THESIS

REGIONAL PATTERNS OF SNOW WATER EQUIVALENT IN THE COLORADO RIVER BASIN USING SNOWPACK TELEMETRY (SNOTEL) DATA

Submitted by

Jeffrey Edward Derry

Department of Forest, Rangeland and Watershed Stewardship

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2008

COLORADO STATE UNIVERSITY

October 31, 2007

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY JEFF DERRY ENTITLED REGIONAL PATTERNS OF SNOW WATER EQUIVALENT IN THE COLORADO RIVER BASIN USING SNOWPACK TELEMETRY (SNOTEL) DATA BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

Committee on Graduate Work

[redacted]

Nolan J. Doesken

[redacted]

John D. Stednick

[redacted]

Adviser: Steven R. Fassnacht

[redacted]

Department Head: Michael Manfredo

ABSTRACT OF THESIS

REGIONAL PATTERNS OF SNOW WATER EQUIVALENT IN THE COLORADO RIVER BASIN USING SNOWPACK TELEMETRY (SNOTEL) DATA

Identifying regions of homogeneity of precipitation data is often a crucial preliminary step in natural resource investigations. Previous clustering of station based snow water equivalent (SWE) data has typically grouped stations based on spatial proximity, political boundaries, or watershed boundaries, and has been restricted due to the temporal resolution of snow course data. This investigation utilized daily data from 216 snowpack telemetry (SNOTEL) stations located in and around the Colorado River Basin over a 15-year period (1991-2005) to cluster stations, i.e., identify regions of homogeneity, based on the patterns and variability of SWE. To achieve this, data were submitted to a self-organizing map (SOM), a particular application of artificial neural networks. This methodology represents a learning algorithm that is non-linear, non-parametric, unsupervised, and learns through an iterative training process.

The number of clusters can be specified to the SOM based on the level of generalization desired. A SOM consisting of a 4, 6, 9, and 16-cluster were constructed from daily values as well as a 6-cluster derived from snowpack descriptor variables (peak SWE, length of snow season, etc.) and physical variables (elevation, aspect, distance to moisture source, etc.) for each station. Areas of homogeneity derived from daily SWE values, annual peak SWE, and physiography were used for multivariate regression analysis to determine the physical variables that best explain variability in peak SWE.

Results showed an unbiased clustering of stations defined not by station location, but by each station's specific SWE variability over the period of study. The established snow climatologies derived from daily values show general homogenous coarse-scale clusters along a north/south gradient with spatial coherence improving at finer resolutions, but overall there are no definitive spatial patterns to the climatologies, indicating that complex local-scale variables dominate variability of daily SWE. Climatologies derived from descriptor variables showed improved spatial coherence which reflected larger scale influences. Descriptor variables that best represent daily time-step classifications were peak SWE (50% similarity), April 1st SWE (43% similarity), and physical variables (41% similarity).

Regression results showed a consistent increase in predictability as cluster size went from more general (4-cluster) to less general (16-cluster). Key physical variables are elevation, southwest barrier height, regional northness, and southwest shield height. These key variables were consistently used in the regression model, although the degree of importance of the variable depends on resolution and general location of the climatology.

Jeff Derry
Department of Forest, Rangeland and Watershed Stewardship
Colorado State University
Fort Collins, CO 80523
Spring 2008

ACKNOWLEDGEMENTS

I would like to acknowledge members of my graduate committee, Dr. John Stednick, Nolan Doesken, and Dr. Steven Fassnacht, who provided insightful advice and guidance. Special thanks to my advisor, Steven R. Fassnacht, for his patience, enthusiasm, and open door policy throughout the entire process, it has been very much appreciated. Appreciation to Chuck Anderson and James Knight for assisting with the use and application of Self Organizing Maps. Thank you to Andrew Slater for suggesting the use of Self Organizing Maps. My gratitude to my family and friends as they have provided unwavering support in my personal and academic endeavors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The Western United States is dependent upon the water melting from seasonal snowpacks with 70-80% of the annual discharge in most streams originating from snowmelt [Doesken and Judson, 1996]. Throughout the Rocky Mountain region, a majority of the precipitation originates in the form of snow, with 64% as snow in Northwest Wyoming, 63% in Colorado, 62% in Idaho/western Montana, and 39% in New Mexico and Arizona [Serreze *et al.* 1999]. For the Colorado River Basin, mountains are a crucial water source since the proportion of discharge generated in mountain zones exceeds 90% of total discharge for the entire Basin, thereby making higher elevations/mountains a vital source of surface water for lowland populations [Viviroli *et al.* 2003].

Precipitation in the form of snow is temporarily stored in mountains over the winter providing reliable discharge during the spring and summer seasons. Snow accumulation and ablation processes are spatially and temporally heterogeneous principally due to complex mountain topography and landscape [Bales *et al.* 2006]. Understanding the temporal and spatial variability of snow accumulation, as well as the magnitude and timing of spring runoff are of crucial importance in developing accurate discharge forecasts to meet various needs including irrigation, hydropower generation, timing of storage on reservoirs, domestic consumption, recreation, interstate compacts, and international treaties. Furthermore, efforts to model regional responses to large-scale climate processes with global circulation models are difficult to downscale to a local level since these are limited by the uncertainty of regional hydroclimate forecasts [McGinnis, 1997; Derksen and LeDrew, 2000; Clark *et al.* 2001].

To understand the spatiotemporal distributions of snow accumulation and ablation, topographic variables have been used as predictor variables in modeling precipitation and snow water equivalent [Daly *et al.* 1994; Elder *et al.* 1998; Chang & Li 2000; Fassnacht *et al.* 2003]. Although much research has explored the relationships between topography and snow water equivalent (SWE), results have varied with scale [Bales *et al.* 2006]. At larger scales, such as the Colorado Basin, Fassnacht *et al.* [2003] suggested the need to create sub-basins, or regions of hydroclimatic homogeneity. The ability to accurately identify like hydroclimate regions can improve our capacity to relate topography with SWE.

The headwaters of the Colorado Basin lie in Western Colorado, Utah, Wyoming, Arizona, and New Mexico where snow is the dominate form of precipitation. In order to estimate the distribution of snow, it is crucial to have spatially representative measurements of SWE, or snow depth and snow density measurements. Traditionally, station based SWE data used in snow studies have been monthly snow course data. More recently, snow telemetry (SNOTEL) data with the advantage of having a daily resolution have become available [NRCS, undated].

This study examines if SNOTEL stations, within and in close proximity to the Colorado Basin, can be regionalized based on patterns and variability of SWE data in order to establish snowpack climatologies. The established climatologies will be examined to see if they will improve the multivariate regressions with physiography as the predictor variables and peak SWE as the response variable. Regressions are conducted on regions of homogeneity considering different resolutions, i.e., different number of groupings, and regions derived from different snowpack datasets.

The second chapter of this thesis will summarize operational snow measurements, regionalization of snow data, grouping techniques, physiography for snow mapping, and ends with objectives. The study area and data set are presented in chapter three. The methods are described in chapter four, the results are presented in chapter five and discussed in chapter six. The conclusions and discussions are summarized in the last chapter.

CHAPTER 2: BACKGROUND

2.1 Snow Measurements

Snow water equivalent is defined as:

$$SWE = ds * \rho s \qquad (1),$$

where ρs is snow density, and ds is snow depth. At any given point in time or space a snowpack's density multiplied by its snow depth result in the snow water equivalent of that particular column of snow.

Field techniques typically involve measuring SWE directly. This entails taking a hollow tube with a known inner cross-sectional area, initially a Mt. Rose and now in the Western U.S. a Federal Sampler, and inserting it vertically into the snow. The depth is noted on the outside of the tube. The tube containing the snow sample is then weighed in-situ with the tare weight subtracted from the total amount, yielding the SWE for that column of snow.

2.1.1 Snow Course

Historically, the snowpack SWE measurements used to estimate spring runoff have been collected at established snow course sites on or around the first of every month during the winter and spring season. For a snow course, ten measurements are typically taken in clearings along an approximately 300 m transect meant to represent the surrounding area [Viessman and Lewis, 2003]. Initial snow surveys began in the early 1900's in California [Church, 1914], and became more widespread by the late 1930's after a devastating drought in the early 1930's [NRCS, undated]. Approximately 2000 snow courses are surveyed monthly (or semi-monthly) across the Western U.S. several times each winter [NRCS, undated; Serreze *et al.* 1999], often January or February 1st through May or June 1st. April 1st is the most important date as it is generally considered representative of peak SWE.

2.1.2 Snow Telemetry

As an alternate to the labor intensive process of conducting snow course surveys, the USDA Soil Conservation Service (now the Natural Resources Conservation Service, NRCS) began installing year round automated Snow Telemetry (SNOTEL) stations in the late 1970's [Doesken and Schaefer, 1987]. SNOTEL stations are designed to provide economically feasible point data at daily or higher temporal resolution providing SWE, precipitation, temperature, and at some locations soil temperature/moisture and snow depth [NRCS, undated]. SWE is determined with a snow pillow made of rubber, hypalon, or recently stainless steel containing an antifreeze solution that measures the mass of the overlying snowpack using an electronic pressure transducer converts this mass to an equivalent voltage. All data are then telemetered using meteor burst technology to a central NRCS facility [Palmer, 1986].

There are approximately 722 SNOTEL sites across the western United States [NRCS, undated]. While SNOTEL data have a much higher temporal resolution than snow course data, they have seen limited use in snowpack investigations [Serreze *et al*. 1999]. This may be due to the short period of records for SNOTEL data (28 years or less) compared to snow courses (often 70 years or more) [NRCS, undated].

Another concern involves data quality. For instance, incorrect measurements can occur when the soil and sensor interface are warmer then freezing, resulting in a disparity in melt rates. Another source of error can occur when snow overlying the pillow forms a hard crust, causing a bridge to form over the pillow. In addition, branches and other debris can fall on the pillow increasing the mass being measured [Johnson and Schaefer, 2002].

The point nature of a SNOTEL station may be less representative of its surrounding area compared to the transect of snow course points. Molotch *et al.* [2001] showed SWE could vary significantly 500 m beyond a SNOTEL site due to terrain influences on snow accumulation and ablation. However at the larger scale, such as the Colorado River Basin, Dressler *et al.* [2006] determined that even though snow course data are more representative of SWE elevation distributions, both datasets are sufficient for estimating peak basin-wide SWE. Overall, snow course data provide higher spatial coverage and low temporal resolution while SNOTEL data provide low spatial coverage and high temporal resolution.

2.2 Regionalization of Snow Data

Upscaling from station to regional data requires designating data into like groups, i.e., identifying regions of homogeneity. This is a crucial first step when conducting regional flood frequency analysis, modeling unmonitored basins, or in the development of global circulation models [McGinnis, 1997], since grouping data into like groups often results in regional relationships with lower standard errors than otherwise [Hann, 2002]. There is no uniquely objective approach to delineate areas of homogeneity, yet a common approach is to define regions by pre-established boundaries, such as political, catchment, geologic, climatic, physiographic, or physical proximity to measurement stations. For instance, Serreze *et al*. [1999] separated SNOTEL stations in and around the Colorado River Basin into four regions approximately by state boundaries, i.e. Northwest Wyoming, Utah, Colorado, Arizona/New Mexico (Figure 2.1), while Pitlick [1994] divided the Colorado region into "alpine" and "foothills". In the Canadian Rockies Moore and McKendry [1996] grouped snow course stations by major geographic delineations.

Others have defined regions based on statistical techniques, including correlation analysis, probability distributions, coefficient of variation, and residual analysis. Changnon *et al*. [1991] combined snow course, streamflow, and precipitation data using a modified coefficient of variation in select watersheds to define areas of homogeneity in the Rocky Mountains. Based on these groupings, Changnon *et al*. [1993] used averaged snow course values to relate snowpack with synoptic patterns. Clark *et al.* [2001] classified snow courses and stream gauges into drainage basins based on major subbasins in the Columbia and Colorado River system to investigate El Nino and La Nina effects on seasonal snowpack evolution.

Non-classical statistical approaches include multivariate techniques, such as principle component analysis (PCA) and cluster analysis. These techniques require less subjective decision making relative to other methods and have the ability to facilitate data interpretation. Cayan [1996] utilized April 1st snow course SWE to show a covariance on large geographic scales using rotating principle component analysis (PCA). For the Colorado Basin, McGinnis [1997] subjected daily SNOTEL data to a PCA and from the established regions used averaged daily time-steps to simulate potential climate change impacts. While there is no strictly objective approach to the delineation of homogeneous regions, most investigations require defining areas of homogeneity prior to the start of an analysis (e.g., Changnon *et al*. 1993; Cayan, 1996; Moore and McKendry, 1996; McGinnis, 1997; Serreze *et al*. 1999; Clark *et al.* 2001).

2.3 Clustering

2.3.1 Ward's Minimum Variance

Conventional clustering techniques, such as exclusive or hierarchical, are the procedures most often used to group like data. Ward's Minimum Variance is a linear, hierarchical and agglomerative clustering procedure. The Ward's procedure along with other common clustering techniques, such as centroid or average linkage, are available through various statistical software packages. Ward's minimum variance method is the most frequently used hierarchical clustering technique for climatic classification [Kalkstein *et al*. 1987].

As with all hierarchical clustering techniques, Ward's possesses unique characteristics that determine data categorization. Clusters are generated that minimize the squared Euclidean distance to the center mean value for all objects in a cluster (or to the centroid). One important distinction is how the number of observations influences the cluster groupings. If the distance between two clusters is the same, an observation will be joined with the cluster with the fewest observations. Additionally, when considering distance between clusters, the value is proportional to the squared distance between the new centroid and the original centroid. These two factors result in priority to merging clusters with fewer observations and encouraging similar sized clusters.

Other hierarchical clustering procedures have related limitations. For example, the centroid method calculates the squared Euclidean distance between the centroid of clusters and joins clusters that minimize this distance, resulting in one or two large clusters encompassing most observations, and smaller clusters having only one or two observations. Conversely the average linkage method attempts to minimize the within-group variance and maximize the between group variance. Use of variance, rather than sum of squares, reduces the influence of cluster size on merging of clusters. The result is an equilibrium between having a large cluster as in centroid and having many similar sized clusters as in Ward's. Regardless of the hierarchical clustering technique employed, all possess unique characteristics which can equate to large interprocedural differences.

2.3.2 Self-Organizing Maps

A relatively recent method of data classification uses a form of artificial neural networks (ANN), known as the Kohonen Self-Organizing Map (SOM) [Kohonen, 1995]. The SOM provides a methodology that extracts common patterns of regional variability from patterns of individual variability of each station through an automated, iterative training process.

Crane and Hewitson [2003] used SOMs to upscale precipitation data in the Northeastern/Mid-Atlantic region of the United States, demonstrating the utility of the methodology in grouping multi-dimensional spatially distributed precipitation data. Lin and Chen [2006] used precipitation data with SOMs to identify regions of homogeneity for regional frequency analysis. Cavazos [2000] SOMs coupled with an artificial neural network to investigate wintertime extreme climate events in the Balkans. Cavazos [1999] also used SOMs in Texas to classify winter circulation and humidity fields and used the results as input to a feed-forward artificial network to downscale grid-scale precipitation. These investigations using SOMs offer encouraging results in terms of improving the spatial classification of hydroclimate variables. It has been shown that SOMs produce better results compared to the conventional linear methods of clustering techniques (K-means, etc.), particularly when applied to spatial data [Mangiameli *et al.* 1996; Lin & Chen, 2006].

SOMs are a form of cluster analysis that have the ability to reduce multi-dimensional data, which may relate to each other in a nonlinear fashion, into two dimensions [Cavazos, 1999]. Yet, the SOM is not primarily concerned with identifying clusters. Instead, SOMs convert a continuous input space to a finite output space, called the SOM map. Multi-dimensional data are projected onto the map as generalized states by means of an iterative training process. Nodes (or clusters) in the measurement space (or map) represent the cloud of observations and, all together, describe the multi-dimensional distribution of the data set [Hewitson and Crane, 2002]. The emphasis of the algorithm is to plot similar input vectors close to one another on the two dimensional map. The vector of an individual data record is compared to the reference vector of each

SOM node. The SOM node with the closest match, calculated from the Euclidean distance, to the data vector is referred to as the best matching unit (BMU). This node is updated to resemble a closer match to the input vector with each surrounding node slightly adjusted in the direction of the input vector with the amount of adjustment decreasing further away from the reference node. SOM training consists of alternating between mapping input records and adjusting reference vectors. This process is repeated for each data record with multiple iterations through the dataset to train the SOM until the nodes no longer require adjustment with continued iterations. For this unsupervised automated learning, prior parametric constraints are not required, i.e., the network learns to organize the data by recognizing different patterns.

The number of output nodes (or clusters) needs to be specified before the learning process. This is analogous to K-means clustering where the number of clusters (K) is determined beforehand. For example, if 12 clusters are desired in a traditional cluster analysis this is specified beforehand. Similarly, for a SOM, a 3×4 (=12) map is specified beforehand. The level of generalization depends on the size of the map. A 2×3 map (6 nodes) is more general than a 4×6 map (24 nodes). Nonetheless, the SOM will remain consistent regardless of map size. For instance, nodes that mapped opposite of each other on a 2×3 map will map opposite on a 4×6.

A major difference between SOM's and traditional cluster analysis is that a SOM provides a visualization of the distribution function of multi-dimensional data revealing relationships not obvious from direct observation. As mentioned, a SOM will arrange the distribution of nodes onto a two dimensional map, where similar nodes are located close together on the map and dissimilar nodes are further apart. Applied to SNOTEL data, stations with similar patterns of variability will be close together on the SOM map and vice-versa. This is an advantage to other clustering approaches in that the results are less dependent on the input data conforming to a specific distribution or preconceived model, because SOMs are not linearly spaced. They represent an even coverage of the probability distribution function, and do not force data into orthogonal linear combinations [Hewiston and Crane, 2002].

The SOM assumes the data are continuous, and if there are missing data it provides a means to interpolate. This interpolation replaces missing values in the data with values from the SOM prototypes. Fewer SOM nodes are allocated where there are sparse data and more nodes on the map where there is increased information of the data set. Essentially the SOM attempts to represent the detailed points of the data at whatever level of generalization is used [Crane and Hewitson, 2003]. Crane and Hewitson [2003] tested the SOM methodology in relation to the degree of missing precipitation data in the Northeastern U.S. and compared to a complete dataset, noted an improvement in the regional signal when up to 80% of the data is missing from 80% of stations or when there is a bias in the missing data.

SOM classifications are both a subjective and quantitative process. Training can be considered complete (subjectively) once nodes near the edges of the map have stabilized because these nodes are often the most distinct. Quantization error, the average distance between each data vector and its best matching unit (node), is a useful objective measure of a SOM. The topographic error, the proportion of all data vectors for which first and second best matching unit are not adjacent, also assists in measuring topology preservation.

Another adjustable parameter in construction of a SOM is the "learning rate". Each input value is matched to the closest reference vector, and in turn the reference vector is adjusted towards the value of the input record. The amount of adjustment is determined by the dimensionless learning rate. A learning rate of zero makes no adjustments and a rate of one applies the entire difference between the reference and input vector. For instance, a learning rate of 0.05 applies this amount of the difference between an input value and its closest reference vector. Typically, an analysis involves two training phases: initialization and fine-tuning. During the initialization phase, a learning rate begins around 0.5 and decreases exponentially to 0.00 upon completion of training.

The learning rate is adjusted together with the update radius. The update radius typically begins at the maximum value of the entire map dimension (e.g. a value of 5 for a 5 x 3 map). Towards the end of the fine-tuning phase the radius is steadily reduced to 0.5 or half the amount of one radius.

Two different training algorithms are available; sequential and batch. The sequential training algorithm trains the SOM iteratively. For each training step, one sample from the input data set is chosen randomly and the distances between it and all weight vectors are calculated with the closest weight vector being the BMU. The batch training is also iterative, but instead of using a single data vector at a time, all the data is presented to the SOM before any adjustments are made.

2.4 Physiography For Snow Mapping

Physiography has received much attention as a significant spatial and temporal control on the snowpack. Using binary regression trees at the hill-slope scale, Elder *et al.* [1998] found that net radiation (net flux of all radiation), elevation, and slope accounted for 60-70% of the variance in snow depth. At the regional scale, Fassnacht *et al.* [2003] evaluated inverse weighted distance and regression methods using physiography, i.e., location (x, y and z), distance to ocean, barrier height, shield height, and aspect, as predictive parameters for interpolating SWE. Throughout a five state area (Colorado, Utah, Idaho, Montana, and Wyoming) in the Northern Rocky Mountains, spatial variability of SWE was best explained by aspect in relation to moist airflows, and snow was a better monitor of winter climate than streamflow or precipitation [Changnon *et al.* 1991].

Investigating interannual variability of mountain snow accumulation in the 11 western states using snow course data, Cayan [1996] found SWE anomaly patterns covary on large regional scales with loss of coherence occurring between the Cascade Mountains and the Sierra Nevada (around the Oregon-California border), and between the Northern and Southern Rocky Mountains. Serreze *et al.* [2001] examined spatiotemporal characteristics of large snowfall events using SNOTEL data and found large events at a particular site strongly reflected the local topographical setting and that continental effects on leading event size can be masked by local topographic effects. They reported lower coherence frequencies, i.e., spatial proximity of stations with like results, than Cayan [1996], possibly due to the dataset used.

A majority of studies have utilized snow course data usually with a monthly time step, and these data do not have the temporal resolution required to understand the variability over the winter season [Serreze *et al.* 2001]. This is important since the observed variability of snow distributions changes with the scale of the observation [Blöschl, 1999]. Thus, consideration of temporal resolution as well as spatial scale is vital in any assessment of the distribution of snow.

2.5 Purpose and Objectives

The intent of this study is to identify areas of homogeneity using high temporal resolution (daily) SWE data over a large spatial extent, in this case the Colorado River Basin. Homogeneous regions will be grouped based on similar trends of variability in SWE among SNOTEL stations. This investigation will address how snow climatologies based on groupings of station data are partitioned in the Basin at different levels of resolution.

To reduce the amount of data used in clustering, rather than using all the daily data in an annual niveograph, four descriptor variables will be used as surrogates for the daily data. For this study, a niveograph is defined as a plot of SWE versus time (Figure 2.2). The four niveograph descriptor variables are peak SWE, cumulative SWE, length of snow season and date of peak SWE (Figure 2.2). Cumulative SWE is used as it represents the persistence of snow cover together with the magnitude of SWE and is analogous to runoff volume for a streamflow hydrograph. April 1st SWE will also be used since it is the traditional peak SWE surrogate measured at snow courses.

The snow climatologies will be established using four approaches (Table 2.1): i) from daily data at different levels of generalization, i.e., number of clusters, ii) using the surrogate descriptor variables, iii) change of clustering methodologies, i.e., SOM vs. Ward's, and iv) clustering with physiography. These clusterings will all be compared to groupings based on political boundaries, as defined by Serreze *et al*. [1999].

The established regions are then utilized as areas of homogeneity for the purpose of identifying the dominant physical influences on snow accumulation/ablation within each region. This will consist of applying a multivariate regression between physiography as the independent variable and peak SWE as the dependent variable (Table 2.2).

Table 2.1. Combination of cluster analyses compared. In parenthesis indicates number of variables for each station used in the clustering analysis. The letter in the right hand columns are reference letters for results in Appendix A.

			number of clusters				
method	data	number of data points	4	6	9	12	16
wards	peak SWE	(15 years x 1 descriptor variable per season = 15 variables		V			
wards	length of season	(15 years x 1 descriptor variable per season = 15 variables		X			
wards	date of peak SWE	(15 years x 1 descriptor variable per season = 15 variables		W			
wards	cumulative SWE	(15 years x 1 descriptor variable per season = 15 variables		Y			
SOM	daily for 15-yrs	(15 years x 273 days per season = 4095 variables)	B	C	E	F	G
SOM	daily for 25-yrs	(25 years x 273 days per season = 6825 variables)		H			
SOM	peak SWE	(15 years x 1 descriptor variable per season = 15 variables)	I	J	K		L
SOM	length of season	(15 years x 1 descriptor variable per season = 15 variables)		U			
SOM	date of peak SWE	(15 years x 1 descriptor variable per season = 15 variables)		R			
SOM	April 1st SWE	(15 years x 1 descriptor variable per season = 15 variables)		Q			
SOM	peak & length	(15 years x 2 descriptor variable per season = 30 variables)		T			
SOM	4 descriptors	(15 years x 4 descriptor variables per season = 60 variables)		S			
SOM	raw, unscaled data	(15 years x 273 days per season = 4095 variables)		D			
SOM	physiography	(same 18 variables for 15 years)	M	N	O		P
political/physical	as per Serreze *et al.* [1999]	none	A				

Table 2.2. Derivation and size of the clusters included in the regression analysis.

physiography	daily data	annual peak SWE
4 cluster	4 cluster	4 cluster
6 cluster	6 cluster	6 cluster
16 cluster	16 cluster	16 cluster

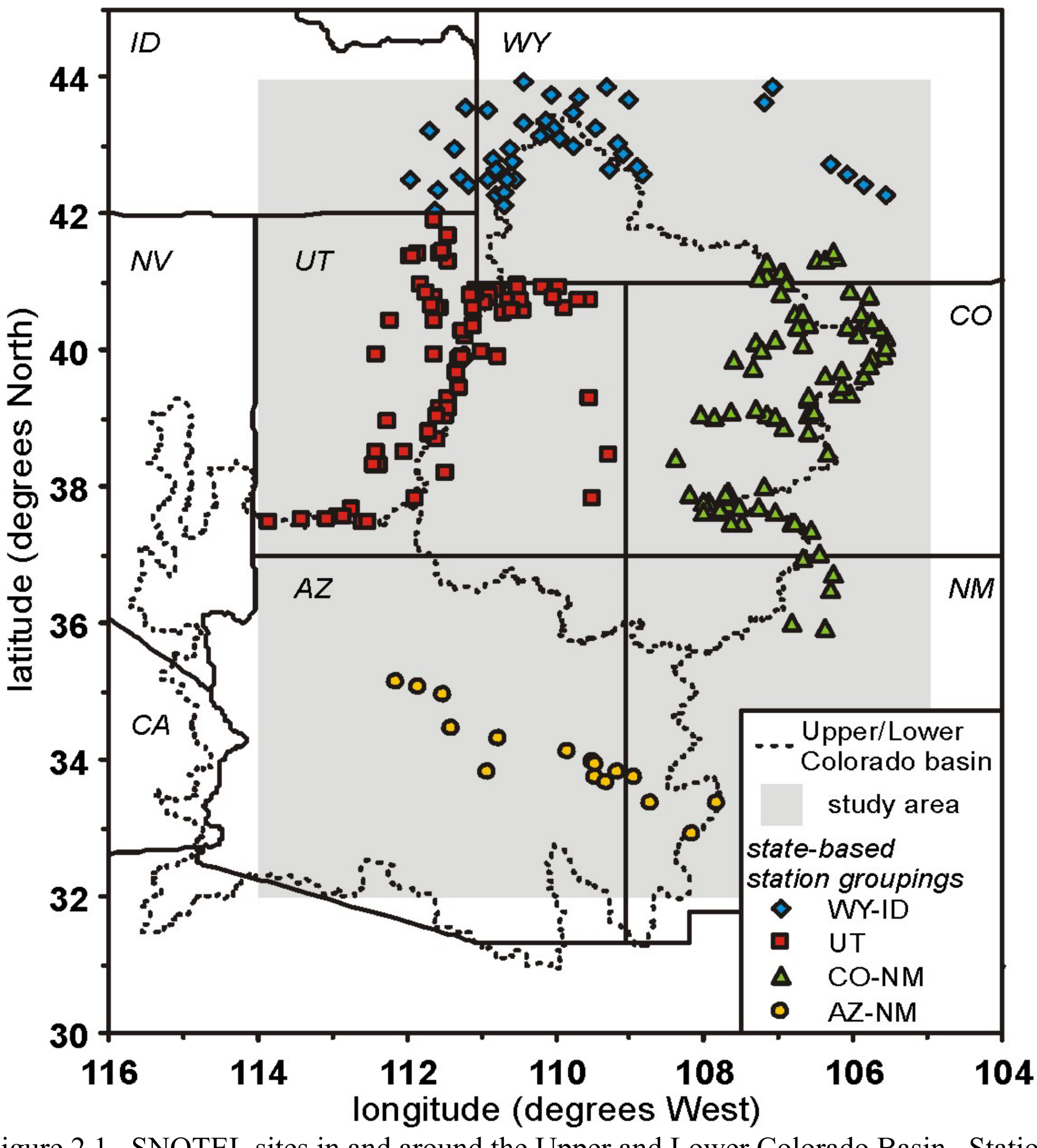

Figure 2.1. SNOTEL sites in and around the Upper and Lower Colorado Basin. Stations are grouped according to state-based delineations defined by Serreze *et al.* [1999].

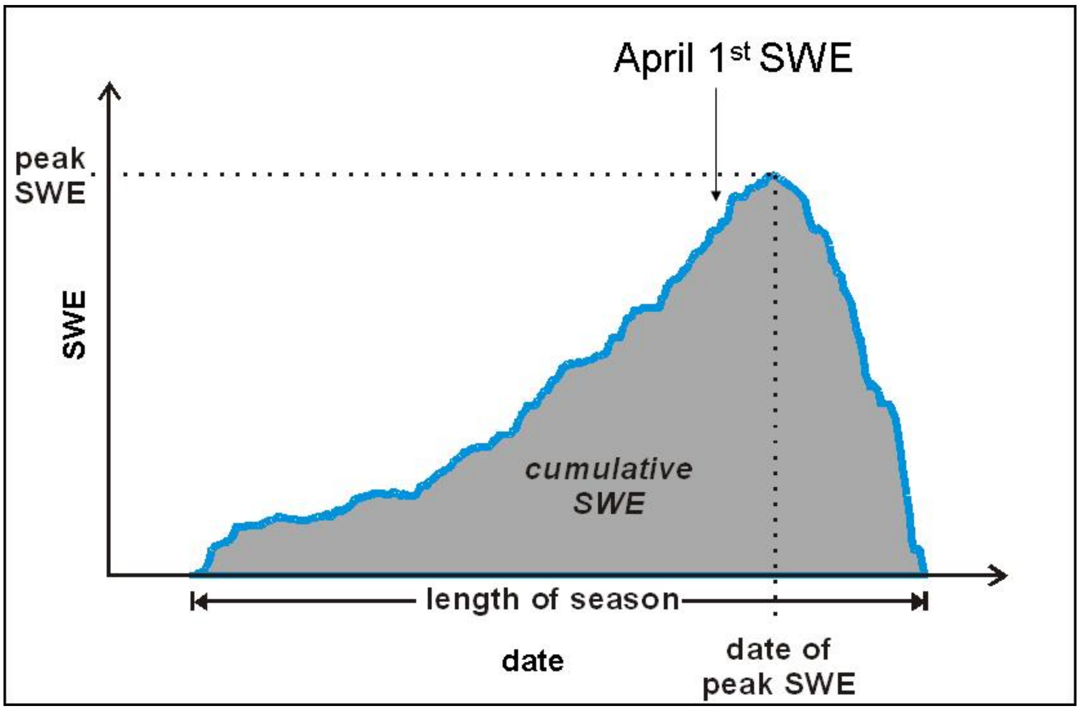

Figure 2.2. Niveograph (season snow accumulation/ablation over time) illustrating key snowpack descriptor variables.

CHAPTER 3: STUDY SITE AND DATA

3.1 Study Basin

This investigation considers regions of homogeneity defined in and around the Colorado Basin (Figure 2.1). The majority of snowfall occurs in the Upper Colorado Basin where the elevation range is 975 to 4260 m (average 2150 m) in a drainage area of 277,000 km^2. The Lower Colorado has an elevation range of 0 to 3771 m (average 1310 m) in a drainage area of 346,000 km^2. Approximately 60% of the Upper Colorado Basin is above 2000 m, but only 16% of the Lower Basin reaches this height [Fassnacht *et al.* 2003]. Principal areas that receive snow in the Lower Basin are located along the Mogollan Rim in eastern and central Arizona, western New Mexico, and the Colorado Plateau near the Grand Canyon.

3.2 Dataset

This study employed SNOTEL data acquired from the NRCS [undated]. With the majority of SNOTEL stations coming online in the early 1980's a compromise was required between number of stations and number of years included in this study. There were 240 stations identified by Fassnacht *et al.* [2003] in and around the Colorado Basin with the range of complete records varying considerably. To avoid missing data, 216 stations were identified that had continuous data for a 15-year period from 1991 to 2005. The investigation also utilized 25-years of data from 1981 to 2005 for the same 216 stations with 10% of these SWE values missing, to access the SOM ability to group data with missing time steps.

Since October 1 marks the beginning of the water year and missing or erroneous data values are common during the summer months, the snow year is taken from October 1 to June 30. Even though some snow may fall outside this period, it is the main accumulation and ablation period for SNOTEL stations in and around the Colorado Basin.

3.2.1 Data Quality Control

Serreze *et al.* [1999] developed quality control procedures for the SNOTEL data for seasons 1963/1964 though 1995/1996 and the methodology was applied to this data set. The procedures identified days with the following criteria that were then removed from the time series:

- Changes in SWE with an absolute value greater than 10 inches.
- A large accumulation event (SWE > 2.5 in) followed the next day by a large loss event (SWE < 2.5 in); or conversely, a large loss event followed by a large accumulation event.
- SWE values more than five standard deviations above or below the monthly mean. However, if the positive extreme event was also confirmed by the precipitation gage at the site, it was assumed to be a valid event and the value was retained.

The above checks resulted in discarding less than 0.13% of the data. In addition to the above quality control procedures, approximately 16 days of data were missing from the 884,088 station days (15 years x 273 days/year x 216 stations). These missing days were evenly distributed throughout the data set and typically only amounted to one day out of a station's record with only a few stations having missing values for 3-4 consecutive days. Missing data were replaced with the average of the previous data record and next complete data record. It was judged replacement of such a small amount of data records with average values was more than offset by the advantage of including as many stations as possible in the study. Moreover, especially with the methodology used (utilization of daily values), several averaged days out of a station's 4093 days of data will likely be negligible.

CHAPTER 4: METHODS

4.1 Data

4.1.1 Data preparation

To determine snow climatologies, the objective is to compare a particular station with itself to decide what constitutes a dry, average, or wet year and then stations with similar yearly patterns will be grouped together. Magnitude of SWE is important but this varies locally, often as a function of elevation [Fassnacht *et al*. 2003], so annual patterns of accumulation and ablation compared to the average of a particular station are more indicative of snow climatology. To scale the daily data to reflect an individual station's variability, each station's daily value was divided by that station's 15-year mean peak SWE value. This transforms the daily values into fractions of the average maximum peak SWE yielding standardized SWE values that are scaled to mean peak SWE. This procedure allows stations with low accumulation in the southern regions or at lower elevations to be directly compared with stations of relatively high accumulation in the northerly regions or at higher elevations. Stations with the same snow accumulation and ablation patterns should thus be grouped together, due to similar atmospheric controls, even if absolute amounts differ.

Standardization of daily SWE data depends on the normality of the data. Generally, stations exhibit relatively normal distributions in terms frequency of magnitude of peak SWE (Figure 4.1). The justification to scale data in this investigation is twofold. First, since different data derived from each SNOTEL station were used in clustering, it was not prudent to assume that all niveograph descriptor variables as well as the daily data values for the fifteen year period of study was normally distributed. Second, the scale of the variables is very different, and clustering methods are very sensitive to such scale differences [Lin and Chen, 2006]. Hence, variables must be transformed so that their ranges are comparable.

To explore how the SOM performs with a high amount of missing data, groupings of the complete 15-year data set were compared with a longer 25-year (1981 – 2005) data set with some missing data for the same stations. A total of 10% of the data were missing and evenly dispersed throughout the data set. This longer dataset was also scaled this time with the 25-year average peak SWE value.

4.1.2 Descriptor Variables

Peak SWE is the maximum amount of snow water equivalent for a particular season (Figure 2.2). Cumulative SWE is the sum of additional daily SWE for a particular season, analogous to runoff volume in a hydrograph. Length of season is the number of days between the first day and last day of snow being present. Date of peak SWE is the date of maximum SWE. April 1st SWE represents the date that snow course campaigns are typically conducted and used to forecast runoff. Therefore, for the purpose of this study, the value reported on April 1^{st} by the SNOTEL stations is analogous to using snow course data.

A scaling procedure was also required for the descriptor variables in order for station comparisons to be made. To achieve this, each descriptor variable for each station was normalized with a mean of zero and standard deviation of one. Hence, when normalized values were submitted to the SOM process, comparisons were made based on degree of variability from a station's mean.

4.1.3 Physiographic Variables

The last method of identifying homogeneous regions established groups based solely on the physiographic variables pertaining to a particular SNOTEL station. Clustering of these variables allows stations to be grouped by the topographical attributes of the particular location. The variables are listed in table 4.1 with the method of their derivation. Multivariate regressions of physiographic variables has been used for large-scale climate data gridding [Solomon *et al.* 1968; Daly *et al.* 1997; Fassnacht *et al.* 2003]. Variables in this study were those used by Fassnacht *et al.* [2003] for large scale SWE interpolation. A brief overview of the physiographic variables is included here. For a more detailed description the reader is referred to Fassnacht *et al.* [2003].

Parameters were computed for a 1-km pixel using a 100 m digital elevation model. Variables are station based parameters; latitude (X), longitude (Y), elevation (Z), local slope (1 km^2 around pixel), local eastness (sine slope $*$ sine aspect), local northness (sine slope $*$ cosine aspect), regional slope (9 km x 9 km = 81 km^2 around pixel), regional northness, and forest density (represented as percent density). Also derived parameters

(Figure 4.2), which are west/southwest/northwest distance to ocean, west/southwest/northwest barrier height (elevation difference between maximum barrier in direction of the ocean and the station), west/southwest/northwest shield height (cumulative elevation increase between the ocean and the station).

Variability in storm-track direction does exist, but predominantly storm systems move in a westerly to easterly direction. Therefore, for this study the moisture source is considered to prevail out of the southwest to northwest principally from the Pacific Ocean.

4.1.4 Political Boundaries

The results of defining snow climatologies from this study were compared with Serreze *et al.* [1999] who delineated areas of homogeneity by state boundaries (Figure 2.1). Serreze *et al.* [1999] divided the Colorado Basin into four regions that were compared to the 4 cluster SOM derived from daily values.

4.2 Clustering

4.2.1 SOM procedure

The software package used for implementing SOM's and documentation are freely available [CIS, 2007]. The software was executed in MATLAB 7®.

For comparison of climatologies based on the 15-years of daily SWE data based on resolution, SOM's of size 4, 6, 9, and 16 were constructed. The reference clustering was chosen as a 6 cluster SOM with 15-years of daily SWE data. Climatologies of size 6 were the clustering resolution used for comparisons between different data sources (Table

2.1) since an analysis of this size is general enough to allow for visual comparisons yet still subdivides the Basin adequately for distinct regions to be established. SOM's were also produced for descriptor variables, combinations thereof, and the physiographic variables.

The finest temporal resolution of SNOTEL data is a daily time step. It is assumed clusters based on daily data result in the most appropriate and optimal groupings. For this reason, the 6 cluster SOM made with the daily data are considered the "base-case", or reference groupings (Table 2.1).

The SOM process includes a number of adjustable parameters. To determine when the training process is completed, this investigation combines mapping stability, i.e., are the groups changing significantly, and the quantization error to determine when training parameters and training length are optimal. Training iterations of 50,000 resulted in no further improvements to the map stability or quantization error. However, depending on the input variables, slight differences were still seen in what constituted each station's best matching unit after a few hundred thousand iterations. It was found that a predictable convergent solution was obtained with all the variables under consideration after approximately 500,000 iterations. Therefore, all of the SOM classifications in this work were done with 700,000 iterations to ensure sufficient iterations of training steps.

Default learning rates were used which begin at 0.5 and have a final value of 0.02 at the completion of training. The optimum initial radius proved to be the maximum value of the map size, and the final training radius to be 0.5. Lastly, with sufficient iterations, both the sequential and batch algorithm produced essentially the same results.

However, since the time requirements for the sequential training were 3-4 times more than the batch training, batch training was used.

4.2.2 Ward's Minimum Variance Procedure

Cluster size and data submitted to the Ward's Minimum Variance program were the same as for the SOM clustering of the normalized descriptor variables. Since traditional clustering methods cannot cluster multidimensional data, but rather only a single variable, comparison of the SOM results with Ward's Minimum Variance method was limited to snowpack descriptor variables.

4.3 Method Comparison

The cluster number resulting from the SOM procedure is somewhat arbitrary. The cluster number was reassigned for consistency and to allow for comparisons by keeping the cluster number the same for the general region it is located in the Basin.

Comparison of different cluster analyses was twofold. First, comparison of different cluster resolutions was made via the spatial distribution of clusters from each analysis (e.g., map of clusters). Second, for clusters derived from descriptor variables, the percent of stations that cluster to the same cluster from one analysis to another analysis was considered. This is similar to reporting the correlation between two different analyses, but since correlation of categorical data can be misleading, the percent of stations that were in the same cluster for different analyses was reported.

For all the analyses, comparisons were made based on the niveograph, derived from the average daily value, for each cluster in an analysis. Categorization of the niveographs was based on the timing and amount of peak SWE, as well as accumulation and ablation limbs.

4.4 Regression Analysis

For each analysis, each cluster, and every year of clustering, a linear stepwise regression was used between physiography as the independent variables and peak SWE as the dependent variable. Taking the four cluster analysis derived from daily SWE values as an example, for all stations that comprise each individual cluster, annual peak SWE was regressed with physiography every year for the 15 years of study.

Regressions were first conducted on all the data without clustering (cluster size one) to illustrate improvements in the regressions at finer levels of clustering. The coefficient of determination, or R^2, was reported to access predictive differences between regression models conducted on regions derived from different data sources as well as clustering at different resolutions. Regressions were performed using the SPLUS® statistical software package except when the number of independent variables exceeded the number of dependent variables. When this occurred, SAS® statistical software was used since it allows regressions to proceed under these conditions.

To avoid over-fitting the model and to reduce the independent variables to the least amount possible, the following three procedures were adopted:

- All independent variables were initially submitted to the regression model. If two or more variables incorporated into the model had a correlation of 0.70 or higher, then the variable(s) was removed that least decreased the coefficient of determination.

- After an initial run of the regression model to determine essential independent variables, the number of independent variables included in the model was kept at less than half of the number of dependent variables. This was done to avoid creating an artificially high R^2 due to over-fitting of the model.
- A variable was only retained if it increased the R^2 value by more than 3%. This threshold value was chosen because experimental runs of the model indicated it would help keep too many independent variables from being included in the model.

Table 4.1. Summary of regression parameters used, method of their derivation, and units. Notes: * the distance to the ocean, barrier height, and shield height were measured from the west, northwest and southwest.

variable name	source/method	units
longitude	X from Natural Resources Conservation Service (NRCS) data	decimal degree
latitude	Y component from above (NRCS data)	decimal degree
elevation	Z from DEM	metres
local slope	sin slope (Δz) from 1x1 pixels	-
local eastness	sin aspect (Δx) from 1x1 pixels	-
local northness	cos aspect (Δy) from 1x1 pixels	-
regional slope	Δz for 9 km swath around pixel	-
regional northness	regional Δy	-
W/NW/SW distance to ocean *	distance to ocean computed from west, northwest, and southwest	kilometres
W/NW/SW barrier height *	elevation difference between maximum barrier in direction of ocean and pixel	metres
W/NW/SW shield height *	cumulative elevation increase between ocean and pixel	metres
forest density	US Forest Service density maps from AVHRR imagery [U.S. Forest Service, 2001]	%

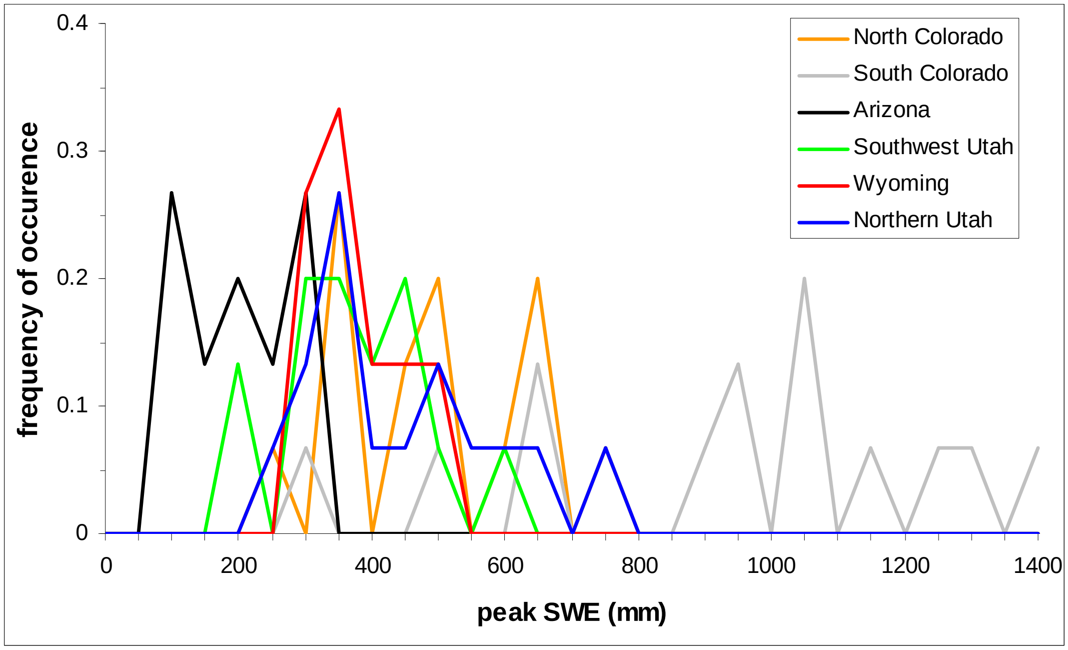

Figure 4.1. Frequency of occurrence of magnitude of peak SWE for six SNOTEL stations in the Colorado Basin.

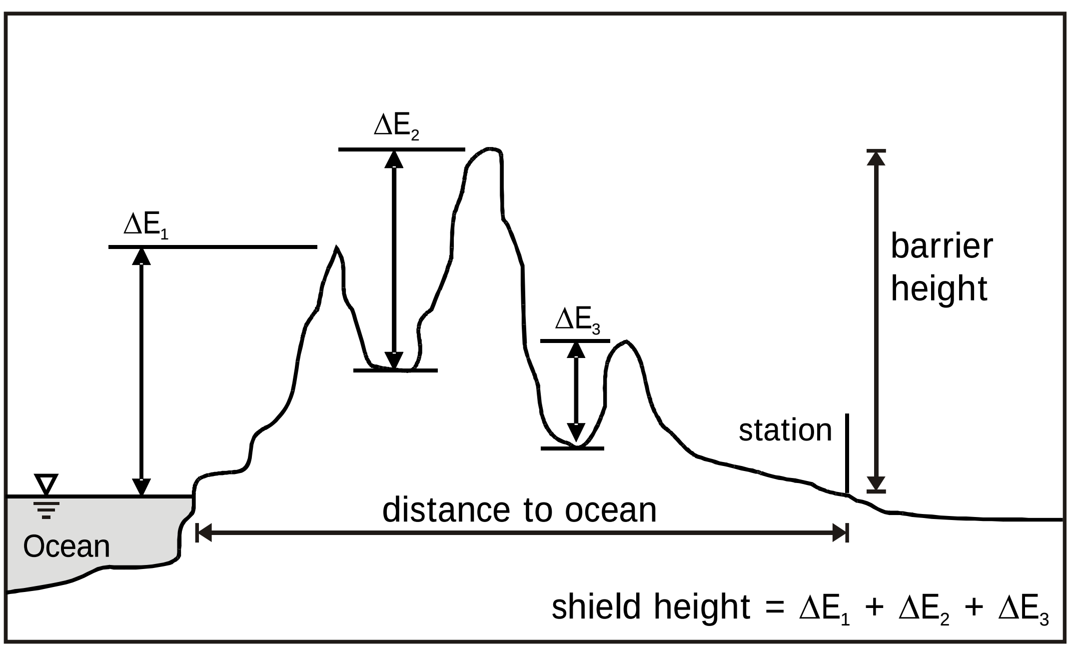

Figure 4.2. Schematic of derived variables used as independent variables in regression analysis.

CHAPTER 5: RESULTS

5.1 Clustering

5.1.1 Self Organizing Map

5.1.1a 15-Year Daily Data

The spatial distributions for various cluster sizes and data are presented in Figure 5.1. Stations cluster on a north/south gradient, with Arizona and New Mexico being a homogeneous area (Figure 5.1b). Wyoming, Idaho, Northern Colorado, and Northern Utah group primarily to two clusters (cluster 3 and 4). The majority of stations in cluster 6 are in mid/north Utah around the Wasatch Mountains. Cluster 5 dominates in the southern region of Colorado but includes a number of stations in northern Utah. The remainder of Colorado and Utah tend to maintain a certain level of heterogeneity with cluster 2 covering a more east/west transect located at lower to mid-latitudes of Utah and Colorado. Colorado maintains a division of northern and southern clusters approximately around the latitude of the town of Leadville, near the headwaters of three river basins (Colorado, Arkansas, and Platte).

The average niveographs for each cluster for various cluster sizes and data are presented in Figure 5.2. The six-cluster daily data niveograph (Figure 5.2b) illustrates clusters 2 and 3 as being similar and clusters 4 and 5 as being similar with the differences being attributed to total SWE as a result of the dissimilar ablation limbs. Average peak SWE for clusters 1 thru 6 occurs on March 7, April 1, April 1, May 1, April 10, March 15, respectively.

At the coarsest resolution of four clusters (Figure 5.1a), stations located in the southern area of Arizona and New Mexico already show a high degree of spatial coherence (cluster 1). Cluster 2 and 4 constitute all of Wyoming and northern Colorado, while cluster 3 includes the majority of southern Colorado, Utah, and Idaho.

Going from cluster 1 to cluster 4 (Figure 5.2a), the niveograph is equally divided in terms of total SWE and peak SWE. Peak SWE for clusters 1 thru 4 occurs on March 8, April 1, April 9, and May 1, respectively. The cumulative average SWE increases from cluster 1 in the southerly region; to cluster 2 in the northerly region of Colorado, Wyoming, and Utah; to cluster 3 in the mid-latitude area of Colorado and Utah, and finally to cluster 4 again in the northerly region of Colorado, Wyoming, and Utah.

Progressing to a finer-scale of nine clusters, sub-divisions are seen across the entire Basin (Figure 5.1c). The SOM at this level of generalization creates another cluster in the southern area of the Basin still with an average peak SWE on March 1st as in the six-cluster, yet a third less in magnitude (Figure 5.2c). From there each cluster has a different peak SWE in terms of timing and magnitude with varying accumulation and ablation limbs. Cluster 8 has the latest date of peak SWE of May 1 and the maximum peak SWE of 647 mm. In addition, we see the creation of cluster 7, located along the eastern section of the Colorado Basin from central Colorado to Northern Wyoming, which has a peak SWE in April but a relatively low cumulative SWE – indicative of colder, drier regions. Average peak SWE for clusters 1 thru 9 occur approximately on

March 15, March 10, March 1, April 1, April 1, April 1, April 24, May 1, April 14, respectively (Figure 5.2c).

The finest cluster analysis conducted in this study is a sixteen-cluster (Figure 5.1i) that shows a further division of the niveograph in terms of the clusters range in timing and amount (Figure 5.3) of peak SWE (Figure 5.2 f & g). One cluster has a peak SWE in February, nine clusters with a peak in March and six with a peak in April. Clusters 2, 3 and 5 are nearly identical. Overall, the clusters become more spatially coherent as the level of generalization goes from 4 to the 16 cluster analyses (Figure 5. a, b, c, i).

5.1.1b 25-Year Daily Data

The six-cluster 25-year daily data analysis derived from the same stations (Figure 5.1d), produced very similar results as the 15-year data set (Figure 5.1b). There is 78% similarity based on the percent of stations that cluster to the same cluster from one analysis to another analysis (Table 5.1 and 5.2). Going from the complete 15-year data set to the 25-year data set, cluster 3 becomes more spatially consolidated, while cluster 6 is slightly more dispersed. Changes in clusters 5 and 6 are typically due to a shifting from one of these clusters to cluster 3. Clusters 1 and 4 remained essentially unchanged. Cluster 2 retains stations located in southern/eastern Utah and southern/western Colorado. The change in groupings result in a slightly improved spatial coherence and delineation of clusters in most areas, such as southern Colorado and central Utah, which consists of three different clusters as opposed to four in the 15-year analysis.

5.1.1c Niveograph Descriptor Variables

The spatial distribution of the six-cluster analysis for clusters based on the niveograph descriptor variable peak SWE is illustrated in Figure 5.1e. Peak SWE and cumulative SWE have a high correlation (98%) and produced almost identical clusters, discussion focuses on peak SWE results. Clusters derived from peak SWE are more spatial coherent compared than those identified from the daily data. There is a 50% similarity of peak SWE groupings with daily data groupings (Table 5.1). Cluster 1 defines stations in the southwest comprising of Arizona, New Mexico, and Southern Utah. Cluster 2 incorporates stations from mid and southern Colorado as well as the mid section of Utah. Cluster 3 is exclusively in the northern section of the basin which is Wyoming and Idaho. Cluster 4 is exclusively located in Northern Colorado and eastern Wyoming. Cluster 5 is a more spread out compared to the other clusters with stations in Utah and the San Juan Mountains of Colorado. Cluster 6 dominates Northern Utah.

The niveograph for the peak SWE clusters are less distinct than that of the daily data clusters (Figure 5.2 b). Excluding cluster 1 (Figure 5.2e) which compares well, the range of average peak SWE derived from the peak SWE clusters is 368mm to 475mm, whereas the range for the daily data clusters is 277mm to 580mm.

Figure 5.4 and 5.5 illustrates frequency of occurrence of quantity and timing of peak SWE derived from the daily data and peak SWE analyses. Individual clusters from the daily data analysis are more concentrated in terms of frequency of peak SWE amounts and date of peak SWE, and are more normally distributed. Peak SWE clusters, as seen in the niveograph (Figure 5.2e) and frequency plots are more variable in terms of when peak SWE (Figure 5.4b) and the date of peak SWE (Figure 5.5b) can be expected to occur.

Clusters derived from April 1st SWE (Figure 5.1f) show an even larger spatial coherence than the daily data or peak SWE clusters. So much so, that cluster 5 appears to become a default cluster that contains only a few spatially dispersed stations not grouped in any other cluster (Table 5.2). In general, groupings resemble clusters derived from peak SWE data yet constitute a larger spatially homogenous area. April 1st clusters have a 43% similarity with daily data groupings and a 72% similarity with peak SWE groupings (Table 5.1).

Clustering based on date of peak SWE or length of snow season appears devoid of any significant spatial coherence (not shown). On average, the percentage of stations that grouped to the same cluster as date of peak SWE length of snow season is 30% and 22%, respectively (Table 5.1). General spatial coherence exists for stations comprising the Arizona clusters the Wyoming clusters, while the remaining clusters are spread over the remainder of the Basin.

Clustering derived from the combination of the two descriptor variables peak SWE and length of snow season are essentially a combination of the individual peak SWE and length of snow season maps. The result is more coherent clusters since the analysis based on the variable peak SWE show more coherencies. Clustering based on peak SWE and the length of snow season agrees slightly better with the daily data clusters compared to other descriptor variable analyses.

Clustering based on the 4-descriptors that are used to define the niveograph, are similar to the peak SWE analysis. This may be expected given that peak SWE and cumulative SWE are highly correlated thereby weighing the outcome to reflect the peak SWE groupings.

5.1.1d Physiography

Clustering based on physiography (Figure 5.1g) has a high percent of similarity to the peak SWE, April 1st, and daily data groupings (57%, 64%, and 41% respectively as seen in Table 5.1). As with the April 1st clusters, the large spatial coherence of the physiography groupings result in one cluster (cluster 5) being a default cluster (Table 5.2). These clusters derived from completely different data maintain similar divisions in the Basin, namely, northern/southern Colorado, northern/southern Utah, Wyoming, and Arizona/New Mexico. This similarity is also seen in the niveographs for each of these cluster approaches as well (not shown).

5.1.1e Non-Normalized Data

Applying the SOM method to the daily raw (non-normalized) SNOTEL values, show much less coherency compared to the daily scaled values (Figure 5.1b and 5.1j). The two clusters that to some extent resemble one another between the two analyses are cluster 1 and cluster 4. Cluster 1 of the non-normalized data consists of Arizona and New Mexico, similar to the normalized, but also incorporates stations from the entire basin. Cluster 4 maintains the same locations in the Basin between the two analyses – Northern Colorado and Wyoming.

5.1.2 Political Groupings

Qualitative inspection of the regions defined by political boundaries show little similarities with groups defined by daily data or niveograph variables. The only comparable groupings are in the southern region consisting of Arizona and New Mexico (Figure 2.1 versus 5.1a). For the rest of the Basin, the spatial heterogeneity and natural geographic divisions are not reflected in groupings based on political divisions. Figure 5.2a and Figure 5.2d illustrates the SOM's ability to divide stations that result in an equally distinct niveograph, and that of politically based divisions which do not capture natural hydroclimate divisions of stations.

5.1.3 Wards Minimum Variance

Quantitatively, Ward's Minimum Variance method of clustering closely resembles the SOM approach in grouping the snowpack surrogate variables. For instance, the peak SWE clusters based on Ward's have a 77% similarity with peak SWE clusters based on the SOM yet only a 46% similarity with daily data clusters (Table 5.1). Qualitatively, however, the Ward's method creates clusters that have a larger spatial coherence (Figure 5.1h), resembling more the SOM derived clusters using April 1st and physiography data. For example, the Ward's method considers the San Juan Mountains in southern Colorado as a homogeneous group, whereas the SOM methodology divides this region into groups.

The Ward's method of clustering the variable length-of-snow season has the lowest overall percent of comparability (22%) with other analyses (Table 5.1). Clustering the variable date-of-peak SWE compares similarly as that as the SOM method (30% vs. 28%).

5.2 Regressions

5.2.1 Regression Fit

Table 5.3 summarizes the coefficient of determination (R^2) values resulting from regressing peak SWE as the dependent variable and physiography as the independent variable for clusters derived from daily data, peak SWE, and physiography. For these analyses, average predictability improves as the clusters progress from one cluster (all data) to coarse (4 cluster), to medium (6 cluster), to finer (16 cluster) resolution with R^2 values of 0.38, 0.37, 0.40, 0.70, respectively (Table 5.3 and Figure 5.6). On average, R^2 values are consistently higher for clusters derived from physiography, followed by peak SWE, and then daily data (Figure5.7).

Average R^2 results from the peak SWE clusters are 0.45 for the 4-cluster, 0.53 for the 6-cluster (Table 5.7), and 0.74 for the 16-cluster level of generalization. Average R^2 results from clusters derived from physiography are 0.59 for the 4-cluster, 0.61 for the 6-cluster, and 0.76 for the 16-cluster level of generalization.

5.2.2 Physiographic Parameters Identified

Regardless of cluster size or data used to derive cluster, results are consistent in terms of the independent variables incorporated into the regression model (Table 5.4). Primary predictor variables are chosen based on the number of occurrences (the top five summarized in Table 5.4), in order are elevation, SW barrier, regional northness, SW shield, and W barrier. Elevation is the first or second parameter incorporated in the regression model in seven of the nine analyses. Regional northness, i.e. aspect, is the first or second parameter in five analyses. Southwest barrier height proves to be consistently important (Table 5.5), with the variable being incorporated as a primary predictor in all but one model and is the top predictor in two analyses. In terms of relative importance, primary predictors in the regression models are elevation, SW barrier height, regional northness, SW shield height, and W barrier height.

Table 5.1. Comparison of each six cluster analysis. Percent of stations that grouped to the same cluster number between analyses. Highlighted entry indicates a similarity greater than 50%. 25-yrs is analysis using 25 year data set. Peak is analysis using annual variable peak SWE. April 1^{st} is analysis using April 1st SWE. [W] is analysis using Ward's Minimum Variance clustering methodology. Physiography is analysis using physiographic variables to derive clusters. Pk/Lnth is analysis using peak SWE and length of snow season. Date is analysis using date of peak SWE. Length is analysis using variable length of snow season. 4 variables is analysis using peak SWE, cumulative SWE, date of peak SWE, length of snow season.

	25-yrs	peak SWE	April 1st	4 variables	physiography	pk/lnth	length	date	peak [W]	length [W]	date [W]
daily [15-yrs]	0.78	0.50	0.43	0.50	0.41	0.29	0.26	0.26	0.46	0.23	0.34
daily [25-yrs]		0.49	0.47	0.52	0.45	0.33	0.30	0.25	0.47	0.21	0.36
peak SWE			0.72	0.71	0.57	0.44	0.31	0.35	0.77	0.25	0.35
April 1st				0.71	0.51	0.38	0.27	0.33	0.65	0.23	0.34
4 variables					0.54	0.34	0.27	0.31	0.67	0.24	0.27
physiography						0.34	0.25	0.21	0.66	0.20	0.29
pk/lnth							0.32	0.34	0.43	0.19	0.27
length								0.31	0.28	0.22	0.25
date									0.25	0.20	0.25
peak [W]										0.24	0.34
length [W]											0.25

Table 5.2. Number of stations per cluster for six cluster analysis.

	cluster 1	cluster 2	cluster 3	cluster 4	cluster 5	cluster 6
SOM: daily [15 years]	24	28	61	48	28	27
SOM: daily [25 years]	22	43	26	41	35	49
SOM: peak SWE	24	29	38	42	42	41
SOM: April 1st	16	58	41	43	16	42
SOM: four descriptor variables	34	29	34	46	40	33
SOM: physiography	26	37	40	45	11	57
SOM: peak SWE & length of season	26	43	37	44	37	29
SOM: length of season	23	40	37	39	36	41
SOM: date of peak SWE	43	21	42	28	50	32
Wards peak SWE	18	36	37	43	27	55
Wards length of season	62	21	28	40	37	28
Wards date of peak SWE	13	15	53	55	41	39

Table 5.3. Summary of R^2 values for each cluster analysis performed.

	data						
# of clusters	daily	peak SWE	physiography	Mean	Max	Min	Median
1	X			0.38	0.53	0.20	0.40
4	X			0.37	0.75	0.10	0.36
4		X		0.45	0.66	0.16	0.45
4			X	0.59	0.86	0.37	0.59
6	X			0.40	0.76	0.01	0.40
6		X		0.53	0.80	0.19	0.53
6			X	0.61	0.92	0.25	0.61
16	X			0.70	0.94	0.30	0.70
16		X		0.74	0.98	0.53	0.71
16			X	0.76	0.94	0.46	0.75

Table 5.4. List of top five independent variables used in each regression. In parenthesis is number of times variable was used in the regression model.

	data							
# of clusters	daily	peak SWE	physiography	1	2	3	4	5
1	X			sw barrier (15)	w barrier h (14)	regional north (13)	Z (12)	Y (9)
4	X			sw barrier h (30)	regional north (21)	regional slope (19)	sw shield h (18)	forest density (18)
4		X		local north (34)	sw barrier h (33)	nw barrier h (26)	w barrier h (23)	regional slope (21)
4			X	Z (56)	regional north (50)	sw barrier h (42)	regional slope (34)	local north (22)
6	X			sw shield h (34)	Z (27)	local east (23)	local north (23)	sw barrier h (22)
6		X		Z (47)	w barrier h (43)	regional north (42)	sw barrier h (42)	sw shield h (38)
6			X	regional north (61)	Z (58)	sw barrier h (44)	nw barrier h (43)	sw shield h (31)
16	X			sw barrier h (70)	Z (52)	local east (52)	regional north (51)	forest density (39)
16		X		Z (82)	regional north (71)	sw barrier h (61)	local north (54)	w barrier h (53)
16			X	Z (81)	regional north (62)	w barrier h (61)	nw dist ocean (57)	forest density (57)

Table 5.5. Summary of all independent variables used in regression models in table 5.4. Number under each cluster size and method used to derive cluster indicate amount of times each variable (left column) was used in regression equation. Fraction (right column) indicates relative importance considering all regression equations.

variable	no cluster	4 cluster			6 cluster			16 cluster			
		daily data	peak SWE	physiography	daily data	peak SWE	physiography	daily data	peak SWE	physiography	fraction
number of equations	15	60			90			240			-
Z	12	15	19	56	27	47	58	52	82	81	0.38
Y	9	17	12	19	18	27	26	36	35	49	0.21
X	2	11	2	18	11	11	7	31	10	21	0.10
W shield	0	6	8	9	11	3	28	26	37	34	0.14
W distance	9	8	6	8	4	9	1	30	11	19	0.09
W barrier	14	13	23	11	9	43	13	28	53	61	0.23
SW shield	6	30	21	8	34	38	31	33	45	49	0.25
SW distance	0	5	15	2	9	7	6	23	30	21	0.10
SW barrier	15	25	33	42	22	42	44	70	61	52	0.34
slope	1	19	7	4	10	5	20	32	29	22	0.13
regional slope	8	14	21	34	15	26	26	32	39	33	0.21
regional northness	13	19	14	50	14	42	61	51	71	62	0.34
NW shield	2	9	13	8	21	7	9	15	45	33	0.14
NW distance	0	10	10	2	6	3	2	10	9	57	0.09
NW barrier	6	5	26	15	14	15	43	31	37	19	0.18
local northness	0	6	34	22	23	27	8	33	54	33	0.20
local eastness	0	4	5	9	23	18	8	52	41	41	0.17
forest density	2	32	8	14	5	21	9	39	33	53	0.18

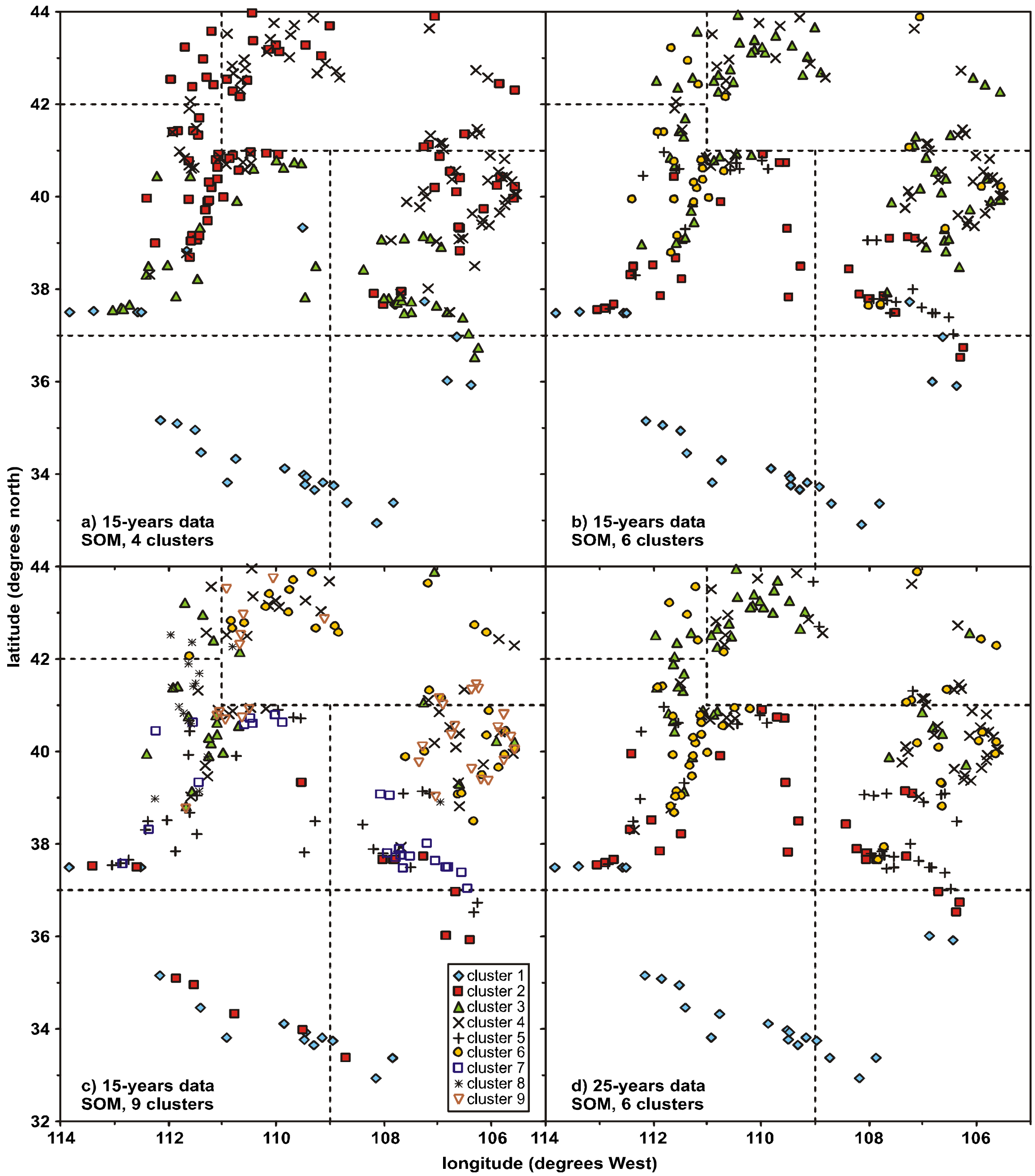

Figure 5.1. Spatial distribution of stations for clusters derived from daily SWE values for a) four cluster analysis derived from daily SWE values b) "Base-case" six cluster analysis derived from daily SWE values c) nine cluster analysis derived from daily SWE values, and d) six cluster analysis from 25-year period (1981-2005) derived from daily SWE values with 10% missing data values.

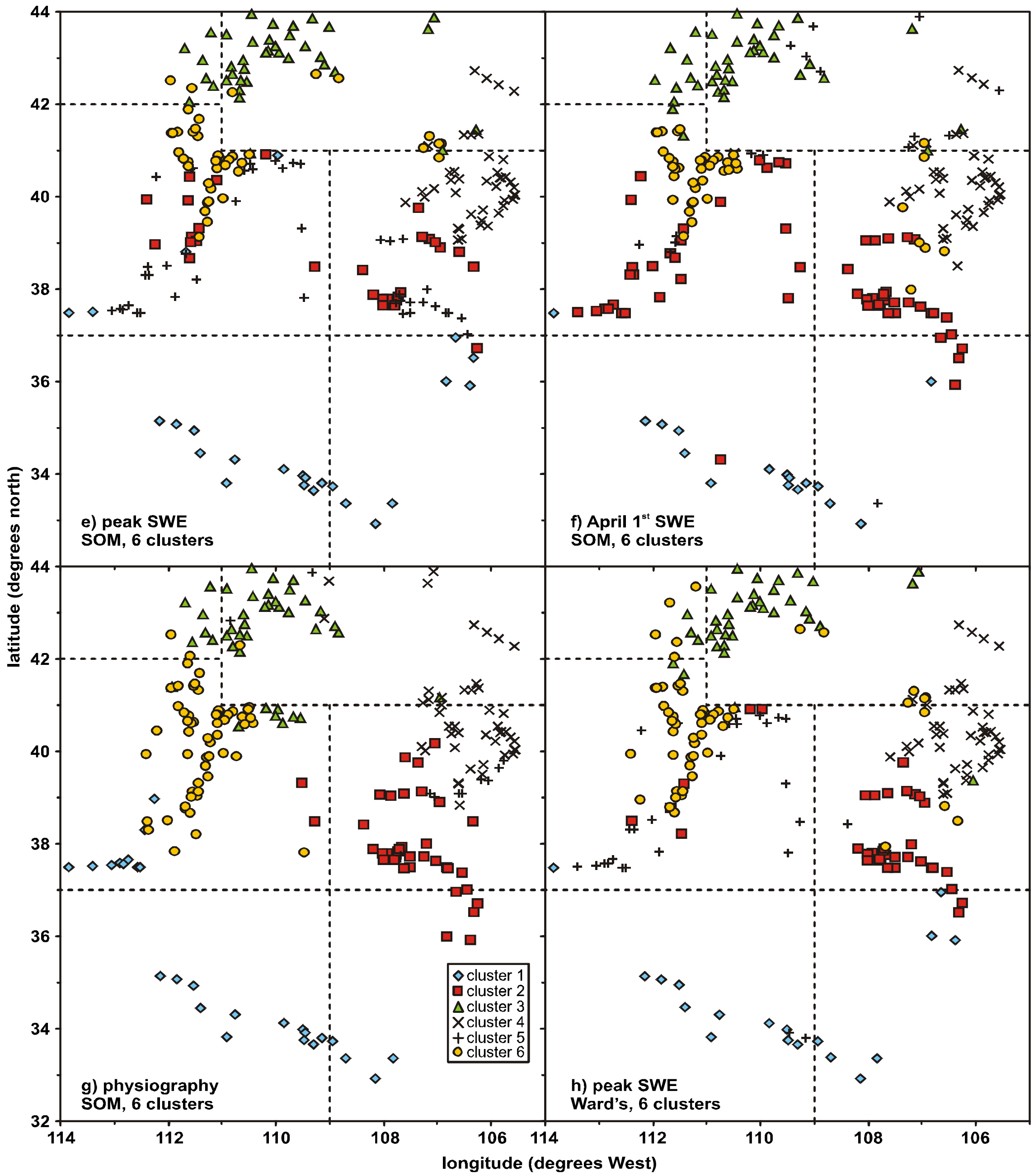

Figure 5.1. Spatial distribution of stations for e) six cluster analysis derived from peak SWE values f) six cluster analysis derived from April 1st SWE values g) six cluster analysis derived from physiography, and h) six cluster analysis derived from peak SWE values using Ward's Minimum Variance clustering procedure.

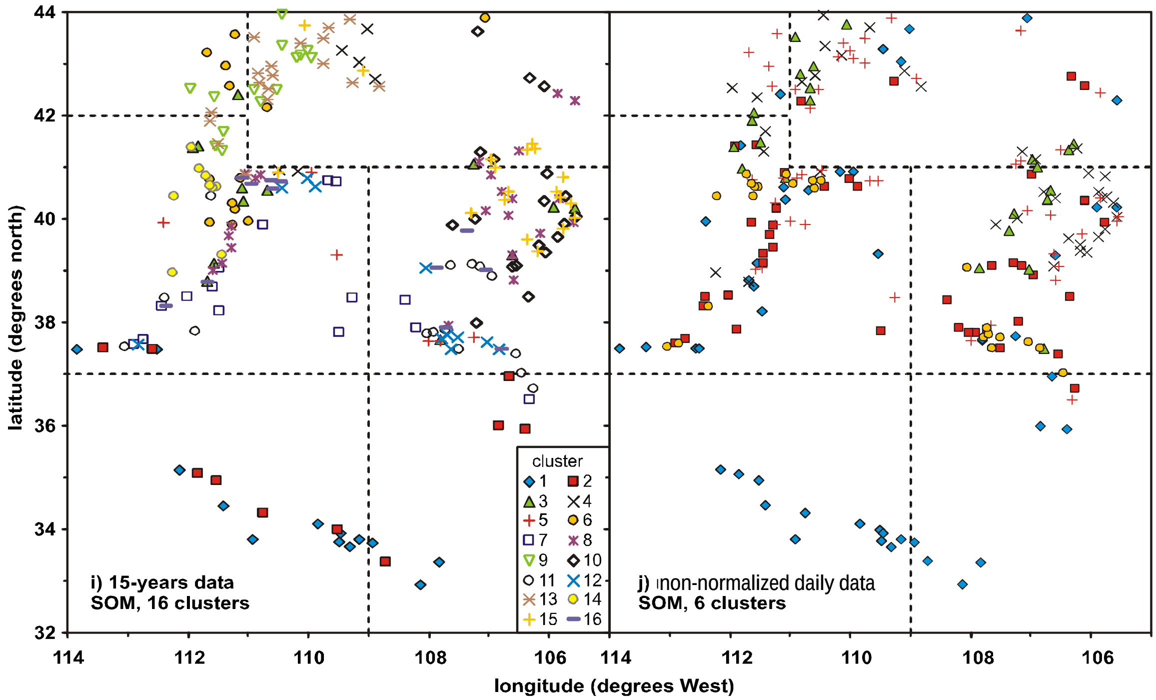

Figure 5.1. Spatial distribution of stations for i) sixteen cluster analysis derived from daily SWE values, and j) six cluster analysis derived from raw (non-normalized) daily SWE values.

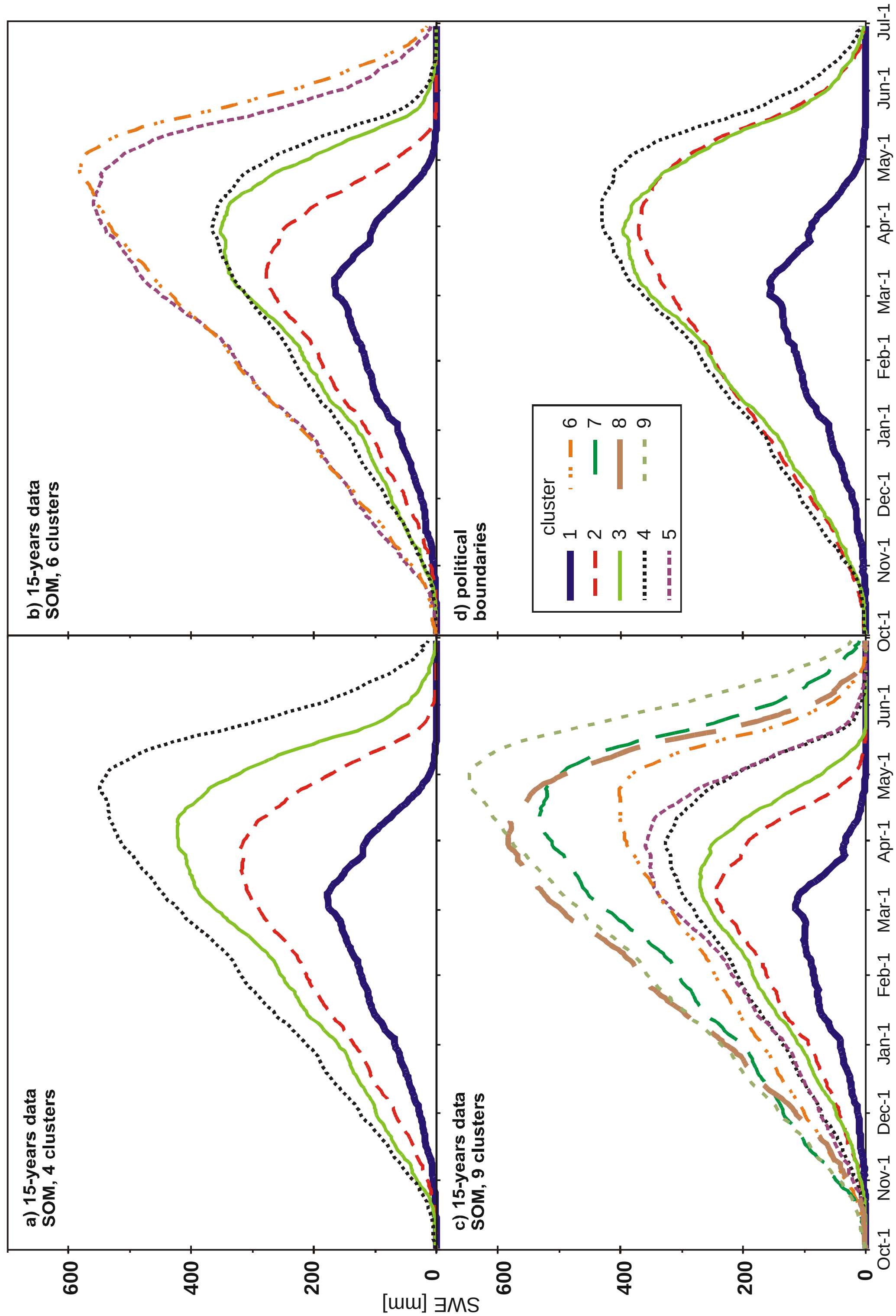

Figure 5.2. Fifteen year (1991-2005) average daily SWE for all stations within each cluster, for a) four-cluster analysis based on daily SWE values b) six-cluster analysis based on daily SWE values c) nine-cluster analysis based on daily SWE values and d) four-cluster state-based groupings based on Serreze *et al.* [1999].

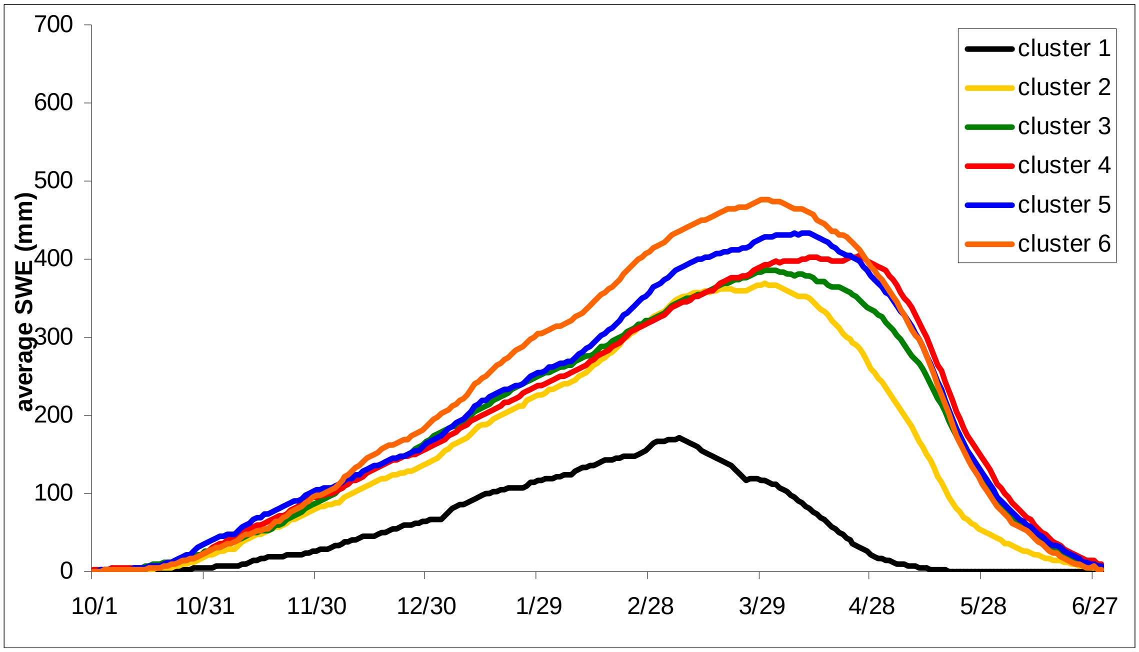

Figure 5.2. e) Fifteen year (1991-2005) average daily SWE for all stations within each cluster. Six-cluster analysis based on peak SWE values.

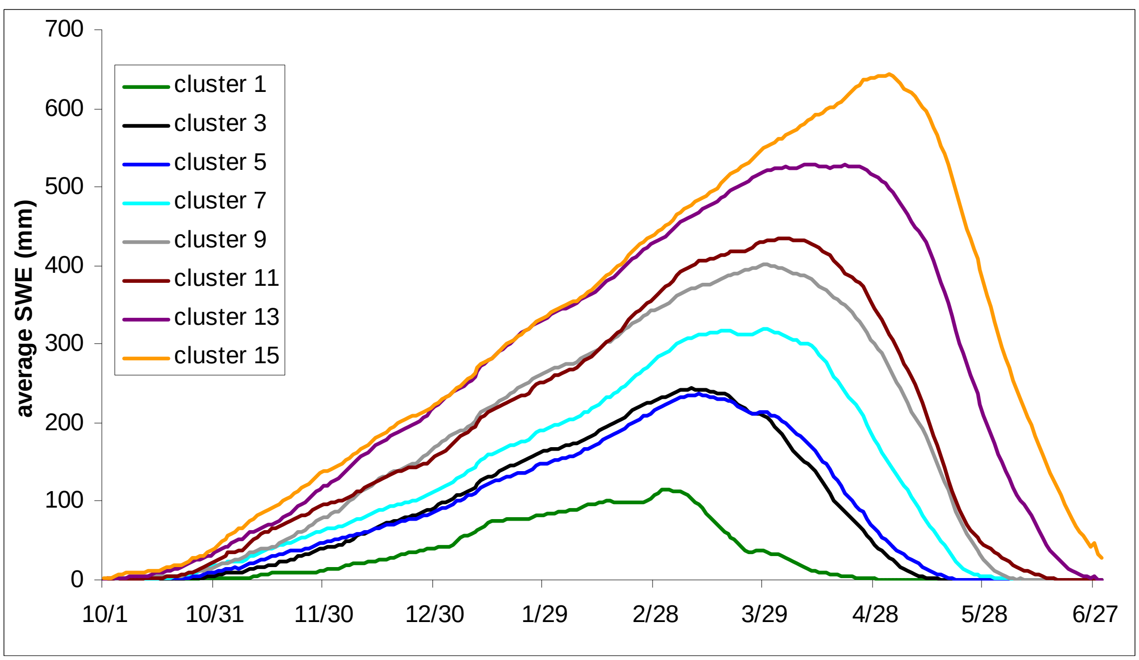

Figure 5.2. f) Fifteen year (1991-2005) average daily SWE for all stations within each cluster. Odd clusters sixteen-cluster based on daily values.

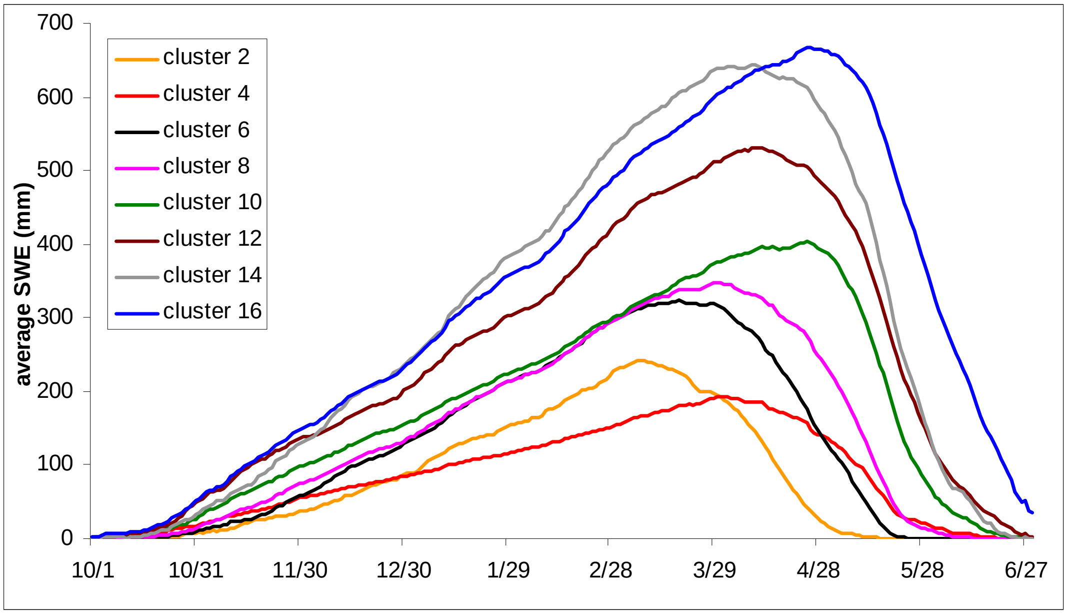

Figure 5.2. g) Fifteen year (1991-2005) average daily SWE for all stations within each cluster. Even clusters sixteen-cluster based on daily values.

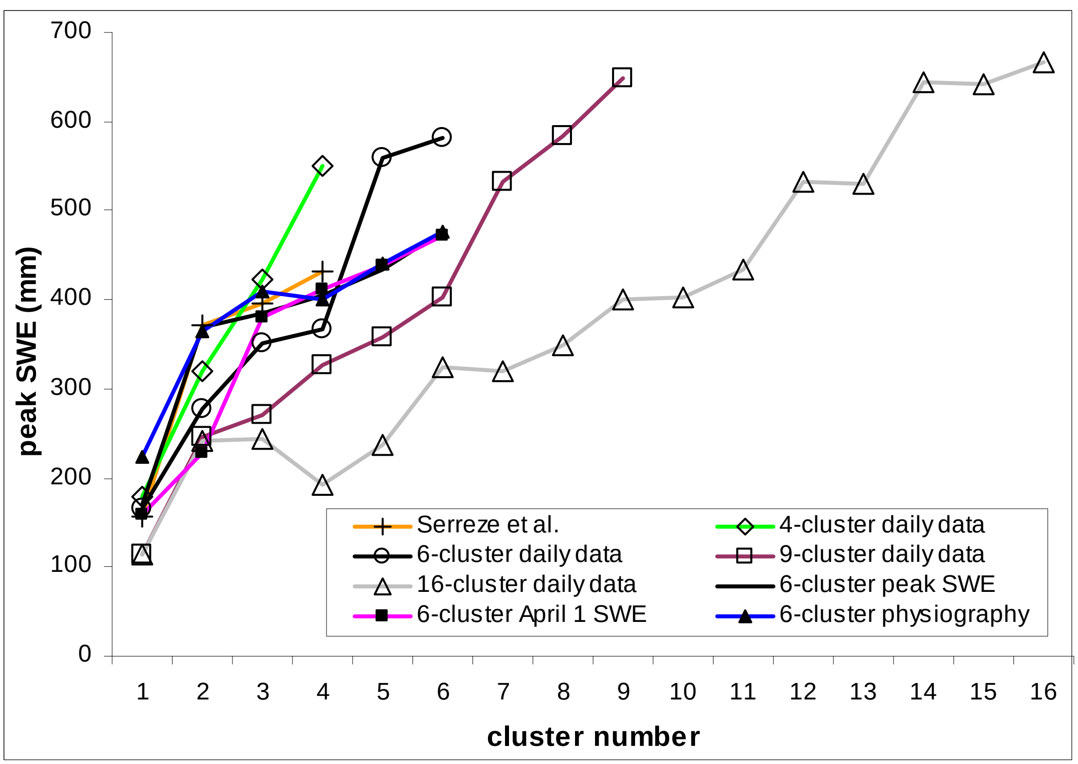

Figure 5.3. Peak SWE for clusters derived from daily data for different size analyses and for 6-cluster analyses derived from descriptor variables.

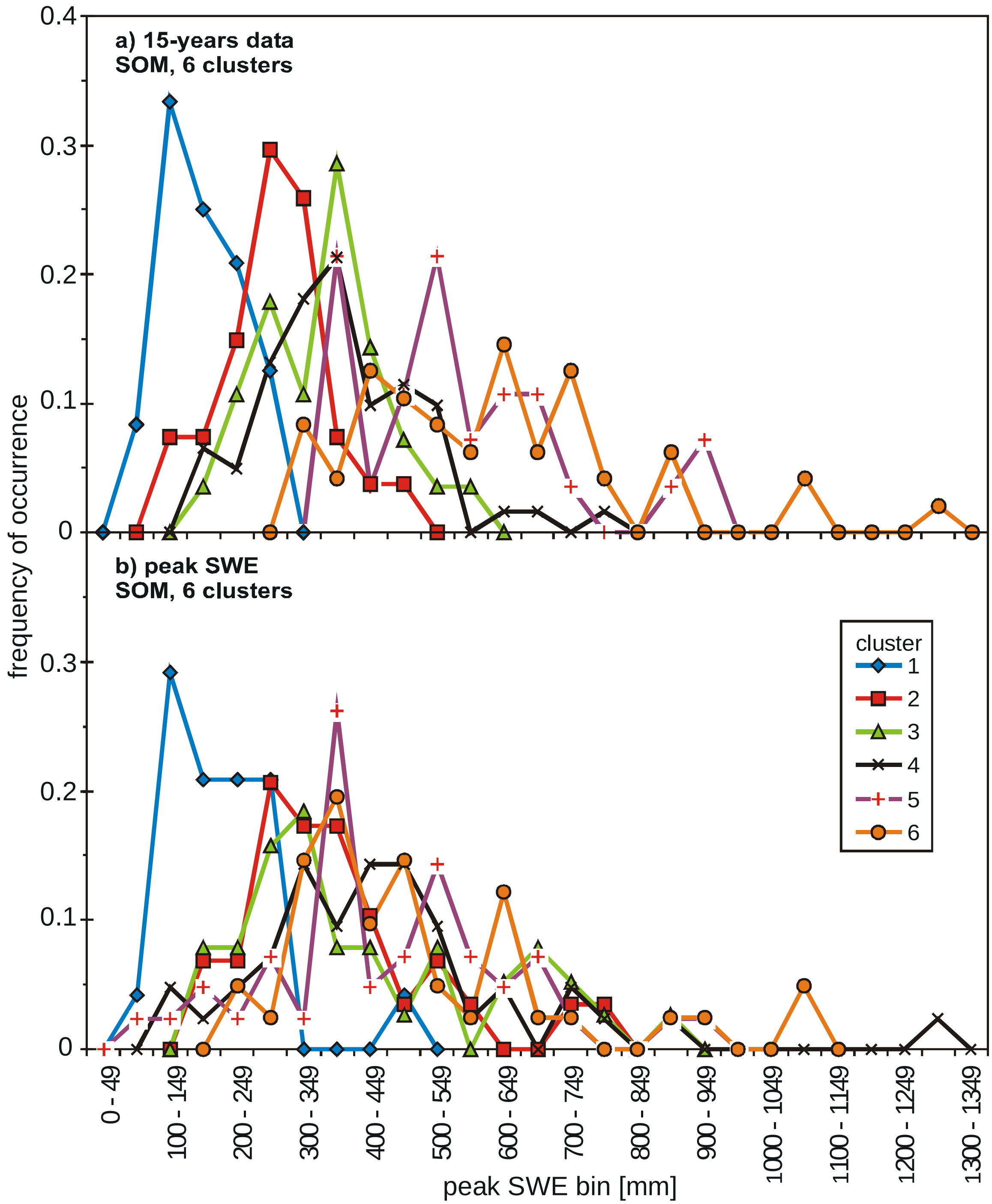

Figure 5.4. Histogram of clusters in six-cluster daily data and peak SWE analysis of frequency of magnitude of peak SWE for each cluster.

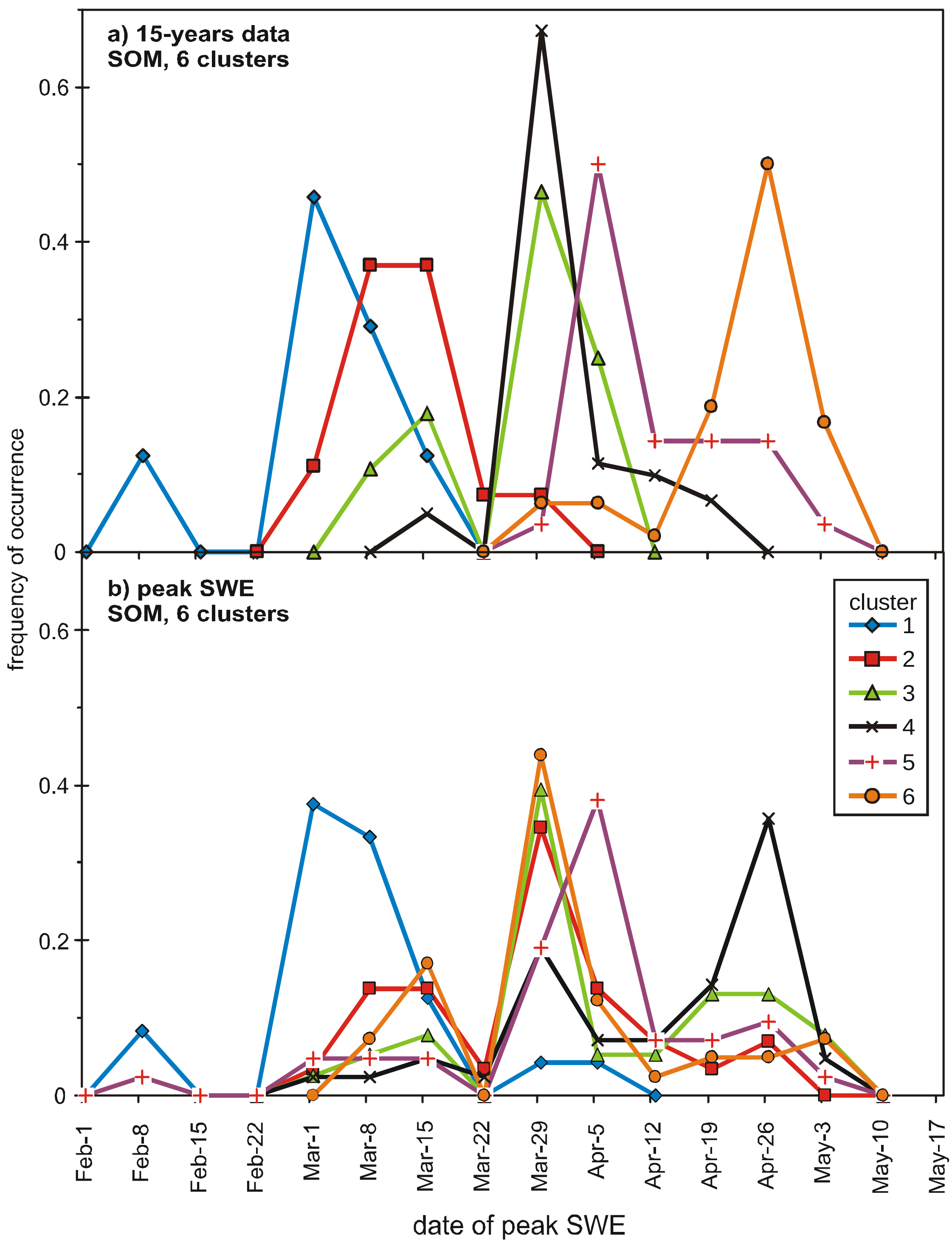

Figure 5.5. Histogram of clusters in six-cluster daily data and peak SWE analysis of frequency of magnitude of date of peak SWE for each cluster.

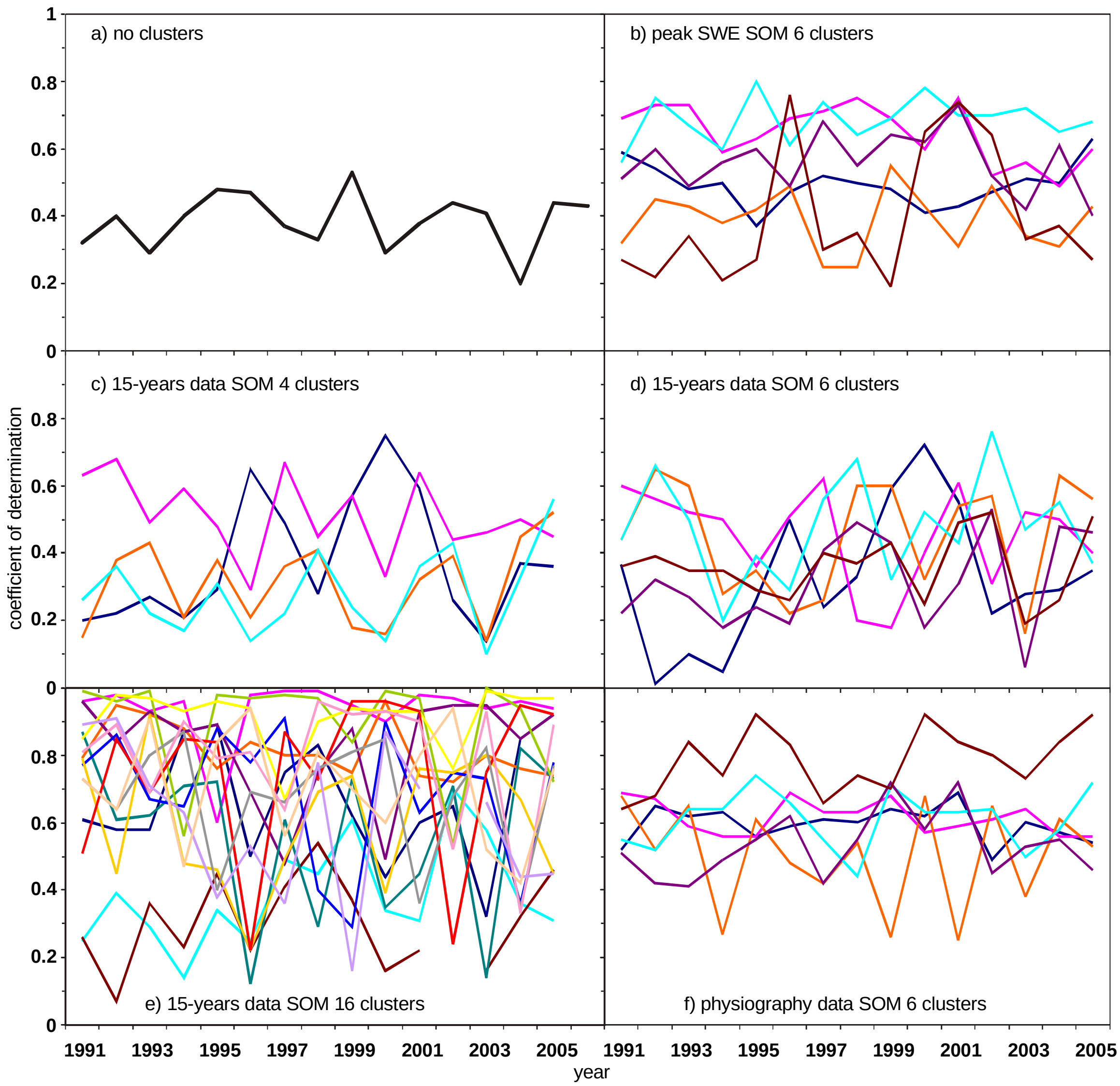

Figure 5.6. Average coefficient of determination for each cluster from 1991-2005 for a) no clusters b) peak SWE six-cluster c) daily data four-cluster d) daily data six-cluster e) daily data sixteen-cluster and f) physiography six-cluster.

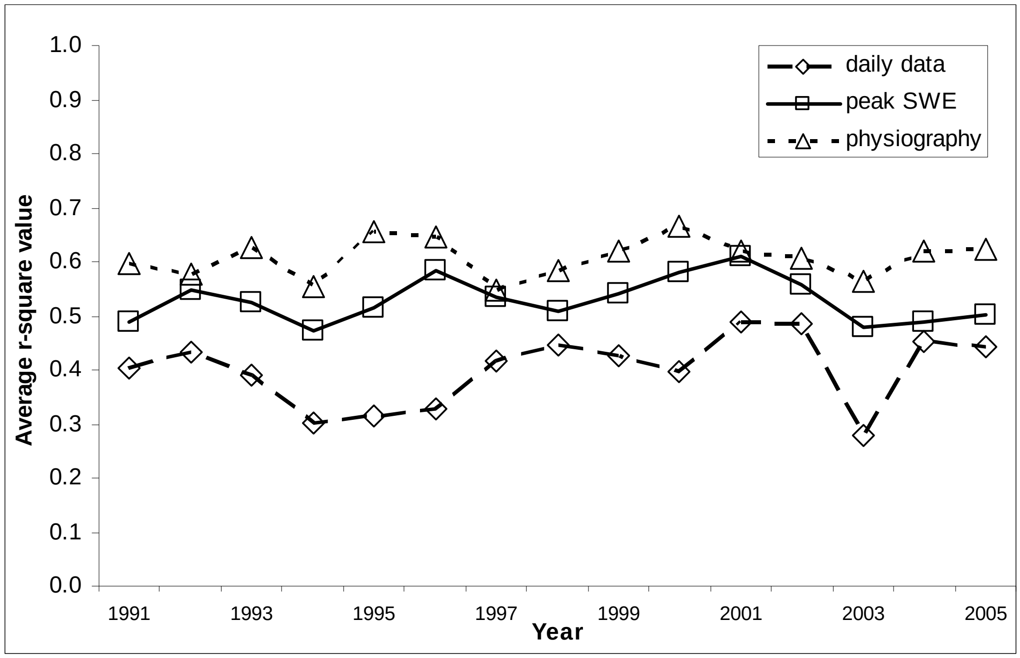

Figure 5.7. Average R^2 value for six-cluster SOM analysis of daily data, peak SWE, and physiography.

CHAPTER 6: DISCUSSION

6.1 Defining Areas of Homogeneity

Results show general similarities exist regardless of data used to derive clusters (i.e. physiography vs. daily SWE) or cluster size. There is a north/south gradient with generally at least one cluster comprising a large area in the mid-latitude region of Utah and Colorado. A consistent dissimilarity exists between northern and southern Colorado with the divide being approximately near the latitude 39º to 40º at the boundary of the headwaters of the Arkansas, Platte, Gunnison-Dolores, and Colorado River Basins. This discontinuity echoes results from previous studies [Changnon *et al.* 1991; Cayan, 1996]. There is also a consistent separation of northern Utah and southern Utah. Arizona and western New Mexico is consistently considered a similar region.

In addition to characterizing a particular region, clusters show a correspondingly distinct variability in terms of deviations from their annual mean peak SWE. For example, cluster 1 located in Arizona and New Mexico is close to one deviation above its mean peak SWE in 1992 (Figure 6.1 and 6.2), whereas the remaining clusters are 0.5 to 1.5 deviations below their mean. Differences exist to some degree between each of the six clusters with stations further apart in a north/south direction tending to have opposite tendencies in terms of patterns of variability.

For the clusters derived from daily data, the niveographs (Figure 5.2) show distinctive mean daily SWE values for all levels of generalization. This illustrates that larger dissimilarities exist where the SOM created a new cluster and that information is subsumed according to the degree of generality imposed. For example, going from the four-cluster to the six-cluster analysis, with more dissimilarity in the northern region, the area was subdivided as seen in the spatial distribution of clusters (Figure 5.1a and 5.1b) and niveographs (Figure 5.2a and 5.2b). At the nine-cluster and sixteen-cluster level of generalization, subdivisions are seen across the entire basin and generally clusters appear to be more defined by smaller-scale rather than more larger-scale regional influences (Figure 5.1c and 5.1i).

The transformation of clusters at the four-cluster resolution to the sixteen-cluster resolution is informative (Figure 5.3). At more generalized resolutions, clustering results are the product of larger dissimilarities of variability between stations, typically stemming from larger-scale controls. At finer resolutions, clusters are defined by finer differences in variance. These finer, smaller magnitudes of variance are typically influenced by local-scale variability yielding more localized clusters. Results in this study reflect those of other studies. Serreze *et al.* [2001] found in moisture limited areas such as Colorado, Wyoming, and Utah, local topographic setting plays an important role in enhancing snowfall. Which may explain the large inhomogeneity in the clustering results. From coarse to finer-scale, clusters develop based on site specifics when continental effects begin to be masked by local topographical influences. For example, the subtle influence of local aspect on SWE may not be apparent at the four-cluster resolution because other, larger scale factors such as SW shield height dominate, but at finer scales the influence of aspect on SWE variability may be more influential.

Using representative niveograph variables to derive clusters has varying results. The best solo descriptor compared to the daily data is peak SWE. Figure 5.1 and Table 5.1 illustrate clustering via peak SWE capturing annual trends (deviations from each clusters mean) of individual clusters. Regionalization based on peak SWE allows stations with the same patterns of snow accumulation to be grouped together due to similar synoptic-scale atmospheric controls such as storm tracks and continental barriers, even if absolute amounts differ and covariance exists due to localized physiographic influences. The Ward's method for clustering peak SWE results in a larger spatial coherence than the peak SWE SOM, since stations that comprise a cluster are within close proximity to each other. While the Ward's clustering has a 77% similarity to peak SWE SOM, it does not capture some of its subtleties.

The length of snow season groupings do not represent daily data or peak SWE groupings well based on percent of similarity (Table 5.1). This result is surprising since the number of days a station has snow is typically indicative of it's location in terms of the temperature of warmer Arizona or colder Wyoming, i.e., or high/low accumulation areas. In addition to this, Serreze *et al.* [2001] found a high correlation between annual precipitation in the form of snow and number of snow days (80% or higher at over 70% of the stations).

Additionally, groupings of date of peak SWE are not a considerable improvement compared to length of snow season. This may occur since the data are scaled prior to being subjected to the SOM or Ward's program, deviations from a station's average is relative to that particular station. Using all four niveograph descriptor variables in the analysis improves the representativeness in terms of percent similarity with classifications based on daily data, except for peak SWE, over a single descriptor variable. Thus peak SWE better represents the daily data clustering than other niveograph descriptor variables or combination of variables.

When regressions with SWE are based on clusters derived from physiographic variables, the correlation in the regressions can be excellent (Figure 5.1 and Figure 5.6). For example, at the 6-cluster level of generalization, classifications defined by physiography represent daily data, peak SWE, and April 1st SWE classifications well (Table 5.1). Physiography is comparable to April 1st data in terms of representing classifications derived from daily values based on percent of similarity and spatial location of climatologies. The nature of the physical variables used in the study is a combination of smaller and larger-scale attributes capturing the relative combination of scale needed to represent April 1st SWE. This result may be useful in data sparse regions of the world since physiographic variables and their relationship with SWE are significant, as seen by Solomon *et al.* [1968], Daly *et al.* [1997], and Fassnacht *et al.* [2003].

Overall, clusters based on daily values, peak SWE, April 1st, physiography, and the four descriptors show a good relationship with each other. The results of the peak SWE analysis show less coherency of clusters compared to April 1st SWE and physiography. This stronger spatial coherency of clusters may occur since April 1st SWE and physiography are fixed in time. Also, snow accumulation and ablation information is not utilized when one niveograph descriptor variable is selected to integrate the snow season, illustrating larger scale influences as opposed to more localized influences from the daily data.

Comparing clusters from daily data vs. April 1st SWE may be more clearly understood by examining results from the SOM descriptor variables. From SNOTEL data Serreze *et al.* [2001] found large snowfall events have a relatively low coherence frequency in the Colorado Basin, while Cayan [1996] showed April 1st SWE covaried on larger regional scales, although still maintaining some localized spatial SWE correlations.

6.1.1 San Juan Mountain Region

To illustrate the temporal and spatial variability of peak SWE within the Colorado Basin, the 25-years of standardized peak SWE from the 216 Snow Telemetry (SNOTEL) sites were plotted in Figure 6.3. For any given year, there is a wide range of values the Basin exhibits as far as what constitutes a wet, dry, or normal year. For example, 1983 was a very wet year in most areas [Rhodes *et al.* 1984] but with some stations being three deviations above their mean, others were as much as one standard deviation below their mean. This lack of homogeneity is apparent even with stations in close proximity to each other. To examine this on a more localized scale, seven SNOTEL sites in the San Juan Mountains of Southwestern Colorado, in and around the headwaters of the Rio Grande Basin (Figure 6.4) were examined. In 1993 the Upper Rio Grande SNOTEL station was close to one deviation above its mean while nearby stations Slumgullion, Wolf Creek, and Beartown were below their mean (Figure 6.5). The difference is even greater in 1996, when the peak SWE at the Upper Rio Grande site was one standard deviation below its mean and the other three stations were one deviation above their mean.

At the small scale, the SOM technique of clustering daily data showed that a particular SNOTEL station could cluster to a seemingly inappropriate climatology, reflecting local-scale influences on that station. For example, for the 4-clustering of daily data the Upper Rio Grande station groups to stations in Arizona (cluster 1) whereas Slumgullion and Wolf Creek clustered to group four and the others to group three. However, for the 6-clusters, all stations but Upper Rio Grande (cluster 1) grouped to cluster 5 (Table 6.1). The Upper Rio Grande station is at a lower elevation (Table 6.1) and all others are 1150 to 1870 ft higher. Hence, clusters derived from daily-time steps better reflect the local-scale environment, which appear to dominate over larger-scale influences.

The peak SWE descriptor variable is reasonably spatially defined yet reveals some of the subtleties seen in the daily data clustering. In the San Juan Mountains (Figure 6.4 and 5.1e), stations east of the Continental Divide, including the Upper Rio Grande station, group to cluster 5 (average peak SWE 497.7 mm), while stations west of the divide group to cluster 2 (average peak SWE 426.9 mm). Larger scale influences, such as atmospheric patterns, position relative to moisture sources and mountain barriers, rather than local-scale influences and likely dominate variability of peak SWE for a particular station. Similarly, Molotch and Bales [2006] found SNOTEL stations in the Rio Grande headwaters over represent the western barrier distance, which is the Continental Divide.

Compared to other descriptor variables, clusters derived from April 1st and physiography result in the San Juan Mountain region being considered a homogeneous region where southern Colorado, northern New Mexico, and southern Utah is a single climatology. This result may be intuitive if one considers that April 1st SWE varies in space but is fixed in time, peak SWE varies in space and time while physiography does not vary. The differences between datasets are important, since different clustering results may occur in what is considered a like region.

6.1.2 Missing Data

Clusters derived from a complete 15-year complete data set show a 78% similarity to clusters derived from a 10% missing 25-year data set (Table 5.1). Hewitson and Crane [2003] also showed that SOM's could tolerate missing data. To identify temporal modes of precipitation records with improvements seen when up to 80% of the data are missing from 80% of stations. For this data set, the improved coherency of the 25-year analysis (Figure 5.1) indicate that with more data, clusters could potentially take on an even greater coherency and accuracy. However, the high degree of similarity between the two analyses implies SOMs can be used on data records of relatively shorter duration and still adequately identify areas of homogeneity.

6.1.3 Comparison with Political Boundaries

Classifications based on daily data offer contrasting results to classifications of political boundaries as defined by Serreze *et al.* [1999]. The SOM methodology groups like stations regardless of location, whereas political divisions did not acknowledge natural divisions. This is evident in inspecting the spatial distribution of clusters based on daily data, peak SWE, and April 1st (Figure 5.1a, e and f). Divisions between northern/southern Colorado, northern/southern Utah, and some of the similarities of southern Utah/Colorado, northern Colorado/Wyoming were not recognized from political boundaries (Figure 2.1). Niveographs of individual clusters (Figures 5.2a and d) illustrate how the SOM methodology creates groupings that are equally divided in terms of each cluster's average daily SWE. The SOM captures a distinct peak SWE range of 178 mm, 320 mm, 423 mm, 550 mm, for clusters 1 thru 4**;** while political boundary classifications yielded less variation at 156 mm, 371 mm, 396 mm, and 431 mm (Figure 5.2a and d). The latter combined dissimilar stations that yielded more averaged niveographs. Local and regional uniqueness is lost when groupings based on political boundaries.

6.2 Multivariate Regression Analysis

6.2.1 Regression Coefficient of Determination

Regression results for all analyses show overall R^2 values increased as the number of clusters increased from the four to sixteen (Table 5.3). While the predictive ability improved for some years, the R^2 values are generally low until there are 16 clusters. Considerable scatter exists for the 16-cluster based on daily time-steps, but only 2 of the 16 clusters have an overall average R^2 of 0.5 or less (Figure 5.6). The increase in predictability can be attributed to an increase in sub-divisions put stations of higher similarity in a single cluster and; since fewer stations comprise a group, it was simpler for one or more independent variables to explain variability in the dependent variables.

Improvement of predictive ability from increasing the number of clusters is seen in the southern extent of the study area. Six clustering produces cluster 1 consisting of stations mainly in Arizona and New Mexico (Figure 5.1b) with an associated R^2 of 0.33

(Figure 5.6), yet a 16-cluster subdivides this into clusters 1 and 2 (Figure 5.1i), resulting in an average R^2 of 0.68 and 0.78, respectively (Figure 5.6).

6.2.2 Physical Parameters Identified

The regression models identified the physiographic variables that influence SWE. Physiographic variables can be used to estimate SWE, but due to the different scales of variability in climate that influence SWE, such a relationship is not strong for the entire Colorado Basin or even at smaller scales e.g., Molotch and Bales, 2006. For example, for the 16-cluster derived from daily-values (Figure 5.1i and Figure 5.2f and g), cluster 14 (located almost exclusively around the Great Salt Lake in Utah) identified elevation, latitude, and aspect as principal variables and had an average R^2 value of 0.91. This possibly suggests relative positioning to a considerable moisture source (Great Salt Lake, Pacific storm tracks) as a key determinant of SWE for this area. Another example is cluster 4 in Wyoming, a colder dryer area that usually has a low peak SWE occurring in March or April where the model identified southwest barrier height, shield height, (i.e. natural barriers in direction dominate storm tracks) and regional slope as predictors of SWE with an average R^2 of 0.94. Similarly, cluster 10, located mostly in the northern half of Colorado and Wyoming, has southwest barrier height, aspect and elevation as predictors with an average R^2 of 0.57. Hence, physiographic influence on SWE can vary depending on many factors and at many scales which impacts the prediction of SWE.

Regression results for groupings derived from peak SWE are different than the daily data groupings since peak SWE clustering typically have a higher spatial coherence, due to larger-scale controls on peak SWE. In the 16-cluster analysis derived from peak

SWE (not shown), cluster 16 in the southern, warmer part of the basin, identifies elevation and western shield height as primary predictors. Alternatively, in northern Colorado (cluster 1) local aspect, and regional aspect, and western barrier height are the main predictors. SWE at SNOTEL stations located in warmer (lower and/or southerly) regions are controlled primarily by elevation, whereas stations that are located at cooler (higher elevations and/or more northerly) regions are controlled by factors such as location relative to large scale weather systems.

6.2.3 Cross-Correlation of Variables

Clustering based on physiography and peak SWE show better correlation in the regression models than from daily data clusters. This may occur due to over-fitting of models since the same variables are used as independent variables as those used to defined the clusters. The total number of independent variables used in each regression analysis of the three different cluster types is shown in Table 6.2. The peak SWE cluster analyses contain on average 18% more independent variables in the regressions than the daily data, while clusters derived from physiography contain on average 24% more independent variables (Table 6.2). Variability is better explained when more independent variables included in regression model.

Cross-correlation of independent variables may result in over-fitting the regression model. Higher than realistic R^2 value can result when two or more independent variables are highly correlated in a multivariate regression model. To reduce this cross-correlation, only the first variable that is cross-correlated by 0.70 or higher was included in the regression.

6.3 Summary

As water resources become increasingly scarce, better defining snow climatologies can improve estimations of the timing and amount of snowmelt discharge. Modeling the relationships between snowpack accumulation/ablation and physiography, atmospheric circulation patterns, and larger-scale climate processes can also benefit.

The ability of the Self-Organizing Map methodology has seen limited use with large, multifaceted, incomplete, multidimensional datasets where it is not sufficient to simply average available station records within a region. This is often the case [Crane and Hewitson, 2003] in attempts to capture and understand complex and subtle processes.

The amount of information garnered from a SOM analysis depends on the level of generality imposed and the temporal resolution of the input data. The SOMs ability to capture the daily variability of SNOTEL stations to derive areas of homogeneity offers the opportunity to observe complex and subtle interrelationships. Regressing physiography as the independent variable with peak SWE as the dependent variable shows there are dominate physical influences on SWE. The relative importance of these variables changes with scale and location.

The optimal niveograph surrogate variable that best represents the variability of the daily data is peak SWE. Peak SWE varies annually in timing and amount, but can be determined from the SNOTEL time series. The controls on peak SWE, such as, large scale weather patterns, position relative to a moisture source and natural barriers, are better reflected in the peak SWE climatologies than by other surrogate variables (Table 5.1).

Climatologies based on physiography, April 1st SWE, or the Ward's Minimum Variance method have a larger spatial coherence compared to peak SWE. However, results based on physiography are promising given that at certain scales physiography can represent daily SWE variability better than April 1st SWE or Ward's Minimum Variance. Nonetheless, daily data provides more information than single or multiple surrogate variables as they integrate the snow season. The approach used to establish climatologies in future investigations should depend on data availability, level of generalization desired, and goals of the researcher.

Table 6.1. SNOTEL station cluster designation in the Upper Rio Grande Basin from different cluster analyses.

				analysis method								
				SOM								Ward's
	elevation		political	daily data (15 years)				25 years	peak SWE	April 1st	physiography	peak
station	[ft]	# clusters	4	4	6	9	16	6	6	6	6	6
Upper San Juan	10,130		7	3	5	7	12	5	5	2	2	2
Wolf Creek	11,000		7	4	5	7	16	5	5	2	2	2
Upper Rio Grande	9,400		7	1	1	2	5	2	5	2	2	2
Middle Creek	11,250		7	3	5	7	12	5	5	2	2	2
Slumgullion	11,440		7	4	5	7	10	5	5	6	2	2
Beartown	11,600		7	3	5	7	12	5	5	2	2	2
Lily Pond	11,000		7	3	5	7	11	5	5	2	2	2

Table 6.2. Total number of independent variables used in regression analysis.

	daily-data	peak SWE	physiography
4 cluster	243	277	331
6 cluster	276	391	400
16 cluster	624	722	740

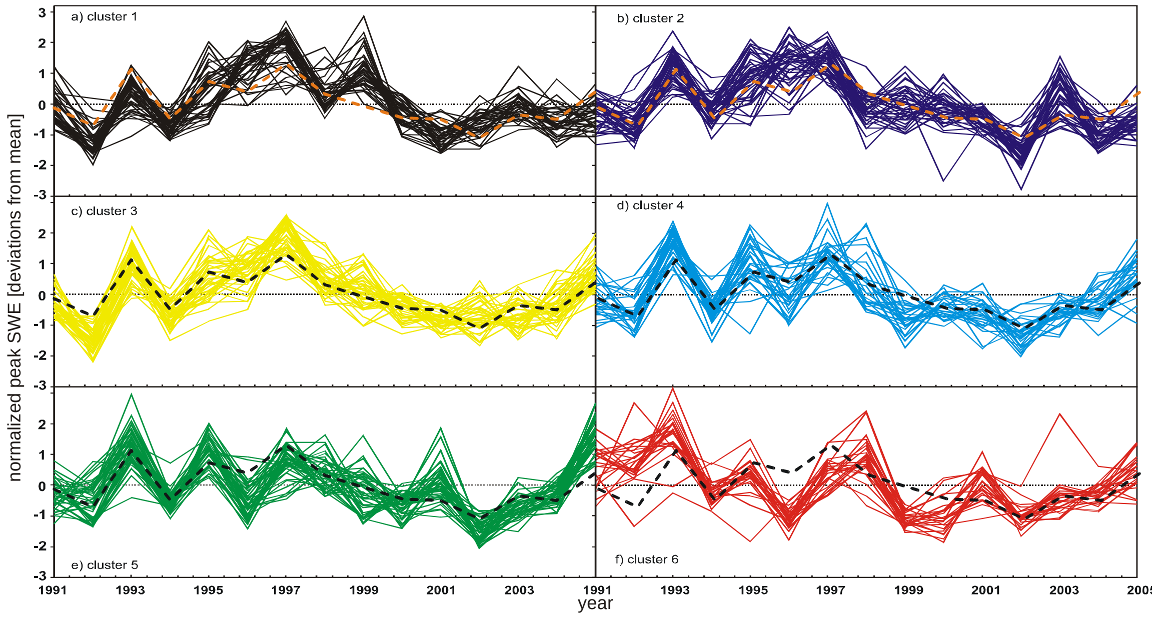

Figure 6.1. SNOTEL stations standardized about its mean with a standard deviation of one of annual peak SWE from 1991 – 2005. Results for individual clusters for SOM 6 cluster analysis derived from daily data. Dashed line resembles the overall mean.

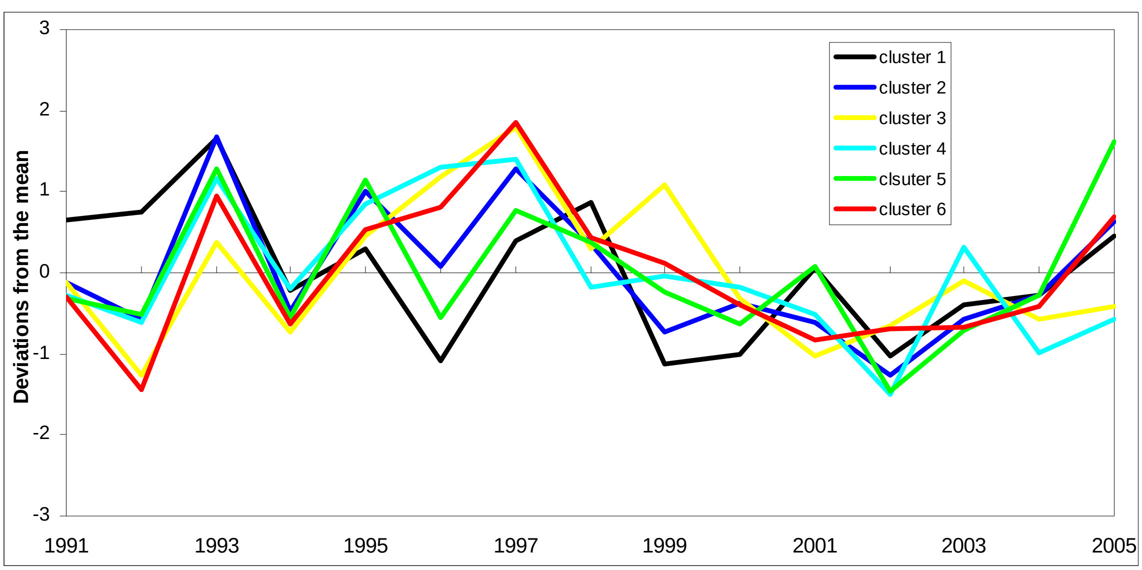

Figure 6.2. Annual average peak SWE of SNOTEL stations standardized with a mean of zero and standard deviation of one for each cluster of the six cluster analysis derived from daily values.

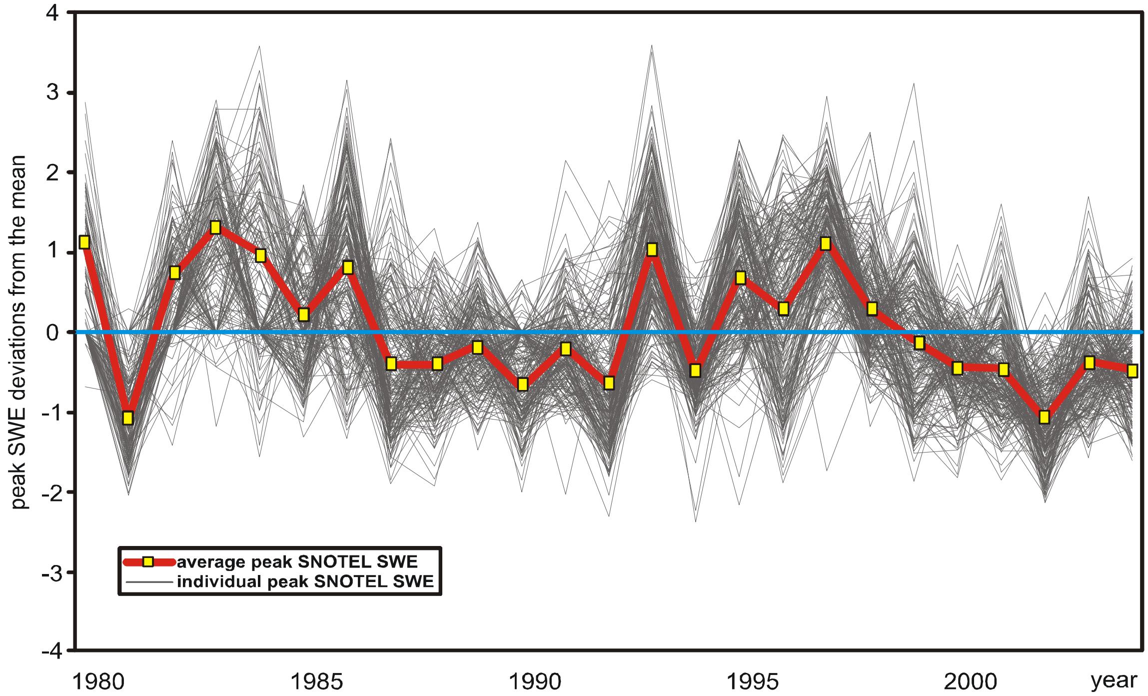

Figure 6.3. Individual SNOTEL stations standardized with a mean zero and standard deviation of one from 1991 – 2005

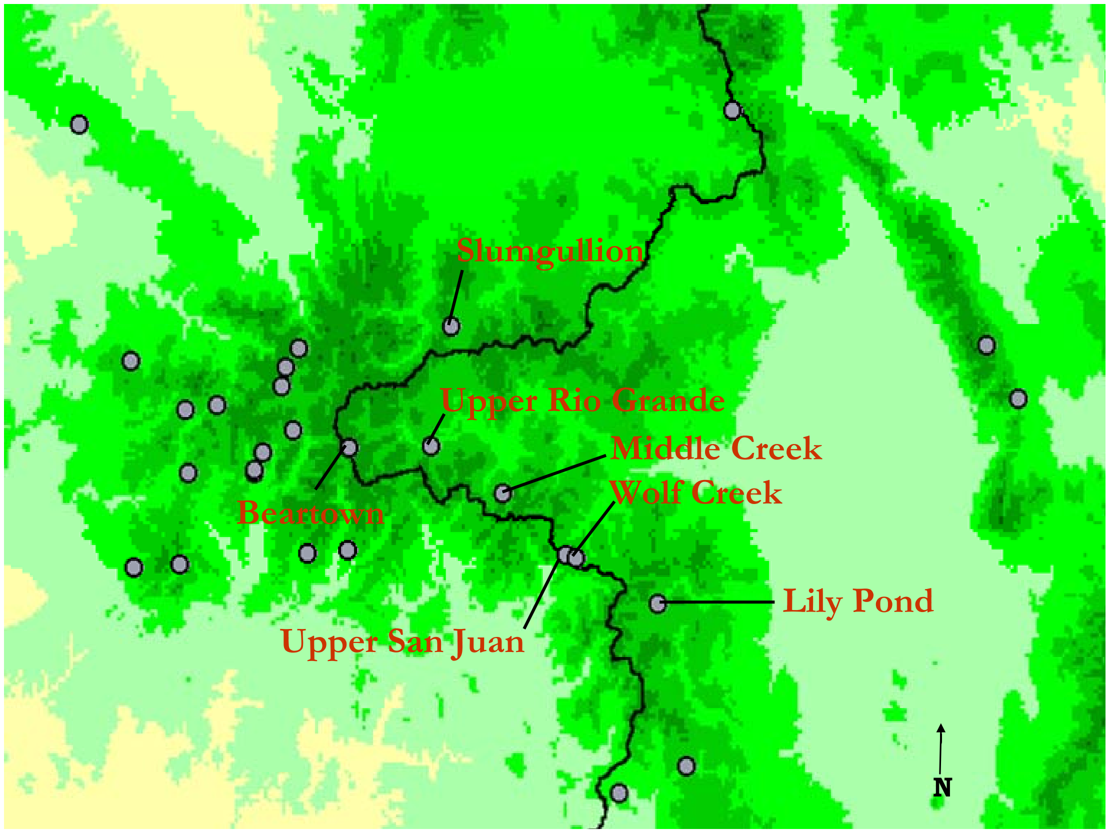

Figure 6.4. Location of SNOTEL stations located in the San Juan Mountains, CO.

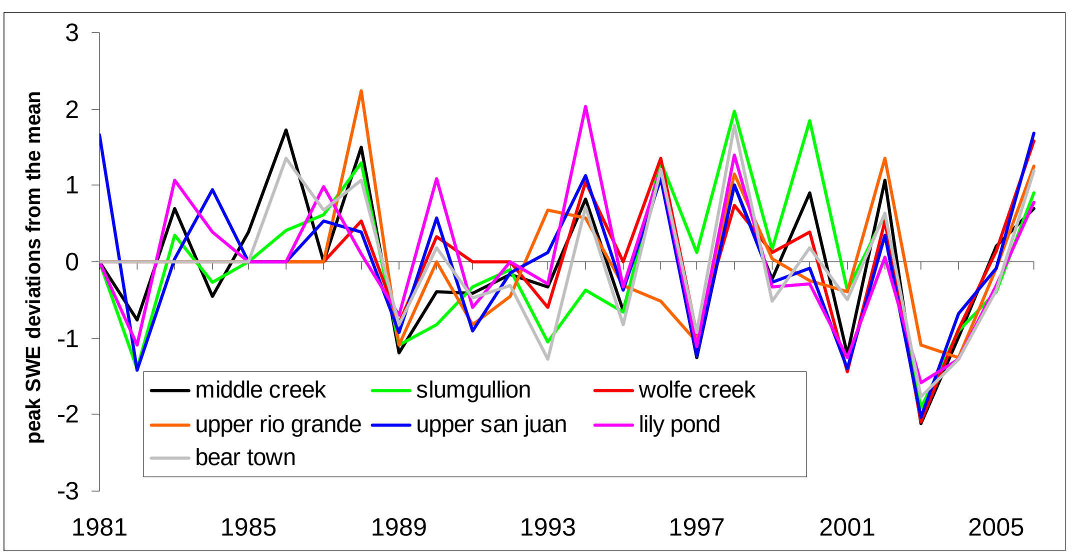

Figure 6.5. Peak SWE of SNOTEL stations in the San Juan Mountains standardized with a mean zero and standard deviation of one from 1981 – 2005.

CHAPTER 7: CONCLUSIONS

Defining regions of homogeneity is a crucial first step for many hydroclimatic investigations and spatial variability is a major component of hydroclimatic processes. Previous clustering of snow accumulation/ablation data has grouped measurement stations based on spatial proximity, or been temporally restricted due to the monthly data resolution of snow courses. This study investigated defining regions of homogeneity based on SNOTEL station daily SWE values with resultant groupings used in multivariate regressions to predict peak SWE with physiography as the independent variable.

The established snow climatologies showed general homogenous coarse-scale clusters along a north/south gradient. There are no definitive spatial patterns to the climatologies as a result of complex smaller-scale variables dominating daily SWE variability. Climatologies derived from niveograph descriptor variables showed improved spatial coherence as a result of larger scale influences. At the 6-cluster level of generalization, daily data clusters are best represented by the surrogate variable peak SWE (50% similarity), followed by April 1st (43% similarity) and physiography (41% similarity). Peak SWE clusters showed a higher spatially defined coherence than the daily data clusters, and the April 1st and physiography clusters showed even a greater coherence coupled with larger sized clusters at the expense of one smaller default cluster. The more traditional clustering procedure, Wards Minimum Variance, compared well on a quantitative basis (46%) with the SOM procedure, although qualitatively it failed to capture important subtle distinctions. Political groupings (state boundaries) compared poorly with the SOM and Ward's procedure.

Regressions showed improved predictive ability with an increase in cluster resolution (4, 6, and 16) and consistently improved results for clusters derived from daily data, peak SWE, and physiography (R^2 = 0.37, 0.45, 0.59 for the 4-cluster; 0.40, 0.53, 0.61 for the 6-cluster; and 0.70, 0.74, 0.76 for the 16-cluster, respectively). Elevation is a crucial parameter in the regression equations (included in 0.38 of the regressions). Other consistently vital variables were southwest barrier height (included in 0.34 of the regressions), regional northness (included in 0.34 of the regressions), and southwest shield height (included in 0.25 of the regressions).

Results showed that consideration of the data source (i.e. peak SWE vs. April 1st SWE) is crucial. Different data sources can lend way to different characterizations and delineations pertaining to a particular region and possibly lead to different conclusions in regards to interpretive work.

Daily data derived climatologies reflected local scale influences; whereas, climatologies based on the descriptor variables peak SWE, April 1st SWE, and physiography reflected larger scale influences. There were physical variables that were consistently used in determining SWE, although the degree of importance of these variables depends on resolution and general location of the climatology.

Further exploration is required using the SOM technique to improve understanding of interrelationships and processes. However, the responsibility is on the investigator when using automated, "black box" methods in ensuring the validity of the results.

Through utilization of the self-organizing map, regions of homogeneity were identified based on each station's SWE variability. Also, from these climatologies, the principle physical variables that best explain variability in peak SWE determined, supporting the hypothesis. The objective of this investigation to identify snow climatologies through the utilization of daily data and from these climatologies examine the relationship between peak SWE and physiography was attained.

CHAPTER 8: RECOMMENDATIONS

Further exploration is recommended using the SOM methodology, as follows:

- Incorporate high temporal SNOTEL values with synpotic-scale climate data, meteorological data, and atmospheric teleconnection indices.
- Explore whether additional stations should be used in the analysis, such as Medano Pass and Ute Creek in the Sangre de Cristos Mountain range.
- Cluster station data using alternative groupings of data, such as physiography and peak SWE.
- Cluster stations data using longer than daily time steps, such as weekly, monthly.
- Combine snow course and SNOTEL data to construct a longer and higher resolution dataset, especially for April 1st SWE. A monthly time step over the winter would provide an intra-annual time series with 4 to 6 time steps. Investigate how clustering using only snow course data combines to using only SNOTEL and the combination of the two datasets.
- Statistically test the heterogeneity of SOM groupings, such as undertaken by Lin and Chen (2006).
- Integrate satellite-based snow covered area datasets with SOM clustering to further explore the physiographic influence on SWE.

- Explore discretization of individual clusters by performing a SOM analysis on individual groupings. For example, from the SOM four cluster analysis using 15-years of daily data, further cluster number 4 since it spans Utah, Wyoming, and Colorado.

CHAPTER 9: LITERATURE CITED

Bales R., N. Molotch, T. Painter, M. Dettinger, R. Rice, and J Dozier, Mountain hydrology of the Western United States, *Water Resources Research*, 42: 1208 (doi:10.1029/2005WR004387), 2006.

Blöschl, G., Scaling Issues in Snow Hydrology, *Hydrological Processes*, 13, 2149-2175, 1999.

Cavazos T., Large-Scale circulation anomalies conductive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas, *Journal of Climate*, 12, 1506-1523, 1999.

Cavazos T., Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in the Balkans, *Journal of Climate*, 13, 1718-1732, 2000.

Cayan, D.R., Interannual climate variability and snowpack in the Western United States, *Journal of Climate*, 9, 928-947, 1996.

Chang, K., and Z. Li, Modeling snow accumulation with a geographic information system, *International Journal of Geographical Information Science*, 14 (7), 693-707, 2000.

Changnon, D., T.B. McKee, and N.J. Doesken, Hydroclimate variability in the Rocky Mountains, *Water Resources Bulletin*, 27, 733-743, 1991.

Changnon, D., T.B. McKee, and N.J. Doesken, Annual snowpack patterns across the Rockies: Long-term trends and associated 500-mb synoptic patterns, *Monthly Weather Review*, 121, 633-647, 1993.

Church, J. E., Recent studies of snow in the United States, *Quarterly Journal of the Royal Meteorological Society,* 169, 1914.

CIS, February 2006, *Self-Organizing Map Research*, Laboratory of Computer and Information Science Adaptive Informatics Research Centre, Helsinki University of

Technology. URL: < http://www.cis.hut.fi/research/som-research/> last accessed: 15 January 2008.

Clark, M.P., M.C. Serreze, and G.J. McCabe, Historical effects of El Nino and La Nina events on the seasonal evolution of the montane snowpack in the Columbia and Colorado River Basins, *Water Resources Research,* 37, 741-757, 2001.

Crane, R.G., and B.C. Hewitson, Clustering and upscaling of station precipitation records to regional patterns using self-organizing maps (SOMs), *Climate Research*, 25, 95-107, 2003.

Daly, C., R. P. Neilson, and D. L. Phillips, A statistical topographic model for mapping climatological precipitation over mountainous terrain, *Journal of Applied Meteorology*, 33, 140-158, 1994.

Daly, C., G. Taylor, and W. Gibson, The PRISM approach to mapping precipitation and temperature, Proceedings of the 10th Conference on Applied Climatology, Reno, Nevada, *American Meteorology Society*, 10-12, 1997.

Derksen, C., and LeDrew, Variability and change in terrestrial snow cover: Data acquisition and links to the atmosphere, *Progress in Physical Geography* 24, 469-498, 2000.

Doesken, N.J., and A. Judson, *The Snow Booklet: A Guide to the Science, Climatology, and Measurement of Snow in the United States*, Department of Atmospheric Sciences, Colorado State University, Fort Collins, 1996**,** 86 pp.

Doesken, N.J., and G.L. Schaefer, The contribution of SNOTEL precipitation measurements to climate analysis, monitoring and research in Colorado, in *Proceedings of the Western Snow Conference*, 55, 20-30, 1987.

Dressler, K.A., S.R. Fassnacht, and R.C. Bales, A comparison of snow telemetry and snow course measurements in the Colorado River Basin, *Journal of Hydrometeorology*, 7, 705-712, 2006.

Elder, K., W. Rosenthal, and R.E. Davis, Estimating the spatial distribution of snow water equivalence in a montane watershed, *Hydrological Processes*, 12, 1793-1808, 1998.

Fassnacht, S.R., K.A. Dressler, and R.C.Bales, Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data, *Water Resources Research*, 39(8): 1208, 2003, (doi:10.1029/2002WR001512).

Hann, C.T., *Statistical Methods in Hydrology*, Iowa State Press, 2002, 378 pp.

Hewitson, B.C. and R.G. Crane, Self-organizing maps: applications to synoptic climatology, *Climate Research*, 22, 13-26, 2002.

Hewitson, B.C. and R.G. Crane, (eds), *Neural Nets: Applications in Geography*, Kluwer Academic, Boston, 1994, 212 pp.

Johnson, J.B., and G.L. Schaefer, The influence of thermal, hydrologic, and snow deformation mechanisms on snow water equivalent pressure sensor accuracy, *Hydrological Processes*, 16, 3529-3542, 2002.

Kalkstein, L., G. Tan, J. Skindlov, An evaluation of three clustering procedures for use in Synoptic Climatological Classification, *Journal of Climate and Applied Meteorology,* 26, 717 – 730, 1987.

Kohonen, T., *Self-organizing Maps*, Springer-Verlag, Heidelberg, 1995, pp. 534.

Lin, G., and L. Chen, Identification of homogeneous regions for regional frequency analysis using the self-organizing map, *Journal of Hydrology,* 2006.

Mangiameli, P., S.K. Chen, and D. West, A comparison of SOM neural network and hierarchical clustering methods, *European Journal of Operational Research* 93, 402 – 417, 1996.

McGinnis, D.L., Estimating climate-change impacts on the Colorado Plateau snowpack using downscaling methods, *Professional Geographer*, 49(1), 117-125, 1997.

Moore, R.D. and I.G. McKendry, Spring snowpack anomaly patterns and winter climatic variability, British Columbia, Canada, *Water Resources Research*, 32(3), 623-632, 1996.

Molotch, N.P., and R.C. Bales, SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence, *Hydrological Processes*, 20, 723 – 739, 2006.

Molotch, N.P., S.R. Fassnacht, M.T. Colee, T. Bardsley, and R.C. Bales, A comparison of spatial statistical techniques for the development of a validation dataset for mesoscale modeling of snow water equivalent, *Eos Trans. AGU*, 82(47), Fall Meet. Supplemental, F553, 2001.

NRCS, *Snow surveys and water supply forecasting*, Agricultural Information Bulletin No. 536, National Resource Conservation Service, US Dept. of Agriculture, Portland, Oregon, undated.

NRCS, *National Water and Climate Center- SNOTEL data network*, National Resource Conservation Service, US Dep. Of Agric. URL: <http://www.wcc.nrcs.usda.gov/snow/> last accessed: 25 October 2007, undated.

Palmer, P.L., Estimating snow course water equivalent from SNOTEL pillow telemetry: An analysis of accuracy, *Proceedings of the Western Snow Conference*, 54, 81-86, 1986.

Pitlick, J., Relation between peak flows, precipitation, and physiography for five mountainaous regions in the western USA, *Journal of Hydrology*, 158, 219-240, 1994.

Rhodes, S.L., D. Ely, and J.A. Dracup, Climate and the Colorado River: The limits of Management, *Bulletin American Meteorological Society*, 65 (7), 682 – 691, 1984.

Serreze, M.C., M.P. Clark, R.L. Armstrong, D.A. McGinnis, and R.S. Pulwarty, Charateristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resources Research*, 35(7), 2145-2160, 1999.

Serreze, M.C., M.P. Clark, and A. Frei, Characteristics of large snowfall events in the montane western United States as examined using snowpack telemetry (SNOTEL) data, *Water Resources Research*, 37(3), 675-688, 2001.

Solomon , S. I., J. P. Denouvilliez, E. J. Chart, J. A. Woolley, and C. Cadou, The use of a square grid system for computer estimation of precipitation, temperature, and runoff, *Water Resources Research*, 4, 919 – 929, 1968.

U.S. Forest Service, *Forest land distribution data for the United States*, Washington, D.C., 2001.

Viessman, W., and L. Lewis, *Introduction to Hydrology 5th ed.*, Pearson Education Inc., New Jersey, 2003, 624 pp.

Viviroli, D., R. Weingartner, and B. Messerli, Assessing the hydrological significance of the world's mountains, *Mountain Research and Development*, 23(1), 32-40, 2003.

APPENDIX A:

Station cluster designation for all cluster analyses conducted.

Serreze *et al* : 4	a
All data: 4	b
All data: 6	c
Original, unscaled data: 6	d
All data: 9	e
All data: 12	f
All data: 16	g
All data 25-yr 10% missing values: 6	h
Peak SWE: 4	i
Peak SWE: 6	j
Peak SWE: 9	k
Peak SWE: 16	l
Physiography: 4	m
Physiography: 6	n
Physiography: 9	o
Physiography: 16	p
April SWE: 6	q
Date peak SWE: 6	r
4-variables: 6	s
Peak and length of season: 6	t
Length of season: 6	u
Wards peak SWE: 6	v
Wards date of peak SWE: 6	w
Wards length of season: 6	x
Wards cumulative SWE: 6	y

Station	Longitude	Latitude	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y
05G04S	-105.860826	42.43583	5	2	3	5	4	5	8	6	3	4	5	6	1	4	7	1	4	1	4	6	5	4	3	5	6
05G05S	-105.578164	42.280198	5	2	3	6	4	5	8	6	4	4	5	5	1	4	7	1	5	5	4	3	6	4	4	5	6
05J04S	-105.847565	40.399368	7	2	3	5	4	5	8	6	1	4	6	14	1	4	7	1	4	2	4	4	5	4	2	5	6
05J06S	-105.769931	40.805712	7	4	4	3	9	11	15	4	4	4	4	2	1	4	7	2	4	6	4	4	2	4	4	6	6
05J08S	-105.576147	40.032785	7	4	4	3	9	11	15	4	4	4	4	1	1	4	7	5	4	1	4	4	3	4	4	6	6
05J10S	-105.819799	40.414324	7	4	4	3	9	11	15	4	4	4	4	1	1	4	7	5	4	3	4	5	3	4	4	6	6
05J12S	-105.91977	40.225432	7	2	6	6	3	3	3	6	1	4	3	15	1	4	7	5	4	6	4	4	3	4	2	5	6
05J18S	-105.566667	40.207777	7	2	6	6	3	3	3	6	4	4	5	6	1	4	7	1	4	5	4	5	3	4	4	5	6
05J37S	-105.887004	40.532148	7	4	4	3	9	11	15	4	4	4	4	1	1	4	7	5	4	5	4	5	3	4	4	6	6
05J39S	-105.644833	40.311177	7	4	4	3	9	11	15	4	4	4	4	1	1	4	7	5	4	4	4	5	4	4	4	6	6
05J40S	-105.733367	40.432542	7	4	4	3	6	8	10	4	4	4	4	1	1	4	7	5	4	1	4	2	2	4	4	6	6
05J41S	-105.589558	39.936784	7	2	3	5	4	5	8	6	3	4	9	4	1	4	7	5	4	5	4	5	3	4	5	3	6
05J42S	-105.544261	40.035232	7	4	4	5	6	8	10	4	4	4	7	5	1	4	7	5	4	1	6	5	3	4	4	1	6
05K06S	-105.760833	39.915498	7	4	3	4	6	8	10	4	4	4	4	1	1	4	7	5	4	6	4	2	2	4	4	6	6
05K09S	-105.869731	39.646311	7	4	4	3	6	8	10	4	4	4	4	1	2	5	1	13	4	4	4	4	2	4	4	5	6
05K14S	-105.777892	39.803918	7	4	4	3	9	11	15	4	4	4	4	1	1	5	1	13	4	1	4	5	3	4	4	6	6
06G01S	-106.318611	42.733889	5	4	4	4	6	8	10	4	3	4	9	4	1	4	7	1	4	3	4	5	4	4	5	5	6
06G02S	-106.088888	42.571111	5	4	3	4	6	8	10	3	3	4	9	8	1	4	7	1	4	5	4	6	5	4	5	5	6
06H09S	-106.927143	41.159463	7	4	4	3	6	8	10	4	3	6	9	7	1	4	7	2	4	2	4	6	5	6	5	6	3
06H10S	-106.969167	41.154722	7	4	4	1	9	11	15	4	4	6	4	2	1	3	8	2	6	6	6	4	2	6	4	6	3
06H13S	-106.232087	41.358848	7	4	4	3	9	11	15	4	4	4	4	1	1	4	7	2	4	5	4	6	5	4	4	6	6
06H19S	-106.502475	41.32943	7	2	3	5	4	5	8	6	2	4	9	8	1	4	7	2	5	5	4	6	5	4	6	6	6
06H20S	-106.375553	41.330803	7	4	4	1	9	11	15	4	4	4	8	2	1	4	7	2	4	5	4	6	5	4	4	6	6
06H22S	-106.908475	41.002887	7	4	4	1	9	11	15	4	3	3	9	8	1	4	7	5	3	5	4	6	5	4	5	6	6
06H23S	-106.281074	41.462574	7	4	4	1	9	11	15	4	4	3	8	2	1	4	7	2	3	5	4	6	5	4	4	6	6
06J01S	-106.781296	40.533972	7	2	3	3	4	5	8	3	2	4	1	12	1	4	7	2	4	5	4	4	4	4	6	6	6
06J03S	-106.604077	40.394798	7	2	3	3	4	5	8	3	2	4	1	8	1	4	7	2	4	1	4	1	1	4	6	5	6
06J05S	-106.094328	40.347033	7	4	4	4	6	8	10	4	4	4	4	1	1	4	7	5	4	6	4	5	3	4	3	5	6
06J06S	-106.670277	40.078055	7	2	3	5	4	5	8	6	1	4	3	14	1	4	7	2	4	5	4	4	3	4	2	5	6
06J09S	-106.740376	40.367826	7	4	4	1	9	11	15	4	3	4	9	8	1	4	7	2	4	2	4	6	5	4	5	5	6
06J12S	-106.046027	40.875021	7	4	4	3	6	8	10	4	4	4	4	1	1	4	7	2	4	5	4	4	4	4	4	6	6
06J15S	-106.968707	40.847812	7	2	3	4	4	5	8	3	2	6	1	12	1	4	7	2	6	1	4	1	1	6	6	6	3
06J29S	-106.676797	40.537427	7	4	4	1	9	11	15	4	4	4	4	2	1	4	7	6	4	5	4	5	6	4	4	6	6
06K01S	-106.059781	39.361268	7	4	4	3	9	8	10	4	4	4	7	5	2	5	1	13	4	6	4	5	3	3	4	5	4
06K04S	-106.611694	39.075387	7	4	3	3	6	8	10	5	4	4	4	1	2	5	1	13	4	6	4	4	3	4	4	5	6
06K06S	-106.606944	39.297222	7	2	6	6	3	3	3	6	2	4	1	12	1	4	7	5	4	2	4	4	3	4	6	5	6
06K08S	-106.19681	39.379914	7	4	4	3	9	11	15	4	4	4	4	1	2	5	1	13	4	5	4	5	3	4	4	6	6
06K14S	-106.158015	39.717958	7	2	3	5	4	5	8	3	3	4	9	8	1	4	7	5	4	5	4	5	5	4	5	6	6
06K24S	-106.170954	39.489542	7	4	4	3	6	8	10	4	3	4	9	4	1	4	7	5	4	5	4	5	3	4	5	5	6
06K30S	-106.614527	39.317236	7	2	3	5	4	5	8	6	2	4	1	12	1	4	7	5	4	5	4	4	3	4	6	5	6
06K39S	-106.380059	39.616764	7	4	4	3	9	8	15	4	4	4	4	1	1	4	7	5	4	6	4	4	2	4	4	6	6
06K40S	-106.541699	39.087661	7	4	3	5	6	8	10	5	4	4	9	7	2	5	4	9	4	1	4	1	1	4	5	5	6
06L02S	-106.589743	38.819955	7	2	3	5	4	6	8	6	4	2	7	5	1	4	7	5	6	2	6	2	2	6	4	3	3
06L03S	-106.339654	38.488842	7	4	3	4	6	8	10	5	4	2	7	5	2	2	1	13	4	3	6	5	3	6	4	3	3
06L11S	-106.953	38.894327	7	3	3	4	8	6	11	5	4	2	7	5	2	2	1	13	6	2	6	2	2	2	4	3	2
06M03S	-106.835348	37.48576	7	3	5	2	7	9	12	5	3	5	5	6	2	2	1	14	2	2	3	3	6	2	3	2	2
06M17S	-106.801702	37.479215	7	4	5	1	7	12	16	5	4	5	8	2	2	2	1	13	2	4	3	3	4	2	4	2	2
06M22S	-106.451792	37.018776	7	3	5	2	7	9	11	5	3	5	5	6	2	2	2	14	2	5	3	3	6	2	3	2	2
06M23S	-106.548351	37.379287	7	3	5	4	7	9	11	5	4	5	7	5	2	2	1	13	2	6	3	3	6	2	4	2	2
06N03S	-106.65682	36.956356	7	1	1	6	2	2	2	2	1	1	3	16	2	2	2	14	2	2	1	3	4	1	2	2	5
06N04S	-106.321286	36.512161	7	3	2	5	5	4	7	2	1	1	3	15	2	2	2	14	2	2	1	3	6	2	2	2	2
06N14S	-106.263479	36.716309	7	3	2	4	5	6	11	2	4	2	7	5	2	2	2	14	2	3	3	5	3	2	4	2	2
06P01S	-106.392032	35.92244	7	1	1	6	2	2	2	1	1	1	3	13	2	2	2	10	2	2	1	5	4	1	2	2	5
06P10S	-106.833481	36.001983	7	1	1	6	2	2	2	1	1	1	3	13	2	2	2	14	1	4	1	3	4	1	1	1	5

Station	Longitude	Latitude	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y
07F01S	-107.066667	43.883333	5	2	6	6	3	3	6	6	1	3	3	15	1	4	7	1	5	5	2	6	6	3	2	6	4
07F02S	-107.183333	43.633333	5	4	4	5	6	8	10	4	3	3	9	8	1	4	7	1	3	2	4	6	5	3	5	6	4
07F03S	-107.170602	41.11172	7	2	3	5	4	5	8	6	2	4	5	6	1	4	7	2	4	2	4	3	6	4	3	6	6
07F04S	-107.266089	41.054128	7	2	6	5	3	3	3	6	1	6	6	14	1	4	7	1	5	3	5	5	3	6	2	4	3
07H05S	-107.152406	41.303981	7	4	3	3	6	8	10	5	3	6	9	4	1	4	7	2	5	4	6	2	2	6	5	6	3
07J04S	-107.057499	40.167454	7	2	3	5	4	5	8	6	2	4	1	12	2	2	2	14	4	4	4	4	2	4	6	5	6
07J05S	-107.294113	40.108123	7	4	4	1	9	11	15	4	4	4	4	1	1	4	7	6	4	6	4	5	3	4	4	6	6
07K01S	-107.143887	39.078134	7	3	2	4	5	6	11	2	2	2	1	12	2	5	4	9	2	2	3	3	6	2	6	5	2
07K02S	-107.598533	39.87505	7	4	3	3	6	8	10	3	3	4	9	7	2	2	2	6	4	1	4	1	1	4	5	6	6
07K06S	-107.874135	39.046441	7	4	5	1	7	12	16	5	4	5	7	5	2	2	2	10	2	6	3	5	5	2	4	3	2
07K09S	-107.288062	39.128967	7	3	2	4	5	6	11	2	3	2	9	8	2	2	1	13	2	6	3	2	2	2	4	3	2
07K11S	-107.048766	39.01522	7	4	4	1	9	12	16	4	4	2	8	1	2	5	1	13	6	5	6	6	5	2	4	3	2
07K12S	-107.356812	39.764865	7	4	4	1	9	12	16	4	4	2	4	1	2	2	2	6	6	1	6	3	4	2	4	6	2
07K13S	-107.236198	39.998839	7	4	4	3	6	8	10	4	3	4	9	8	1	4	7	6	4	3	4	4	2	4	5	5	6
07K14S	-107.634722	39.090555	7	3	2	4	5	6	11	5	2	5	1	12	2	2	2	10	2	6	3	2	2	2	6	3	2
07M05S	-107.805063	37.650962	7	1	6	6	2	2	5	2	1	2	3	16	2	2	1	14	2	5	1	3	4	2	2	2	2
07M11S	-107.777146	37.698663	7	3	5	2	7	9	12	5	2	5	5	6	2	2	1	10	2	2	3	3	6	2	3	2	2
07M12S	-107.688655	37.749324	7	3	5	2	7	9	12	5	4	5	7	5	2	5	1	13	2	6	3	3	5	2	4	2	2
07M14S	-107.726571	37.847472	7	3	2	4	5	6	11	5	4	2	7	5	2	2	1	13	2	2	3	3	4	2	4	2	2
07M16S	-107.260145	37.721944	7	1	1	6	2	2	5	2	1	5	3	15	2	2	1	14	2	2	1	3	6	2	2	1	2
07M21S	-107.034818	37.619784	7	3	5	2	7	9	12	5	4	5	7	5	2	2	1	13	2	2	3	6	5	2	4	2	2
07M27S	-107.675523	37.933903	7	2	3	5	4	6	8	6	3	2	5	6	2	2	1	13	2	1	6	2	2	6	3	3	3
07M29S	-107.92426	37.799256	7	3	5	4	7	6	11	5	4	2	7	5	2	2	1	10	2	4	3	2	2	2	4	2	2
07M30S	-107.204141	37.991517	7	4	5	4	7	10	10	5	4	5	7	5	2	2	2	10	6	2	6	3	6	2	4	3	2
07M31S	-107.506801	37.485095	7	3	2	4	5	9	11	5	2	5	5	11	2	2	1	10	2	6	3	3	6	2	3	2	2
07M32S	-107.512123	37.714091	7	3	5	2	7	9	12	5	4	5	7	5	2	2	1	13	2	4	3	3	4	2	4	2	2
07M33S	-107.713417	37.891803	7	4	5	2	7	12	16	5	4	5	4	1	2	2	1	13	2	2	3	3	6	2	4	2	2
07M34S	-107.632951	37.476211	7	3	5	2	7	9	12	5	4	5	7	5	2	2	1	10	2	4	3	2	4	2	4	2	2
07M35S	-107.802681	37.658	7	1	6	6	2	2	3	6	1	2	3	8	2	2	1	14	2	2	5	3	5	2	2	1	2
07S04S	-107.83122	33.360265	8	1	1	6	1	1	1	1	1	1	3	16	4	1	3	16	5	2	1	3	5	1	2	1	5
08F01S	-109.016667	43.666667	5	2	3	6	4	5	4	5	3	3	8	3	1	4	8	1	5	3	6	6	5	3	5	6	4
08G03S	-108.833333	42.566667	5	4	4	3	6	10	13	4	4	6	8	3	1	3	8	3	3	2	6	4	4	6	5	6	3
08G07S	-108.9	42.7	5	4	3	5	6	8	4	5	4	3	7	5	1	3	8	2	5	5	6	3	6	3	4	6	4
08K04S	-108.058349	39.058307	7	3	5	2	7	9	12	5	4	5	4	1	2	2	2	10	2	4	3	2	2	2	4	3	2
08L02S	-108.382489	38.417945	7	3	2	4	5	4	7	2	1	2	3	16	2	2	2	10	2	4	1	3	6	5	6	2	1
08M06S	-108.021549	37.786167	7	3	2	4	5	6	11	2	2	2	1	12	2	2	1	10	2	2	3	3	6	2	6	2	2
08M07S	-108.195442	37.891825	7	2	2	4	5	6	7	2	2	2	1	12	2	2	2	10	2	2	3	3	6	2	6	2	2
08M08S	-108.007859	37.645555	7	2	6	5	2	2	5	2	1	2	3	16	2	2	1	14	2	1	1	5	3	2	2	2	2
08S01S	-108.945023	33.736461	8	1	1	6	1	1	1	1	4	1	7	13	4	1	3	16	1	2	1	2	2	1	1	1	5
08S08S	-108.706178	33.371058	8	1	1	6	2	2	2	1	1	1	3	13	4	1	3	14	1	2	1	2	4	1	1	1	5
08T01S	-108.145381	32.924009	8	1	1	6	1	1	1	1	1	1	3	16	4	1	3	16	1	4	1	3	6	1	2	1	5
09F04S	-109.666667	43.7	5	4	4	3	6	8	13	3	3	3	8	3	1	3	8	3	3	3	2	6	5	3	5	6	4
09F08S	-109.75	43.5	5	4	3	5	6	7	13	3	3	3	9	4	1	3	8	6	3	4	2	3	6	3	5	6	4
09F21S	-109.949167	43.113056	5	2	3	5	4	7	9	3	3	3	6	10	3	3	8	3	3	4	2	3	6	3	3	6	4
09F23S	-109.758889	43.006667	5	4	4	5	6	8	13	3	3	3	8	3	1	3	8	6	3	5	2	6	5	3	5	6	4
09F24S	-109.316667	43.866667	5	4	4	5	6	11	13	4	4	3	8	3	1	5	4	9	3	5	2	6	5	3	4	6	4
09F25S	-109.45	43.266667	5	2	3	6	4	5	4	3	3	3	9	7	1	3	8	3	5	3	6	6	5	3	5	6	4
09F27S	-109.166667	43.033333	5	2	3	6	4	5	4	3	3	3	9	4	1	3	8	3	5	2	6	6	6	3	5	6	4
09G03S	-109.1	42.866667	5	4	4	3	9	11	15	4	4	3	8	3	1	4	8	6	3	2	6	6	5	3	5	6	4
09G09S	-109.259444	42.645833	5	4	3	4	6	10	13	3	3	6	8	3	1	3	8	6	3	6	6	4	3	6	5	6	3

Station	Longitude	Latitude	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y
09J01S	-109.543833	40.7165	6	3	2	5	5	4	7	2	1	5	6	10	3	3	8	7	2	2	3	5	4	5	3	2	1
09J05S	-109.883333	40.616667	6	3	5	4	7	9	12	5	3	5	9	7	3	3	5	7	2	2	3	3	6	5	5	2	1
09J08S	-109.9625	40.906833	6	2	2	6	5	4	5	2	1	1	3	14	3	3	5	7	5	1	1	2	1	2	2	2	2
09J16S	-109.673167	40.738333	6	3	2	5	5	4	7	2	3	5	6	14	3	3	8	7	2	1	3	2	2	5	3	2	1
09K01S	-109.531667	39.312167	6	1	2	6	2	2	5	2	1	5	3	16	2	2	1	14	2	2	1	2	2	5	2	1	1
09L03S	-109.271333	38.482	6	3	2	5	5	4	7	2	2	2	1	12	2	2	2	10	2	1	1	1	2	5	2	2	1
09M02S	-109.487333	37.8135	6	3	2	4	5	4	7	2	1	5	3	16	2	6	6	15	2	1	1	3	4	5	2	2	1
09S01S	-109.503437	33.978825	8	1	1	6	2	2	2	1	1	1	3	16	4	1	3	16	1	1	1	1	1	1	2	1	5
09S02S	-110.754312	34.31202	8	1	1	6	2	2	2	1	1	1	3	16	4	1	3	16	2	1	1	1	1	1	2	1	5
09S07S	-109.847619	34.114116	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
09S11S	-111.84374	35.072973	8	1	1	6	2	2	2	1	1	1	3	16	4	1	3	16	1	1	1	1	2	1	2	2	5
09S18S	-109.479727	33.758347	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
10F02S	-110.443611	43.944167	5	2	3	3	4	7	9	3	3	3	8	9	1	3	8	3	3	6	2	5	3	3	3	4	4
10F09S	-110.057778	43.748611	5	4	4	1	9	11	15	4	4	3	8	2	1	3	8	6	3	5	2	6	6	3	4	6	4
10F15S	-110.016389	43.249444	5	2	3	5	4	7	9	3	3	3	6	9	3	3	8	3	3	5	2	6	4	3	3	4	4
10F16S	-110.139444	43.173611	5	2	3	3	4	7	9	3	3	3	6	10	3	3	8	3	3	3	2	5	3	3	3	6	4
10F17S	-110.195833	43.130833	5	4	3	5	6	7	9	3	3	3	9	4	3	3	8	3	3	4	2	3	6	3	5	6	4
10F19S	-110.130278	43.390278	5	4	3	5	6	7	13	3	3	3	9	4	1	3	8	3	3	5	2	6	6	3	5	6	4
10F23S	-110.912778	43.516944	5	4	4	1	9	11	13	4	3	3	9	4	3	3	8	3	3	4	2	3	6	3	5	6	4
10F29S	-110.434722	43.343056	5	2	3	3	4	7	9	3	3	3	6	9	3	3	9	4	3	3	2	6	5	3	3	4	4
10G02S	-110.608889	42.964444	5	4	4	1	9	11	13	4	3	3	8	3	3	3	8	3	3	5	2	6	5	3	5	6	4
10G08S	-110.91	42.506389	5	2	3	5	4	7	9	3	3	3	6	10	3	3	9	4	3	5	2	6	6	3	3	4	4
10G12S	-110.801667	42.265278	5	2	3	4	8	7	9	3	1	6	6	14	3	3	8	7	3	3	2	2	2	3	3	6	4
10G13S	-110.531389	42.494722	5	2	3	5	4	7	9	3	3	3	6	14	3	3	8	3	3	4	2	6	4	3	3	4	4
10G15S	-110.591389	42.763889	5	4	3	3	6	7	13	3	3	3	9	4	3	3	8	3	3	4	2	6	6	3	5	6	4
10G20S	-110.660833	42.525	5	4	4	1	9	11	13	4	3	3	8	3	3	3	8	3	3	5	2	6	5	3	5	6	4
10G22S	-110.6775	42.301389	5	4	4	1	9	11	13	4	4	3	8	3	3	6	5	7	3	2	2	6	6	3	5	6	4
10G23S	-110.835	42.815	5	4	4	1	6	8	13	4	3	3	9	4	3	5	4	9	3	4	2	3	6	3	5	6	4
10G24S	-110.677778	42.145278	5	2	6	5	3	3	6	6	1	3	2	10	3	3	8	3	3	3	2	6	5	3	3	4	4
10G25S	-110.814167	42.645556	5	4	3	3	6	7	13	3	3	3	6	14	3	3	9	4	3	5	2	6	6	3	3	6	4
10J01S	-110.185333	40.921833	6	2	3	6	4	5	4	6	3	2	9	7	3	3	5	7	5	1	6	2	2	2	5	3	2
10J04S	-110.483667	40.950333	6	2	3	5	4	5	4	6	3	3	9	7	3	6	5	7	4	5	2	6	5	3	5	3	4
10J10S	-110.433167	40.597333	6	3	5	4	7	9	12	5	4	5	7	5	3	6	5	11	6	3	3	3	4	5	4	3	1
10J18S	-110.692833	40.548833	6	2	6	6	3	3	3	6	3	6	2	10	3	3	9	4	6	2	5	4	4	6	2	4	3
10J20S	-110.504167	40.909333	6	4	4	3	9	11	15	4	4	6	4	2	3	6	5	7	6	1	6	1	1	6	4	6	3
10J25S	-110.620833	40.738333	6	4	4	2	9	12	16	4	4	6	8	2	3	6	5	11	6	4	6	5	3	6	4	3	3
10J26S	-110.467667	40.717333	6	4	5	2	7	12	16	4	4	5	4	1	3	6	5	11	6	6	3	3	4	5	4	3	1
10J30S	-110.584833	40.580667	6	4	5	2	7	12	16	4	4	5	4	2	3	6	5	11	6	6	3	2	2	5	4	3	1
10J35S	-110.798167	40.864667	6	2	3	5	4	5	8	3	3	6	9	4	3	6	5	7	6	3	6	5	3	6	5	3	3
10J43S	-110.009833	40.774833	6	3	5	4	7	9	12	5	3	5	9	7	3	3	5	7	2	2	3	2	4	5	5	3	1
10J44S	-110.884833	40.796667	6	2	3	5	4	5	8	3	2	6	5	11	3	6	5	7	6	3	6	5	5	6	3	6	3
10J52S	-110.948333	40.679333	6	4	4	2	9	10	16	4	3	6	9	4	3	6	5	11	6	1	6	1	2	6	5	6	3
10K01S	-110.746167	39.891667	6	3	2	5	5	4	7	2	2	5	1	12	3	6	5	11	2	6	3	2	4	5	2	4	1
10K02S	-110.9875	39.964333	6	2	6	5	3	6	6	6	2	6	1	12	3	6	5	8	6	4	5	2	4	6	6	4	3
10R04S	-109.309524	33.65387	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
10S01S	-110.917734	33.812421	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
11F02S	-111.212222	43.570833	5	2	3	5	4	7	6	6	3	3	6	9	3	3	9	3	3	5	2	3	6	6	3	4	3
11F11S	-111.687778	43.209444	5	2	6	5	3	3	6	6	2	3	2	8	3	3	9	4	3	5	5	6	5	6	2	4	3
11G01S	-111.358889	42.9525	5	2	6	5	3	3	6	6	3	3	8	3	3	3	9	4	3	3	2	5	3	3	2	4	4
11G05S	-111.298333	42.562222	5	2	3	5	4	7	6	3	3	3	6	9	3	3	9	4	3	5	2	6	5	3	3	4	4

Station	Longitude	Latitude	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y
11G06S	-111.560833	42.360278	5	2	3	3	8	7	9	3	3	6	6	14	3	3	9	4	3	1	5	5	3	6	3	4	3
11G30S	-111.956111	42.524722	5	2	3	3	8	7	9	3	3	6	6	14	3	6	9	4	3	6	2	5	3	6	3	6	3
11G32S	-111.603333	42.051111	5	4	4	1	6	10	13	3	3	3	9	4	3	6	5	7	3	4	2	6	5	6	5	6	3
11G33S	-111.165833	42.412778	5	2	6	6	3	3	3	6	1	3	3	15	3	3	9	4	3	3	2	6	5	3	2	4	4
11H08S	-111.944	41.376	6	4	5	1	8	10	14	3	3	6	9	7	3	6	6	8	6	1	5	5	3	6	5	4	3
11H21S	-111.445833	41.313667	6	2	3	3	4	7	9	3	3	6	6	9	3	6	5	11	3	4	2	4	4	6	3	4	3
11H25S	-111.826333	41.405833	6	2	6	6	3	3	3	6	1	6	3	14	3	6	9	8	6	1	5	5	3	6	2	1	3
11H30S	-111.921	41.383333	6	2	6	4	3	3	3	6	1	6	3	15	3	3	9	4	6	6	5	4	4	6	2	1	3
11H36S	-111.629333	41.898333	6	4	4	1	8	10	13	3	3	6	9	4	3	6	5	8	3	4	2	6	4	3	5	6	4
11H37S	-111.419167	41.684667	6	2	3	3	8	7	9	3	3	6	5	6	3	6	5	7	3	5	2	4	3	3	3	6	4
11H55S	-111.537667	41.413	6	2	3	4	8	7	9	3	3	6	6	14	3	6	5	8	6	6	5	4	2	6	3	4	3
11H57S	-111.496667	41.465833	6	4	4	1	8	10	13	4	3	6	9	4	3	6	5	11	6	3	6	5	3	6	5	6	3
11J01S	-111.047833	40.855	6	4	4	2	9	10	13	4	3	6	9	4	3	6	5	11	6	5	6	4	4	6	5	3	3
11J02S	-111.069167	40.885	6	2	3	4	4	5	8	3	3	6	6	10	3	6	5	8	3	2	5	4	5	6	3	3	3
11J08S	-111.216667	40.183333	6	2	6	4	3	3	6	6	1	6	2	10	3	6	6	12	6	4	5	5	5	6	3	4	3
11J11S	-111.809333	40.974833	6	4	5	1	8	10	14	5	3	6	9	4	3	6	6	15	6	5	5	5	5	6	5	4	3
11J21S	-111.616333	40.428167	6	3	2	2	5	6	11	3	1	2	2	10	3	6	5	8	6	1	5	1	1	6	2	4	3
11J23S	-111.256667	40.295333	6	2	6	5	3	3	6	6	2	6	2	15	3	6	6	8	6	1	5	2	2	6	2	4	3
11J32S	-111.089833	40.357333	6	2	6	6	3	3	3	6	1	2	3	15	3	6	6	8	6	3	5	5	4	6	2	4	3
11J42S	-111.092167	40.789167	6	4	4	1	9	12	16	4	4	6	8	2	3	6	5	15	6	1	6	1	1	6	4	3	3
11J46S	-111.097667	40.6125	6	2	6	6	3	3	3	6	1	6	2	15	3	6	6	11	6	4	5	5	3	6	2	4	3
11J52S	-111.628667	40.762333	6	2	6	5	3	3	6	6	2	6	2	15	3	6	9	8	6	1	5	1	1	6	2	4	3
11J53S	-111.117667	40.79	6	2	6	5	3	3	6	6	1	6	2	10	3	6	9	8	6	4	5	4	4	6	2	4	3
11J56S	-111.532833	40.620167	6	4	5	2	7	10	14	5	3	5	9	4	3	6	5	11	6	2	6	5	4	5	5	3	1
11J57S	-111.582667	40.599167	6	4	5	2	8	10	14	5	3	5	9	4	3	6	5	11	6	2	6	2	2	6	5	3	3
11J64S	-111.709167	40.837333	6	4	3	2	8	10	14	3	3	6	9	4	3	6	5	8	6	6	5	5	3	6	5	6	3
11J65S	-111.636667	40.658167	6	4	3	2	8	10	14	3	3	6	9	4	3	6	4	8	6	2	5	2	2	6	5	3	3
11K03S	-111.318833	39.683	6	2	3	4	4	6	8	6	3	6	9	7	3	6	6	8	6	6	5	4	3	6	5	4	3
11K09S	-111.433167	39.310167	6	3	5	4	7	10	14	5	4	2	4	1	3	6	5	11	2	1	3	1	2	2	4	4	2
11K13S	-111.558167	39.137167	6	2	6	6	3	3	3	6	1	2	3	16	3	6	6	8	5	3	5	5	3	6	2	4	3
11K15S	-111.468333	39.045833	6	2	2	5	5	6	7	6	2	2	2	11	3	6	6	12	2	5	5	5	3	6	3	4	3
11K21S	-111.283	39.866333	6	2	3	4	8	6	8	3	3	6	5	6	3	6	5	11	6	5	5	5	4	6	3	4	3
11K22S	-111.250667	39.892833	6	2	6	5	3	3	6	6	1	6	2	10	3	6	5	8	6	4	5	5	3	6	3	4	3
11K28S	-111.272333	39.451833	6	2	3	4	4	6	8	6	3	6	9	7	3	6	5	11	6	5	5	4	3	6	5	4	3
11K31S	-111.4365	39.1335	6	2	3	4	8	6	8	3	3	6	9	7	3	6	6	11	6	6	5	4	2	6	5	3	3
11K39S	-111.5825	39.012833	6	2	3	5	4	5	8	6	3	2	9	7	3	6	6	8	5	2	5	2	2	6	5	4	3
11K52S	-111.630833	39.929667	6	2	6	4	5	6	6	6	2	2	2	12	3	6	6	8	6	2	5	1	1	6	3	4	3
11L01S	-111.6765	38.772667	6	4	5	3	9	12	16	4	4	5	7	1	3	6	6	11	2	2	6	2	2	2	4	3	2
11L04S	-111.597167	38.679667	6	2	2	6	5	6	7	6	2	2	1	12	4	6	6	12	2	6	5	5	3	6	6	4	3
11L05S	-111.4745	38.208167	6	3	2	6	5	4	7	2	4	5	7	5	4	6	6	12	2	2	3	2	2	2	4	2	2
11L12S	-111.683333	38.800333	6	1	6	6	3	3	3	6	1	1	3	16	4	6	6	12	5	1	1	2	2	6	2	1	3
11M03S	-111.881667	37.8365	6	3	2	4	5	9	11	2	1	5	7	5	4	6	6	12	2	4	3	2	2	5	2	2	1
11P13S	-109.152819	33.803924	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	5	1	1	1
11R03S	-109.458805	33.921211	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	5	1	1	1

Station	Longitude	Latitude	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y
11R06S	-111.406433	34.456602	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
11R10S	-111.518485	34.941094	8	1	1	6	2	2	2	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
12J06S	-112.224	40.4425	6	3	5	2	7	10	14	5	3	5	9	7	3	6	6	15	2	2	3	3	6	5	5	4	1
12K01S	-112.414667	39.937	6	2	6	6	3	3	5	2	1	2	3	15	3	6	6	8	2	1	5	1	1	6	2	1	3
12L04S	-112.0185	38.508	6	3	2	4	5	6	7	2	3	5	9	7	4	6	6	12	2	2	3	2	2	5	5	4	1
12L06S	-112.392833	38.484	6	3	2	4	5	6	11	5	3	5	5	6	3	6	6	8	2	6	3	2	4	2	3	2	2
12L07S	-112.356667	38.302	6	4	5	2	7	12	16	4	2	5	2	2	4	6	6	15	2	2	3	3	6	5	4	3	1
12L12S	-112.4365	38.303167	6	3	2	4	5	6	7	2	3	5	5	6	4	1	6	12	2	4	3	3	6	5	3	2	1
12L15S	-112.249	38.965333	6	2	3	3	8	6	14	5	3	2	9	7	4	1	6	12	5	1	5	1	1	6	5	4	3
12M03S	-112.900333	37.574	6	3	2	4	5	4	7	2	2	5	1	16	4	1	6	12	2	2	1	2	4	5	6	2	1
12M05S	-112.575333	37.49	6	1	1	6	2	2	2	1	1	5	3	16	4	1	6	12	2	2	1	3	4	5	2	1	1
12M06S	-112.514167	37.4875	6	1	1	6	1	1	1	1	1	5	3	16	4	1	6	12	2	2	1	2	4	5	2	1	1
12M13S	-112.740667	37.660667	6	3	2	4	5	4	7	2	2	5	2	11	4	1	6	12	2	4	3	2	4	5	3	2	1
12M23S	-112.837833	37.569167	6	3	5	2	7	9	12	5	2	5	2	11	4	1	6	12	2	2	3	2	2	5	3	2	1
12P02S	-112.149147	35.142024	8	1	1	6	1	1	1	1	1	1	3	13	4	1	3	16	1	1	1	1	1	1	1	1	5
13M02S	-113.396833	37.513	6	1	1	6	2	2	2	1	1	1	2	13	4	1	6	12	2	1	1	2	2	5	1	1	1
13M04S	-113.846	37.486333	6	1	1	6	1	1	1	1	1	1	3	13	4	1	6	12	1	4	1	3	4	1	2	1	5
13M05S	-113.05	37.533333	6	3	2	2	5	9	11	2	2	5	2	11	4	1	6	12	2	2	3	4	4	5	3	2	1

APPENDIX B:

Regression R^2 results and physical parameters identified, for daily data, peak SWE, and physiography SOM analyses of size four, six, and sixteen.

Regression results for the SOM 4-cluster derived from daily data.

parameter	cluster				total
	1	2	3	4	
X	0	5	1	5	11
Y	4	5	3	5	17
Z	6	3	3	3	15
Slp	1	4	10	4	19
Local-E	0	0	4	0	4
Loc-N	3	0	3	0	6
Reg-Slp	4	1	8	1	14
Reg-N	5	6	2	6	19
W-Distance	2	3	0	3	8
W-Barrier	2	0	11	0	13
W-Shield	1	0	5	0	6
NW-Distance	1	4	1	4	10
NW-Barrier	2	1	1	1	5
NW-Shield	2	1	5	1	9
SW-Distance	0	0	5	0	5
SW-Barrier	1	6	12	6	25
SW-Shield	2	13	2	13	30
Forest	3	14	1	14	32
R-square	0.38	0.31	0.51	0.28	
N	28	72	40	76	

Regression results for the SOM 4-cluster derived from peak SWE.

parameter	cluster				total
	1	2	3	4	
Z	5	4	3	7	19
Y	7	0	2	3	12
X	0	1	1	0	2
W-Shield	0	5	1	2	8
W-Distance	2	1	0	3	6
W-Barrier	7	10	6	0	23
SW-Shield	2	5	14	0	21
SW-Distance	2	2	1	10	15
SW-Barrier	0	6	14	13	33
Slp	0	2	3	2	7
Reg-Slp	9	1	10	1	21
Reg-N	0	1	9	4	14
NW-Shield	4	8	1	0	13
NW-Distance	5	0	2	3	10
NW-Barrier	10	8	3	5	26
Loc-N	0	5	14	15	34
Local-E	0	0	0	5	5
Forest	1	2	4	1	8
R-square	0.41	0.46	0.42	0.49	
N	52	29	76	59	

Regression results for the SOM 4-cluster derived from physiography.

parameter	cluster				total
	1	2	3	4	
Z	15	13	13	15	56
Y	4	4	10	1	19
X	1	0	10	7	18
W-Shield	8	0	0	1	9
W-Distance	0	2	2	4	8
W-Barrier	0	8	3	0	11
SW-Shield	0	3	5	0	8
SW-Distance	0	1	0	1	2
SW-Barrier	14	15	11	2	42
Slp	0	2	1	1	4
Reg-Slp	14	2	15	3	34
Reg-N	11	13	12	14	50
NW-Shield	0	0	1	7	8
NW-Distance	0	0	1	1	2
NW-Barrier	3	3	7	2	15
Loc-N	0	8	7	7	22
Local-E	0	1	5	3	9
Forest	0	0	0	14	14
R-square	0.54	0.56	0.57	0.69	
N	59	46	79	32	

Regression results for the SOM 6-cluster derived from daily data.

parameter	cluster						total
	1	2	3	4	5	6	
X	1	0	6	0	0	4	11
Y	4	3	2	7	2	0	18
Z	8	0	3	1	11	4	27
Slp	1	1	0	0	7	1	10
Local-E	3	10	2	4	4	0	23
Loc-N	3	2	7	2	1	8	23
Reg-Slp	1	6	4	4	0	0	15
Reg-N	2	4	2	0	3	3	14
W-Distance	0	1	1	0	2	0	4
W-Barrier	1	3	1	2	2	0	9
W-Shield	1	0	1	9	0	0	11
NW-Distance	0	0	0	1	1	4	6
NW-Barrier	2	5	5	0	1	1	14
NW-Shield	2	5	3	7	0	4	21
SW-Distance	0	2	2	4	0	1	9
SW-Barrier	0	5	1	11	5	0	22
SW-Shield	2	6	13	2	9	2	34
Forest	2	0	0	0	2	1	5
R-square	0.32	0.45	0.45	0.48	0.32	0.36	
N	24	28	61	48	28	27	

Regression results for the SOM 6-cluster derived from peak data.

parameter	cluster						total
	1	2	3	4	5	6	
X	0	10	0	1	0	0	11
Y	9	0	10	2	4	2	27
Z	1	14	15	7	1	9	47
Slp	0	2	0	1	1	1	5
Local-E	9	2	1	2	0	4	18
Loc-N	2	4	0	15	1	5	27
Reg-Slp	12	2	7	5	0	0	26
Reg-N	3	6	14	15	3	1	42
W-Distance	1	4	0	0	1	3	9
W-Barrier	6	11	0	8	15	3	43
W-Shield	0	2	0	0	1	0	3
NW-Distance	1	0	0	1	0	1	3
NW-Barrier	4	2	0	0	8	1	15
NW-Shield	1	0	0	0	3	3	7
SW-Distance	0	5	2	0	0	0	7
SW-Barrier	14	3	0	12	13	0	42
SW-Shield	6	5	14	5	6	2	38
Forest	4	11	2	0	0	4	21
R-Square	0.49	0.65	0.39	0.69	0.56	0.38	
N	24	29	38	42	42	41	

Regression results for the SOM 6-cluster derived from physiography.

parameter	cluster						total
	1	2	3	4	5	6	
Z	14	15	0	6	9	14	58
Y	1	4	0	12	0	9	26
X	0	1	0	0	1	5	7
W-Shield	2	15	5	0	6	0	28
W-Distance	0	0	0	1	0	0	1
W-Barrier	3	2	1	1	4	2	13
SW-Shield	0	15	0	7	1	8	31
SW-Distance	0	0	0	3	2	1	6
SW-Barrier	15	0	0	12	14	3	44
Slp	0	15	0	1	3	1	20
Reg-Slp	7	0	1	14	2	2	26
Reg-N	8	15	10	15	8	5	61
NW-Shield	1	3	3	0	2	0	9
NW-Distance	0	0	0	1	0	1	2
NW-Barrier	9	3	12	13	2	4	43
Loc-N	0	2	0	3	0	3	8
Local-E	0	0	0	4	2	2	8
Forest	0	2	0	1	0	6	9
R-square	0.6	0.62	0.5	0.61	0.53	0.79	
N	26	37	40	45	11	57	

Regression results for the SOM 16-cluster derived from daily-data.

parameter	cluster																total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Z	2	5	8	0	3	2	4	0	4	3	8	2	0	9	0	2	52
Y	3	2	0	2	0	1	3	0	1	1	6	7	2	6	2	0	36
X	2	0	1	1	2	5	2	2	0	1	0	0	12	1	0	2	31
W-Shield	1	1	2	1	4	1	4	1	1	3	1	0	1	3	1	1	26
W-Distance	1	0	0	0	0	6	0	0	14	1	7	0	0	1	0	0	30
W-Barrier	1	3	6	0	2	4	0	0	1	3	0	0	3	3	0	2	28
SW-Shield	1	3	4	5	1	1	5	1	0	2	1	2	0	2	0	5	33
SW-Distance	6	0	0	1	1	3	0	0	1	0	0	3	2	0	3	3	23
SW-Barrier	0	8	4	7	2	3	3	0	0	13	11	2	2	2	10	3	70
Slp	1	2	5	0	1	1	5	4	0	0	7	4	1	0	0	1	32
Reg-Slp	1	3	1	4	0	1	1	0	2	1	1	5	2	1	4	5	32
Reg-N	2	1	3	1	6	2	2	1	2	1	8	0	2	4	9	7	51
NW-Shield	2	3	0	1	0	3	2	0	0	0	0	1	1	1	1	0	15
NW-Distance	1	0	0	0	0	2	0	3	0	0	0	0	0	2	0	2	10
NW-Barrier	1	0	0	2	0	1	4	9	0	1	2	0	7	0	0	4	31
Loc-N	1	5	0	0	6	2	0	1	5	7	2	1	1	1	1	0	33
Local-E	1	0	0	0	3	2	4	0	0	6	13	6	11	2	0	4	52
Forest	0	7	0	4	3	4	2	1	0	1	12	1	0	0	3	1	39
R-square	0.68	0.79	0.60	0.94	0.89	0.83	0.61	0.30	0.67	0.57	0.71	0.75	0.81	0.91	0.39	0.69	
N	13	10	13	6	6	13	15	23	14	19	15	11	18	9	19	12	

Regression results for the SOM 16-cluster derived from peak SWE.

parameter	cluster																total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Z	0	3	5	7	6	7	13	7	3	2	0	4	3	1	7	14	82
Reg-N	11	5	7	9	2	2	7	4	13	1	4	1	1	0	2	2	71
SW-Barrier	8	5	5	6	8	9	6	2	0	2	6	0	0	1	2	1	61
Loc-N	14	0	1	0	2	2	3	4	0	1	2	6	6	2	8	3	54
W-Barrier	9	0	3	8	4	6	0	1	4	1	7	3	1	3	1	2	53
SW-Shield	4	3	2	4	2	0	9	0	0	4	0	8	5	4	0	0	45
NW-Shield	0	10	1	1	4	6	0	0	0	2	0	5	4	0	7	5	45
Local-E	2	10	4	1	2	5	0	1	0	2	2	2	7	0	1	2	41
Reg-Slp	0	1	5	11	1	0	1	3	4	5	1	2	1	2	1	1	39
W-Shield	1	3	2	1	0	0	3	1	0	3	1	3	2	3	7	7	37
NW-Barrier	3	0	1	2	3	1	1	5	1	6	0	2	5	1	3	3	37
Y	1	3	1	1	5	0	1	0	0	0	0	3	3	10	5	2	35
Forest	4	3	2	0	0	2	2	1	1	3	1	1	2	1	9	1	33
SW-Distance	4	0	7	1	3	2	0	2	0	0	4	4	0	0	1	2	30
Slp	4	2	1	1	1	6	0	1	3	1	0	0	4	2	1	2	29
W-Distance	0	0	0	0	0	0	1	0	2	2	0	0	2	0	0	4	11
X	2	1	0	1	2	0	0	0	1	0	0	0	1	2	0	0	10
NW-Distance	1	0	1	0	2	0	0	0	0	0	1	1	0	0	2	1	9
R-square	0.69	0.81	0.82	0.66	0.53	0.92	0.72	0.59	0.98	0.65	0.9	0.65	0.59	0.64	0.85	0.77	
N	22	12	11	22	20	11	16	11	6	11	6	14	15	12	11	16	

Regression results for the SOM 16-cluster derived from physiography.

parameter	cluster																total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Z	0	14	1	3	4	12	5	15	0	4	9	5	2	7	0	7	81
Reg-N	0	7	8	5	3	4	2	5	2	5	1	13	1	6	0	4	62
W-Barrier	1	1	11	2	2	1	1	11	0	2	8	10	8	2	1	1	61
NW-Distance	0	2	2	1	0	14	8	3	12	0	9	4	1	1	0	1	57
Forest	1	0	0	10	11	0	0	15	0	0	1	14	0	0	1	2	53
SW-Barrier	0	0	1	0	14	1	5	5	0	5	4	0	14	3	0	0	52
Y	3	2	0	7	1	0	4	13	0	3	4	1	0	0	11	1	49
SW-Shield	14	0	2	6	12	1	1	9	0	0	2	1	0	1	0	5	49
Local-E	0	7	5	1	1	0	0	5	0	0	2	3	5	1	11	1	41
W-Shield	1	10	1	1	3	7	0	6	0	0	0	0	3	2	0	5	34
Reg-Slp	2	2	7	0	0	0	7	3	0	0	1	1	9	0	1	3	33
NW-Shield	3	2	3	1	1	2	0	10	0	0	0	1	2	8	0	1	33
Loc-N	0	2	3	0	0	0	0	1	0	3	3	0	8	13	0	4	33
Slp	1	3	5	2	0	0	1	2	0	0	7	1	0	0	0	3	22
X	0	0	0	6	5	0	0	0	1	0	3	5	0	1	0	0	21
SW-Distance	0	0	0	0	2	1	5	2	0	1	4	2	3	1	0	3	21
W-Distance	0	6	2	1	3	0	0	0	0	3	2	0	1	0	1	0	19
NW-Barrier	1	0	1	1	1	0	3	4	0	0	0	0	3	5	0	0	19
R-square	0.65	0.67	0.66	0.68	0.83	0.94	0.75	0.89	0.93	0.46	0.74	0.82	0.73	0.87	0.91	0.62	
N	10	16	17	11	17	10	14	20	4	13	17	17	17	13	5	15	