[,3]~timep, na.action=na.omit)
time.pba <-seq(1:length(bayankhongor.lm$residuals))
plot(time.pba, bayankhongor.lm$residuals, pch=19)
abline(0,0)

erdene.lm <-lm(khp.day[,4]~timep, na.action=na.omit)
time.pe <-seq(1:length(erdene.lm$residuals))
plot(time.pe, erdene.lm$residuals, pch=19)
abline(0,0)

galuut.lm <-lm(khp.day[,5]~timep, na.action=na.omit)
time.pg <-seq(1:length(galuut.lm$residuals))
plot(time.pg, galuut.lm$residuals, pch=19)
abline(0,0)

#what about using Tariat rather than Khoriult- as it is more northerly and has less missing data?
khoriult.lm <-lm(khp.day[,6]~timep, na.action=na.omit)
time.pk <-seq(1:length(khoriult.lm$residuals))
plot(time.pk, khoriult.lm$residuals, pch=19)
abline(0,0)


tsetser.lm <-lm(khp.day[,7]~timep, na.action=na.omit)
time.pt <-seq(1:length(tsetser.lm$residuals))
plot(time.pt, tsetser.lm$residuals, pch=19)
abline(0,0)

#transform
khp.daysqrt<- sqrt(khp.day[,2:7])
```

```
#qqplots
qqnorm(khp.daysqrt[,1])
qqline(khp.daysqrt[,1])
qqnorm(khp.daysqrt[,2])
qqline(khp.daysqrt[,2])
qqnorm(khp.daysqrt[,3])
qqline(khp.daysqrt[,3])
qqnorm(khp.daysqrt[,4])
qqline(khp.daysqrt[,4])
qqnorm(khp.daysqrt[,5])
qqline(khp.daysqrt[,5])
qqnorm(khp.daysqrt[,6])
qqline(khp.daysqrt[,6])

#skewness and kurtosis
pskew <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  psk <-skewness(khp.daysqrt[,i], na.rm=TRUE)
  pskew[i,]<-psk
}
colnames(pskew) <-c("skewness")
rownames(pskew) <-colnames(khp.daysqrt[,1:6])

pkurt <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  pku <-kurtosis(khp.daysqrt[,i], na.rm=TRUE)
  pkurt[i,]<-pku
}
colnames(pkurt) <-c("kurtosis")
rownames(pkurt) <-colnames(khp.daysqrt[,1:6])

#autocorrelation
pacf <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6) {
  pcorr <-acf(khp.daysqrt[,i], type="correlation", na.action=na.pass)
  pacf[i,] <-pcorr$acf[[2]]
}
colnames(pacf) <-"lag1"
rownames(pacf) <-colnames(khp.daysqrt[,1:6])

#check residuals
baidragsqrt.lm <-lm(khp.daysqrt[,1]~timep, na.action=na.omit)
time.p1 <-seq(1:length(baidragsqrt.lm$residuals))
plot(time.p1, baidragsqrt.lm$residuals, pch=19)
abline(0,0)


bayankhongorsqrt.lm <-lm(khp.daysqrt[,2]~timep, na.action=na.omit)
time.p2 <-seq(1:length(bayankhongorsqrt.lm$residuals))
plot(time.p2, bayankhongorsqrt.lm$residuals, pch=19)
abline(0,0)

erdenesqrt.lm <-lm(khp.daysqrt[,3]~timep, na.action=na.omit)
time.p3 <-seq(1:length(erdenesqrt.lm$residuals))
plot(time.p3, erdenesqrt.lm$residuals, pch=19)
```

```
abline(0,0)

galuutsqrt.lm <-lm(khp.daysqrt[,4]~timep, na.action=na.omit)
time.p4 <-seq(1:length(galuutsqrt.lm$residuals))
plot(time.p4, galuutsqrt.lm$residuals, pch=19)
abline(0,0)

khoriultsqrt.lm <-lm(khp.daysqrt[,5]~timep, na.action=na.omit)
time.p5 <-seq(1:length(khoriultsqrt.lm$residuals))
plot(time.p5, khoriultsqrt.lm$residuals, pch=19)
abline(0,0)

tsetsersqrt.lm <-lm(khp.daysqrt[,6]~timep, na.action=na.omit)
time.p6 <-seq(1:length(tsetsersqrt.lm$residuals))
plot(time.p6, tsetsersqrt.lm$residuals, pch=19)
abline(0,0)

#transform using log e +0.1
khp.daylog<- log(khp.day[,2:7]+0.1)

#qqplots
qqnorm(khp.daylog[,1])
qqline(khp.daylog[,1])
qqnorm(khp.daylog[,2])
qqline(khp.daylog[,2])
qqnorm(khp.daylog[,3])
qqline(khp.daylog[,3])
qqnorm(khp.daylog[,4])
qqline(khp.daylog[,4])
qqnorm(khp.daylog[,5])
qqline(khp.daylog[,5])
qqnorm(khp.daylog[,6])
qqline(khp.daylog[,6])

#skewness and kurtosis
pskew <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  psk <-skewness(khp.daylog[,i], na.rm=TRUE)
  pskew[i,]<-psk
}
colnames(pskew) <-c("skewness")
rownames(pskew) <-colnames(khp.daylog[,1:6])

pkurt <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  pku <-kurtosis(khp.daylog[,i], na.rm=TRUE)
  pkurt[i,]<-pku
}
colnames(pkurt) <-c("kurtosis")
rownames(pkurt) <-colnames(khp.daylog[,1:6])

#autocorrelation
pacf <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6) {
  pcorr <-acf(khp.daylog[,i], type="correlation", na.action=na.pass)
```

```
  pacf[i,] <-pcorr$acf[[2]]
}
colnames(pacf) <-"lag1"
rownames(pacf) <-colnames(khp.daylog[,1:6])

#check residuals
baidraglog.lm <-lm(khp.daylog[,1]~timep, na.action=na.omit)
time.p1 <-seq(1:length(baidraglog.lm$residuals))
plot(time.p1, baidraglog.lm$residuals, pch=19)
abline(0,0)

bayankhongorlog.lm <-lm(khp.daylog[,2]~timep, na.action=na.omit)
time.p2 <-seq(1:length(bayankhongorlog.lm$residuals))
plot(time.p2, bayankhongorlog.lm$residuals, pch=19)
abline(0,0)

erdenelog.lm <-lm(khp.daylog[,3]~timep, na.action=na.omit)
time.p3 <-seq(1:length(erdenelog.lm$residuals))
plot(time.p3, erdenelog.lm$residuals, pch=19)
abline(0,0)

galuutlog.lm <-lm(khp.daylog[,4]~timep, na.action=na.omit)
time.p4 <-seq(1:length(galuutlog.lm$residuals))
plot(time.p4, galuutlog.lm$residuals, pch=19)
abline(0,0)

khoriultlog.lm <-lm(khp.daylog[,5]~timep, na.action=na.omit)
time.p5 <-seq(1:length(khoriultlog.lm$residuals))
plot(time.p5, khoriultlog.lm$residuals, pch=19)
abline(0,0)

tsetserlog.lm <-lm(khp.daylog[,6]~timep, na.action=na.omit)
time.p6 <-seq(1:length(tsetserlog.lm$residuals))
plot(time.p6, tsetserlog.lm$residuals, pch=19)
abline(0,0)

#log is NOT better for P, sqrt may be adequate from plots
#-------------------------------------
#monthly values tend to have less autocorrelation than daily, particularly for Q.
#monthly p
#-------------------------------------------------------------------------------------
#import QC'd station-based monthly precipitation (what was used in LOR paper, but swap Tariat for Khoriult
(check 4 SD for errors))
khp <-read.csv("Khangai_Monthly_P.csv")

#convert to wide format
khp.l <-dcast(khp, yrmon~name, value.var="monP")
khp.l$Year <-substr(khp.l$yrmon, 1,4)
khp.l$Month <-substr(khp.l$yrmon, 6,7)

#select only stations of interest (not Khoriult)

#truncate precip dataset to overlap period with streamflow (given that will get data for 2011/2012
streamflow)
khp.li <-khp.l[,c(1:5, 7:10)]
```

```
khp.t <- khp.li[312:755,]

#avg annual precip (no years with missing months allowed in overall avgs)
yearlyp <-ddply(khp.t, .(Year), summarize, BaidP=sum(Baidrag), BayanP=sum(Bayankhongor),
ErdeneP=sum(Erdenemandal), GalP=sum(Galuut), TarP=sum(Tariat), TsetP=sum(Tsetserleg))
colMeans(yearlyp[2:7], na.rm=TRUE)

#perform EDA
psummary <-matrix(ncol=7, nrow=6, NA)
for (i in 1:6){
  psum <-summary(khp.t[,i+1])
  psummary[i,]<-psum
}
colnames(psummary) <-names(psum)
rownames(psummary) <-colnames(khp.t[,2:7])

qqnorm(khp.t[,2])
qqline(khp.t[,2])
qqnorm(khp.t[,3])
qqline(khp.t[,3])
qqnorm(khp.t[,4])
qqline(khp.t[,4])
qqnorm(khp.t[,5])
qqline(khp.t[,5])
qqnorm(khp.t[,6])
qqline(khp.t[,6])
qqnorm(khp.t[,7])
qqline(khp.t[,7])

#skew and kurtosis
pskew <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  psk <-skewness(khp.t[,i+1], na.rm=TRUE)
  pskew[i,]<-psk
}
colnames(pskew) <-c("skewness")
rownames(pskew) <-colnames(khp.t[,2:7])

pkurt <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  pku <-kurtosis(khp.t[,i+1], na.rm=TRUE)
  pkurt[i,]<-pku
}
colnames(pkurt) <-c("kurtosis")
rownames(pkurt) <-colnames(khp.t[,2:7])


#autocorrelation
pacf <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6) {
  pcorr <-acf(khp.t[,i+1], type="correlation", na.action=na.pass)
  pacf[i,] <-pcorr$acf[[2]]
}
colnames(pacf) <-"lag1"
rownames(pacf) <-colnames(khp.t[,2:7])
```

```
#check residuals
timep <-seq(1:length(khp.t[,2]))
baidrag.lm <-lm(khp.t[,2]~timep, na.action=na.omit)
time.pb <-seq(1:length(baidrag.lm$residuals))
plot(time.pb, baidrag.lm$residuals, pch=19)
abline(0,0)

bayankhongor.lm <-lm(khp.t[,3]~timep, na.action=na.omit)
time.pba <-seq(1:length(bayankhongor.lm$residuals))
plot(time.pba, bayankhongor.lm$residuals, pch=19)
abline(0,0)

erdene.lm <-lm(khp.t[,4]~timep, na.action=na.omit)
time.pe <-seq(1:length(erdene.lm$residuals))
plot(time.pe, erdene.lm$residuals, pch=19)
abline(0,0)

galuut.lm <-lm(khp.t[,5]~timep, na.action=na.omit)
time.pg <-seq(1:length(galuut.lm$residuals))
plot(time.pg, galuut.lm$residuals, pch=19)
abline(0,0)

tariat.lm <-lm(khp.t[,6]~timep, na.action=na.omit)
time.pk <-seq(1:length(tariat.lm$residuals))
plot(time.pk, tariat.lm$residuals, pch=19)
abline(0,0)

tsetser.lm <-lm(khp.t[,7]~timep, na.action=na.omit)
time.pt <-seq(1:length(tsetser.lm$residuals))
plot(time.pt, tsetser.lm$residuals, pch=19)
abline(0,0)

#transform
khp.tsqrt<- sqrt(khp.t[,2:7])

#qqplots
qqnorm(khp.tsqrt[,1])
qqline(khp.tsqrt[,1])
qqnorm(khp.tsqrt[,2])
qqline(khp.tsqrt[,2])
qqnorm(khp.tsqrt[,3])
qqline(khp.tsqrt[,3])
qqnorm(khp.tsqrt[,4])
qqline(khp.tsqrt[,4])
qqnorm(khp.tsqrt[,5])
qqline(khp.tsqrt[,5])
qqnorm(khp.tsqrt[,6])
qqline(khp.tsqrt[,6])

#skewness and kurtosis
pskew <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  psk <-skewness(khp.tsqrt[,i], na.rm=TRUE)
  pskew[i,]<-psk
}
```

```
colnames(pskew) <-c("skewness")
rownames(pskew) <-colnames(khp.tsqrt[,1:6])

pkurt <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
 pku <-kurtosis(khp.tsqrt[,i], na.rm=TRUE)
 pkurt[i,]<-pku
}
colnames(pkurt) <-c("kurtosis")
rownames(pkurt) <-colnames(khp.tsqrt[,1:6])

#autocorrelation
pacf <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6) {
 pcorr <-acf(khp.tsqrt[,i], type="correlation", na.action=na.pass)
 pacf[i,] <-pcorr$acf[[2]]
}
colnames(pacf) <-"lag1"
rownames(pacf) <-colnames(khp.tsqrt[,1:6])

#check residuals
baidragsqrt.lm <-lm(khp.tsqrt[,1]~timep, na.action=na.omit)
time.p1 <-seq(1:length(baidragsqrt.lm$residuals))
plot(time.p1, baidragsqrt.lm$residuals, pch=19)
abline(0,0)

bayankhongorsqrt.lm <-lm(khp.tsqrt[,2]~timep, na.action=na.omit)
time.p2 <-seq(1:length(bayankhongorsqrt.lm$residuals))
plot(time.p2, bayankhongorsqrt.lm$residuals, pch=19)
abline(0,0)

erdenesqrt.lm <-lm(khp.tsqrt[,3]~timep, na.action=na.omit)
time.p3 <-seq(1:length(erdenesqrt.lm$residuals))
plot(time.p3, erdenesqrt.lm$residuals, pch=19)
abline(0,0)

galuutsqrt.lm <-lm(khp.tsqrt[,4]~timep, na.action=na.omit)
time.p4 <-seq(1:length(galuutsqrt.lm$residuals))
plot(time.p4, galuutsqrt.lm$residuals, pch=19)
abline(0,0)

tariatsqrt.lm <-lm(khp.tsqrt[,5]~timep, na.action=na.omit)
time.p5 <-seq(1:length(tariatsqrt.lm$residuals))
plot(time.p5, tariatsqrt.lm$residuals, pch=19)
abline(0,0)

tsetsersqrt.lm <-lm(khp.tsqrt[,6]~timep, na.action=na.omit)
time.p6 <-seq(1:length(tsetsersqrt.lm$residuals))
plot(time.p6, tsetsersqrt.lm$residuals, pch=19)
abline(0,0)

#transform using log e +0.1
khp.tlog<- log(khp.t[,2:7]+0.1)

#qqplots
```

```
qqnorm(khp.tlog[,1])
qqline(khp.tlog[,1])
qqnorm(khp.tlog[,2])
qqline(khp.tlog[,2])
qqnorm(khp.tlog[,3])
qqline(khp.tlog[,3])
qqnorm(khp.tlog[,4])
qqline(khp.tlog[,4])
qqnorm(khp.tlog[,5])
qqline(khp.tlog[,5])
qqnorm(khp.tlog[,6])
qqline(khp.tlog[,6])

#skewness and kurtosis
pskew <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  psk <-skewness(khp.tlog[,i], na.rm=TRUE)
  pskew[i,]<-psk
}
colnames(pskew) <-c("skewness")
rownames(pskew) <-colnames(khp.tlog[,1:6])

pkurt <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6){
  pku <-kurtosis(khp.tlog[,i], na.rm=TRUE)
  pkurt[i,]<-pku
}
colnames(pkurt) <-c("kurtosis")
rownames(pkurt) <-colnames(khp.tlog[,1:6])


#autocorrelation
pacf <-matrix(ncol=1, nrow=6, NA)
for (i in 1:6) {
  pcorr <-acf(khp.tlog[,i], type="correlation", na.action=na.pass)
  pacf[i,] <-pcorr$acf[[2]]
}
colnames(pacf) <-"lag1"
rownames(pacf) <-colnames(khp.tlog[,1:6])

#check residuals
baidraglog.lm <-lm(khp.tlog[,1]~timep, na.action=na.omit)
time.p1 <-seq(1:length(baidraglog.lm$residuals))
plot(time.p1, baidraglog.lm$residuals, pch=19)
abline(0,0)

bayankhongorlog.lm <-lm(khp.tlog[,2]~timep, na.action=na.omit)
time.p2 <-seq(1:length(bayankhongorlog.lm$residuals))
plot(time.p2, bayankhongorlog.lm$residuals, pch=19)
abline(0,0)

erdenelog.lm <-lm(khp.tlog[,3]~timep, na.action=na.omit)
time.p3 <-seq(1:length(erdenelog.lm$residuals))
plot(time.p3, erdenelog.lm$residuals, pch=19)
abline(0,0)
```

```
galuutlog.lm <-lm(khp.tlog[,4]~timep, na.action=na.omit)
time.p4 <-seq(1:length(galuutlog.lm$residuals))
plot(time.p4, galuutlog.lm$residuals, pch=19)
abline(0,0)

tariatlog.lm <-lm(khp.tlog[,5]~timep, na.action=na.omit)
time.p5 <-seq(1:length(tariatlog.lm$residuals))
plot(time.p5, tariatlog.lm$residuals, pch=19)
abline(0,0)

tsetserlog.lm <-lm(khp.tlog[,6]~timep, na.action=na.omit)
time.p6 <-seq(1:length(tsetserlog.lm$residuals))
plot(time.p6, tsetserlog.lm$residuals, pch=19)
abline(0,0)
#-------
#log transform skews the data to the negative, and even though residual plots are better distributed, qq plots
are more negative.
#square root transform should be sufficient for precip data
#test significance of results

#monthly q
#----------------------------------------------------------------
#compare to already aggregate monthly flow (from LOR paper and new 2011-2012 data provided by Odko)
converted to CM
khq <-read.csv("Mean_Monthly_Q_All.csv")

#add yrmon column to streamflow
khq$Month_Num <-rep(seq(1:12),42)
khq$mon <-sprintf("%02d", khq$Month_Num)
khq$yrmon <-paste(khq$Year, khq$mon, sep="-")

#truncate streamflow to period of interest
khq.t <-khq[61:504,c(9,3:6,1,2)]
colnames(khq.t) <- c("yrmon","BayanB","BayanT","ErdK", "IkhKT", "Year", "Month")

#convert Q dataset to cubic meters per month (should not effect stats as it's just multiplying by a constant?!)
#each series is Jan-Dec x 37 years, each month has a differing number of days, and assume 28 days uniformly
for February (as it is a low flow month anyhow).
monthday <-rep(c(31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31), 37)
mdsec <-86400*monthday
khq.cmcon <-khq.t[,2:5]*mdsec

#divide by million to get MCM per month
khq.mc <-khq.cmcon/1000000
khq.mcm <-cbind(khq.t[,1],khq.mc, khq.t[,6:7])
colnames(khq.mcm) <-colnames(khq.t)

#sum over months per year to get MCM/annual and compare to converted daily estimates
#khq.mcmann <-ddply(khq.mcm, .(Year), summarize, BayanB=sum(BayanB, na.rm=TRUE),
BayanT=sum(BayanT, na.rm=TRUE), ErdK=sum(ErdK, na.rm=TRUE), IkhKT=sum(IkhKT, na.rm=TRUE))
#colMeans(khq.mcmann[,2:5])
#  BayanB    BayanT      ErdK     IkhKT  results are on the order of size of "headwaters" rivers studied in
Watson et al 2009
#309.71769  83.08605 121.65355 153.39193
```

```
#run EDA
qsummary <-matrix(ncol=7, nrow=4, NA)
for (i in 1:4){
  qsum <-summary(khq.mcm[,i+1])
  qsummary[i,]<-qsum
}
colnames(qsummary) <-names(qsum)
rownames(qsummary) <-colnames(khq.mcm[,2:5])


#check distribution of data
qqnorm(khq.mcm[,2])
qqline(khq.mcm[,2])
qqnorm(khq.mcm[,3])
qqline(khq.mcm[,3])
qqnorm(khq.mcm[,4])
qqline(khq.mcm[,4])
qqnorm(khq.mcm[,5])
qqline(khq.mcm[,5])

#check autocorrelation of data
qacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
  qcorr <-acf(khq.mcm[,i+1], type="correlation", na.action=na.pass)
  qacf[i,] <-qcorr$acf[[2]]
}
colnames(qacf) <-"lag1"
rownames(qacf) <-colnames(khq.mcm[,2:5])

#High levels of autocorr for all basins

#check for skewness and kurtosis
qskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qsk <-skewness(khq.mcm[,i+1], na.rm=TRUE)
  qskew[i,]<-qsk
}
colnames(qskew) <-c("skewness")
rownames(qskew) <-colnames(khq.mcm[,2:5])

qkurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qku <-kurtosis(khq.mcm[,i+1], na.rm=TRUE)
  qkurt[i,]<-qku
}
colnames(qkurt) <-c("kurtosis")
rownames(qkurt) <-colnames(khq.mcm[,2:5])

#check distribution of residuals
time <-seq(1:length(khq.mcm[,2]))
bayanb <-khq.mcm[,2]
bayanb.lm <-lm(bayanb~time, na.action=na.omit)
time.r <-seq(1:length(bayanb.lm$residuals))
plot(time.r, bayanb.lm$residuals, pch=19)
abline(0,0)
```

```
bayant <-khq.mcm[,3]
bayant.lm <-lm(bayant~time, na.action=na.omit)
time.r2 <-seq(1:length(bayant.lm$residuals))
plot(time.r2, bayant.lm$residuals, pch=19)
abline(0,0)

erdk <-khq.mcm[,4]
erdk.lm <-lm(erdk~time, na.action=na.omit)
time.r4 <-seq(1:length(erdk.lm$residuals))
plot(time.r4, erdk.lm$residuals, pch=19)
abline(0,0)

ikhkt <-khq.mcm[,3]
ikhkt.lm <-lm(ikhkt~time, na.action=na.omit)
time.r6 <-seq(1:length(ikhkt.lm$residuals))
plot(time.r6, ikhkt.lm$residuals, pch=19)
abline(0,0)

#try transform using sqrt first
khq.mcmsqrt<- sqrt(khq.mcm[,2:5])

#autocorr in transformed variables
qacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
  qcorr <-acf(khq.mcmsqrt[,i], type="correlation", na.action=na.pass)
  qacf[i,] <-qcorr$acf[[2]]
}
colnames(qacf) <-"lag1"
rownames(qacf) <-colnames(khq.mcmsqrt[,1:4])

#transformed with square root
qskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qsk <-skewness(khq.mcmsqrt[,i], na.rm=TRUE)
  qskew[i,]<-qsk
}
colnames(qskew) <-c("skewness")
rownames(qskew) <-colnames(khq.mcmsqrt[,1:4])

qkurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
  qku <-kurtosis(khq.mcmsqrt[,i], na.rm=TRUE)
  qkurt[i,]<-qku
}
colnames(qkurt) <-c("kurtosis")
rownames(qkurt) <-colnames(khq.mcmsqrt[,1:4])

#check normalcy
qqnorm(khq.mcmsqrt[,1])
qqline(khq.mcmsqrt[,1])
qqnorm(khq.mcmsqrt[,2])
qqline(khq.mcmsqrt[,2])
qqnorm(khq.mcmsqrt[,3])
qqline(khq.mcmsqrt[,3])
qqnorm(khq.mcmsqrt[,4])
```

```
qqline(khq.mcmsqrt[,4])

#check residuals
bayanb.sqrt <-khq.mcmsqrt[,1]
bayanb.sqrtlm <-lm(bayanb.sqrt~time, na.action=na.omit)
time.r1 <-seq(1:length(bayanb.sqrtlm$residuals))
plot(time.r1, bayanb.sqrtlm$residuals, pch=19)
abline(0,0)

#plot of absolute residuals
plot(time.r1, abs(bayanb.sqrtlm$residuals), pch=19)

bayant.sqrt <-khq.mcmsqrt[,2]
bayant.sqrtlm <-lm(bayant.sqrt~time, na.action=na.omit)
time.r3 <-seq(1:length(bayant.sqrtlm$residuals))
plot(time.r3, bayant.sqrtlm$residuals, pch=19)
abline(0,0)

erdk.sqrt <-khq.mcmsqrt[,3]
erdk.sqrtlm <-lm(erdk.sqrt~time, na.action=na.omit)
time.r5 <-seq(1:length(erdk.sqrtlm$residuals))
plot(time.r5, erdk.sqrtlm$residuals, pch=19)
abline(0,0)

ikhkt.sqrt <-khq.mcmsqrt[,2]
ikhkt.sqrtlm <-lm(ikhkt.sqrt~time, na.action=na.omit)
time.r7 <-seq(1:length(ikhkt.sqrtlm$residuals))
plot(time.r7, ikhkt.sqrtlm$residuals, pch=19)
abline(0,0)

#transformed with log +0.1
khq.mcmlog <-log(khq.mcm[,2:5]+0.1)

#autocorr in transformed variables
qacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
 qcorr <-acf(khq.mcmlog[,i], type="correlation", na.action=na.pass)
 qacf[i,] <-qcorr$acf[[2]]
}
colnames(qacf) <-"lag1"
rownames(qacf) <-colnames(khq.mcmlog[,1:4])

#skewness and kurtosis
qskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
 qsk <-skewness(khq.mcmlog[,i], na.rm=TRUE)
 qskew[i,]<-qsk
}
colnames(qskew) <-c("skewness")
rownames(qskew) <-colnames(khq.mcmlog[,1:4])
qkurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
 qku <-kurtosis(khq.mcmlog[,i], na.rm=TRUE)
 qkurt[i,]<-qku
}
```

```
colnames(qkurt) <-c("kurtosis")
rownames(qkurt) <-colnames(khq.mcmlog[,1:4])

#check normalcy of distribution
qnorm(khq.mcmlog[,1])
qqline(khq.mcmlog[,1])
qqnorm(khq.mcmlog[,2])
qqline(khq.mcmlog[,2])
qqnorm(khq.mcmlog[,3])
qqline(khq.mcmlog[,3])
qqnorm(khq.mcmlog[,4])
qqline(khq.mcmlog[,4])

#check residuals transform as log (add a constant and transform to log)
bayanblogt <-log(bayanb+0.1)
bayanb.loglm <-lm(bayanblogt~time, na.action=na.omit)
time.r9 <-seq(1:length(bayanb.loglm$residuals))
plot(time.r9, bayanb.loglm$residuals, pch=19)
abline(0,0)

bayantlogt <-log(bayant+0.1)
bayant.loglm <-lm(bayantlogt~time, na.action=na.omit)
time.r10 <-seq(1:length(bayant.loglm$residuals))
plot(time.r10, bayant.loglm$residuals, pch=19)
abline(0,0)

erdklogt <-log(erdk+0.1)
erdk.loglm <-lm(erdklogt~time, na.action=na.omit)
time.r12 <-seq(1:length(erdk.loglm$residuals))
plot(time.r12, erdk.loglm$residuals, pch=19)
abline(0,0)

ikhktlogt <-log(ikhkt+0.1)
ikhkt.loglm <-lm(ikhktlogt~time, na.action=na.omit)
time.r8 <-seq(1:length(ikhkt.loglm$residuals))
plot(time.r8, ikhkt.loglm$residuals, pch=19)
abline(0,0)
#----
#log transform of monthly data introduces more negative skew square root transform may be adequate

#GPCC
#----------------------------------------------------------------------------------------
#Load monthly GPCC values for each basin using average of pixels within each basin area (no gaps, thru 2010
only)
#also load distances to centroid (latlong of basin)
gpcc <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPC
C_Extract_R/GPCC_Mean_P_1976_2010_Basins.csv")

gpcc$yrmon <-khq.t$yrmon[1:420]
gpcc.t <-gpcc[,c(6,2:5)]

#distance to centroid
cent.dist <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQ_Khangai/GPC
```

```
C_Extract_R/Station_Distances_to_Basin_Centroid.csv")

gpccsummary <-matrix(ncol=6, nrow=4, NA)
for (i in 1:4){
 gpccsum <-summary(gpcc.t[,i+1])
 gpccsummary[i,]<-gpccsum
}
colnames(gpccsummary) <-names(gpccsum)
rownames(gpccsummary) <-colnames(gpcc.t[,2:5])

#check normality
#skewness
gpccskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
 gpccsk <-skewness(gpcc.t[,i+1], na.rm=TRUE)
 gpccskew[i,]<-gpccsk
}
colnames(gpccskew) <-c("skewness")
rownames(gpccskew) <-colnames(gpcc.t[,2:5])

#kurtosis
gpcckurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
 gpccku <-kurtosis(gpcc.t[,i+1], na.rm=TRUE)
 gpcckurt[i,]<-gpccku
}
colnames(gpcckurt) <-c("kurtosis")
rownames(gpcckurt) <-colnames(gpcc.t[,2:5])

#autocorrelation
gpccacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
 gpcccorr <-acf(gpcc.t[,i+1], type="correlation", na.action=na.pass)
 gpccacf[i,] <-gpcccorr$acf[[2]]
}
colnames(gpccacf) <-"lag1"
rownames(gpccacf) <-colnames(gpcc.t[,2:5])

#examine residuals and qq plots
#check normalcy of distribution
qnorm(gpcc.t[,1])
qqline(gpcc.t[,1])
qqnorm(gpcc.t[,2])
qqline(gpcc.t[,2])
qqnorm(gpcc.t[,3])
qqline(gpcc.t[,3])
qqnorm(gpcc.t[,4])
qqline(gpcc.t[,4])

#qq plots are similar, due to smoothing effects of interpolation procedure?

#check residuals
time <-seq(1:length(gpcc.t[,2]))
bayanb.gpcclm <-lm(gpcc.t[,2]~time, na.action=na.omit)
time.r9 <-seq(1:length(bayanb.gpcclm$residuals))
```

```
plot(time.r9, bayanb.gpcclm$residuals, pch=19)
abline(0,0)

bayant.gpcclm <-lm(gpcc.t[,3]~time, na.action=na.omit)
time.r10 <-seq(1:length(bayant.gpcclm$residuals))
plot(time.r10, bayant.gpcclm$residuals, pch=19)
abline(0,0)

erdk.gpcclm <-lm(gpcc.t[,4]~time, na.action=na.omit)
time.r12 <-seq(1:length(erdk.gpcclm$residuals))
plot(time.r12, erdk.gpcclm$residuals, pch=19)
abline(0,0)

ikhkt.gpcclm <-lm(gpcc.t[,5]~time, na.action=na.omit)
time.r8 <-seq(1:length(ikhkt.gpcclm$residuals))
plot(time.r8, ikhkt.gpcclm$residuals, pch=19)
abline(0,0)

#transform using sqrt
gpcc.sqrt <-sqrt(gpcc.t[,2:5])

#skewness
gpccskew <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
 gpccsk <-skewness(gpcc.sqrt[,1], na.rm=TRUE)
 gpccskew[i,]<-gpccsk
}
colnames(gpccskew) <-c("skewness")
rownames(gpccskew) <-colnames(gpcc.sqrt[,1:4])

#kurtosis
gpcckurt <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4){
 gpccku <-kurtosis(gpcc.sqrt[,i], na.rm=TRUE)
 gpcckurt[i,]<-gpccku
}
colnames(gpcckurt) <-c("kurtosis")
rownames(gpcckurt) <-colnames(gpcc.sqrt[,1:4])


#autocorrelation
gpccacf <-matrix(ncol=1, nrow=4, NA)
for (i in 1:4) {
 gpcccorr <-acf(gpcc.sqrt[,i], type="correlation", na.action=na.pass)
 gpccacf[i,] <-gpcccorr$acf[[2]]
}
colnames(gpccacf) <-"lag1"
rownames(gpccacf) <-colnames(gpcc.sqrt[,1:4])

#examine residuals and qq plots
#check normalcy of distribution
qnorm(gpcc.sqrt[,1])
qqline(gpcc.sqrt[,1])
qqnorm(gpcc.sqrt[,2])
qqline(gpcc.sqrt[,2])
```

```
qqnorm(gpcc.sqrt[,3])
qqline(gpcc.sqrt[,3])
qqnorm(gpcc.sqrt[,4])
qqline(gpcc.sqrt[,4])

#qq plots are similar, due to smoothing effects of interpolation procedure?

#check residuals
time <-seq(1:length(gpcc.sqrt[,2]))
bayanb.gpcclm <-lm(gpcc.sqrt[,1]~time, na.action=na.omit)
time.r9 <-seq(1:length(bayanb.gpcclm$residuals))
plot(time.r9, bayanb.gpcclm$residuals, pch=19)
abline(0,0)

bayant.gpcclm <-lm(gpcc.sqrt[,2]~time, na.action=na.omit)
time.r10 <-seq(1:length(bayant.gpcclm$residuals))
plot(time.r10, bayant.gpcclm$residuals, pch=19)
abline(0,0)

erdk.gpcclm <-lm(gpcc.sqrt[,3]~time, na.action=na.omit)
time.r12 <-seq(1:length(erdk.gpcclm$residuals))
plot(time.r12, erdk.gpcclm$residuals, pch=19)
abline(0,0)

ikhkt.gpcclm <-lm(gpcc.sqrt[,4]~time, na.action=na.omit)
time.r8 <-seq(1:length(ikhkt.gpcclm$residuals))
plot(time.r8, ikhkt.gpcclm$residuals, pch=19)
abline(0,0)

#-------------------------------------------------------------------------------------------------------
#collate results of transformations (sqrt) and export
khq.mcmsqrtf <-data.frame(khq.mcm[,1], khq.mcmsqrt[,], khq.mcm[,6:7])
colnames(khq.mcmsqrtf) <-colnames(khq.mcm)
write.csv(khq.mcmsqrtf, "Transformed_Q_1976-2012.csv")

khp.tsqrtf <-data.frame(khp.t[,1], khp.tsqrt[,], khp.t[,8:9])
colnames(khp.tsqrtf) <-colnames(khp.t)
write.csv(khp.tsqrtf, "Transformed_P_1976-2012.csv")

gpcc.tsqrt <-data.frame(gpcc.t[,1], gpcc.sqrt[,])
colnames(gpcc.tsqrt) <-colnames(gpcc.t)
write.csv(gpcc.tsqrt, "Transformed_GPCC_1976-2010.csv")

#---------------------------------------------------------------------------------
#testing of variables
#khq.mcmsqrtf, khp.tsqrtf, gpcc.tsqrt

#durbin watson test for autocorrelation significantly different than zero
#must create lagged dataset to test one on the other- verfies if residuals from a linear model are correlated or not.
#test on least autocorrelated variable in dataset
bayanburd.z <-zoo(khq.mcmsqrtf[,2], khq.mcmsqrtf[,1])
bayanburd.zno <-na.omit(bayanburd.z)
bayanburd.zlag <-lag(bayanburd.zno, k=-1)
dwtest(bayanburd.zno[2:437]~bayanburd.zlag, alternative="two.sided")
```

```
#statistic is 0, pvalue is 2.2e-16 so can reject the null that there is no correlation among residuals (they are independent)
#alternative is that they are autocorrelated (GPCC would be the same since it uses the same underlying data)

baidrag.z <-zoo(khp.tsqrtf[,2], khp.tsqrtf[,1])
baidrag.zno <-na.omit(baidrag.z)
baidrag.zlag <-lag(baidrag.zno, k=-1)
dwtest(baidrag.zno[2:220]~baidrag.zlag, alternative="two.sided")
#statistic is 0, pvalue is 2.2e-16 so can reject the null that there is no correlation among residuals (they are independent)
#alternative is that they are autocorrelatedc

#d'agostino test for skew significantly different from zero (default 2 tailed)
bayanburd.a <-agostino.test(bayanburd.zno,alternative="two.sided")
#skew = 0.3602, z = 3.0358, p-value = 0.002399
#skew significant from zero for bayanburd but less than for untransformed data

baidrag.a <-agostino.test(baidrag.zno,alternative="two.sided")
#skew = 0.9900, z = 5.2849, p-value = 1.258e-07
#skew significant from zero

#shapiro-wilk test for from a normal population
bayanburd.sw <-shapiro.test(khq.mcmsqrtf[,2])
#W = 0.9688, p-value = 4.972e-08
#reject null

baidrag.sw <-shapiro.test(khp.tsqrtf[,2])
#W = 0.8884, p-value = 1.082e-11
#reject null

#the square root transformed datasets still fail tests for normalcy, lack of skew and insignificant autocorrelation,
#but are more "normal" than the untransformed datasets

#---------------------------------------------------------------------------------
#Correlations between monthly streamflow and monthly precipitation both for stations and gridded data
#import data from csvs
khq <-read.csv("Transformed_Q_1976-2012.csv")
khp <-read.csv("Transformed_P_1976-2012.csv")
gpcc <-read.csv("Transformed_GPCC_1976-2010.csv")

#for each Q site, check which P stations corr
khq1p <-data.frame(khq$BayanB, khp[,2:7])
khq2p <-data.frame(khq$BayanT, khp[,2:7])
khq3p <-data.frame(khq$ErdK, khp[,2:7])
khq4p <-data.frame(khq$IkhKT, khp[,2:7])

khq1g <-data.frame(khq[1:420,2], gpcc[,2:5])
khq2g <-data.frame(khq[1:420,3], gpcc[,2:5])
khq3g <-data.frame(khq[1:420,4], gpcc[,2:5])
khq4g <-data.frame(khq[1:420,5], gpcc[,2:5])

round(cor(khq1p, use = "pair"), 3)

#lagging correlations by 1 month
```

```
bayanb.z <-zoo(khq[,2], khq[,1])
bayanb.lag <-lag(bayanb.z, k=1) #feb Q is labled as Jan to match with P
bayant.z <-zoo(khq[,3], khq[,1])
bayant.lag <-lag(bayant.z, k=1)
erdk.z <-zoo(khq[,4], khq[,1])
erdk.lag <-lag(erdk.z, k=1)
ikhkt.z <-zoo(khq[,5], khq[,1])
ikhkt.lag <-lag(ikhkt.z, k=1)

bayanbg.z <-zoo(gpcc[,2], gpcc[,1])
bayanbg.lag <-lag(bayanbg.z, k=1)
bayantg.z <-zoo(gpcc[,3], gpcc[,1])
bayantg.lag <-lag(bayantg.z, k=1)
erdkg.z <-zoo(gpcc[,4], gpcc[,1])
erdkg.lag <-lag(erdkg.z, k=1)
ikhktg.z <-zoo(gpcc[,5], gpcc[,1])
ikhktg.lag <-lag(ikhktg.z, k=1)

khq1pl <-data.frame(bayanb.lag, khp[1:443,2:7])
khq2pl <-data.frame(bayant.lag, khp[1:443,2:7])
khq3pl <-data.frame(erdk.lag, khp[1:443,2:7])
khq4pl <-data.frame(ikhkt.lag, khp[1:443,2:7])

khq1gl <-data.frame(bayanbg.lag, gpcc[1:419,2:5])
khq2gl <-data.frame(bayantg.lag, gpcc[1:419,2:5])
khq3gl <-data.frame(erdkg.lag, gpcc[1:419,2:5])
khq4gl <-data.frame(ikhktg.lag, gpcc[1:419,2:5])

round(cor(khq1pl, use = "pair"), 3)

#lagging correlations by 2 months just to check
bayanb.z <-zoo(khq[,2], khq[,1])
bayanb.lag <-lag(bayanb.z, k=2) #mar Q is labled as Jan to match with P
bayant.z <-zoo(khq[,3], khq[,1])
bayant.lag <-lag(bayant.z, k=2)
erdk.z <-zoo(khq[,4], khq[,1])
erdk.lag <-lag(erdk.z, k=2)
ikhkt.z <-zoo(khq[,5], khq[,1])
ikhkt.lag <-lag(ikhkt.z, k=2)

khq1pl <-data.frame(bayanb.lag, khp[1:442,2:7])
khq2pl <-data.frame(bayant.lag, khp[1:442,2:7])
khq3pl <-data.frame(erdk.lag, khp[1:442,2:7])
khq4pl <-data.frame(ikhkt.lag, khp[1:442,2:7])

round(cor(khq1pl, use = "pair"), 3)

#--------------------------------------------------------------------------------------
#Filling Q data using PMM and multiple linear regression to model/predict values for Q to fill missingness
#arrange P and Q data for imputation
#stations and gages
khp1q <-data.frame(khp[,2:7],khq$BayanB)
colnames(khp1q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","BayanB")
khp2q <-data.frame(khp[,2:7],khq$BayanT)
colnames(khp2q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","BayanT")
```

```
khp3q <-data.frame(khp[,2:7],khq$ErdK)
colnames(khp3q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","ErdK")
khp4q <-data.frame(khp[,2:7],khq$IkhKT)
colnames(khp4q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","IkhKT")

#grids and gages
khqt <-khq[1:420,]
khg1q <-data.frame(gpcc[,2:5],khqt$BayanB)
colnames(khg1q) <-c("BayanB","BayanT","ErdK", "IkhKT", "BayanBQ")
khg2q <-data.frame(gpcc[,2:5],khqt$BayanT)
colnames(khg2q) <-c("BayanB","BayanT","ErdK", "IkhKT", "BayanTQ")
khg3q <-data.frame(gpcc[,2:5],khqt$ErdK)
colnames(khg3q) <-c("BayanB","BayanT","ErdK", "IkhKT", "ErdKQ")
khg4q <-data.frame(gpcc[,2:5],khqt$IkhKT)
colnames(khg4q) <-c("BayanB","BayanT","ErdK", "IkhKT", "IkhKTQ")

#complete data model (P components arranged by correlation value to Q)
#khp1q.fit <-lm(BayanB~Baidrag+Tsetserleg+Galuut+Tariat+Erdenemandal+Bayankhongor,khp1q
,na.action="na.omit")

#check quickpred
khp1q.qui <-quickpred(khp1q,minpuc=0.10)
khp2q.qui <-quickpred(khp2q,minpuc=0.10)
khp3q.qui <-quickpred(khp3q, minpuc=0.10)
khp4q.qui <-quickpred(khp4q, minpuc=0.10)

khg1q.qui <-quickpred(khg1q,minpuc=0.10)
khg2q.qui <-quickpred(khg2q,minpuc=0.10)
khg3q.qui <-quickpred(khg3q, minpuc=0.10)
khg4q.qui <-quickpred(khg4q, minpuc=0.10)



#imputations done with default of 5
khp1q.imp <-mice(khp1q, pred=khp1q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp2q.imp <-mice(khp2q, pred=khp2q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp3q.imp <-mice(khp3q, pred=khp3q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp4q.imp <-mice(khp4q, pred=khp4q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)

khg1q.imp <-mice(khg1q, pred=khg1q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khg2q.imp <-mice(khg2q, pred=khg2q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khg3q.imp <-mice(khg3q, pred=khg3q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khg4q.imp <-mice(khg4q, pred=khg4q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)

#plot of convergence
plot(khp1q.imp) #all look fair, no points of flatlining

#stripplots to look at imputed data
stripplot(khp2q.imp, pch=20, cex=1.2)
#---------------------------------------------------------------------
#fitted a model, but really this isn't what I needed,
#it would only be applicable if you had observed data to compare it to.
#what is more interesting/important are the imputed data values above.
#model fit (with P/GPCC predictors arranged by correlation with Q (ties broken by distance))
#khp1q.fit <-with(khp1q.imp,
```

```
lm(BayanB~Baidrag+Tsetserleg+Galuut+Tariat+Erdenemandal+Bayankhongor))
#khp2q.fit <-with(khp2q.imp,
lm(BayanT~Tsetserleg+Tariat+Bayankhongor+Erdenemandal+Galuut+Baidrag))
#khp3q.fit <-with(khp3q.imp, lm(ErdK~Erdenemandal+Tsetserleg+Tariat+Baidrag+Galuut+Bayankhongor))
#khp4q.fit <-with(khp4q.imp,
lm(IkhKT~Tsetserleg+Erdenemandal+Galuut+Baidrag+Tariat+Bayankhongor))

#khg1q.fit <-with(khg1q.imp, lm(BayanBQ~IkhKT+ErdK+BayanB+BayanT))
#khg2q.fit <-with(khg2q.imp, lm(BayanTQ~BayanT+IkhKT+BayanB+ErdK))
#khg3q.fit <-with(khg3q.imp, lm(ErdKQ~ErdK+IkhKT+BayanT+BayanB))
#khg4q.fit <-with(khg4q.imp, lm(IkhKTQ~IkhKT+ErdK+BayanT+BayanB))

#pooled fits
#print(pool(khp1q.fit))
#round(summary(pool(khp1q.fit)), 2)

#other info
#attributes(pool(khp1q.fit)) #same as est objects above

#-----------------------------------------------------------------------------
#collate and calculate/export the PMM results

#Make 500 imputations and collate in a df and find the mean for each to fill missing values
khp1q.imp <-mice(khp1q, m=500, pred=khp1q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp2q.imp <-mice(khp2q, m=500, pred=khp2q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp3q.imp <-mice(khp3q, m=500, pred=khp3q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp4q.imp <-mice(khp4q, m=500, pred=khp4q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)

khg1q.imp <-mice(khg1q, m=500, pred=khg1q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khg2q.imp <-mice(khg2q, m=500, pred=khg2q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khg3q.imp <-mice(khg3q, m=500, pred=khg3q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khg4q.imp <-mice(khg4q, m=500, pred=khg4q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)

#extract imputation for each basin
#vector of original data
bayanbp <-khp1q.imp$data$BayanB
bayantp <-khp2q.imp$data$BayanT
erdkp <-khp3q.imp$data$ErdK
ikhktp <-khp4q.imp$data$IkhKT

bayanbg <-khg1q.imp$data$BayanBQ
bayantg <-khg2q.imp$data$BayanTQ
erdkg <-khg3q.imp$data$ErdKQ
ikhktg <-khg4q.imp$data$IkhKTQ

#make each set a zoo object
time.z<-khq[,1]
bayanbp.z <-zoo(bayanbp, time.z)
bayantp.z <-zoo(bayantp, time.z)
erdkp.z <-zoo(erdkp, time.z)
ikhktp.z <-zoo(ikhktp, time.z)

time.z2 <-gpcc[,1]
bayanbg.z <-zoo(bayanbg, time.z2)
bayantg.z <-zoo(bayantg, time.z2)
```

```
erdkg.z <-zoo(erdkg, time.z2)
ikhktg.z <-zoo(ikhktg, time.z2)

#extract missing values and dates as irregular zoo objects
bayanbp.na <-subset(bayanbp.z, is.na(bayanbp.z))
bayantp.na <-subset(bayantp.z, is.na(bayantp.z))
erdkp.na <-subset(erdkp.z, is.na(erdkp.z))
ikhktp.na <-subset(ikhktp.z, is.na(ikhktp.z))

bayanbg.na <-subset(bayanbg.z, is.na(bayanbg.z))
bayantg.na <-subset(bayantg.z, is.na(bayantg.z))
erdkg.na <-subset(erdkg.z, is.na(erdkg.z))
ikhktg.na <-subset(ikhktg.z, is.na(ikhktg.z))

#give converse
bayanbp.nona <-subset(bayanbp.z, !is.na(bayanbp.z))
bayantp.nona <-subset(bayantp.z, !is.na(bayantp.z))
erdkp.nona <-subset(erdkp.z, !is.na(erdkp.z))
ikhktp.nona <-subset(ikhktp.z, !is.na(ikhktp.z))

bayanbg.nona <-subset(bayanbg.z, !is.na(bayanbg.z))
bayantg.nona <-subset(bayantg.z, !is.na(bayantg.z))
erdkg.nona <-subset(erdkg.z, !is.na(erdkg.z))
ikhktg.nona <-subset(ikhktg.z, !is.na(ikhktg.z))
#gives values for only the missing data in order (in this case 7 for BayanB!)
bayanbp.imp <-khp1q.imp$imp$BayanB
bayantp.imp <-khp2q.imp$imp$BayanT
erdkp.imp <-khp3q.imp$imp$ErdK
ikhktp.imp <-khp4q.imp$imp$IkhKT

bayanbg.imp <-khg1q.imp$imp$BayanBQ
bayantg.imp <-khg2q.imp$imp$BayanTQ
erdkg.imp <-khg3q.imp$imp$ErdKQ
ikhktg.imp <-khg4q.imp$imp$IkhKTQ

#find mean of all these values
bayanbp.mean <-rowMeans(bayanbp.imp)
bayantp.mean <-rowMeans(bayantp.imp)
erdkp.mean <-rowMeans(erdkp.imp)
ikhktp.mean <-rowMeans(ikhktp.imp)

bayanbg.mean <-rowMeans(bayanbg.imp)
bayantg.mean <-rowMeans(bayantg.imp)
erdkg.mean <-rowMeans(erdkg.imp)
ikhktg.mean <-rowMeans(ikhktg.imp)

#use index from na values for mean values
bayanbp.meanz <-zoo(bayanbp.mean, index(bayanbp.na))
bayantp.meanz <-zoo(bayantp.mean, index(bayantp.na))
erdkp.meanz <-zoo(erdkp.mean, index(erdkp.na))
ikhktp.meanz <-zoo(ikhktp.mean, index(ikhktp.na))

bayanbg.meanz <-zoo(bayanbg.mean, index(bayanbg.na))
bayantg.meanz <-zoo(bayantg.mean, index(bayantg.na))
erdkg.meanz <-zoo(erdkg.mean, index(erdkg.na))
```

```
ikhktg.meanz <-zoo(ikhktg.mean, index(ikhktg.na))

#join orginal values and mean values and collate and export
bayanbp.j <-zoo(rbind(bayanbp.meanz, bayanbp.nona), index(bayanbp.z))
bayantp.j <-zoo(rbind(bayantp.meanz, bayantp.nona), index(bayantp.z))
erdkp.j <-zoo(rbind(erdkp.meanz, erdkp.nona), index(erdkp.z))
ikhktp.j <-zoo(rbind(ikhktp.meanz, ikhktp.nona), index(ikhktp.z))

bayanbg.j <-zoo(rbind(bayanbg.meanz, bayanbg.nona), index(bayanbg.z))
bayantg.j <-zoo(rbind(bayantg.meanz, bayantg.nona), index(bayantg.z))
erdkg.j <-zoo(rbind(erdkg.meanz, erdkg.nona), index(erdkg.z))
ikhktg.j <-zoo(rbind(ikhktg.meanz, ikhktg.nona), index(ikhktg.z))

khqfillp <-data.frame(bayanbp.j,bayantp.j,erdkp.j,ikhktp.j)
colnames(khqfillp) <-c("BayanB", "BayanT", "ErdK", "IkhKT")
write.csv(khqfillp, "Filled_SQRT_Q_from_P_1976-2012.csv")

khqfillgpcc <-data.frame(bayanbg.j,bayantg.j,erdkg.j,ikhktg.j)
colnames(khqfillgpcc) <-c("BayanB", "BayanT", "ErdK", "IkhKT")
write.csv(khqfillgpcc, "Filled_SQRT_Q_from_GPCC_1976-2010.csv")

#collate mean filled values for comparison
bayanbfillpgpcc <- data.frame(bayanbp.meanz,bayanbg.meanz)
colnames(bayanbfillpgpcc) <-c("Station", "Grid")
write.csv(bayanbfillpgpcc, "Filled_P_GPCC_Bayanburd.csv")

bayantfillpgpcc <-data.frame(bayantp.meanz,bayantg.meanz)
colnames(bayantfillpgpcc) <-c("Station", "Grid")
write.csv(bayantfillpgpcc, "Filled_P_GPCC_Bayankhongor.csv")

#truncate to same length
erdkp.meanz2 <-erdkp.meanz[1:97,]
erdkfillpgpcc <-data.frame(erdkp.meanz2,erdkg.meanz)
colnames(erdkfillpgpcc) <-c("Station", "Grid")
write.csv(erdkfillpgpcc, "Filled_P_GPCC_Erdenemandal.csv")

ikhktp.meanz2 <-ikhktp.meanz[1:96,]
ikhktfillpgpcc <-data.frame(ikhktp.meanz2,ikhktg.meanz)
colnames(ikhktfillpgpcc) <-c("Station", "Grid")
write.csv(ikhktfillpgpcc, "Filled_P_GPCC_Ikhtamir.csv")

#-----------------------------------------------------------------------------------------------
#Are there significant differences between the mean filled values in each dataset?
#import csvs for comparison
bayanb.c <-read.csv("Filled_P_GPCC_Bayanburd.csv")
bayant.c <-read.csv("Filled_P_GPCC_Bayankhongor.csv")
erdk.c <-read.csv("Filled_P_GPCC_Erdenemandal.csv")
ikhkt.c <-read.csv("Filled_P_GPCC_Ikhtamir.csv")

#use Kolmogorov-Smirnov test calcs max absolute difference between empirical cdfs of each dataset
plot(ecdf(bayanb.c[,2]))
plot(ecdf(bayanb.c[,3]), add=TRUE)
ks.test(bayanb.c[,2], bayanb.c[,3])
ks.test(bayant.c[,2], bayant.c[,3])
ks.test(erdk.c[,2], erdk.c[,3])
```

```
ks.test(ikhkt.c[,2], ikhkt.c[,3])

#results of all testing for liklihood of two samples being drawn from the same distribution
#suggest that at the 0.05 or even smaller level, cannot reject the null.
#that the two samples were drawn from the same continous distribution (which is logical since the GPCC values were created from the P values of the region)

#PMM method therefore performs reasonably well with full filled predictors (GPCC) and predictors with missing data to fill Q data.
#-------------------------------------------------------
#check results of PMM on untransformed data
#Filling Q data using PMM
#arrange P and Q data for imputation
#using khq.mcm and khp.t and not checking GPCC as it was similar with transformed data.
#Comparing transformed and untransformed results.

#stations and gages
khp1q <-data.frame(khp.t[,2:7],khq.mcm$BayanB)
colnames(khp1q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","BayanB")
khp2q <-data.frame(khp.t[,2:7],khq.mcm$BayanT)
colnames(khp2q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","BayanT")
khp3q <-data.frame(khp.t[,2:7],khq.mcm$ErdK)
colnames(khp3q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","ErdK")
khp4q <-data.frame(khp.t[,2:7],khq.mcm$IkhKT)
colnames(khp4q) <-c("Baidrag","Bayankhongor","Erdenemandal", "Galuut", "Tariat","Tsetserleg","IkhKT")

#check quickpred
khp1q.qui <-quickpred(khp1q,minpuc=0.10)
khp2q.qui <-quickpred(khp2q,minpuc=0.10)
khp3q.qui <-quickpred(khp3q, minpuc=0.10)
khp4q.qui <-quickpred(khp4q, minpuc=0.10)

#Make 500 imputations and collate in a df and find the mean for each to fill missing values
khp1q.imp <-mice(khp1q, m=500, pred=khp1q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp2q.imp <-mice(khp2q, m=500, pred=khp2q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp3q.imp <-mice(khp3q, m=500, pred=khp3q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)
khp4q.imp <-mice(khp4q, m=500, pred=khp4q.qui, method="pmm",pri=FALSE, seed=210, maxit=100)

#extract imputation for each basin
#vector of original data
bayanbp <-khp1q.imp$data$BayanB
bayantp <-khp2q.imp$data$BayanT
erdkp <-khp3q.imp$data$ErdK
ikhktp <-khp4q.imp$data$IkhKT

#make each set a zoo object
time.z<-khq.mcm[,1]
bayanbp.z <-zoo(bayanbp, time.z)
bayantp.z <-zoo(bayantp, time.z)
erdkp.z <-zoo(erdkp, time.z)
ikhktp.z <-zoo(ikhktp, time.z)

#extract missing values and dates as irregular zoo objects
bayanbp.na <-subset(bayanbp.z, is.na(bayanbp.z))
bayantp.na <-subset(bayantp.z, is.na(bayantp.z))
```

```
erdkp.na <-subset(erdkp.z, is.na(erdkp.z))
ikhktp.na <-subset(ikhktp.z, is.na(ikhktp.z))

#give converse
bayanbp.nona <-subset(bayanbp.z, !is.na(bayanbp.z))
bayantp.nona <-subset(bayantp.z, !is.na(bayantp.z))
erdkp.nona <-subset(erdkp.z, !is.na(erdkp.z))
ikhktp.nona <-subset(ikhktp.z, !is.na(ikhktp.z))

#gives values for only the missing data in order (in this case 7 for BayanB!)
bayanbp.imp <-khp1q.imp$imp$BayanB
bayantp.imp <-khp2q.imp$imp$BayanT
erdkp.imp <-khp3q.imp$imp$ErdK
ikhktp.imp <-khp4q.imp$imp$IkhKT


#find mean of all these values
bayanbp.mean <-rowMeans(bayanbp.imp)
bayantp.mean <-rowMeans(bayantp.imp)
erdkp.mean <-rowMeans(erdkp.imp)
ikhktp.mean <-rowMeans(ikhktp.imp)

#use index from na values for mean values
bayanbp.meanz <-zoo(bayanbp.mean, index(bayanbp.na))
bayantp.meanz <-zoo(bayantp.mean, index(bayantp.na))
erdkp.meanz <-zoo(erdkp.mean, index(erdkp.na))
ikhktp.meanz <-zoo(ikhktp.mean, index(ikhktp.na))

#join orginal values and mean values and collate and export
bayanbp.j <-zoo(rbind(bayanbp.meanz, bayanbp.nona), index(bayanbp.z))
bayantp.j <-zoo(rbind(bayantp.meanz, bayantp.nona), index(bayantp.z))
erdkp.j <-zoo(rbind(erdkp.meanz, erdkp.nona), index(erdkp.z))
ikhktp.j <-zoo(rbind(ikhktp.meanz, ikhktp.nona), index(ikhktp.z))

khqfillp <-data.frame(bayanbp.j,bayantp.j,erdkp.j,ikhktp.j)
colnames(khqfillp) <-c("BayanB", "BayanT", "ErdK", "IkhKT")
write.csv(khqfillp, "Untransformed_Filled_Q_from_P_1976-2012.csv")

#collate mean filled values for comparison from transformed (inverted) filled and untransformed
#read transformed values and invert

trans <-read.csv("Filled_Q_from_P_1976-2012.csv")

bayanbt.meanz <-(trans[,2])^2
bayantt.meanz <-(trans[,3])^2
erdkt.meanz <-(trans[,4])^2
ikhktt.meanz <-(trans[,5])^2

bayanbfillcomp <- data.frame(bayanbt.meanz,khqfillp$BayanB)
colnames(bayanbfillcomp) <-c("Transformed", "Untransformed")
write.csv(bayanbfillcomp, "Filled_Transformed_Untransformed_Bayanburd.csv")

bayantfillcomp <-data.frame(bayantt.meanz,khqfillp$BayanT)
colnames(bayantfillcomp) <-c("Transformed", "Untransformed")
write.csv(bayantfillcomp, "Filled_Transformed_Untransformed_Bayankhongor.csv")
```

```
erdkfillcomp <-data.frame(erdkt.meanz,erdkp.meanz)
colnames(erdkfillcomp) <-c("Transformed", "Untransformed")
write.csv(erdkfillcomp, "Filled_Transformed_Untransformed_Erdenemandal.csv")

ikhktfillcomp <-data.frame(ikhktt.meanz,ikhktp.meanz)
colnames(ikhktfillcomp) <-c("Transformed", "Untransformed")
write.csv(ikhktfillcomp, "Filled_Transformed_Untransformed_Ikhtamir.csv")

#comparison of results using kolmogorov-smirnov
bayanb.c <-read.csv("Filled_Transformed_Untransformed_Bayanburd.csv")
bayant.c <-read.csv("Filled_Transformed_Untransformed_Bayankhongor.csv")
erdk.c <-read.csv("Filled_Transformed_Untransformed_Erdenemandal.csv")
ikhkt.c <-read.csv("Filled_Transformed_Untransformed_Ikhtamir.csv")


#use Kolmogorov-Smirnov test calcs max absolute difference between empirical cdfs of each dataset
plot(ecdf(bayanb.c[,2]))
plot(ecdf(bayanb.c[,3]), add=TRUE)
ks.test(bayanb.c[,2], bayanb.c[,3])
ks.test(bayant.c[,2], bayant.c[,3])
ks.test(erdk.c[,2], erdk.c[,3])
ks.test(ikhkt.c[,2], ikhkt.c[,3])
```

D.1.3 Khangai Mountain Region Tree-Ring Crossdating and Analyses

```
#------------------------------------------------------------------------
#  TITLE: Khangai_TR_CrossTrend
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2015-08-19
#  LAST REVISION: 2015-10-13
#  DESCRIPTION: This script provides code for checking crossdating of Khangai total ring width
measurements and detrending to chronologies
#  PACKAGES REQUIRED: dplR
#  VARIABLES/DATA USED:
#  NAME:
# TYPE:
# COMMENT:
#-------------------------------------------------------------------------

#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detren
d/")

#libraries
library(dplR)

#note: goal here is to create final chronologies from series that are most representative of/sensitive to the
common signal (assuming that signal is climatic in nature).
#--------------------------------------------------------------
#import raw ring widths and resolve issues with files/series
#per conversation with Peter and Cari- want longest segment lengths possible see if can trncate to 200 years
or more
#include everything that isn't a complete problem as biweight mean will minimize these problems
#truncate methods from below (at end of code) to those applicable to ideas from changes above
```

```
#JGB
jgb.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Raw_Cores/JGB/JGB_Total_Widths.rwl")

#correlation with master
jgb.corrpw <-corr.rwl.seg(jgb.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

jgb.corrplotpw21 <-series.rwl.plot(jgb.rwl, "JB_21B",seg.length=20, bin.floor=20)
jgb.corrplotpw22 <-series.rwl.plot(jgb.rwl, "JB_22B",seg.length=20, bin.floor=20)
jgb.corrplotpw08 <-series.rwl.plot(jgb.rwl, "JB_08A",seg.length=20, bin.floor=20)
jgb.corrplotpw12 <-series.rwl.plot(jgb.rwl, "JB_12B",seg.length=50, bin.floor=50)
jgb.corrplotpw23 <-series.rwl.plot(jgb.rwl, "JB_23B",seg.length=20, bin.floor=20)

#12B seems most likely an error
#extract values for series and master from 1673 to 1750 and compare
plot(jgb.corrplotpw12$series[1:100], type="b", col="red", pch=19)
lines(jgb.corrplotpw12$master[1:100], type="l", col="blue")

#extract values for series and master from 1960 to 2009 and compare
plot(jgb.corrplotpw23$series[232:282], type="b", col="red", pch=19)
lines(jgb.corrplotpw23$master[232:282], type="b", col="blue")

#truncate series 12B prior to 1750 by making NA to that point
jgb.rwlt <-jgb.rwl
jgb.rwlt[1:291,46] <-NA

#check result
jgb.corrpw <-corr.rwl.seg(jgb.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)

rwl.stats(jgb.rwlt)

plot(jgb.rwlt, plot.type="spag")

#correct series names to 1-6 characters
colnames(jgb.rwlt)
colnames(jgb.rwlt) <-c("JB21A","JB21B", "JB17B",  "JB17A", "JB20A","JB20B","JB22B",
"JB22A","JB06A","JB06B","JB05A","JB05B", "JB02A","JB02B",  "JB15A",  "JB15B","JB08A",  "JB08B",  "JB10A1",
"JB10A2","JB10B", "JB16B",  "JB14A",  "JB14B", "JB13A",  "JB13B",  "JB01A",  "JB01B","JB18A",  "JB18B",
"JB19A",  "JB19B", "JB24A",  "JB24B",  "JB16A",  "JB25A1", "JB25A2", "JB25B",  "JB25C",  "JB09A", "JB09C",
"JB11A",  "JB11B",  "JB04C","JB12A",  "JB12B",  "JB23B",  "JB23A", "JB09B",  "JB07A",  "JB07B",  "JB04B",
"JB04A",  "JB03A",  "JB03B" )

#Subdivide JGB to 2 datasets, one with all cores beginning prior to 1800 and all after 1800.
#find rownumber for 1800-1801
which(rownames(jgb.rwlt)>1800)

jgb.1800 <-jgb.rwlt[342,]
#select those series that are NA at 1800
which(is.na(jgb.1800))
#note the 25A1 section is not really "young"- remove it from final split
jgb.young <-jgb.rwlt[,c(1:18, 20, 23:24, 26:32, 37, 43, 52, 54:55)]
jgb.youngstats <-rwl.stats(jgb.young)
jgb.corrpwyoung <-corr.rwl.seg(jgb.young, seg.length=50, bin.floor=20, pcrit=0.01)
```

```
plot(jgb.young[343:553,])

plot(jgb.young[343:553,], plot.type="spag")

which(!is.na(jgb.1800))
jgb.old <-jgb.rwlt[,c(19,21:22,25,33:35, 38:42, 44:51, 53)]
jgb.corrpwold <-corr.rwl.seg(jgb.old, seg.length=50, bin.floor=20, pcrit=0.01)

jgb.oldstats <-rwl.stats(jgb.old)

#series lengths
jgb.series <-jgb.oldstats$year
names(jgb.series) <-jgb.oldstats$series

plot(jgb.old, plot.type="spag")

#common period "old" cores: 1727-2010
jgbold.comm <-common.interval(jgb.old, type="years")

#note no truncation by correlation amount
write.rwl(jgb.old,"JGB_AdjTWf.rwl" )

#KHU
khu.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detr
end/Raw_Cores/KHU/KHU_Total_Widths_Final.rwl")
khu.corrpw <-corr.rwl.seg(khu.rwl, seg.length=50, bin.floor=20, pcrit=0.01)
#no problems in cores shown at 50-year segment lengths, if shorten to 30 years prewhitened a few problems
show up, checking them below.
khu.corrpw30 <-corr.rwl.seg(khu.rwl, seg.length=30, bin.floor=20, pcrit=0.01)

#examine correlation of the master and each problematic series using series.rwl.plot
khu.corrplotpw07 <-series.rwl.plot(khu.rwl, "KU_07A",seg.length=30, bin.floor=20)
khu.corrplotpw05 <-series.rwl.plot(khu.rwl, "KU_05A",seg.length=30, bin.floor=20)

#extract values for series and master 1835-1899 and compare
plot(khu.corrplotpw07$series[36:100], type="b", col="red", pch=19)
lines(khu.corrplotpw07$master[36:100], type="l", col="blue")

#extract values for series and master 1800-1839 and compare
plot(khu.corrplotpw05$series[1:40], type="b", col="red", pch=19)
lines(khu.corrplotpw05$master[1:40], type="l", col="blue")

#truncate series 05A and 05B prior to 1830 by making NA to that point
khu.rwlt2 <-khu.rwl
khu.rwlt2[1:30,35] <-NA
khu.rwlt2[1:30,36] <-NA

#check result
khu.corrpw2 <-corr.rwl.seg(khu.rwlt2, seg.length=30, bin.floor=20, pcrit=0.01)

plot(khu.rwlt2, plot.type="spag")

#correct headings
colnames(khu.rwlt2) <-c("KU02A", "KU02B", "KU03A", "KU03B", "KU04A","KU04B", "KU06A", "KU06B",
```

```
"KU08A", "KU08B", "KU10A", "KU10B", "KU12A", "KU12B", "KU13A","KU13B", "KU14A", "KU14B", "KU15A",
"KU15B","KU16A", "KU16B", "KU17A", "KU17B", "KU18A", "KU18B", "KU19A", "KU19B", "KU20A",
"KU20B","KU21A", "KU21B", "KU22A", "KU22B", "KU05A","KU05B", "KU07C", "KU07B", "KU07A")

#find common period
khu.comm <-common.interval(khu.rwlt2, type="years")

khu.stats <-rwl.stats(khu.rwlt2)

#series lengths
khu.series <-khu.stats$year
names(khu.series) <-khu.stats$series

#truncate to longer cores (100+ years at minimum) but try to retain some depth of series
khu.1911 <-khu.rwlt2[112,]

#remove shorter than 100 years
khu.older <-khu.rwlt2[,c(3:8, 11:14, 17,19:20, 23:32,35:36,39)]
khu.olderstats <-rwl.stats(khu.older)

#series lengths
khu.series <-khu.olderstats$last-khu.olderstats$first
names(khu.series) <-khu.olderstats$series

#find common period #1911-2011
khu.commold <-common.interval(khu.older, type="years")

khu.corrpw2 <-corr.rwl.seg(khu.older, seg.length=50, bin.floor=20, pcrit=0.01)

write.rwl(khu.older, "KHU_AdjTWf.rwl")

khu.corrpwold <-corr.rwl.seg(khu.older, seg.length=50, bin.floor=20, pcrit=0.01)

#KLL
kll.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detr
end/Raw_Cores/KLL/KLL_mong015.rwl")
kll.corrpw <-corr.rwl.seg(kll.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

#notes:
#few problems in cores shown at 50-year segment lengths
#TB02E and TB02N- added to coreset later, would have to be re-named to include
#check 43N and 43E- possible endogenous distrurbance patterns?

#examine correlation of the master and each problematic series using series.rwl.plot
kll.corrplotpw43n <-series.rwl.plot(kll.rwl, "KL43N",seg.length=50, bin.floor=20)
kll.corrplotpw43e <-series.rwl.plot(kll.rwl, "KL43E",seg.length=50, bin.floor=20)

#extract values for series and master 1790-1840 and compare
plot(kll.corrplotpw43e$series[147:197], type="b", col="red", pch=19)
lines(kll.corrplotpw43e$master[147:197], type="l", col="blue")

#extract values for series and master 1940-1988 and compare
plot(kll.corrplotpw43n$series[280:328], type="b", col="red", pch=19)
lines(kll.corrplotpw43n$master[280:328], type="l", col="blue")
```

```
lines(kll.corrplotpw43e$series[297:345], type="l", col="green")

#remove TB02E and TB02N, leave 43 N and E in for now.
kll.rwlt <-kll.rwl[,3:66]

#check result
kll.corrpw2 <-corr.rwl.seg(kll.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)

#correct headings
#Naming convention will be used that first core in series is A, next is B, etc.
colnames(kll.rwlt) <-c("KL01A","KL01B", "KL03A", "KL04A","KL04B","KL05A","KL06A","KL07A",
"KL07B","KL08A","KL10A",  "KL10B",  "KL10C","KL12A",  "KL13A","KL15A","KL15B",
"KL21A","KL21B","KL22A","KL22B",  "KL23A",  "KL24A", "KL24B",  "KL25A", "KL26A",  "KL26B",   "KL27A",
"KL27B","KL27C","KL28A","KL29A", "KL30A", "KL35A","KL36A","KL38A",  "KL38B","KL40A", "KL40B",
"KL43A","KL43B", "KL45A","KL44A","KL45B", "KL48A", "KL48B","KL49A","KL49B", "KL50A",
"KL51A","KL51B","KL53A","KL53B" ,"KL54A","KL54B",  "KL55A",  "KL55B", "KL56A", "KL56B",
"KL62A","KL63A",  "KL63B",  "KL64A",  "KL64B")

#series lengths
kll.stats <-rwl.stats(kll.rwlt)
kll.series <-kll.stats$year
names(kll.series) <-kll.stats$series

#remove any series less than 200 years long

kll.rwlt1 <-kll.rwlt[ ,c(3:16, 18:28,30:33,35:56,58:59,63:64)]
kll.stats <-rwl.stats(kll.rwlt1)

write.rwl(kll.rwlt1, "KLL_AdjTWf.rwl")

kll.corrpwold <-corr.rwl.seg(kll.rwlt1, seg.length=50, bin.floor=20, pcrit=0.01)

#KLP
klp.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detr
end/Raw_Cores/KLP/KLP_mong039.rwl")

#truncate chronology to those cores that at least extend to the 21st century to compare to other core sets
klp.rwlt <-klp.rwl[,c(1:34, 42)]

#truncate years for better plotting
klp.rwltp <-klp.rwlt[1507:2118,]
plot(klp.rwltp, plot.type="spag")

#method for correlation = Spearman by default even though COFECHA uses Pearson as the data may not be
normally distributed.
klp.corrpw <-corr.rwl.seg(klp.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)
klp.corrpw <-corr.rwl.seg(klp.rwltp, seg.length=50, bin.floor=20, pcrit=0.01)

#correct headings
colnames(klp.rwltp) <-c("KP001A", "KP001B", "KP002A", "KP003A","KP003B","KP004B",
"KP007A","KP007B", "KP011A","KP011B", "KP013A", "KP013B","KP014A", "KP014B", "KP015A", "KP015B",
"KP016A", "KP018A", "KP018B", "KP050A","KP052B","KP054A", "KP054B", "KP056A","KP056B", "KP058A",
"KP058B", "KP061B", "KP062A", "KP063A", "KP064A", "KP065B","KP066B", "KP069A", "KP209A")
```

```
#find common period 1760-1993
klp.comm <-common.interval(klp.rwltp, type="years")

#series lengths
klp.stats <-rwl.stats(klp.rwlt)

#no core series removed for adj set
write.rwl(klp.rwltp, "KLP_AdjTWf.rwl")

klp.corrpwold <-corr.rwl.seg(klp.rwltp, seg.length=50, bin.floor=20, pcrit=0.01)

#MHM
mhm.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detr
end/Raw_Cores/MHM/MHM_mong026.rwl")

#truncate chronology to those cores that at least extend to the 21st century to compare to other core sets
mhm.rwlt <-mhm.rwl[,c(1:20)]

mhm.corrpw <-corr.rwl.seg(mhm.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)

#examine correlation of the master and each problematic series using series.rwl.plot
mhm.corrplotpw10 <-series.rwl.plot(mhm.rwlt, "MU10A",seg.length=50, bin.floor=20)
mhm.corrplotpw06 <-series.rwl.plot(mhm.rwlt, "MU06N",seg.length=50, bin.floor=20)

#extract values for series and master 1545-1595 and compare-seems 1 yr off
plot(mhm.corrplotpw10$series[96:146], type="b", col="red", pch=19)
lines(mhm.corrplotpw10$master[96:146], type="l", col="blue")

#extract values for series and master 1905-1935 and compare- seems overall ok
plot(mhm.corrplotpw06$series[35:65], type="b", col="red", pch=19)
lines(mhm.corrplotpw06$master[35:65], type="l", col="blue")

#truncate MU10A at 1585
mhm.rwlt2 <-mhm.rwlt
mhm.rwlt2[1:135,17] <-NA


#correct headings
colnames(mhm.rwlt2) <-c("MU01A", "MU01B","MU02A",  "MU02B",  "MU03A","MU03B",  "MU04A", "MU04B",
"MU05B",  "MU05A",  "MU06A",  "MU07A",  "MU07B",  "MU08A",  "MU08B",  "MU09A",  "MU10A",  "MU12A",
"MU12B",  "MU14A")

mhm.stats <-rwl.stats(mhm.rwlt2)

#retain all cores >185 years
mhm.rwlt3 <-mhm.rwlt2[,c(1:2,5:10,12:13,15:20)]
mhm.stats <-rwl.stats(mhm.rwlt3)

write.rwl(mhm.rwlt3, "MHM_AdjTWf.rwl")

mhm.corrpwold <-corr.rwl.seg(mhm.rwlt3, seg.length=50, bin.floor=20, pcrit=0.01)

#NDB
ndb.rwl <-
```

```
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Raw_Cores/NDB/NDB_mong027.rwl")

ndb.corrpw <-corr.rwl.seg(ndb.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

#examine correlation of the master and each problematic series using series.rwl.plot
ndb.corrplotpw4 <-series.rwl.plot(ndb.rwl, "ND04N",seg.length=50, bin.floor=20)
ndb.corrplotpw5 <-series.rwl.plot(ndb.rwl, "ND05N",seg.length=50, bin.floor=20)# probs in the 1830's-maybe a year different than the master?
ndb.corrplotpw14 <-series.rwl.plot(ndb.rwl, "ND14BX",seg.length=50, bin.floor=20)
ndb.corrplotpw16 <-series.rwl.plot(ndb.rwl, "ND16A",seg.length=50, bin.floor=20)
ndb.corrplotpw7 <-series.rwl.plot(ndb.rwl, "ND07N",seg.length=50, bin.floor=20)

#extract values for series and master 1810-1840 and compare- possibly 1 or more years off
plot(ndb.corrplotpw5$series[143:173], type="b", col="red", pch=19)
lines(ndb.corrplotpw5$master[143:173], type="l", col="blue")

#series and master 1790-1830 and compare- overall better agreement than 5N
plot(ndb.corrplotpw16$series[172:212], type="b", col="red", pch=19)
lines(ndb.corrplotpw16$master[172:212], type="b", col="blue")

#series and master 1710-1760 and compare- overall better agreement than 5N
plot(ndb.corrplotpw7$series[90:140], type="b", col="red", pch=19)
lines(ndb.corrplotpw7$master[90:140], type="b", col="blue")

#correct headings
colnames(ndb.rwl) <-c("ND01A", "ND01B", "ND02A", "ND02B", "ND03A", "ND04A" ,"ND05A", "ND06A",
"ND07A", "ND11A","ND11B", "ND12A", "ND12B", "ND13B", "ND13A","ND14B", "ND14A", "ND15A",
"ND15B", "ND16B", "ND16A")

#reexamine 5N
#extract values for series and master 1800-1860 and compare- possibly 1 or more years off between 1810 and 1840
plot(ndb.corrplotpw5$series[133:193], type="b", col="red", pch=19)
lines(ndb.corrplotpw5$master[133:193], type="l", col="blue")

ndb.stats <-rwl.stats(ndb.rwl)

#retain all cores >198 years
ndb.rwlt <-ndb.rwl[,c(1:7,9:14,16,18:21)]
ndb.stats <-rwl.stats(ndb.rwlt)

#Do not remove 5N (5A) as only core for that tree- hopefully robust mean will reduce influence
write.rwl(ndb.rwlt, "NDB_AdjTWf.rwl")

ndb.corrpwold <-corr.rwl.seg(ndb.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)

#OGH
#Differing precision used in ring measurements may be able to convert file tpes and combine and import (per Peter)
#Import OGH as two "sites" dividing based on precision of record
ogh1.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Raw_Cores/OGH/OGH_Split/OGH_Final1_ImportR.rwl")
ogh2.rwl <-
```

```
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detr
end/Raw_Cores/OGH/OGH_Split/OGH_Final2_ImportR.rwl")

#plot each (note 15 a has a string of 0 values- missing data- should be truncated?)
plot(ogh1.rwl, plot.type="spag")
plot(ogh2.rwl, plot.type="spag")

#export in compact file format with 0.01 precision and append second file
write.compact(ogh1.rwl, "OGH1_compact.rwl", prec=0.01)
write.compact(ogh2.rwl, "OGH1_compact.rwl", prec=0.01, append=TRUE)

#read in
ogh.rwl<-
read.compact("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_
Detrend/Raw_Cores/OGH/OGH_Split/OGH1_compact.rwl")
plot(ogh.rwl, plot.type="spag")

ogh.corrpw <-corr.rwl.seg(ogh.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

ogh.corrplotpw1 <-series.rwl.plot(ogh.rwl, "OG01b",seg.length=50, bin.floor=20)

#extract values for series and master 1550-1610 and compare-
plot(ogh.corrplotpw1$series[1:68], type="b", col="red", pch=19)
lines(ogh.corrplotpw1$master[1:68], type="l", col="blue")
#overall agreement so not altering

#correct headings
colnames(ogh.rwl) <-c("OG01A",  "OG02A", "OG04A","OG09B", "OG09C","OG13B","OG17A",  "OG17B",
"OG01B", "OG02B","OG03A", "OG05A", "OG07A",  "OG07B",  "OG08A1", "OG08A2", "OG08B1", "OG08B2",
"OG09A",  "OG11A",  "OG11B",  "OG13A", "OG15A",  "OG15B",  "OG16A", "OG18A",  "OG18B",  "OG18C",
"OG19A")

#truncate 15A at 0's
ogh.rwl[1:259,23] <-NA

#check ending lengths
ogh.rwl[,c(15,17,19:21)]

#remove 08 cores
ogh.rwl1 <-ogh.rwl[,c(1:14,19:29)]

plot(ogh.rwl1, plot.type="spag")

#find common period 1810-1991
ogh.comm <-common.interval(ogh.rwl1, type="years")

#check lengths
ogh.stats <-rwl.stats(ogh.rwl1)

#output as adjusted rwl file
write.rwl(ogh.rwl1, "OGH_AdjTWf.rwl")

ogh.corrpwold <-corr.rwl.seg(ogh.rwl1, seg.length=50, bin.floor=20, pcrit=0.01)

#SLB
```

```
slb.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Raw_Cores/SLB/SLB_mong010.rwl")

slb.corrpw <-corr.rwl.seg(slb.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

#no problems in cores at 50 year segments, remove short segment
slb.rwlt <-slb.rwl[,c(1:22, 24:33)]

colnames(slb.rwlt) <-c("SB01A",  "SB01B", "SB03A","SB03B",  "SB07A", "SB07B", "SB08A",
"SB08B","SB09A","SB09B", "SB11A", "SB11B",  "SB11C",  "SB13A",  "SB13B",  "SB13C",  "SB15A",  "SB15B",
"SB20A",  "SB20B", "SB21A", "SB21B",  "SB22A",  "SB22B",  "SB22C", "SB24A",  "SB24B",  "SB25A", "SB25B",
"SB25C", "SB27A",  "SB27B")

slb.stats<-rwl.stats(slb.rwlt)

write.rwl(slb.rwlt, "SLB_AdjTWf.rwl")

slb.corrpwold <-corr.rwl.seg(slb.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)

#ZSM
#possible differing precision in record, but may just be due to fast-growng younger trees. Divided one core
(prob rot)
zsm.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Raw_Cores/ZSM/ZSM_mong011.rwl")

zsm.corrpw <-corr.rwl.seg(zsm.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

zsm.corrplotpw2e <-series.rwl.plot(zsm.rwl, "ZS02E",seg.length=50, bin.floor=20)

#extract values for series and master 1527-1660 and compare-
plot(zsm.corrplotpw2e$series[1:134], type="b", col="red", pch=19)
lines(zsm.corrplotpw2e$master[1:134], type="l", col="blue")
#decent agreement so not altering

#correct headings
colnames(zsm.rwl) <-c("ZS01A",  "ZS01B", "ZS02A",  "ZS02B",  "ZS03A","ZS03B", "ZS04A", "ZS04B", "ZS05A",
"ZS05B","ZS06A",  "ZS06B",  "ZS07A", "ZS07B", "ZS08A", "ZS08B",  "ZS09A",  "ZS09B",  "ZS10A",  "ZS10B",
"ZS11A", "ZS11B",  "ZS12A",  "ZS13A",  "ZS13B", "ZS14A",  "ZS14B",  "ZS15A",  "ZS15B",  "ZS20A", "ZS20B",
"ZS22A",  "ZS23A",  "ZS24A",  "ZS25A", "ZS25B")

#find common period 1800-2000
zsm.comm <-common.interval(zsm.rwl, type="years")

zsm.stats <-rwl.stats(zsm.rwl)

#output as adjusted rwl file
write.rwl(zsm.rwl, "ZSM_AdjTWf.rwl")

zsm.corrpwold <-corr.rwl.seg(zsm.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

#possible differing precision in record, but may just be due to fast-growng younger trees. Divided one core
(prob rot)
ztg.rwl <-
```

```
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Raw_Cores/ZTG/ZTG_mong032.rwl")

ztg.corrpw <-corr.rwl.seg(ztg.rwl, seg.length=50, bin.floor=20, pcrit=0.01)

#many possible problems with cores but could be an just issue with the site, leaving all in for now

#correct headings
colnames(ztg.rwl) <-c("ZT01A", "ZT01B", "ZT02A",  "ZT02B",  "ZT02C", "ZT02D",  "ZT03A",  "ZT03B",
"ZT03C",  "ZT03D",  "ZT04A",  "ZT04B",  "ZT04C",  "ZT04D",  "ZT05A", "ZT05B",  "ZT05C", "ZT06A",  "ZT06B",
"ZT07G", "ZT07A",  "ZT07B",  "ZT07C",  "ZT07D",  "ZT07E","ZT07F",  "ZT08D",   "ZT08A",  "ZT08B",  "ZT08C",
"ZT09A",  "ZT10A",  "ZT10B",  "ZT11A",  "ZT11B", "ZT12A", "ZT12B", "ZT13C",  "ZT14A",  "ZT14B",  "ZT15A",
"ZT15B", "ZT16A", "ZT17A", "ZT17B",  "ZT18A","ZT19A"  )

#find common period 1876-2002
ztg.comm <-common.interval(ztg.rwl, type="years")

ztg.stats <-rwl.stats(ztg.rwl)

#truncate to at least 193 years
ztg.rwlt<- ztg.rwl[,c(3:4,7:8,16:19,21:22,24:29,31:47)]
ztg.stats <-rwl.stats(ztg.rwlt)

#output as adjusted rwl file
write.rwl(ztg.rwlt, "ZTG_AdjTWf.rwl")

ztg.corrpwold <-corr.rwl.seg(ztg.rwlt, seg.length=50, bin.floor=20, pcrit=0.01)

#-----------------------------------------------------------------------------------------------------------
#load adjusted raw ring width files for detrending
jgb.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/JGB_AdjTWf.rwl")
khu.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/KHU_AdjTWf.rwl")
kll.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/KLL_AdjTWf.rwl")
klp.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/KLP_AdjTWf.rwl")
mhm.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/MHM_AdjTWf.rwl")
ndb.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/NDB_AdjTWf.rwl")
ogh.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/OGH_AdjTWf.rwl")
slb.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/SLB_AdjTWf.rwl")
zsm.rwl <-
```

```
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/ZSM_AdjTWf.rwl")
ztg.rwl <-
read.rwl("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Crossdating_Detrend/Adjust_Raw_Ring_Widths_Final/ZTG_AdjTWf.rwl")

#detrend each series, use a uniform method for all for reproducability (per Peter)
#a spline fit with a long period of 220 years

jgb.rwi <-detrend(jgb.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
khu.rwi <-detrend(khu.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
kll.rwi <-detrend(kll.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
klp.rwi <-detrend(klp.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
mhm.rwi <-detrend(mhm.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
ndb.rwi <-detrend(ndb.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
ogh.rwi <-detrend(ogh.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
slb.rwi <-detrend(slb.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
zsm.rwi <-detrend(zsm.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )
ztg.rwi <-detrend(ztg.rwl, make.plot=FALSE, method="Spline", nyrs=220, f=0.5 )

#write detrended files
write.csv(ztg.rwi, "ZTG_Spline_220.csv")

#create ids for each site from rwl for stats
jgb.ids <-read.ids(jgb.rwl, stc=c(2,2,3))
khu.ids <-read.ids(khu.rwl, stc=c(2,2,3))
kll.ids <-read.ids(kll.rwl, stc=c(2,2,3))
klp.ids <-read.ids(klp.rwl, stc=c(2,3,3))
mhm.ids <-read.ids(mhm.rwl, stc=c(2,2,3))
ndb.ids <-read.ids(ndb.rwl, stc=c(2,2,3))
ogh.ids <-read.ids(ogh.rwl, stc=c(2,2,3))
slb.ids <-read.ids(slb.rwl, stc=c(2,2,3))
zsm.ids <-read.ids(zsm.rwl, stc=c(2,2,3))
ztg.ids <-read.ids(ztg.rwl, stc=c(2,2,3))

#get detrend stats (over whole series)
jgb.stat <-rwi.stats(jgb.rwi, jgb.ids)
khu.stat <-rwi.stats(khu.rwi, khu.ids)
kll.stat <-rwi.stats(kll.rwi, kll.ids)
klp.stat <-rwi.stats(klp.rwi, klp.ids)
mhm.stat <-rwi.stats(mhm.rwi, mhm.ids)
ndb.stat <-rwi.stats(ndb.rwi, ndb.ids)
ogh.stat <-rwi.stats(ogh.rwi, ogh.ids)
slb.stat <-rwi.stats(slb.rwi, slb.ids)
zsm.stat <-rwi.stats(zsm.rwi, zsm.ids)
ztg.stat <-rwi.stats(ztg.rwi, ztg.ids)

#detrend stats over 50 year windows
#also needed for variance adjustment
jgb.stats <-rwi.stats.running(jgb.rwi, jgb.ids, running.window=TRUE, window.length=50)
khu.stats <-rwi.stats.running(khu.rwi, khu.ids, running.window=TRUE, window.length=50)
kll.stats <-rwi.stats.running(kll.rwi, kll.ids, running.window=TRUE, window.length=50)
klp.stats <-rwi.stats.running(klp.rwi, klp.ids, running.window=TRUE, window.length=50)
mhm.stats <-rwi.stats.running(mhm.rwi, mhm.ids, running.window=TRUE, window.length=50)
ndb.stats <-rwi.stats.running(ndb.rwi, ndb.ids, running.window=TRUE, window.length=50)
```

```
ogh.stats <-rwi.stats.running(ogh.rwi, ogh.ids, running.window=TRUE, window.length=50)
slb.stats <-rwi.stats.running(slb.rwi, slb.ids, running.window=TRUE, window.length=50)
zsm.stats <-rwi.stats.running(zsm.rwi, zsm.ids, running.window=TRUE, window.length=50)
ztg.stats <-rwi.stats.running(ztg.rwi, ztg.ids, running.window=TRUE, window.length=50)

#generate standard and residual
#chronologies using tukey robust mean
jgb.std <-chron(jgb.rwi, prefix="JB", biweight=TRUE, prewhiten=FALSE)
jgb.res <-chron(jgb.rwi, prefix="JB", biweight=TRUE, prewhiten=TRUE)

khu.std <-chron(khu.rwi, prefix="KU", biweight=TRUE, prewhiten=FALSE)
khu.res <-chron(khu.rwi, prefix="KU", biweight=TRUE, prewhiten=TRUE)

kll.std <-chron(kll.rwi, prefix="KL", biweight=TRUE, prewhiten=FALSE)
kll.res <-chron(kll.rwi, prefix="KL", biweight=TRUE, prewhiten=TRUE)

klp.std <-chron(klp.rwi, prefix="KP", biweight=TRUE, prewhiten=FALSE)
klp.res <-chron(klp.rwi, prefix="KP", biweight=TRUE, prewhiten=TRUE)

mhm.std <-chron(mhm.rwi, prefix="MU", biweight=TRUE, prewhiten=FALSE)
mhm.res <-chron(mhm.rwi, prefix="MU", biweight=TRUE, prewhiten=TRUE)

ndb.std <-chron(ndb.rwi, prefix="ND", biweight=TRUE, prewhiten=FALSE)
ndb.res <-chron(ndb.rwi, prefix="ND", biweight=TRUE, prewhiten=TRUE)

ogh.std <-chron(ogh.rwi, prefix="OG", biweight=TRUE, prewhiten=FALSE)
ogh.res <-chron(ogh.rwi, prefix="OG", biweight=TRUE, prewhiten=TRUE)

slb.std <-chron(slb.rwi, prefix="SB", biweight=TRUE, prewhiten=FALSE)
slb.res <-chron(slb.rwi, prefix="SB", biweight=TRUE, prewhiten=TRUE)

zsm.std <-chron(zsm.rwi, prefix="ZS", biweight=TRUE, prewhiten=FALSE)
zsm.res <-chron(zsm.rwi, prefix="ZS", biweight=TRUE, prewhiten=TRUE)

ztg.std <-chron(ztg.rwi, prefix="ZT", biweight=TRUE, prewhiten=FALSE)
ztg.res <-chron(ztg.rwi, prefix="ZT", biweight=TRUE, prewhiten=TRUE)

#------------------------------------------------------------------------------
#as in cook and peters 1997:
#comparison of ratios versus residuals fitted with a power transformation
#step 1 compare ratios and residuals in plot
#Use JGB04A with variance decrease in raw rings over time
#JB09C similar variance through time with some higher earlier in time
#JB11A less variance in center of time series than at ends

jbtest.raw <-jgb.rwl[,c(21,11,12)]
jbtest.raw1 <-jbtest.raw[,1][!is.na(jbtest.raw[,1])]
jbtest.raw2 <-jbtest.raw[,2][!is.na(jbtest.raw[,2])]
jbtest.raw3 <-jbtest.raw[,3][!is.na(jbtest.raw[,3])]


#apply spline to each to detrend
yearsx <-as.numeric(rownames(jbtest.raw))
yearsx1 <-yearsx[27:553]
yearsx2 <-yearsx[167:553]
```

```
yearsx3 <-yearsx[235:553]
jbtest.spl1<-ffcsaps(jbtest.raw1, x=yearsx1, nyrs=220, f=0.5)
jbtest.spl2<-ffcsaps(jbtest.raw2, x=yearsx2, nyrs=220, f=0.5)
jbtest.spl3<-ffcsaps(jbtest.raw3, x=yearsx3, nyrs=220, f=0.5)

#apply linear model to each to detrend
jbtest.lm1<-lm(jbtest.raw1~seq(1:length(yearsx1)))
jbtest.lm2<-lm(jbtest.raw2~seq(1:length(yearsx2)))
jbtest.lm3<-lm(jbtest.raw3~seq(1:length(yearsx3)))

#calculate ratios(indices=actual/expected)
jbtest1.ratiolm <-jbtest.raw1/jbtest.lm1$fitted.values
jbtest2.ratiolm <-jbtest.raw2/jbtest.lm2$fitted.values
jbtest3.ratiolm <-jbtest.raw3/jbtest.lm3$fitted.values

#calculate residuals (observed minus fitted)
jbtest1.resid <-jbtest.raw1-jbtest.spl1
jbtest2.resid <-jbtest.raw2-jbtest.spl2
jbtest3.resid <-jbtest.raw3-jbtest.spl3

jbtest1.residlm <-jbtest.lm1$residuals
jbtest2.residlm <-jbtest.lm2$residuals
jbtest3.residlm <-jbtest.lm3$residuals

#ratios
jbtest.ratio<-jgb.rwi[,c(21,11,12)]
jbtest1.ratio <-jbtest.ratio[,1][!is.na(jbtest.ratio[,1])]
jbtest2.ratio <-jbtest.ratio[,2][!is.na(jbtest.ratio[,2])]
jbtest3.ratio <-jbtest.ratio[,3][!is.na(jbtest.ratio[,3])]

#plot raw series and ratio indices for linear detrend
plot(jbtest.raw1, type="l", ylab="mm", main="JB04A", xlab="Age(Years)")
abline(jbtest.lm1, col="green")

plot(jbtest1.ratiolm, type="l", ylim=c(0,2), , ylab="RWI", main="JB04A", xlab="Age(Years)", col="green")
abline(h=1)

plot(jbtest.raw2, type="l", ylab="mm", main="JB09C", xlab="Age(Years)")
abline(jbtest.lm2, col="green")

plot(jbtest2.ratiolm, type="l", ylim=c(0,2), ylab="RWI", main="JB09C", xlab="Age(Years)", col="green")
abline(h=1)

plot(jbtest.raw3, type="l", ylab="mm", main="JB11A", xlab="Age(Years)")
abline(jbtest.lm3, col="green")

plot(jbtest3.ratiolm, type="l", ylim=c(0,2), ylab="RWI", main="JB11A", xlab="Age(Years)", col="green")
abline(h=1)

#plot ratios and residuals
#from spline
plot(jbtest1.ratio, type="l", ylim=c(-1,2), xaxt="n", ylab="Departures", main="JB04A", xlab="Year")
lines(jbtest1.resid, type="l", col="dark blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])
```

```
plot(jbtest2.ratio, type="l", ylim=c(-1,2), xaxt="n", ylab="Departures", main="JB09C", xlab="Year")
lines(jbtest2.resid, type="l", col="dark blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.ratio, type="l", ylim=c(-1,2), xaxt="n", ylab="Departures", main="JB11A", xlab="Year")
lines(jbtest3.resid, type="l", col="dark blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#from linear
plot(jbtest1.ratiolm, type="l", ylim=c(-1,2), xaxt="n", ylab="Departures", main="JB04A", xlab="Year")
lines(jbtest1.residlm, type="l", col="dark blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.ratiolm, type="l", ylim=c(-1,2), xaxt="n", ylab="Departures", main="JB09C", xlab="Year")
lines(jbtest2.residlm, type="l", col="dark blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.ratiolm, type="l", ylim=c(-1,2), xaxt="n", ylab="Departures", main="JB11A", xlab="Year")
lines(jbtest3.residlm, type="l", col="dark blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#difference (indices-residuals)
jbtest1.diff <-jbtest1.ratio-jbtest1.resid
jbtest2.diff <-jbtest2.ratio-jbtest2.resid
jbtest3.diff <-jbtest3.ratio-jbtest3.resid

jbtest1.difflm <-jbtest1.ratiolm-jbtest1.residlm
jbtest2.difflm <-jbtest2.ratiolm-jbtest2.residlm
jbtest3.difflm <-jbtest3.ratiolm-jbtest3.residlm

#dataframes of time and data
jbtest1.df <-data.frame(seq(1:length(yearsx1)), jbtest1.diff)
colnames(jbtest1.df) <-c("x", "y")
jbtest2.df <-data.frame(seq(1:length(yearsx2)), jbtest2.diff)
colnames(jbtest2.df) <-c("x", "y")
jbtest3.df <-data.frame(seq(1:length(yearsx3)), jbtest3.diff)
colnames(jbtest3.df) <-c("x", "y")

jbtest1.dflm <-data.frame(seq(1:length(yearsx1)), jbtest1.difflm)
colnames(jbtest1.dflm) <-c("x", "y")
jbtest2.dflm <-data.frame(seq(1:length(yearsx2)), jbtest2.difflm)
colnames(jbtest2.dflm) <-c("x", "y")
jbtest3.dflm <-data.frame(seq(1:length(yearsx3)), jbtest3.difflm)
colnames(jbtest3.dflm) <-c("x", "y")

#get a loess fit to each difference (using seq not years)
loess1 <-loess(y~x,jbtest1.df )
loess2 <-loess(y~x,jbtest2.df )
loess3 <-loess(y~x,jbtest3.df )

loess1lm <-loess(y~x,jbtest1.dflm )
loess2lm <-loess(y~x,jbtest2.dflm )
loess3lm <-loess(y~x,jbtest3.dflm )

#plot differences (with loess curve fit for reference)
```

```
#differences from spline
plot(jbtest1.diff, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB04A")
lines(jbtest1.df$x, predict(loess1), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.diff, type="l", ylim=c(0,2),xaxt="n", ylab="Difference", xlab="Year", main="JB09C" )
lines(jbtest2.df$x, predict(loess2), col = "blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.diff, type="l", ylim=c(0,2),xaxt="n", ylab="Difference", xlab="Year", main="JB11A"  )
lines(jbtest3.df$x, predict(loess3), col = "blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#differences from linear
plot(jbtest1.difflm, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB04A")
lines(jbtest1.dflm$x, predict(loess1lm), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.difflm, type="l", ylim=c(0,2),xaxt="n", ylab="Difference", xlab="Year", main="JB09C" )
lines(jbtest2.dflm$x, predict(loess2lm), col = "blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.difflm, type="l", ylim=c(0,2),xaxt="n", ylab="Difference", xlab="Year", main="JB11A"  )
lines(jbtest3.dflm$x, predict(loess3lm), col = "blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#now see about power transforming the raw ring widths and comparing the differences in the residuals
#calculating slope of power transform
#spread (y=log(std dev)) versus level (x=log(local mean of ring widths) relation
#where local std dev is s=absolute value of (actual width at time t minus actual width at previous time)
#and mean is (actual width at time t plus actual width at previous time)/2

#create lagged time series
jbtest.raw1lag <-c(NA, jbtest.raw1)
jbtest.raw1lag2 <-c(jbtest.raw1,NA)
jbtest1.lagdf <-data.frame(jbtest.raw1lag2,jbtest.raw1lag)
colnames(jbtest1.lagdf) <-c("t", "tminus1")
jbtest1.lagdft <-jbtest1.lagdf[2:527,]
jbtest1.lagsd <-abs(jbtest1.lagdft$t-jbtest1.lagdft$tminus1)
jbtest1.lagm <-(jbtest1.lagdft$t+jbtest1.lagdft$tminus1)/2
jbtest1.lagsdlog <-log(jbtest1.lagsd)
jbtest1.lagmlog <-log(jbtest1.lagm)

jbtest.raw2lag <-c(NA, jbtest.raw2)
jbtest.raw2lag2 <-c(jbtest.raw2,NA)
jbtest2.lagdf <-data.frame(jbtest.raw2lag2,jbtest.raw2lag)
colnames(jbtest2.lagdf) <-c("t", "tminus1")
jbtest2.lagdft <-jbtest2.lagdf[2:387,]
jbtest2.lagsd <-abs(jbtest2.lagdft$t-jbtest2.lagdft$tminus1)
jbtest2.lagm <-(jbtest2.lagdft$t+jbtest2.lagdft$tminus1)/2
jbtest2.lagsdlog <-log(jbtest2.lagsd)
jbtest2.lagmlog <-log(jbtest2.lagm)

jbtest.raw3lag <-c(NA, jbtest.raw3)
jbtest.raw3lag2 <-c(jbtest.raw3,NA)
```

```
jbtest3.lagdf <-data.frame(jbtest.raw3lag2,jbtest.raw3lag)
colnames(jbtest3.lagdf) <-c("t", "tminus1")
jbtest3.lagdft <-jbtest3.lagdf[2:319,]
jbtest3.lagsd <-abs(jbtest3.lagdft$t-jbtest3.lagdft$tminus1)
jbtest3.lagm <-(jbtest3.lagdft$t+jbtest3.lagdft$tminus1)/2
jbtest3.lagsdlog <-log(jbtest3.lagsd)
jbtest3.lagmlog <-log(jbtest3.lagm)

#plot scatterplot
plot(jbtest1.lagmlog,jbtest1.lagsdlog, pch=19, ylim=c(-5,0), xlim=c(-2.5,0.5),xlab="log(level)",
ylab="log(spread)", main="JB04A")
#add lm line
abline(jbtest1.lm)

plot(jbtest2.lagmlog,jbtest2.lagsdlog, pch=19, ylim=c(-5,0), xlim=c(-2.5,0.5),xlab="log(level)",
ylab="log(spread)", main="JB09C")
abline(jbtest2.lm)

plot(jbtest3.lagmlog,jbtest3.lagsdlog, pch=19, ylim=c(-5,0), xlim=c(-2.5,0.5),xlab="log(level)",
ylab="log(spread)", main="JB11A")
abline(jbtest3.lm)

#fit lm
#replace -Inf with NA
jbtest1.lagsdlog[jbtest1.lagsdlog==-Inf] <-NA
jbtest1.lagmlog[jbtest1.lagmlog==-Inf] <-NA

jbtest2.lagsdlog[jbtest2.lagsdlog==-Inf] <-NA
jbtest2.lagmlog[jbtest2.lagmlog==-Inf] <-NA

jbtest3.lagsdlog[jbtest3.lagsdlog==-Inf] <-NA
jbtest3.lagmlog[jbtest3.lagmlog==-Inf] <-NA

jbtest1.lm <-lm(jbtest1.lagsdlog~jbtest1.lagmlog, na.action="na.omit")
jbtest1.lm$coefficients
#(Intercept) jbtest1.lagmlog
#-1.8673047      0.6793456

jbtest2.lm <-lm(jbtest2.lagsdlog~jbtest2.lagmlog, na.action="na.omit")
jbtest2.lm$coefficients
#(Intercept) jbtest2.lagmlog
#-1.9764193      0.4096667

jbtest3.lm <-lm(jbtest3.lagsdlog~jbtest3.lagmlog, na.action="na.omit")
jbtest3.lm$coefficients
# (Intercept) jbtest3.lagmlog
#-2.4119779      0.2815063

#p value for power transform is 1-b
jbtest1.p <-1-jbtest1.lm$coefficients[2]
jbtest1.transraw <-jbtest.raw1^abs(jbtest1.p)

jbtest2.p <-1-jbtest2.lm$coefficients[2]
jbtest2.transraw <-jbtest.raw2^abs(jbtest2.p)
```

```
jbtest3.p <-1-jbtest3.lm$coefficients[2]
jbtest3.transraw <-jbtest.raw3^abs(jbtest3.p)

#plot transformed and raw values
plot(jbtest.raw1, type="l", , xaxt="n", ylab="mm", main="JB04A", xlab="Year")
lines(jbtest1.transraw, type="l", col="dark blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest.raw2, type="l", , xaxt="n", ylab="mm", main="JB09C", xlab="Year")
lines(jbtest2.transraw, type="l", col="dark blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest.raw3, type="l", , xaxt="n", ylab="mm", main="JB11A", xlab="Year")
lines(jbtest3.transraw, type="l", col="dark blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#detrend using spline on raw and power transformed raw
#spline on raw:
#jbtest.spl1, jbtest.spl2, jbtest.spl3

jbtest.spl1trans<-ffcsaps(jbtest1.transraw, x=yearsx1, nyrs=220, f=0.5)

jbtest.spl2trans<-ffcsaps(jbtest2.transraw, x=yearsx2, nyrs=220, f=0.5)

jbtest.spl3trans<-ffcsaps(jbtest3.transraw, x=yearsx3, nyrs=220, f=0.5)

#calcuate residuals
#residuals from untransformed raw with spline detrend
#jbtest1.resid,jbtest2.resid,jbtest3.resid

jbtest1.residtrans <-jbtest1.transraw-jbtest.spl1trans
jbtest2.residtrans <-jbtest2.transraw-jbtest.spl2trans
jbtest3.residtrans <-jbtest3.transraw-jbtest.spl3trans

#untransformed residuals from linear fit
#jbtest1.residlm
#jbtest2.residlm
#jbtest3.residlm

#transformed residuals with lm applied
jbtest.lm1trans<-lm(jbtest1.transraw~seq(1:length(yearsx1)))
jbtest.lm2trans<-lm(jbtest2.transraw~seq(1:length(yearsx2)))
jbtest.lm3trans<-lm(jbtest3.transraw~seq(1:length(yearsx3)))

jbtest1.residtranslm <-jbtest.lm1trans$residuals
jbtest2.residtranslm <-jbtest.lm2trans$residuals
jbtest3.residtranslm <-jbtest.lm3trans$residuals

#plot transformed and untransformed residuals
#spline detrend
plot(jbtest1.resid, type="l", ylim=c(-1,1), xaxt="n", ylab="Departures", main="JB04A", xlab="Year")
lines(jbtest1.residtrans, type="l", col="dark blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.resid, type="l", ylim=c(-1,1), xaxt="n", ylab="Departures", main="JB09C", xlab="Year")
```

```
lines(jbtest2.residtrans, type="l", col="dark blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.resid, type="l", ylim=c(-1,1), xaxt="n", ylab="Departures", main="JB11A", xlab="Year")
lines(jbtest3.residtrans, type="l", col="dark blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#lm detrend
plot(jbtest1.residlm, type="l", ylim=c(-1,1), xaxt="n", ylab="Departures", main="JB04A", xlab="Year")
lines(jbtest1.residtranslm, type="l", col="dark blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.residlm, type="l", ylim=c(-1,1), xaxt="n", ylab="Departures", main="JB09C", xlab="Year")
lines(jbtest2.residtranslm, type="l", col="dark blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.residlm, type="l", ylim=c(-1,1), xaxt="n", ylab="Departures", main="JB11A", xlab="Year")
lines(jbtest3.residtranslm, type="l", col="dark blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#calculate differences in residuals
#from spline detrend
jbtest1.difftrans <-jbtest1.resid-jbtest1.residtrans
jbtest1.dftrans <-data.frame(seq(1:length(yearsx1)), jbtest1.difftrans)
colnames(jbtest1.dftrans) <-c("x", "y")
loess1trans <-loess(y~x,jbtest1.dftrans )

jbtest2.difftrans <-jbtest2.resid-jbtest2.residtrans
jbtest2.dftrans <-data.frame(seq(1:length(yearsx2)), jbtest2.difftrans)
colnames(jbtest2.dftrans) <-c("x", "y")
loess2trans <-loess(y~x,jbtest2.dftrans )

jbtest3.difftrans <-jbtest3.resid-jbtest3.residtrans
jbtest3.dftrans <-data.frame(seq(1:length(yearsx3)), jbtest3.difftrans)
colnames(jbtest3.dftrans) <-c("x", "y")
loess3trans <-loess(y~x,jbtest3.dftrans )

#using linear
jbtest1.difftranslm <-jbtest1.residlm-jbtest1.residtranslm
jbtest1.dftranslm <-data.frame(seq(1:length(yearsx1)), jbtest1.difftranslm)
colnames(jbtest1.dftranslm) <-c("x", "y")
loess1translm <-loess(y~x,jbtest1.dftranslm )

jbtest2.difftranslm <-jbtest2.residlm-jbtest2.residtranslm
jbtest2.dftranslm <-data.frame(seq(1:length(yearsx2)), jbtest2.difftranslm)
colnames(jbtest2.dftranslm) <-c("x", "y")
loess2translm <-loess(y~x,jbtest2.dftranslm )

jbtest3.difftranslm <-jbtest3.residlm-jbtest3.residtranslm
jbtest3.dftranslm <-data.frame(seq(1:length(yearsx3)), jbtest3.difftranslm)
colnames(jbtest3.dftranslm) <-c("x", "y")
loess3translm <-loess(y~x,jbtest3.dftranslm )

#plot differences (with loess curve fit for reference)
#from spline
```

```
plot(jbtest1.difftrans, type="l", ylim=c(-1,1), xaxt="n", ylab="Difference", xlab="Year", main="JB04A")
lines(jbtest1.dftrans$x, predict(loess1trans), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.difftrans, type="l", ylim=c(-1,1), xaxt="n", ylab="Difference", xlab="Year", main="JB09C")
lines(jbtest2.dftrans$x, predict(loess2trans), col = "blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.difftrans, type="l", ylim=c(-1,1), xaxt="n", ylab="Difference", xlab="Year", main="JB11A")
lines(jbtest3.dftrans$x, predict(loess3trans), col = "blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#from linear
plot(jbtest1.difftranslm, type="l", ylim=c(-1,1), xaxt="n", ylab="Difference", xlab="Year", main="JB04A")
lines(jbtest1.dftranslm$x, predict(loess1translm), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.difftranslm, type="l", ylim=c(-1,1), xaxt="n", ylab="Difference", xlab="Year", main="JB09C")
lines(jbtest2.dftranslm$x, predict(loess2translm), col = "blue")
axis(1, at=seq(from=1, to=387, by=50), labels=yearsx2[c(1,  51, 101, 151, 201, 251, 301, 351)])

plot(jbtest3.difftranslm, type="l", ylim=c(-1,1), xaxt="n", ylab="Difference", xlab="Year", main="JB11A")
lines(jbtest3.dftranslm$x, predict(loess3translm), col = "blue")
axis(1, at=seq(from=1, to=319, by=50), labels=yearsx3[c( 1,  51, 101, 151, 201, 251, 301)])

#final comparison of power detrended residuals and ratios (non-power detrended)
#ratios from splines
#jbtest1.ratio,jbtest2.ratio,jbtest3.ratio

#ratios from linear
#jbtest1.ratiolm,jbtest2.ratiolm,jbtest3.ratiolm

#power transformed residuals from splines
#jbtest1.residtrans,jbtest2.residtrans,jbtest3.residtrans

#power tranformed residuals from linear
#jbtest1.residtranslm,jbtest1.residtranslm,jbtest1.residtranslm

#calculate difference of ratios and power transformed residuals and fit loess curve
#for splines
jbtest1.diffratiotrans <-jbtest1.ratio-jbtest1.residtrans
jbtest1.dfratiotrans <-data.frame(seq(1:length(yearsx1)), jbtest1.diffratiotrans)
colnames(jbtest1.dfratiotrans) <-c("x", "y")
loess1ratiotrans <-loess(y~x,jbtest1.dfratiotrans )

jbtest2.diffratiotrans <-jbtest2.ratio-jbtest2.residtrans
jbtest2.dfratiotrans <-data.frame(seq(1:length(yearsx2)), jbtest2.diffratiotrans)
colnames(jbtest2.dfratiotrans) <-c("x", "y")
loess2ratiotrans <-loess(y~x,jbtest2.dfratiotrans )

jbtest3.diffratiotrans <-jbtest3.ratio-jbtest3.residtrans
jbtest3.dfratiotrans <-data.frame(seq(1:length(yearsx3)), jbtest3.diffratiotrans)
colnames(jbtest3.dfratiotrans) <-c("x", "y")
loess3ratiotrans <-loess(y~x,jbtest3.dfratiotrans )
```

```
#for linear
jbtest1.diffratiotranslm <-jbtest1.ratiolm-jbtest1.residtranslm
jbtest1.dfratiotranslm <-data.frame(seq(1:length(yearsx1)), jbtest1.diffratiotranslm)
colnames(jbtest1.dfratiotranslm) <-c("x", "y")
loess1ratiotranslm <-loess(y~x,jbtest1.dfratiotranslm )

jbtest2.diffratiotranslm <-jbtest2.ratiolm-jbtest2.residtranslm
jbtest2.dfratiotranslm <-data.frame(seq(1:length(yearsx2)), jbtest2.diffratiotranslm)
colnames(jbtest2.dfratiotranslm) <-c("x", "y")
loess2ratiotranslm <-loess(y~x,jbtest2.dfratiotranslm )

jbtest3.diffratiotranslm <-jbtest3.ratiolm-jbtest3.residtranslm
jbtest3.dfratiotranslm <-data.frame(seq(1:length(yearsx3)), jbtest3.diffratiotranslm)
colnames(jbtest3.dfratiotranslm) <-c("x", "y")
loess3ratiotranslm <-loess(y~x,jbtest3.dfratiotranslm )

#plot differences for splines
plot(jbtest1.diffratiotrans, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB04A")
lines(jbtest1.dfratiotrans$x, predict(loess1ratiotrans), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.diffratiotrans, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB09C")
lines(jbtest2.dfratiotrans$x, predict(loess2ratiotrans), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest3.diffratiotrans, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB11A")
lines(jbtest3.dfratiotrans$x, predict(loess3ratiotrans), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

#plot differences for linear
plot(jbtest1.diffratiotranslm, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB04A")
lines(jbtest1.dfratiotranslm$x, predict(loess1ratiotranslm), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest2.diffratiotranslm, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB09C")
lines(jbtest2.dfratiotranslm$x, predict(loess2ratiotranslm), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])

plot(jbtest3.diffratiotranslm, type="l", ylim=c(0,2), xaxt="n", ylab="Difference", xlab="Year", main="JB11A")
lines(jbtest3.dfratiotranslm$x, predict(loess3ratiotranslm), col = "blue")
axis(1, at=seq(from=1, to=527, by=50), labels=yearsx1[c(1,51,101,151,201,251,301,351,401,451,501)])
#-----------------------------------------------------------------------
#Applying variance correction factor to chronologies
#using chronologies times the square root of the correction
#then scaling by 1/rbar to get estimate of mean in original units

#generate rwi rownames
jgb.chronyr <-as.numeric(rownames(jgb.rwi))
khu.chronyr <-as.numeric(rownames(khu.rwi))
kll.chronyr <-as.numeric(rownames(kll.rwi))
klp.chronyr <-as.numeric(rownames(klp.rwi))
mhm.chronyr <-as.numeric(rownames(mhm.rwi))
ndb.chronyr <-as.numeric(rownames(ndb.rwi))
ogh.chronyr <-as.numeric(rownames(ogh.rwi))
slb.chronyr <-as.numeric(rownames(slb.rwi))
```

```
zsm.chronyr <-as.numeric(rownames(zsm.rwi))
ztg.chronyr <-as.numeric(rownames(ztg.rwi))
#----------
#create running stats matrix that is filled to the same number of years as rwi file
jgb.statsfillmat <-matrix(ncol=4, nrow=length(jgb.chronyr), NA)
colnames(jgb.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(jgb.chronyr)) {
 if (jgb.chronyr[i]<jgb.stats$start.year[1]) {
   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[1], jgb.stats$end.year[1],jgb.stats$rbar.tot[1])
 } else if (jgb.chronyr[i]>=jgb.stats$start.year[1] & jgb.chronyr[i]<=jgb.stats$end.year[1]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[1], jgb.stats$end.year[1],jgb.stats$rbar.tot[1])
 } else if (jgb.chronyr[i]>jgb.stats$end.year[1] & jgb.chronyr[i]<=jgb.stats$end.year[3]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[3], jgb.stats$end.year[3],jgb.stats$rbar.tot[3])
 } else if (jgb.chronyr[i]>jgb.stats$end.year[3] & jgb.chronyr[i]<=jgb.stats$end.year[5]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[5], jgb.stats$end.year[5],jgb.stats$rbar.tot[5])
 } else if (jgb.chronyr[i]>jgb.stats$end.year[5] & jgb.chronyr[i]<=jgb.stats$end.year[7]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[7], jgb.stats$end.year[7],jgb.stats$rbar.tot[7])
 }else if (jgb.chronyr[i]>jgb.stats$end.year[7] & jgb.chronyr[i]<=jgb.stats$end.year[9]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[9], jgb.stats$end.year[9],jgb.stats$rbar.tot[9])
 }else if (jgb.chronyr[i]>jgb.stats$end.year[9] & jgb.chronyr[i]<=jgb.stats$end.year[11]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[11], jgb.stats$end.year[11],jgb.stats$rbar.tot[11])
 }else if (jgb.chronyr[i]>jgb.stats$end.year[11] & jgb.chronyr[i]<=jgb.stats$end.year[13]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[13], jgb.stats$end.year[13],jgb.stats$rbar.tot[13])
 }else if (jgb.chronyr[i]>jgb.stats$end.year[13] & jgb.chronyr[i]<=jgb.stats$end.year[15]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[15], jgb.stats$end.year[15],jgb.stats$rbar.tot[15])
 }else if (jgb.chronyr[i]>jgb.stats$end.year[15] & jgb.chronyr[i]<=jgb.stats$end.year[17]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[17], jgb.stats$end.year[17],jgb.stats$rbar.tot[17])
 }else if (jgb.chronyr[i]>jgb.stats$end.year[17] & jgb.chronyr[i]<=jgb.stats$end.year[19]) {

   jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[19], jgb.stats$end.year[19],jgb.stats$rbar.tot[19])
 }else {jgb.statsfill <-cbind(jgb.chronyr[i], jgb.stats$start.year[21],
jgb.stats$end.year[21],jgb.stats$rbar.tot[21])
  }
 jgb.statsfillmat[i,] <-jgb.statsfill
}

#--------------------
khu.statsfillmat <-matrix(ncol=4, nrow=length(khu.chronyr), NA)
colnames(khu.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(khu.chronyr)) {
 if (khu.chronyr[i]<khu.stats$start.year[1]) {
   khu.statsfill <-cbind(khu.chronyr[i], khu.stats$start.year[1], khu.stats$end.year[1],khu.stats$rbar.tot[1])
 } else if (khu.chronyr[i]>=khu.stats$start.year[1] & khu.chronyr[i]<=khu.stats$end.year[1]) {
```

```
  khu.statsfill <-cbind(khu.chronyr[i], khu.stats$start.year[1], khu.stats$end.year[1],khu.stats$rbar.tot[1])
 } else if (khu.chronyr[i]>khu.stats$end.year[1] & khu.chronyr[i]<=khu.stats$end.year[3]) {

  khu.statsfill <-cbind(khu.chronyr[i], khu.stats$start.year[3], khu.stats$end.year[3],khu.stats$rbar.tot[3])
 } else if (khu.chronyr[i]>khu.stats$end.year[3] & khu.chronyr[i]<=khu.stats$end.year[5]) {

  khu.statsfill <-cbind(khu.chronyr[i], khu.stats$start.year[5], khu.stats$end.year[5],khu.stats$rbar.tot[5])
 } else {  khu.statsfill <-cbind(khu.chronyr[i], khu.stats$start.year[6],
khu.stats$end.year[6],khu.stats$rbar.tot[6])
 }
 khu.statsfillmat[i,] <-khu.statsfill
}
#----------------------
kll.statsfillmat <-matrix(ncol=4, nrow=length(kll.chronyr), NA)
colnames(kll.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(kll.chronyr)) {
 if (kll.chronyr[i]<kll.stats$start.year[1]) {
  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[1], kll.stats$end.year[1],kll.stats$rbar.tot[1])
 } else if (kll.chronyr[i]>=kll.stats$start.year[1] & kll.chronyr[i]<=kll.stats$end.year[1]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[1], kll.stats$end.year[1],kll.stats$rbar.tot[1])
 } else if (kll.chronyr[i]>kll.stats$end.year[1] & kll.chronyr[i]<=kll.stats$end.year[3]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[3], kll.stats$end.year[3],kll.stats$rbar.tot[3])
 } else if (kll.chronyr[i]>kll.stats$end.year[3] & kll.chronyr[i]<=kll.stats$end.year[5]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[5], kll.stats$end.year[5],kll.stats$rbar.tot[5])
 } else if (kll.chronyr[i]>kll.stats$end.year[5] & kll.chronyr[i]<=kll.stats$end.year[7]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[7], kll.stats$end.year[7],kll.stats$rbar.tot[7])
 }else if (kll.chronyr[i]>kll.stats$end.year[7] & kll.chronyr[i]<=kll.stats$end.year[9]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[9], kll.stats$end.year[9],kll.stats$rbar.tot[9])
 }else if (kll.chronyr[i]>kll.stats$end.year[9] & kll.chronyr[i]<=kll.stats$end.year[11]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[11], kll.stats$end.year[11],kll.stats$rbar.tot[11])
 }else if (kll.chronyr[i]>kll.stats$end.year[11] & kll.chronyr[i]<=kll.stats$end.year[13]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[13], kll.stats$end.year[13],kll.stats$rbar.tot[13])
 }else if (kll.chronyr[i]>kll.stats$end.year[13] & kll.chronyr[i]<=kll.stats$end.year[15]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[15], kll.stats$end.year[15],kll.stats$rbar.tot[15])
 }else if (kll.chronyr[i]>kll.stats$end.year[15] & kll.chronyr[i]<=kll.stats$end.year[17]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[17], kll.stats$end.year[17],kll.stats$rbar.tot[17])
 }else if (kll.chronyr[i]>kll.stats$end.year[17] & kll.chronyr[i]<=kll.stats$end.year[19]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[19], kll.stats$end.year[19],kll.stats$rbar.tot[19])
 }else if (kll.chronyr[i]>kll.stats$end.year[19] & kll.chronyr[i]<=kll.stats$end.year[21]) {

  kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[21], kll.stats$end.year[21],kll.stats$rbar.tot[21])
 } else {kll.statsfill <-cbind(kll.chronyr[i], kll.stats$start.year[23], kll.stats$end.year[23],kll.stats$rbar.tot[23])
 }
 kll.statsfillmat[i,] <-kll.statsfill }
```

```
#-------
klp.statsfillmat <-matrix(ncol=4, nrow=length(klp.chronyr), NA)
colnames(klp.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(klp.chronyr)) {
 if (klp.chronyr[i]<klp.stats$start.year[1]) {
  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[1], klp.stats$end.year[1],klp.stats$rbar.tot[1])
 } else if (klp.chronyr[i]>=klp.stats$start.year[1] & klp.chronyr[i]<=klp.stats$end.year[1]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[1], klp.stats$end.year[1],klp.stats$rbar.tot[1])
 } else if (klp.chronyr[i]>klp.stats$end.year[1] & klp.chronyr[i]<=klp.stats$end.year[3]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[3], klp.stats$end.year[3],klp.stats$rbar.tot[3])
 } else if (klp.chronyr[i]>klp.stats$end.year[3] & klp.chronyr[i]<=klp.stats$end.year[5]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[5], klp.stats$end.year[5],klp.stats$rbar.tot[5])
 } else if (klp.chronyr[i]>klp.stats$end.year[5] & klp.chronyr[i]<=klp.stats$end.year[7]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[7], klp.stats$end.year[7],klp.stats$rbar.tot[7])
 }else if (klp.chronyr[i]>klp.stats$end.year[7] & klp.chronyr[i]<=klp.stats$end.year[9]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[9], klp.stats$end.year[9],klp.stats$rbar.tot[9])
 }else if (klp.chronyr[i]>klp.stats$end.year[9] & klp.chronyr[i]<=klp.stats$end.year[11]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[11], klp.stats$end.year[11],klp.stats$rbar.tot[11])
 }else if (klp.chronyr[i]>klp.stats$end.year[11] & klp.chronyr[i]<=klp.stats$end.year[13]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[13], klp.stats$end.year[13],klp.stats$rbar.tot[13])
 }else if (klp.chronyr[i]>klp.stats$end.year[13] & klp.chronyr[i]<=klp.stats$end.year[15]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[15], klp.stats$end.year[15],klp.stats$rbar.tot[15])
 }else if (klp.chronyr[i]>klp.stats$end.year[15] & klp.chronyr[i]<=klp.stats$end.year[17]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[17], klp.stats$end.year[17],klp.stats$rbar.tot[17])
 }else if (klp.chronyr[i]>klp.stats$end.year[17] & klp.chronyr[i]<=klp.stats$end.year[19]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[19], klp.stats$end.year[19],klp.stats$rbar.tot[19])
 }else if (klp.chronyr[i]>klp.stats$end.year[19] & klp.chronyr[i]<=klp.stats$end.year[21]) {

  klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[21], klp.stats$end.year[21],klp.stats$rbar.tot[21])
 } else {klp.statsfill <-cbind(klp.chronyr[i], klp.stats$start.year[22],
klp.stats$end.year[22],klp.stats$rbar.tot[22])
 }
 klp.statsfillmat[i,] <-klp.statsfill
}

#------------
mhm.statsfillmat <-matrix(ncol=4, nrow=length(mhm.chronyr), NA)
colnames(mhm.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(mhm.chronyr)) {
 if (mhm.chronyr[i]<mhm.stats$start.year[1]) {
  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[1],
mhm.stats$end.year[1],mhm.stats$rbar.tot[1])
 } else if (mhm.chronyr[i]>=mhm.stats$start.year[1] & mhm.chronyr[i]<=mhm.stats$end.year[1]) {
```

```
  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[1],
mhm.stats$end.year[1],mhm.stats$rbar.tot[1])
 } else if (mhm.chronyr[i]>mhm.stats$end.year[1] & mhm.chronyr[i]<=mhm.stats$end.year[3]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[3],
mhm.stats$end.year[3],mhm.stats$rbar.tot[3])
 } else if (mhm.chronyr[i]>mhm.stats$end.year[3] & mhm.chronyr[i]<=mhm.stats$end.year[5]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[5],
mhm.stats$end.year[5],mhm.stats$rbar.tot[5])
 } else if (mhm.chronyr[i]>mhm.stats$end.year[5] & mhm.chronyr[i]<=mhm.stats$end.year[7]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[7],
mhm.stats$end.year[7],mhm.stats$rbar.tot[7])
 }else if (mhm.chronyr[i]>mhm.stats$end.year[7] & mhm.chronyr[i]<=mhm.stats$end.year[9]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[9],
mhm.stats$end.year[9],mhm.stats$rbar.tot[9])
 }else if (mhm.chronyr[i]>mhm.stats$end.year[9] & mhm.chronyr[i]<=mhm.stats$end.year[11]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[11],
mhm.stats$end.year[11],mhm.stats$rbar.tot[11])
 }else if (mhm.chronyr[i]>mhm.stats$end.year[11] & mhm.chronyr[i]<=mhm.stats$end.year[13]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[13],
mhm.stats$end.year[13],mhm.stats$rbar.tot[13])
 }else if (mhm.chronyr[i]>mhm.stats$end.year[13] & mhm.chronyr[i]<=mhm.stats$end.year[15]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[15],
mhm.stats$end.year[15],mhm.stats$rbar.tot[15])
 }else if (mhm.chronyr[i]>mhm.stats$end.year[15] & mhm.chronyr[i]<=mhm.stats$end.year[17]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[17],
mhm.stats$end.year[17],mhm.stats$rbar.tot[17])
 }else if (mhm.chronyr[i]>mhm.stats$end.year[17] & mhm.chronyr[i]<=mhm.stats$end.year[19]) {

  mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[19],
mhm.stats$end.year[19],mhm.stats$rbar.tot[19])
 }else { mhm.statsfill <-cbind(mhm.chronyr[i], mhm.stats$start.year[21],
mhm.stats$end.year[21],mhm.stats$rbar.tot[21])
  }
 mhm.statsfillmat[i,] <-mhm.statsfill
}

#--------
ndb.statsfillmat <-matrix(ncol=4, nrow=length(ndb.chronyr), NA)
colnames(ndb.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(ndb.chronyr)) {
 if (ndb.chronyr[i]<ndb.stats$start.year[1]) {
  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[1], ndb.stats$end.year[1],ndb.stats$rbar.tot[1])
 } else if (ndb.chronyr[i]>=ndb.stats$start.year[1] & ndb.chronyr[i]<=ndb.stats$end.year[1]) {

  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[1], ndb.stats$end.year[1],ndb.stats$rbar.tot[1])
 } else if (ndb.chronyr[i]>ndb.stats$end.year[1] & ndb.chronyr[i]<=ndb.stats$end.year[3]) {
```

```
ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[3], ndb.stats$end.year[3],ndb.stats$rbar.tot[3])
 } else if (ndb.chronyr[i]>ndb.stats$end.year[3] & ndb.chronyr[i]<=ndb.stats$end.year[5]) {

  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[5], ndb.stats$end.year[5],ndb.stats$rbar.tot[5])
 } else if (ndb.chronyr[i]>ndb.stats$end.year[5] & ndb.chronyr[i]<=ndb.stats$end.year[7]) {

  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[7], ndb.stats$end.year[7],ndb.stats$rbar.tot[7])
 }else if (ndb.chronyr[i]>ndb.stats$end.year[7] & ndb.chronyr[i]<=ndb.stats$end.year[9]) {

  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[9], ndb.stats$end.year[9],ndb.stats$rbar.tot[9])
 }else if (ndb.chronyr[i]>ndb.stats$end.year[9] & ndb.chronyr[i]<=ndb.stats$end.year[11]) {

  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[11],
ndb.stats$end.year[11],ndb.stats$rbar.tot[11])
 }else if (ndb.chronyr[i]>ndb.stats$end.year[11] & ndb.chronyr[i]<=ndb.stats$end.year[13]) {

  ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[13],
ndb.stats$end.year[13],ndb.stats$rbar.tot[13])
 }else { ndb.statsfill <-cbind(ndb.chronyr[i], ndb.stats$start.year[15],
ndb.stats$end.year[15],ndb.stats$rbar.tot[15])
  }
 ndb.statsfillmat[i,] <-ndb.statsfill
}

#----------
ogh.statsfillmat <-matrix(ncol=4, nrow=length(ogh.chronyr), NA)
colnames(ogh.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(ogh.chronyr)) {
 if (ogh.chronyr[i]<ogh.stats$start.year[1]) {
  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[1], ogh.stats$end.year[1],ogh.stats$rbar.tot[1])
 } else if (ogh.chronyr[i]>=ogh.stats$start.year[1] & ogh.chronyr[i]<=ogh.stats$end.year[1]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[1], ogh.stats$end.year[1],ogh.stats$rbar.tot[1])
 } else if (ogh.chronyr[i]>ogh.stats$end.year[1] & ogh.chronyr[i]<=ogh.stats$end.year[3]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[3], ogh.stats$end.year[3],ogh.stats$rbar.tot[3])
 } else if (ogh.chronyr[i]>ogh.stats$end.year[3] & ogh.chronyr[i]<=ogh.stats$end.year[5]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[5], ogh.stats$end.year[5],ogh.stats$rbar.tot[5])
 } else if (ogh.chronyr[i]>ogh.stats$end.year[5] & ogh.chronyr[i]<=ogh.stats$end.year[7]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[7], ogh.stats$end.year[7],ogh.stats$rbar.tot[7])
 }else if (ogh.chronyr[i]>ogh.stats$end.year[7] & ogh.chronyr[i]<=ogh.stats$end.year[9]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[9], ogh.stats$end.year[9],ogh.stats$rbar.tot[9])
 }else if (ogh.chronyr[i]>ogh.stats$end.year[9] & ogh.chronyr[i]<=ogh.stats$end.year[11]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[11], ogh.stats$end.year[11],ogh.stats$rbar.tot[11])
 }else if (ogh.chronyr[i]>ogh.stats$end.year[11] & ogh.chronyr[i]<=ogh.stats$end.year[13]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[13], ogh.stats$end.year[13],ogh.stats$rbar.tot[13])
 }else if (ogh.chronyr[i]>ogh.stats$end.year[13] & ogh.chronyr[i]<=ogh.stats$end.year[15]) {

  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[15], ogh.stats$end.year[15],ogh.stats$rbar.tot[15])
```

```
 }else if (ogh.chronyr[i]>ogh.stats$end.year[15] & ogh.chronyr[i]<=ogh.stats$end.year[17]) {

   ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[17], ogh.stats$end.year[17],ogh.stats$rbar.tot[17])
 }else{  ogh.statsfill <-cbind(ogh.chronyr[i], ogh.stats$start.year[19],
ogh.stats$end.year[19],ogh.stats$rbar.tot[19])
 }
 ogh.statsfillmat[i,] <-ogh.statsfill
}

#-------------------------------------
slb.statsfillmat <-matrix(ncol=4, nrow=length(slb.chronyr), NA)
colnames(slb.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(slb.chronyr)) {
 if (slb.chronyr[i]<slb.stats$start.year[1]) {
   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[1], slb.stats$end.year[1],slb.stats$rbar.tot[1])
 } else if (slb.chronyr[i]>=slb.stats$start.year[1] & slb.chronyr[i]<=slb.stats$end.year[1]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[1], slb.stats$end.year[1],slb.stats$rbar.tot[1])
 } else if (slb.chronyr[i]>slb.stats$end.year[1] & slb.chronyr[i]<=slb.stats$end.year[3]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[3], slb.stats$end.year[3],slb.stats$rbar.tot[3])
 } else if (slb.chronyr[i]>slb.stats$end.year[3] & slb.chronyr[i]<=slb.stats$end.year[5]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[5], slb.stats$end.year[5],slb.stats$rbar.tot[5])
 } else if (slb.chronyr[i]>slb.stats$end.year[5] & slb.chronyr[i]<=slb.stats$end.year[7]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[7], slb.stats$end.year[7],slb.stats$rbar.tot[7])
 }else if (slb.chronyr[i]>slb.stats$end.year[7] & slb.chronyr[i]<=slb.stats$end.year[9]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[9], slb.stats$end.year[9],slb.stats$rbar.tot[9])
 }else if (slb.chronyr[i]>slb.stats$end.year[9] & slb.chronyr[i]<=slb.stats$end.year[11]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[11], slb.stats$end.year[11],slb.stats$rbar.tot[11])
 }else if (slb.chronyr[i]>slb.stats$end.year[11] & slb.chronyr[i]<=slb.stats$end.year[13]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[13], slb.stats$end.year[13],slb.stats$rbar.tot[13])
 }else if (slb.chronyr[i]>slb.stats$end.year[13] & slb.chronyr[i]<=slb.stats$end.year[15]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[15], slb.stats$end.year[15],slb.stats$rbar.tot[15])
 }else if (slb.chronyr[i]>slb.stats$end.year[15] & slb.chronyr[i]<=slb.stats$end.year[17]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[17], slb.stats$end.year[17],slb.stats$rbar.tot[17])
 }else if (slb.chronyr[i]>slb.stats$end.year[17] & slb.chronyr[i]<=slb.stats$end.year[19]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[19], slb.stats$end.year[19],slb.stats$rbar.tot[19])
 }else if (slb.chronyr[i]>slb.stats$end.year[19] & slb.chronyr[i]<=slb.stats$end.year[21]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[21], slb.stats$end.year[21],slb.stats$rbar.tot[21])
 } else if (slb.chronyr[i]>slb.stats$end.year[21] & slb.chronyr[i]<=slb.stats$end.year[23]) {

   slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[23], slb.stats$end.year[23],slb.stats$rbar.tot[23])
 }else { slb.statsfill <-cbind(slb.chronyr[i], slb.stats$start.year[24],
slb.stats$end.year[24],slb.stats$rbar.tot[24])
 }
```

```
  slb.statsfillmat[i,] <-slb.statsfill
}

#----------------------------------
zsm.statsfillmat <-matrix(ncol=4, nrow=length(zsm.chronyr), NA)
colnames(zsm.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(zsm.chronyr)) {
  if (zsm.chronyr[i]<zsm.stats$start.year[1]) {
    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[1], zsm.stats$end.year[1],zsm.stats$rbar.tot[1])
  } else if (zsm.chronyr[i]>=zsm.stats$start.year[1] & zsm.chronyr[i]<=zsm.stats$end.year[1]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[1], zsm.stats$end.year[1],zsm.stats$rbar.tot[1])
  } else if (zsm.chronyr[i]>zsm.stats$end.year[1] & zsm.chronyr[i]<=zsm.stats$end.year[3]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[3], zsm.stats$end.year[3],zsm.stats$rbar.tot[3])
  } else if (zsm.chronyr[i]>zsm.stats$end.year[3] & zsm.chronyr[i]<=zsm.stats$end.year[5]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[5], zsm.stats$end.year[5],zsm.stats$rbar.tot[5])
  } else if (zsm.chronyr[i]>zsm.stats$end.year[5] & zsm.chronyr[i]<=zsm.stats$end.year[7]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[7], zsm.stats$end.year[7],zsm.stats$rbar.tot[7])
  }else if (zsm.chronyr[i]>zsm.stats$end.year[7] & zsm.chronyr[i]<=zsm.stats$end.year[9]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[9], zsm.stats$end.year[9],zsm.stats$rbar.tot[9])
  }else if (zsm.chronyr[i]>zsm.stats$end.year[9] & zsm.chronyr[i]<=zsm.stats$end.year[11]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[11],
zsm.stats$end.year[11],zsm.stats$rbar.tot[11])
  }else if (zsm.chronyr[i]>zsm.stats$end.year[11] & zsm.chronyr[i]<=zsm.stats$end.year[13]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[13],
zsm.stats$end.year[13],zsm.stats$rbar.tot[13])
  }else if (zsm.chronyr[i]>zsm.stats$end.year[13] & zsm.chronyr[i]<=zsm.stats$end.year[15]) {

    zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[15],
zsm.stats$end.year[15],zsm.stats$rbar.tot[15])
  }else {  zsm.statsfill <-cbind(zsm.chronyr[i], zsm.stats$start.year[17],
zsm.stats$end.year[17],zsm.stats$rbar.tot[17])
  }
  zsm.statsfillmat[i,] <-zsm.statsfill
}

#--------
ztg.statsfillmat <-matrix(ncol=4, nrow=length(ztg.chronyr), NA)
colnames(ztg.statsfillmat) <-c("chronyr", "win.startyr", "win.endyr", "rbar.tot")

for (i in 1:length(ztg.chronyr)) {
  if (ztg.chronyr[i]<ztg.stats$start.year[1]) {
    ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[1], ztg.stats$end.year[1],ztg.stats$rbar.tot[1])
  } else if (ztg.chronyr[i]>=ztg.stats$start.year[1] & ztg.chronyr[i]<=ztg.stats$end.year[1]) {

    ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[1], ztg.stats$end.year[1],ztg.stats$rbar.tot[1])
  } else if (ztg.chronyr[i]>ztg.stats$end.year[1] & ztg.chronyr[i]<=ztg.stats$end.year[3]) {
```

```
  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[3], ztg.stats$end.year[3],ztg.stats$rbar.tot[3])
 } else if (ztg.chronyr[i]>ztg.stats$end.year[3] & ztg.chronyr[i]<=ztg.stats$end.year[5]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[5], ztg.stats$end.year[5],ztg.stats$rbar.tot[5])
 } else if (ztg.chronyr[i]>ztg.stats$end.year[5] & ztg.chronyr[i]<=ztg.stats$end.year[7]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[7], ztg.stats$end.year[7],ztg.stats$rbar.tot[7])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[7] & ztg.chronyr[i]<=ztg.stats$end.year[9]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[9], ztg.stats$end.year[9],ztg.stats$rbar.tot[9])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[9] & ztg.chronyr[i]<=ztg.stats$end.year[11]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[11], ztg.stats$end.year[11],ztg.stats$rbar.tot[11])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[11] & ztg.chronyr[i]<=ztg.stats$end.year[13]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[13], ztg.stats$end.year[13],ztg.stats$rbar.tot[13])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[13] & ztg.chronyr[i]<=ztg.stats$end.year[15]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[15], ztg.stats$end.year[15],ztg.stats$rbar.tot[15])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[15] & ztg.chronyr[i]<=ztg.stats$end.year[17]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[17], ztg.stats$end.year[17],ztg.stats$rbar.tot[17])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[17] & ztg.chronyr[i]<=ztg.stats$end.year[19]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[19], ztg.stats$end.year[19],ztg.stats$rbar.tot[19])
 }else if (ztg.chronyr[i]>ztg.stats$end.year[19] & ztg.chronyr[i]<=ztg.stats$end.year[21]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[21], ztg.stats$end.year[21],ztg.stats$rbar.tot[21])
 } else if (ztg.chronyr[i]>ztg.stats$end.year[21] & ztg.chronyr[i]<=ztg.stats$end.year[23]) {

  ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[23], ztg.stats$end.year[23],ztg.stats$rbar.tot[23])
 }else {   ztg.statsfill <-cbind(ztg.chronyr[i], ztg.stats$start.year[24],
ztg.stats$end.year[24],ztg.stats$rbar.tot[24])
 }
 ztg.statsfillmat[i,] <-ztg.statsfill
}

#----------------------------
#use filled matrices for calculating effective sample size at each timestep
#add rbar from matrices to rwi files

neff <-function (x,y) {
 sum(!is.na(x))/(1+(sum(!is.na(x))-1)*y)
}

#-----
#JGB
jgb.neff <-matrix(ncol=2, nrow=length(jgb.statsfillmat[,1]), NA)
colnames(jgb.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(jgb.statsfillmat[,1])){
 neffec <-neff(jgb.rwi[i,], jgb.statsfillmat[i,4])
 rbar <-jgb.statsfillmat[i,4]
 jgb.neff[i,] <-cbind(neffec,rbar)
}
```

```
jgb.correctstd <-jgb.std[,1]*sqrt(jgb.neff[,1])
jgb.correctrstd <-jgb.correctstd*1/(sqrt(1/jgb.neff[,2]))

jgb.stdc<- jgb.std
jgb.stdc$JBstd<-jgb.correctrstd

#write chronologies
write.crn(jgb.std, "JGB_Std.crn")
write.crn(jgb.stdc, "JGB_Std_VarC.crn")

plot(jgb.std)
plot(jgb.stdc)

#compare plots
plot(seq(1:length(jgb.chronyr)), jgb.std[,1], type="l")
lines(jgb.stdc[,1], col="blue")

#for prewhitened chronologies
plot(jgb.res[,2:3])

jgb.correctres <-jgb.res[,2]*sqrt(jgb.neff[,1])
jgb.correctrres <-jgb.correctres*1/(sqrt(1/jgb.neff[,2]))

jgb.resc<- jgb.res
jgb.resc$JBres<-jgb.correctrres
jgb.rest <-jgb.res[,2:3]
jgb.resct <-jgb.resc[,2:3]

#write chronologies
write.crn(jgb.rest, "JGB_Res.crn")
write.crn(jgb.resct, "JGB_Res_VarC.crn")

plot(jgb.res[,2:3])
plot(jgb.resc[,2:3])

#compare plots
plot(seq(1:length(jgb.chronyr)), jgb.res[,2], type="l")
lines(jgb.resc[,2], col="blue")

#-----
#KHU
khu.neff <-matrix(ncol=2, nrow=length(khu.statsfillmat[,1]), NA)
colnames(khu.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(khu.statsfillmat[,1])){
  neffec <-neff(khu.rwi[i,], khu.statsfillmat[i,4])
  rbar <-khu.statsfillmat[i,4]
  khu.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
khu.correctstd <-khu.std[,1]*sqrt(khu.neff[,1])
khu.correctrstd <-khu.correctstd*1/(sqrt(1/khu.neff[,2]))

khu.stdc<- khu.std
khu.stdc$KUstd<-khu.correctrstd
```

```
#write chronologies
write.crn(khu.std, "KHU_Std.crn")
write.crn(khu.stdc, "KHU_Std_VarC.crn")

#res chrons
khu.correctres <-khu.res[,2]*sqrt(khu.neff[,1])
khu.correctrres <-khu.correctres*1/(sqrt(1/khu.neff[,2]))

khu.resc<- khu.res
khu.resc$KUres<-khu.correctrres
khu.rest <-khu.res[,2:3]
khu.resct <-khu.resc[,2:3]

#write chronologies
write.crn(khu.rest, "KHU_Res.crn")
write.crn(khu.resct, "KHU_Res_VarC.crn")

#--------------------------
#KLL
kll.neff <-matrix(ncol=2, nrow=length(kll.statsfillmat[,1]), NA)
colnames(kll.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(kll.statsfillmat[,1])){
 neffec <-neff(kll.rwi[i,], kll.statsfillmat[i,4])
 rbar <-kll.statsfillmat[i,4]
 kll.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
kll.correctstd <-kll.std[,1]*sqrt(kll.neff[,1])
kll.correctrstd <-kll.correctstd*1/(sqrt(1/kll.neff[,2]))

kll.stdc<- kll.std
kll.stdc$KLstd<-kll.correctrstd

#write chronologies
write.crn(kll.std, "KLL_Std.crn")
write.crn(kll.stdc, "KLL_Std_VarC.crn")

#res chrons
kll.correctres <-kll.res[,2]*sqrt(kll.neff[,1])
kll.correctrres <-kll.correctres*1/(sqrt(1/kll.neff[,2]))

kll.resc<- kll.res
kll.resc$KLres<-kll.correctrres
kll.rest <-kll.res[,2:3]
kll.resct <-kll.resc[,2:3]

#write chronologies
write.crn(kll.rest, "KLL_Res.crn")
write.crn(kll.resct, "KLL_Res_VarC.crn")

#--------
#KLP
klp.neff <-matrix(ncol=2, nrow=length(klp.statsfillmat[,1]), NA)
colnames(klp.neff) <-c("Neff", "rbar.tot")
```

```
for (i in 1:length(klp.statsfillmat[,1])){
 neffec <-neff(klp.rwi[i,], klp.statsfillmat[i,4])
 rbar <-klp.statsfillmat[i,4]
 klp.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
klp.correctstd <-klp.std[,1]*sqrt(klp.neff[,1])
klp.correctrstd <-klp.correctstd*1/(sqrt(1/klp.neff[,2]))

klp.stdc<- klp.std
klp.stdc$KPstd<-klp.correctrstd

#write chronologies
write.crn(klp.std, "KLP_Std.crn")
write.crn(klp.stdc, "KLP_Std_VarC.crn")

#res chrons
klp.correctres <-klp.res[,2]*sqrt(klp.neff[,1])
klp.correctrres <-klp.correctres*1/(sqrt(1/klp.neff[,2]))

klp.resc<- klp.res
klp.resc$KPres<-klp.correctrres
klp.rest <-klp.res[,2:3]
klp.resct <-klp.resc[,2:3]

#write chronologies
write.crn(klp.rest, "KLP_Res.crn")
write.crn(klp.resct, "KLP_Res_VarC.crn")

#-----
#MHM
mhm.neff <-matrix(ncol=2, nrow=length(mhm.statsfillmat[,1]), NA)
colnames(mhm.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(mhm.statsfillmat[,1])){
 neffec <-neff(mhm.rwi[i,], mhm.statsfillmat[i,4])
 rbar <-mhm.statsfillmat[i,4]
 mhm.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
mhm.correctstd <-mhm.std[,1]*sqrt(mhm.neff[,1])
mhm.correctrstd <-mhm.correctstd*1/(sqrt(1/mhm.neff[,2]))

mhm.stdc<- mhm.std
mhm.stdc$MUstd<-mhm.correctrstd

#write chronologies
write.crn(mhm.std, "MHM_Std.crn")
write.crn(mhm.stdc, "MHM_Std_VarC.crn")

#res chrons
mhm.correctres <-mhm.res[,2]*sqrt(mhm.neff[,1])
mhm.correctrres <-mhm.correctres*1/(sqrt(1/mhm.neff[,2]))
```

```
mhm.resc<- mhm.res
mhm.resc$MUres<-mhm.correctrres
mhm.rest <-mhm.res[,2:3]
mhm.resct <-mhm.resc[,2:3]

#write chronologies
write.crn(mhm.rest, "MHM_Res.crn")
write.crn(mhm.resct, "MHM_Res_VarC.crn")

#---------
#NDB
ndb.neff <-matrix(ncol=2, nrow=length(ndb.statsfillmat[,1]), NA)
colnames(ndb.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(ndb.statsfillmat[,1])){
 neffec <-neff(ndb.rwi[i,], ndb.statsfillmat[i,4])
 rbar <-ndb.statsfillmat[i,4]
 ndb.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
ndb.correctstd <-ndb.std[,1]*sqrt(ndb.neff[,1])
ndb.correctrstd <-ndb.correctstd*1/(sqrt(1/ndb.neff[,2]))

ndb.stdc<- ndb.std
ndb.stdc$NDstd<-ndb.correctrstd

#write chronologies
write.crn(ndb.std, "NDB_Std.crn")
write.crn(ndb.stdc, "NDB_Std_VarC.crn")

#res chrons
ndb.correctres <-ndb.res[,2]*sqrt(ndb.neff[,1])
ndb.correctrres <-ndb.correctres*1/(sqrt(1/ndb.neff[,2]))

ndb.resc<- ndb.res
ndb.resc$NDres<-ndb.correctrres
ndb.rest <-ndb.res[,2:3]
ndb.resct <-ndb.resc[,2:3]

#write chronologies
write.crn(ndb.rest, "NDB_Res.crn")
write.crn(ndb.resct, "NDB_Res_VarC.crn")

#------
#OGH
ogh.neff <-matrix(ncol=2, nrow=length(ogh.statsfillmat[,1]), NA)
colnames(ogh.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(ogh.statsfillmat[,1])){
 neffec <-neff(ogh.rwi[i,], ogh.statsfillmat[i,4])
 rbar <-ogh.statsfillmat[i,4]
 ogh.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
ogh.correctstd <-ogh.std[,1]*sqrt(ogh.neff[,1])
```

```
ogh.correctrstd <-ogh.correctstd*1/(sqrt(1/ogh.neff[,2]))

ogh.stdc<- ogh.std
ogh.stdc$OGstd<-ogh.correctrstd

#write chronologies
write.crn(ogh.std, "OGH_Std.crn")
write.crn(ogh.stdc, "OGH_Std_VarC.crn")

#res chrons
ogh.correctres <-ogh.res[,2]*sqrt(ogh.neff[,1])
ogh.correctrres <-ogh.correctres*1/(sqrt(1/ogh.neff[,2]))

ogh.resc<- ogh.res
ogh.resc$OGres<-ogh.correctrres
ogh.rest <-ogh.res[,2:3]
ogh.resct <-ogh.resc[,2:3]

#write chronologies
write.crn(ogh.rest, "OGH_Res.crn")
write.crn(ogh.resct, "OGH_Res_VarC.crn")
#------
#SLB
slb.neff <-matrix(ncol=2, nrow=length(slb.statsfillmat[,1]), NA)
colnames(slb.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(slb.statsfillmat[,1])){
  neffec <-neff(slb.rwi[i,], slb.statsfillmat[i,4])
  rbar <-slb.statsfillmat[i,4]
  slb.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
slb.correctstd <-slb.std[,1]*sqrt(slb.neff[,1])
slb.correctrstd <-slb.correctstd*1/(sqrt(1/slb.neff[,2]))

slb.stdc<- slb.std
slb.stdc$SBstd<-slb.correctrstd

#write chronologies
write.crn(slb.std, "SLB_Std.crn")
write.crn(slb.stdc, "SLB_Std_VarC.crn")

#res chrons
slb.correctres <-slb.res[,2]*sqrt(slb.neff[,1])
slb.correctrres <-slb.correctres*1/(sqrt(1/slb.neff[,2]))

slb.resc<- slb.res
slb.resc$SBres<-slb.correctrres
slb.rest <-slb.res[,2:3]
slb.resct <-slb.resc[,2:3]

#write chronologies
write.crn(slb.rest, "SLB_Res.crn")
write.crn(slb.rest, "SLB_Res_VarC.crn")
```

```
#----
#ZSM
zsm.neff <-matrix(ncol=2, nrow=length(zsm.statsfillmat[,1]), NA)
colnames(zsm.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(zsm.statsfillmat[,1])){
 neffec <-neff(zsm.rwi[i,], zsm.statsfillmat[i,4])
 rbar <-zsm.statsfillmat[i,4]
 zsm.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
zsm.correctstd <-zsm.std[,1]*sqrt(zsm.neff[,1])
zsm.correctrstd <-zsm.correctstd*1/(sqrt(1/zsm.neff[,2]))

zsm.stdc<- zsm.std
zsm.stdc$ZSstd<-zsm.correctrstd

#write chronologies
write.crn(zsm.std, "ZSM_Std.crn")
write.crn(zsm.stdc, "ZSM_Std_VarC.crn")

#res chrons
zsm.correctres <-zsm.res[,2]*sqrt(zsm.neff[,1])
zsm.correctrres <-zsm.correctres*1/(sqrt(1/zsm.neff[,2]))

zsm.resc<- zsm.res
zsm.resc$ZSres<-zsm.correctrres
zsm.rest <-zsm.res[,2:3]
zsm.resct <-zsm.resc[,2:3]

#write chronologies
write.crn(zsm.rest, "ZSM_Res.crn")
write.crn(zsm.resct, "ZSM_Res_VarC.crn")
#-----
#ZTG
ztg.neff <-matrix(ncol=2, nrow=length(ztg.statsfillmat[,1]), NA)
colnames(ztg.neff) <-c("Neff", "rbar.tot")
for (i in 1:length(ztg.statsfillmat[,1])){
 neffec <-neff(ztg.rwi[i,], ztg.statsfillmat[i,4])
 rbar <-ztg.statsfillmat[i,4]
 ztg.neff[i,] <-cbind(neffec,rbar)
}

#std chrons
ztg.correctstd <-ztg.std[,1]*sqrt(ztg.neff[,1])
ztg.correctrstd <-ztg.correctstd*1/(sqrt(1/ztg.neff[,2]))

ztg.stdc<- ztg.std
ztg.stdc$ZTstd<-ztg.correctrstd

#write chronologies
write.crn(ztg.std, "ZTG_Std.crn")
write.crn(ztg.stdc, "ZTG_Std_VarC.crn")

#res chrons
```

```
ztg.correctres <-ztg.res[,2]*sqrt(ztg.neff[,1])
ztg.correctrres <-ztg.correctres*1/(sqrt(1/ztg.neff[,2]))

ztg.resc<- ztg.res
ztg.resc$ZTres<-ztg.correctrres
ztg.rest <-ztg.res[,2:3]
ztg.resct <-ztg.resc[,2:3]

#write chronologies
write.crn(ztg.rest, "ZTG_Res.crn")
write.crn(ztg.resct, "ZTG_Res_VarC.crn")

#--------
#truncate corrected chronologies based on between-tree EPS calculated in 50 year moving windows.
#50-yr windows moving back from start of chronology to establish general period

#----
#JGB
jgb.statsmat <-list()
for (i in 1:25){
stats<-rwi.stats.running(jgb.rwi, jgb.ids, running.window=TRUE, window.length=50, first.start=i)
jgb.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
jgb.statsmat1 <-data.frame(jgb.statsmat[1:4])
jgb.statsmat2 <-data.frame(jgb.statsmat[5:25])


#truncate at year
#JGB 1650
which(rownames(jgb.stdc)=="1650")
jgb.stdceps <-jgb.stdc[192:553,]
jgb.resceps <-jgb.resct[192:553,]

write.crn(jgb.stdceps, "JGB_Stdf.crn")
write.crn(jgb.resceps, "JGB_Resf.crn")

#get median length of series contributing to final chron
jgb.truncraw <-jgb.rwi[192:553,]

jgb.colcount <-list()
jgb.truncrawsum <-
  for (i in 1:ncol(jgb.rwi)) {
  colcounts <-sum(!is.na(jgb.truncraw[,i]))
jgb.colcount[[i]] <-colcounts
  }
jgb.colco <-data.frame(jgb.colcount)
jgb.med <-median(as.numeric(jgb.colco))

#------
#KHU
khu.statsmat <-list()
for (i in 1:25){
  stats<-rwi.stats.running(khu.rwi, khu.ids, running.window=TRUE, window.length=50, first.start=i)
```

```
  khu.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
khu.statsmat1 <-data.frame(khu.statsmat[1:13])
khu.statsmat2 <-data.frame(khu.statsmat[14:25])


#truncate at year
#khu 1829
which(rownames(khu.stdc)=="1829")
khu.stdceps <-khu.stdc[5:187,]
khu.resceps <-khu.resct[5:187,]

write.crn(khu.stdceps, "KHU_Stdf.crn")
write.crn(khu.resceps, "KHU_Resf.crn")

#get median length of series contributing to final chron
khu.truncraw <-khu.rwi[5:187,]

khu.colcount <-list()
khu.truncrawsum <-
  for (i in 1:ncol(khu.rwi)) {
    colcounts <-sum(!is.na(khu.truncraw[,i]))
    khu.colcount[[i]] <-colcounts
  }
khu.colco <-data.frame(khu.colcount)
khu.med <-median(as.numeric(khu.colco))

#----
#KLL
kll.statsmat <-list()
for (i in 1:25){
  stats<-rwi.stats.running(kll.rwi, kll.ids, running.window=TRUE, window.length=50, first.start=i)
  kll.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
kll.statsmat1 <-data.frame(kll.statsmat[1:12])
kll.statsmat2 <-data.frame(kll.statsmat[13:25])


#truncate at year
#kll
which(rownames(kll.stdc)=="1405")
kll.stdceps <-kll.stdc[66:661,]
kll.resceps <-kll.resct[66:661,]

write.crn(kll.stdceps, "KLL_Stdf.crn")
write.crn(kll.resceps, "KLL_Resf.crn")

#get median length of series contributing to final chron
kll.truncraw <-kll.rwi[66:661,]

kll.colcount <-list()
```

```
kll.truncrawsum <-
  for (i in 1:ncol(kll.rwi)) {
    colcounts <-sum(!is.na(kll.truncraw[,i]))
    kll.colcount[[i]] <-colcounts
  }
kll.colco <-data.frame(kll.colcount)
kll.med <-median(as.numeric(kll.colco))

#----
#KLP
klp.statsmat <-list()
for (i in 1:25){
  stats<-rwi.stats.running(klp.rwi, klp.ids, running.window=TRUE, window.length=50, first.start=i)
  klp.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
klp.statsmat1 <-data.frame(klp.statsmat[1:16])
klp.statsmat2 <-data.frame(klp.statsmat[17:25])


#truncate at year
#klp
which(rownames(klp.stdc)=="1448")
klp.stdceps <-klp.stdc[27:590,]
klp.resceps <-klp.resct[27:590,]

write.crn(klp.stdceps, "KLP_Stdf.crn")
write.crn(klp.resceps, "KLP_Resf.crn")

#get median length of series contributing to final chron
klp.truncraw <-klp.rwi[27:590,]

klp.colcount <-list()
klp.truncrawsum <-
  for (i in 1:ncol(klp.rwi)) {
    colcounts <-sum(!is.na(klp.truncraw[,i]))
    klp.colcount[[i]] <-colcounts
  }
klp.colco <-data.frame(klp.colcount)
klp.med <-median(as.numeric(klp.colco))

#-----
#MHM
mhm.statsmat <-list()
for (i in 1:25){
  stats<-rwi.stats.running(mhm.rwi, mhm.ids, running.window=TRUE, window.length=50, first.start=i)
  mhm.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
mhm.statsmat1 <-data.frame(mhm.statsmat[1:21])
mhm.statsmat2 <-data.frame(mhm.statsmat[22:25])
```

```
#truncate at year
#mhm
which(rownames(mhm.stdc)=="1576")
mhm.stdceps <-mhm.stdc[144:570,]
mhm.resceps <-mhm.resct[144:570,]

write.crn(mhm.stdceps, "MHM_Stdf.crn")
write.crn(mhm.resceps, "MHM_Resf.crn")

#get median length of series contributing to final chron
mhm.truncraw <-mhm.rwi[,]

mhm.colcount <-list()
mhm.truncrawsum <-
 for (i in 1:ncol(mhm.rwi)) {
   colcounts <-sum(!is.na(mhm.truncraw[,i]))
   mhm.colcount[[i]] <-colcounts
 }
mhm.colco <-data.frame(mhm.colcount)
mhm.med <-median(as.numeric(mhm.colco))

#----
#NDB
ndb.statsmat <-list()
for (i in 1:25){
 stats<-rwi.stats.running(ndb.rwi, ndb.ids, running.window=TRUE, window.length=50, first.start=i)
 ndb.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
ndb.statsmat1 <-data.frame(ndb.statsmat[1:4])
ndb.statsmat2 <-data.frame(ndb.statsmat[5:25])


#truncate at year
#ndb
which(rownames(ndb.stdc)=="1630")
ndb.stdceps <-ndb.stdc[32:403,]
ndb.resceps <-ndb.resct[32:403,]

write.crn(ndb.stdceps, "NDB_Stdf.crn")
write.crn(ndb.resceps, "NDB_Resf.crn")

#get median length of series contributing to final chron
ndb.truncraw <-ndb.rwi[32:403,]

ndb.colcount <-list()
ndb.truncrawsum <-
 for (i in 1:ncol(ndb.rwi)) {
   colcounts <-sum(!is.na(ndb.truncraw[,i]))
   ndb.colcount[[i]] <-colcounts
 }
ndb.colco <-data.frame(ndb.colcount)
ndb.med <-median(as.numeric(ndb.colco))
```

```
#----
#OGH
ogh.statsmat <-list()
for (i in 1:25){
  stats<-rwi.stats.running(ogh.rwi, ogh.ids, running.window=TRUE, window.length=50, first.start=i)
  ogh.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
ogh.statsmat1 <-data.frame(ogh.statsmat[1:8])
ogh.statsmat2 <-data.frame(ogh.statsmat[9:25])


#truncate at year
#ogh
which(rownames(ogh.stdc)=="1582")
ogh.stdceps <-ogh.stdc[80:507,]
ogh.resceps <-ogh.resct[80:507,]

write.crn(ogh.stdceps, "OGH_Stdf.crn")
write.crn(ogh.resceps, "OGH_Resf.crn")

#get median length of series contributing to final chron
ogh.truncraw <-ogh.rwi[80:507,]

ogh.colcount <-list()
ogh.truncrawsum <-
  for (i in 1:ncol(ogh.rwi)) {
    colcounts <-sum(!is.na(ogh.truncraw[,i]))
    ogh.colcount[[i]] <-colcounts
  }
ogh.colco <-data.frame(ogh.colcount)
ogh.med <-median(as.numeric(ogh.colco))

#----
#SLB
slb.statsmat <-list()
for (i in 1:25){
  stats<-rwi.stats.running(slb.rwi, slb.ids, running.window=TRUE, window.length=50, first.start=i)
  slb.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
slb.statsmat1 <-data.frame(slb.statsmat[1:13])
slb.statsmat2 <-data.frame(slb.statsmat[14:25])


#truncate at year
#slb
which(rownames(slb.stdc)=="1455")
slb.stdceps <-slb.stdc[93:637,]
slb.resceps <-slb.resct[93:637,]

write.crn(slb.stdceps, "SLB_Stdf.crn")
write.crn(slb.resceps, "SLB_Resf.crn")
```

```
#get median length of series contributing to final chron
slb.truncraw <-slb.rwi[93:637,]

slb.colcount <-list()
slb.truncrawsum <-
 for (i in 1:ncol(slb.rwi)) {
   colcounts <-sum(!is.na(slb.truncraw[,i]))
   slb.colcount[[i]] <-colcounts
 }
slb.colco <-data.frame(slb.colcount)
slb.med <-median(as.numeric(slb.colco))

#----
#ZSM
zsm.statsmat <-list()
for (i in 1:25){
 stats<-rwi.stats.running(zsm.rwi, zsm.ids, running.window=TRUE, window.length=50, first.start=i)
 zsm.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
zsm.statsmat1 <-data.frame(zsm.statsmat[1:15])
zsm.statsmat2 <-data.frame(zsm.statsmat[16:25])


#truncate at year
#zsm
which(rownames(zsm.stdc)=="1564")
zsm.stdceps <-zsm.stdc[52:489,]
zsm.resceps <-zsm.resct[52:489,]

write.crn(zsm.stdceps, "ZSM_Stdf.crn")
write.crn(zsm.resceps, "ZSM_Resf.crn")

#get median length of series contributing to final chron
zsm.truncraw <-zsm.rwi[52:489,]

zsm.colcount <-list()
zsm.truncrawsum <-
 for (i in 1:ncol(zsm.rwi)) {
   colcounts <-sum(!is.na(zsm.truncraw[,i]))
   zsm.colcount[[i]] <-colcounts
 }
zsm.colco <-data.frame(zsm.colcount)
zsm.med <-median(as.numeric(zsm.colco))

#----
#ZTG
ztg.statsmat <-list()
for (i in 1:25){
 stats<-rwi.stats.running(ztg.rwi, ztg.ids, running.window=TRUE, window.length=50, first.start=i)
 ztg.statsmat[[i]]<-stats[,c(1, 15)]
}

#aggregate files in data frames
```

```
ztg.statsmat1 <-data.frame(ztg.statsmat[1:18])
ztg.statsmat2 <-data.frame(ztg.statsmat[19:25])
#truncate at year
#ztg
which(rownames(ztg.stdc)=="1639")
ztg.stdceps <-ztg.stdc[379:742,]
ztg.resceps <-ztg.resct[379:742,]

write.crn(ztg.stdceps, "ZTG_Stdf.crn")
write.crn(ztg.resceps, "ZTG_Resf.crn")

#get median length of series contributing to final chron
ztg.truncraw <-ztg.rwi[379:742,]

ztg.colcount <-list()
ztg.truncrawsum <-
 for (i in 1:ncol(ztg.rwi)) {
  colcounts <-sum(!is.na(ztg.truncraw[,i]))
  ztg.colcount[[i]] <-colcounts
 }
ztg.colco <-data.frame(ztg.colcount)
ztg.med <-median(as.numeric(ztg.colco))
```

D.1.4 Khangai Mountain Region Hydroclimate Data and Tree-Ring Analyses

```
#------------------------------------------------------------------------
#  TITLE: PQ_TR_Khangai_for_Reconstructions
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2015-10-13
#  LAST REVISION: 2015-11-03
#  DESCRIPTION: This script provides code for analyzing Hydromet data and tree ring chronologies.
#  PACKAGES REQUIRED:
#  VARIABLES/DATA USED:
#  NAME:
# TYPE:
# COMMENT:
#------------------------------------------------------------------------
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/")

#libraries
library(dplR)
library(zoo)
#library(Hmisc) #dont load this library when using ggplot2 due to conflicts with discrete
library(reshape2)
library(plyr)
library(ggplot2)

#-------------------------------------------------------------------------------
#import finalized chronologies
#----
#std
jgb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
```

```
nal_Crns/Std/JGB_Stdf.crn")
khu.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KHU_Stdf.crn")
kll.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLL_Stdf.crn")
klp.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLP_Stdf.crn")
mhm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/MHM_Stdf.crn")
ndb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/NDB_Stdf.crn")
ogh.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/OGH_Stdf.crn")
slb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/SLB_Stdf.crn")
zsm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/ZSM_Stdf.crn")
ztg.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/ZTG_Stdf.crn")

#res
jgb.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/JGB_Resf.crn")
khu.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/KHU_Resf.crn")
kll.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/KLL_Resf.crn")
klp.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/KLP_Resf.crn")
mhm.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/MHM_Resf.crn")
ndb.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/NDB_Resf.crn")
ogh.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/OGH_Resf.crn")
slb.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Res/SLB_Resf.crn")
zsm.res <-
```

```
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Final_Crns/Res/ZSM_Resf.crn")
ztg.res <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Final_Crns/Res/ZTG_Resf.crn")

#convert each set to zoo series and collate
#std zoo, KHU and no KHU
#no KHU
std.chron <-list(jgb.std[,1], kll.std[,1], klp.std[,1], mhm.std[,1], ndb.std[,1], ogh.std[,1], slb.std[,1], zsm.std[,1],
ztg.std[,1])
std.chronrows <-list(as.numeric(rownames(jgb.std)), as.numeric(rownames(kll.std)),
as.numeric(rownames(klp.std)), as.numeric(rownames(mhm.std)), as.numeric(rownames(ndb.std)),
as.numeric(rownames(ogh.std)), as.numeric(rownames(slb.std)), as.numeric(rownames(zsm.std)),
as.numeric(rownames(ztg.std)))
std.chronzall <-list()

for(i in 1:length(std.chron)){
std.ind <-std.chronrows[[i]]
std.chronz <-zoo(std.chron[[i]], std.chronrows[[i]])
std.chronzall[[i]] <-std.chronz
}

#matchup listed zoo chrons into one object
std.chronnokhu <-
merge(std.chronzall[[1]],std.chronzall[[2]],std.chronzall[[3]],std.chronzall[[4]],std.chronzall[[5]],std.chronzal
l[[6]],std.chronzall[[7]],std.chronzall[[8]],std.chronzall[[9]])
colnames(std.chronnokhu) <-c("JGB.std", "KLL.std", "KLP.std", "MHM.std", "NDB.std","OGH.std", "SLB.std",
"ZSM.std", "ZTG.std")

#with KHU
std.chronk <-list(jgb.std[,1], khu.std[,1],kll.std[,1], klp.std[,1], mhm.std[,1], ndb.std[,1], ogh.std[,1], slb.std[,1],
zsm.std[,1], ztg.std[,1])
std.chronrowsk <-list(as.numeric(rownames(jgb.std)),
as.numeric(rownames(khu.std)),as.numeric(rownames(kll.std)), as.numeric(rownames(klp.std)),
as.numeric(rownames(mhm.std)), as.numeric(rownames(ndb.std)), as.numeric(rownames(ogh.std)),
as.numeric(rownames(slb.std)), as.numeric(rownames(zsm.std)), as.numeric(rownames(ztg.std)))
std.chronzallk <-list()

for(i in 1:length(std.chronk)){
 std.indk <-std.chronrowsk[[i]]
 std.chronzk <-zoo(std.chronk[[i]], std.chronrowsk[[i]])
 std.chronzallk[[i]] <-std.chronzk
}

#matchup listed zoo chrons into one object
std.chronkhu <-
merge(std.chronzallk[[1]],std.chronzallk[[2]],std.chronzallk[[3]],std.chronzallk[[4]],std.chronzallk[[5]],std.ch
ronzallk[[6]],std.chronzallk[[7]],std.chronzallk[[8]],std.chronzallk[[9]],std.chronzallk[[10]] )
colnames(std.chronkhu) <-c("JGB.std", "KHU.std","KLL.std", "KLP.std", "MHM.std", "NDB.std","OGH.std",
"SLB.std", "ZSM.std", "ZTG.std")

#res zoo, KHU and no KHU
#no KHU
res.chron <-list(jgb.res[,1], kll.res[,1], klp.res[,1], mhm.res[,1], ndb.res[,1], ogh.res[,1], slb.res[,1], zsm.res[,1],
```

```
ztg.res[,1])
res.chronrows <-list(as.numeric(rownames(jgb.res)), as.numeric(rownames(kll.res)),
as.numeric(rownames(klp.res)), as.numeric(rownames(mhm.res)), as.numeric(rownames(ndb.res)),
as.numeric(rownames(ogh.res)), as.numeric(rownames(slb.res)), as.numeric(rownames(zsm.res)),
as.numeric(rownames(ztg.res)))
res.chronzall <-list()

for(i in 1:length(res.chron)){
 res.ind <-res.chronrows[[i]]
 res.chronz <-zoo(res.chron[[i]], res.chronrows[[i]])
 res.chronzall[[i]] <-res.chronz
}

#matchup listed zoo chrons into one object
res.chronnokhu <-
merge(res.chronzall[[1]],res.chronzall[[2]],res.chronzall[[3]],res.chronzall[[4]],res.chronzall[[5]],res.chronzal
l[[6]],res.chronzall[[7]],res.chronzall[[8]],res.chronzall[[9]])
colnames(res.chronnokhu) <-c("JGB.res", "KLL.res", "KLP.res", "MHM.res", "NDB.res","OGH.res", "SLB.res",
"ZSM.res", "ZTG.res")

#with KHU
res.chronk <-list(jgb.res[,1], khu.res[,1],kll.res[,1], klp.res[,1], mhm.res[,1], ndb.res[,1], ogh.res[,1], slb.res[,1],
zsm.res[,1], ztg.res[,1])
res.chronrowsk <-list(as.numeric(rownames(jgb.res)),
as.numeric(rownames(khu.res)),as.numeric(rownames(kll.res)), as.numeric(rownames(klp.res)),
as.numeric(rownames(mhm.res)), as.numeric(rownames(ndb.res)), as.numeric(rownames(ogh.res)),
as.numeric(rownames(slb.res)), as.numeric(rownames(zsm.res)), as.numeric(rownames(ztg.res)))
res.chronzallk <-list()

for(i in 1:length(res.chronk)){
 res.indk <-res.chronrowsk[[i]]
 res.chronzk <-zoo(res.chronk[[i]], res.chronrowsk[[i]])
 res.chronzallk[[i]] <-res.chronzk
}

#matchup listed zoo chrons into one object
res.chronkhu <-
merge(res.chronzallk[[1]],res.chronzallk[[2]],res.chronzallk[[3]],res.chronzallk[[4]],res.chronzallk[[5]],res.ch
ronzallk[[6]],res.chronzallk[[7]],res.chronzallk[[8]],res.chronzallk[[9]],res.chronzallk[[10]] )
colnames(res.chronkhu) <-c("JGB.res", "KHU.res","KLL.res", "KLP.res", "MHM.res", "NDB.res","OGH.res",
"SLB.res", "ZSM.res", "ZTG.res")

#------------------------------------------------------
#import Q and P data (use GPCC and BogdT (fill procedure slightly different but should be adequate)) too as
that's what I said for AGU)
#streamflow
transq1 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/M
onthlyQ/Bogd_FilledMeanMonthly_QTrans.csv")
transq1$mon <-sprintf("%02d", transq1$Month)
transq2 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/M
onthlyQ/Filled_SQRT_Q_from_P_1976-2012.csv")

transq <-data.frame(transq2[,1:5],transq1[5],transq1[,1], transq1[,6])
```

```
colnames(transq) <-c("yrmon", "BayanB", "BayanT", "ErdK", "IkhKT", "BogdT", "Year", "Month")

transqtrans <-data.frame(Year=transq[,7],(transq[,2:6]^2))

yearlyq <-ddply(transqtrans, .(Year), summarize, BayanB=sum(BayanB), BayanT=sum(BayanT),
ErdK=sum(ErdK), IkhKT=sum(IkhKT), BogdT=sum(BogdT))

avgq <-colMeans(yearlyq[,2:6], na.rm=TRUE)
#BayanB   BayanT    ErdK    IkhKT    BogdT
#317.48061  94.59820 154.98084 192.09194  27.59341

#aggregate is correct due to untransformation

#precip
transp1 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/M
onthlyP/GPCC_Mean_P_1976_2010_Bogd.csv")
transp2 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/M
onthlyP/Transformed_GPCC_1976-2010.csv")

transp <-data.frame(transp2[,1:5],transp1[,3],transq1[1:420,1], transq1[1:420,6])
colnames(transp) <-c("yrmon", "BayanB", "BayanT", "ErdK", "IkhKT", "BogdT", "Year", "Month")

#correlate!
#note that should be looking at lagged effects too. i.e. previous April thru current November
#can use with KHU datasets here as the period of correlation is 1976-2011 (or shorter for 1999, etc chrons)

#create corr matrices
#truncate to period of interest 1976-2011 (std.chronkhu, res.chronkhu)
std.chronkhut <-window(std.chronkhu,start=1976)
res.chronkhut <-window(res.chronkhu,start=1976)

#extend precip to 2011 for corr
transp11 <-data.frame(rep(NA,12),rep(NA,12),rep(NA,12),rep(NA,12),rep(NA,12),rep(2011,
12),rep(sprintf("%02d",seq(1,12,1))))
transp11c <-cbind((paste(transp11[,6], transp11[,7], sep="-")), transp11)
colnames(transp11c) <-colnames(transp)
transpf <-rbind(transp, transp11c)

#truncate q for corr
transqf<-transq[1:432,]

#check autocorrelation
acf(transpf[,2], plot=FALSE, na.action=na.omit)

#check corr between gages/grids/tr
pcorrmat <-as.matrix(transpf[,2:6])
pcorrgrid <-rcorr(pcorrmat, type="pearson")

qcorrmat <-as.matrix(transqf[,2:6])
qcorrgrid <-rcorr(qcorrmat, type="pearson")

#check correlation of tr time series
trmat <-as.matrix(std.chronkhut[,1:10])
```

```
trcorr<-rcorr(trmat, type="pearson")

#make zoo objects for lagging
transpfz <-zoo(transpf[,2:8], transpf[,1])
transqfz <-zoo(transqf[,2:8], transqf[,1])

#lag by month
transpl <-lag(transpfz, k=-12)
transql <-lag(transqfz, k=-12)
colnames(transpl) <-c("BayanB_P","BayanT_P", "ErdK_P", "IkhKT_P", "BogdT_P", "Year_P", "Month_P")
colnames(transql) <-c("BayanB_P","BayanT_P", "ErdK_P", "IkhKT_P", "BogdT_P", "Year_P", "Month_P")

#convert back to df to merge as merge.zoo isn't working
transpldf <-data.frame(index(transpl),transpl)
colnames(transpldf) <-c("yrmon", "BayanB_P",  "BayanT_P", "ErdK_P","IkhKT_P", "BogdT_P", "Year_P",
"Month_P")
transplf <-merge(transpf,transpldf, by.x="yrmon", by.y="yrmon", all.x=TRUE )

transqldf <-data.frame(index(transql),transql)
colnames(transqldf) <-c("yrmon", "BayanB_P",  "BayanT_P", "ErdK_P","IkhKT_P", "BogdT_P", "Year_P",
"Month_P")
transqlf <-merge(transqf,transqldf, by.x="yrmon", by.y="yrmon", all.x=TRUE )

#convert factors to numeric
transplf$BayanB_P <-as.numeric(levels(transplf$BayanB_P))[transplf$BayanB_P]
transplf$BayanT_P <-as.numeric(levels(transplf$BayanT_P))[transplf$BayanT_P]
transplf$ErdK_P <-as.numeric(levels(transplf$ErdK_P))[transplf$ErdK_P]
transplf$IkhKT_P <-as.numeric(levels(transplf$IkhKT_P))[transplf$IkhKT_P]
transplf$BogdT_P <-as.numeric(levels(transplf$BogdT_P))[transplf$BogdT_P]

transqlf$BayanB_P <-as.numeric(levels(transqlf$BayanB_P))[transqlf$BayanB_P]
transqlf$BayanT_P <-as.numeric(levels(transqlf$BayanT_P))[transqlf$BayanT_P]
transqlf$ErdK_P <-as.numeric(levels(transqlf$ErdK_P))[transqlf$ErdK_P]
transqlf$IkhKT_P <-as.numeric(levels(transqlf$IkhKT_P))[transqlf$IkhKT_P]
transqlf$BogdT_P <-as.numeric(levels(transqlf$BogdT_P))[transqlf$BogdT_P]

#fill yearp and monthp
transplf$Year_P[1:12]<-transplf$Year[1:12]
transplf$Month_P[1:12]<-transplf$Month[1:12]

transqlf$Year_P[1:12]<-transqlf$Year[1:12]
transqlf$Month_P[1:12]<-transqlf$Month[1:12]

#create matrices of TR year and each month to test for p and q
#pull out individual months by year
pjan <-transplf[seq(1,432,12),]
pfeb <-transplf[seq(2,432,12),]
pmar <-transplf[seq(3,432,12),]
papr <-transplf[seq(4,432,12),]
pmay <-transplf[seq(5,432,12),]
pjun <-transplf[seq(6,432,12),]
pjul <-transplf[seq(7,432,12),]
paug <-transplf[seq(8,432,12),]
psep <-transplf[seq(9,432,12),]
poct <-transplf[seq(10,432,12),]
```

```
pnov <-transplf[seq(11,432,12),]
pdec <-transplf[seq(12,432,12),]

qjan <-transqlf[seq(1,432,12),]
qfeb <-transqlf[seq(2,432,12),]
qmar <-transqlf[seq(3,432,12),]
qapr <-transqlf[seq(4,432,12),]
qmay <-transqlf[seq(5,432,12),]
qjun <-transqlf[seq(6,432,12),]
qjul <-transqlf[seq(7,432,12),]
qaug <-transqlf[seq(8,432,12),]
qsep <-transqlf[seq(9,432,12),]
qoct <-transqlf[seq(10,432,12),]
qnov <-transqlf[seq(11,432,12),]
qdec <-transqlf[seq(12,432,12),]

monp <-list(pjan,pfeb,pmar,papr,pmay,pjun,pjul,paug,psep,poct,pnov,pdec)
monq <-list(qjan,qfeb,qmar,qapr,qmay,qjun,qjul,qaug,qsep,qoct,qnov,qdec)

#extract each TR site, each column of hydroclim and match up in list then test and return result
monthp <-list()
for (i in 1:length(monp)){
 transplff <-monp[[i]][,c(2:6,9:13)]
 rownames(transplff) <-monp[[i]][,1]
 monthp[[i]]<-transplff
}

monthq<-list()
for (i in 1:length(monq)){
transqlff <-monq[[i]][,c(2:6,9:13)]
rownames(transqlff) <-monq[[i]][,1]
monthq[[i]]<-transqlff
}

#correlation precip
monthhead <-factor(c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"),
levels=c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"))

#std chron
corrmatprs <-list()
corrmatpps <-list()

for (i in 1:length(monthp)){
trs <-as.matrix(std.chronkhut)
climps <-as.matrix(monthp[[i]])
climatps <-rcorr(trs, climps)
corrmatprs[[i]] <-climatps$r
corrmatpps[[i]] <-climatps$P
}
names(corrmatprs) <-monthhead
names(corrmatpps) <-monthhead
#residual chron
corrmatprr <-list()
corrmatppr <-list()
```

```
for (i in 1:length(monthp)){
 trs <-as.matrix(res.chronkhut)
 climpr <-as.matrix(monthp[[i]])
 climatpr <-rcorr(trs, climpr)
 corrmatprr[[i]] <-climatpr$r
 corrmatppr[[i]] <-climatpr$P
}
names(corrmatprr) <-monthhead
names(corrmatppr) <-monthhead

#for streamflow

#std chron
corrmatqrs <-list()
corrmatqps <-list()

for (i in 1:length(monthq)){
 trs <-as.matrix(std.chronkhut)
 climqs <-as.matrix(monthq[[i]])
 climatqs <-rcorr(trs, climqs)
 corrmatqrs[[i]] <-climatqs$r
 corrmatqps[[i]] <-climatqs$P
}
names(corrmatqrs) <-monthhead
names(corrmatqps) <-monthhead

#residual chron
corrmatqrr <-list()
corrmatqpr <-list()

for (i in 1:length(monthq)){
 trr <-as.matrix(res.chronkhut)
 climqr <-as.matrix(monthp[[i]])
 climatqr <-rcorr(trr, climqr)
 corrmatqrr[[i]] <-climatqr$r
 corrmatqpr[[i]] <-climatqr$P
}
names(corrmatqrr) <-monthhead
names(corrmatqpr) <-monthhead

#sift through the findings
#note that each set of correlations also provides information on how well each site correlates with all the
others!
#truncate each dataset to unique values for sites and hydromet data
#for std chrons
corrmatprst <-list()

for (i in 1:length(corrmatprs)){
 corrmatprst[[i]] <-corrmatprs[[i]][1:10, 11:20]
}
corrmatppst <-list()
for (i in 1:length(corrmatpps)){
 corrmatppst[[i]] <-corrmatpps[[i]][1:10, 11:20]
}
```

```
names(corrmatprst) <-monthhead
names(corrmatppst) <-monthhead

corrmatqrst <-list()
for (i in 1:length(corrmatqrs)){
  corrmatqrst[[i]] <-corrmatqrs[[i]][1:10, 11:20]
}

corrmatqpst <-list()
for (i in 1:length(corrmatqps)){
  corrmatqpst[[i]] <-corrmatqps[[i]][1:10, 11:20]
}

names(corrmatqrst) <-monthhead
names(corrmatqpst) <-monthhead

#for residual chrons
corrmatprrt <-list()

for (i in 1:length(corrmatprr)){
  corrmatprrt[[i]] <-corrmatprr[[i]][1:10, 11:20]
}

corrmatpprt <-list()
for (i in 1:length(corrmatppr)){
  corrmatpprt[[i]] <-corrmatppr[[i]][1:10, 11:20]
}

names(corrmatprrt) <-monthhead
names(corrmatpprt) <-monthhead

corrmatqrrt <-list()
for (i in 1:length(corrmatqrr)){
  corrmatqrrt[[i]] <-corrmatqrr[[i]][1:10, 11:20]
}

corrmatqprt <-list()
for (i in 1:length(corrmatqpr)){
  corrmatqprt[[i]] <-corrmatqpr[[i]][1:10, 11:20]
}

names(corrmatqrrt) <-monthhead
names(corrmatqprt) <-monthhead

#for std chrons
#precip: corrmatprst corrmatppst
#streamflow:corrmatqrst,corrmatqpst

#for residual chrons
#precip: corrmatprrt, corrmatpprt
#streamflow:corrmatqrrt, corrmatqprt

#write files for later
write.csv(corrmatprst, "corrmat_p_rval_std.csv")
write.csv(corrmatppst, "corrmat_p_pval_std.csv")
```

```
write.csv(corrmatqrst, "corrmat_q_rval_std.csv")
write.csv(corrmatqpst, "corrmat_q_pval_std.csv")

write.csv(corrmatprrt, "corrmat_p_rval_res.csv")
write.csv(corrmatpprt, "corrmat_p_pval_res.csv")
write.csv(corrmatqrrt, "corrmat_q_rval_res.csv")
write.csv(corrmatqprt, "corrmat_q_pval_res.csv")

#look at significant values, rest NA, significance at p<0.05 (and p<0.10?)

#gives NA for whole frame- need to increment! Think about it!

#try to replace all values p<0.05 with NA
#compare to replacement with NA for p<0.10

#corsigps5 <-list()
 # for (i in 1:length(corrmatppst)){
  #corsigpsl5 <-corrmatppst[[i]]
  #corsigpsl5[corsigpsl5>0.05] <-NA
  #corsigps5[[i]] <-corsigpsl5
#}
#names(corsigps5) <-monthhead

#corsigps1 <-list()
#for (i in 1:length(corrmatqpst)){
#  corsigpsl1 <-corrmatppst[[i]]
 # corsigpsl1[corsigpsl1>0.10] <-NA
  #corsigps1[[i]] <-corsigpsl1
#}
#names(corsigps1) <-monthhead

#almost half again values deemed significant with 0.1 rather than 0.5 criteria.
#use 0.05 here and pvalue files
#for std chrons
corsigps <-list()
for (i in 1:length(corrmatppst)){
  corsigpsl <-corrmatppst[[i]]
  corsigpsl[corsigpsl>0.05] <-NA
  corsigps[[i]] <-corsigpsl
}
names(corsigps) <-monthhead

corsigqs<-list()

for (i in 1:length(corrmatqpst)){
  corsigqsl <-corrmatqpst[[i]]
  corsigqsl[corsigqsl>0.05] <-NA
  corsigqs[[i]] <-corsigqsl
}
names(corsigqs) <-monthhead
#for residual chrons
corsigpr <-list()
for (i in 1:length(corrmatpprt)){
  corsigprl <-corrmatpprt[[i]]
  corsigprl[corsigprl>0.05] <-NA
```

```
  corsigpr[[i]] <-corsigprl
}
names(corsigpr) <-monthhead

corsigqr<-list()

for (i in 1:length(corrmatqprt)){
  corsigqrl <-corrmatqprt[[i]]
  corsigqrl[corsigqrl>0.05] <-NA
  corsigqr[[i]] <-corsigqrl
}
names(corsigqr) <-monthhead

#apply NA's to correlation values and add column of month names
#for std chrons
corsigpsr <-list()
for (i in 1:length(corrmatprst)){
  corsigpslr <-corrmatprst[[i]]
  corsigpslr[is.na(corsigps[[i]])] <-NA
  corsigpsr[[i]] <-corsigpslr
}
names(corsigpsr) <-monthhead

corsigqsr<-list()

for (i in 1:length(corrmatqrst)){
  corsigqslr <-corrmatqrst[[i]]
  corsigqslr[is.na(corsigqs[[i]])] <-NA
  corsigqsr[[i]] <-corsigqslr
}
names(corsigqsr) <-monthhead

#for residual chrons
corsigprr <-list()
for (i in 1:length(corrmatprrt)){
  corsigprlr <-corrmatprrt[[i]]
  corsigprlr[is.na(corsigpr[[i]])] <-NA
  corsigprr[[i]] <-corsigprlr
}
names(corsigprr) <-monthhead

corsigqrr<-list()

for (i in 1:length(corrmatqrrt)){
  corsigqrlr <-corrmatqrrt[[i]]
  corsigqrlr[is.na(corsigqr[[i]])] <-NA
  corsigqrr[[i]] <-corsigqrlr
}
names(corsigqrr) <-monthhead

#extract remaining significant correlations and group for plotting
#stdchrons p:corsigpsr q:corsigqsr
#resid chrons p:corsigprr q:corsigqrr

#make into stacked dataframe with months as a column identifier
```

```
#std chron
signames <-rep(rownames(corsigpsr[[1]]), 12)
sigcorpstd <-ldply(corsigpsr, data.frame, .id="Month")
sigcorpstd$Chron <-signames

sigcorqstd <-ldply(corsigqsr, data.frame, .id="Month")
sigcorqstd$Chron <-signames

#res chron
signamer <-rep(rownames(corsigprr[[1]]), 12)
sigcorpres <-ldply(corsigprr, data.frame, .id="Month")
sigcorpres$Chron <-signamer

sigcorqres <-ldply(corsigqrr, data.frame, .id="Month")
sigcorqres$Chron <-signamer

#write as csv
write.csv(sigcorpstd, "Signif_Corr_Precip_Std.csv")
write.csv(sigcorqstd, "Signif_Corr_Q_Std.csv")
write.csv(sigcorpres, "Signif_Corr_Precip_Res.csv")
write.csv(sigcorqres, "Signif_Corr_Q_Res.csv")

#--------------------------------
#create correlations to seasonal data
#as previous code, but aggregate to seasonal
# need to back transform then transform
#use monthly to get seasonal through aggregation, can use sum as it's MCM and total P
#monp <-list(pjan,pfeb,pmar,papr,pmay,pjun,pjul,paug,psep,poct,pnov,pdec)
#monq <-list(qjan,qfeb,qmar,qapr,qmay,qjun,qjul,qaug,qsep,qoct,qnov,qdec)

#for spring, summer, and fall can use sums, will give na is na's in col
#-----
#for spring
#precipitation
#current years
seasmatcp <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
col2sumcp <-data.frame((pmar[,i+1])^2, (papr[,i+1])^2, (pmay[,i+1])^2)
sumcolcp <-rowSums(col2sumcp)
seasmatcp[,i] <-sqrt(sumcolcp)
}

colnames(seasmatcp) <-colnames(pmar[,2:6])

#previous years
seasmatpp <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
 col2sumpp <-data.frame((as.numeric(pmar[,i+8])^2), (as.numeric(papr[,i+8])^2),
(as.numeric(pmay[,i+8])^2))
 sumcolpp <-rowSums(col2sumpp)
 seasmatpp[,i] <-sqrt(sumcolpp)}
```

```
colnames(seasmatpp) <-colnames(pmar[,9:13])

sprp <-data.frame(pmar[,1], seasmatcp, pmar[,7:8], seasmatpp, pmar[,14:15])
colnames(sprp)[1] <-"yrmon"

#streamflow
#current years
seasmatcq <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
 col2sumcq <-data.frame((qmar[,i+1]^2), (qapr[,i+1]^2), (qmay[,i+1]^2))
 sumcolcq <-rowSums(col2sumcq)
 seasmatcq[,i] <-sqrt(sumcolcq)
}

colnames(seasmatcq) <-colnames(pmar[,2:6])

#previous years

seasmatpq <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
 col2sumpq <-data.frame((as.numeric(qmar[,i+8])^2), (as.numeric(qapr[,i+8])^2),
(as.numeric(qmay[,i+8])^2))
 sumcolpq <-rowSums(col2sumpq)
 seasmatpq[,i] <-sqrt(sumcolpq)
}
colnames(seasmatpq) <-colnames(pmar[,9:13])

sprq <-data.frame(pmar[,1], seasmatcq, pmar[,7:8], seasmatpq, pmar[,14:15])
colnames(sprq)[1] <-"yrmon"

#-----
#for summer
#current years
seasmatcps <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
 col2sumcps <-data.frame(pjun[,i+1]^2, pjul[,i+1]^2, paug[,i+1]^2)
 sumcolcps <-rowSums(col2sumcps)
 seasmatcps[,i] <-sqrt(sumcolcps)
}

colnames(seasmatcps) <-colnames(pjun[,2:6])

#previous years
seasmatpps <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
 col2sumpps <-data.frame(as.numeric(pjun[,i+8])^2, as.numeric(pjul[,i+8])^2, as.numeric(paug[,i+8])^2)
 sumcolpps <-rowSums(col2sumpps)
```

```
  seasmatpps[,i] <-sqrt(sumcolpps)
}
colnames(seasmatpps) <-colnames(pjun[,9:13])

sump <-data.frame(pjun[,1], seasmatcps, pjun[,7:8], seasmatpps, pjun[,14:15])
colnames(sump)[1] <-"yrmon"

#streamflow
#current years
seasmatcqs <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
  col2sumcqs <-data.frame(qjun[,i+1]^2, qjul[,i+1]^2, qaug[,i+1]^2)
  sumcolcqs <-rowSums(col2sumcqs)
  seasmatcqs[,i] <-sqrt(sumcolcqs)
}

colnames(seasmatcqs) <-colnames(pjun[,2:6])

#previous years

seasmatpqs <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
  col2sumpqs <-data.frame(as.numeric(qjun[,i+8])^2, as.numeric(qjul[,i+8])^2, as.numeric(qaug[,i+8])^2)
  sumcolpqs <-rowSums(col2sumpqs)
  seasmatpqs[,i] <-sqrt(sumcolpqs)
}
colnames(seasmatpqs) <-colnames(pjun[,9:13])

sumq <-data.frame(pjun[,1], seasmatcqs, pjun[,7:8], seasmatpqs, pjun[,14:15])
colnames(sumq)[1] <-"yrmon"

#---------
#for fall
#current years
seasmatcpf <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
  col2sumcpf <-data.frame(psep[,i+1]^2, poct[,i+1]^2, pnov[,i+1]^2)
  sumcolcpf <-rowSums(col2sumcpf)
  seasmatcpf[,i] <-sqrt(sumcolcpf)
}

colnames(seasmatcpf) <-colnames(psep[,2:6])

#previous years

seasmatppf <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
```

```
 col2sumppf <-data.frame(as.numeric(psep[,i+8])^2, as.numeric(poct[,i+8])^2, as.numeric(pnov[,i+8])^2)
 sumcolppf <-rowSums(col2sumppf)
 seasmatppf[,i] <-sqrt(sumcolppf)
}
colnames(seasmatppf) <-colnames(psep[,9:13])

falp <-data.frame(psep[,1], seasmatcpf, psep[,7:8], seasmatppf, psep[,14:15])
colnames(falp)[1] <-"yrmon"

#streamflow
#current years
seasmatcqf <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
 col2sumcqf <-data.frame(qsep[,i+1]^2, qoct[,i+1]^2, qnov[,i+1]^2)
 sumcolcqf <-rowSums(col2sumcqf)
 seasmatcqf[,i] <-sqrt(sumcolcqf)
}

colnames(seasmatcqf) <-colnames(psep[,2:6])

#previous years

seasmatpqf <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
 col2sumpqf <-data.frame(as.numeric(qsep[,i+8])^2, as.numeric(qoct[,i+8])^2, as.numeric(qnov[,i+8])^2)
 sumcolpqf <-rowSums(col2sumpqf)
 seasmatpqf[,i] <-sqrt(sumcolpqf)
}
colnames(seasmatpqf) <-colnames(psep[,9:13])

falq <-data.frame(psep[,1], seasmatcqf, psep[,7:8], seasmatpqf, psep[,14:15])
colnames(falq)[1] <-"yrmon"

#--------------
#for winter

#group previous dec with current jan and feb
#insert row of NA and remove last row
pdecnextfirst<-pdec[1,]
pdecnextfirst[,2:6]<-NA
pdect <-pdec[1:35,]
pdecnext <-data.frame(rbind(pdecnextfirst, pdect))

qdecnextfirst<-qdec[1,]
qdecnextfirst[,2:6]<-NA
qdect <-qdec[1:35,]
qdecnext <-data.frame(rbind(qdecnextfirst, qdect))

#current years

seasmatcpw <-matrix(ncol=5, nrow=36, NA)
```

```
#collate columns from 3 df then sum then return in df
for (i in 1:5){
 col2sumcpw <-data.frame(pdecnext[,i+1]^2, pjan[,i+1]^2, pfeb[,i+1]^2)
 sumcolcpw <-rowSums(col2sumcpw)
 seasmatcpw[,i] <-sqrt(sumcolcpw)
}

colnames(seasmatcpw) <-colnames(pjan[,2:6])

#previous years
seasmatppw <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
 col2sumppw <-data.frame(as.numeric(pdecnext[,i+8])^2, as.numeric(pjan[,i+8])^2,
as.numeric(pfeb[,i+8])^2)
 sumcolppw <-rowSums(col2sumppw)
 seasmatppw[,i] <-sqrt(sumcolppw)
}
colnames(seasmatppw) <-colnames(pjan[,9:13])

winp <-data.frame(pjan[,1], seasmatcpw, pjan[,7:8], seasmatppw, pjan[,14:15])
colnames(winp)[1] <-"yrmon"

#streamflow
#current years
seasmatcqw <-matrix(ncol=5, nrow=36, NA)

#collate columns from 3 df then sum then return in df

for (i in 1:5){
 col2sumcqw <-data.frame(qdecnext[,i+1]^2, qjan[,i+1]^2, qfeb[,i+1]^2)
 sumcolcqw <-rowSums(col2sumcqw)
 seasmatcqw[,i] <-sqrt(sumcolcqw)
}

colnames(seasmatcqw) <-colnames(pjan[,2:6])

#previous years

seasmatpqw <-matrix(ncol=5, nrow=36, NA)

for (i in 1:5){
 col2sumpqw <-data.frame(as.numeric(qdecnext[,i+8])^2, as.numeric(qjan[,i+8])^2,
as.numeric(qfeb[,i+8])^2)
 sumcolpqw <-rowSums(col2sumpqw)
 seasmatpqw[,i] <-sqrt(sumcolpqw)
}
colnames(seasmatpqw) <-colnames(pjan[,9:13])

winq <-data.frame(pjan[,1], seasmatcqw, psep[,7:8], seasmatpqw, pjan[,14:15])
colnames(winq)[1] <-"yrmon"

#make last winter current value for 2011 NA's as only is dec 2011
#already that for precip (due to using gpcc data)
#for streamflow
```

```
winq[36, 2:6] <-NA

#-------------------------------------
#seasonal files
#precip
#sprp, sump, falp, winp
#streamflow
#sprq,sumq, falq, winq

#string together in a list for further analysis
seasp <-list(sprp, sump, falp, winp)
seasq <-list(sprq, sumq, falq, winq)
#--------------------------------------------------------------
#extract each TR site, each column of hydroclim and match up in list then test and return result
seasonp <-list()
for (i in 1:length(seasp)){
 transplff <-seasp[[i]][,c(2:6,9:13)]
 rownames(transplff) <-seasp[[i]][,1]
 seasonp[[i]]<-transplff
}

seasonq<-list()
for (i in 1:length(seasq)){
 transqlff <-seasq[[i]][,c(2:6,9:13)]
 rownames(transqlff) <-seasq[[i]][,1]
 seasonq[[i]]<-transqlff
}

#correlation precip


seashead <-factor(c("Spr", "Sum", "Fal", "Win"), levels=c("Win","Spr", "Sum", "Fal"))

#std chron
corrmatprs <-list()
corrmatpps <-list()

for (i in 1:length(seasonp)){
 trs <-as.matrix(std.chronkhut)
 climps <-as.matrix(seasonp[[i]])
 climatps <-rcorr(trs, climps)
 corrmatprs[[i]] <-climatps$r
 corrmatpps[[i]] <-climatps$P
}
names(corrmatprs) <-seashead
names(corrmatpps) <-seashead

#residual chron
corrmatprr <-list()
corrmatppr <-list()


for (i in 1:length(seasonp)){
 trs <-as.matrix(res.chronkhut)
 climpr <-as.matrix(seasonp[[i]])
```

```
  climatpr <-rcorr(trs, climpr)
  corrmatprr[[i]] <-climatpr$r
  corrmatppr[[i]] <-climatpr$P
}
names(corrmatprr) <-seashead
names(corrmatppr) <-seashead

#for streamflow

#std chron
corrmatqrs <-list()
corrmatqps <-list()

for (i in 1:length(seasonq)){
  trs <-as.matrix(std.chronkhut)
  climqs <-as.matrix(seasonq[[i]])
  climatqs <-rcorr(trs, climqs)
  corrmatqrs[[i]] <-climatqs$r
  corrmatqps[[i]] <-climatqs$P
}
names(corrmatqrs) <-seashead
names(corrmatqps) <-seashead

#residual chron
corrmatqrr <-list()
corrmatqpr <-list()

for (i in 1:length(seasonq)){
  trr <-as.matrix(res.chronkhut)
  climqr <-as.matrix(seasonp[[i]])
  climatqr <-rcorr(trr, climqr)
  corrmatqrr[[i]] <-climatqr$r
  corrmatqpr[[i]] <-climatqr$P
}
names(corrmatqrr) <-seashead
names(corrmatqpr) <-seashead

#sift through the findings
#note that each set of correlations also provides information on how well each site correlates with all the others!
#truncate each dataset to unique values for sites and hydromet data
#for std chrons
corrmatprst <-list()

for (i in 1:length(corrmatprs)){
  corrmatprst[[i]] <-corrmatprs[[i]][1:10, 11:20]
}

corrmatppst <-list()
for (i in 1:length(corrmatpps)){
  corrmatppst[[i]] <-corrmatpps[[i]][1:10, 11:20]
}

names(corrmatprst) <-seashead
names(corrmatppst) <-seashead
```

```
corrmatqrst <-list()
for (i in 1:length(corrmatqrs)){
  corrmatqrst[[i]] <-corrmatqrs[[i]][1:10, 11:20]
}

corrmatqpst <-list()
for (i in 1:length(corrmatqps)){
  corrmatqpst[[i]] <-corrmatqps[[i]][1:10, 11:20]
}

names(corrmatqrst) <-seashead
names(corrmatqpst) <-seashead

#for residual chrons
corrmatprrt <-list()

for (i in 1:length(corrmatprr)){
  corrmatprrt[[i]] <-corrmatprr[[i]][1:10, 11:20]
}

corrmatpprt <-list()
for (i in 1:length(corrmatppr)){
  corrmatpprt[[i]] <-corrmatppr[[i]][1:10, 11:20]
}

names(corrmatprrt) <-seashead
names(corrmatpprt) <-seashead

corrmatqrrt <-list()
for (i in 1:length(corrmatqrr)){
  corrmatqrrt[[i]] <-corrmatqrr[[i]][1:10, 11:20]
}

corrmatqprt <-list()
for (i in 1:length(corrmatqpr)){
  corrmatqprt[[i]] <-corrmatqpr[[i]][1:10, 11:20]
}

names(corrmatqrrt) <-seashead
names(corrmatqprt) <-seashead

#for std chrons
#precip: corrmatprst corrmatppst
#streamflow:corrmatqrst,corrmatqpst

#for residual chrons
#precip: corrmatprrt, corrmatpprt
#streamflow:corrmatqrrt, corrmatqprt

#write files for later
write.csv(corrmatprst, "corrmat_p_rval_std_seas.csv")
write.csv(corrmatppst, "corrmat_p_pval_std_seas.csv")
write.csv(corrmatqrst, "corrmat_q_rval_std_seas.csv")
write.csv(corrmatqpst, "corrmat_q_pval_std_seas.csv")
```

```
write.csv(corrmatprrt, "corrmat_p_rval_res_seas.csv")
write.csv(corrmatpprt, "corrmat_p_pval_res_seas.csv")
write.csv(corrmatqrrt, "corrmat_q_rval_res_seas.csv")
write.csv(corrmatqprt, "corrmat_q_pval_res_seas.csv")

#use 0.05 here and pvalue files
#for std chrons
corsigps <-list()
for (i in 1:length(corrmatppst)){
 corsigpsl <-corrmatppst[[i]]
 corsigpsl[corsigpsl>0.05] <-NA
 corsigps[[i]] <-corsigpsl
}
names(corsigps) <-seashead

corsigqs<-list()

for (i in 1:length(corrmatqpst)){
 corsigqsl <-corrmatqpst[[i]]
 corsigqsl[corsigqsl>0.05] <-NA
 corsigqs[[i]] <-corsigqsl
}
names(corsigqs) <-seashead

#for residual chrons
corsigpr <-list()
for (i in 1:length(corrmatpprt)){
 corsigprl <-corrmatpprt[[i]]
 corsigprl[corsigprl>0.05] <-NA
 corsigpr[[i]] <-corsigprl
}
names(corsigpr) <-seashead

corsigqr<-list()

for (i in 1:length(corrmatqprt)){
 corsigqrl <-corrmatqprt[[i]]
 corsigqrl[corsigqrl>0.05] <-NA
 corsigqr[[i]] <-corsigqrl
}
names(corsigqr) <-seashead


#apply NA's to correlation values and add column of month names
#for std chrons
corsigpsr <-list()
for (i in 1:length(corrmatprst)){
 corsigpslr <-corrmatprst[[i]]
 corsigpslr[is.na(corsigps[[i]])] <-NA
 corsigpsr[[i]] <-corsigpslr
}
names(corsigpsr) <-seashead

corsigqsr<-list()
```

```
for (i in 1:length(corrmatqrst)){
  corsigqslr <-corrmatqrst[[i]]
  corsigqslr[is.na(corsigqs[[i]])] <-NA
  corsigqsr[[i]] <-corsigqslr
}
names(corsigqsr) <-seashead

#for residual chrons
corsigprr <-list()
for (i in 1:length(corrmatprrt)){
  corsigprlr <-corrmatprrt[[i]]
  corsigprlr[is.na(corsigpr[[i]])] <-NA
  corsigprr[[i]] <-corsigprlr
}
names(corsigprr) <-seashead

corsigqrr<-list()

for (i in 1:length(corrmatqrrt)){
  corsigqrlr <-corrmatqrrt[[i]]
  corsigqrlr[is.na(corsigqr[[i]])] <-NA
  corsigqrr[[i]] <-corsigqrlr
}
names(corsigqrr) <-seashead

#extract remaining significant correlations and group for plotting
#stdchrons p:corsigpsr q:corsigqsr
#resid chrons p:corsigprr q:corsigqrr

#make into stacked dataframe with months as a column identifier
#std chron
signames <-rep(rownames(corsigpsr[[1]]), 4)
sigcorpstd <-ldply(corsigpsr, data.frame, .id="Season")
sigcorpstd$Chron <-signames

sigcorqstd <-ldply(corsigqsr, data.frame, .id="Season")
sigcorqstd$Chron <-signames

#res chron
signamer <-rep(rownames(corsigprr[[1]]), 4)
sigcorpres <-ldply(corsigprr, data.frame, .id="Season")
sigcorpres$Chron <-signamer

sigcorqres <-ldply(corsigqrr, data.frame, .id="Season")
sigcorqres$Chron <-signamer

#write as csv
write.csv(sigcorpstd, "Signif_Corr_Precip_Std_Seas.csv")
write.csv(sigcorqstd, "Signif_Corr_Q_Std_Seas.csv")
write.csv(sigcorpres, "Signif_Corr_Precip_Res_Seas.csv")
write.csv(sigcorqres, "Signif_Corr_Q_Res_Seas.csv")

#----------
#import std chron correlations results monthly and seasonal
#divide several ways to better understand correlations
```

```
#north: ErdK, IkhKT
#south: BayanB, BayanT, BogdT

monpstd <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Correlations/Final/Signif_Corr_Precip_Std.csv", colClasses=(c("factor",rep("numeric", 10),"factor")))
monqstd <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Correlations/Final/Signif_Corr_Q_Std.csv",colClasses=(c("factor",rep("numeric", 10),"factor")))
seaspstd <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Correlations/Final/Signif_Corr_Precip_Std_Seas.csv",colClasses=(c("factor",rep("numeric", 10),"factor")))
seasqstd <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Correlations/Final/Signif_Corr_Q_Std_Seas.csv",colClasses=(c("factor",rep("numeric", 10),"factor")))

#-----------------------------------------------------------
#stack previous months/seasons
#monthlyp
monpstd$Site <-substr(monpstd$Chron,1,3)
monpstd.c <-monpstd[,c(2:6,13)]
monpstd.p <-monpstd[,c(7:11,13)]
colnames(monpstd.p) <-colnames(monpstd.c)
monpstd.s <-rbind(monpstd.p,monpstd.c )
monpstd.sd <-rep(monpstd[,1],2)
#monp.c <-rep("c", length(monpstd[,1]))
#monp.p <-rep("p", length(monpstd[,1]))
#monpstd.cp <-c(monp.p, monp.c)
#monpstd.s$month <-paste(monpstd.cp, monpstd.sd, sep="")
monpstd.s$month <-monpstd.sd
#monpstd.s$month <-factor(monpstd.s$month, levels=c("pJan", "pFeb", "pMar", "pApr", "pMay", "pJun",
"pJul", "pAug", "pSep", "pOct", "pNov","pDec","cJan", "cFeb", "cMar", "cApr", "cMay", "cJun", "cJul", "cAug",
"cSep", "cOct", "cNov","cDec"))
monpstd.s$month <-factor(monpstd.s$month, levels=c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug",
"Sep", "Oct", "Nov","Dec"))
monpstd.s$Site <-factor(monpstd.s$Site, levels=c("KHU", "KLL", "KLP", "OGH", "SLB", "ZSM", "JGB", "MHM",
"NDB", "ZTG"))

#stack basins dividing north and south
#south
monpstd.stbb <-monpstd.s[,c(1,6:7)]
colnames(monpstd.stbb) <-c("Correlation","Site", "Month")
monpstd.stbt <-monpstd.s[,c(2,6:7)]
colnames(monpstd.stbt) <-c("Correlation","Site", "Month")
monpstd.stbot <-monpstd.s[,c(5,6:7)]
colnames(monpstd.stbot) <-c("Correlation","Site", "Month")
basinss <-c(rep("Baidrag R. at Bayanburd P",240), rep("Tuin R. at Bayankhongor P",240),rep("Tuin R. at Bogd
P",240))
monpstd.sts <-rbind(monpstd.stbb,monpstd.stbt,monpstd.stbot)
monpstd.sts$Basins <-basinss
monpstd.sts$prevcur <-c(rep("Previous Year", 120), rep("Current Year", 120), rep("Previous Year", 120),
rep("Current Year", 120),rep("Previous Year", 120), rep("Current Year", 120))
monpstd.sts$prevcur <-factor(monpstd.sts$prevcur, levels=c("Previous Year","Current Year"))

#also divide by previous and current year's months
```

```
which(monpstd.sts$Month=="pJan"|monpstd.sts$Month=="pFeb"|monpstd.sts$Month=="pMar"|monpstd.sts
$Month=="pApr"|monpstd.sts$Month=="pMay"|monpstd.sts$Month=="pJun"|monpstd.sts$Month=="pJul"|m
onpstd.sts$Month=="pAug"|monpstd.sts$Month=="pSep"|monpstd.sts$Month=="pOct"|monpstd.sts$Month=
="pNov"|monpstd.sts$Month=="pDec")

monpstd.stsp <-monpstd.sts[c(1:120,241:360,481:600),]
monpstd.stsc <-monpstd.sts[c(121:240,361:480,601:720),]

#north
monpstd.stek <-monpstd.s[,c(3,6:7)]
colnames(monpstd.stek) <-c("Correlation","Site", "Month")
monpstd.stikt <-monpstd.s[,c(4,6:7)]
colnames(monpstd.stikt) <-c("Correlation","Site", "Month")
basinsn <-c(rep("Khanui R. at Erdenemandal P",240),rep("Khoid Tamir R. at Ikhtamir P",240))
monpstd.stn <-rbind(monpstd.stek,monpstd.stikt)
monpstd.stn$Basins <-basinsn
monpstd.stn$prevcur <-c(rep("Previous Year", 120), rep("Current Year", 120), rep("Previous Year", 120),
rep("Current Year", 120))
monpstd.stn$prevcur <-factor(monpstd.stn$prevcur, levels=c("Previous Year","Current Year"))


#Divide previous and current months
monpstd.stnp <-monpstd.stn[c(1:120,241:360),]
monpstd.stnc <-monpstd.stn[c(121:240,361:480),]

#monthlyq
monqstd$Site <-substr(monqstd$Chron,1,3)
monqstd.c <-monqstd[,c(2:6,13)]
monqstd.p <-monqstd[,c(7:11,13)]
colnames(monqstd.p) <-colnames(monqstd.c)
monqstd.s <-rbind(monqstd.p,monqstd.c )
monqstd.sd <-rep(monqstd[,1],2)
#monq.c <-rep("c", length(monqstd[,1]))
#monq.p <-rep("p", length(monqstd[,1]))
#monqstd.cp <-c(monq.p, monq.c)
#monqstd.s$month <-paste(monqstd.cp, monqstd.sd, sep="")
monqstd.s$month <-monqstd.sd
#monqstd.s$month <-factor(monqstd.s$month, levels=c("pJan", "pFeb", "pMar", "pApr", "pMay", "pJun",
"pJul", "pAug", "pSep", "pOct", "pNov","pDec","cJan", "cFeb", "cMar", "cApr", "cMay", "cJun", "cJul", "cAug",
"cSep", "cOct", "cNov","cDec"))
monqstd.s$month <-factor(monqstd.s$month, levels=c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug",
"Sep", "Oct", "Nov","Dec"))
monqstd.s$Site <-factor(monqstd.s$Site, levels=c("KHU", "KLL", "KLP", "OGH", "SLB", "ZSM", "JGB", "MHM",
"NDB", "ZTG"))

#stack basins dividing north and south
#south
monqstd.stbb <-monqstd.s[,c(1,6:7)]
colnames(monqstd.stbb) <-c("Correlation","Site", "Month")
monqstd.stbt <-monqstd.s[,c(2,6:7)]
colnames(monqstd.stbt) <-c("Correlation","Site", "Month")
monqstd.stbot <-monqstd.s[,c(5,6:7)]
colnames(monqstd.stbot) <-c("Correlation","Site", "Month")
basinss <-c(rep("Baidrag R. at Bayanburd",240), rep("Tuin R. at Bayankhongor",240),rep("Tuin R. at
Bogd",240))
```

```
monqstd.sts <-rbind(monqstd.stbb,monqstd.stbt,monqstd.stbot)
monqstd.sts$Basins <-basinss
monqstd.sts$prevcur <-c(rep("Previous Year", 120), rep("Current Year", 120), rep("Previous Year", 120),
rep("Current Year", 120),rep("Previous Year", 120), rep("Current Year", 120))
monqstd.sts$prevcur <-factor(monqstd.sts$prevcur, levels=c("Previous Year","Current Year"))

#divide to previous and current months
monqstd.stsp <-monqstd.sts[c(1:120,241:360,481:600),]
monqstd.stsc <-monqstd.sts[c(121:240,361:480,601:720),]

#north
monqstd.stek <-monqstd.s[,c(3,6:7)]
colnames(monqstd.stek) <-c("Correlation","Site", "Month")
monqstd.stikt <-monqstd.s[,c(4,6:7)]
colnames(monqstd.stikt) <-c("Correlation","Site", "Month")
basinsn <-c(rep("Khanui R. at Erdenemandal",240),rep("Khoid Tamir R. at Ikhtamir",240))
monqstd.stn <-rbind(monqstd.stek,monqstd.stikt)
monqstd.stn$Basins <-basinsn
monqstd.stn$prevcur <-c(rep("Previous Year", 120), rep("Current Year", 120), rep("Previous Year", 120),
rep("Current Year", 120))
monqstd.stn$prevcur <-factor(monqstd.stn$prevcur, levels=c("Previous Year", "Current Year"))

#divide to previous and current months
monqstd.stnp <-monpstd.stn[c(1:120,241:360),]
monqstd.stnc <-monpstd.stn[c(121:240,361:480),]

#-----------------------------------------------
#seasonalp
seaspstd$Site <-substr(seaspstd$Chron,1,3)
seaspstd.c <-seaspstd[,c(2:6,13)]
seaspstd.p <-seaspstd[,c(7:11,13)]
colnames(seaspstd.p) <-colnames(seaspstd.c)
seaspstd.s <-rbind(seaspstd.p,seaspstd.c )
seaspstd.sd <-rep(seaspstd[,1],2)
#seasp.c <-rep("c", length(seaspstd[,1]))
#seasp.p <-rep("p", length(seaspstd[,1]))
#seaspstd.cp <-c(seasp.p, seasp.c)
#seaspstd.s$seasth <-paste(seaspstd.cp, seaspstd.sd, sep="")
seaspstd.s$seasth <-seaspstd.sd
#seaspstd.s$seasth <-factor(seaspstd.s$seasth, levels=c("pWin", "pSpr", "pSum", "pFal", "cWin", "cSpr",
"cSum", "cFal"))
seaspstd.s$seasth <-factor(seaspstd.s$seasth, levels=c("Win", "Spr", "Sum", "Fal"))
seaspstd.s$Site <-factor(seaspstd.s$Site, levels=c("KHU", "KLL", "KLP", "OGH", "SLB", "ZSM", "JGB", "MHM",
"NDB", "ZTG"))

#stack basins dividing north and south
#south
seaspstd.stbb <-seaspstd.s[,c(1,6:7)]
colnames(seaspstd.stbb) <-c("Correlation","Site", "Season")
seaspstd.stbt <-seaspstd.s[,c(2,6:7)]
colnames(seaspstd.stbt) <-c("Correlation","Site", "Season")
seaspstd.stbot <-seaspstd.s[,c(5,6:7)]
colnames(seaspstd.stbot) <-c("Correlation","Site", "Season")
basinss <-c(rep("Baidrag R. at Bayanburd P",80), rep("Tuin R. at Bayankhongor P",80),rep("Tuin R. at Bogd
P",80))
```

```
seaspstd.sts <-rbind(seaspstd.stbb,seaspstd.stbt,seaspstd.stbot)
seaspstd.sts$Basins <-basinss
seaspstd.sts$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40), rep("Previous Year", 40),
rep("Current Year", 40),rep("Previous Year", 40), rep("Current Year", 40))
seaspstd.sts$prevcur <-factor(seaspstd.sts$prevcur, levels=c("Previous Year", "Current Year"))

#north
seaspstd.stek <-seaspstd.s[,c(3,6:7)]
colnames(seaspstd.stek) <-c("Correlation","Site", "Season")
seaspstd.stikt <-seaspstd.s[,c(4,6:7)]
colnames(seaspstd.stikt) <-c("Correlation","Site", "Season")
basinsn <-c(rep("Khanui R. at Erdenemandal P",80),rep("Khoid Tamir R. at Ikhtamir P",80))
seaspstd.stn <-rbind(seaspstd.stek,seaspstd.stikt)
seaspstd.stn$Basins <-basinsn
seaspstd.stn$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40), rep("Previous Year", 40),
rep("Current Year", 40))
seaspstd.stn$prevcur <-factor(seaspstd.stn$prevcur, levels=c("Previous Year", "Current Year"))

#seasonalq
seasqstd$Site <-substr(seasqstd$Chron,1,3)
seasqstd.c <-seasqstd[,c(2:6,13)]
seasqstd.p <-seasqstd[,c(7:11,13)]
colnames(seasqstd.p) <-colnames(seasqstd.c)
seasqstd.s <-rbind(seasqstd.p,seasqstd.c )
seasqstd.sd <-rep(seasqstd[,1],2)
#seasq.c <-rep("c", length(seasqstd[,1]))
#seasq.p <-rep("p", length(seasqstd[,1]))
#seasqstd.cp <-c(seasq.p, seasq.c)
#seasqstd.s$seasth <-paste(seasqstd.cp, seasqstd.sd, sep="")
seasqstd.s$seasth <-seasqstd.sd
#seasqstd.s$seasth <-factor(seasqstd.s$seasth, levels=c("pWin", "pSpr", "pSum", "pFal", "cWin", "cSpr",
"cSum", "cFal"))
seasqstd.s$seasth <-factor(seasqstd.s$seasth, levels=c("Win", "Spr", "Sum", "Fal"))
seasqstd.s$Site <-factor(seasqstd.s$Site, levels=c("KHU", "KLL", "KLP", "OGH", "SLB", "ZSM", "JGB", "MHM",
"NDB", "ZTG"))

#stack basins dividing north and south
#south
seasqstd.stbb <-seasqstd.s[,c(1,6:7)]
colnames(seasqstd.stbb) <-c("Correlation","Site", "Season")
seasqstd.stbt <-seasqstd.s[,c(2,6:7)]
colnames(seasqstd.stbt) <-c("Correlation","Site", "Season")
seasqstd.stbot <-seasqstd.s[,c(5,6:7)]
colnames(seasqstd.stbot) <-c("Correlation","Site", "Season")
basinss <-c(rep("Baidrag R. at Bayanburd",80), rep("Tuin R. at Bayankhongor",80),rep("Tuin R. at Bogd",80))
seasqstd.sts <-rbind(seasqstd.stbb,seasqstd.stbt,seasqstd.stbot)
seasqstd.sts$Basins <-basinss
seasqstd.sts$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40), rep("Previous Year", 40),
rep("Current Year", 40),rep("Previous Year", 40), rep("Current Year", 40))
seasqstd.sts$prevcur <-factor(seasqstd.sts$prevcur, levels=c("Previous Year", "Current Year"))

#north
seasqstd.stek <-seasqstd.s[,c(3,6:7)]
colnames(seasqstd.stek) <-c("Correlation","Site", "Season")
seasqstd.stikt <-seasqstd.s[,c(4,6:7)]
```

```
colnames(seasqstd.stikt) <-c("Correlation","Site", "Season")
basinsn <-c(rep("Khanui R. at Erdenemandal",80),rep("Khoid Tamir R. at Ikhtamir",80))
seasqstd.stn <-rbind(seasqstd.stek,seasqstd.stikt)
seasqstd.stn$Basins <-basinsn
seasqstd.stn$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40), rep("Previous Year", 40),
rep("Current Year", 40))
seasqstd.stn$prevcur <-factor(seasqstd.stn$prevcur, levels=c("Previous Year", "Current Year"))

#--------------------------------------
#plots
colScale <-
scale_fill_manual(values=c("#c7e9b4","#7fcdbb","#41b6c4","#1d91c0","#225ea8","#0c2c84","#fed976","#fd
8d3c","#f03b20","#bd0026"))
#------------------------------------------
#monthly plots
monthpsp <-ggplot(monpstd.stsp, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
 axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
 legend.text=element_text(size=12),legend.title=element_text(size=14), strip.text.x=element_text(size=14))

monthpsc <-ggplot(monpstd.stsc, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
      axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
      legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))

monthpnp <-ggplot(monpstd.stnp, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
      axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
      legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))

monthpnc <-ggplot(monpstd.stnc, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
```

```
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
     axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
     legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))

monthqsp <-ggplot(monqstd.stsp, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
     axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
     legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))

monthqsc <-ggplot(monqstd.stsc, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
     axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
     legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))
monthqnp <-ggplot(monqstd.stnp, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
     axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
     legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))

monthqnc <-ggplot(monpstd.stnc, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(-0.8, 0.8), breaks=c(-0.8,-0.6, -
0.4, -0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_wrap(~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
     axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
     legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14))
#-------------------------------------------------
#using whole dataset
```

```
#monpstd.sts,monpstd.stn,monqstd.stn,monqstd.stn
#--------------------------------------------------------
#positive corr plots monthly with stacked previous year's corr and current year's corr
monthps <-ggplot(monpstd.sts, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
    axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
    legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14),strip.text.y=element_text(size=14))

monthpn <-ggplot(monpstd.stn, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
    axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
    legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14), strip.text.y=element_text(size=14))
monthqs <-ggplot(monqstd.sts, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
    axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
    legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14),strip.text.y=element_text(size=14))

monthqn <-ggplot(monqstd.stn, aes(x=Month, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Month", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
    axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
    legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14), strip.text.y=element_text(size=14))

#----------------------------------------------------
#seasonal plots
seasps <- ggplot(seaspstd.sts, aes(x=Season, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
```

```
 labs(x="Season", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
      axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
      legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14), strip.text.y=element_text(size=14))

seaspn <- ggplot(seaspstd.stn, aes(x=Season, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Season", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
      axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
      legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14), strip.text.y=element_text(size=14))


seasqs <- ggplot(seasqstd.sts, aes(x=Season, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Season", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
      axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
      legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14), strip.text.y=element_text(size=14))

seasqn <- ggplot(seasqstd.stn, aes(x=Season, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Season", fill="Site")+facet_grid(prevcur~Basins) +colScale+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=90, hjust=0, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
      axis.text.y=element_text(size=14),axis.title.y=element_text(size=14,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
      legend.text=element_text(size=12),legend.title=element_text(size=14),
strip.text.x=element_text(size=14), strip.text.y=element_text(size=14))

#Want a plot for fig 3 in paper of Khoid Tamir and Tuin at Bayankhongor
#for previous summer through current summer precip and streamflow (positive and negative)
#symbolize sites as previously decided, but think about grayscale adaptations
#make note of no significant neg corr (also maybe exclude win from plot)

#seasonalq
seasqstd$Site <-substr(seasqstd$Chron,1,3)
seasqstd.c <-seasqstd[,c(2:6,13)]
seasqstd.p <-seasqstd[,c(7:11,13)]
```

```
colnames(seasqstd.p) <-colnames(seasqstd.c)
seasqstd.s <-rbind(seasqstd.p,seasqstd.c )
seasqstd.sd <-rep(seasqstd[,1],2)
seasq.c <-rep("c", length(seasqstd[,1]))
seasq.p <-rep("p", length(seasqstd[,1]))
seasqstd.cp <-c(seasq.p, seasq.c)
seasqstd.s$seasth <-paste(seasqstd.cp, seasqstd.sd, sep="")
#seasqstd.s$seasth <-seasqstd.sd
seasqstd.s$seasth <-factor(seasqstd.s$seasth, levels=c("pWin", "pSpr", "pSum", "pFal", "cWin", "cSpr", "cSum",
"cFal"))
#seasqstd.s$seasth <-factor(seasqstd.s$seasth, levels=c("Win", "Spr", "Sum", "Fal"))
seasqstd.s$Site <-factor(seasqstd.s$Site, levels=c("JGB","KHU", "KLL", "KLP", "MHM","NDB","OGH", "SLB",
"ZSM", "ZTG"))

#seasonalp
seaspstd$Site <-substr(seaspstd$Chron,1,3)
seaspstd.c <-seaspstd[,c(2:6,13)]
seaspstd.p <-seaspstd[,c(7:11,13)]
colnames(seaspstd.p) <-colnames(seaspstd.c)
seaspstd.s <-rbind(seaspstd.p,seaspstd.c )
seaspstd.sd <-rep(seaspstd[,1],2)
seasp.c <-rep("c", length(seaspstd[,1]))
seasp.p <-rep("p", length(seaspstd[,1]))
seaspstd.cp <-c(seasp.p, seasp.c)
seaspstd.s$seasth <-paste(seaspstd.cp, seaspstd.sd, sep="")
#seaspstd.s$seasth <-seaspstd.sd
seaspstd.s$seasth <-factor(seaspstd.s$seasth, levels=c("pWin", "pSpr", "pSum", "pFal", "cWin", "cSpr", "cSum",
"cFal"))
#seaspstd.s$seasth <-factor(seaspstd.s$seasth, levels=c("Win", "Spr", "Sum", "Fal"))
seaspstd.s$Site <-factor(seaspstd.s$Site, levels=c("JGB","KHU", "KLL", "KLP", "MHM","NDB","OGH", "SLB",
"ZSM", "ZTG"))

#reformat plot data to only that needed
#Tuin at Bayankhongor
seasqstd.stbt <-seasqstd.s[,c(2,6:7)]
colnames(seasqstd.stbt) <-c("Correlation","Site", "Season")
basinssbt <-c( rep("Tuin R. at Bayankhongor",80))
seasqstd.stsbt <-seasqstd.stbt
seasqstd.stsbt$Basins <-basinssbt
seasqstd.stsbt$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40))
seasqstd.stsbt$prevcur <-factor(seasqstd.stsbt$prevcur, levels=c("Previous Year", "Current Year"))

#select previous summer, current spring and current summer
btseasq <-seasqstd.stsbt[c(11:20,41:60),]

seaspstd.stbtp <-seaspstd.s[,c(2,6:7)]
colnames(seaspstd.stbtp) <-c("Correlation","Site", "Season")
basinssbtp <-c( rep("Tuin R. at Bayankhongor",80))
seaspstd.stsbtp <-seaspstd.stbtp
seaspstd.stsbtp$Basins <-basinssbtp
seaspstd.stsbtp$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40))
seaspstd.stsbtp$prevcur <-factor(seaspstd.stsbtp$prevcur, levels=c("Previous Year", "Current Year"))

#select previous spring, current spring and current summer
btseasp <-seaspstd.stsbtp[c(11:20,41:60),]
```

```
#north
seasqstd.stikt <-seasqstd.s[,c(4,6:7)]
colnames(seasqstd.stikt) <-c("Correlation","Site", "Season")
basinsnkt <-c(rep("Khoid Tamir R. at Ikhtamir",80))
seasqstd.stnkt <-seasqstd.stikt
seasqstd.stnkt$Basins <-basinsnkt
seasqstd.stnkt$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40))
seasqstd.stnkt$prevcur <-factor(seasqstd.stnkt$prevcur, levels=c("Previous Year", "Current Year"))

#select previous spring, current spring and current summer
ktseasq <-seasqstd.stnkt[c(11:20,41:60),]

seaspstd.stikt <-seaspstd.s[,c(4,6:7)]
colnames(seaspstd.stikt) <-c("Correlation","Site", "Season")
basinsnktp <-c(rep("Khoid Tamir R. at Ikhtamir",80))
seaspstd.stnktp <-seaspstd.stikt
seaspstd.stnktp$Basins <-basinsnktp
seaspstd.stnktp$prevcur <-c(rep("Previous Year", 40), rep("Current Year", 40))
seaspstd.stnktp$prevcur <-factor(seaspstd.stnktp$prevcur, levels=c("Previous Year", "Current Year"))

#select previous spring, current spring and current summer
ktseasp <-seaspstd.stnktp[c(11:20,41:60),]

#sites correlated
#south- JGBp, KHU,KLP,OGH,ZSM,JGB,MHM
#north- KHUp,OGHp,SLBp,KHU,KLL,KLP,OGH,SLB,ZSM,MHM

#alphabetical and seasonal
#KHUp,JGBp,JGB,KHU,KLL,KLP,MHM,OGH,SLB,ZSM

#reduce sets to sites of interest
btseasq2 <-btseasq[c(1:5,7:9,11:15,17:19,21:25,27:29),]
btseasp2 <-btseasp[c(1:5,7:9,11:15,17:19,21:25,27:29),]

ktseasq2 <-ktseasq[c(1:5,7:9,11:15,17:19,21:25,27:29),]
ktseasp2 <-ktseasp[c(1:5,7:9,11:15,17:19,21:25,27:29),]

#conjoin p and q for facet plotting and add season
btseasq2$hydromet <-rep("Streamflow",length(btseasq2[,1]))
btseasp2$hydromet <-rep("Precipitation",length(btseasp2[,1]))

ktseasq2$hydromet <-rep("Streamflow",length(ktseasq2[,1]))
ktseasp2$hydromet <-rep("Precipitation",length(ktseasp2[,1]))

#stack
btseas2 <-rbind(btseasp2,btseasq2)
names(btseas2)[names(btseas2) == "Season"] <- "ShSeason"
btseas2$Season <-c(rep("pSummer", 8), rep("cSpring", 8), rep("cSummer", 8),rep("pSummer", 8),
rep("cSpring", 8), rep("cSummer", 8))
btseas2$Season <-factor(btseas2$Season, levels=c("pSummer", "cSpring", "cSummer"))

ktseas2 <-rbind(ktseasp2,ktseasq2)
names(ktseas2)[names(ktseas2) == "Season"] <- "ShSeason"
ktseas2$Season <-c(rep("pSummer", 8), rep("cSpring", 8), rep("cSummer", 8),rep("pSummer", 8),
rep("cSpring", 8), rep("cSummer", 8))
```

```
ktseas2$Season <-factor(ktseas2$Season, levels=c("pSummer", "cSpring", "cSummer"))

#plots
#detach(package:Hmisc, unload=TRUE)

colScale2 <-scale_fill_manual(values=c("#999999", "#E69F00", "#56B4E9", "#CC79A7", "#996633",
"#0072B2", "#D55E00", "#006633"))

btseasplot <- ggplot(btseas2, aes(x=Season, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(0.0,0.2, 0.4,
0.6, 0.8))+
 labs(x="Season", fill="Site")+facet_grid(hydromet~Basins) +colScale2+
 geom_hline(aes(yintercept=0.0), col="grey80")+theme_bw()+
 theme(axis.text.x=element_text(hjust=0.5, vjust=0.5, size=18),axis.title.x=element_text(size=18,
face="bold"),
	axis.text.y=element_text(size=18),axis.title.y=element_text(size=18, face="bold"),
	legend.text=element_text(size=16),legend.title=element_text(size=18),
strip.text.x=element_text(size=18), strip.text.y=element_text(size=18),
panel.grid.major.x=element_blank(),panel.grid.minor.x=element_blank(),panel.grid.major.y=element_line(siz
e=0.1, color="grey60"))

ktseasplot <- ggplot(ktseas2, aes(x=Season, y=Correlation, fill=Site)) +
 geom_bar(stat="identity", position="dodge") + scale_y_continuous(limits=c(0.0, 0.8), breaks=c(-0.8,-0.6, -0.4,
-0.2, 0,0.2, 0.4, 0.6, 0.8))+
 labs(x="Season", fill="Site")+facet_grid(hydromet~Basins) +colScale2+theme_bw()+
 geom_hline(aes(yintercept=0.0), col="grey80")+
 theme(axis.text.x=element_text(angle=0, hjust=0.5, vjust=0.5, size=14),axis.title.x=element_text(size=14,
face="bold"),
	axis.text.y=element_text(size=18),axis.title.y=element_text(size=18,
face="bold"),panel.grid.major.x=element_blank(),panel.grid.major.y=element_line(size=0.1, color="grey90"),
	legend.text=element_text(size=16),legend.title=element_text(size=18),
strip.text.x=element_text(size=18), strip.text.y=element_text(size=18),
panel.grid.major.x=element_blank(),panel.grid.minor.x=element_blank(),panel.grid.major.y=element_line(siz
e=0.1, color="grey60"))
```

D.1.5 Khangai Mountain Region Streamflow Reconstructions

```
#--------------------------------------------------------------------------
#  TITLE: Khangai_Reconstructions
#  AUTHOR: Niah Venable
#  DATE WRITTEN: 2015-10-19
#  LAST REVISION: 2015-11-04
#  DESCRIPTION: This script provides code for analyzing Hydromet data and tree ring chonologies.
#  PACKAGES REQUIRED:
#  VARIABLES/DATA USED:
#  NAME:
# TYPE:
# COMMENT:
#--------------------------------------------------------------------------
#Set your working directory where the input file is located
setwd("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/")

#libraries
library(dplR)
```

```
library(zoo)
library(plyr)
library(ggplot2)
#library(pls) only needed for PCR analysis
library(hydroGOF)
library(leaps)
library(car)
library(DAAG)
library(reshape2)

#---------------------------------------------------------------------------------
#import monthly correlation matrix
#truncate to correlation months of previous May to current September
sigcorr <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Signif_Corr_Q_Std.csv" )
#truncate to months of interest
sigcorr.t <-sigcorr[41:90,]

#divide into current and previous
sigcorr.c <-sigcorr.t[,c(1:6,12)]
sigcorr.p <-sigcorr.t[,c(1,7:12)]

#sort by highest corr for each basin and each period
bbcor.p <-sigcorr.p[order(-sigcorr.p$BayanB_P),]
bbcor.c <-sigcorr.c[order(-sigcorr.c$BayanB),]

btcor.p <-sigcorr.p[order(-sigcorr.p$BayanT_P),]
btcor.c <-sigcorr.c[order(-sigcorr.c$BayanT),]

ekcor.p <-sigcorr.p[order(-sigcorr.p$ErdK_P),]
ekcor.c <-sigcorr.c[order(-sigcorr.c$ErdK),]

ikcor.p <-sigcorr.p[order(-sigcorr.p$IkhKT_P),]
ikcor.c <-sigcorr.c[order(-sigcorr.c$IkhKT),]

#create order from ranking of chrons and significance
#use for order of final predictors
bbcor.ptr <-unique(bbcor.p$Chron)
bbcor.ctr <-unique(bbcor.c$Chron)

btcor.ptr <-unique(btcor.p$Chron)
btcor.ctr <-unique(btcor.c$Chron)

ekcor.ptr <-unique(ekcor.p$Chron)
ekcor.ctr <-unique(ekcor.c$Chron)

ikcor.ptr <-unique(ikcor.p$Chron)
ikcor.ctr <-unique(ikcor.c$Chron)

#interleaved prior and current
#stack prior and current, add column to denote which is which
#sigcorr.c <-sigcorr.t[,c(1:6,12)]
#sigcorr.p <-sigcorr.t[,c(1,7:12)]
colnames(sigcorr.p) <-colnames(sigcorr.c)
```

```
monstd.s <-rbind(sigcorr.p,sigcorr.c )
mon.c <-rep("c", length(sigcorr.c[,1]))
mon.p <-rep("p", length(sigcorr.p[,1]))
monstd.cp <-c(mon.p, mon.c)
monstd.s$month <-paste(monstd.cp, monstd.s[,1], sep="")
monstd.s$month <-factor(monstd.s$month, levels=c("pJan", "pFeb", "pMar", "pApr", "pMay", "pJun", "pJul",
"pAug", "pSep", "pOct", "pNov","pDec","cJan", "cFeb", "cMar", "cApr", "cMay", "cJun", "cJul", "cAug", "cSep",
"cOct", "cNov","cDec"))

#sort by highest corr for each basin and each period
bbcor <-monstd.s[order(-monstd.s$BayanB),]
btcor <-monstd.s[order(-monstd.s$BayanT),]
ekcor <-monstd.s[order(-monstd.s$ErdK),]
ikcor <-monstd.s[order(-monstd.s$IkhKT),]


#create order from ranking of chrons and significance
#use for order of predictors
bbcor.tr <-unique(bbcor$Chron)
btcor.tr <-unique(btcor$Chron)
ekcor.tr <-unique(ekcor$Chron)
ikcor.tr <-unique(ikcor$Chron)

#-------------------------------------
#import finalized std chronologies
#----
#std- do not use KHU here due to short series lengths
jgb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/JGB_Stdf.crn")
#khu.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KHU_Stdf.crn")
kll.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLL_Stdf.crn")
klp.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLP_Stdf.crn")
mhm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/MHM_Stdf.crn")
ndb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/NDB_Stdf.crn")
ogh.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/OGH_Stdf.crn")
slb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/SLB_Stdf.crn")
zsm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/ZSM_Stdf.crn")
ztg.std <-
```

```
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/ZTG_Stdf.crn")

#convert each set to zoo series and collate
#std zoo, KHU and no KHU
#no KHU
std.chron <-list(jgb.std[,1], kll.std[,1], klp.std[,1], mhm.std[,1], ndb.std[,1], ogh.std[,1], slb.std[,1], zsm.std[,1],
ztg.std[,1])
std.chronrows <-list(as.numeric(rownames(jgb.std)), as.numeric(rownames(kll.std)),
as.numeric(rownames(klp.std)), as.numeric(rownames(mhm.std)), as.numeric(rownames(ndb.std)),
as.numeric(rownames(ogh.std)), as.numeric(rownames(slb.std)), as.numeric(rownames(zsm.std)),
as.numeric(rownames(ztg.std)))
std.chronzall <-list()

for(i in 1:length(std.chron)){
std.ind <-std.chronrows[[i]]
std.chronz <-zoo(std.chron[[i]], std.chronrows[[i]])
std.chronzall[[i]] <-std.chronz
}

#matchup listed zoo chrons into one object
std.chronnokhu <-
merge(std.chronzall[[1]],std.chronzall[[2]],std.chronzall[[3]],std.chronzall[[4]],std.chronzall[[5]],std.chronzal
l[[6]],std.chronzall[[7]],std.chronzall[[8]],std.chronzall[[9]])
colnames(std.chronnokhu) <-c("JGB", "KLL", "KLP", "MHM", "NDB","OGH", "SLB", "ZSM", "ZTG")

#truncate to common period for first analysis 1650-1999 and add lagged predictors
std.chron1 <-window(std.chronnokhu, start=1650, end=1999)

#add lagged predictors
std.chron1l <-lag(std.chron1, k=1)
colnames(std.chron1l) <-c("JGB1", "KLL1", "KLP1", "MHM1", "NDB1","OGH1", "SLB1", "ZSM1", "ZTG1")

#merge datasets together
std.chron1m <-merge(std.chron1,std.chron1l)

#truncate to remove NAs and create calibration set
#common period becomes 1650-1998 due to lag
std.chron1t <-window(std.chron1m, start=1650, end=1998)

#truncate to model period to match common perod with Q data (1977 (dendroyear)-1998 only 22 years!)
std.chron1tc <-window(std.chron1t, start=1977, end=1998)

#df for analysis
df1.tr <-data.frame(coredata(std.chron1tc))

#------------------------------------------------------
#import Q and P data (use GPCC and BogdT (fill procedure slightly different but should be adequate)) too as
that's what I said for AGU)
#----
#streamflow not including Bogd (was for poster analysis)
transq <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Filled_S
QRT_Q_from_P_1976-2012.csv")
transq$Year <-substr(transq$yrmon, 1, 4)
```

```
transq$Month <-substr(transq$yrmon, 6,7)

#aggregate Q by dendroclimatic year, previous aug to current july
#classify as no and year ie.e dc yr #1= 1977-01
dcyrdc <-c(rep(1, 7), rep(2:37, each=12), rep(37, 5))
dcyryr <-c(rep(1976, 7), rep(1977:2012, each=12), rep(2013, 5))
transq$dcyr <-dcyryr
transq$dcdc <-dcyrdc

#untransform before aggregation
transqtt <-data.frame(transq[,1], transq[,2:5]^2, transq[,6:9])
colnames(transqtt) <-colnames(transq)


transqag <-ddply(transqtt, .(dcyr, dcdc), summarize, BayanB=sum(BayanB), BayanT=sum(BayanT),
ErdK=sum(ErdK), IkhKT=sum(IkhKT))

#re-transform aggregate values
transqagt <-data.frame(transqag[,1:2], sqrt(transqag[,3:6]))
colnames(transqagt) <-colnames(transqag)

#truncate to calibration period
transqt <-transqagt[2:23,]

#data for analysis
df1.q <-data.frame(transqt[,3:6])

#PCR-Not Used In Final Analyses- only as a test here.
#-----------
#Note: I am not sure PCR is the best method for ananlysis with the high level of correlation seen between
moisture and growth at these sites.
#I feel stepwise regression with judicious use of diagnostics to avoid offitting may provide resonable results.
#Also the results may be more defensible.
#PCR
#testing

#combine to give dataset  (need I() to generate matrix within data frame)
df1.tr <-matrix(coredata(std.chron1tc), ncol=ncol(coredata(std.chron1tc)),
nrow=nrow(coredata(std.chron1tc)))
pcr1.df <-data.frame(df1.q, A=I(df1.tr))

#pcr2 <-pcr(BayanB~ A,ncomp=2,data=pcr1.df, validation="CV", segments=segs)#this one
#pcrf <-pcr(BayanB~ A,ncomp=2,data=pcr1.df)

#question of if scaling is needed as TR chrons should be mean=1 and std dev=0, but Q is not so would need to
be scaled.
#Use a rolling window for CV, that adds one time period onto a another, starting with half data and moving to
end.
#note, not enough data to perform split calibration (traing and test datasets) confidently.
#using CV will give an idea of model error so a "best" fitting model can be chosen.
#Not truly verifiable though (no model is!) and k-fold won't help either with the short length of record.
#--------------------------------
#alternate method of assessing PC to retain
#look at variance of errors
```

```
#for 4 PC
pcr1.sc <-scores(pcr1)
pcr1.load <-t(loadings(pcr1))
pcr1.rep <-pcr1.sc%*%pcr1.load
pcr1.sqdiff <-(pcr1.df[,1]-pcr1.rep)^2
pcr1.sumsq <-sum(pcr1.sqdiff)
pcr1.var <-pcr1.sumsq/(22*1)

pcr2.sc <-scores(pcr2)
pcr2.load <-t(loadings(pcr2))
pcr2.rep <-pcr2.sc%*%pcr2.load
pcr2.sqdiff <-(pcr1.df[,1]-pcr2.rep)^2
pcr2.sumsq <-sum(pcr2.sqdiff)
pcr2.var <-pcr2.sumsq/(22*1)

pcr3.sc <-scores(pcr3)
pcr3.load <-t(loadings(pcr3))
pcr3.rep <-pcr3.sc%*%pcr3.load
pcr3.sqdiff <-(pcr1.df[,1]-pcr3.rep)^2
pcr3.sumsq <-sum(pcr3.sqdiff)
pcr3.var <-pcr3.sumsq/(22*1)

pcr4.sc <-scores(pcr4)
pcr4.load <-t(loadings(pcr4))
pcr4.rep <-pcr4.sc%*%pcr4.load
pcr4.sqdiff <-(pcr1.df[,1]-pcr4.rep)^2
pcr4.sumsq <-sum(pcr4.sqdiff)
pcr4.var <-pcr4.sumsq/(22*1)

#workflow for PCR modeling

#fit model using CV scheme to identify number of components to include
segs <-list(12:22, 14:22, 16:22, 18:22, 20:22)

#initially include more PCA's to check fits
#pcr1 <-pcr(BayanB~ A,ncomp=4,data=pcr1.df, validation="CV", segments=segs)
#pcr1 <-pcr(BayanT~ A,ncomp=4,data=pcr1.df, validation="CV", segments=segs)
pcr1 <-pcr(ErdK~ A,ncomp=4,data=pcr1.df, validation="CV", segments=segs)
#pcr1 <-pcr(IkhKT~ A,ncomp=4,data=pcr1.df, validation="CV", segments=segs)

#gives percent of variance explained for each PC
summary(pcr1)

#validation results- check first minima of error
plot(RMSEP(pcr1), legendpos="topright")

#prediction plot
plot(pcr1, ncomp=1, asp=1, line=TRUE)

#correlation loadings plot
corrplot(pcr1, comps=1:2, label="names")

#plot of scores and loadings
biplot(pcr1)
```

```
#plot the residuals- so few it's difficult to judge skew etc
pcr1.res <-residuals(pcr1, comp=4)
plot(seq(1:22),pcr1.res[,,1])
abline(h=0)

#pick components to retain and re-fit model to whole time series
pcr1.bb <-pcr(BayanB~A, ncomp=2,data=pcr1.df)
summary(pcr1.bb)

pcr1.bt <-pcr(BayanT~A, ncomp=1,data=pcr1.df)
summary(pcr1.bt)

#check fit criteria
#to get values
RMSEP(pcr1.bt)
MSEP(pcr1.bt)

# for mae, retrieve the fitted values
pcrf.fit <-pcr1.bt$fitted[,,1]
pcrf.or <-pcr1.df[,1]
mae(pcrf.fit, pcrf.or)

#for adjR2
adjr2 <-function(r2,n, p){
  1-((1-r2)*(n-1))/(n-p-1)
}

r2 <-R2(pcr1.bt)$val[2]
n <-22
p <-2

adjr2(r2,n,p)

#retrieve regression coefficients
pcr1.bbcoefi<-coef(pcr1.bb, ncomps=2, intercept=TRUE)
pcr1.bbcoef<-coef(pcr1.bb, ncomps=2)

#Only make reconstructions with "best" models
#apply regression coefficients to predict Q
#test with original values and fitted values
pcr2.coefmat <-matrix(nrow=nrow(pcr2.coef),ncol=1, pcr2.coef)
pcr2.or[1] #original value
pcr2.fit[1]
pcr2.coeffit <-t(pcr2.coefmat)*df1.tr[1,]#will need automation for calculation over reconstructed periods
pcr2.coeffitf <-pcr2.coefi[1]+sum(pcr2.coeffit)#is equal to first fitted value

#must apply coefficients (plus intercept) using actual and lagged time series 1650-1998.
#this will yield reconstructed Q

#------------------------------
#Stepwise regression analysis per reuslts of San Luis work
#want to get correlations for each basin sorted by highest significant and site for entry into model
#initial model set up for each basin is unique and based on correlation analysis
#current year correlations will be used first here, followed by prior year in next order
```

```
#combine to give dataset- unsorted
mlr1 <-data.frame(df1.q,df1.tr)

#step() usses AICp criterion at each step, get AIC with extractAIC(), first element is equiv deg freedonm, second is AIC (wnat minimum)

#-------------------------------------
#BayanB

#full model, all postively correlated, unique non-NA predictors
#MHM, KLL,MHM1, SLB, OGH1,OGH, ZSM, NDB1,ZSM1, ZTG1, JGB,SLB1, KLP, KLP1,JGB1
mlr1.bb <-data.frame(df1.q$BayanB,df1.tr$MHM,df1.tr$KLL,df1.tr$MHM1,
df1.tr$SLB,df1.tr$OGH1,df1.tr$OGH,df1.tr$ZSM,df1.tr$NDB1,df1.tr$ZSM1,df1.tr$ZTG1,df1.tr$JGB,df1.tr$SLB1,
df1.tr$KLP,df1.tr$KLP1,df1.tr$JGB1 )
colnames(mlr1.bb) <-c("BayanB","MHM","KLL","MHM1", "SLB","OGH1","OGH","ZSM", "NDB1", "ZSM1",
"ZTG1","JGB","SLB1","KLP", "KLP1", "JGB1" )
bb.full <-lm(BayanB~., data=mlr1.bb)

#base model
bb.base <-lm(BayanB~MHM, data=mlr1.bb)

#step selection
bb.step <-step(bb.base, scope=list(lower=bb.base, upper=bb.full), direction="forward")

#2 steps
bb.stepr <-lm(BayanB ~ MHM + KLL+OGH+JGB, data=mlr1.bb)

#model selection using leaps from full model
#pick lowest Cp
#then which gives the model (bb.leaps$Cp, bbleaps$which[mod#,])
bb.leaps <-leaps(x=mlr1.bb[2:10], y=mlr1.bb[,1],
method=c("Cp"),names=c("MHM","KLL","SLB","OGH1","ZSM", "NDB1", "ZTG1","JGB","KLP"))

#model 30
bb.leapr <-lm(BayanB ~ KLL + SLB+NDB1+KLP, data=mlr1.bb)

#use leaps from all predictors avaliable
bb.leapsall <-leaps(x=mlr1[,5:22], y=mlr1[,1], method=c("Cp"), names=c( "JGB","KLL","KLP","MHM", "NDB",
"OGH", "SLB", "ZSM", "ZTG","JGB1","KLL1","KLP1","MHM1", "NDB1", "OGH1", "SLB1", "ZSM1", "ZTG1" ))
#JGB,NDB,KLL1,

#bb.leapsall2 <-leaps(x=mlr1[,5:22], y=mlr1[,1], method=c("adjr2"), names=c( "JGB","KLL","KLP","MHM",
"NDB", "OGH", "SLB", "ZSM", "ZTG","JGB1","KLL1","KLP1","MHM1", "NDB1", "OGH1", "SLB1", "ZSM1", "ZTG1"
))
#too many predictors

#model 21
bb.leapa <-lm(BayanB ~ JGB + NDB + KLL1, data=mlr1)

#check VIF
vif(bb.leapa)

#press criterion for each model
sum((bb.leapa$residuals/(1-hatvalues(bb.leapa)))^2)
sum((bb.leapr$residuals/(1-hatvalues(bb.leapr)))^2)
```

```
#try to improve model as a almagamation of methods and reduce vif of predictors
bb.combo1 <-lm(BayanB ~ JGB+NDB + KLL1, data=mlr1)
sum((bb.combo1$residuals/(1-hatvalues(bb.combo1)))^2)

bb.combo2 <-lm(BayanB ~ JGB + NDB+ ZSM1, data=mlr1)
sum((bb.combo2$residuals/(1-hatvalues(bb.combo2)))^2)



#BayanT
#full model- no na/neg predictors
mlr1.bt <-data.frame(df1.q$BayanT,df1.tr$NDB1,df1.tr$OGH,df1.tr$ZTG,df1.tr$NDB,df1.tr$JGB1 )
colnames(mlr1.bt) <-c("BayanT","NDB1","OGH","ZTG","NDB","JGB1")
bt.full <-lm(BayanT~., data=mlr1.bt)

#base model
bt.base <-lm(BayanT~NDB1, data=mlr1.bt)

#step
bt.step <-step(bt.base, scope=list(lower=bt.base, upper=bt.full), direction="forward")

#2 steps
bt.stepr <-lm(BayanT ~ NDB1 + OGH, data=mlr1.bt)

#model selection using leaps from full model
bt.leaps <-leaps(x=mlr1.bt[2:6], y=mlr1.bt[,1],
method=c("Cp"),names=c("NDB1","OGH","ZTG","NDB","JGB1"))
bt.leaps$Cp
bt.leaps$which

#model 6
bt.leapr <-lm(BayanT ~ OGH+JGB1, data=mlr1.bt)

#use leaps from all predictors avaliable
bt.leapsall <-leaps(x=mlr1[,5:22], y=mlr1[,2], method=c("Cp"), names=c( "JGB","KLL","KLP","MHM", "NDB",
"OGH", "SLB", "ZSM", "ZTG","JGB1","KLL1","KLP1","MHM1", "NDB1", "OGH1", "SLB1", "ZSM1", "ZTG1" ))

#model 41
bt.leapa <-lm(BayanT ~ JGB + SLB+JGB1+NDB1+SLB1, data=mlr1)

lm(BayanT ~ MHM + JGB1, data=mlr1)

#check VIF for best 2 models
vif(bt.leapa)

#press criterion for each model
sum((bt.leapa$residuals/(1-hatvalues(bt.leapa)))^2)
sum((bt.full$residuals/(1-hatvalues(bt.full)))^2)

#try to improve model as a almagamation of methods and reduce predictors
bt.combo1 <-lm(BayanT ~ NDB1+SLB1+JGB1, data=mlr1)

summary(bt.combo1)
sum((bt.combo1$residuals/(1-hatvalues(bt.combo1)))^2)
```

```
bt.combo2 <-lm(BayanT ~ OGH+JGB1+SLB1, data=mlr1)
summary(bt.combo2)
sum((bt.combo2$residuals/(1-hatvalues(bt.combo2)))^2)

bt.combo3 <-lm(BayanT ~ZTG+SLB1+JGB1, data=mlr1)
summary(bt.combo3)

sum((bt.combo3$residuals/(1-hatvalues(bt.combo3)))^2)
bt.combo4 <-lm(BayanT ~MHM+JGB1+SLB1, data=mlr1)
summary(bt.combo4)

sum((bt.combo4$residuals/(1-hatvalues(bt.combo4)))^2)

#ErdK
#full model-
mlr1.ek <-
data.frame(df1.q$ErdK,df1.tr$OGH,df1.tr$ZSM,df1.tr$OGH1,df1.tr$SLB,df1.tr$KLP,df1.tr$MHM,df1.tr$JGB1,df
1.tr$KLP1)
colnames(mlr1.ek) <-c("ErdK","OGH","ZSM","OGH1","SLB","KLP","MHM", "JGB1", "KLP1" )
ek.full <-lm(ErdK~., data=mlr1.ek)

#base model
ek.base <-lm(ErdK~OGH , data=mlr1.ek)

#step
ek.step <-step(ek.base, scope=list(lower=ek.base, upper=ek.full), direction="forward")

#1 steps
ek.stepr <-lm(ErdK ~OGH, data=mlr1.ek)

#model selection using leaps from full model
ek.leaps <-leaps(x=mlr1.ek[,2:9], y=mlr1.ek[,1],
method=c("Cp"),names=c("OGH","ZSM","OGH1","SLB","KLP","MHM", "JGB1", "KLP1"))
ek.leaps$Cp
ek.leaps$which[19,]

#model 19
ek.leapr <-lm(ErdK ~ OGH + ZSM + KLP1, data=mlr1.ek)

sum((ek.leapr$residuals/(1-hatvalues(ek.leapr)))^2)

#use leaps from all predictors avaliable
ek.leapsall <-leaps(x=mlr1[,5:22], y=mlr1[,3], method=c("Cp"), names=c( "JGB","KLL","KLP","MHM", "NDB",
"OGH", "SLB", "ZSM", "ZTG","JGB1","KLL1","KLP1","MHM1", "NDB1", "OGH1", "SLB1", "ZSM1", "ZTG1" ))

#model 71
ek.leapa <-lm(ErdK ~ JGB+KLL+MHM+NDB+ZSM+KLL1+MHM1+ZSM1, data=mlr1)

#press criterion for each model
sum((ek.leapa$residuals/(1-hatvalues(ek.leapa)))^2)
sum((ek.step$residuals/(1-hatvalues(ek.step)))^2)

#try to improve model as a almagamation of methods and reduce predictors
ek.combo1 <-lm(ErdK ~ OGH+ZSM+SLB1 , data=mlr1)
summary(ek.combo1)
```

```
#check VIF and press
vif(ek.combo1)
sum((ek.combo1$residuals/(1-hatvalues(ek.combo1)))^2)

#IkhKT
#full model-OGH ZSM, SLB,SLB1,OGH1, KLP,ZSM1, MHM,KLP1, KLL, JGB
mlr1.ik <-
data.frame(df1.q$IkhKT,df1.tr$OGH,df1.tr$ZSM,df1.tr$SLB,df1.tr$SLB1,df1.tr$OGH1,df1.tr$KLP,df1.tr$ZSM1,
df1.tr$MHM,df1.tr$KLP1,df1.tr$KLL,df1.tr$JGB)
colnames(mlr1.ik) <-c("IKhKT","OGH","ZSM","SLB","SLB1","OGH1", "KLP","ZSM1","MHM", "KLP1", "JGB" )
ik.full <-lm(IKhKT ~., data=mlr1.ik)

#base model
ik.base <-lm(IKhKT ~OGH, data=mlr1.ik)

#step
ik.step <-step(ik.base, scope=list(lower=ik.base, upper=ik.full), direction="forward")

#3 steps
ik.stepr <-lm(IKhKT ~ OGH + ZSM+ JGB, data=mlr1.ik)

#model selection using leaps from full model
ik.leaps <-leaps(x=mlr1.ik[2:11], y=mlr1.ik[,1], method=c("Cp"),names=c("OGH","ZSM","SLB","SLB1","OGH1",
"KLP","ZSM1","MHM", "KLP1", "JGB" ))
ik.leaps$Cp
ik.leaps$which

#model 31
ik.leapr <-lm(IKhKT ~ OGH + SLB+ OGH1+MHM, data=mlr1.ik)

#use leaps from all predictors avaliable
ik.leapsall <-leaps(x=mlr1[,5:22], y=mlr1[,4], method=c("Cp"), names=c( "JGB","KLL","KLP","MHM", "NDB",
"OGH", "SLB", "ZSM", "ZTG","JGB1","KLL1","KLP1","MHM1", "NDB1", "OGH1", "SLB1", "ZSM1", "ZTG1" ))

#model 71
ik.leapa <-lm(IkhKT ~ OGH + SLB + JGB1+ KLP1+MHM1+ OGH1+SLB1+ZTG1, data=mlr1)

#check VIF for best 2 models
vif(ik.leapa)
vif(ik.step)

#press criterion for each model
sum((ik.leapa$residuals/(1-hatvalues(ik.leapa)))^2)
sum((ik.step$residuals/(1-hatvalues(ik.step)))^2)

#check std error
ik.leapase <- sqrt(diag(vcov(ik.leapa)))

#try to improve model as a almagamation of methods and reduce predictors
ik.combo1 <-lm(mlr1[,4] ~ OGH + SLB+JGB, data=mlr1)

summary(ik.combo1)
sum((ik.combo1$residuals/(1-hatvalues(ik.combo1)))^2)

ik.combo2 <-lm(mlr1[,4] ~ OGH + SLB+ZTG, data=mlr1)
```

```
summary(ik.combo2)
sum((ik.combo2$residuals/(1-hatvalues(ik.combo2)))^2)
#-------------
#use newest years to test models (include KHU and not)
#test first no KHU because chrons already collated- use orginal common timeperiod
#0.52 3 pred
bb.leapsl <-leaps(x=mlr1[,c(5,7,10,14,16,19)], y=mlr1[,1], method=c("Cp"),names=c("JGB","KLP", "OGH",
"JGB1", "KLP1", "OGH1"))
bb.leapl <-lm(mlr1[,1] ~ JGB +KLP +OGH, data=mlr1 )
summary(bb.leapl)
sum((bb.leapl$residuals/(1-hatvalues(bb.leapl)))^2)

#0.15 1 pred
bt.leapsl <-leaps(x=mlr1[,c(5,7,10,14,16,19)], y=mlr1[,2], method=c("Cp"),names=c("JGB","KLP", "OGH",
"JGB1", "KLP1", "OGH1"))
bt.leapl <-lm(mlr1[,2] ~ OGH, data=mlr1 )
summary(bt.leapl)

#0.47 3 pred
ek.leapsl <-leaps(x=mlr1[,c(5,7,10,14,16,19)], y=mlr1[,3], method=c("Cp"),names=c("JGB","KLP", "OGH",
"JGB1", "KLP1", "OGH1"))
ek.leapl <-lm(mlr1[,3] ~ JGB+ OGH +JGB1, data=mlr1 )
summary(ek.leapl)

#0.67 but 5 predictors!
ik.leapsl <-leaps(x=mlr1[,c(5,7,10,14,16,19)], y=mlr1[,4], method=c("Cp"),names=c("JGB","KLP", "OGH",
"JGB1", "KLP1", "OGH1"))
ik.leaplf <-lm(mlr1[,4] ~ JGB+ OGH +JGB1+KLP1 +OGH1, data=mlr1 )
summary(ik.leaplf)

ik.leapl <-lm(mlr1[,4] ~ OGH +JGB1+KLP1, data=mlr1 )
summary(ik.leapl)
sum((ik.leapl$residuals/(1-hatvalues(ik.leapl)))^2)

#does adding KHU improve the fits? Use improved common time period
#------------------------------------------
#std-
jgb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/JGB_Stdf.crn")
khu.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KHU_Stdf.crn")
kll.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLL_Stdf.crn")
klp.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLP_Stdf.crn")
mhm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/MHM_Stdf.crn")
ndb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/NDB_Stdf.crn")
```

```
ogh.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Final_Crns/Std/OGH_Stdf.crn")
slb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Final_Crns/Std/SLB_Stdf.crn")
zsm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Final_Crns/Std/ZSM_Stdf.crn")
ztg.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Final_Crns/Std/ZTG_Stdf.crn")

#convert each set to zoo series and collate
#std zoo, KHU and no KHU
#KHU
std.chronk <-list(jgb.std[,1], khu.std[,1],kll.std[,1], klp.std[,1], mhm.std[,1], ndb.std[,1], ogh.std[,1], slb.std[,1],
zsm.std[,1], ztg.std[,1])
std.chronrowsk <-list(as.numeric(rownames(jgb.std)),
as.numeric(rownames(khu.std)),as.numeric(rownames(kll.std)), as.numeric(rownames(klp.std)),
as.numeric(rownames(mhm.std)), as.numeric(rownames(ndb.std)), as.numeric(rownames(ogh.std)),
as.numeric(rownames(slb.std)), as.numeric(rownames(zsm.std)), as.numeric(rownames(ztg.std)))
std.chronzallk <-list()

for(i in 1:length(std.chronk)){
 std.ind <-std.chronrowsk[[i]]
 std.chronz <-zoo(std.chronk[[i]], std.chronrowsk[[i]])
 std.chronzallk[[i]] <-std.chronz
}

#matchup listed zoo chrons into one object
std.chronkhu <-
merge(std.chronzallk[[1]],std.chronzallk[[2]],std.chronzallk[[3]],std.chronzallk[[4]],std.chronzallk[[5]],std.chronzallk[[6]],std.chronzallk[[7]],std.chronzallk[[8]],std.chronzallk[[9]], std.chronzallk[[10]])
colnames(std.chronkhu) <-c("JGB", "KHU","KLL", "KLP", "MHM", "NDB","OGH", "SLB", "ZSM", "ZTG")

#truncate to common period for first analysis 1650-2009 and add lagged predictors
std.chron2 <-window(std.chronkhu, start=1650, end=2009)

#add lagged predictors
std.chron2l <-lag(std.chron2, k=1)
colnames(std.chron2l) <-c("JGB1","KHU1", "KLL1", "KLP1", "MHM1", "NDB1","OGH1", "SLB1", "ZSM1",
"ZTG1")

#merge datasets together
std.chron2m <-merge(std.chron2,std.chron2l)

#truncate to remove NAs and create calibration set
#common period becomes 1650-1998 due to lag
std.chron2t <-window(std.chron2m, start=1650, end=2008)

#truncate to model period to match common perod with Q data (1977 (dendroyear)-1998 only 22 years!)
std.chron2tc <-window(std.chron2t, start=1977, end=2008)

#df for analysis
```

```
df2.tr <-data.frame(coredata(std.chron2tc))
#-------------------------------------------
#streamflow
#use transformed, untransformed, retransformed values
#truncate to calibration period
transqt2 <-transqagt[2:33,]

#data for analysis
df2.q <-data.frame(transqt2[,3:6])

#merge into mlrdf
mlr2 <-data.frame(df2.q,df2.tr)
#----
#0.39 2 pred
bb.leapsl2 <-leaps(x=mlr2[,c(5,6,8,11,15:16,18,21)], y=mlr2[,1],
method=c("Cp"),names=c("JGB","KHU","KLP", "OGH", "JGB1", "KHU1","KLP1", "OGH1"))
bb.leapl2 <-lm(mlr2[,1] ~ JGB+KLP+OGH+KLP1, data=mlr2 )
summary(bb.leapl2)

bb.leapl2 <-lm(mlr2[,1] ~ JGB+KLP+OGH, data=mlr2 )
summary(bb.leapl2)

#0.14 1 pred
bt.leapsl2 <-leaps(x=mlr2[,c(5,6,8,11,15:16,18,21)], y=mlr2[,2], method=c("Cp"),names=c("JGB","KHU","KLP",
"OGH", "JGB1", "KHU1","KLP1", "OGH1"))
bt.leapl2 <-lm(mlr2[,2] ~ OGH, data=mlr2 )
summary(bt.leapl2)

#0.53 2 pred
ek.leapsl2 <-leaps(x=mlr2[,c(5,6,8,11,15:16,18,21)], y=mlr2[,3],
method=c("Cp"),names=c("JGB","KHU","KLP", "OGH", "JGB1", "KHU1","KLP1", "OGH1"))
ek.leapl2 <-lm(mlr2[,3] ~ KHU + OGH, data=mlr2 )
summary(ek.leapl2)

sum((ek.leapl2$residuals/(1-hatvalues(ek.leapl2)))^2)

#4 predictors
ik.leapsl2 <-leaps(x=mlr2[,c(5,6,8,11,15:16,18,21)], y=mlr2[,4], method=c("Cp"),names=c("JGB","KHU","KLP",
"OGH", "JGB1", "KHU1","KLP1", "OGH1"))
ik.leapl2 <-lm(mlr2[,4] ~ OGH +JGB1+KLP1+OGH1, data=mlr2 )
summary(ik.leapl2)

sum((ik.leapl2$residuals/(1-hatvalues(ik.leapl2)))^2)

#3 predictors
ik.leapl2 <-lm(mlr2[,4] ~ OGH +JGB1+KLP1, data=mlr2 )
summary(ik.leapl2)

sum((ik.leapl2$residuals/(1-hatvalues(ik.leapl2)))^2)

#2 pred
ik.leapl3 <-lm(mlr2[,4] ~ OGH +KLP1, data=mlr2 )
summary(ik.leapl3)

sum((ik.leapl3$residuals/(1-hatvalues(ik.leapl3)))^2)
```

```
#-----------------
#examine residuals from models and test for normality etc
#models:1977-1998: bb.combo1, bt.combo1, ek.combo1, ik.step, 1977-2008: ek.leapl2, ik.leapl3

bb.res <-bb.combo1$residuals
bt.res <-bt.combo2$residuals
ek.res <-ek.combo1$residuals
ik.res <-ik.combo2$residuals
ek2.res <-ek.leapl2$residuals
ik2.res <-ik.leapl2$residuals
btalt.res <-bt.combo3$residuals
ikalt.res <-ik.combo1$residuals
btalt2.res <-bt.combo4$residuals

#check distribution of residuals
plot(bt.res, ylim=-c(-20,20))
abline(0,0)

#use kolmogorov-smirnov test
#gives max diff between cdf's-what is chance the statitic would be as large or larger than observed?
#if pval is >0.05 H0 cannot be rejected and values are stat similar to normal dist.
ks.test(ek2.res, pnorm)
#Some are not much different (bigger pvals)
#-----
#get fitted values from models
bb.fit <-bb.combo1$fitted
bt.fit <-bt.combo2$fitted
ek.fit <-ek.combo1$fitted
ik.fit <-ik.combo2$fitted
ek2.fit <-ek.leapl2$fitted
ik2.fit <-ik.leapl2$fitted
ik3.fit <-ik.leapl3$fitted

btalt.fit <-bt.combo3$fitted
ikalt.fit <-ik.combo1$fitted
btalt2.fit <-bt.combo4$fitted

#get observed values from models
bb.obs <-mlr1[,1]
bt.obs <-mlr1[,2]
ek.obs <-mlr1[,3]
ik.obs <-mlr1[,4]
ek2.obs <-mlr2[,3]
ik2.obs <-mlr2[,4]
ik3.obs <-mlr2[,4]
btalt.obs <-mlr1[,2]
ikalt.obs <-mlr1[,4]
btalt2.obs <--mlr1[,2]

#use rmse() and mae() on fitted and observed datasets to get calibration stats.

#put in df and make zoo with dates
mod1.ind <-seq(1977, 1998)
mod2.ind <-seq(1977, 2008)
```

```
mod1.z <-zoo(data.frame(bb.fit,bb.obs,bt.fit,bt.obs,ek.fit,ek.obs,ik.fit,ik.obs),mod1.ind)
mod2.z <-zoo(data.frame(ek2.fit,ek2.obs,ik2.fit,ik2.obs, ik3.fit, ik3.obs),mod2.ind)
modalt.z <-zoo(data.frame(btalt.fit, bt.obs), mod1.ind)
modaltbt.z <-zoo(data.frame(btalt2.fit, btalt2.obs), mod1.ind)
modaltik.z <-zoo(data.frame(ikalt.fit, ik.obs), mod1.ind)

write.csv(modalt.z, "Fitted_Observed_BayanTALT.csv")
write.csv(modaltbt.z, "Fitted_Observed_BayanTALT2.csv")
write.csv(modaltik.z, "Fitted_Observed_IkhKTALT.csv")
write.csv(mod1.z, "Fitted_Observed_4Basins.csv")
write.csv(mod2.z, "Fitted_Observed_AltBasins.csv")

#----------------------
#Run cross validation on selected models

#LOOCV for each model

#make df of Q and predictors
bb.df <-data.frame(mlr1$BayanB, mlr1$JGB,mlr1$NDB,mlr1$KLL1)
bt.df <-data.frame(mlr1$BayanT, mlr1$OGH,mlr1$JGB1,mlr1$SLB1)
ek.df <-data.frame(mlr1$ErdK, mlr1$OGH,mlr1$ZSM,mlr1$SLB1)
ik.df <-data.frame(mlr1$IkhKT, mlr1$OGH,mlr1$SLB,mlr1$ZTG)
ek2.df <-data.frame(mlr2$ErdK, mlr2$KHU,mlr2$OGH)
ik2.df <-data.frame(mlr2$IkhKT, mlr2$OGH,mlr2$JGB1,mlr2$KLP1)
ik3.df <-data.frame(mlr2$IkhKT, mlr2$OGH, mlr2$KLP1)

btalt.df <-data.frame(mlr1$BayanT, mlr1$ZTG,mlr1$SLB1, mlr1$JGB1)
btalt2.df <-data.frame(mlr1$BayanT, mlr1$MHM,mlr1$JGB1, mlr1$SLB1)
ikalt.df <-data.frame(mlr1$IkhKT, mlr1$OGH, mlr1$SLB, mlr1$JGB)

#fit model, leaving one observation out each time
#generate errors, GOF stats and average for output
#--------
#BayanB
statsmod.bb <-matrix(ncol=5, nrow=length(bb.df[,1]), NA)
statsfit.bb <-matrix(ncol=(length(bb.df[,1])), nrow=(length(bb.df[,1])-1), NA)
colnames(statsfit.bb) <-paste("LOO", seq(1,ncol(statsfit.bb)), sep="")

for (i in 1:length(bb.df[,1])){
newdf <-bb.df[-i,]
n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
r2.lm <-summary(n.lm)$r.squared
adjr2.lm <-summary(n.lm)$adj.r.squared
se.lm <-summary(n.lm)$sigma
press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
aic.lm <-extractAIC(n.lm)
stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
statsmod.bb[i,] <-stats.lm
statsfit.bb[,i] <-n.lm$fitted
}

statsmodavg.bb <-colMeans(statsmod.bb)
names(statsmodavg.bb) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )
```

```
#match up fitted to observed for each CV round
mod1cv.bbl <-list()

for (i in 1:length(bb.df[,1])){
 newdf <-bb.df[-i,1]
 newmat <-statsfit.bb[,i]
 newdfmat <-cbind(newmat,newdf)
 colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("BayanB", rep(i,1),sep=""))
mod1cv.bbl[[i]] <-newdfmat
}
#------
#BayanT
statsmod.bt <-matrix(ncol=5, nrow=length(bt.df[,1]), NA)
statsfit.bt <-matrix(ncol=(length(bt.df[,1])), nrow=(length(bt.df[,1])-1), NA)
colnames(statsfit.bt) <-paste("LOO", seq(1,ncol(statsfit.bt)), sep="")

for (i in 1:length(bt.df[,1])){
 newdf <-bt.df[-i,]
 n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
 r2.lm <-summary(n.lm)$r.squared
 adjr2.lm <-summary(n.lm)$adj.r.squared
 se.lm <-summary(n.lm)$sigma
 press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
 aic.lm <-extractAIC(n.lm)
 stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
 statsmod.bt[i,] <-stats.lm
 statsfit.bt[,i] <-n.lm$fitted
}

statsmodavg.bt <-colMeans(statsmod.bt)
names(statsmodavg.bt) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod1cv.btl <-list()

for (i in 1:length(bt.df[,1])){
 newdf <-bt.df[-i,1]
 newmat <-statsfit.bt[,i]
 newdfmat <-cbind(newmat,newdf)
 colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("BayanT", rep(i,1),sep=""))
 mod1cv.btl[[i]] <-newdfmat
}

#BT Alt
statsmod.btalt <-matrix(ncol=5, nrow=length(btalt.df[,1]), NA)
statsfit.btalt <-matrix(ncol=(length(btalt.df[,1])), nrow=(length(btalt.df[,1])-1), NA)
colnames(statsfit.btalt) <-paste("LOO", seq(1,ncol(statsfit.btalt)), sep="")

for (i in 1:length(btalt.df[,1])){
 newdf <-btalt.df[-i,]
 n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
 r2.lm <-summary(n.lm)$r.squared
 adjr2.lm <-summary(n.lm)$adj.r.squared
 se.lm <-summary(n.lm)$sigma
 press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
```

```
  aic.lm <-extractAIC(n.lm)
  stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
  statsmod.btalt[i,] <-stats.lm
  statsfit.btalt[,i] <-n.lm$fitted
}

statsmodavg.btalt <-colMeans(statsmod.btalt)
names(statsmodavg.btalt) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod1cv.btlalt <-list()

for (i in 1:length(btalt.df[,1])){
  newdf <-btalt.df[-i,1]
  newmat <-statsfit.btalt[,i]
  newdfmat <-cbind(newmat,newdf)
  colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("BayanT", rep(i,1),sep=""))
  mod1cv.btlalt[[i]] <-newdfmat
}

#BT Alt2
statsmod.btalt2 <-matrix(ncol=5, nrow=length(btalt2.df[,1]), NA)
statsfit.btalt2 <-matrix(ncol=(length(btalt2.df[,1])), nrow=(length(btalt2.df[,1])-1), NA)
colnames(statsfit.btalt2) <-paste("LOO", seq(1,ncol(statsfit.btalt2)), sep="")

for (i in 1:length(btalt2.df[,1])){
  newdf <-btalt2.df[-i,]
  n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
  r2.lm <-summary(n.lm)$r.squared
  adjr2.lm <-summary(n.lm)$adj.r.squared
  se.lm <-summary(n.lm)$sigma
  press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
  aic.lm <-extractAIC(n.lm)
  stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
  statsmod.btalt2[i,] <-stats.lm
  statsfit.btalt2[,i] <-n.lm$fitted
}

statsmodavg.btalt2 <-colMeans(statsmod.btalt2)
names(statsmodavg.btalt2) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod1cv.btlalt2 <-list()

for (i in 1:length(btalt2.df[,1])){
  newdf <-btalt2.df[-i,1]
  newmat <-statsfit.btalt2[,i]
  newdfmat <-cbind(newmat,newdf)
  colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("BayanT", rep(i,1),sep=""))
  mod1cv.btlalt2[[i]] <-newdfmat
}

#------------------------------
#ErdK
statsmod.ek <-matrix(ncol=5, nrow=length(ek.df[,1]), NA)
```

```
statsfit.ek <-matrix(ncol=(length(ek.df[,1])), nrow=(length(ek.df[,1])-1), NA)
colnames(statsfit.ek) <-paste("LOO", seq(1,ncol(statsfit.ek)), sep="")

for (i in 1:length(ek.df[,1])){
 newdf <-ek.df[-i,]
 n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
 r2.lm <-summary(n.lm)$r.squared
 adjr2.lm <-summary(n.lm)$adj.r.squared
 se.lm <-summary(n.lm)$sigma
 press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
 aic.lm <-extractAIC(n.lm)
 stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
 statsmod.ek[i,] <-stats.lm
 statsfit.ek[,i] <-n.lm$fitted
}

statsmodavg.ek <-colMeans(statsmod.ek)
names(statsmodavg.ek) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod1cv.ekl <-list()

for (i in 1:length(ek.df[,1])){
 newdf <-ek.df[-i,1]
 newmat <-statsfit.ek[,i]
 newdfmat <-cbind(newmat,newdf)
 colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("ErdK", rep(i,1),sep=""))
 mod1cv.ekl[[i]] <-newdfmat
}
#---------------------------------
#IkhKT
statsmod.ik <-matrix(ncol=5, nrow=length(ik.df[,1]), NA)
statsfit.ik <-matrix(ncol=(length(ik.df[,1])), nrow=(length(ik.df[,1])-1), NA)
colnames(statsfit.ik) <-paste("LOO", seq(1,ncol(statsfit.ik)), sep="")

for (i in 1:length(ik.df[,1])){
 newdf <-ik.df[-i,]
 n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
 r2.lm <-summary(n.lm)$r.squared
 adjr2.lm <-summary(n.lm)$adj.r.squared
 se.lm <-summary(n.lm)$sigma
 press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
 aic.lm <-extractAIC(n.lm)
 stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
 statsmod.ik[i,] <-stats.lm
 statsfit.ik[,i] <-n.lm$fitted
}

statsmodavg.ik <-colMeans(statsmod.ik)
names(statsmodavg.ik) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod1cv.ikl <-list()

for (i in 1:length(ik.df[,1])){
```

```
  newdf <-ik.df[-i,1]
  newmat <-statsfit.ik[,i]
  newdfmat <-cbind(newmat,newdf)
  colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("IkhKT", rep(i,1),sep=""))
  mod1cv.ikl[[i]] <-newdfmat
}

#IkhKT Alt
statsmod.ikalt <-matrix(ncol=5, nrow=length(ikalt.df[,1]), NA)
statsfit.ikalt <-matrix(ncol=(length(ikalt.df[,1])), nrow=(length(ikalt.df[,1])-1), NA)
colnames(statsfit.ikalt) <-paste("LOO", seq(1,ncol(statsfit.ikalt)), sep="")

for (i in 1:length(ikalt.df[,1])){
  newdf <-ikalt.df[-i,]
  n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3]+newdf[,4], data=newdf)
  r2.lm <-summary(n.lm)$r.squared
  adjr2.lm <-summary(n.lm)$adj.r.squared
  se.lm <-summary(n.lm)$sigma
  press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
  aic.lm <-extractAIC(n.lm)
  stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
  statsmod.ikalt[i,] <-stats.lm
  statsfit.ikalt[,i] <-n.lm$fitted
}

statsmodavg.ikalt <-colMeans(statsmod.ikalt)
names(statsmodavg.ikalt) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod1cv.iklalt <-list()

for (i in 1:length(ikalt.df[,1])){
  newdf <-ikalt.df[-i,1]
  newmat <-statsfit.ikalt[,i]
  newdfmat <-cbind(newmat,newdf)
  colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("IkhKT", rep(i,1),sep=""))
  mod1cv.iklalt[[i]] <-newdfmat
}

#--------------------------------
#ErdK2
statsmod.ek2 <-matrix(ncol=5, nrow=length(ek2.df[,1]), NA)
statsfit.ek2 <-matrix(ncol=(length(ek2.df[,1])), nrow=(length(ek2.df[,1])-1), NA)
colnames(statsfit.ek2) <-paste("LOO", seq(1,ncol(statsfit.ek2)), sep="")

for (i in 1:length(ek2.df[,1])){
  newdf <-ek2.df[-i,]
  n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3], data=newdf)
  r2.lm <-summary(n.lm)$r.squared
  adjr2.lm <-summary(n.lm)$adj.r.squared
  se.lm <-summary(n.lm)$sigma
  press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
  aic.lm <-extractAIC(n.lm)
  stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
  statsmod.ek2[i,] <-stats.lm
```

```
  statsfit.ek2[,i] <-n.lm$fitted
}

statsmodavg.ek2 <-colMeans(statsmod.ek2)
names(statsmodavg.ek2) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod2cv.ek2l <-list()

for (i in 1:length(ek2.df[,1])){
  newdf <-ek2.df[-i,1]
  newmat <-statsfit.ek2[,i]
  newdfmat <-cbind(newmat,newdf)
  colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("ErdK2", rep(i,1),sep=""))
  mod2cv.ek2l[[i]] <-newdfmat
}

#----------------------------------------------
#IkhKT3
statsmod.ik3 <-matrix(ncol=5, nrow=length(ik3.df[,1]), NA)
statsfit.ik3 <-matrix(ncol=(length(ik3.df[,1])), nrow=(length(ik3.df[,1])-1), NA)
colnames(statsfit.ik3) <-paste("LOO", seq(1,ncol(statsfit.ik3)), sep="")

for (i in 1:length(ik3.df[,1])){
  newdf <-ik3.df[-i,]
  n.lm <-lm(newdf[,1]~ newdf[,2] +newdf[,3], data=newdf)
  r2.lm <-summary(n.lm)$r.squared
  adjr2.lm <-summary(n.lm)$adj.r.squared
  se.lm <-summary(n.lm)$sigma
  press.lm <-sum((n.lm$residuals/(1-hatvalues(n.lm)))^2)
  aic.lm <-extractAIC(n.lm)
  stats.lm <-c(r2.lm,adjr2.lm,aic.lm[2],se.lm,press.lm)
  statsmod.ik3[i,] <-stats.lm
  statsfit.ik3[,i] <-n.lm$fitted
}

statsmodavg.ik3 <-colMeans(statsmod.ik3)
names(statsmodavg.ik3) <-c("R2","adjR2","AIC" ,"SE", "PRESS" )

#match up fitted to observed for each CV round
mod2cv.ik3l <-list()

for (i in 1:length(ik3.df[,1])){
  newdf <-ik3.df[-i,1]
  newmat <-statsfit.ik3[,i]
  newdfmat <-cbind(newmat,newdf)
  colnames(newdfmat) <-c(paste("LOOCV", rep(i,1), sep=""),paste("IkhKT3", rep(i,1),sep=""))
  mod2cv.ik3l[[i]] <-newdfmat
}

#-----------------------------------------------------------

#make df for fitted/obs
mod1cv.bb <-data.frame(mod1cv.bbl)
mod1cv.bt <-data.frame(mod1cv.btl)
```

```
mod1cv.ek <-data.frame(mod1cv.ekl)
mod1cv.ik <-data.frame(mod1cv.ikl)
mod2cv.ek2 <-data.frame(mod2cv.ek2l)
mod2cv.ik2 <-data.frame(mod2cv.ik2l)
mod2cv.ik3 <-data.frame(mod2cv.ik3l)

mod1cv.btalt <-data.frame(mod1cv.btlalt)
mod1cv.btalt2 <-data.frame(mod1cv.btlalt2)
mod1cv.ikalt <-data.frame(mod1cv.iklalt)

#df of gof stats
mod1.stats <-data.frame(rbind(statsmodavg.bb,statsmodavg.bt, statsmodavg.ek, statsmodavg.ik))
colnames(mod1.stats) <-c("R2", "adjR2", "AIC", "SE","PRESS")
rownames(mod1.stats) <-c("BayanB", "BayanT", "ErdK", "IkhKT")

mod2.stats <-data.frame(rbind(statsmodavg.ek2, statsmodavg.ik2, statsmodavg.ik3))
colnames(mod2.stats) <-c("R2", "adjR2", "AIC", "SE","PRESS")
rownames(mod2.stats) <-c("ErdK2", "IkhKT2", "IkhKT3")

#---------------------------------------------------
#save results

write.csv(mod1cv.bb, "Fitted_Observed_CV_BayanB.csv")
write.csv(mod1cv.bt, "Fitted_Observed_CV_BayanT.csv")
write.csv(mod1cv.ek, "Fitted_Observed_CV_ErdK.csv")
write.csv(mod1cv.ik, "Fitted_Observed_CV_IkhKT.csv")
write.csv(mod2cv.ek2, "Fitted_Observed_CV_ErdK2.csv")
write.csv(mod2cv.ik2, "Fitted_Observed_CV_IkhKT2.csv")
write.csv(mod2cv.ik3, "Fitted_Observed_CV_IkhKT3.csv")

write.csv(mod1cv.btalt, "Fitted_Observed_CV_BayanTALT.csv")
write.csv(mod1cv.btalt2, "Fitted_Observed_CV_BayanTALT2.csv")
write.csv(mod1cv.ikalt, "Fitted_Observed_CV_IkhKTALT.csv")


write.csv(mod1.stats, "Model_Stats_4_Models.csv")
write.csv(mod2.stats, "Model_Stats_2_Models.csv")

#---------------------------------------------------------
#get regression coefficients from models
#bb.combo1, bt.combo1,ek.combo1,ik.step, ek.leapl2,ik.leapl3
bb.co <-summary(bb.combo1)$coefficients[,1]
bt.co <-summary(bt.combo2)$coefficients[,1]
ek.co <-summary(ek.combo1)$coefficients[,1]
ik.co <-summary(ik.combo2)$coefficients[,1]
ek2.co <-summary(ek.leapl2)$coefficients[,1]
ik2.co <-summary(ik.leapl2)$coefficients[,1]

btalt.co <-summary(bt.combo3)$coefficients[,1]
btalt2.co <-summary(bt.combo4)$coefficients[,1]
ikalt.co <-summary(ik.combo1)$coefficients[,1]

#two predictor models
twopred <-matrix(ncol=3, nrow=1, ek2.co)
rownames(twopred) <-c("ErdK2")
```

```
colnames(twopred) <-c("Intercept", "Pred1", "Pred2")
write.csv(twopred, "Model_Coeffs_TwoPred.csv")

#three predictor models
threepred <-matrix(ncol=4, nrow=8, byrow=TRUE, c(bb.co, bt.co,ek.co, ik.co, ik2.co,btalt.co,btalt2.co,
ikalt.co))
rownames(threepred) <-c("BayanB", "BayanT", "ErdK", "IkhKT", "IkhKT2", "BayanTAlt",
"BayanTAlt2","IkhKTAlt")
colnames(threepred) <-c("Intercept", "Pred1", "Pred2", "Pred3")

write.csv(threepred, "Model_Coeffs_ThreePred.csv")

#------------------------------------------------------
#import model results
mod1o <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_4Basins.csv")
mod2o <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_AltBasins.csv")
mod1o.bb<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_BayanB.csv")
mod1o.bt<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_BayanT.csv")
mod1o.ek<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_ErdK.csv")
mod1o.ik<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_IkhKT.csv")
mod2o.ek <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_ErdK2.csv")
mod2o.ik <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_IkhKT2.csv")
mod2o.ik3 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_IkhKT3.csv")

mod1o.btalt <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_BayanTALT.csv")
mod1o.btalt2 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_BayanTALT2.csv")
mod1o.ikalt <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_CV_IkhKTALT.csv")

#backtrans values for stats calc
mod1 <-mod1o^2
mod2 <-mod2o^2
```

```
mod1.bb <-mod1o.bb^2
mod1.bt<-mod1o.bt^2
mod1.ek<-mod1o.ek^2
mod1.ik<-mod1o.ik^2
mod2.ek <-mod2o.ek^2
mod2.ik <-mod2o.ik ^2
mod2.ik3 <-mod2o.ik3 ^2


mod1.btalt<-mod1o.btalt^2
mod1.btalt2<-mod1o.btalt2^2
mod1.ikalt <-mod1o.ikalt^2

#apply RMSE and MAE stats
rmma1.mat<-matrix(ncol=2,nrow=7, NA)
for (i in 1:7){
rootmean <-rmse(mod1[,i], mod1[, i+1])
meanabs <-mae(mod1[,i], mod1[,i+1])
rmma1.mat[i,]<-c(rootmean,meanabs)
}

rmma1.fin <-data.frame(rmma1.mat[c(1,3,5,7),])
colnames(rmma1.fin) <-c("RMSE", "MAE")
rownames(rmma1.fin) <-c("BayanB","BayanT","ErdK", "IkhKT")

rmma2.mat<-matrix(ncol=2,nrow=5, NA)
for (i in 1:3){
 rootmean <-rmse(mod2[,i], mod2[, i+1])
 meanabs <-mae(mod2[,i], mod2[,i+1])
 rmma2.mat[i,]<-c(rootmean,meanabs)
}
rmma2.fin <-data.frame(rmma2.mat[c(1,3,5),])
colnames(rmma2.fin) <-c("RMSE", "MAE")
rownames(rmma2.fin) <-c( "ErdK2", "IkhKT2","IkhKT3")

#for cv need to run stats then take mean of stats
#BayanB
rmma1.bbmat<-matrix(ncol=2,nrow=ncol(mod1.bb)/2, NA)
for (i in 1:(ncol(mod1.bb)/2)){
 rootmean <-rmse(mod1.bb[,i], mod1.bb[, i+1])
 meanabs <-mae(mod1.bb[,i], mod1.bb[,i+1])
 rmma1.bbmat[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.bbmat) <-c("RMSE", "MAE")
rmma.bb <-colMeans(rmma1.bbmat)

#BayanT
rmma1.btmat<-matrix(ncol=2,nrow=ncol(mod1.bt)/2, NA)
for (i in 1:(ncol(mod1.bt)/2)){
 rootmean <-rmse(mod1.bt[,i], mod1.bt[, i+1])
 meanabs <-mae(mod1.bt[,i], mod1.bt[,i+1])
 rmma1.btmat[i,]<-c(rootmean,meanabs)

}
```

```
colnames(rmma1.btmat) <-c("RMSE", "MAE")
rmma.bt <-colMeans(rmma1.btmat)

#btalt
rmma1.btaltmat<-matrix(ncol=2,nrow=ncol(mod1.btalt)/2, NA)
for (i in 1:(ncol(mod1.btalt)/2)){
  rootmean <-rmse(mod1.btalt[,i], mod1.btalt[, i+1])
  meanabs <-mae(mod1.btalt[,i], mod1.btalt[,i+1])
  rmma1.btaltmat[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.btaltmat) <-c("RMSE", "MAE")
rmma.btalt <-colMeans(rmma1.btaltmat)

#btalt2
rmma1.btaltmat2<-matrix(ncol=2,nrow=ncol(mod1.btalt2)/2, NA)
for (i in 1:(ncol(mod1.btalt2)/2)){
  rootmean <-rmse(mod1.btalt2[,i], mod1.btalt2[, i+1])
  meanabs <-mae(mod1.btalt2[,i], mod1.btalt2[,i+1])
  rmma1.btaltmat2[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.btaltmat2) <-c("RMSE", "MAE")
rmma.btalt2 <-colMeans(rmma1.btaltmat2)

#ErdK
rmma1.ekmat<-matrix(ncol=2,nrow=ncol(mod1.ek)/2, NA)
for (i in 1:(ncol(mod1.ek)/2)){
  rootmean <-rmse(mod1.ek[,i], mod1.ek[, i+1])
  meanabs <-mae(mod1.ek[,i], mod1.ek[,i+1])
  rmma1.ekmat[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.ekmat) <-c("RMSE", "MAE")
rmma.ek <-colMeans(rmma1.ekmat)

#IkhKT
rmma1.ikmat<-matrix(ncol=2,nrow=ncol(mod1.ik)/2, NA)
for (i in 1:(ncol(mod1.ik)/2)){
  rootmean <-rmse(mod1.ik[,i], mod1.ik[, i+1])
  meanabs <-mae(mod1.ik[,i], mod1.ik[,i+1])
  rmma1.ikmat[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.ikmat) <-c("RMSE", "MAE")
rmma.ik <-colMeans(rmma1.ikmat)

#ikalt
rmma1.ikaltmat<-matrix(ncol=2,nrow=ncol(mod1.ikalt)/2, NA)
for (i in 1:(ncol(mod1.ikalt)/2)){
  rootmean <-rmse(mod1.ikalt[,i], mod1.ikalt[, i+1])
  meanabs <-mae(mod1.ikalt[,i], mod1.ikalt[,i+1])
  rmma1.ikaltmat[i,]<-c(rootmean,meanabs)

}
```

```
colnames(rmma1.ikaltmat) <-c("RMSE", "MAE")
rmma.ikalt <-colMeans(rmma1.ikaltmat)


#ErdK2
rmma1.ek2mat<-matrix(ncol=2,nrow=ncol(mod2.ek)/2, NA)
for (i in 1:(ncol(mod2.ek)/2)){
 rootmean <-rmse(mod2.ek[,i], mod2.ek[, i+1])
 meanabs <-mae(mod2.ek[,i], mod2.ek[,i+1])
 rmma1.ek2mat[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.ek2mat) <-c("RMSE", "MAE")
rmma.ek2 <-colMeans(rmma1.ek2mat)

#IkhKT2
rmma1.ik2mat<-matrix(ncol=2,nrow=ncol(mod2.ik)/2, NA)
for (i in 1:(ncol(mod2.ik)/2)){
 rootmean <-rmse(mod2.ik[,i], mod2.ik[, i+1])
 meanabs <-mae(mod2.ik[,i], mod2.ik[,i+1])
 rmma1.ik2mat[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.ik2mat) <-c("RMSE", "MAE")
rmma.ik2 <-colMeans(rmma1.ik2mat)

#IkhKT3
rmma1.ik2mat3<-matrix(ncol=2,nrow=ncol(mod2.ik3)/2, NA)
for (i in 1:(ncol(mod2.ik3)/2)){
 rootmean <-rmse(mod2.ik3[,i], mod2.ik3[, i+1])
 meanabs <-mae(mod2.ik3[,i], mod2.ik3[,i+1])
 rmma1.ik2mat3[i,]<-c(rootmean,meanabs)

}
colnames(rmma1.ik2mat3) <-c("RMSE", "MAE")
rmma.ik23 <-colMeans(rmma1.ik2mat3)

#---------------
#reconstructions

#import observed data
obsfit4 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_4Basins.csv")
obsfit2 <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_AltBasins.csv")

obsaltbt<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_BayanTALT.csv")
obsalt <-obsaltbt[,2]
fitalt <-obsaltbt[,1]
```

```
obsaltbt2<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_BayanTALT.csv")
obsaltbt <-obsaltbt2[,2]
fitaltbt <-obsaltbt2[,1]

obsaltikt<-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Fit_Obs/Fitted_Observed_IkhKTALT.csv")
obsaltik <-obsaltikt[,2]
fitaltik <-obsaltikt[,1]

#extract obs and make zoo
obs4 <-obsfit4[,c(2,4,6,8)]
obs2 <-obsfit2[,c(2,4)]
obs.ind4 <-seq(1977,1998,1 )
obs.ind2 <-seq(1977,2008,1 )
obs4.z <-zoo(obs4,obs.ind4)
obs2.z <-zoo(obs2,obs.ind2)

obsalt.z <-zoo(obsalt, obs.ind4)
obsaltbt.z <-zoo(obsaltbt, obs.ind4)
obsaltik.z <-zoo(obsaltik, obs.ind4)

#import tr chrons
#std
jgb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/JGB_Stdf.crn")
khu.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KHU_Stdf.crn")
kll.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLL_Stdf.crn")
klp.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/KLP_Stdf.crn")
mhm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/MHM_Stdf.crn")
ndb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/NDB_Stdf.crn")
ogh.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/OGH_Stdf.crn")
slb.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/SLB_Stdf.crn")
zsm.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/ZSM_Stdf.crn")
```

```
ztg.std <-
read.crn("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/PQTR_Khangai/Fi
nal_Crns/Std/ZTG_Stdf.crn")

#convert each set to zoo series and collate
#std zoo, KHU and no KHU
#no KHU
std.chron <-list(jgb.std[,1], kll.std[,1], klp.std[,1], mhm.std[,1], ndb.std[,1], ogh.std[,1], slb.std[,1], zsm.std[,1],
ztg.std[,1])
std.chronrows <-list(as.numeric(rownames(jgb.std)), as.numeric(rownames(kll.std)),
as.numeric(rownames(klp.std)), as.numeric(rownames(mhm.std)), as.numeric(rownames(ndb.std)),
as.numeric(rownames(ogh.std)), as.numeric(rownames(slb.std)), as.numeric(rownames(zsm.std)),
as.numeric(rownames(ztg.std)))
std.chronzall <-list()

for(i in 1:length(std.chron)){
  std.ind <-std.chronrows[[i]]
  std.chronz <-zoo(std.chron[[i]], std.chronrows[[i]])
  std.chronzall[[i]] <-std.chronz
}

#matchup listed zoo chrons into one object
std.chronnokhu <-
merge(std.chronzall[[1]],std.chronzall[[2]],std.chronzall[[3]],std.chronzall[[4]],std.chronzall[[5]],std.chronzal
l[[6]],std.chronzall[[7]],std.chronzall[[8]],std.chronzall[[9]])
colnames(std.chronnokhu) <-c("JGB", "KLL", "KLP", "MHM", "NDB","OGH", "SLB", "ZSM", "ZTG")

#truncate to longest chronologies and add lagged predictors
std.chron1 <-window(std.chronnokhu, start=1405, end=1999)

#add lagged predictors
std.chron1l <-lag(std.chron1, k=1)
colnames(std.chron1l) <-c("JGB1", "KLL1", "KLP1", "MHM1", "NDB1","OGH1", "SLB1", "ZSM1", "ZTG1")

#merge datasets together
std.chron1m <-merge(std.chron1,std.chron1l)

#truncate to remove NAs and create sets for recons
std.chron1t <-window(std.chron1m, start=1582, end=1998)

#with KHU
std.chronkhu <-list(jgb.std[,1], khu.std[,1],kll.std[,1], klp.std[,1], mhm.std[,1], ndb.std[,1], ogh.std[,1],
slb.std[,1], zsm.std[,1], ztg.std[,1])
std.chronrowskhu <-list(as.numeric(rownames(jgb.std)),
as.numeric(rownames(khu.std)),as.numeric(rownames(kll.std)), as.numeric(rownames(klp.std)),
as.numeric(rownames(mhm.std)), as.numeric(rownames(ndb.std)), as.numeric(rownames(ogh.std)),
as.numeric(rownames(slb.std)), as.numeric(rownames(zsm.std)), as.numeric(rownames(ztg.std)))
std.chronzallkhu <-list()

for(i in 1:length(std.chronkhu)){
  std.indkhu <-std.chronrowskhu[[i]]
  std.chronzkhu <-zoo(std.chronkhu[[i]], std.chronrowskhu[[i]])

  std.chronzallkhu[[i]] <-std.chronzkhu
}
```

```
#matchup listed zoo chrons into one object
std.chronkhu <-
merge(std.chronzallkhu[[1]],std.chronzallkhu[[2]],std.chronzallkhu[[3]],std.chronzallkhu[[4]],std.chronzallk
hu[[5]],std.chronzallkhu[[6]],std.chronzallkhu[[7]],std.chronzallkhu[[8]],std.chronzallkhu[[9]],std.chronzallk
hu[[10]])
colnames(std.chronkhu) <-c("JGB", "KHU","KLL", "KLP", "MHM", "NDB","OGH", "SLB", "ZSM", "ZTG")

#truncate to common period for first analysis 1650-1999 and add lagged predictors
std.chron2 <-window(std.chronkhu, start=1650, end=2009)

#add lagged predictors
std.chron2l <-lag(std.chron2, k=1)
colnames(std.chron2l) <-c("JGB1", "KHU1","KLL1", "KLP1", "MHM1", "NDB1","OGH1", "SLB1", "ZSM1",
"ZTG1")

#merge datasets together
std.chron2m <-merge(std.chron2,std.chron2l)

#truncate to remove NAs and create calibration set
std.chron2t <-window(std.chron2m, start=1582, end=2008)

#import reconst coeff
threepred <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Model_Coeffs_ThreePred.csv")
twopred <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Model_Coeffs_TwoPred.csv")

#make time series for each model
std.chron1bb <-window(std.chron1t[,c(1,5,11)], start=1650, end=1998)
std.chron1bt <-window(std.chron1t[,c(6,10,16)], start=1650, end=1998)
std.chron1ek <-window(std.chron1t[,c(6,8,16)], start=1582, end=1998)
std.chron1ik <-window(std.chron1t[,c(6,7,9)], start=1639, end=1998)
std.chron2ek2 <-window(std.chron2t[,c(2,7)], start=1829, end=2008)
std.chron2ik2 <-window(std.chron2t[,c(7,11,14)], start=1650, end=2008)
std.chron1btalt <-window(std.chron1t[,c(9,16,10)], start=1650,end=1998)
std.chron1btalt2 <-window(std.chron1t[,c(4,10,16)], start=1650,end=1998)
std.chron1ikalt <-window(std.chron1t[,c(6,7,1)], start=1650,end=1998)

#apply regression coefs
#BayanB
#create df for estimators

bb.reg <-
threepred[1,2]+(threepred[1,3]*std.chron1bb[,1])+(threepred[1,4]*std.chron1bb[,2])+(threepred[1,5]*std.c
hron1bb[,3])

#mean Q recon
bb.regmean <-mean(bb.reg)


plot(bb.reg)
lines(obs4.z[,1], col="red")
abline(h=bb.regmean, col="gray")
```

```
#BackTransformed
bb.regt <-
(threepred[1,2]+(threepred[1,3]*std.chron1bb[,1])+(threepred[1,4]*std.chron1bb[,2])+(threepred[1,5]*std.c
hron1bb[,3]))^2

#mean Q recon
bb.regmean <-mean(bb.regt)

plot(bb.regt)
lines((obs4.z[,1])^2, col="red")
abline(h=bb.regmean, col="gray50")

#BayanT
#create df for estimators
bt.reg <-
threepred[2,2]+(threepred[2,3]*std.chron1bt[,1])+(threepred[2,4]*std.chron1bt[,2])+(threepred[2,5]*std.ch
ron1bt[,3])

#mean Q recon
bt.regmean <-mean(bt.reg)

plot(bt.reg, xlim=c(1650,1998))
lines(obs4.z[,2], col="red")
abline(h=bt.regmean, col="gray")

#BackTransformed
bt.regt <-
(threepred[2,2]+(threepred[2,3]*std.chron1bt[,1])+(threepred[2,4]*std.chron1bt[,2])+(threepred[2,5]*std.c
hron1bt[,3]))^2

#mean Q recon
bt.regmean <-mean(bt.regt)

plot(bt.regt, xlim=c(1650,1998))
lines((obs4.z[,2])^2, col="red")
abline(h=bt.regmean, col="gray50")

#BayanT Alt
#create df for estimators
btalt.reg <-
threepred[6,2]+(threepred[6,3]*std.chron1btalt[,1])+(threepred[6,4]*std.chron1btalt[,2])+(threepred[6,5]*s
td.chron1btalt[,3])

#mean Q recon
btalt.regmean <-mean(btalt.reg)

plot(btalt.reg, xlim=c(1650, 1998))
lines(obsalt.z, col="red")
abline(h=btalt.regmean, col="gray")


#BackTransformed
btalt.regt <-
(threepred[6,2]+(threepred[6,3]*std.chron1btalt[,1])+(threepred[6,4]*std.chron1btalt[,2])+(threepred[6,5]*
std.chron1btalt[,3]))^2
```

```
#mean Q recon
btalt.regmean <-mean(btalt.regt)

plot(btalt.regt, xlim=c(1650, 1998))
lines((obsalt.z)^2, col="red")
abline(h=btalt.regmean, col="gray50")

#compare bt and btalt
bt.comp <-merge(bt.regt, btalt.regt)

plot(bt.regt)
lines(btalt.regt, col="blue")
lines(btalt2.regt, col="red")

#BayanT Alt2
#create df for estimators
btalt2.reg <-
threepred[7,2]+(threepred[7,3]*std.chron1btalt2[,1])+(threepred[7,4]*std.chron1btalt2[,2])+(threepred[7,5]
*std.chron1btalt2[,3])

#mean Q recon
btalt2.regmean <-mean(btalt2.reg)

plot(btalt2.reg, xlim=c(1650, 1998))
lines(obsalt.z, col="red")
abline(h=btalt2.regmean, col="gray")

#BackTransformed
btalt2.regt <-
(threepred[7,2]+(threepred[7,3]*std.chron1btalt2[,1])+(threepred[7,4]*std.chron1btalt2[,2])+(threepred[7,
5]*std.chron1btalt2[,3]))^2

#mean Q recon
btalt2.regmean <-mean(btalt2.regt)

plot(btalt2.regt, xlim=c(1650, 1998))
lines((obsalt.z)^2, col="red")
abline(h=btalt2.regmean, col="gray50")

#ErdK
#create df for estimators
ek.reg <-
threepred[3,2]+(threepred[3,3]*std.chron1ek[,1])+(threepred[3,4]*std.chron1ek[,2])+(threepred[3,5]*std.ch
ron1ek[,3])

#mean Q recon
ek.regmean <-mean(ek.reg)

plot(ek.reg)
lines(obs4.z[,3], col="red")
abline(h=ek.regmean, col="gray")

#BackTransformed
ek.regt <-
(threepred[3,2]+(threepred[3,3]*std.chron1ek[,1])+(threepred[3,4]*std.chron1ek[,2])+(threepred[3,5]*std.c
```

```
hron1ek[,3]))^2

#mean Q recon
ek.regmean <-mean(ek.regt)

plot(ek.regt)
lines((obs4.z[,3])^2, col="red")
abline(h=ek.regmean, col="gray")

#IkhKT
#create df for estimators
ik.reg <-
threepred[4,2]+(threepred[4,3]*std.chron1ik[,1])+(threepred[4,4]*std.chron1ik[,2])+(threepred[4,5]*std.chr
on1ik[,3])

#mean Q recon
ik.regmean <-mean(ik.reg)

plot(ik.reg)
lines(obs4.z[,4], col="red")
abline(h=ik.regmean, col="gray")

#BackTransformed
ik.regt <-
(threepred[4,2]+(threepred[4,3]*std.chron1ik[,1])+(threepred[4,4]*std.chron1ik[,2])+(threepred[4,5]*std.ch
ron1ik[,3]))^2

#mean Q recon
ik.regmean <-mean(ik.regt)

plot(ik.regt)
lines((obs4.z[,4])^2, col="red")
abline(h=ik.regmean, col="gray")


#IkhKT Alt
#create df for estimators
ikalt.reg <-
threepred[8,2]+(threepred[8,3]*std.chron1ikalt[,1])+(threepred[8,4]*std.chron1ikalt[,2])+(threepred[8,5]*st
d.chron1ikalt[,3])

#mean Q recon
ikalt.regmean <-mean(ikalt.reg)

plot(ikalt.reg)
lines(obsaltik.z, col="red")
abline(h=ikalt.regmean, col="gray")

#compare ik and ikalt
ik.comp <-merge(ik.reg, ikalt.reg)
plot(ik.reg)
lines(ikalt.reg, col="blue")
lines(obsaltik.z, col="red")

#BackTransformed
```

```
ikalt.regt <-
(threepred[8,2]+(threepred[8,3]*std.chron1ikalt[,1])+(threepred[8,4]*std.chron1ikalt[,2])+(threepred[8,5]*s
td.chron1ikalt[,3]))^2

#mean Q recon
ikalt.regmean <-mean(ikalt.regt)

plot(ikalt.regt)
lines((obsaltik.z)^2, col="red")
abline(h=ikalt.regmean, col="gray")

plot(ik.regt)
lines(ikalt.regt, col="blue")

#ErdK2
#create df for estimators
ek2.reg <-twopred[1,2]+(twopred[1,3]*std.chron2ek2[,1])+(twopred[1,4]*std.chron2ek2[,2])

#mean Q recon
ek2.regmean <-mean(ek2.reg)

plot(ek2.reg)
lines(obs2.z[,1], col="red")
abline(h=ek2.regmean, col="gray")

#BackTransformed
ek2.regt <-(twopred[1,2]+(twopred[1,3]*std.chron2ek2[,1])+(twopred[1,4]*std.chron2ek2[,2]))^2

#mean Q recon
ek2.regmean <-mean(ek2.regt)

plot(ek2.regt)
lines((obs2.z[,1])^2, col="red")
abline(h=ek2.regmean, col="gray")

plot(ek.regt)
lines(ek2.regt, col="blue")

#IkhKT2
#create df for estimators
ik2.reg <-threepred[5,2]+(threepred[5,3]*std.chron2ik2[,1])+(threepred[5,4]*std.chron2ik2[,2])

#mean Q recon
ik2.regmean <-mean(ik2.reg)

plot(ik2.reg)
lines(obs2.z[,2], col="red")
abline(h=ik2.regmean, col="gray")


#BackTransformed
ik2.regt <-(threepred[5,2]+(threepred[5,3]*std.chron2ik2[,1])+(threepred[5,4]*std.chron2ik2[,2]))^2

#mean Q recon
ik2.regmean <-mean(ik2.regt)
```

```
plot(ik2.regt)
lines((obs2.z[,2])^2, col="red")
abline(h=ik2.regmean, col="gray")

plot(ik.regt)
plot(ik2.regt, col="blue")#what's wrong with this model?
lines(ikalt.regt, col="red")
lines(ik.regt)
lines(ik3.regt, col="green")

#check results of final model (leaps 3 IkhKT)
#IkhKT3
#BackTransformed
ik3.regt <-(4.154847+(8.140614*std.chron2t[,7])+(2.235638*std.chron2t[,14]))^2

#mean Q recon
ik3.regmean <-mean(ik3.regt)

plot(ik3.regt)
lines((obs2.z[,2])^2, col="red")
abline(h=ik2.regmean, col="gray")

#export final reconstruction values (merge df)
basin.recon <-merge(bb.reg, bt.reg, ek.reg, ik.reg, ek2.reg, ik2.reg, btalt.reg,btalt2.reg, ikalt.reg)
basin.obs <-merge(obs4.z, obs2.z)

write.csv(basin.recon, "Basin_Reconst.csv")
write.csv(basin.obs, "Basin_Obs.csv")

#export backtransformed values
basin.recont <-merge(bb.regt, bt.regt, ek.regt, ik.regt, ek2.regt, ik2.regt,ik3.regt, btalt.regt,btalt2.regt,
ikalt.regt)
basin.obst <-merge(obs4.z, obs2.z,obsalt.z , obsaltik.z)
basin.obstt <-basin.obst^2
write.csv(basin.recont, "Basin_Reconst_BackTrans.csv")
write.csv(basin.obstt, "Basin_Obs_BackTrans.csv")

#------------------------------------
#run this section seperately in case of overlapping variable names

#import final reconstruction values for plotting
basin.recono <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Basin_Reconst_Backtrans.csv")
basin.obso <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Basin_Obs_Backtrans.csv")

#make zoo objects
basin.reconind <-seq(1582,2008,1)
basin.recon <-zoo(basin.recono, basin.reconind)

basin.obsind <-seq(1977,2008,1)
basin.obs <-zoo(basin.obso, basin.obsind)
```

```
#plotting reconstruction results

#break out for different plots
#four basins
four.recon <-basin.recon[,c(1:4)]
four.recont <-window(four.recon, start=1582, end=1998)
four.recondf <-data.frame(coredata(four.recont))
four.recondf$Year <-index(four.recont)
colnames(four.recondf) <-c("Baidrag R. at Bayanburd","Tuin R. at Bayankhongor", "Khanui R. at
Erdenemandal", "Khoid Tamir R. at Ikhtamir","Year")

four.basin <-melt(four.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))
four.basin$Basin <-factor(four.basin$Basin, levels=c("Khanui R. at Erdenemandal", "Khoid Tamir R. at
Ikhtamir","Baidrag R. at Bayanburd","Tuin R. at Bayankhongor"))

#get obs
four.obs <-basin.obs[,c(1:4)]
four.obstime <-merge(four.obs, four.recon)
four.obst <-window(four.obstime[,1:4], start=1582, end=1998)
four.obsdf <-data.frame(coredata(four.obst))
four.obsdf$Year <-index(four.obst)
colnames(four.obsdf) <-c("Baidrag R. at Bayanburd","Tuin R. at Bayankhongor", "Khanui R. at
Erdenemandal", "Khoid Tamir R. at Ikhtamir","Year")

four.basinobs <-melt(four.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))

#make rolling mean for plotting
#20 yr
basin.roll <-rollapply(four.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
four.basinmeandf<-data.frame(coredata(basin.roll))
four.basinmean <-melt(four.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

four.basin$RollMean <-four.basinmean$Mean

#add mean for each recon
four.basin$Mean <-c(rep(NA, 68),rep(417.9, 349),rep(NA, 68),rep(85.4, 349), rep(162.8, 417), rep(NA, 57),
rep(223.0, 360))

#add obs values
four.basin$MCMObs <-four.basinobs$MCMObs

#add point value for Khanui 1810.
#four.basin$Point <-c(rep(NA, 1080), 500, rep(NA, 611))

#make all into mm values for plotting
#convert flows to mm for comparison across basins
#basin size conversion to mm from m/sec
#(((xxm^3/sec*86400sec)/1,000,000m^2)*1000mm)/xxxkm^2
#basin sizes at gages:Baidrag 1622 km^2, Bayanburd 5887 km^2, Bayankhongor 2436 km^2, Bogd 7564
km^2, Erdenenmandal 5799 km^2, Ikhtamir 1963 km^2
#basin <-function(x,y) {
#  ((((x*86400)/1000000)*1000)/y)
#}
```

```
#divide each volume by an area for each basin given in million square meters
#area in km^2
#baid.b <-1622
#bayanb.b <-5887
#bayank.t <- 2436
#bogd.t <-7564
#erd.k <-5799
#ikh.kt <- 1963

#area in km^2*1000000= sq M/1000000= Million SQ Meters so each km^2 is equal to MSQM

#divide each MCM by the area of each basin to get MM then multiply by
four.basin$Area <-c(rep(5887, 417),rep(2436, 417), rep(5799, 417), rep(1963, 417))

four.basinmcm <-four.basin$MCM
four.basinMM <-four.basinmcm/four.basin$Area
four.basinmm <-four.basinMM*1000
four.basin$mm <-four.basinmm

#add mean for mm
four.basin$meanmm <-c(rep(NA, 68),rep(71.0, 349),rep(NA, 68), rep(35.0, 349), rep(28.1, 417), rep(NA, 57),
rep(113.6, 360))

#create rolling mean for mm data
#divide data to basins
four.basinmmb <-
data.frame(cbind(four.basinmm[1:417],four.basinmm[418:834],four.basinmm[835:1251],four.basinmm[125
2:1668]))
four.basinrm <-rollapply(four.basinmmb, 5, fill=NA, mean)
#stack to vector
four.basin$rollmm <-c(four.basinrm[,1], four.basinrm[,2], four.basinrm[,3], four.basinrm[,4])

#for observed
four.basin$mmobs <-(four.basin$MCMObs/four.basin$Area)*1000

#alt basins
#BayanT Alt1 and Alt2 (have similar means even tho slightly different model stats)
alt1.recon <-basin.recon[,c(2,9)]
alt1.recont <-window(alt1.recon, start=1650, end=1998)
alt1.recondf <-data.frame(coredata(alt1.recont))
alt1.recondf$Year <-index(alt1.recont)
colnames(alt1.recondf) <-c("Tuin R. at Bayankhongor Model 1", "Tuin R. at Bayankhongor Model 2", "Year")

alt1.basin <-melt(alt1.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))

#get obs
alt1.obs <-basin.obs[,c(2,2)]
alt1.obso <-merge(alt1.obs, alt1.recon)
alt1.obst <-window(alt1.obso[,1:2], start=1650, end=1998)
alt1.obsdf <-data.frame(coredata(alt1.obst))
alt1.obsdf$Year <-index(alt1.obst)
colnames(alt1.obsdf) <-c("Tuin R. at Bayankhongor Model 1", "Tuin R. at Bayankhongor Model 2","Year")

alt1.basinobs <-melt(alt1.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))
```

```
#make rolling mean for plotting
#20 yr
alt1.roll <-rollapply(alt1.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
alt1.basinmeandf<-data.frame(coredata(alt1.roll))
alt1.basinmean <-melt(alt1.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

alt1.basin$RollMean <-alt1.basinmean$Mean

#add mean for each recon
alt1.basin$Mean <-c(rep(85.4, 349),rep(85.0, 349))

#add obs values
alt1.basin$MCMObs <-alt1.basinobs$MCMObs

#Khoid Tamir Alt2
alt2.recon <-basin.recon[,c(4,10)]
alt2.recont <-window(alt2.recon, start=1639, end=1998)
alt2.recondf <-data.frame(coredata(alt2.recont))
alt2.recondf$Year <-index(alt2.recont)
colnames(alt2.recondf) <-c("Khoid Tamir R. at Ikhtamir Model 1", "Khoid Tamir R. at Ikhtamir Model 2",
"Year")

alt2.basin <-melt(alt2.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))

#get obs
alt2.obs <-basin.obs[,c(4,4)]
alt2.obso <-merge(alt2.obs, alt2.recon)
alt2.obst <-window(alt2.obso[,1:2], start=1639, end=1998)
alt2.obsdf <-data.frame(coredata(alt2.obst))
alt2.obsdf$Year <-index(alt2.obst)
colnames(alt2.obsdf) <-c("Khoid Tamir R. at Ikhtamir Model 1", "Khoid Tamir R. at Ikhtamir Model 2", "Year")

alt2.basinobs <-melt(alt2.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))


#make rolling mean for plotting
#20 yr
alt2.roll <-rollapply(alt2.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
alt2.basinmeandf<-data.frame(coredata(alt2.roll))
alt2.basinmean <-melt(alt2.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

alt2.basin$RollMean <-alt2.basinmean$Mean

#add mean for each recon
alt2.basin$Mean <-c(rep(223.0, 360),rep(NA,11),rep(215.5, 349))

#add obs values
alt2.basin$MCMObs <-alt2.basinobs$MCMObs

#short and long truncated to short
#ErdK2
```

```
two1.recon <-basin.recon[,c(3,5)]
two1.recont <-window(two1.recon, start=1829, end=2008)
two1.recondf <-data.frame(coredata(two1.recont))
two1.recondf$Year <-index(two1.recont)
colnames(two1.recondf) <-c( "Khanui R. at Erdenemandal Model 1",  "Khanui R. at Erdenemandal Model
2","Year")

two1.basin <-melt(two1.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))

#get obs
two1.obs <-basin.obs[,c(3,5)]
two1.obszm <-merge(two1.obs,basin.recon)
two1.obst <-window(two1.obszm[,1:2], start=1829, end=2008)
two1.obsdf <-data.frame(coredata(two1.obst))
two1.obsdf$Year <-index(two1.obst)
colnames(two1.obsdf) <-c("Khanui R. at Erdenemandal Model 1",  "Khanui R. at Erdenemandal Model 2",
"Year")

two1.basinobs <-melt(two1.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))

#make rolling mean for plotting
#20 yr
two1.roll <-rollapply(two1.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
two1.basinmeandf<-data.frame(coredata(two1.roll))
two1.basinmean <-melt(two1.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

two1.basin$RollMean <-two1.basinmean$Mean

#add mean for each recon
two1.basin$Mean <-c(rep(162.8, 170), rep(NA, 10),rep(175.2, 180))

#add obs values
two1.basin$MCMObs <-two1.basinobs$MCMObs

#IkhKT2
two2.recon <-basin.recon[,c(4,7)]
two2.recont <-window(two2.recon, start=1582, end=2008)
two2.recondf <-data.frame(coredata(two2.recont))
two2.recondf$Year <-index(two2.recont)
colnames(two2.recondf) <-c("Khoid Tamir R. at Ikhtamir Model 1","Khoid Tamir R. at Ikhtamir Model
2","Year")

two2.basin <-melt(two2.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))

#get obs

two2.obszm <-merge(basin.obs[,c(4,6)],basin.recon)

two2.obs <-two2.obszm[,1:2]
two2.obst <-window(two2.obs, start=1582, end=2008)
two2.obsdf <-data.frame(coredata(two2.obst))
two2.obsdf$Year <-index(two2.obst)
```

```
colnames(two2.obsdf) <-c("Khoid Tamir R. at Ikhtamir Model 1", "Khoid Tamir R. at Ikhtamir Model 2",
"Year")

two2.basinobs <-melt(two2.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))

#make rolling mean for plotting
#20 yr
two2.roll <-rollapply(two2.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
two2.basinmeandf<-data.frame(coredata(two2.roll))
two2.basinmean <-melt(two2.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

two2.basin$RollMean <-two2.basinmean$Mean

#add mean for each recon
two2.basin$Mean <-c(rep(NA, 57),rep(223.0,360 ),rep(NA,21),rep(207.8, 416))

#add obs values
two2.basin$MCMObs <-two2.basinobs$MCMObs

#plot
#all basins on same graph?

#individual graphs
fourbasin <- ggplot(four.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1600, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+
  scale_x_continuous(breaks=c(seq(1575,2000, by=25)))+theme_bw()+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
    axis.text.y=element_text(size=12),axis.title.y=element_text(size=14,
face="bold"),strip.text.y=element_text(size=10))

fourbasinlog <- ggplot(four.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1600, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+geom_point(aes(x=Year, y=Point), pch="*", size=10)+
  facet_grid(Basin~.)+scale_y_log10(limits=c(100,10000), breaks=c(10,100,1000,1000,
10000))+annotation_logticks(sides="l")+
  scale_x_continuous(breaks=c(seq(1575,2000, by=25)))+theme_bw()+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
      axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
      strip.text.y=element_text(size=10), panel.grid.minor.y =element_blank(),panel.grid.minor.x
=element_blank())

fourbasinmm <- ggplot(four.basin, aes(x=Year, y=mm)) +geom_vline(xintercept=seq(1600, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=rollmm), colour="red")+geom_line(aes(x=Year, y=meanmm),
colour="grey50")+
  geom_line(aes(x=Year, y=mmobs), colour="blue")+facet_grid(Basin~.)+
  scale_x_continuous(breaks=c(seq(1575,2000, by=25)))+theme_bw()+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
```

```
   axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
   strip.text.y=element_text(size=10),panel.grid.minor.x =element_blank())

#Alt models
alt1bt <- ggplot(alt1.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
 geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
 geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
 scale_x_continuous(breaks=c(seq(1650,2000,
by=25)))+scale_y_continuous(limits=c(0,175),breaks=c(seq(25,175, by=50)))+
 theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
    axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
    strip.text.y=element_text(size=14),panel.grid.minor.x =element_blank())

alt2 <- ggplot(alt2.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
 geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
 geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
 scale_x_continuous(breaks=c(seq(1625,2000, by=25)))+
 theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
     axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
     strip.text.y=element_text(size=14), panel.grid.minor.x =element_blank())

#longer models
two1 <- ggplot(two1.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1900, 2000, by=100),
colour="grey80")+
 geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
 geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
 scale_x_continuous(breaks=c(seq(1825,2000, by=25)))+
 theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
     axis.text.y=element_text(size=12),axis.title.y=element_text(size=14,
face="bold"),strip.text.y=element_text(size=14))

two2 <- ggplot(two2.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1600, 2000, by=100),
colour="grey80")+
 geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
 geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
 scale_x_continuous(breaks=c(seq(1575,2000, by=25)))+
 theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
     axis.text.y=element_text(size=12),axis.title.y=element_text(size=14,
face="bold"),strip.text.y=element_text(size=14))

#make last plot into model #3 standing alone for comparison
#IkhKT2 Full
two22.recon <-basin.recon[,7]
two22.recont <-window(two22.recon, start=1582, end=2008)
two22.recondf <-data.frame(coredata(two22.recont))
two22.recondf$Year <-index(two22.recont)
colnames(two22.recondf) <-c("Khoid Tamir R. at Ikhtamir Model 3","Year")

two22.basin <-melt(two22.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))
```

```
#get obs

two22.obszm <-merge(basin.obs[,6],basin.recon)

two22.obs <-two22.obszm[,1]
two22.obst <-window(two22.obs, start=1582, end=2008)
two22.obsdf <-data.frame(coredata(two22.obst))
two22.obsdf$Year <-index(two22.obst)
colnames(two22.obsdf) <-c("Khoid Tamir R. at Ikhtamir Model 3", "Year")

two22.basinobs <-melt(two22.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))

#make rolling mean for plotting
#20 yr
two22.roll <-rollapply(two22.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
two22.basinmeandf<-data.frame(coredata(two22.roll))
two22.basinmean <-melt(two22.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

two22.basin$RollMean <-two22.basinmean$Mean

#add mean for each recon
two22.basin$Mean <-rep(207.8, 427)

#add obs values
two22.basin$MCMObs <-two22.basinobs$MCMObs

two2 <- ggplot(two22.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1600, 2000, by=100),
colour="grey80")+
 geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
 geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
 scale_x_continuous(breaks=c(seq(1575,2000, by=25)))+
 theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
      axis.text.y=element_text(size=12),axis.title.y=element_text(size=14,
face="bold"),strip.text.y=element_text(size=14))

#IkhKT2 1800-present
two23.recon <-basin.recon[,7]
two23.recont <-window(two23.recon, start=1775, end=2008)
two23.recondf <-data.frame(coredata(two23.recont))
two23.recondf$Year <-index(two23.recont)
colnames(two23.recondf) <-c("Khoid Tamir R. at Ikhtamir Model 3","Year")

two23.basin <-melt(two23.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))

#get obs
two23.obszm <-merge(basin.obs[,6],basin.recon)

two23.obs <-two23.obszm[,1]
two23.obst <-window(two23.obs, start=1775, end=2008)
two23.obsdf <-data.frame(coredata(two23.obst))
two23.obsdf$Year <-index(two23.obst)
colnames(two23.obsdf) <-c("Khoid Tamir R. at Ikhtamir Model 3", "Year")
```

```
two23.basinobs <-melt(two23.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))

#make rolling mean for plotting
#20 yr
two23.roll <-rollapply(two23.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
two23.basinmeandf<-data.frame(coredata(two23.roll))
two23.basinmean <-melt(two23.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

two23.basin$RollMean <-two23.basinmean$Mean

#add mean for each recon
two23.basin$Mean <-rep(207.8, 234)

#add obs values
two23.basin$MCMObs <-two23.basinobs$MCMObs

two3 <- ggplot(two23.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1800, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
  scale_x_continuous(breaks=c(seq(1775,2000, by=25)))+scale_y_continuous(limits=c(0,570),
breaks=c(seq(0,500, by=100)))+

theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
      axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
      strip.text.y=element_text(size=14), panel.grid.minor.x =element_blank())

#------------------------------------
#calculate top 10 wet/dry periods using non-overlapping 5 year periods.
#using reconstructed values for each model.
#four.recont, alt1.recont[,2], alt2.recont[,2], two1.recont[,2], two2.recont[,2]
#5-year means for each basin main model and for longer model for khoid tamir basin
four.roll5 <-data.frame(rollapply(four.recont,5 , mean, fill=NA, by=5))

#divide to each basin and rank as a df to keep years
four.roll5$Year <-index(four.recont)

bb.roll5 <-four.roll5[,c(1,5)]
bt.roll5 <-four.roll5[,c(2,5)]
ek.roll5 <-four.roll5[,c(3,5)]
ik.roll5 <-four.roll5[,c(4,5)]

#sort by value
#dry
bb.dry <-bb.roll5[order(bb.roll5[,1]),]
bt.dry <-bt.roll5[order(bt.roll5[,1]),]
ek.dry <-ek.roll5[order(ek.roll5[,1]),]
ik.dry <-ik.roll5[order(ik.roll5[,1]),]

#collate values
basin.dry <-data.frame(c(bb.dry[1:2],bt.dry[,1:2],ek.dry[,1:2],ik.dry[,1:2]))
colnames(basin.dry) <-c("BayanB", "BBYr", "BayanT", "BTYr", "Erdk", "EKYr", "IkhKT", "IKYr")
```

```
write.csv(basin.dry, "Basin_Dry_5Yrs.csv")

#wet
bb.wet <-bb.roll5[order(-bb.roll5[,1]),]
bt.wet <-bt.roll5[order(-bt.roll5[,1]),]
ek.wet <-ek.roll5[order(-ek.roll5[,1]),]
ik.wet <-ik.roll5[order(-ik.roll5[,1]),]

basin.wet <-data.frame(c(bb.wet[,1:2],bt.wet[,1:2],ek.wet[,1:2],ik.wet[,1:2]))
colnames(basin.wet) <-c("BayanB", "BBYr", "BayanT", "BTYr", "Erdk", "EKYr", "IkhKT", "IKYr")
write.csv(basin.wet, "Basin_Wet_5Yrs.csv")

#examine BayanB model for extending streamflow (same declines?)
#predictors: JGB, KLP, OGH,
#vif:  1.1780, 1.0598, 1.1859
bb.obs <-sqrt(basin.obs[,1])
bb.obst <-window(bb.obs, start=1977, end=1998)
bb.chrons <-std.chron1t[,c(1,3,6)]
bb.chronst <-window(bb.chrons, start=1977, end=1998)
bb.ext <-lm(bb.obst ~bb.chronst$JGB+bb.chronst$KLP+bb.chronst$OGH)
bb.fit <-bb.ext$fitted
rmse(bb.fit^2, bb.obst^2 )
mae(bb.fit^2, bb.obst^2 )
summary(bb.ext) #adjR2: 0.52 SE 1.97,AIC 33.32,RMSE 66.55, MAE 54.59
bb.coef <-summary(bb.ext)$coefficients[,1]

#add last 10 years to bb.obs
last <-
read.csv("/Users/niah/Documents/CSU/Mongolia/dendrocores/TR_Paper/Khangai_TR_R/Reconst/Stepwis
e/Last10Yrs_Obs.csv")
lastind <-index(basin.obs[23:32])
last.z <-zoo(last[,2:5], lastind)
bb.obstr <-window(basin.obs[,1:4], start=1977,end=1998)
colnames(bb.obstr) <-colnames(last.z)
bb.obsa <-rbind( bb.obstr,last.z) #note these are actual
bb.obsasqrt <-sqrt(bb.obsa)
bb.chron2 <-std.chron2t[,c(1,4,7)]
#use regression coeffs
bb.extmod <-(7.645891+(5.995396*bb.chron2[,1])+(1.839592*bb.chron2[,2])+(3.403622*bb.chron2[,3]))^2

#plot
#BayanB2
bb2.recon <-bb.extmod
bb2.recont <-window(bb2.recon, start=1650, end=2008)
bb2.recondf <-data.frame(coredata(bb2.recont))
bb2.recondf$Year <-index(bb2.recont)
colnames(bb2.recondf) <-c("Baidrag R. at Bayanburd Model 2", "Year")

bb2.basin <-melt(bb2.recondf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCM"))

#get obs
bb2.obszm <-merge(bb.obsa[,1],bb.extmod)

bb2.obs <-bb2.obszm[,1]
bb2.obst <-window(bb2.obs, start=1650, end=2008)
```

```
bb2.obsdf <-data.frame(coredata(bb2.obst))
bb2.obsdf$Year <-index(bb2.obst)
colnames(bb2.obsdf) <-c("Baidrag R. at Bayanburd Model 2", "Year")

bb2.basinobs <-melt(bb2.obsdf, id.vars=c("Year"),variable.name=c("Basin"),value.name=c("MCMObs"))

#make rolling mean for plotting
#20 yr
bb2.roll <-rollapply(bb2.recont,20 , mean, fill=NA)

#melt rolling mean for appending to longwise
bb2.basinmeandf<-data.frame(coredata(bb2.roll))
bb2.basinmean <-melt(bb2.basinmeandf, variable.name=c("Basin"),value.name=c("Mean"))

bb2.basin$RollMean <-bb2.basinmean$Mean

#add mean for each recon
bb2.basin$Mean <-c(rep(345.9,359))

#add obs values
bb2.basin$MCMObs <-bb2.basinobs$MCMObs

bb2 <- ggplot(bb2.basin, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+facet_grid(Basin~.)+theme_bw()+
  scale_x_continuous(breaks=c(seq(1650,2000, by=25)))+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
      axis.text.y=element_text(size=12),axis.title.y=element_text(size=14,
face="bold"),strip.text.y=element_text(size=14))
#no BayanT model for extending streamflow, model fits too poor.
#looking at mean flows from 1999 to 2008 fitted and observed for BayanB2, ErdK2, and IkhKT2
#BayanB
bb2.fitmean <-cbind(mean(bb2.basin[350:354,3]), mean(bb2.basin[355:359,3]))
colnames(bb2.fitmean) <-c("1999-2003", "2004-2008")
bb2.obsmean <-cbind(mean(bb2.basin[350:354,6]), mean(bb2.basin[355:359,6]))
colnames(bb2.obsmean) <-c("1999-2003", "2004-2008")

#ErdK
ek2.fitmean <-cbind(mean(two1.basin[352:355,3]), mean(two1.basin[356:360,3]))
colnames(ek2.fitmean) <-c("1999-2003", "2004-2008")
ek2.obsmean <-cbind(mean(two1.basin[352:355,6]), mean(two1.basin[356:360,6]))
colnames(ek2.obsmean) <-c("1999-2003", "2004-2008")

#IkhKT
ik2.fitmean <-cbind(mean(two2.basin[845:849,3]), mean(two2.basin[850:854,3]))
colnames(ik2.fitmean) <-c("1999-2003", "2004-2008")
ik2.obsmean <-cbind(mean(two2.basin[845:849,6]), mean(two2.basin[850:854,6]))
colnames(ik2.obsmean) <-c("1999-2003", "2004-2008")

#BayanT
bt2.obsmean <-cbind(mean(bb.obsa[23:27,2]), mean(bb.obsa[28:32,2]))
colnames(bt2.obsmean) <-c("1999-2003", "2004-2008")
```

```r
#plots of each basin for comparison
#divide four.basin to each basin
bb.4 <-four.basin[1:417,]
bt.4 <-four.basin[418:834,]
ek.4 <-four.basin[835:1251,]
ik.4 <-four.basin[1252:1668,]

#plot
bb.4 <- ggplot(bb.4, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+theme_bw()+
  scale_x_continuous(breaks=c(seq(1650,2000, by=25)))+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
        axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
        strip.text.y=element_text(size=14),panel.grid.minor.x =element_blank())

bt.4 <- ggplot(bt.4, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+theme_bw()+
  scale_x_continuous(breaks=c(seq(1650,2000, by=25)))+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
        axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
        strip.text.y=element_text(size=14),panel.grid.minor.x =element_blank())

ek.4 <- ggplot(ek.4, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+theme_bw()+
  scale_x_continuous(breaks=c(seq(1650,2000, by=25)))+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
        axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
        strip.text.y=element_text(size=14),panel.grid.minor.x =element_blank())

ik.4 <- ggplot(ik.4, aes(x=Year, y=MCM)) +geom_vline(xintercept=seq(1700, 2000, by=100),
colour="grey80")+
  geom_line()+geom_line(aes(x=Year, y=RollMean), colour="red")+geom_line(aes(x=Year, y=Mean),
colour="grey50")+
  geom_line(aes(x=Year, y=MCMObs), colour="blue")+theme_bw()+
  scale_x_continuous(breaks=c(seq(1650,2000, by=25)))+
  theme(axis.text.x=element_text(size=12),axis.title.x=element_text(size=14, face="bold"),
        axis.text.y=element_text(size=12),axis.title.y=element_text(size=14, face="bold"),
        strip.text.y=element_text(size=14),panel.grid.minor.x =element_blank())

#make barcharts of above and below mean conditions for better interp.
bb.4wd <-data.frame(ifelse(bb.4$MCM>bb.4$Mean, 1, -1))
bb.4wd$Year <-as.factor(bb.4$Year)
colnames(bb.4wd) <-c("WD", "Year")
bb.4wdt <-bb.4wd[69:417,]
barplot(bb.4wdt$WD,axes=FALSE, space=0, main="Baidrag R.")
```

```
axis(side=1, at=c(1, 51, 101, 151, 201, 251, 301, 341), labels=c(1650, 1700, 1750, 1800, 1850, 1900, 1950,
1990))

bt.4wd <-data.frame(ifelse(bt.4$MCM>bt.4$Mean, 1, -1))
bt.4wd$Year <-as.factor(bt.4$Year)
colnames(bt.4wd) <-c("WD", "Year")
bt.4wdt <-bt.4wd[69:417,]
barplot(bt.4wdt$WD,axes=FALSE, space=0, main="Tuin R.")
axis(side=1, at=c(1, 51, 101, 151, 201, 251, 301, 341), labels=c(1650, 1700, 1750, 1800, 1850, 1900, 1950,
1990))

ek.4wd <-data.frame(ifelse(ek.4$MCM>ek.4$Mean, 1, -1))
ek.4wd$Year <-as.factor(ek.4$Year)
colnames(ek.4wd) <-c("WD", "Year")
ek.4wdt <-ek.4wd[69:417,]
barplot(ek.4wdt$WD,axes=FALSE, space=0, main="Khanui R.")
axis(side=1, at=c(1, 51, 101, 151, 201, 251, 301, 341), labels=c(1650, 1700, 1750, 1800, 1850, 1900, 1950,
1990))

ik.4wd <-data.frame(ifelse(ik.4$MCM>ik.4$Mean, 1, -1))
ik.4wd$Year <-as.factor(ik.4$Year)
colnames(ik.4wd) <-c("WD", "Year")
ik.4wdt <-ik.4wd[69:417,]
barplot(ik.4wdt$WD,axes=FALSE, space=0, main="Khoid Tamir R.")
axis(side=1, at=c(1, 51, 101, 151, 201, 251, 301, 341), labels=c(1650, 1700, 1750, 1800, 1850, 1900, 1950,
1990))
```

REFERENCES

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http:// www.R-project.org/.*