DISSERTATION

METHODOLOGIES FOR TRANSFORMING DATA TO INFORMATION AND ADVANCING THE UNDERSTANDING OF WATER RESOURCES SYSTEMS TOWARDS INTEGRATED WATER RESOURCES MANAGEMENT

Submitted by

Panagiotis D. Oikonomou

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2017

Doctoral Committee:

Advisor: Darrell G. Fontane
Co-Advisor: Reagan M. Waskom

Neil S. Grigg
Christos A. Karavitis
Charles W. Anderson

ABSTRACT

METHODOLOGIES FOR TRANSFORMING DATA TO INFORMATION AND ADVANCING THE UNDERSTANDING OF WATER RESOURCES SYSTEMS TOWARDS INTEGRATED WATER RESOURCES MANAGEMENT

The majority of river basins in the world, have undergone a great deal of transformations in terms of infrastructure and water management practices in order to meet increasing water needs due to population growth and socio-economic development. Surface water and groundwater systems are interwoven with environmental and socio-economic ones. The systems' dynamic nature, their complex interlinkages and interdependencies are inducing challenges for integrated water resources management. Informed decision-making process in water resources is deriving from a systematic analysis of the available data with the utilization of tools and models, by examining viable alternatives and their associated tradeoffs under the prism of a set of prudent priorities and expert knowledge.

In an era of increasing volume and variety of data about natural and anthropogenic systems, opportunities arise for further enhancing data integration in problem-solving approaches and thus support decision-making for water resources planning and management. Although there is a plethora of variables monitored in various spatial and temporal scales, particularly in the United States, in real life, for water resources applications there are rarely, if ever, perfect data. Developing more systematic procedures to integrate the available data and harness their full

potential of generating information, will improve the understanding of water resources systems and assist at the same time integrated water resources management efforts.

The overarching objective of this study is to develop tools and approaches to overcome data obstacles in water resources management. This required the development of methodologies that utilize a wide range of water and environmental datasets in order to transform them into reliable and valuable information, which would address unanswered questions about water systems and water management practices, contributing to implementable efforts of integrated water resources management. More specifically, the objectives of this research are targeted in three complementary topics: drought, water demand, and groundwater supply. In this regard, their unified thread is the common quest for integrated river basin management (IRBM) under changing water resources conditions. All proposed methodologies have a common area of application namely the South Platte basin, located within Colorado. The area is characterized by limited water resources with frequent drought intervals. A system's vulnerability to drought due to the different manifestations of the phenomenon (meteorological, agricultural, hydrological, socio-economic and ecological) and the plethora of factors affecting it (precipitation patterns, the supply and demand trends, the socioeconomic background etc.) necessitates an integrated approach for delineating its magnitude and spatiotemporal extent and impacts. Thus, the first objective was to develop an implementable drought management policy tool based on the standardized drought vulnerability index framework and expanding it in order to capture more of drought's multifaceted effects. This study illustrated the advantages of a more transparent data rigorous methodology, which minimizes the need for qualitative information replacing it with a more quantitative one. It is believed that such approach may convey drought information to decision makers in a holistic manner and at the same time avoid the existing practices of broken linkages and fragmentation of reported drought impacts.

Secondly, a multi-scale (well, HUC-12, and county level) comparative analysis framework was developed to identify the characteristics of the emergent water demand for unconventional oil and gas development. This effort revealed the importance of local conditions in well development patterns that influence water demand, the magnitude of water consumption in local scales in comparison to other water uses, the strategies of handling flowback water, and the need for additional data, and improved data collection methods for a detailed water life-cycle analysis including the associated tradeoffs. Finally, a novel, easy to implement, and computationally low cost methodology was developed for filling gaps in groundwater level time series. The proposed framework consists of four main components, namely: groundwater level time series; data (groundwater level, recharge and pumping) from a regional physically-based groundwater flow model; autoregressive integrated moving average with external inputs modeling; and the Ensemble Smoother (ES) technique. The methodology's efficacy to predict accurately groundwater levels was tested by conducting three numerical experiments at eighteen alluvial wells. The results suggest that the framework could serve as a valuable tool in gaining further insight of alluvium aquifer dynamics by filling missing groundwater level data in an intermittent or continuous (with relative short span) fashion. Overall, it is believed that this research has important implications in water resources decision making by developing implementable frameworks which advance further the understanding of water systems and may aid in integrated river basin management efforts.

ACKNOWLEDGEMENTS

This Ph.D. process was a magnificent endeavor, which allowed me to expand my horizons. Kavafis writes in his poem *Ithaca,* “As you set out for Ithaca, hope the voyage is a long one, full of adventure, full of discovery”. My journey would not be possible if certain people had not believed in my capacity for success, and of course, it would not have been the same if I had not crossed paths with inspiring and supportive people that I met later on during the process. I would like to take the opportunity to express my sincere gratitude and appreciation to these wonderful people.

First and foremost, I am indebted to Professor Christos A. Karavitis, who was the first to believe in my capacity for success and who encouraged me to pursue doctoral studies at Colorado State University (CSU), his alma mater. Our frequent communication continued throughout these years, and his useful academic advices and research feedbacks kept me in the right direction. I am enormously thankful to Professor Karavitis for all he has done so far and for his trust in me.

I will remain forever grateful to Professor Evan C. Vlachos, my former co-advisor. He was always available since the first day I arrived in Fort Collins, and his tremendous support enabled and enriched this process. This journey would not have been the same without him and I own him much more than few words can convey. Unfortunately, life circumstances did not allow Professor Vlachos to be a part of my effort up until the end, however I feel extremely lucky to have crossed paths with him, and thankful for his mentorship and friendship.

To my advisor Professor Darrell G. Fontane, I owe many thanks for his encouragement and understanding over these years. I appreciate his great help and guidance, which were pivotal for the completion of this dissertation.

Professor Reagan M. Waskom, the Director of the Colorado Water Institute (CWI), and co-advisor in my research, had a key role in my success. During the years I worked as a research assistant at CWI under his supervision, he gave me the opportunity to be involved in several research projects and taught me a lot about the water resources challenges that the western United States faces. Professor Waskom empowered me to address some of these issues that amalgamated to this dissertation. I am extremely grateful for his continuous help, support, guidance, and mentorship throughout the course of this research.

I would like to thank my committee member Professor Neil S. Grigg for his insightful comments, which drastically enhanced the whole dissertation, as well as my committee member Professor Charles W. Anderson for always being available.

All in all, this dissertation would not have been completed without the contributions from the important individuals discussed above.

Furthermore, I would like to acknowledge *National Science Foundation* (under Grant No. CBET-1240584, AirWaterGas Sustainability Research Network) and *The State of Colorado* (through the House Bill HB12-1278, South Platte Alluvial Aquifer Study) who partially funded this research.

I would like to extend my thanks to Ms. Nancy J. Grice and Ms. Julie A. Kallenberger for their help and collaboration, as well as to the rest of the CWI staff, which was a pleasure to work with over the years.

A great gift from this journey is the long-lasting friendships. Within this period, I had the chance to meet exceptional people from all around the world that enriched my life and expanded my perspective. I want to acknowledge Dr. Ayman Alzraiee, the first person I met at CSU, and a great friend ever since. He provided encouragement and support on the idea of the groundwater level time series filling methodology and supplied me with the data from the regional groundwater model. Colleagues in Greece were also of great help, especially Dr. Demetrios Tsesmelis, Dr. Nikolaos Skondras, and Dr. Vassilia Fassouli. In addition, many thanks go to Dr. Haw Yen, Professor Jorge Gironás, Dr. Iordanis Moustakidis, Dr. Francesco Scavone, Mr. Prajwal Gautam, Dr. Nikolaos Chorianopoulos, Dr. Andreas Kakolyris, Dr. Galatios Moschonas, Mr. Dragan Krstić, Dr. Spyridon Thanasoulas, Dr. Marotessa Voultsidou, Mr. Demetrios Kounalakis, Dr. Melina Zempila, Mr. Athanasios Panagiotidis, and Mr. Vassilıs Angelis for their support when it was needed the most, and for making life outside graduate school more enjoyable. Additionally, I extend my thanks to Professor Mohammed Abdulrazzak, Professor Elpida Kolokytha, Professor Athanasios Loukas, Professor Stavros Alexandris, Professor Tapio Katko and Professor Rodrigo Maia for their friendship and moral support. My amazing friends back home also provided essential moral support including: Mr. Nikos Mallis, Mr. Spyridon Naksakis, Mr. Loukas Papathanasiou, Ms. Athanasia-Nancy Oikonomou, Mr. Charalampos Kakolyris, Mr. Panagiotis Mpekatoros, Ms. Georgia Mouriadou, and Ms. Katerina Matragkou.

My special thanks are due to Dr. Didem Alkan for proof reading the manuscript and her tremendous support during the last stages of this dissertation.

Finally, I would like to express my everlasting gratitude to my parents Demetrios and Triantafyllia Oikonomou, and my sister Elpida Oikonomou, for their unconditional love and the sacrifices made throughout these years, as well as for their moral and financial support.

DEDICATION

To my beloved parents and sister, Demetrios, Triantafyllia, and Elpida,

(Στους αγαπημένους μου γονείς και αδελφή, Δημήτριο, Τριανταφυλλιά και Ελπίδα)

and

to Professor Evan C. Vlachos,

a wonderful and kind man, an exemplary scholar and empowering mentor. I will always treasure everything he has taught me and be grateful for his guidance, support, and friendship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF TERMS

The following glossary provides the necessary information to familiarize the reader to specific notions in Colorado's water law, as well as defining terms used in chapter 3.

Municipal Use	Water from a municipal or quasi-municipal entity that is treated or considered wastewater effluent. Leased water must be used inside existing service area.
Industrial Use	Water that is leased or purchased from an existing industrial water use.
Agricultural Use	Agricultural water rights often undergo a change in type of use from irrigation to municipal and industrial through the water court process. Water can be leased or purchased, but changed use must include industrial.
Free River Diversion	Water that is diverted during times of excess water. The diversion needs to be in priority and is likely not a consistently reliable source of water for oil and gas development. Some reservoir owners have stored water during times of free river for oil and gas companies. Because this type of water is diverted and stored at a time during which there is no call on the river, the water can be used for any beneficial use, including oil and gas activities.
Tributary Groundwater	Groundwater that has a hydrological connection to the surface stream system. Each well permit requires an augmentation plan or substitute water supply plan for replacement of water.
Nontributary Groundwater	As defined by Colorado Senate Bill 213, effective July 6, 1973 is the "groundwater, located outside the boundaries of any designated ground water basin in existence on January 1, 1985, the withdrawal of which will not, within 100 years, deplete the flow of a natural stream, ...at a rate greater than one tenth of one percent of the annual rate of withdrawal." [C.R.S. 37-90-137(4) and 37-90-107(7)]
Nontributary Groundwater located in the Denver Basin and Designated Basins	New sources of nontributary groundwater which have a limited hydrological connection to surface water. Nontributary groundwater is not administered within the legal priority system and does not require replacement of water.

New Determinations of Nontributary Groundwater	Nontributary groundwater located outside of the Groundwater located in designated groundwater basins which has a limited hydrological connection to surface water. This water is not administered within the legal priority system and does not require replacement of water. This is a relatively new source of water for oil and gas operations in the Denver-Julesburg Basin.
Multi-use water	Return flows from the first use of water are claimed by the Northern Water District for successive uses by other downstream water users
Produced water	Water that is native to the surrounding formation and is produced throughout the oil and gas recovery process. Produced groundwater is a diversion that is subject to administration and well permitting by the Colorado Division of Water Resources. The State Engineer has authority to conduct rulemaking to determine whether groundwater in specific oil and gas formations underlying certain geographic areas is nontributary. Pursuant to section 37-90-137(7), C.R.S., an oil and gas company who withdraws nontributary groundwater during the mining of minerals is not required to obtain a water well permit, unless the produced groundwater being removed will be beneficially used [a,b].
Flowback water	Flowback is water that is returned to the surface during hydraulic fracturing and may be recycled and reused for subsequent needs, depending on its quality. Recycled flowback water is not a new source of water because it was previously acquired for oil and gas well development. Note: Form 5A classifies "Flowback Volume Recovered" to include treatment fluids and produced water.

[a] This statement is not applicable to coalbed methane wells because the Colorado Supreme Court found that the withdrawal of water in the coalbed methane extraction process is, in and of itself, a beneficial use [*Colorado Supreme Court*, 2009]. Therefore, each coalbed methane well that produces water must be permitted as a water well by the Colorado Division of Water Resources.

[b] While a water well permit is required if the nontributary ground water is put to a beneficial use, a provision, in section 37-90-137(7) C.R.S., states that no water well permit is required if the nontributary ground water is used for the mining of minerals in the same geological basin.

1 General Introduction

1.1 Overview

Water permeates natural, social, and economic systems, thus being the link between many sectors and needs [Grigg, 2017]. Early on, planning and management of water resources revolved around the themes of development and preservation [*Vlachos and Hendricks*, 1977]. Over the last decades, the concept of Integrated Water Resources Management (IWRM) has gained worldwide acceptance and its adoption has fundamentally changed water planning and management processes. Water planners, managers, and engineers acknowledging the increasing complexity and the intensifying water quantity and quality challenges, which were originating from economic development, expanding populations and extreme climatic events, have shifted from practicing fragmented approaches towards the paradigm of holistic management that the IWRM framework is representing. The integrated notion is underscoring that the water, land and related resources are interconnected sub-systems, but also emphasizes that the common development and management goals should not compromise the ecosystems and that the maximized socio-economic benefits of this process need to be allocated in an equitable way [*Global Water Partnership*, 2000]. The IWRM framework accounts for the interlinkages of the different water uses, resulting in cross-sectoral integration (Figure 1).

The implementation of the IWRM principles at the river basin level it has come to be known also as Integrated River Basin Management (IRBM) and can also assist in sustainable development efforts by delivering a triple bottom line of economic, social, and environmental benefits

[*Grigg*, 2017]. The intricate interconnections of the physical and socio-economic systems along with the diverse natural, socio-political and economic challenges, as well as development priorities from case to case, do not allow for a common implementation blueprint [*Global Water Partnership*, 2000; *Lenton and Muller*, 2009; *Pegram et al.*, 2013; *Grigg*, 2017]. The challenges can be found in the institutional part of IRBM which the stakeholders and institutional organizations can differ drastically in each basin. Furthermore, the political boundaries rarely coincide with river basin boundaries, hampering integration efforts [*Coelho et al.*, 2011]. At the same time, the transformations taking place in a river basin level are constant and manifold, affecting different subsystems, creating a complex and rapidly changing environment that water managers and planners need to respond to. Vlachos [1999] characterized such conditions with the composite term of "raplexity", which reflects the complexity and the rapidity of change. Given the raplexity taking place in river basins it is difficult to attain the triple bottom line of economic efficiency, social equity and environmental sustainability, which renders IWRM implementation an iterative process.

Decision Support Systems (DSS) are among the tools water scientists and water managers use in order to inform the general public, stakeholders and decision-makers when they need to make water resources related choices and describe the associated tradeoffs [*Grigg*, 1996, 2017; *Global Water Partnership*, 2013; *Fontane*, 2016]. Monitoring programs are the main source of providing raw data to the DSS [*Fontane*, 2000]. Indeed, a great volume of raw data regarding the environment, are collected and stored in various databases in order to be utilized for planning and management purposes. Technological advances improved the quality and accuracy of measurements and have made inexpensive data collection and storage possible. This has allowed the sampling of datasets that would be considered as secondary a few decades ago. Likewise, the temporal and spatial sampling resolution of data measured by satellites has increased, now permitting real-time analysis for small geographical extents [*Thenkabail*, 2015].

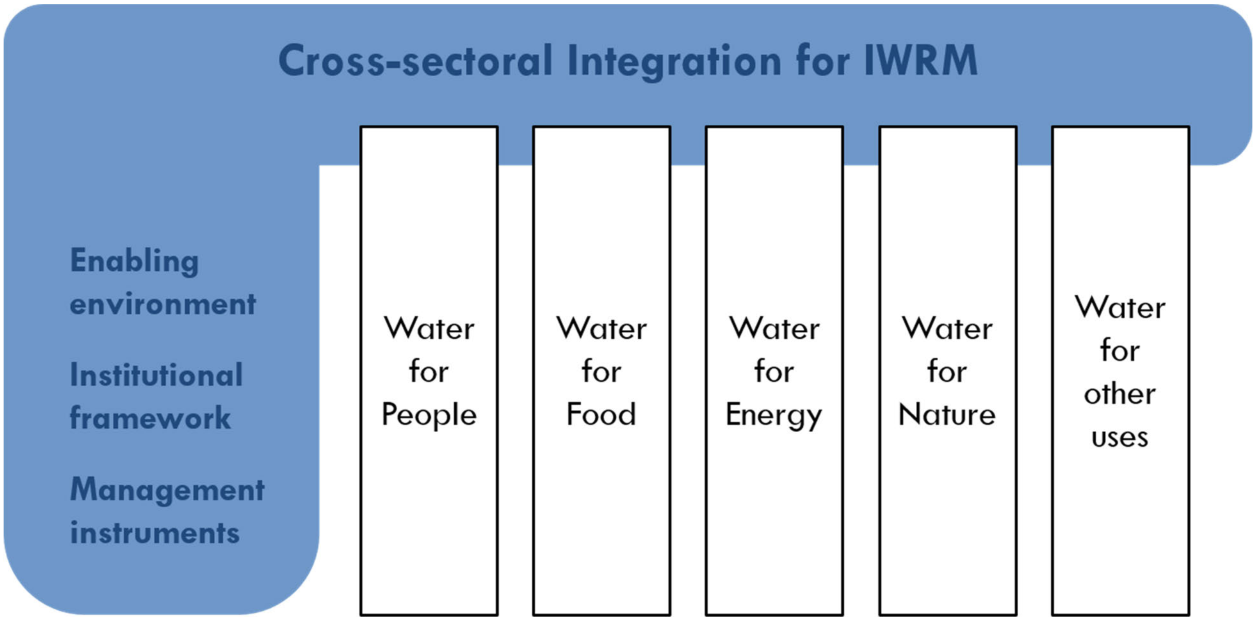


Figure 1: Integrated water resources management (IWRM) and its relations to sub-sectors [Global Water Partnership, 2000]

Public participation and stakeholder involvement is a key element in the IRBM process for deciding the development priorities and the common goals in a river basin [*COROADO*, 2014]. Fontane [2000] stresses that the collected data and the information must serve all the divergent communities that might need it. Transparency in data collection, the methods used for their transformation to information, and the dissemination of the information is a way to build trust. Thus, due to the growing raplexity and increasing amount of data collected, along with the different needs of the various groups involved in the decision process, information systems are an essential element in IRBM. The creation of a knowledge base in basin level study that would contain information about the hydrological cycle, the ecosystem, hazards, socio-economic activities etc., and would assist in water resources assessment, is key to effective water management [*Global Water Partnership*, 2013]. Despite the modern technological capabilities and the continuously increasing quality of the collected data, the complexity of managing the information is also increasing [*Dalcanale et al.*, 2011].

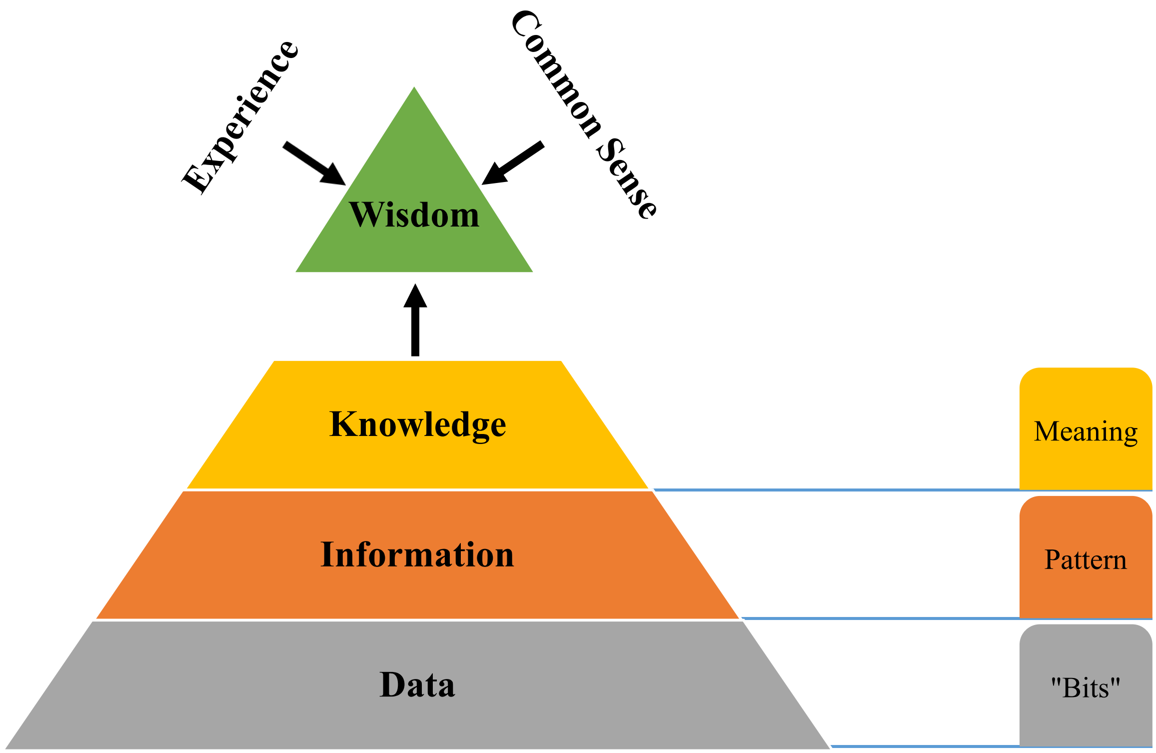


*Figure 2: Hierarchy of information and knowledge [*Grigg*, 2005; *Karavitis*, 2007; *Vlachos*, 2008]*

The IRBM process has decision steps in which information and knowledge generated by data and models reveals the implications of an array of alternative actions and examine the associated trade-offs. In order to move up into the "information and knowledge" hierarchy [*Grigg*, 2005; *Vlachos*, 2008], illustrated in Figure 2, and have an informed decision making process, the available raw data must be transformed first into patterns (isolated information), and then into meaningful knowledge (general trends). The attained knowledge from data analysis, modeling and interpretation can ultimately lead to wise decisions for the problem/issue at study, through experience and common sense. Thus, the tremendous amounts of data in an array of scientific disciplines (hydrological, atmospheric, environmental, ecological etc.) and the availability of operational, financial, and socio-economic datasets are presenting unique problem-solving opportunities for water resources planning and management issues. At the same time these opportunities are being followed by new challenges.

While the fact that the awareness has increased dramatically regarding water resources related issues and multiple paradigms, contexts and methods that have being introduced, there are still great challenges in water resources management in practice [*Vlachos*, 2008]. One of the sources of these challenges in IWRM process has to do with data and information issues. There are several data inefficiencies or the data are such that they cannot be used, or cannot be accessed. Quite often water resources information is scarce; fragmented; outdated; not suitable; wrong data analysis methods are used; lack of evenly spatial-temporal information of water resource conditions; which in practice hinders proper assessment of the resources and increases the risk of making suboptimal decisions and taking management actions that could have adverse impacts [*Global Water Partnership*, 2013]. In the context of IWRM and data use, another important aspect is the challenges of integrating different sources of data. All aforementioned factors may affect the transformation of data to accurate and meaningful information for decision-making.

The overarching challenge, still present, is to extract the most information from data and models and utilize them in the decision making process [*Grigg*, 1985]. The challenges of ascending the "information knowledge" hierarchy related to data hindrances could be classified in six main types, as illustrated in Figure 3. Despite the abundance of data, there are cases where when the data exist but cannot be accessed or the wrong kind of data are collected (data out of institutional scope, low institutional priority, etc.). The second category of hindrances is related to data quality. It is very common for the datasets of interest to include gaps (intermittent or continuous) or their collection does not cover the intended period of study, which impedes their use in decision making processes and their statistical analysis without proper treatment to avoid spurious results. Furthermore, data quality as a term can refer also to the spatial and temporal resolution of the data, and/or the existence of errors and duplicates entries. Technology has transitioned hard copy data into digital formats. Sources of data may not be used effectively due to poor documentation and/or metadata, which may result in their misuse and misapplication. Despite the collected data being available online, there are cases that do not offer ways for user-friendly acquisition. At the same time, when online queries are possible, the available digital data can be in databases structured in such a way that make the retrieval of specific attributes difficult due to fragmentation. Another common challenge is data uncertainty, including uncertainty inherent in the datasets and uncertainty introduced through data manipulation (interpolation, etc.). In the era of "big data", a critical challenge is surfacing on how much sense can be made out of huge volumes of data and identify opportunities for their integration, allowing to further the understanding of water systems and their interconnections with social systems. Therefore, there is still the need to develop more systematic procedures for both data collection and methodologies, in order to improve the transformation of data to information and knowledge, with the ultimate goal to assist integrated water resources planning and management.

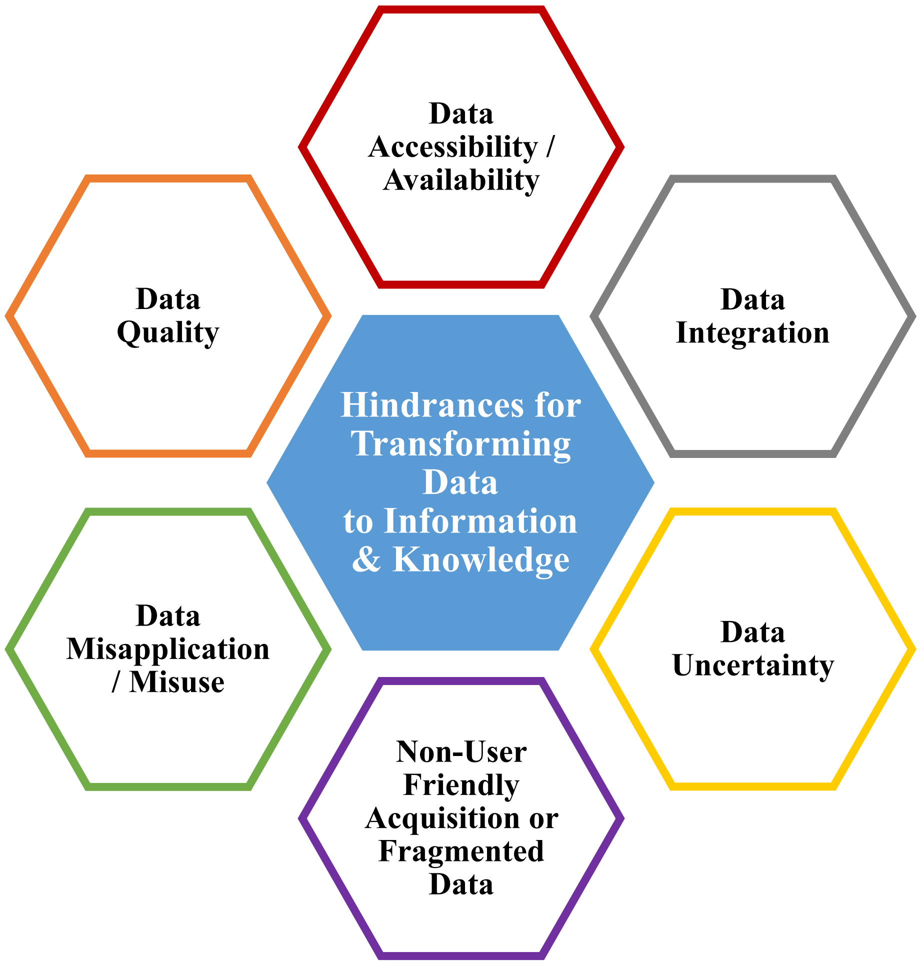


Figure 3: Hindrances for transforming data to information and knowledge

Vlachos [1999, 2008] has underlined the importance of new methodologies that would improve the understanding of the current complex setting and the need of combining data, information and judgment as ways to deepen the integration process in water management efforts. In this context, the pertinent research effort, investigates methodologies, in different - but connected - areas of water resources management for advancing the understanding of water systems through transforming data to information with an ultimate scope to aid IRBM practices. The areas comprising the present research are within the broader topics of drought, water demand, and groundwater supply drought, water demand and groundwater supply. The diversity of the studied topics and their real-world case studies give the opportunity to confront a variety of the aforementioned hindrances for transforming data to useful information. However, their unified thread is the common quest for IRBM under changing water resources conditions. The three research cases described in this dissertation have been presented to academic conferences and one has been already published to a refereed scientific journal. More specifically, initial aspects of the formulation of a more quantitative integrated way of capturing drought vulnerability were presented in the annual conference of the American Water Resources Association (AWRA) [*Karavitis et al.*, 2015]. The second research effort is part of a larger National Science Foundation research project and different parts of this work have been presented in AWRA and American Geophysical Union (AGU) conferences [*Oikonomou et al.*, 2015c, 2015b]. Some of the results have been published at the Journal of Environmental Management [*Oikonomou et al.*, 2016]. Initial results for part of the last research effort was presented at a national AWRA conference [*Oikonomou et al.*, 2015a].

IRBM attempts to take into account the interdependencies of water systems with other systems such as ecological, economic, infrastructure and social, when decisions are made. This decision-making environment is highly complex, requiring the use of a wide array of different datasets and their transformation into useful information. Hydroclimatic uncertainty is an additional factor that contributes to the complexity of water resources planning and management. In the context of stressed systems, drought events are challenging our ability to cope with their impacts that spread into the different components of human systems. Vlachos and Braga [2001] emphasize that improvements should be made into methods for risk assessment and vulnerability analysis related to water resources. The difficulty of capturing the intensity and the extent of the drought hazard hinders the identification of areas that are most vulnerable. Hence, information about drought vulnerability can assist decision makers in taking anticipatory actions and developing drought strategies. At the same time, a system's vulnerability is affected also by nonphysical components such as the efficacy of the available water infrastructure and supply capacity and demand patterns. Since a deeper understanding of the different dimensions of the system is needed in order to inform appropriate decisions, data play a critical role in describing the conditions and understanding the nature of the constraints (physical and anthropogenic).

While extreme dry variations of the physical system are imposing multifaceted risks, impacts and challenges onto human systems, the increasing population growth, conflicting and competing water uses within river basins with limited water supplies, are further stretching the system's ability to meet all the different water demands in an efficient and the effective manner. Data that potentially influence water demand trends in a river basin are important knowledge that can aid in decisions that focus on balancing water supply and demand. Such an example is the recent boom of unconventional oil and gas exploration in the United States adding another water demand on systems already in stress. This "new" water demand that has emerged as a result of technological advancements in unconventional oil and gas development and has received high attention since it is considered fully consumptive, it rarely returns to the system, but also due to concerns about potential risk for groundwater contamination [*U.S. Environmental Protection Agency*, 2012; *Vengosh et al.*, 2014].

During drought periods, sustainable conjunctive use of surface and groundwater resources is often essential for meeting water demands. Furthermore, groundwater and surface water systems integration is one of the main objectives of IWRM. Despite increased computational prowess and data availability of a plethora of relevant information (measurements, satellite derived data etc.) supporting the advancement of novel tools and methods in water resources, there are still gaps that must be filled. Information about groundwater resources is still an area where data are lacking both in spatial and temporal extent. The development of methods and tools that enhance our comprehension through data augmentation is essential for better water planning and management actions. The western United States is a region characterized by limited water resources with frequent drought intervals. The vulnerability of the western physical and human systems to drought is an open question.

The present research is inspired by the aforementioned topics, and tries to close some of the knowledge gaps by proposing methodologies that advance the understanding of water systems from an IWRM perspective, and provides solutions to existing challenges and offers recommendations for further research.

1.2 Research Objectives

The **overarching objective** of this study is to develop tools and approaches to overcome data obstacles in water resources management. The methodologies developed in this research take advantage of the constantly increasing available water resources related datasets in a way that address the aforementioned challenges in a manner that would be useful to water resources planners and managers. The interrelations and interactions of the elements of the study are illustrated in Figure 4.

Drought management is a key element in regional planning, and understanding the way that different components contribute to drought vulnerability is an important step towards the development of integrated management strategies and policies. The **first objective** of this study is to develop a framework that advances the quantification of drought vulnerability based on an integrated approach including aspects of the physical, structural, and socio-economic components.

This was achieved by evolving the Standardized Drought Vulnerability index [*Karavitis et al.*, 2014], from a qualitative measure to a quantitative one. Components that contribute to drought vulnerability include demand, supply, infrastructure and impacts were advanced to incorporate satellite derived information, operational model outputs, measured data and some qualitative information. The approach not only allows for increased spatial representation of the combined drought vulnerability, but it can also be used to trace weak components of the system and try to create action to enhance their individual vulnerability status. In addition, it was attempted to create a spatiotemporal unified representation of the multifaceted drought impacts that could inform decision makers and planners about their magnitude and extent.

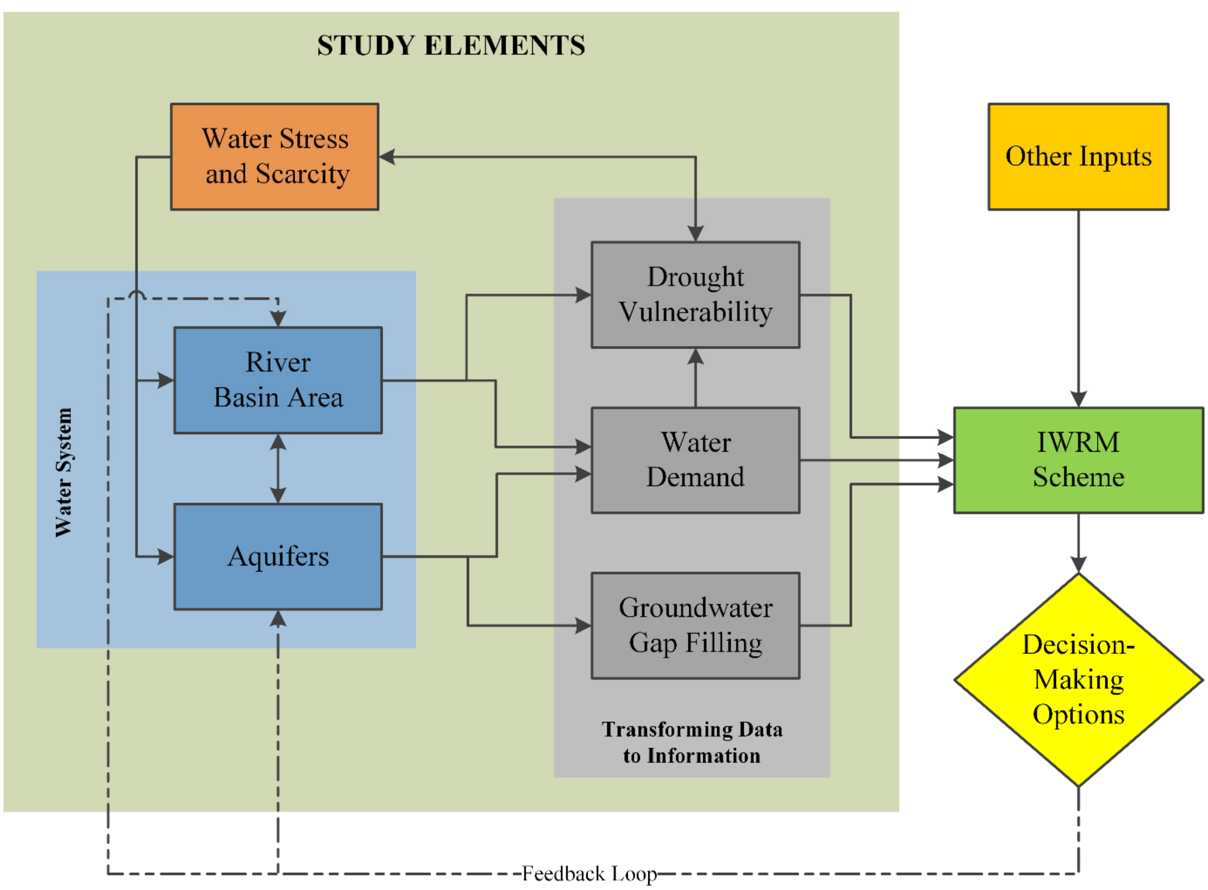


Figure 4: Central components of the study with their interrelations and interactions

New trends and emerging water demands can change the setting and lead towards greater uncertainty about the efficacy of established drought plans. The quantification of water demand, the analysis of its characteristics, the potential effects and tradeoffs between traditional water demands is crucial knowledge. The **second objective** of the research is to develop a methodology for quantifying and analyzing water demand for unconventional energy development. Since there is uncertainty regarding how much water is needed for the completion of a hydraulic fracturing operation, first the characteristics affecting water use intensity were identified. By demarcating the extracted information into standardized measures, the comparison of wells in different regions is possible. Analyzing a new water demand in different scales (county and HUC-12 level) gives the ability to quantify its share and understand better possible tradeoffs and effects on a local scale. In addition, the average water recovered for each formation is quantified and its handling was identified. The last two are important pieces of the analysis since can indicate reuse opportunities and treatment challenges.

Furthermore, conjunctive use of both surface and groundwater resources during occasions of water stress and scarcity is essential, but many times the actions taken are limited by imperfect knowledge of groundwater dynamics caused by missing observation measurements. Thus, the **final objective** is to develop and test a novel methodology for bridging data gaps in groundwater level measurements by taking advantage of available groundwater models and their outputs in a framework that includes statistical modeling and the Ensemble Smoother. The testing of the proposed computationally low cost framework was achieved by three numerical experiments performed in 18 spatially dispersed wells. The design of the numerical experiments was created to represent different patterns of missing data.

1.3 Contributions

The contributions of the current research are manifold and are not restricted to one area of water resources planning and management. The **first contribution** is the evolution of an integrated drought vulnerability index to an operational stage through the incorporation of spatially quantitative datasets. Although the application was performed in an area of data abundance, alternatives are presented for expanding the index on less data-rich settings. The new approach for calculating the index can lead toward an efficient, timely and effective decision-making in cases of drought events.

An **additional contribution** demonstrated by the study was the importance of fully understanding the implications of new water demands that compete with traditional ones. Providing ways of evaluating the magnitude of the new water demand in different geographical scales can inform about implications visible only on a local scale. Significant was the demonstration of the different factors influencing water intensity and the comparison of different shale formations. The study was also able to identify data gaps hindering the full understanding of the water cycle (source-use-handling) and the need for additional data and reporting harmonization policies across the US.

Finally, a **major contribution** is the proposal of a novel framework for bridging data gaps in groundwater level measurements in alluvial aquifers. The suggested methodology is an innovative, easy to implement and computationally low cost approach for filling gaps in groundwater level time series. At the same time, it augments the value of readily available regional groundwater models since their information is used in the method. The proposed approach is shown to provide satisfactory results for the first two numerical experiments that were tested, offering a viable solution in retrieving missing information about groundwater levels. The design of the framework is not geographically bound and thus can easily be applied to any modeled alluvial aquifer. Its transferability is guaranteed since no aquifer specific parameters are required. This key characteristic is making the method suitable for providing to water planners and managers a more complete picture about local groundwater variability and trends, which is a vital component for integrated water resources management.

1.4 Structure of the Study

Apart from the described sections, the rest of the manuscript is comprised of four more chapters. The design is schematically presented in Figure 5. Chapter 2 focuses on developing and presenting an implementable drought management policy tool through improving the Standardized Drought Vulnerability Index. The components of the index are presented and then the modification in its calculation is described in detail along with the specific datasets used for its application in the South Platte. The results of the components and the overall index are presented and discussed. A summary of the approach is given in the last section along with further testing and future work propositions.

The next chapter discusses the emergent water demand for unconventional oil and gas in Weld County, since this is the county with highest well concentration in the South Platte basin, and is compared with Garfield County to examine the importance of localities. The water development patterns and water use intensity are presented based on state data. The study moves between well scale, HUC-12 level and county scale for drawing implications in these different levels. The end of the life cycle of the oil and gas water is revealed with the challenges and opportunities for water reuse. Data discrepancies in the publicly available databases are discussed and concrete suggestions are made regarding additional data to be collected and ways to improve data reporting.

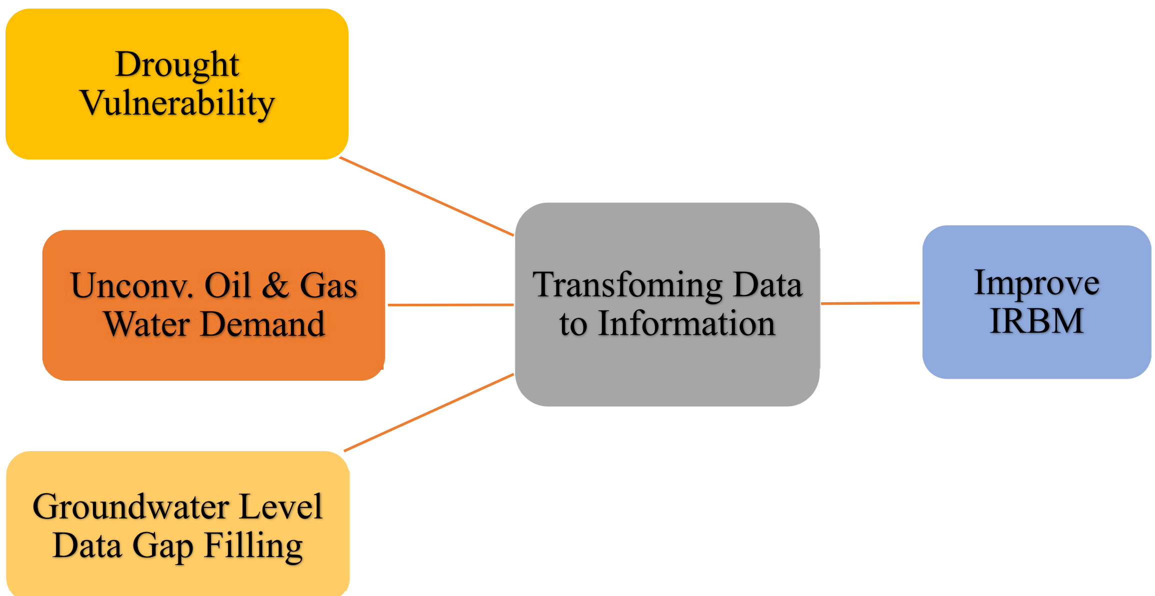


Figure 5: Architecture of the Study

Chapter 4 is introducing a new method for filling data gaps in time series of groundwater head measurements. A thorough review about the use of the techniques in water resources that are included in the framework pinpoint the novelty of the approach. The methodology is taking advantage of an already built regional groundwater physical model, an exogenous seasonal auto-regressive moving average model (SARIMAX) and the Ensemble Smoother. The applicability and the performance of the proposed framework was tested by three numerical experiments in eighteen groundwater wells of the South Platte alluvial aquifer. The results of each numerical experiment are discussed in details followed by alternative schemes to be tested in the future.

The final chapter of this study includes a summary of the research, the key findings of the presented approaches and summarized future research suggestions.

References

Coelho, A. C., D. G. Fontane, E. C. Vlachos, and R. Maia (2011), Delineation of Water Resources Regions to Promote Integrated Water Resources Management and Facilitate Transboundary Water Conflicts Resolution, in *Transboundary Water Resources Management*, edited by J. Ganoulis, A. Aureli, and J. Fried, pp. 261–267, Wiley-VCH Verlag GmbH & Co. KGaA.

COROADO (2014), *Deliverable 2.3. Report on the Workshops Structure, Development and Findings*, http://www.coroado-project.eu/

Dalcanale, F., D. G. Fontane, and J. Csapo (2011), A General Framework for a Collaborative Water Quality Knowledge and Information Network, *Environmental Management*, *47*(3), 443–455, doi:10.1007/s00267-011-9622-7.

Fontane, D. G. (2000), Monitoring Water Systems: The USA Experience, in *Transboundary Water Resources in the Balkans*, edited by J. Ganoulis, I. L. Murphy, and M. Brilly, pp. 193–201, Springer Netherlands.

Fontane, D. G. (2016), *Water Resource Systems Analysis (CIVE 546)*, Class Notes of Spring 2016, Department of Civil & Environmental Engineering, Colorado State University, Fort Collins, CO.

Global Water Partnership (2000), *Integrated Water Resources Management*, Global Water Partnership (GWP), Stockholm.

Global Water Partnership (2013), *The Role of Decision Support Systems and Models in Integrated River Basin Management*, Technical Focus Paper, Global Water Partnership (GWP), Stockholm.

Grigg, N. S. (1985), *Water Resources Planning*, McGraw-Hill, New York, NY.

Grigg, N. S. (1996), *Water Resources Management: Principles, Regulations, and Cases*, 1st edition, McGraw-Hill Professional, New York.

Grigg, N. S. (2005), *Water Manager's Handbook: A Guide to the Water Industry*, AquaMedia Publishing, Fort Collins, CO.

Grigg, N. S. (2017), Integrated River Basin Management, in *Handbook of Applied Hydrology*, edited by V. P. Singh, McGraw-Hill Education.

Karavitis, C. A. (2007), *Water Resources Management,* Class Notes of Spring 2007, Graduate Program of Water Resources and Environmental Management, Departement of Natural Resources Development and Agricultural Engineeering, Agricultural University of Athens, Athens, Greece.

Karavitis, C. A., D. E. Tsesmelis, N. A. Skondras, D. Stamatakos, S. Alexandris, V. Fassouli, C. G. Vasilakou, P. D. Oikonomou, G. Gregorič, N. S. Grigg and E. C. Vlachos (2014), Linking Drought Characteristics to Impacts on a Spatial and Temporal Scale, *Water Policy*, *16*(6), 1172–1197, doi:10.2166/wp.2014.205.

Karavitis, C. A., P. D. Oikonomou, R. M. Waskom, D. E. Tsesmelis, C. G. Vasilakou, N. A. Skondras, D. Stamatakos, S. Alexandris, and N. S. Grigg (2015), Application of the

Standardized Drought Vulnerability Index in the Lower South Platte Basin, Colorado, in *2015 AWRA Annual Water Resources Conference,16-19 November 2015, Denver, CO*.

Lenton, R., and M. Muller (Eds.) (2009), *Integrated Water Resources Management in Practice: Better Water Management for Development*, Earthscan, London, UK.

Oikonomou, P. D., A. H. Alzraiee, and R. M. Waskom (2015a), Evaluating the Performance of the Ensemble Smoother for Filling Missing Groundwater Head Measurements, in *2015 AWRA Annual Water Resources Conference, 16-19 November 2015, Denver, CO*.

Oikonomou, P. D., R. M. Waskom, K. K. Boone, and J. N. Ryan (2015b), Unconventional Oil and Gas Development and its Stresses on Water Resources in the Context of Water-Energy-Food Nexus: The case of Weld County, CO, in *2015 Fall Meeting of the American Geophysical Union, 12-16 December 2015, San Francisco, CA.*

Oikonomou, P. D., R. M. Waskom, K. K. Boone, J. A. Kallenberger, E. N. Plombon, and J. N. Ryan (2015c), Water Use Impacts of Oil and Gas Development in Colorado, in *2015 AWRA Annual Water Resources Conference, 16-19 November 2015, Denver, CO*.

Oikonomou, P. D., J. A. Kallenberger, R. M. Waskom, K. K. Boone, E. N. Plombon, and J. N. Ryan (2016), Water Acquisition and Use during Unconventional Oil and Gas Development and the Existing Data Challenges: Weld and Garfield counties, CO, *Journal of Environmental Management*, *181*, 36–47, doi:10.1016/j.jenvman.2016.06.008.

Pegram, G., Y. Li, T. L. Quesne, R. Speed, J. Li, and F. Shen (2013), *River Basin Planning: Principles, Procedures and Approaches for Strategic Basin Planning*, UNESCO, Paris, France.

Thenkabail, P. S. (Ed.) (2015), *Remote Sensing of Water Resources, Disasters, and Urban Studies*, Remote Sensing Handbook, CRC Press, Boca Raton, FL.

U.S. Environmental Protection Agency (2012), *Study of the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources: Progress Report*, U.S. EPA Office of Research and Development, Washington, D.C.

Vengosh, A., R. B. Jackson, N. Warner, T. H. Darrah, and A. Kondash (2014), A Critical Review of the Risks to Water Resources from Unconventional Shale Gas Development and Hydraulic Fracturing in the United States, *Environmental Science & Technology*, *48*(15), 8334–8348, doi:10.1021/es40511.

Vlachos, E., and B. Braga (2001), The Challenge of Urban Water Management, in *Frontiers in Urban Water Management: Deadlock or Hope*, edited by Č. Maksimović and J. A. Tejada-Guibert, pp. 1–36, IWA Publishing, London, UK.

Vlachos, E. C. (1999), Environmental Issues and the Passage to the 21st Century: Challenges and Opportunities at the End of the Millennium, in *6th International Conference on Environmental Science and Technology, September 1999, Samos, Greece*.

Vlachos, E. C. (2008), *Technology Assessment and Social Forecasting (CIVE 639),* Class Notes of Spring 2008, Department of Civil & Environmental Engineering, Colorado State University, Fort Collins, CO.

Vlachos, E. C., and D. W. Hendricks (1977), *Technology Assessment for Water Supplies*, Water Resources Publications, Fort Collins, CO.

2 Developing Implementable Drought Management Policy Tools: Evolving the standardized drought vulnerability index

2.1 Introduction

The drought phenomenon has diachronically manifested its existence in almost every culture by affecting social and economic welfare to an extent that cumulatively exceeds any other natural disaster. Drought is a recurrent natural phenomenon striking many areas around the world defying their normal climatic conditions. The nature, characteristics and impacts of drought have continuously drawn the attention of the scientific community, state and federal entities, and the general public, resulting in the production of a rich menu of drought literature that mostly provides crucial information on its parameters [*Rosenberg*, 1978, 1980; *Hagman*, 1984; *Grigg and Vlachos*, 1993; *Fontane and Frevert,* 1995*; Karavitis*, 1998; *Wilhite*, 2004; *Cancelliere et al.*, 2005; *Traore and Fontane,* 2007*; Vasiliades and Loukas*, 2009; *Karavitis et al.*, 2012a; *Grigg*, 2014; *Karavitis et al.*, 2015].

Drought is a dynamic creeping phenomenon without a definition that may be widely accepted. As such, its holistic description is very demanding and consequently, once it occurs, it is seemingly difficult to confront being a non-event, that is, the absence of enough water [*Yevjevich et al.*, 1983; *Karavitis*, 1999; *Bordi et al.*, 2006; *Eriyagama et al.*, 2009; *Karavitis et al.*, 2014]. It has been noted that various terms exemplify a confusion among such concepts signifying “dry environments” or “water deficiencies” and these terms vary from desertification, to aridity, to drought and to water shortages [*Vlachos,* 1982] illustrated in *Figure 6*. Vlachos [1982] organized these concepts combining water availability manifestations (permanent vs. temporary) and causes of environmental transformation (human induced vs. nature produced) as is shown in the figure. In this respect, aridity signifies the stable natural state of an area; water shortage is mainly connected with temporary deficiencies and with limited areal extent created mostly by anthropogenic factors; desertification is a permanent anthropogenic phenomenon and finally, drought is a temporary climatic episode, which may occur regularly in a fixed or unpredictable pattern.

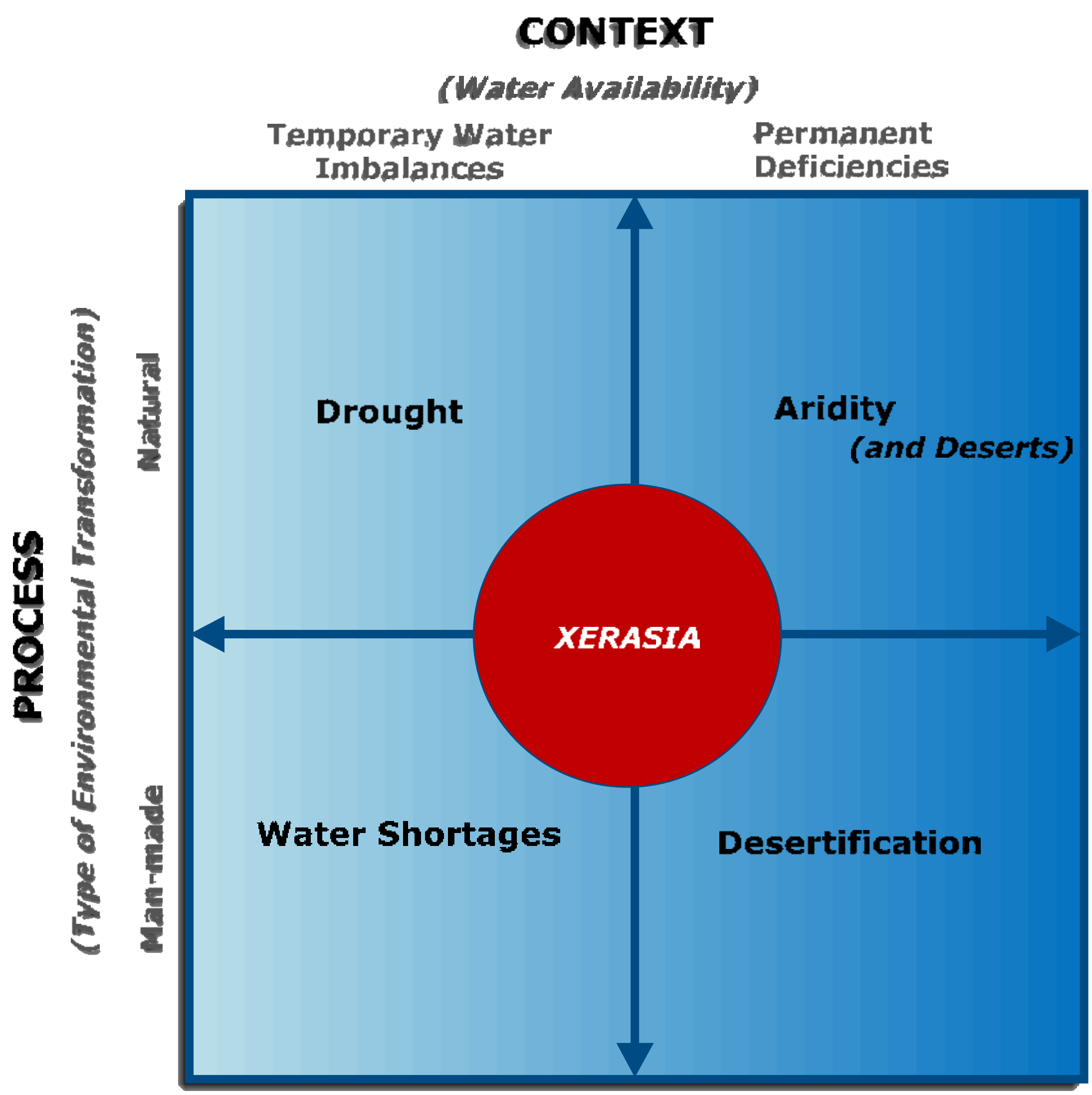


Figure 6: The "Xerasia" processes matrix [Vlachos, 1982; Karavitis, 1992; Karavitis et al., 2014]

Drought constitutes a rather severe hazard to all human activities [*Blauhut et al.*, 2015], and especially to water supply. The vulnerability magnitude of various areas to that particular hazard depends on their exposure to water deficiency and to the existing water management policy framework [*Karavitis*, 2012]. In this context, several attempts have been made in order to describe the different dimensions of the phenomenon such as severity, duration and impacts [*Changnon and Easterling*, 1989; *Byun and Wilhite*, 1999; *Shiau*, 2006; *Esper et al.*, 2007; *Wilhite et al.*, 2007; *Shiau and Modarres*, 2009]. Other attempts have chosen to focus on drought monitoring and forecasting [*Mishra and Desai*, 2005; *Cancelliere et al.*, 2007; *Belayneh and Adamowski*, 2012; *Kavalieratou et al.*, 2012; *Karavitis et al.*, 2015]. In all of these, the common thread seems to be the quest for the formation of holistic drought management schemes, incorporating contingency planning. Such emerging efforts have started to have a pivotal role in the determination of proactive management measures for mitigating the multifaceted adverse drought impacts [*Karavitis et al.*, 2015]. In this pursuit, numerous indices have been developed and some are used quite frequently. The Standardized Precipitation Index (SPI) developed by McKee et al. [1993] serves as such. Furthermore, some recently established indices focus on examining drought within a different context such as the Reconnaissance Drought Index RDI [*Tsakiris and Vangelis*, 2005; *Tsakiris et al.*, 2006] and the Standardized Precipitation Evapotranspiration Index [*Vicente-Serrano et al.*, 2010]; whereas, other more complex ones are referring directly or indirectly to vulnerability to water scarcity concepts, with prominent among them the Water Poverty Index [*Sullivan*, 2002].

Nevertheless, the concept of vulnerability to drought is a challenging one to display due to a disciplinary and/or individually based series of interpretations [*Gallopín*, 2006]. The vulnerability term is composed of two nascent elements, namely; hazard and impacts (Eq. 1) [*Karavitis et al.*, 2014]. In this respect, without a disaster or a system in stress (hazard) there is no vulnerability. Sometimes, exposure may be counted as a vulnerability component. However, it is mainly considered to link the system of interest to a specific disturbance [*Bohle*, 2001]. Hence, exposure may be a risk component [UNISDR, 2004]. Nevertheless, the vulnerability concept mostly refers to the components affecting both a system's capacity to cope and its potential to be harmed, while it is strongly influenced by a plethora of factors. Finally, vulnerability is not a static approach and it changes from time to time, complying with the changes that emerge in the various systems [*O'Brien and Leichenko*, 2001; *O'Brien et al.*, 2004; *Adger*, 2006; *Eakin and Luers*, 2006]. Consequently, the definitions of vulnerability adapt to those changes as well rendering vulnerability assessments quite challenging tasks.

$$\boldsymbol{Vulnerability} = \boldsymbol{F}(\boldsymbol{Hazard}, \boldsymbol{Impacts}) \qquad (1)$$

In the case of vulnerability to drought, the rainfall patterns, the supply and demand trends, and the socioeconomic background are the most important ones. Water demand is an important part of the societal vulnerability to drought and therefore the occurring supply deficits may affect the socio-economic development in some parts of the world. Water demand deficits may increase due to supply failures or because of sudden changes in land use patterns (i.e. increased irrigation requirements) or urgent population needs [*DeFries and Eshleman*, 2004]. At the same time, infrastructure status along with access to technology affect vulnerability levels [*Rosenberg*, 1980], for example efficient irrigation practices could counterbalance some of the effects of drought events. Under such conditions, some regions are more vulnerable than others are.

Drought characterization is a difficult task and thus, a classification system is used [*Grigg*, 2014]. Accordingly, characterization of vulnerability to drought becomes by default challenging since it is dependent to more information about impacts and risk of exposure. Rosenberg [1980] stresses that vulnerability to drought in one water user group has a ripple effect to the society at large. Evidence on this is the increased prices of agricultural products due to shortages attributed to the drought of 2011-2012 in the US [*Grigg*, 2014].

Despite the abundance of environmental information, drought impacts and factors contributing to vulnerability of systems are scarcely reported, even in developed states [*Blauhut et al.*, 2015, 2016]. Grigg et al. [2014] underlines the eluding nature of drought impacts since they could occur in different ways, times and places. Drought reports are generally characterized by fragmentation and lack of synthesis and thus do not give an integrated perspective of impacts [*Grigg*, 2014]. At the same time, statistical data gathered by the relevant authorities are not very detailed and have a coarse spatial scale since they are aggregated.

Duncan et al. [2015] argues that drought planning and management can be benefited by high spatial resolution analysis, since climatic and morphological conditions (e.g. rainfall patterns, plains, mountainous chains) can be masked in regional analysis. Likewise, factors affecting vulnerability to drought, such as demand and supply patterns, infrastructure efficiency, vary spatially too. Therefore, drought vulnerability assessment in high spatial resolution could serve as an additional water resources management tool informing managers and planners on priority actions.

In this respect, the present effort focuses on the assessment of drought vulnerability in the South Platte River Basin during the drought event of 2012. This event was selected since it affected a large portion of the US including a large part of the state of Colorado [*Grigg*, 2014]. The employment of indicators and indices is a common practice for describing complex phenomena and concepts. The SDVI (SPI-based Drought Vulnerability Index) is an integrated attempt for characterizing drought vulnerability based on a classification system. It was constructed by the Agricultural University of Athens, in the context of the Project Drought Management Centre in Southeastern Europe [Karavitis et al., 2011, 2014]. The application of the Index is visualized through geo-statistical methods in a GIS environment.

The index was initially applied in Greece and the procedure and implementation process is presented in detail by Karavitis et al. [2014]. In all previous applications of the SDVI, due to lack of information, the values of the last three components (supply, infrastructure and impacts) are on a basin or sub-basin scale (in the case of infrastructure, general values were used). In that respect, the study is attempting to evolve the SDVI to the next step of its development, which is a less qualitative and more transparent data rigorous method of calculating the index in cases of data abundance, with the ultimate goal to lead in an efficient, timely and effective decision making in cases of drought events. The main argument is that drought planning can be benefited, by the use of such a tool. It could convey drought information to decision makers in a holistic manner, avoiding existing practices of broken linkages and fragmentation of reported impacts. Better understanding the vulnerability levels the different components of a system are experiencing during occurred extreme events can guide decisions and target mitigation and adaptation actions thus allowing for an integrated management approach.

2.2 Area of Application

The assessment of the drought vulnerability index was performed in the South Platte River Basin of northeast Colorado (Figure 7). Its total area reaches 49,000 km^2 [*Dennehy et al.*, 1993], while the total population is estimated approximately at 3,700,000 inhabitants (70% of state population) [*Colorado Department of Local Affairs*, 2016]. Furthermore, the South Platte is a highly developed area with various competing and conflicting water demands. The agricultural sector is large with estimated coverage of the irrigated parcels in the basin in 2010 to be 3,426 km^2 (846,634acres) [*Colorado Decision Support Systems*, 2016]. It is a semi-arid basin with average annual precipitation of about 400 mm. According to McKee at al. [2000] the basin's climate in the upper section (mountainous area) is wet during December-April and dry in June and August-October, while the lower section is dry from November-February and wet from April-July. The dependence on snow accumulation during the winter months is significant since it serves as a natural storage for the basin to meet the demands during the summer period, along with several manmade storage structures.

Colorado has faced the impacts of droughts throughout the 20th century [*McKee et al.*, 2000], and most certainly with the 2002 and 2012 droughts. Additionally, climate change might increase the severity of future droughts [*Colorado Water Conservation Board et al.*, 2013]. During the most recent drought of 2012, Colorado experienced severe impacts, resulting to a damage of approximately $409 million in agricultural revenue [*Pritchett et al.*, 2013]. Thus, August 2012 was selected as representative time frame for the application of vulnerability assessment in the South Platte basin.

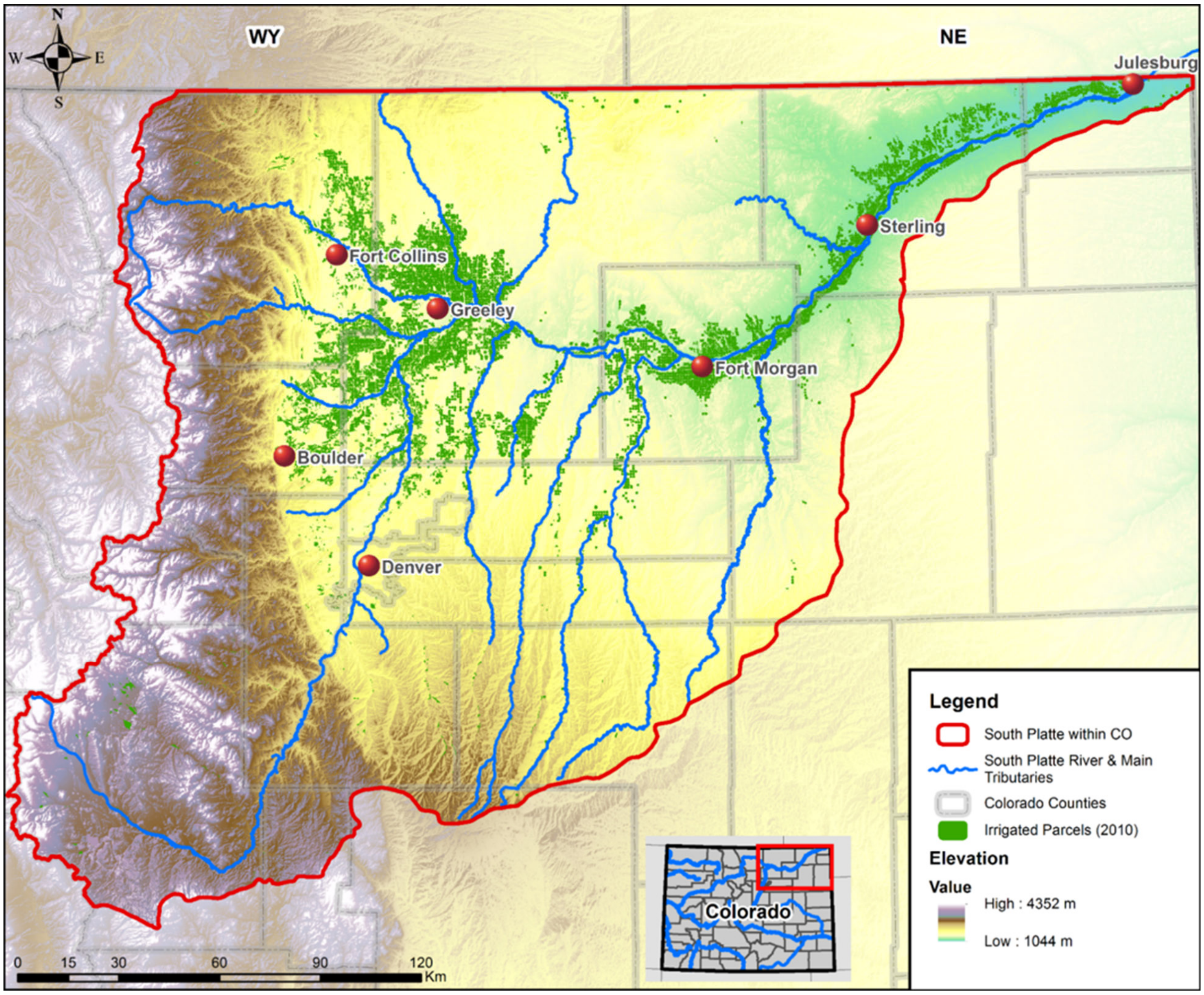


Figure 7: The South Platte basin within Colorado

2.3 Methods and Data

In spite of the uniqueness of decision making in a given time and locale, there is a unifying principle behind every successful decision - communication in a given organizational framework. Indicators are playing the role of the "channel" between slices of a complex reality ("sender") and decision makers ("receiver"). In addition, efficient communication requires clarity and simplicity. Thus, indicators have to simplify the complex interrelations of the reality and convey them in an unambiguous fashion. At the same time, they have to reflect information that should be derived indirectly from the stated properties. Therefore, they should represent the results of the links and the interactions between the natural system and the socio-economic development.

In this context, the advancement of the Standardized Drought Vulnerability Index to accurately depict occurring conditions is essential for use of this index for water resources planning and management under drought conditions. At the same time, incorporating into the information needed in order to calculate the index, datasets that are available for the whole globe are enhancing the value and applicability of the index. The general methodological framework of the SDVI is described below, which is followed by a detailed description of the data and tools used in the current effort.

2.3.1 The initial version of the Drought Vulnerability Index

The SDVI reflects a composite structure (index) which has been produced during the DMCSEE Project [*Karavitis et al.*, 2011, 2012b, 2014]. The SDVI aims at delineating an integrated estimation of drought vulnerability based on four drought manifestations namely: meteorological, hydrological, social and economic. Considering Eq. 1 that expresses vulnerability11, the hazard is expressed with the use of the 6-month and 12-month Standardized Precipitation Index, while the rest of the indicator's components demand, supply, infrastructure and socio-economic impacts are the affecting the impacts of a system.

$$\boldsymbol{SDVI} = \boldsymbol{F}(\boldsymbol{SPI_6}, \boldsymbol{SPI_{12}}, \boldsymbol{Demand}, \boldsymbol{Supply}, \boldsymbol{Infrastructure}, \boldsymbol{Impacts}) \qquad (2)$$

The six components SDVI are classified in four groups. The pertinent components embodied in the index are presented graphically in Figure 8 along with their interconnections in Karavitis et al. [2014].

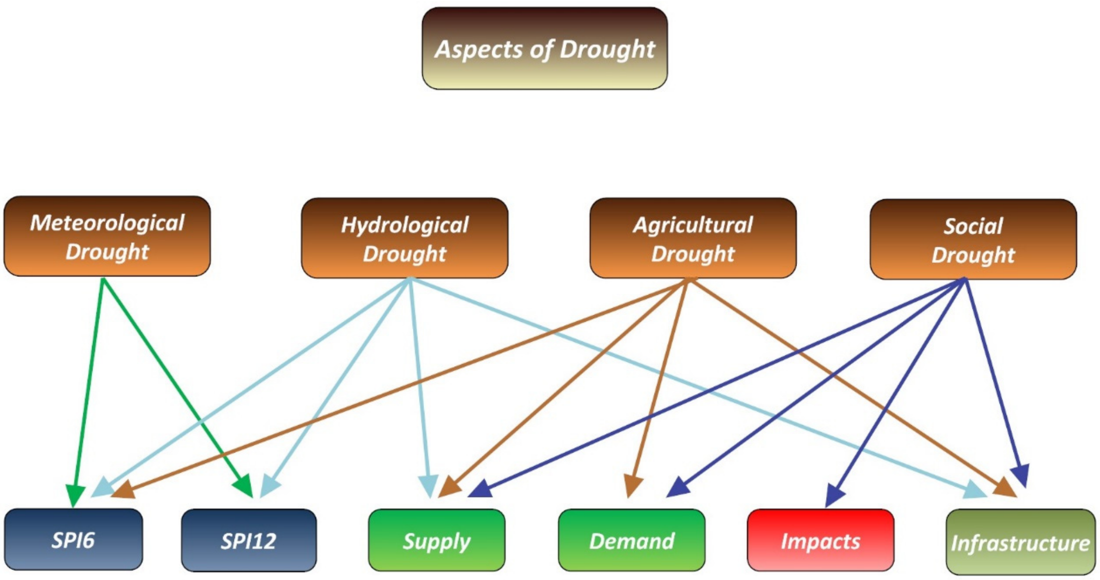


Figure 8: The relation between SDVI components and drought aspects [Karavitis et al., 2014]

i. The 12-month Standard Precipitation Index (SPI-12) provides information about the precipitation patterns over a long period (12 consecutive months) and thus it may reflect the urban/tourism, industrial, reservoir storage and hydropower water availability [*Karavitis et al.*, 2014]. The 6-month Standard Precipitation Index (SPI-6) compares precipitation over 6 consecutive months, thus may display more accurate seasonal variations and also reflects water availability during the crop growing season, especially for rain fed crops [*Karavitis et al.*, 2014; *Karavitis et al.*, 2012a]. Additionally the SPI-6 may display existing seasonality particularly in dry climates [*Wu et al.*, 2007]. It would seem that such a concurrent inclusion may promote links and interdependencies among the climatic precipitation parameters [*Karavitis et al.*, 2014].

The values of the SPI components are computed on a point (meteorological station) scale and then spatially interpolated with geostatistical techniques such as Kriging. Several other areas/points can be included in the process for a more adapted to the existing conditions visual calibration of the index to be deployed. Areas with zero drought vulnerability – such as mountain peaks, bare land etc. – may serve that cause [*Karavitis et al.*, 2014].

On the other hand, the SDVI is including other components (e.g. anthropogenic factors such as water demand, which is independent of the demand values of the adjacent points), thus such a correction is not desired. Hence, the Inverse Distance Weighing (IDW) should be chosen for spatial visualization of SDVI, since IDW assumes that each measured point has a local influence that diminishes with distance. It seems to fit more appropriately the SDVI spatial interpolation than Kriging for the rest of the index's components, when spatially detailed data are not available. In this context, any comparison of precipitation variation (SPI) with the vulnerability variation (SDVI) has to be made by comparing premises and outcomes, and not by applying the same spatial interpolation method, which in this case could not guarantee similar calculation errors. The SDVI may be computed for any given area should no precipitation data exist (no meteorological stations) providing that data for the remaining components of the vulnerability index are available.

ii. The Supply and Demand components illustrate the supply capacity deficits. Their respective scale lies on both the availability of water and the delivered volume for supply, and on changing demand patterns including various lifestyles and cultural traits. During normal conditions, supply satisfies the demand at all times. Then, according to the SDVI classification if no supply deficit occurs, a demand deficit will not also occur. Conversely, if a supply deficit of 15% occurs, an equal demand deficit may occur. However, this will not always trigger the input of a demand deficit in the Index calculation, if despite the reduction in supply; the demand may be still fully satisfied. For example, if the minimum water amount required per inhabitant is 1500 m^3/inhabitant/month and the maximum supply capacity is 1770 m^3/inhabitant/month, then a supply deficit of 15% (265.5 m^3/inhabitant/day) will not cause any deficit in demand coverage, since the remaining water quantity covers the demand ($1504.5 > 1500$ m^3).

iii. Infrastructure that portrays the existing level of infrastructure in terms of inadequacies (e.g. divergence from the designed supply capacity). The terms of infrastructure and supply may cause some times confusion, since e.g. a 15% infrastructure inefficiency may cause an equal deficit in supply. However, infrastructure status may not reflect this information in all cases. For instance, agricultural water infrastructure might be well developed and in excellent operational condition, nevertheless the limited water supply may be rerouted during a drought to an urban network so as to satisfy only the urban demand, leaving agriculture with significant deficits. In this context the two parameters do not reflect each other and have to be differentiated. Thus, the main role of this term is to capture the infrastructure adequacy status.

iv. The component of Impacts describes the inflicted drought damages caused due to the drought deficiencies in the supply/demand equilibrium. The component had focused on the costs forced on society. The impacts which are posed to the environment at previous applications of the index are not taken under consideration, since it was difficult to have a direct monetary representation, particularly with the data at hand [*Karavitis et al.*, 2014].

Table 1: SDVI components vulnerability scale [Karavitis et al., 2014]

Vulnerability Level	Classification of SDVI Components				
	SPI-6 & SPI-12	Supply	Demand	Impact	Infrastructure
Less Vulnerable (0)	≥1.50	0% Deficits	0% Deficits	0% Deficits	0% Deficiency
Vulnerable (1)	0 to 1.49	≤15% Deficits	≤15% Deficits	≤15% Deficits	≤15% Deficiency
Highly Vulnerable (2)	0 to -1.49	16-50% Deficits	16-50% Deficits	16-50% Deficits	16-50% Deficiency
Extremely Vulnerable (3)	≤ -1.50	>50% Deficits	>50% Deficits	>50% Deficits	>50% Deficiency

The main premise of SDVI is that the six components can coexist in an interdependent and complementary way. More specifically, the SPI components represent the hazard/risk element while the remaining four components represent the impacts element within the vulnerability concept. Hence, the SPI components describe the divergence of the accumulated precipitation from normal conditions and the remaining components describe both the effects of the precipitation deficits or surplus and the response conditions within the regions of interest. Thus, as also depicted in the works of Sullivan et al., [2003, 2009], Sullivan and Meigh [2007] and Sullivan [2002, 2011], a region with high positive SPI values may still display high drought vulnerability conditions, if it lacks the mechanisms and the capacity to exploit the available precipitation/water resources. For example, Greece exploits approximately 12% of the annually available water resources [*Barraqué et al.*, 2008]. Consequently, on one hand the country's vulnerability may be greatly affected by precipitation deficits, while on the other hand, precipitation surplus in some time periods cannot guarantee the country's safety against drought impacts. Thus supply and demand deficits are quite possible even when there is an abundance of water available, but it may not be usable for the societal needs due to the undeveloped water infrastructure. In this case, infrastructure development level and demand management may be the cornerstones towards drought impacts and vulnerability mitigation.

Continuing, the six parameters are classified according to their performance into the presented in Table 1 vulnerability categories (0 - 3 scale). Then, the vulnerability to drought is calculated in compliance with Eq. 3 [*Karavitis et al.*, 2014].

$$\mathrm{SDVI} = \sum_{\mathrm{i}=1}^{N} \frac{\text{Scaled Values of the Components}}{\text{Number of Components}} \quad , \text{ where N} = 6 \qquad (3)$$

That equation infers that the SDVI components are of equal importance. That particular technique has been chosen since it one of the most applied ones in the development of composite indicators despite the likelihood for some components to be over or underestimated [*Organisation for Economic Cooperation and Development*, 2008]. Furthermore, assigning weights on the components of a complex phenomenon could be a challenging task. A recent study focusing on assessing the uncertainty caused by different weighting methods on the SDVI index [*Tsesmelis et al.*, 2017] concluded that the equal weighting method is performing equally effective compared to more complex ones. The difference between the required and the supplied water quantities is the factor that determines the effects of the deficits. A supply deficit may cause no impact in demand coverage if the supply is still greater than the demand. Finally, the computed SDVI values are categorized into six classes of vulnerability as demarcated in Table 2.

Table 2: SDVI scaled values [Karavitis et al., 2014]

SDVI	Vulnerability Scale	Signal
0.00 – 0.49	No or Least Vulnerable	
0.50 – 0.99	Low Vulnerability	
1.00 – 1.49	Medium Vulnerability	
1.50 – 1.99	High Vulnerability	
2.00 – 2.49	Very High Vulnerability	
2.50 – 3.00	Extreme Vulnerability	

2.3.2 Index Adjustment and Application

The SDVI application in the South Platte basin is the first application of the index outside the geographical area of Southeast Europe (SEE), which is where it was first tested. The fact that SEE has different climatic conditions and levels of socio-economic development, makes it suitable for developing an index that would not be geographically constrained. At the same time though, it lacks detailed information (except from precipitation data) in order to satisfactorily portray the spatial and temporal variability the SDVI components. The abundance of publicly available state datasets along with the employment of remote sensing information was a great opportunity to revise the way the index was applied until now and thus increase the insight gained from its application. In this respect, the application in the South Platte basin, could not follow the same assumptions of its original application, nor similar datasets. The initial application of the SDVI had certain key limitations, namely: the impacts component was static through the monthly calculation of the index; ecosystem and secondary socio-economic impacts were not considered; the datasets used were aggregated to a basin or sub-basin level which can mask local variations of drought vulnerability levels. The premises adopted for index facilitation, the datasets used and the procedures of spatial representation of each component are described below. Through the proposed approach the above-mentioned limitations were addressed. Drought impacts are now time varying and most importantly, through the incorporation of multiple datasets the impact component now accounts for ecosystem impacts (mainly natural vegetation), recreational and aquatic habitat impacts. The holistic impact representation that was developed could help depict impacts extent and magnitude in a non-fragmented way. Lastly, the inclusion of high resolution spatiotemporal datasets in the calculation of this index offers new capabilities of examining local vulnerability variations and comparing adjacent areas.

2.3.2.1 SPI

The values of SPI-6 and SPI-12 are computed on point scale (for every available meteorological station with the required data). In this context, monthly precipitation records from 24 meteorological stations (Figure 9), that have at least a 50-year record (November 1982 - October 2012), were retrieved from the NOAA's Regional Climate Centers (RCCs) Applied Climate Information System (ACIS) (http://scacis.rcc-acis.org/). The majority of the records had intermittent data gaps which were filled by multiple regression analysis using the Hydrognomon software (version 4.01) [*ITIA N.T.U.A.*, 2010] (for more information regarding the filling results, please see Appendix A).

The monthly precipitation data for the period were used as input to the SPI algorithm of the "SPEI" R package [*Beguería*, 2013]. In order to spatially visualize the calculated SPI values, several geostatistical approaches (Ordinary Kriging, Simple Kriging, Universal Kriging, Cokriging) and model types (Circular, Spherical, Tetraspherical, Pentaspherical, Exponential, Gaussian, Rational Quadratic, Hole Effect, K-Bessel, J-Bessel and Stable) have been tested in an ArcGIS environment.

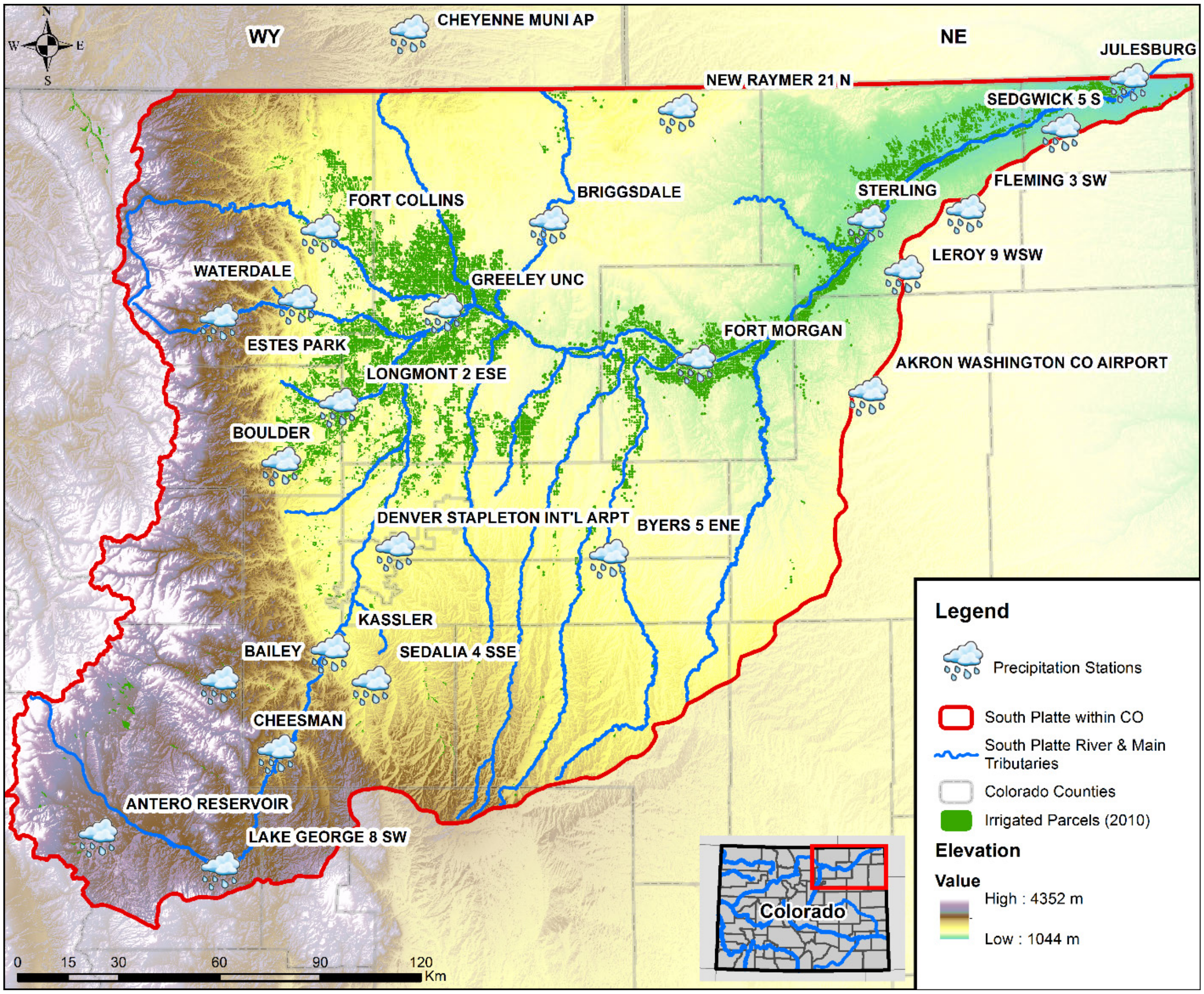


Figure 9: Precipitation stations within and in proximity to the South Platte basin

The combined application of Simple Kriging and Rational Quadratic provided the smallest Root-Mean-Square and Average Standard Errors. Hence this combination has been chosen for the 6-month and 12-month SPI visualization that is shown in Figure 10 and Figure 11 respectively. All in all, the Kriging method provides a way for minimum error distribution, while correcting potential errors within the raw data. The method works in accordance with the premise that two values that are nearby will have the same information. This is a feature presenting a significant advantage, where describing random phenomena such as precipitation patterns.

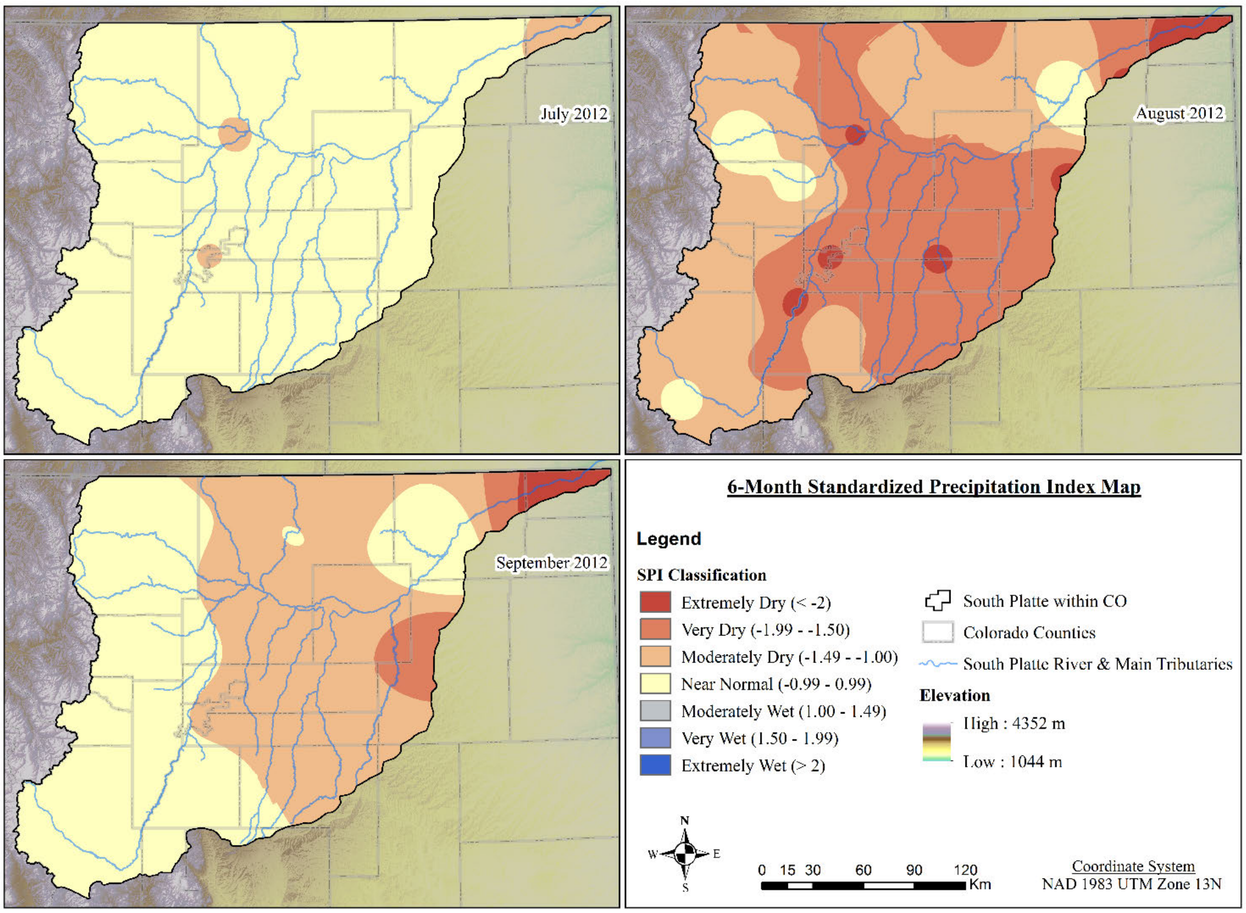


Figure 10: 6-Month SPI for JAS of 2012

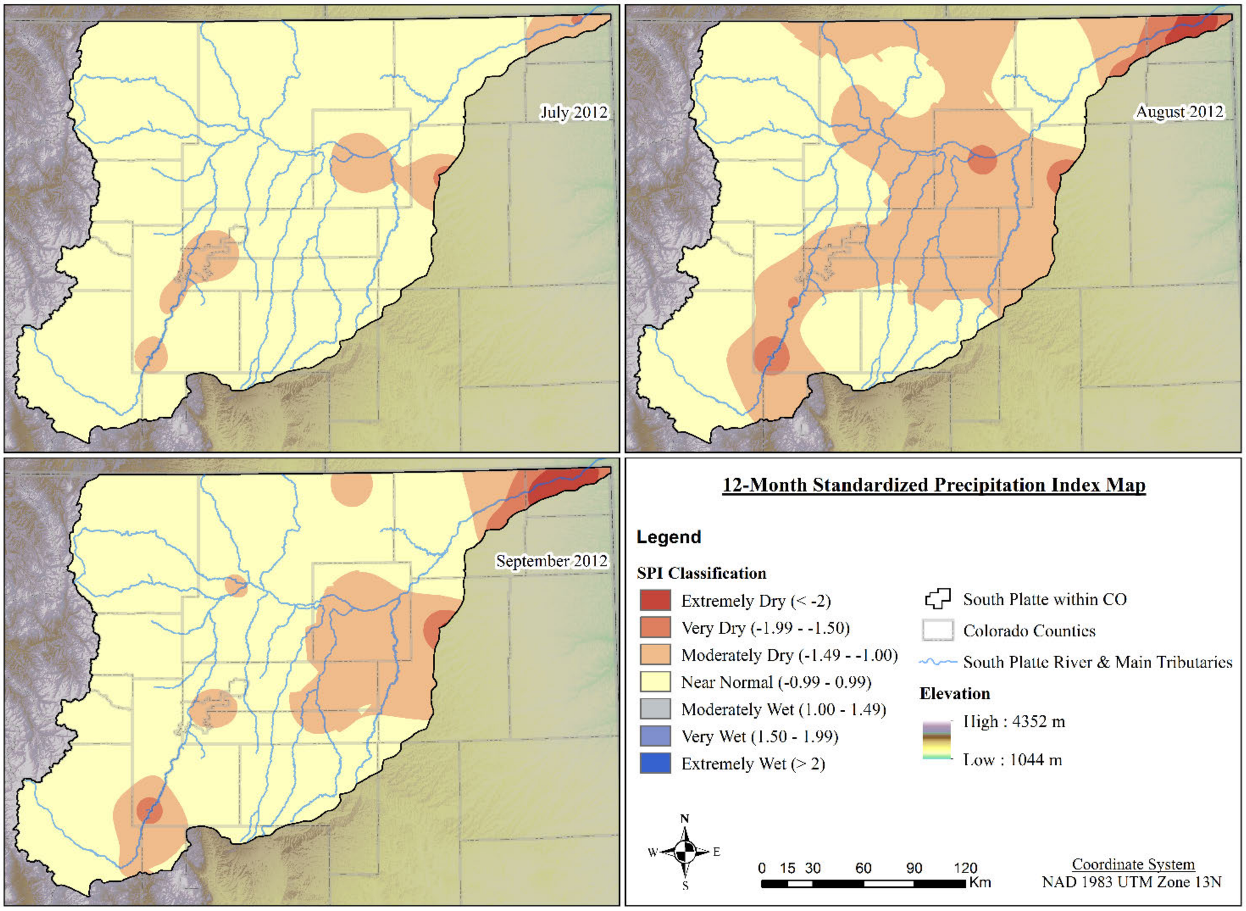


Figure 11: 12-Month SPI for JAS of 2012

2.3.2.2 Demand

Urban centers during drought events, are by default the most vulnerable locations since the impacts could be devastating. In the current effort, instead of showing by default the urban centers as highly vulnerable, it was selected to show the actual vulnerability level they experienced during the drought event. Usually, the monthly domestic water demand for the summer periods although bit higher than the rest of the seasons, is well calculated by the pertinent water authorities, and they are well prepared, not considering it a surprise. The 2011 National Land Cover Database was used in order to identify the exact areal extent of the urban areas.

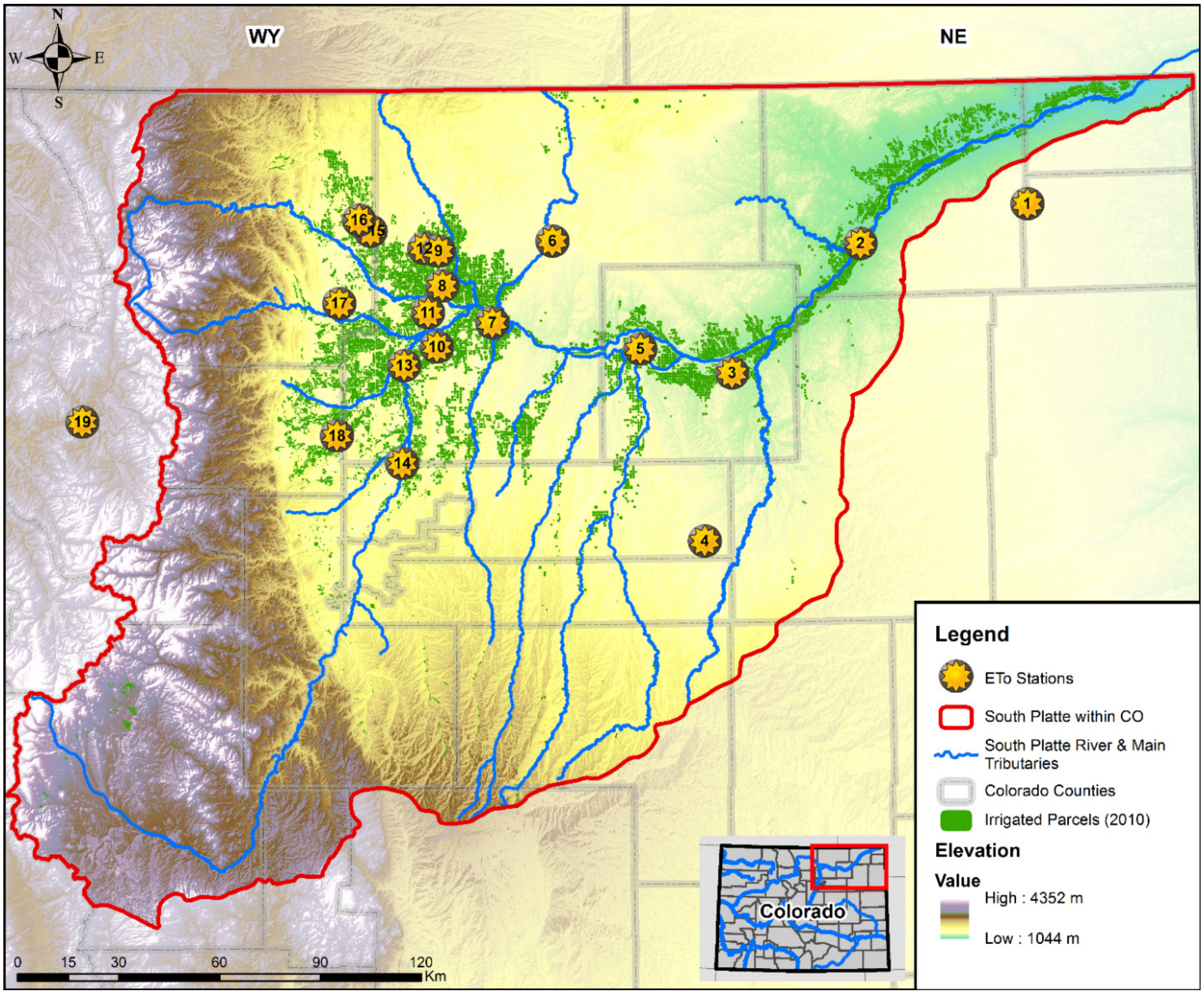


Figure 12: Evapotranspiration stations used in the analysis. (Look in the Appendix D for station names based on the number)

On the other hand, vegetation/crop water requirements vary since evapotranspiration is also dependent on climatic conditions (temperature, humidity, sunshine and wind). Thus during hot and sunny days, evapotranspiration is higher than during cooler and cloudy ones. Optimal plant growth is occurring when an equal amount of the water lost through the evapotranspiration process is applied back to the crop through rainfall and/or irrigation. In this respect, in order to estimate the scale of the demand component for basin's vegetation, reference evapotranspiration (ETo) was considered the most suitable unit of measurement. Due to the high importance of agriculture in the

South Platte basin there is a large number of ETo monitoring stations. But there are only 18 stations with long record (2000-2016). One more station (Windy Gap), with the same record length is located outside the basin and is located on the west slope of the continental divide at 2,418 m, which due to its higher elevation was preferred to approximate ETo conditions for the mountainous area of the South Platte basin. These 19 stations (Figure 12), are part of the CoAgMet (www.coagmet.colostate.edu) and the Northern Colorado Water Conservancy District (NCWCD) (www.northernwater.org) networks. Monthly ETo was selected as an indicative measurement of vegetation water demand. The average monthly ETo for the whole period was compared with July, August and September (JAS) of 2012. The excess of monthly vegetation/crop demand classification is shown in Table 1. ETo point information was transformed to spatial information based on the Thiessen polygons, a simple technique was chosen rather a geostatistical approach, due to the complexity of the phenomenon and the small number of ETo stations at the lower part of the South Platte basin.

The layer of the SDVI demand component was produced combining the calculated urban and vegetation water demand. On top of that classified spatial layer, areas with zero vulnerability (bare land, permanent snow, open water etc.) were superimposed. They were also identified from 2011 National Land Cover Database, and assigned to the low vulnerability class.

2.3.2.3 Supply

In order to create a detailed classification, based on Table 1, for the supply component a variety of information was utilized. Monthly ditch diversion records, within a 50-year period (November 1962 – October 2012), for all ditches in the South Platte basin were extracted from Hydrobase, (http://water.state.co.us/DataMaps/Pages/default.aspx) the Colorado's Division of

Water Resources database. The time series of 310 ditches (see Appendix E for the ditch identification number), which include the major ditches in the basin, were selected with the criterion to have at least 25 years of data, including information for 2012. The average diversion quantity for each month was calculated (excluding 2012) and compared with the amount diverted in 2012, in order to determine if there was a supply deficit. For the ditches that did not fulfill the above criterion, a supply deficit of 0% was assigned. The supply deficit for each ditch was converted to spatial information, utilizing the known ditch service boundaries shown in Figure 13. The cultivated crops within the ditch service polygons were assumed to have a uniform spatial allocation of water supply deficit.

For the lands outside the ditch service areas, deviation from the average monthly soil moisture in the root zone was chosen as an indicator to represent the supply component in the SDVI. Due to lack of measured data, the information from the Noah Land-Surface Model (LSM) [*Chen et al.*, 1996; *Koren et al.*, 1999] was used. The development of the Noah model was to be part of the NOAA NCEP mesoscale Eta model [*Betts et al.*, 1997; *Chen et al.*, 1997; *Ek et al.*, 2003], and it is serving as the LSM to the Weather Research and Forecasting (WRF) regional atmospheric model, the NOAA NCEP coupled Climate Forecast System (CFS) and the Global Forecast System [*Xia et al.*, 2012]. The root zone soil moisture content (measured in kg/m^2) for the pixels within the area of study were extracted from the NLDAS-2 (North American Land Data Assimilation System) Noah Land Surface Model Level-4 Monthly dataset. It has 0.125 decimal degree resolution and root depth is defined as 100cm in the forested areas, and 60 cm in the non-forested ones [*Rui and Mocko*, 2014]. The root zone soil moisture content for three months (JAS) in 2012 was compared to the 30-year (1980 – 2009) monthly averages in order to quantify the difference from the percent of historical average. The classification to the SDVI scales followed that of Table 1.

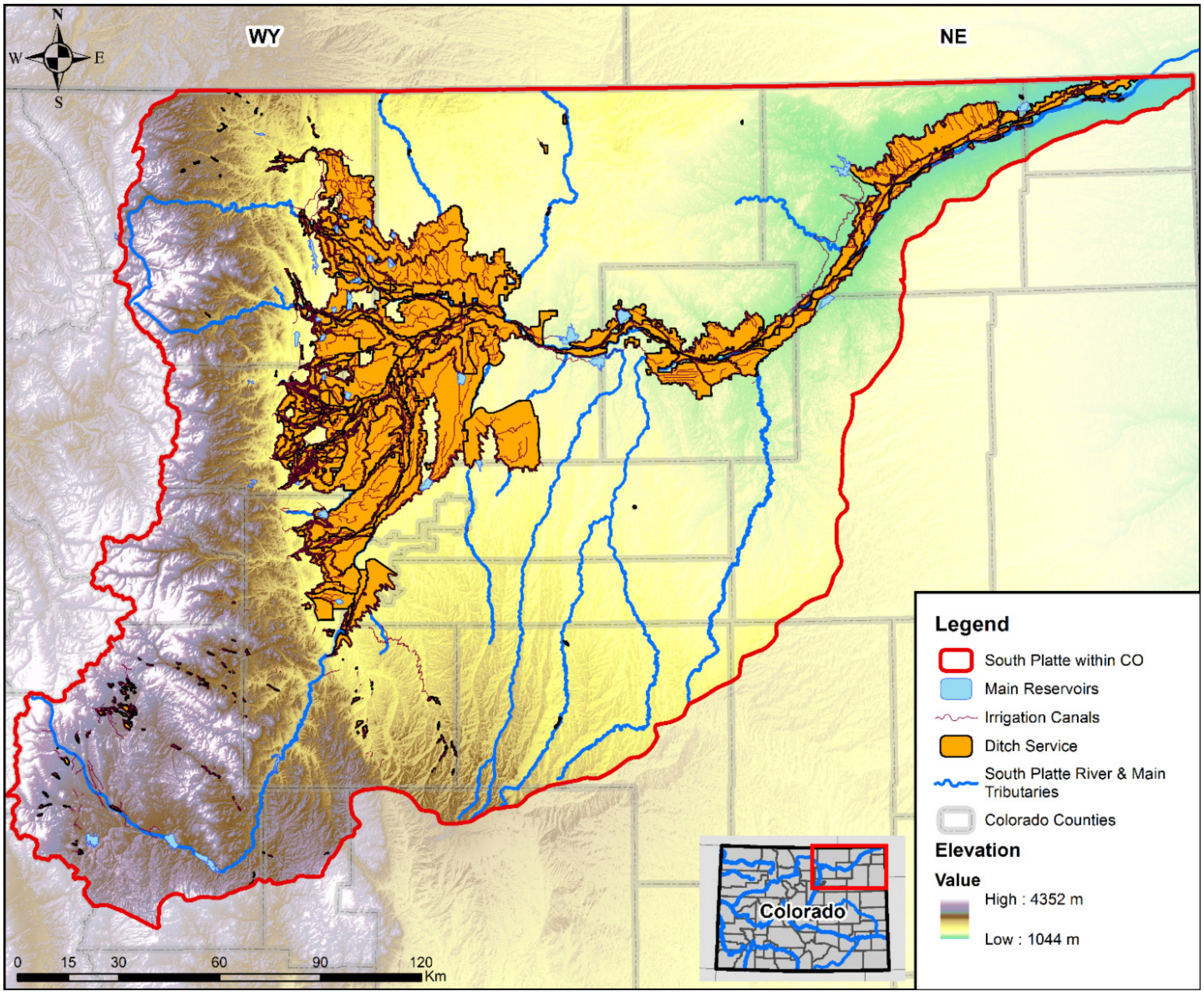


Figure 13: Ditch service areas, irrigation canals and main reservoirs within the South Platte basin

The 2011 National Land Cover Database was used in order to identify areas like bare lands, high urbanized areas, industrial areas and water features (e.g. lakes, reservoirs). The supply component for the SDVI is classified as having low vulnerability to drought, for land uses like bare lands, permanent snow and water features. In order to identify if there was domestic water supply deficit and to classify it, time series of monthly reservoir storage data from the Water and

Climate Center of the USDA Natural Resources Conservation Service National (http://www.wcc.nrcs.usda.gov/) were used. Also, the reservoir storage can be used to depict impacts on hydroelectric capacity and recreation, in cases that applies. Each month's percent average storage is a viable indication of water supply capacity by the water utilities serving these urban centers. Furthermore, in cases with no information, meeting urban water demand was considered to have low vulnerability since the minimum water amount required per inhabitant is always satisfied. A complete list of the reservoirs that were included in the analysis can be found at the Appendix F.

2.3.2.4 Infrastructure

Following the same notion that SDVI infrastructure component should portray the adequacy status of the water supply system, South Platte's systems, both for domestic and agricultural purposes, can be classified as having zero deficiency. The South Platte basin climate is semi-arid, and, the development of adequate water supply infrastructure for agriculture became early on a necessity [*Stenzel and Cech*, 2013]. Agricultural production in northern Colorado is a very important revenue sector, and Weld County is one of highest producing counties in the US. This is possible mainly due to irrigated agriculture which is highly dependent on well operated reservoirs, and canal networks to allocate water based on the prior appropriation doctrine. For the above reasons, the capacity to deliver water in the basin is classified as having 0% deficiency.

2.3.2.5 Impacts

The estimation of drought impacts is inherently difficult to perform since they have spatial and temporal variability. In previous attempts of the SDVI, drought impacts were associated with the societal cost. Either a theoretical high cost of not supplying water to urban center axiomatically depicting urban cities as most vulnerable or the monetary cost due to crop yield loss based on relevant reports were used. Although, the practice of measuring crop yield loss is a tangible way to portray the drought impacts, its disadvantage is that it cannot give any information of the temporal propagation of impacts. At the same time, environmental impacts have not been described due to lack of information at the areas of previous applications. The fragmentation of usually circumstantial impacts reports does not contribute towards their classification into the categories of interest (e.g. social or economic). Instead, the majority of the reports focus on the impacts posed on agriculture leaving great information gaps on the other categories (socio-economic) [*Gregorič*, 2012; *Grigg*, 2014]. This area of interest did not escape such traits, limiting our understanding about the impacts.

Responding to this weakness, and with the motive to also incorporate spatial and temporal effects of drought events, remote sensing information was employed to capture both agricultural and environmental impacts. The Normalized Difference Vegetation Index (NDVI) [*Rouse*, 1974] is a dimensionless transformation of spectral reflectance, that allows one to measure, visualize and evaluate healthy and abundance vegetation. It is expressed (Eq. 4) as the difference of the near-infrared and visible region of the electromagnetic spectrum, divided by their sum. Healthy vegetation compared to stressed or diseased vegetation have different spectrum signatures [*Knipling*, 1970]. Healthy vegetation is absorbing much of the blue and red parts of the electromagnetic spectrum and is reflecting most of the green, while reflecting a large portion of near-infrared [*Curran*, 1980; *Govender et al.*, 2007]. The NDVI has been used in several drought related studies [*Hurcom and Harrison*, 1998; *Brown et al.*, 2008; *Karnieli et al.*, 2010; *Xu et al.*, 2011; *Tadesse et al.*, 2014].

$$\mathrm{NDVI} = \frac{\mathrm{Near\ IR} - \mathrm{VIS}}{\mathrm{Near\ IR} + \mathrm{VIS}} \tag{4}$$

The Moderate Resolution Imaging Spectroradiometer (MODIS) instrument is carried by NASA's Earth Observing System (EOS) satellites of Terra and Aqua, and one the mission's scope is to produce global vegetation indices. The NDVI is reported globally in a 16-day interval with a resolution of 250m from each satellite (products: MOD13Q1 and MYD13Q1). The time-lag (8-day) of Terra and Aqua satellites give the opportunity of combining the data in order to increase temporal resolution to an 8-day time step. Combined MODIS NDVI data, smoothed and gap-filled, for the Conterminous US (CONUS) for the period 2000-2015 in a temporal resolution and spatial of 8-days and 250m respectively [*Spruce et al.*, 2016], were employed to create the vegetation impact. The layer is portraying the NDVI for July, August and September of 2012, as percent of average monthly values for are deviating from long term average values. The NDVI dataset contains 16 files (one for each year) in netCDF-4 format for the whole CONUS. Due to the size of each file exceeding 4 GB, their handling was performed in the Linux Operating System. The netCDF Operator (NCO) toolkit [*Zender*, 2016] was used to manipulate the files in order to obtain a geographical subset that included the area of interest. The cropped files were imported into **R** to extract the gridded information of the dates that fall within JAS, for the period 2000-2015, to a raster format. Raster layers were calculated based on monthly averages, for each year, and then, historical monthly averages for the whole period (excluding 2012). The 2012 NDVI monthly average rasters produced were transformed to express the percentage from the long-term average and were classified according to Table 1.

On the other hand, the delineation of impacts in urbanized areas was only possible through available reports, retrieved from the Drought Impact Reporter of the National Drought Mitigation

Center (droughtreporter.unl.edu). As mentioned in the supply section, reservoir storage could be used to represent impacts on hydroelectric capacity generation in areas that apply (e.g. Colorado River Basin), and for recreation purposes. Under this notion, the area occupied by the reservoirs (Figure 46) is classified according to Table 1, using percent of monthly average storage at the 1st of each month, with a reference period 1981-2010. The vegetation, urban and reservoirs' supplement impacts information can be combined to generate the monthly SDVI impact layer. The satellite-derived information combined with hard data (reservoir storages) and soft data (reports) is an integrated way to depict vegetation (including agricultural crops), hydroelectric (which does not apply to the South Platte study area) and recreation/habitat drought impacts. The classification of different impacts allows one to visualize drought effects temporally and spatially, and thus suggests a way to overcome limitations pressed by information fragmentation and offers a way of synthesis.

2.4 Results and Discussion

The SDVI was conceived as an integrated approach for assessing drought vulnerability and thus the vulnerability levels portrayed represent the major components contributing to the overall vulnerability. The months of July, August and September were selected for assessing the vulnerability of the South Platte basin to the drought event of 2012. The main reason for selecting this period was to include in the assessment the vegetation's growth season. Firstly, in this section, a summary and discussion of the individual calculated SDVI components is provided. Afterwards, the SDVI results calculated for each month, based on the aforementioned framework, are presented and summarized.

2.4.1 SDVI Components' Estimation

According to the previously described methodology (Section 2.3.2), after the collection of the pertinent data for the SDVI components, the next step was their calculation and visualization in an ArcGIS environment, which followed their transformation into their respective scaled vulnerability values. An example of this transformation presented here, is the calculated SPI-6 and SPI-12 (Figure 10 and Figure 11) and their visualization in terms of vulnerability levels is shown in Figure 14 and Figure 15 respectively. From the produced maps for the classified SPI-6 (cSPI-6) and the classified SPI-12 (cSPI-12) presented, a few points may be surfaced. Firstly, the vulnerability due to precipitation deficit for the most part of the examined region is between 0 and -1.46 – portraying high scale vulnerability. During August of 2012, 49.36% of the basin is characterized as extremely vulnerable for the cSPI-6 (Table 3), but it recovers in the next month apart from the area northeast of the city of Sterling and southeast of the city of Fort Morgan. The shift can be explained since precipitation in August was significantly below normal levels. Drought conditions for SPI-12 were below normal (SPI < 0) in all precipitation stations and that is why approximately 95% of the basin (Table 4) is depicted as highly vulnerable. In August and September of 2012, according to cSPI-12 the area northeast of Sterling was classified in the extremely vulnerable level. For both Julesburg and Sedgwick stations, SPI-12 was less than -2 decreasing more in September. Overall, the northeast tip of the South Platte is the most vulnerable area in the basin. The SPI-6 and SPI-12 are representing the hazard component in the function.

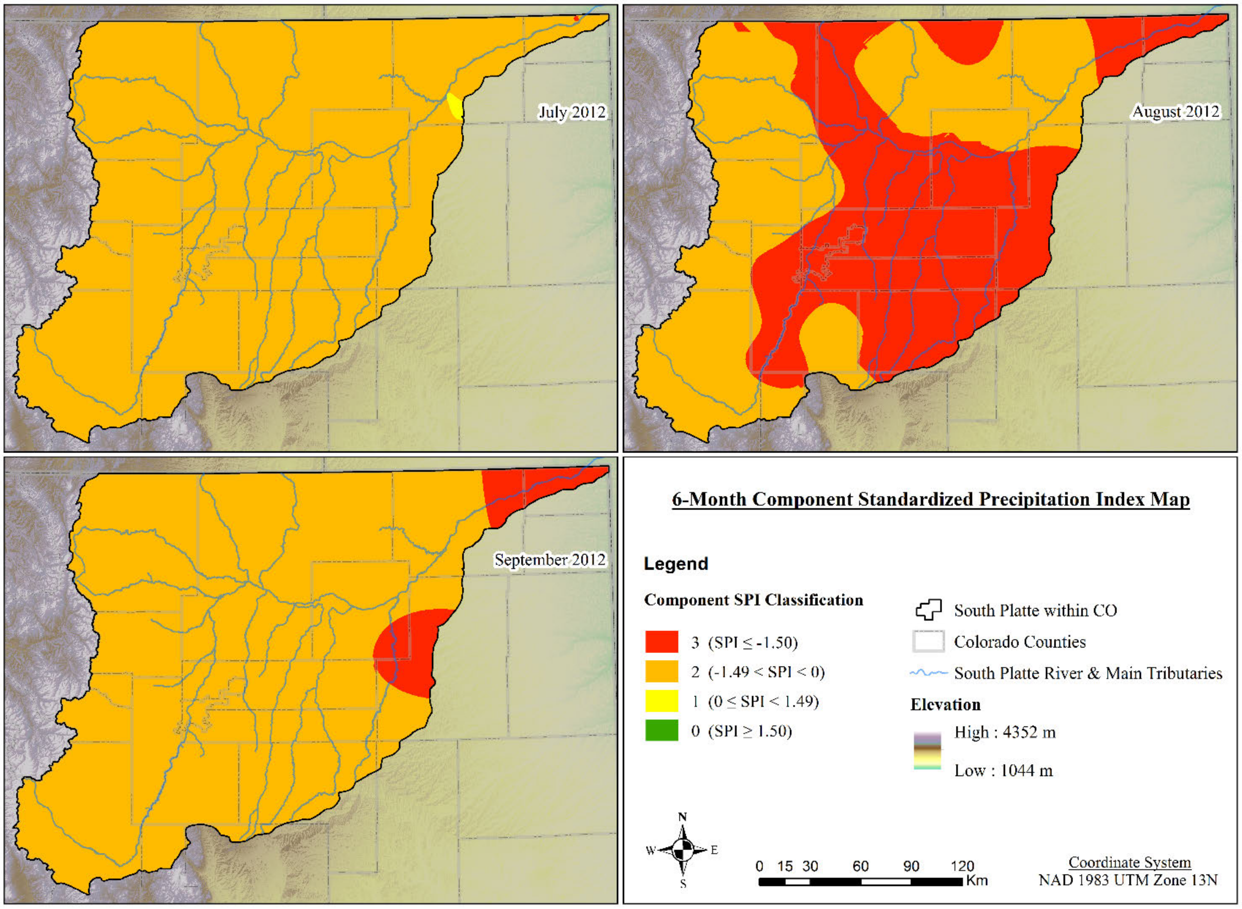


Figure 14: 6-Month component SPI for JAS of 2012

Table 3: Area and percent area of vulnerability levels for the 6-month SPI in the South Platte basin

Component Class	Jul-12		Aug-12		Sep-12	
	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Less Vulnerable (0)	-	0.00%	-	0.00%	-	0.00%
Vulnerable (1)	103.1	0.21%	3.8	0.01%	-	0.00%
Highly Vulnerable (2)	48,825.2	99.77%	24,779.1	50.64%	46,016.6	94.04%
Extremely Vulnerable (3)	7.3	0.01%	24,152.7	49.36%	2,918.9	5.96%
Total Area	48,935.6					

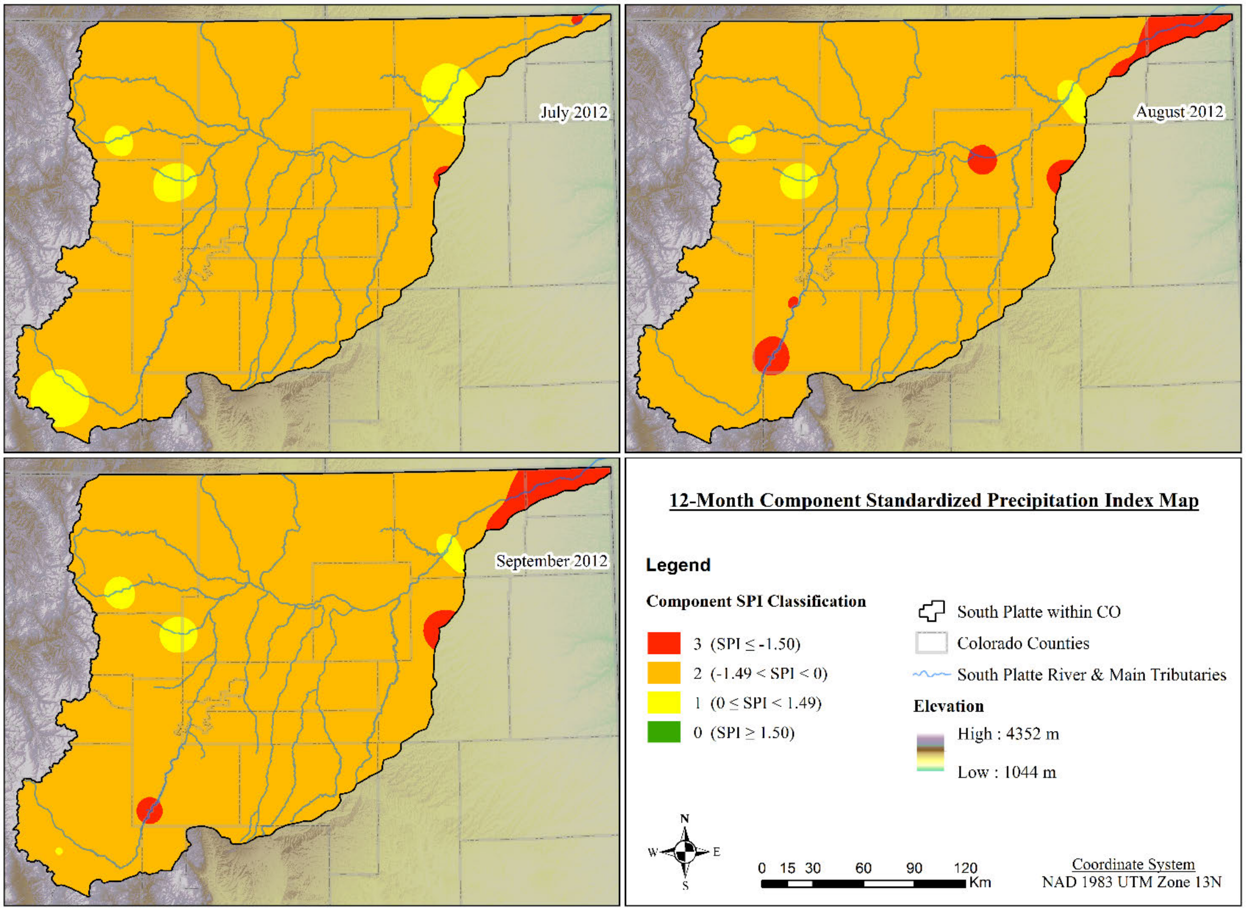


Figure 15: 12-Month component SPI for JAS of 2012 in the South Platte basin

Table 4: Area and percent area of vulnerability levels for the 12-month SPI in the South Platte basin

Component Class	Jul-12		Aug-12		Sep-12	
	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Less Vulnerable (0)	-	0.00%	-	0.00%	-	0.00%
Vulnerable (1)	2,499.5	5.11%	911.7	1.86%	918.5	1.88%
Highly Vulnerable (2)	46,354.0	94.72%	46,328.4	94.67%	46,556.5	95.14%
Extremely Vulnerable (3)	82.1	0.17%	1,695.5	3.46%	1,460.5	2.98%
Total Area	48,935.6					

The infrastructure component is not presented since the South Platte basin is highly developed and with a well-maintained water infrastructure network, thus classified in the less vulnerable condition for the whole period of study.

The demand component for the South Platte is presented in Figure 16. During July of 2012 the vegetation water requirements in the mountainous portion of the basin are classified as average (less vulnerable) and during the following two months are classified as vulnerable since water requirements are not exceeding 15% of average. The "CSU-ARDEC" ETo station had throughout the study period measured conditions of less ETo than the historical average, and that is why the area of influence (Thiessen polygon) including the northeast part of Larimer County and northwest part of Weld County are depicted with green color. The other less vulnerable areas (green color) on the maps of Figure 16 are either urban centers, like the Denver metropolitan area at the lower left side of the basin, or reservoirs and land uses such as bare land. The defining aspect of the 2012 drought was abnormally high temperatures and evapotranspiration rates. During July and August of 2012 the plain portion of the basin is at the most vulnerable state since 41.88% and 51.66% of it was classified in the higher vulnerable conditions for the demand component respectively. The phenomenon dissipated during September and only 4.08% of the basin was depicted with vegetation water requirements more than 15% of the historical average. The area and the percent area of each vulnerability component level is spatially occupying during the three months of study are presented on Table 5.

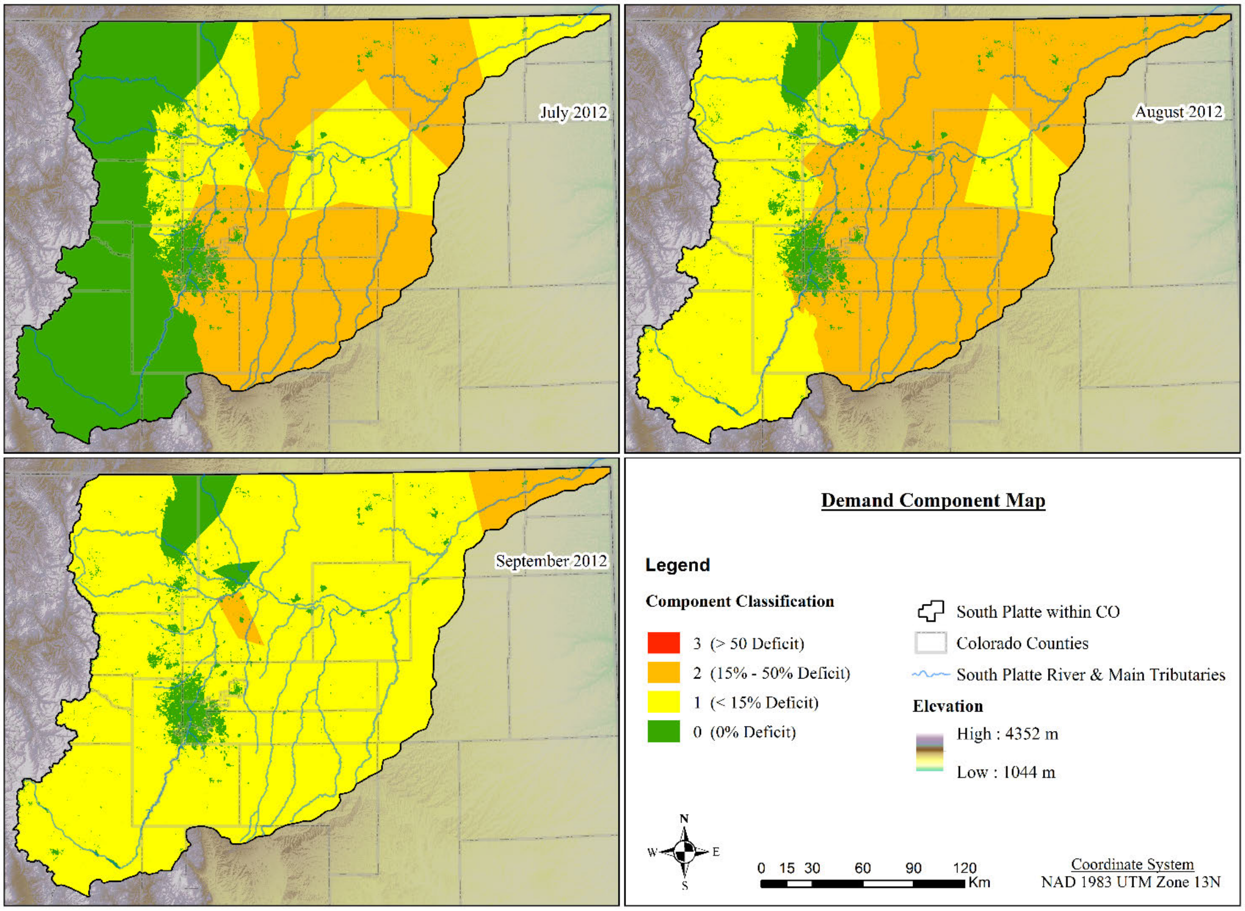


Figure 16: Demand component for JAS of 2012 in the South Platte basin

Table 5: Area and percent area of vulnerability levels for the demand component in the South Platte basin

Component Class	Jul-12		Aug-12		Sep-12	
	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Less Vulnerable (0)	17,669.5	36.11%	3,262.5	6.67%	3,460.8	7.07%
Vulnerable (1)	10,770.3	22.01%	20,393.4	41.67%	43,478.6	88.85%
Highly Vulnerable (2)	20,495.7	41.88%	25,279.6	51.66%	1,996.2	4.08%
Extremely Vulnerable (3)	-	0.00%	-	0.00%	-	0.00%
Total Area	48,935.6					

The SDVI Supply component was calculated for the area of interest, based on information regarding monthly water diversions in irrigation canals, the storage level of the reservoirs used for domestic demand as an indication of capacity to meet the urban water demand and lastly the root zone soil moisture content from the operational Noah LSM model. Urban centers were classified as less vulnerable since storage percent anomaly for the reservoirs with domestic purpose across the basin did not exceed 50% of storage historic average for the whole period. The reservoirs serving Denver and Aurora like Antero and Elevenmile Canyon had above the average storage and the Spinney Mountain almost the average storage. At the same time reservoirs with mainly agricultural water use delivery were experiencing more than 50% lower storage than their monthly historic averages. This could be associated with canal service areas that their delivery was classified in the extremely vulnerable level. Examples of this are the Riverside Reservoir and the North Sterling Reservoir and the water diversions at their respective canals service areas. The deficits regarding the water diversions to the irrigation canals was intensified during September of 2012. Furthermore, about half of the basin's area was classified as highly vulnerable in August and September since the Noah Land-Surface Model estimated the anomaly of the soil moisture content at the root zone for lands outside the ditch service areas to be less than 50% of the historic average. The mountainous area of the South Platte basin according to the data from the Noah LSM indicated that there were vulnerable to drought and only pockets of average conditions were estimated by the model.

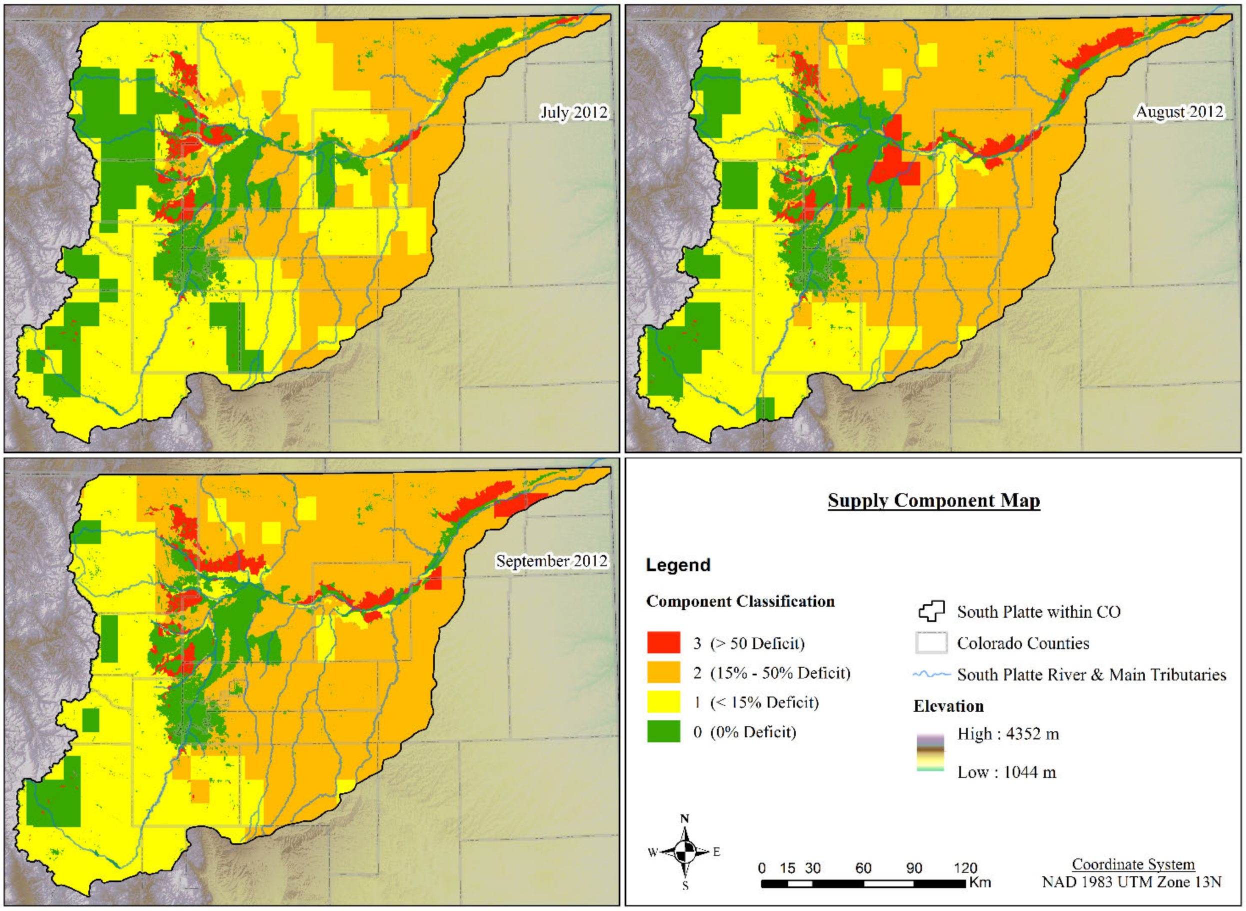


Figure 17: Supply component for JAS of 2012 in the South Platte basin

Table 6: Area and percent area of vulnerability levels for the supply component in the South Platte basin

Component Class	Jul-12		Aug-12		Sep-12	
	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Less Vulnerable (0)	10,736.0	21.94%	8,592.4	17.56%	6,422.1	13.12%
Vulnerable (1)	22,224.9	45.42%	14,431.2	29.49%	15,280.9	31.23%
Highly Vulnerable (2)	14,855.6	30.36%	23,819.5	48.68%	25,107.7	51.31%
Extremely Vulnerable (3)	1,119.1	2.29%	2,092.4	4.28%	2,124.8	4.34%
Total Area	48,935.6					

The last SDVI component is the occurred impacts on a monthly time step. As described in the methodology, satellite derived datasets, reservoirs' storage information and reports from the Drought Impact Reporter database regarding urban impacts were utilized in order to represent a composite spatial layer of multifaceted impacts. Overall, cities on the Front Range did not have any significant impact thus were represented as less vulnerable. From the maps illustrated in Figure 18, it is obvious that the vegetation outside ditch service areas is more vulnerable to drought and fall within the highly vulnerable component class. In the mountainous part of the South Platte basin the vegetation is experiencing some stress but there are also a few areas classified as less vulnerable. The anomaly recorded is less than 15% of average greenness. The only exception in this is the area at the north within the Larimer County, that has highly and extremely vulnerable vegetation. These severe impacts are not credited to drought conditions, but to the Hewlett Gulch Wildfire that burned 31 km^2 on May of 2012 and the High Park fire, that occurred in June of the same year, and burned 350 km^2 [*Writer et al.*, 2014]. This exception stresses the fact that additional information for each area of application need to be taken into account when interpreting the impact's spatial layer in order to identify eluding causes for depicted irregularities.

The comparison of the index produced results with the existing condition show a very good relation. The other spots of extreme vulnerability that are along the South Platte River are reservoirs whose storage was significantly lower than their historic monthly average. Representative are the cases of the Empire, Riverside and Jackson Lake reservoirs that were completely dried out by 1st of August 2012 and North Sterling reservoir by 1st of September. Consequently, affecting not only the irrigation ditches they are serving, but also having recreation and environmental/habitational impacts. One report in the Drought Impact Reporter database is mentioning that on July 23rd action was taken to salvage fish population in the Barr Lake near

Denver, which on the maps of Figure 18 is shown in the extreme vulnerable class. Furthermore, boating at the Prewitt reservoir (July) and Horsetooth (August) reservoirs are affected, which again are portrayed by the index component as vulnerable spots.

A significant part of the irrigated agriculture is under stress with the anomaly on the vegetation greenness in these areas to be classified as vulnerable (<15%) depicting the deficiency of meeting crop water requirements, which the Drought Impact Reporter database is supporting since farmers were asking permission from senior water rights to be allowed to pump groundwater from the alluvial aquifer. At the same time the crops that are within normal conditions can be attributed to their priority on the available water since the land is associated with senior water rights. In Morgan County, the irrigated crops are experiencing of vegetation health since the NDVI percent of average is decreasing steadily from July throughout August.

On the other hand, the rest of the vegetation at the plains is depicted as steadily highly vulnerable, something expected since most of this area is classified as non-irrigated grass/pasture. This high stress observed on grass/pasture lands is in concordance with reports from the media retrieved from the Drought Impact Reporter database. There are several entries mentioning in the area there was not enough grass to feed cattle and thus forced of selling livestock, or in other cases forced to feed alfalfa and hay, which indirectly are displayed in the impacts component maps since grass and pasture lands are severely impacted. All in all, the reports retrieved from the Drought Impact Reporter database are supporting the results from the selected methodology to portray the impacts in a holistic way, but during their interpretation other sources of information (land uses, wildfires, etc.) should be taken into account.

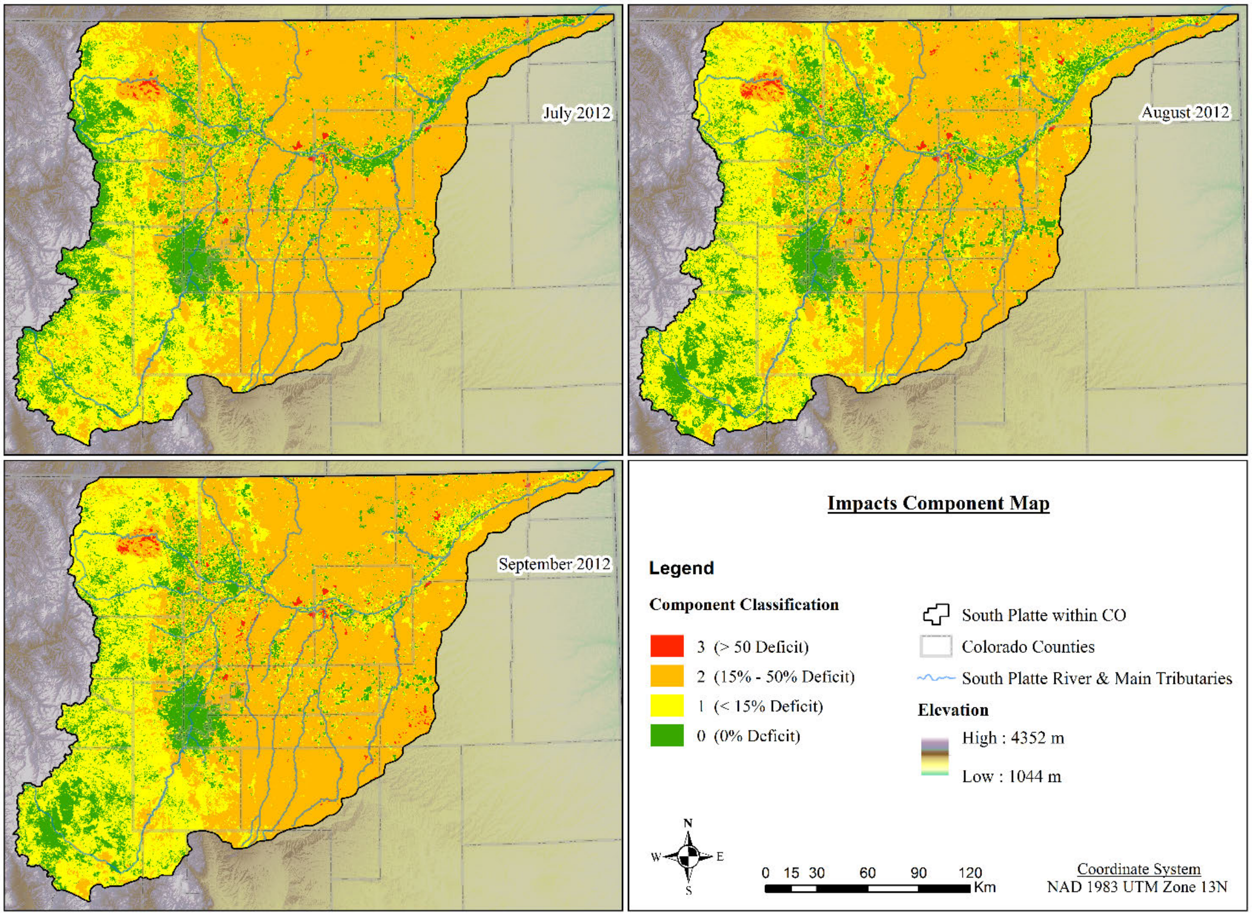


Figure 18: Impacts component for JAS of 2012 in the South Platte basin

Table 7: Area and percent area of vulnerability levels for the impacts component in the South Platte basin

Component Class	Jul-12		Aug-12		Sep-12	
	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Less Vulnerable (0)	7,418.1	15.16%	7,950.5	16.25%	6,326.2	12.93%
Vulnerable (1)	13,643.1	27.88%	16,612.0	33.95%	15,556.1	31.79%
Highly Vulnerable (2)	27,650.0	56.50%	24,011.7	49.07%	26,689.7	54.54%
Extremely Vulnerable (3)	224.4	0.46%	361.3	0.74%	363.6	0.74%
Total Area	48,935.6					

2.4.2 Drought Vulnerability Estimation

Figure 19 portrays the monthly vulnerability magnitude and extent of the South Platte basin, but also informs about the spatiotemporal propagation of the index. The SDVI results may be directly connected to the SPI results (as the SPI values were used in the SDVI estimation representing the hazard component), since the most vulnerable area, northeast of Sterling, displayed a significant precipitation stress, in both SPI components, compared to the least vulnerable ones. Overall, the total extent of the extremely vulnerable class does not surpass 3% of the total area.

The urban areas are characterized as the least vulnerable part of the basin along with parts of the basin with higher elevation. For urban areas this low level of vulnerability during the study period is attributed to the high percent average of reservoirs' storage and the very few reports retrieved from the Drought Impact Reporter database. It is worth mentioning that the urban areas were expected to be classified in one of the least vulnerability classes compared to the other parts of the basin, since the political priority given to these areas for mitigating the effects of drought is the highest due to potentially catastrophic consequences. The mountainous areas display less vulnerability to drought which is accredited to the lower vulnerable demand and supply components. The aforementioned impacts of the two wildfires in the impact component estimation, were somewhat masked by the rest of the SDVI components and also because in the historical NDVI average were included post-fire years, resulting in a classification as one scale more vulnerable compared to adjacent areas. The higher vulnerability of post-fire areas is in accordance with recent findings that during drought years following the wildfires the vulnerability is higher in terms of forest recovery and favored species [*Harvey et al.*, 2016]. Non-irrigated grass/pasture lands, which constitute the vegetation with the greatest extent in the basin, were classified among the most vulnerable parts of the basin for the 2012 drought. This SDVI finding is in accordance and backed up by the numerous reports available in the Drought Impact Reporter database.

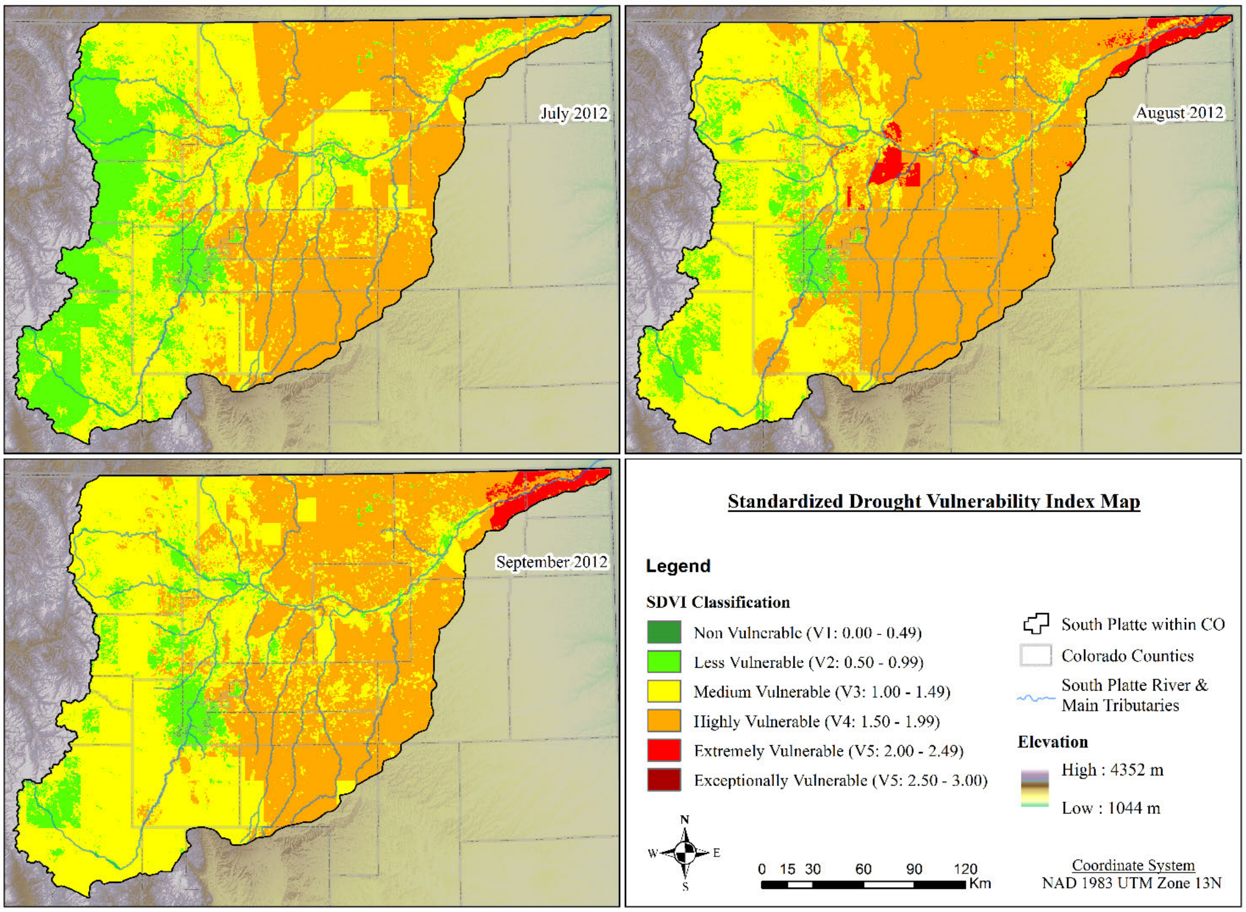


Figure 19: SDVI results for the South Platte basin

Table 8: Area and percent area of SDVI vulnerability levels in the South Platte basin

Component Class	Jul-12		Aug-12		Sep-12	
	(km^2)	(%)	(km^2)	(%)	(km^2)	(%)
Non Vulnerable (1)	-	0.00%	0.4	0.00%	-	0.00%
Less Vulnerable (2)	8,516.7	17.40%	3,095.5	6.33%	3,086.4	6.31%
Medium Vulnerable (3)	20,173.3	41.22%	18,468.4	37.74%	23,462.5	47.95%
Highly Vulnerable (4)	20,245.5	41.37%	25,908.3	52.94%	21,438.2	43.81%
Extremely Vulnerable (5)	-	0.00%	1,463.0	2.99%	948.6	1.94%
Exceptionally Vulnerable (6)	-	0.00%	-	0.00%	-	0.00%
Total Area	48,935.6					

The basin is depicted as more vulnerable during August of 2012, which slightly dissipates the next month. From the components presented and discussed in section 2.4.1, this slight intensification of about 5,000 km^2 shifting to the medium vulnerable class (Table 8) is a result of the combined result of the demand, supply, SPI-6 and impacts components. However, the straight-line borders of vulnerability classes, a non-natural effect, which is visible in almost all three months, is product of the coarser datasets (Demand and Supply) that were used for the calculation of the index. This denotes the need for incorporating and testing alternative datasets, as they become available, with finer resolution. Although, recreation and environmental impacts from reservoirs were well portrayed in the pertinent component maps, the SDVI calculations display lower vulnerability status of the reservoirs. Especially in the lower section of the South Platte where some had zero storage. This is because the minimum demand and supply component for socio-environmental uses, was not considered for the reservoirs. Data on thresholds need to be incorporated for a more accurate display of vulnerability levels, but such information is not easily accessible.

2.5 Conclusions

In drought and its multifaceted impacts lurks a complexity that makes the task of their spatiotemporal assessment challenging. Indices have attempted to describe the duration, magnitude and spatial extent of droughts. Despite the fact that physical and anthropogenic systems are interwoven, most of the time the assessments are following a fragmented approach, and fail to link components (structural and socio-economic) that could intensify vulnerability to drought events. Vulnerability to drought is characterizing the susceptibility of system's components. Vulnerability includes both hazards and impacts and it measures the ability to meet demands at a specific time step. Understanding the weaknesses and interconnections of physical and social systems can better inform drought management strategies. Sine there is inherent difficulty and complexity, the employment of indices that would incorporate different components of vulnerability is a viable way forward to developing implementable drought management tools. SDVI had been conceived and constructed in such a way in order to include aspects of the physical, structural and socio-economic components. Its original development and application though, has been assessed in a region of limited data availability.

The South Platte basin faces great challenges in water resources management that are further intensified due to extreme events. At the same time, it is an area of data abundance compared to other parts of the world and thus suitable for evolving the SDVI to have less qualitative inputs and thus more accurate approximate the system's vulnerability. Furthermore, the approach adopted increases the spatial resolution of the results and thus the index serves as a way of providing a relative measure such that finer scale areas can be compared. In this way, the SDVI evolves to an implementable phase and its results could have more transparency and been better assessed. It should be noted that the value of the index cannot be validated at the ground as in most composite indicators. Instead, the individual components can be validated and compared with different drought events, thus providing information to stakeholders and planners of priority actions at different scales (farm to basin).

The datasets that helped overcome previous limitations of the index include integration of remote sensing and measured data, and, some soft data in cases there was not enough quantitative information. Apart from the SPI-6 and SPI-12 components, the calculation of the rest of the components was reinvented in order to take advantage of the abundant publicly available information. The Demand SDVI component used percent ETo from historical average for the vegetated areas in order to portray plausible increased water requirements during drought events, while domestic and industrial water demands considered as least vulnerable due to their nature. The supply component is calculated based on ditch irrigation diversions for irrigated agriculture. For the urban areas the percent of reservoirs' storage was an indication of domestic water supply capacity. The South Platte basin is a highly developed with water supply networks well maintained, thus the infrastructure component did not have spatial variability and was classified in the less vulnerable level. Lastly, the impact SDVI component was composed by different datasets in an attempt to capture some of the multifaceted drought impacts (agricultural/vegetation, domestic, recreational and environmental). The percent average of NDVI portrayed the impacts in vegetated areas. The reports from the NDMC's Drought Impact Reporter were utilized to classify urban water deficit impacts, while the reservoirs' percent storage compared to the monthly average were used for representing recreational and environmental impacts.

Overall, the SDVI results can be correlated with the SPI components since they are constituting one third of the index's value. That being said, the incorporation of the other index components results in delineating vulnerability levels based on societal, physical and structural factors. The SDVI values produced for the South Platte basin seems to offer a deeper understanding of vulnerability of the different system's components. According to the analysis preformed the urban areas are classified as least vulnerable, along with the forested land uses. The irrigated agriculture is showing less vulnerability than vegetation on the plains located outside ditch service areas. This is attributed to SDVI's supply component since the capacity to meet crop water requirements within the ditch service areas is potentially greater than outside of it. At the same time, the impacts measured are greater in the grass and pasture lands than in crops.

Despite the evolution of the index calculation some limitations still exist. Water supply from groundwater was not possible to be incorporated since data were not available. Uncertainty in the input values for the index calculation is also one of the uncertainty sources affecting the results. The components with the least certainty are the demand and supply. Evapotranspiration measurements tend to be spatially sparse with limited coverage and with usually short time records. A solution to that could be the incorporation of ET data products from satellite observations like the MODIS Global Evapotranspiration Project or from the Thermal InfraRed Sensor onboard the Landsat 8. The approach adopted in the present effort regarding the supply component, could be further enhanced with more accurate representation of the strict water allocation rules in the South Platte due to the prior appropriation doctrine. The combination of allocation models will help reduce the assumption of equal spatial distribution of supply deficit within each ditch service area and thus have a more detailed effect on finer scale. Alternatively, a more efficient way would be the incorporation of soil moisture products from satellites as they become available, e.g. the Soil Moisture Active Passive mission, and thus eliminating the need of intensive modeling of water supply and tracking changes through time in allocation priorities. Relying more on satellite driven data, will lead to a more operational version of the index with the ability to inform about drought vulnerability conditions in near-real time. In addition, the incorporation of demand and supply components for the reservoirs would result to more precise assignment of overall drought vulnerability levels regarding ecological and societal aspects. Detailed representation of infrastructure status is needed to be incorporated for the index calculation since it is one of the main factors that can affect water storage and delivery and thus contribute to system's vulnerability to drought.

The identification of system's vulnerability in an integrated way is crucial to reveal its different contributing underline causes, giving a better understanding of the system's complexities to water planners and managers. Thus, the vulnerability categorization of the system's components based on multiple drought events could lead into triggering targeted actions that could result to a more integrated approach in drought management linking demand, supply and impact focused measures, and at the same time resulting on improving water security.

References

Adger, W. N. (2006), Vulnerability, *Global Environmental Change*, *16*(3), 268–281, doi:10.1016/j.gloenvcha.2006.02.006.

Barraqué, B., C. A. Karavitis, and P. Katsiardi (2008), The Range of Existing Circumstances in the WaterStrategyMan Case Studies, in *Coping with Water Deficiency*, edited by P. Koundouri, pp. 45–112, Springer Netherlands.

Beguería, S. (2013), *SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index,* R package version 1.6.

Belayneh, A., and J. Adamowski (2012), Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression, *Applied Computational Intelligence and Soft Computing*, *2012*, e794061, doi:10.1155/2012/794061.

Betts, A. K., F. Chen, K. E. Mitchell, and Z. I. Janjić (1997), Assessment of the Land Surface and Boundary Layer Models in Two Operational Versions of the NCEP Eta Model Using FIFE Data, *Mon. Wea. Rev.*, *125*(11), 2896–2916, doi:10.1175/1520-0493(1997)125<2896:AOTLSA>2.0.CO;2.

Blauhut, V., L. Gudmundsson, and K. Stahl (2015), Towards pan-European Drought Risk Maps: Quantifying the Link between Drought Indices and Reported Drought Impacts, *Environ. Res. Lett.*, *10*(1), 014008, doi:10.1088/1748-9326/10/1/014008.

Blauhut, V., K. Stahl, J. H. Stagge, L. M. Tallaksen, L. De Stefano, and J. Vogt (2016), Estimating Drought Risk across Europe from Reported Drought Impacts, Drought Indices, and Vulnerability Factors, *Hydrol. Earth Syst. Sci.*, *20*(7), 2779–2800, doi:10.5194/hess-20-2779-2016.

Bohle, H.-G. (2001), Vulnerability and criticality: Perspectives from Social Geography, *International Human Dimensions Programme on Global Environmental Change (IHDP) Newsletter Update*, *02/2001*.

Bordi, I., K. Fraedrich, M. Petitta, and A. Sutera (2006), Large-Scale Assessment of Drought Variability Based on NCEP/NCAR and ERA-40 Re-Analyses, *Water Resour Manage*, *20*(6), 899–915, doi:10.1007/s11269-005-9013-z.

Brown, J. F., B. D. Wardlow, T. Tadesse, M. J. Hayes, and B. C. Reed (2008), The Vegetation Drought Response Index (VegDRI): A New Integrated Approach for Monitoring Drought Stress in Vegetation, *GIScience & Remote Sensing*, *45*(1), 16–46, doi:10.2747/1548-1603.45.1.16.

Byun, H.-R., and D. A. Wilhite (1999), Objective Quantification of Drought Severity and Duration, *J. Climate*, *12*(9), 2747–2756, doi:10.1175/1520-0442(1999)012<2747:OQODSA>2.0.CO;2.

Cancelliere, A., G. Di Mauro, B. Bonaccorso, and G. Rossi (2005), Stochastic Forecasting of Standardized Precipitation Index, in *Proceedings of XXXI IAHR Congress Water Engineering for the future: Choice and Challenges*, pp. 3252–3260, Seoul, Korea.

Cancelliere, A., G. D. Mauro, B. Bonaccorso, and G. Rossi (2007), Drought Forecasting using the Standardized Precipitation Index, *Water Resources Management*, *21*(5), 801–819, doi:10.1007/s11269-006-9062-y.

Changnon, S. A., and W. E. Easterling (1989), Measuring Drought Impacts: The Illinois Case, *JAWRA Journal of the American Water Resources Association*, *25*(1), 27–42, doi:10.1111/j.1752-1688.1989.tb05663.x.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996), Modeling of Land Surface Evaporation by Four Schemes and Comparison with FIFE observations, *J. Geophys. Res.*, *101*(D3), 7251–7268, doi:10.1029/95JD02165.

Chen, T. H. et al. (1997), Cabauw Experimental Results from the Project for Intercomparison of Land-Surface Parameterization Schemes, *J. Climate*, *10*(6), 1194–1215, doi:10.1175/1520-0442(1997)010<1194:CERFTP>2.0.CO;2.

Colorado Decision Support Systems (2016), GIS Data for South Platte, Available from: http://cdss.state.co.us/GIS/Pages/Division1SouthPlatte.aspx (Accessed 5 August 2016)

Colorado Department of Local Affairs (2016), Population Data, Available from: https://demography.dola.colorado.gov/population/

Colorado Water Conservation Board, Western Water Assessment, and University of Colorado at Boulder (2013), *Climate Change in Colorado: A Synthesis to Support Water Resources Management and Adaptation*, BiblioGov, Place of publication not identified.

Curran, P. (1980), Multispectral Remote Sensing of Vegetation Amount, *Progress in Physical Geography*, *4*(3), 315–341, doi:10.1177/030913338000400301.

DeFries, R., and K. N. Eshleman (2004), Land-Use Change and Hydrologic Processes: A Major Focus for the Future, *Hydrol. Process.*, *18*(11), 2183–2186, doi:10.1002/hyp.5584.

Dennehy, K. F., D. W. Litke, C. M. Tate, and J. S. Heiny (1993), South Platte River Basin - Colorado, Nebraska, and Wyoming, *JAWRA Journal of the American Water Resources Association*, *29*(4), 647–683, doi:10.1111/j.1752-1688.1993.tb03231.x.

Eakin, H., and A. L. Luers (2006), Assessing the Vulnerability of Social-Environmental Systems, *Annual Review of Environment and Resources*, *31*(1), 365–394, doi:10.1146/annurev.energy.30.050504.144352.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah Land Surface Model Advances in the National Centers for Environmental Prediction Operational Mesoscale Eta Model, *J. Geophys. Res.*, *108*(D22), 8851, doi:10.1029/2002JD003296.

Eriyagama, N., V. Smakhtin, and N. Gamage (2009), *Mapping Drought Patterns and Impacts a Global Perspective*, IWMI Research Report 133, International Water Management Institute, Colombo, Sri Lanka.

Esper, J., D. Frank, U. Büntgen, A. Verstege, J. Luterbacher, and E. Xoplaki (2007), Long-Term Drought Severity Variations in Morocco, *Geophys. Res. Lett.*, *34*(17), L17702, doi:10.1029/2007GL030844.

Fontane, D. G., and D. K. Frevert (1995), Water Management under Drought Conditions: Overview of Practices, Journal of Irrigation and Drainage Engineering, 121(2), 199–206, doi:10.1061/(ASCE)0733-9437(1995)121:2(199).

Gallopín, G. C. (2006), Linkages between Vulnerability, Resilience, and Adaptive Capacity, *Global Environmental Change*, *16*(3), 293–303, doi:10.1016/j.gloenvcha.2006.02.004.

Govender, M., K. Chetty, and H. Bulcock (2007), A Review of Hyperspectral Remote Sensing and its Application in Vegetation and Water Resource Studies, *Water SA*, *33*(2).

Gregorič, G. (2012), Implementation of Drought Monitoring in DMCSEE, in *Drought Management Centre for South-East Europe – DMCSEE. Summary of the result of the project, co-financed by the South east europe transnational Cooperation programme (contract no. See/a/091/2.2/X)*, edited by G. Gregorič, pp. 11–15, Slovenian Environmental Agency.

Grigg, N. S. (2014), The 2011–2012 Drought in the United States: New Lessons from a Record Event, *International Journal of Water Resources Development*, *30*(2), 183–199, doi:10.1080/07900627.2013.847710.

Grigg, N. S., and E. C. Vlachos (1993), Drought and Water-Supply Management: Roles and Responsibilities, *Journal of Water Resources Planning and Management*, *119*(5), 531–541, doi:10.1061/(ASCE)0733-9496(1993)119:5(531).

Hagman, G. (1984), *Prevention better than cure: report on human and environmental disasters in the Third World*, Swedish Red Cross, Stockholm.

Harvey, B. J., D. C. Donato, and M. G. Turner (2016), High and Dry: Post-Fire Tree Seedling Establishment in Subalpine Forests Decreases with Post-Fire Drought and Large Stand-Replacing Burn Patches, *Global Ecol. Biogeogr.*, *25*(6), 655–669, doi:10.1111/geb.12443.

Hurcom, S. J., and A. R. Harrison (1998), The NDVI and Spectral Decomposition for Semi-Arid Vegetation Abundance Estimation, *International Journal of Remote Sensing*, *19*(16), 3109–3125, doi:10.1080/014311698214217.

ITIA N.T.U.A. (2010), *Hydrognomon*, National Technical University of Athens, Athens.

Karavitis, C. A. (1992), Drought Management Strategies for Urban Water Supplies: The Case of Metropolitan Athens, Ph.D. Dissertation, Colorado State University, Fort Collins, CO.

Karavitis, C. A. (1998), Drought and Urban Water Supplies: The Case of Metropolitan Athens, *Water Policy*, *1*(5), 505–524, doi:10.1016/S1366-7017(99)00009-4.

Karavitis, C. A. (1999), Decision Support Systems for Drought Management Strategies in Metropolitan Athens, *Water International*, *24*(1), 10–21, doi:10.1080/02508069908692129.

Karavitis, C. A. (2012), Drought Vulnerability Assessment - Introduction and Theoretical Background, in *Drought Management Centre for South-East Europe – DMCSEE. Summary of the result of the project, co-financed by the South east europe transnational Cooperation programme (contract no. See/a/091/2.2/X)*, edited by G. Gregorič, pp. 27–32, Slovenian Environmental Agency.

Karavitis, C. A., S. G. Alexandris, V. P. Fassouli, D. Stamatakos, D. E. Tsesmelis, and N. A. Skondras (2011), Vulnerability Assessment, Task 4.2.5, DMCSEE project, in *5th DMCSEE Consortium Meeting and Training, 28 June–1 July 201*, Lasko, Slovenia.

Karavitis, C. A., C. Chortaria, S. G. Alexandris, C. G. Vasilakou, and D. E. Tsesmelis (2012a), Development of the Standardised Precipitation Index for Greece, *Urban Water Journal*, *9*(6), 401–417, doi:10.1080/1573062X.2012.690431.

Karavitis, C. A., N. A. Skondras, D. E. Tsesmelis, D. Stamatakos, S. G. Alexandris, and V. P. Fassouli (2012b), Drought Impacts Archive and Drought Vulnerability Index, in *Drought Management Centre for South-East Europe – DMCSEE. Summary of the result of the project, co-financed by the South east europe transnational Cooperation programme (contract no. See/a/091/2.2/X)*, edited by G. Gregorič, pp. 33–37, Slovenian Environmental Agency.

Karavitis, C. A., D. E. Tsesmelis, N. A. Skondras, D. Stamatakos, S. Alexandris, V. Fassouli, C. G. Vasilakou, P. D. Oikonomou, G. Gregorič, N. S. Grigg and E. C. Vlachos (2014), Linking Drought Characteristics to Impacts on a Spatial and Temporal Scale, *Water Policy*, *16*(6), 1172–1197, doi:10.2166/wp.2014.205.

Karavitis, C. A., C. G. Vasilakou, D. E. Tsesmelis, P. D. Oikonomou, N. A. Skondras, D. Stamatakos, V. Fassouli, and S. Alexandris (2015), Short-Term Drought Forecasting Combining Stochastic and Geo-Statistical Approaches, *European Water*, *49*, 43–63.

Karnieli, A., N. Agam, R. T. Pinker, M. Anderson, M. L. Imhoff, G. G. Gutman, N. Panov, and A. Goldberg (2010), Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations, *J. Climate*, *23*(3), 618–633, doi:10.1175/2009JCLI2900.1.

Kavalieratou, S., D. K. Karpouzos, and C. Babajimopoulos (2012), Drought Analysis and Short-Term Forecast in the Aison River Basin (Greece), *Natural Hazards and Earth System Science*, *12*(5), 1561–1572, doi:10.5194/nhess-12-1561-2012.

Knipling, E. B. (1970), Physical and Physiological Basis for the Reflectance of Visible and Near-Infrared Radiation from Vegetation, *Remote Sensing of Environment*, *1*(3), 155–159, doi:10.1016/S0034-4257(70)80021-9.

Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. M. Baker (1999), A Parameterization of Snowpack and Frozen Ground intended for NCEP Weather and Climate Models, *J. Geophys. Res.*, *104*(D16), 19569–19585, doi:10.1029/1999JD900232.

McKee, T. B., N. J. Doesken, J. Kleist, and others (1993), The Relationship of Drought Frequency and Duration to Time Scales, in *Proceedings of the 8th Conference on Applied Climatology*, vol. 17, pp. 179–183, American Meteorological Society Boston, MA, USA.

McKee, T. B., N. J. Doesken, J. Kleist, C. J. Shrier, and W. P. Stanton (2000), *A History of Drought in Colorado: Lessons learned and what lies ahead*, Colorado State University, Colorado Water Resources Research Institute, Fort Collins, CO.

Mishra, A. K., and V. R. Desai (2005), Drought Forecasting using Stochastic Models, *Stoch Environ Res Ris Assess*, *19*(5), 326–339, doi:10.1007/s00477-005-0238-4.

O'Brien, K., and R. Leichenko (2001), The Dynamics of Vulnerability to Global Change, *International Human Dimensions Programme on Global Environmental Change (IHDP) Newsletter Update*, *01/2001*.

O'Brien, K. et al. (2004), Mapping Vulnerability to Multiple Stressors: Climate Change and Globalization in India, *Global Environmental Change*, *14*(4), 303–313, doi:10.1016/j.gloenvcha.2004.01.001.

Organisation for Economic Cooperation and Development (2008), *Handbook on Constructing Composite Indicators: Methodology and User Guide*, OECD Publishing, Paris.

Pritchett, J., C. Goemans, and R. Nelson (2013), *Estimating the Short and Long-term Economic & Social Impacts of the 2012 Drought in Colorado*, Colorado State University, Department of Agricultural and Resource Economics.

Rosenberg, N. J. (Ed.) (1978), *North American Droughts*, AAAS selected symposium ; 15, Westview Press, Boulder, Colo.

Rosenberg, N. J. (Ed.) (1980), *Drought in the Great Plains: Research on Impacts and Strategies: Proceedings of the Workshop on Research in Great Plains Drought Management Strategies held at the University of Nebraska-Lincoln, March 26-28, 1979*, Water Resources Publications, Littleton, Colo.

Rouse, J. W. (1974), *Monitoring the Vernal Advancement and Retrogradation (green wave effect) of Natural Vegetation*, NASA-CR-144661, NASA, Washington, DC, United States.

Rui, H., and D. Mocko (2014), Readme document *for North America Land Data Assimilation System Phase 2 (NLDAS-2) Products*, https://hydro1.gesdisc.eosdis.nasa.gov/data/NLDAS/README.NLDAS2.pdf.

Shiau, J. T. (2006), Fitting Drought Duration and Severity with Two-Dimensional Copulas, *Water Resour Manage*, *20*(5), 795–815, doi:10.1007/s11269-005-9008-9.

Shiau, J. T., and R. Modarres (2009), Copula-Based Drought Severity-Duration-Frequency Analysis in Iran, *Met. Apps*, *16*(4), 481–489, doi:10.1002/met.145.

Spruce, J. P., G. E. Gasser, and W. W. Hargrove (2016), MODIS NDVI Data, Smoothed and Gap-filled, for the Conterminous US: 2000-2015, *ORNL DAAC, Oak Ridge, Tennessee, USA*, doi:10.3334/ORNLDAAC/1299.

Stenzel, R., and T. V. Cech (2013), *Water, Colorado's Real Gold: A History of the Development of Colorado's Water, the Prior Appropriation Doctrine, and the Colorado Division of Water Resources*, Richard Stenzel, Denver, Colo.

Sullivan, C. (2002), Calculating a Water Poverty Index, *World Development*, *30*(7), 1195–1210, doi:10.1016/S0305-750X(02)00035-9.

Sullivan, C. A. (2011), Quantifying Water Vulnerability: A Multi-Dimensional Approach, *Stoch Environ Res Risk Assess*, *25*(4), 627–640, doi:10.1007/s00477-010-0426-8.

Sullivan, C. A., and J. Meigh (2007), Integration of the Biophysical and Social Sciences using an Indicator Approach: Addressing Water Problems at Different Scales, *Water Resour Manage*, *21*(1), 111–128, doi:10.1007/s11269-006-9044-0.

Sullivan, C. A. et al. (2003), The Water Poverty Index: Development and Application at the Community Scale, *Natural Resources Forum*, *27*(3), 189–199, doi:10.1111/1477-8947.00054.

Sullivan, C. A., N. Diederichs, and M. Mander (2009), *Assessing Water Vulnerability in the Orange River Basin in South Africa*, NeWater technical report, Oxford, UK.

Tadesse, T., B. D. Wardlow, J. F. Brown, M. D. Svoboda, M. J. Hayes, B. Fuchs, and D. Gutzmer (2014), Assessing the Vegetation Condition Impacts of the 2011 Drought across the U.S. Southern Great Plains Using the Vegetation Drought Response Index (VegDRI), *J. Appl. Meteor. Climatol.*, *54*(1), 153–169, doi:10.1175/JAMC-D-14-0048.1.

Traore, Z. N., and D. G. Fontane (2007), Managing Drought Impacts: Case Study of Mali, Africa, Journal of Water Resources Planning and Management, 133(4), 300–308, doi:10.1061/(ASCE)0733-9496(2007)133:4(300).

Tsakiris, G., and H. Vangelis (2004), Towards a Drought Watch System Based on Spatial SPI, *Water Resources Management*, *18*(1), 1–12.

Tsakiris, G., and H. Vangelis (2005), Establishing a Drought Index Incorporating Evapotranspiration, *European Water*, *9/10*, 3–11.

Tsakiris, G., D. Pangalou, and H. Vangelis (2006), Regional Drought Assessment Based on the Reconnaissance Drought Index (RDI), *Water Resour Manage*, *21*(5), 821–833, doi:10.1007/s11269-006-9105-4.

Tsesmelis, D. E., P. D. Oikonomou, C. G. Vasilakou, V. Fassouli, N. A. Skondras, D. Stamatakos, I. I. Gkotsis, S. Alexandris, and C. A. Karavitis (2017), Assessing Model Uncertainty caused by Different Weighting Methods on Composite Indicators: The case of the SPI-based Drought Vulnerability Index, Manuscript submitted for publication.

United Nations Office for Disaster Risk Reduction (UNISDR) (2004), *Living with Risk: A Global Review of Disaster Reduction Initiatives*, United Nations, New York.

Vasiliades, L., and A. Loukas (2009), Hydrological Response to Meteorological Drought using the Palmer Drought Indices in Thessaly, Greece, *Desalination*, *237*(1–3), 3–21, doi:10.1016/j.desal.2007.12.019.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010), A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, *J. Climate*, *23*(7), 1696–1718, doi:10.1175/2009JCLI2909.1.

Vlachos, E. C. (1982), Drought Management Interfaces, in *Annual ASCE Conference*, p. 15, Las Vegas, Nevada.

Wilhite, D. A. (2004), Drought, in *International Perspectives on Natural Disasters: Occurrence, Mitigation, and Consequences*, edited by J. P. Stoltman, J. Lidstone, and L. M. Dechano, pp. 147–162, Springer Netherlands.

Wilhite, D. A., M. D. Svoboda, and M. J. Hayes (2007), Understanding the Complex Impacts of Drought: A Key to Enhancing Drought Mitigation and Preparedness, *Water Resour Manage*, *21*(5), 763–774, doi:10.1007/s11269-006-9076-5.

Writer, J. H., A. Hohner, J. Oropeza, A. Schmidt, K. Cawley, and F. L. Rosario-Ortiz (2014), Water Treatment Implications after the High Park Wildfire, Colorado, *Journal - American Water Works Association*, *106*, E189–E199, doi:10.5942/jawwa.2014.106.0055.

Wu, H., M. D. Svoboda, M. J. Hayes, D. A. Wilhite, and F. Wen (2007), Appropriate Application of the Standardized Precipitation Index in Arid Locations and Dry Seasons, *Int. J. Climatol.*, *27*(1), 65–79, doi:10.1002/joc.1371.

Xia, Y. et al. (2012), Continental-scale Water and Energy Flux Analysis and Validation for the North American Land Data Assimilation System Project Phase 2 (NLDAS-2): 1. Intercomparison and Application of Model Products, *J. Geophys. Res.*, *117*(D3), D03109, doi:10.1029/2011JD016048.

Xu, L., A. Samanta, M. H. Costa, S. Ganguly, R. R. Nemani, and R. B. Myneni (2011), Widespread Decline in Greenness of Amazonian Vegetation due to the 2010 Drought, *Geophysical Research Letters*, *38*(7), n/a-n/a, doi:10.1029/2011GL046824.

Yevjevich, V., L. da Cunha, and E. Vlachos (1983), *Coping with Droughts*, Water Resources Publications, Littleton, Colorado.

Zender, C. S. (2016), *netCDF Operator (NCO) User Guide Version 4.6.2*.

3 Water for Unconventional Oil and Gas Development and the Existing Data Challenges[1]

3.1 Introduction

There is an intricate connection between energy and water resources formulating the water-energy nexus [*Gleick*, 1994]. A subset of this nexus is the supply and use of water for extracting unconventional oil and gas resources. Technological advancements, over the past decade, made it possible to have a boom in unconventional oil and gas sector across the United States [*Murray*, 2013] offering alternative sources of oil and gas and boosting local economies [*Higginbotham et al.*, 2010; *Considine et al.*, 2011]. Studies have indicated that the use of unconventional sources of energy, especially shale gas, could decrease the overall water consumption for energy production if such sources are going to offset the water use of coal power plants [*Grubert et al.*, 2012; *Laurenzi and Jersey*, 2013; *Pacsi et al.*, 2014]. While there are opportunities generated from the unconventional oil and gas development, there are also potential environmental and health related challenges [*Osborn et al.*, 2011; *Clark et al.*, 2012; *McKenzie et al.*, 2012; *Pacsi et al.*, 2013; *Walton and Woocay*, 2013; *Colborn et al.*, 2014; *Vengosh et al.*, 2014; *Mehany and Guggemos*, 2015; *Burton et al.*, 2016; *Goodman et al.*, 2016; *McLaughlin et al.*, 2016]

The rise of unconventional oil and gas exploration has created a "new" water user with an increasing demand, which is competing with traditional water demands. Only recently, few studies have reported water use estimates for the United States [*Gallegos et al.*, 2015; *Kondash and

[1] The research in this chapter was conducted under the AirWaterGas Sustainability Research Network (http://airwatergas.org/) funded by the National Science Foundation under Grant No. CBET-1240584. Any opinion, findings, and conclusions or recommendations expressed in this chapter are those of the author and do not necessarily reflect the views of the National Science Foundation.

Vengosh, 2015; *Chen and Carter*, 2016], but because of their scope of study fail to represent differences localities. Several of the developed regions are exhibiting an arid and semi-arid climate, which can be seen as an additional stress to the limited available water resources. At the same time, the concerns regarding water quality [*Vengosh et al.*, 2014] along with water quantity became a priority on the sociopolitical agenda, since the oil and gas activity, due to the development boom, was taking place in closer proximity to habited areas than it did in the past.

Despite the fact that data need to play a critical role in the complex decision-making surrounding the water-energy nexus, plenty of datasets (i.e. source of water) are either not easily accessible or they are unavailable, hindering trade-off analysis. Openly available data, collected in a reliable and transparent fashion, which answer expert and public concerns are needed to inform decisions made by oil and gas operators, water managers, and policymakers concerning water use related to energy production [*Goodwin et al.*, 2013; *Malone et al.*, 2015; *Nicot and Scanlon*, 2012; *Orford*, 2013]. Such a practice would promote resource sustainability and facilitate integrated water resources planning and management.

In Colorado, the formations of the Denver-Julesburg Basin in the northeast part of the state and the Piceance Basin located in western Colorado have experienced the greatest concentration of new drilling (Figure 20) [*Colorado Oil and Gas Conservation Commission*, 2015]. The Denver-Julesburg Basin is producing oil and gas, where the Piceance Basin is exploited mainly for its natural gas. The 2014 state's crude oil production increased to 95,192 thousand barrels (49,435 in 2012) [*U.S. Energy Information Administration*, 2015b] while the gross natural gas production was 1,631,390 million cubic feet (1,709,376 in 2012) [*U.S. Energy Information Administration*, 2015a]. At the same time, the statewide water use for oil and gas activities is estimated to be approximately 0.1% of overall water use, or 22,202,673 m^3 per year [*Colorado Division of Water*

Resources et al., 2012]. The quantity of water used per well is dependent on several different factors, including the geologic formation, the type of well (vertical, directional, horizontal), the number of hydraulic fracturing stages, the distance of the total reach within the production zone, and the type of hydraulic fracturing fluid (e.g. cross-linked gel or slickwater). Thus aggregating the estimated water use on state-level isn't useful on a local scale since development is taking place in few basins and also different factors can vary significantly at the local level, including competing water demands [*Clark et al.*, 2013; *Nicot and Scanlon*, 2012; *Rahm and Riha*, 2012].

In the Denver-Julesburg Basin, Goodwin *et al*. [2013] found that horizontal reaches in productive shale formations are increasing from 1.6 km to 3.2 km or longer. This has a direct impact on water use consumption per each well, since the length of the reach is associated to the number of stages and subsequently with total quantity of water used during completion. The same study reported the median water requirement in the Denver-Julesburg Basin was 1,363 m^3 per vertical well, 10,868 m^3 per horizontal well while the extended horizontal wells (greater than 25 hydraulically-fractured stages) require more water (median 21,274 m^3). Studies examining water use in the Piceance Basin, and more specifically within Garfield County, report a wide range of water use estimates. A study conducted by the U.S. Environmental Protection Agency [2012], which used FracFocus data, estimated a range of 3,785 to 34,069 m^3 per well (median 4,921 m^3), while another study suggests a median of 6,462 m^3 per well [*U.S. Environmental Protection Agency*, 2013]. Comparing the reported numbers of the aforementioned studies, it becomes apparent that there are data inconsistencies that can create barriers to predicting future water use and associated tradeoffs for making policy decisions and developing regulations.

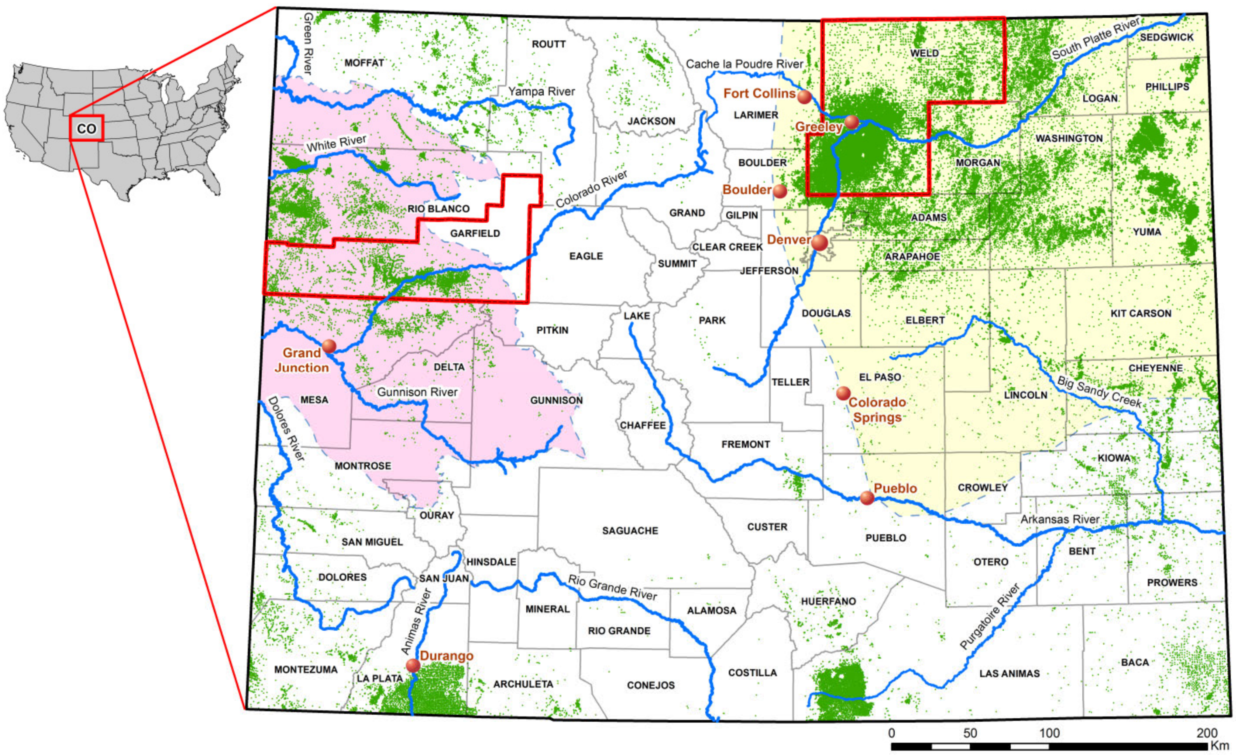


Figure 20: Map of oil and gas wells in Colorado (wells shown as green dots, the polygon in yellow is the Denver-Julesburg Basin and Piceance Basin in pink). The number of producing wells in Weld County and Garfield counties as of 14/7/2015 is 20,918 and 10,499 respectively. [Colorado Oil and Gas Conservation Commission, 2015]

For developing and implementing water resources planning and management that would include water use for unconventional oil and gas, it is imperative to understand the whole cycle of the water used for this industry. Furthermore, identifying the characteristics influencing industry's development patterns decisions that affect water use intensity along with the source of the water is a key to better understand this "new" water demand. The need for accessible and reliable data is the basis of developing better water use predictions, which would support complex decisions that involve local communities and other competing and conflicting water uses, such as agriculture.

The present effort analyzed and compared publicly available water use datasets and identified well development patterns, normalized water use per well, and flowback and disposal practices that are taking place in Garfield and Weld counties. The two Colorado counties, although governed by the same rules, represent two distinctly different basins in terms of geophysical characteristics and producing products, which makes this study more representative. In response to the question of effects on local water resources, this effort explores briefly the sources of water acquisition for oil and gas in these counties, and the relative stress imposed. The additional total water demand from oil and gas within the boundaries of each county was quantified and compared to the total water withdrawals generated locally. Additionally, this work attempts to identify water intensity in a HUC-12 level in Weld County and compare it with other water uses. Finally, the study discusses the existing data limitations and the relevance of the analysis for other regions with similar industrial activity.

3.2 Methodology and Data

This multi-scale analysis (well, county and HUC-12 level) and comparative case study approach used a variety of sources and several publicly available datasets in order to investigate the oil & gas well development patterns and the water that is being used; and what are the channels this water is coming from and how it compares with other uses on a county level and a HUC-12 level. The methodology described below was designed to fill knowledge gaps about the aforementioned goals but also to reveal the existing data challenges and practices that hinder the efficient and effective study of the whole life cycle of the water used for unconventional oil and gas exploration.

3.2.1 Water Use and Oil & Gas Well Development Patterns

Well characteristics and water consumption data per well were extracted from state's organizations publicly available databases and the FracFocus database [*FracFocus Chemical Disclosure Registry*, 2015]. Completed Interval Reports for both counties were extracted from January 2011 through 2014. At the time of the research activity, the available data in 2014 for Garfield County were current up to July 2014 while for Weld County were current until April 2014. It should be mentioned, that the existing administrative protocols of the Colorado Oil and Gas Conservation Commission (http://cogcc.state.co.us/) requires operators to report their water use through the Completed Interval Report form (Form 5A) within 30 days of well completion. Data collected on active oil and gas wells contained "total volume of fluids used in treatment", because the amount of the water was not reported frequently in the completion forms. Furthermore, information was collected, if available, about the type of the water (fresh or recycled) used, the amount of flowback water recovered within the period of well completion and reporting, the disposal methods for flowback water and the number of staged intervals.

The Completed Interval Report dataset was subject to a thorough quality assurance and quality control (QA/QC) process that included consolidation of duplicate well entries with fracture jobs in several geologic formations, correction of entries reporting values in incorrect units, and exclusion of entries with missing values. Furthermore, in order to make sure that discrepancies were not present, several QA/QC tests were performed on each record (e.g., amount of total fluids used should have been greater than flowback water). After consolidating duplicate well entries, the records were 1077 for Garfield County and 2076 for Weld County, which resulted in 712 and 2022 records, respectively for the two counties after the QA/QC procedure. The main reason for discarding a record was lack of reporting total fluids used. Lastly, outliers and spurious data were detected and excluded from further analysis by using an adjusting boxplot method [*Hubert and Vandervieren*, 2008] that accounts for distributional asymmetry. This approach was chosen instead of other methods (extreme Studentized deviation, percentiles) because the study data were not normally distributed and contained outliers.

Since one of the factors influencing the quantity of water used per well is the well type (vertical, directional, and horizontal), it is essential to categorize the wells into different types. The identification of the directional and horizontal wells was possible through the directional well database by Colorado Oil and Gas Conservation Commission. Along with the well type, the length of each well was retrieved. Any well that was not included in the the directional well database was assumed to be a vertical well. The information retrieved from this database (well ID, well type and length) was joined with the COGCC's Completed Interval Report dataset, since the type of each well is not provided at the 5A Forms.

FracFocus (http://fracfocus.org) is an online hydraulic fracturing related database, which covers the largest part of the USA. It is managed by the Ground Water Protection Council (http://www.gwpc.org) and Interstate Oil and Gas Compact Commission (http://iogcc.publishpath.com). Available FracFocus data for the two counties included the combined water volume on a per well basis without specifying the water was fresh or recycled, and if fresh, its source. Information from FracFocus was extracted for the study period of 2011-2014. Prior to April 1, 2012, disclosure of certain well completion information by Colorado oil and gas companies was reported to FracFocus on a volunteer basis. Since then, operators have been required to report data to FracFocus within 60 days following the conclusion of a hydraulic fracturing treatment. The data extracted, were evaluated through a QA/QC process. The screening for outliers resulted in keeping 1,843 producing gas well records in Garfield County from 1852, and 4,046 producing oil and gas well records in Weld County from 4093. Although FracFocus database is easily accessible and gives the option of bulk download, compared to the COGCC, the information provided is less detailed regarding water use. Thus, FracFocus for this scope was used mainly for cross-checking the statistical characteristics of the total water volume used with that of the COGCC dataset.

3.2.2 Water Acquisition Pathways and Water Use on County Scale

Water sourcing for oil and gas development in Colorado is not reported, since there is no relevant state or federal regulation in place at this time, making the gathering of this information challenging. Thus, the information is limited and scattered across several datasets (water right decrees and substitute water supply plans) which hinders a thorough analysis. The water use information gathered at the previous section (3.2.1) was aggregated and was compared to the county water use estimates of the report by Maupin et al. [2014]. The different water source pathways are reported here, in relation to the estimation of the amount of water used in oil and gas on county scale. Furthermore, a significant difference of water administration between Garfield and Weld counties is the existence of nontributary (for definition see **LIST OF TERMS**) groundwater wells in Weld County.

Oil and gas operators see the nontributary groundwater wells located in the South Platte basin as a reliable source of water for the industry. This water is not subject to administration within the prior appropriation law, meaning that there is no augmentation obligation to senior water right owners. The Colorado Division of Water Resources (DWR) open datasets (https://data.colorado.gov/browse?category=Water) and Hydrobase were queried in order to identify the nontributary groundwater wells used for unconventional oil and gas operations. Although the datasets are products of the same organization, the different formatting and attributes made it necessary to use both. A significant obstacle of easy identification of such wells is that the state does not categorize the water decrees used for oil and gas separately from other industrial uses. This fact increases the chances of not capturing all nontributary groundwater wells used to supply water to unconventional oil and gas activities, and thus have uncertain estimates of annual maximum allowed water withdrawals from these wells.

Since there was not a user-friendly and efficient way to acquire the list of the nontributary groundwater wells used for oil and gas exploration and their associated attributes, a three-step process of data collection, screening and analysis was adopted. First, all the recorded active water well structures located within Weld County were identified and the nontributary wells were queried from the Department of Water Resources structures dataset. Then, a code was developed (Visual Basic for Applications in Microsoft Excel) to automatically retrieve the permit attributes (such as permitted use) from the website version of Hydrobase for the water well structures that had an associated permit number. This way it was possible to narrow down the nontributary water wells to those that included a permitted industrial use. The Department of Water Resources permits database was also used to identify industrial water well permits without an associated active well structure and to record the issued nontributary groundwater permits with unknown well construction status. The latter is important because it depicts the current status and gives an accurate picture of what the future water extraction capacity from those wells. The final step involved evaluating the documentation of each final candidate well permit in order to identify the nontributary groundwater wells used in oil and gas operations.

3.2.3 Extreme Oil & Gas Development Scenario in Weld County

Water amounts used in the oil and gas industry are usually aggregated to state level or per shale play, and as a result not depicting localities that might be important. An analysis of water use on a finer scale is central for shedding light on water competition and the tradeoffs between other uses (agricultural, industrial, thermoelectric and domestic), especially in a county like Weld County that is one of the top ten agricultural producing US counties.

For this scope, the FracFocus dataset was used to retrieve information about drilling locations and total water consumption for the 2011-2014 period and manipulated in the ArcGIS environment in order to analyze the per HUC-12 consumption every year. An extreme oil and gas development scenario was selected representing a high development scenario where water demand for each HUC-12 takes the maximum individual value from the period 2011-2014 (not overall). EPA's EnviroAtlas [*Pickard et al.*, 2015] is providing information, among others, on water demand estimates in a HUC-12 level. In each HUC-12, these water demand estimates for domestic, industrial, agricultural and hydroelectric were compared with the water demand for oil and gas. Lastly, assuming a normal water year for irrigation needs for two major crops in Weld County, corn and alfalfa, the net irrigation requirements were estimated based on the difference of average seasonal crop consumptive use (31.58 in for alfalfa and 21.74 in for corn) and effective precipitation (7.32 in) in Greeley, CO [*U.S. Department of Agriculture Soil Conservation Service*, 1988]. The total water demand used for oil and gas in each HUC-12 was converted to fully irrigated acreage of alfalfa or corn as a measure of the tradeoff between water for energy versus water for food.

3.3 Results & Discussion

3.3.1 Development Trends of Oil and Gas Wells

In Colorado, Garfield and Weld counties dominate gas and oil extraction respectively (Figure 21), and different oil and gas development practices are taking place in the two counties. Garfield Country has experienced a declining rate of new well starts (Figure 21) over the last six years, and as shown in Figure 22, directional wells are drilled more than other well types (vertical and horizontal). In Weld County, about 1,500 new well starts were recorded every year since 2011 and the preferred well type shifted during the last years from vertical to horizontal wells (Figure 22). Operational decisions and the geological characteristics of the shale formations, namely the exploitable thickness of the formations in Piceance Basin are greater than in the Denver-Julesburg Basin, are the main reasons for the difference in well types trends between the two counties. Figure 23 is a graphical combination of Box-and-Whisker plot and kernel density estimation (the curves on each side of the Box-and-Whiskers plot), a useful non-parametric technique for visualizing the distribution of the sample data. It portrays the increase of horizontal reach length from 1.5 km median length (2011) to more than 1.8 km (2014). In many cases it is extending beyond 3.0 km, which is in agreement with the findings of Goodwin *et al*. [2013]. The rise of median staged intervals per well in Weld County (Figure 24) can be explained by the increased numbers of horizontal wells drilled and longer lateral lengths. The different shape of the stage interval kernel density estimation for 2014 in Weld County (Figure 24) can be explained due to the fact that there were very few vertical and directional wells completed by April, 2014. There is no significant change in the median number of staged intervals in Garfield County (Figure 24). The outliers in Figure 24 depict the high number of staged intervals of horizontal wells.

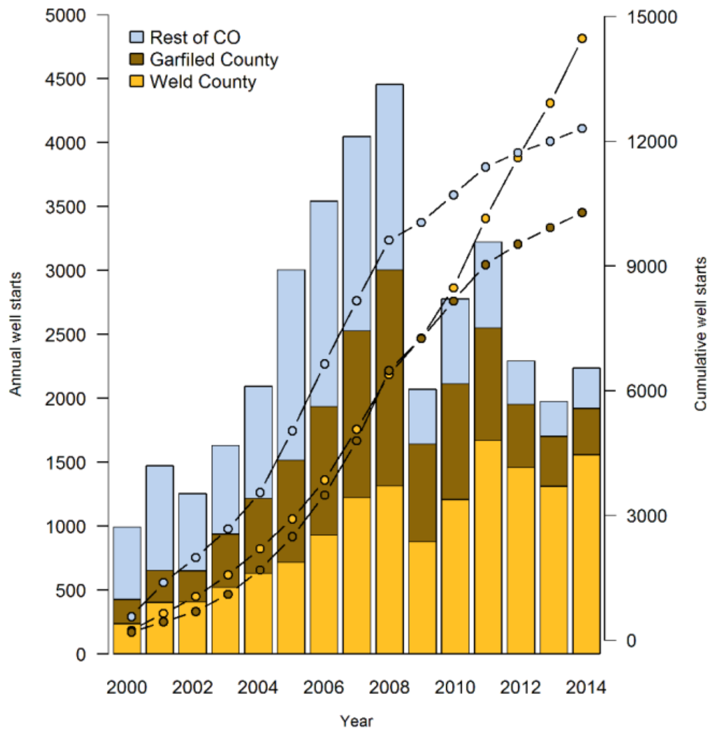


Figure 21: Annual (bars) and cumulative (lines) hydraulic fracturing well starts (COGCC 2015)

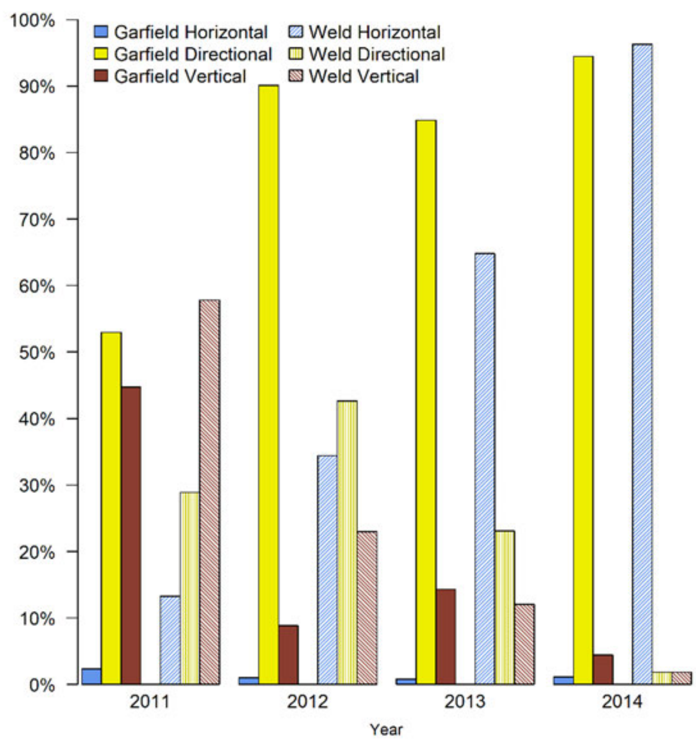


Figure 22: Well type trends in Garfield and Weld counties, 2011-2014 (COGCC dataset)

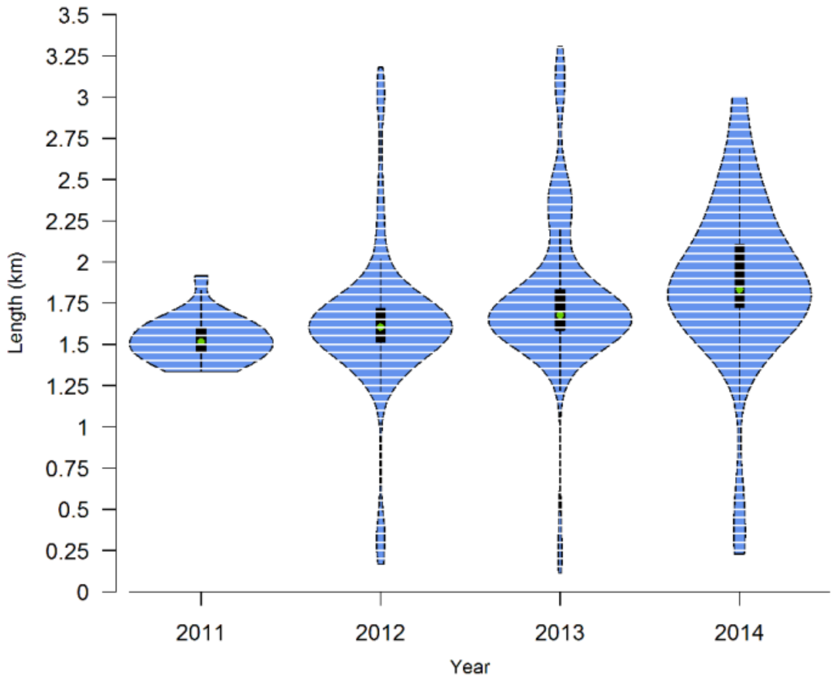


Figure 23: Length of horizontal wells in Weld County and their rotated kernel densities on each side (COGCC dataset)

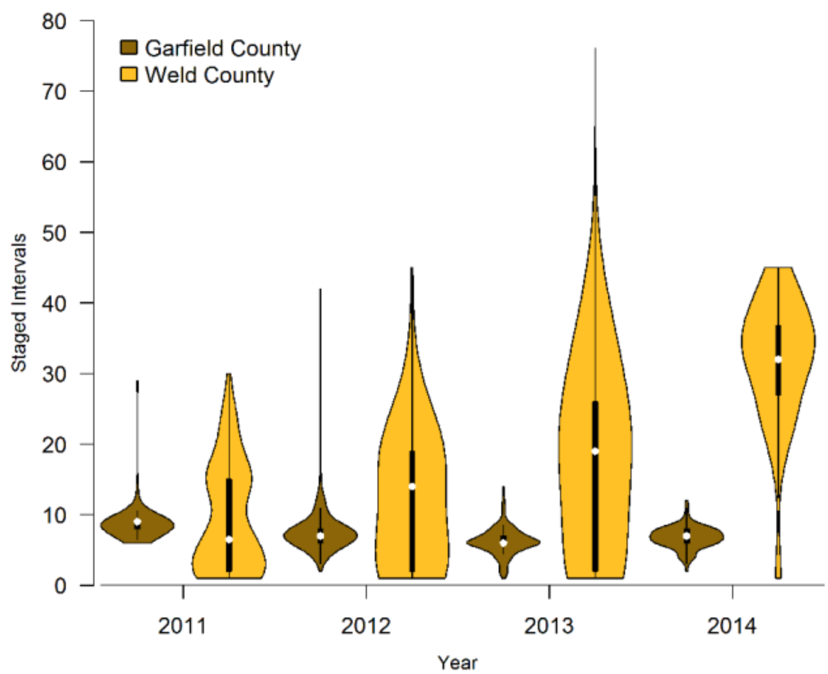


Figure 24: Stage intervals for Garfield and Weld counties and their rotated kernel densities on each side (COGCC dataset)

3.3.2 Water Use for Oil and Gas Industry

Several different factors are responsible for the water needed in oil and gas well completions, including the geologic formation, the type of well, the number of hydraulic fracturing stages, the length of the reach within the production zone, and the type of hydraulic fracturing fluid (cross-linked gel or slickwater). Reporting a normalized water volume on a per well basis is more appropriate rather than lumping records with dissimilar attributes that could skew the results. The calculation of the water amount used per stage results in more precise determination of total water use per well as suggested by Goodwin et al. [2013].

Furthermore, the disaggregation of water use information into single-stage wells vs. multi-stage wells is essential, because they represent different forms of development. In Weld County, the normalized total fluid amount used per staged interval in multi-staged wells is less than for single-staged wells. Figure 25 presents the annual normalized total fluids amount used per staged interval for Weld County. The water used per stage for single-staged wells is 783.01 m^3 and for multi-staged wells 542.62 m^3/stage. In Garfield County, due to very few records available for single-staged wells in Garfield County for the period 2011-2014, it was not possible to estimate the statistical characteristics of the normalized total fluids per stage. The median normalized total fluid amount used per staged interval at multi-staged wells in Garfield County is 682.02 m^3 per stage.

The investigation of the total volume of fluid per length for directional and horizontal wells is key metric in order to understand the effects of different development practices on water requirements. The difference in the normalized total volume of fluid per length in both counties for horizontal and directional wells within the study period are portrayed in Figure 26. From the analysis of the data, in both counties, horizontal wells tend to require more water than directional wells. In Garfield County, the median normalized total volume of fluid per length for horizontal wells is 19.66 m^3 per m and for directional wells is 11.02 m^3 per m. There are cases where directional wells exceeded 50 m^3/ m. In Weld County, the median normalized total volume of fluid per well length for horizontals is 7.25 (m^3 per m) and for directional wells is 3.43 (m^3 per m). The upward trend in horizontal well length (Figure 23) increases the number of stages, and subsequently, the total quantity of water used during completion.

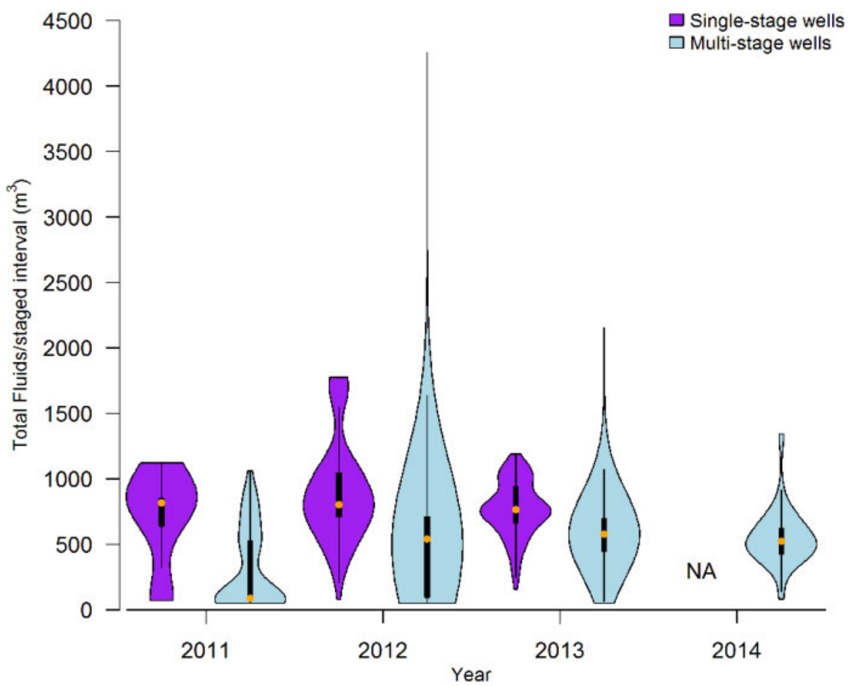


Figure 25: Normalized total volume of fluid used per staged interval in Weld County and their rotated kernel densities on each side (NA: not enough data) (COGCC dataset)

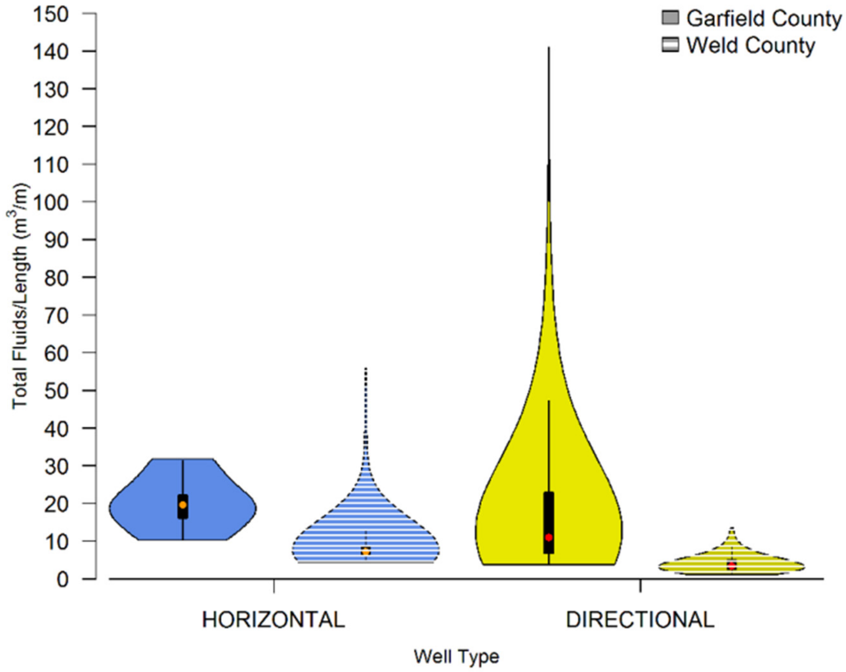


Figure 26: Normalized total volume of fluid used per length of well and their rotated kernel densities on each side (COGCC dataset)

The comparison of COGCC and FracFocus datasets for Weld and Garfield counties provides a quality check for discrepancies between the two. At the same time it should be underlined that the water consumed for oil and gas development reported to COGCC database includes all liquids used (treatment fluid, acid and water), while the FracFocus database includes the total amount of water. In Garfield County, the median volume of water per well reported in the COGCC database is 5,441.02 m^3 and in FracFocus is 7,138.39 m^3. For Weld County, the median volume of water per well is 2,740.62 m^3 for the COGCC dataset and 1,706.94 m^3 for FracFocus. Some small inconsistency was expected be present between the two datasets because the reported amounts are not exactly the same. Figure 27 presents the annual distributions and the quantiles in both counties for both datasets. The differences depicted in Figure 27 between the datasets raise questions about the quality of the reported data.

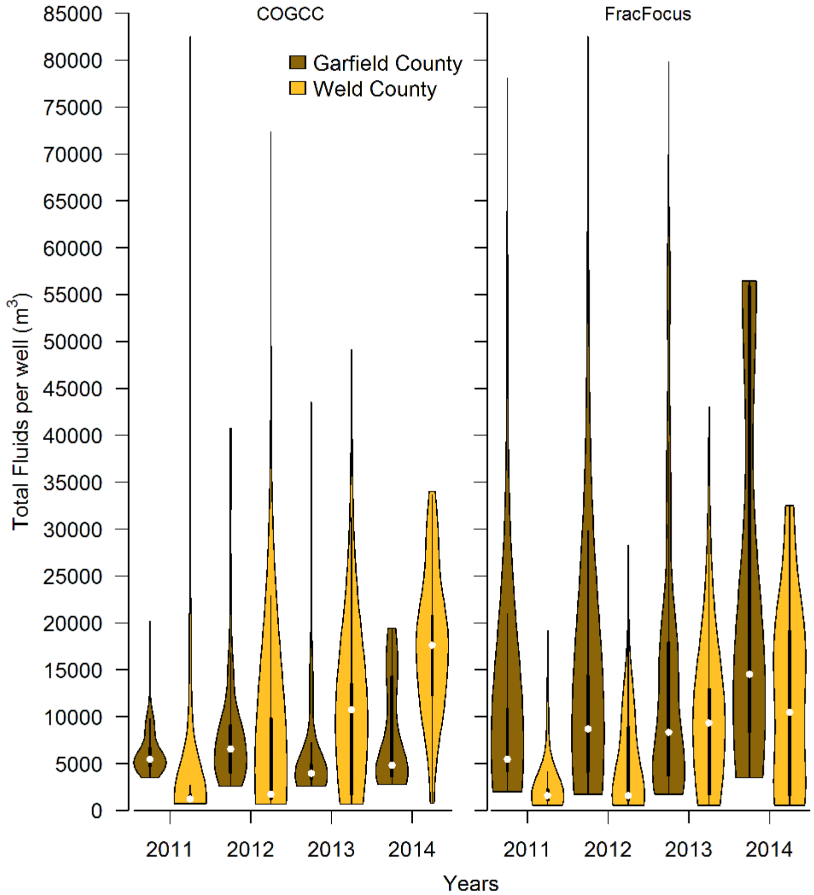


Figure 27: Total volume of fluid consumed for hydraulic fracturing on a per well basis and their rotated kernel densities on each side (COGCC and FracFocus dataset)

3.3.3 Flowback Water and Disposition Strategies

The main weakness of the flowback water quantities reported at the COGCC through the Form 5A could be skewed since it is not representing the water retrieved based on a fix number of days after completion. In reality, this means that in every oil and gas well the operator could take into account different amount of days in order to compute the flowback water, but never exceeding 30 days from the well completion date. For the purpose of this effort, the flowback water is the water volume that returns to the surface between completion and reporting to COGCC (max of 30 days). Figure 28 presents, for both counties, the annual flowback water as a percentage of total fluids used. COGCC data show that the flowback water is about 48% of the total volume of fluid injected in Garfield County. On the other hand, flowback water in Weld County is only 11% of the total volume of fluid injected.

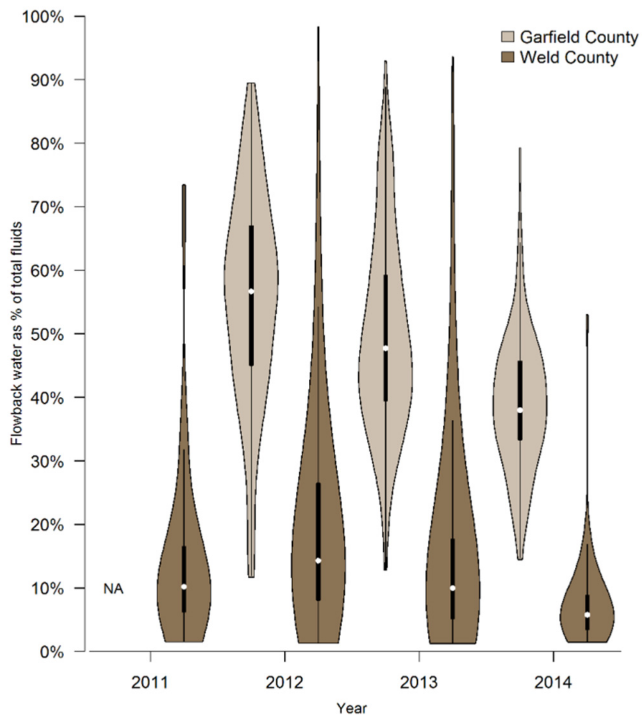


Figure 28: Flowback water as percentage of total fluids used and their rotated kernel densities on each side (NA: not enough data) (COGCC dataset)

Figure 29 portrays the fate of the flowback water expressed as percentages. The reported amount of flowback water recycled in Garfield County is 96% compared to only 28% in Weld County. The percentage of flowback water retrieved could be linked with how much of the flowback water is reused. The data indicate that in Garfield County a significant amount of the fluids recovered could have created incentives for the industry to adopt a pro-reuse strategy. At the same time, companies in Garfield County that recycle nearly all of their flowback water are also acquiring new water rights. This could be explained from the fact that more water is needed for new hydraulic fracturing jobs than the readily available recovered flowback water. Although there is a shift towards reuse in Weld County, the magnitude of the 2014 shift illustrated in Figure 29 could be an artifact of incomplete data. However, the adoption of pro-reusing and pro-recycling practices for the flowback water appears to be of interest to oil and gas companies as these practices can decrease the diversions of freshwater which would be beneficial especially during drought periods, but also it would reduce the amounts going to deep well injection disposal and thus reducing the risk of seismic activity. The COGCC well completion records indicate that fresh water is the primary water source for new wells in Weld County. On the contrary, in Garfield County the water used in hydraulic fracturing is mainly recycled water.

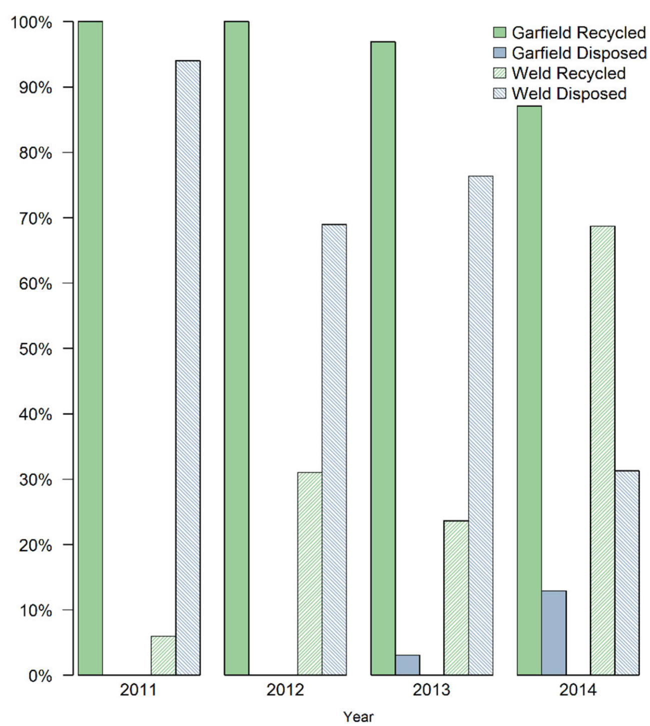


Figure 29: Percentage of annual flowback water recycled or disposed in Garfield and Weld counties (COGCC dataset)

3.3.4 Water Acquisition Sources in Weld and Garfield Counties

All water rights in Colorado have a legal designated use. Water acquired to be used in oil and gas development must be adjudicated as industrial or multi-use water (see LIST OF TERMS). Understanding how water is acquired, by the oil and gas industry gives us insight into the changing nature of water use, in this case in the South Platte and Colorado River Basins, and thus the assessment of possible impacts on these water systems is possible [*Nicot et al.*, 2014]. Because available surface and groundwater supplies in both basins are for the most part fully appropriated, water for oil and gas development is typically acquired from existing users and sources. These sources include municipalities, agricultural organizations and producers, tributary and nontributary groundwater, native produced water, and although rare, acquisitions through new appropriations during times of excess water [*Colorado Division of Water Resources et al.*, 2012; *Colorado Division of Water Resources*, 2014]. The major sources of water for gas development in Garfield County are the Colorado River and its tributaries. In Weld County, oil and gas companies are meeting their needs acquiring water from the South Platte River and groundwater aquifers. The water generated in South Platte River Basin is been augmented by diversions across the continental divide. One of these diversions is the Colorado Big Thompson Project, which has a multi-use designation and can be used by the oil and gas operators. Due to the fact that the state does not require from oil and gas operators to record and report the source of the water used, there is limited available information about water sourcing. Thus, a comprehensive study that would assign a value of total water quantity to each of the sources is not feasible. An alternative to understand the possible effects of this water demand to local scale is to combine the information about the general pathways of water acquisition in each county, with the total annual water used for oil and gas, and compare it with traditional uses.

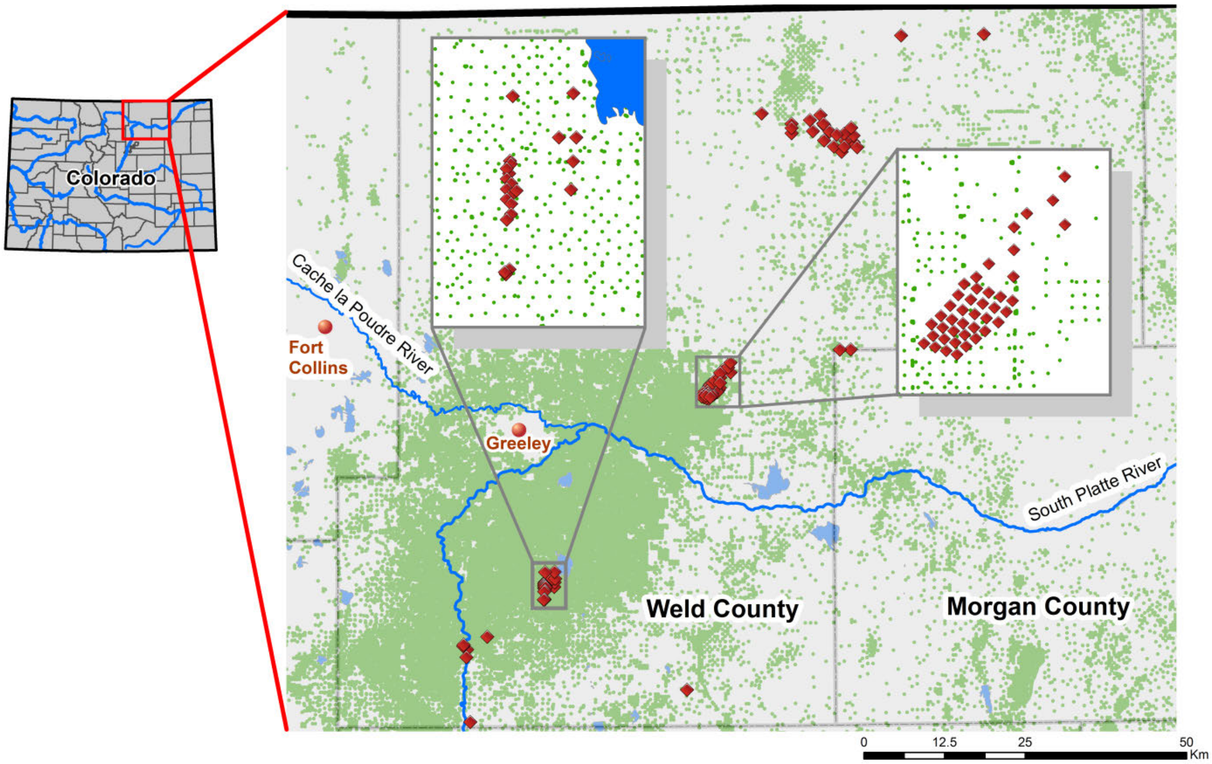


Figure 30: Nontributary groundwater wells used for oil and gas in Weld County (red diamonds; includes permitted with unknown status of completion, and permitted with constructed status) [Colorado Division of Water Resources, 2015].

In Weld County most companies prefer to lease needed water from municipalities, private entities such as water service companies, as well as irrigation and reservoir companies. There are few oil and gas companies who actually own water rights. Oil and gas companies in Weld County, due to the short-term nature of exploration and drilling, are implementing flexible strategies to ensure reliable water supplies for their operations. This includes acquiring water from nontributary freshwater aquifers. This helps reduce competition for fully-appropriated surface water and tributary groundwater sources. Nontributary groundwater has a limited hydrological connection to surface water and it is considered a viable source for oil and gas companies, because it is not administered within the priority system. Currently, data from the Colorado Division of Water Resources indicate that there are 103 records (77 permitted and constructed; 26 permitted with unknown status of completion) of nontributary groundwater wells related to oil and gas development in Weld County (Figure 30). The annual maximum average withdrawal for all 103 nontributary well permits could yield about $15x10^6$ m^3. If they were to be used exclusively for oil and gas development they could meet 2.5 times the current estimated water demand in the county. New well permits are issued for properties which have already secured a nontributary groundwater right, allowing the water rights holder to utilize the full permitted amount. In addition to industrial use, all of the wells have additional uses (e.g., irrigation, commercial, municipal).

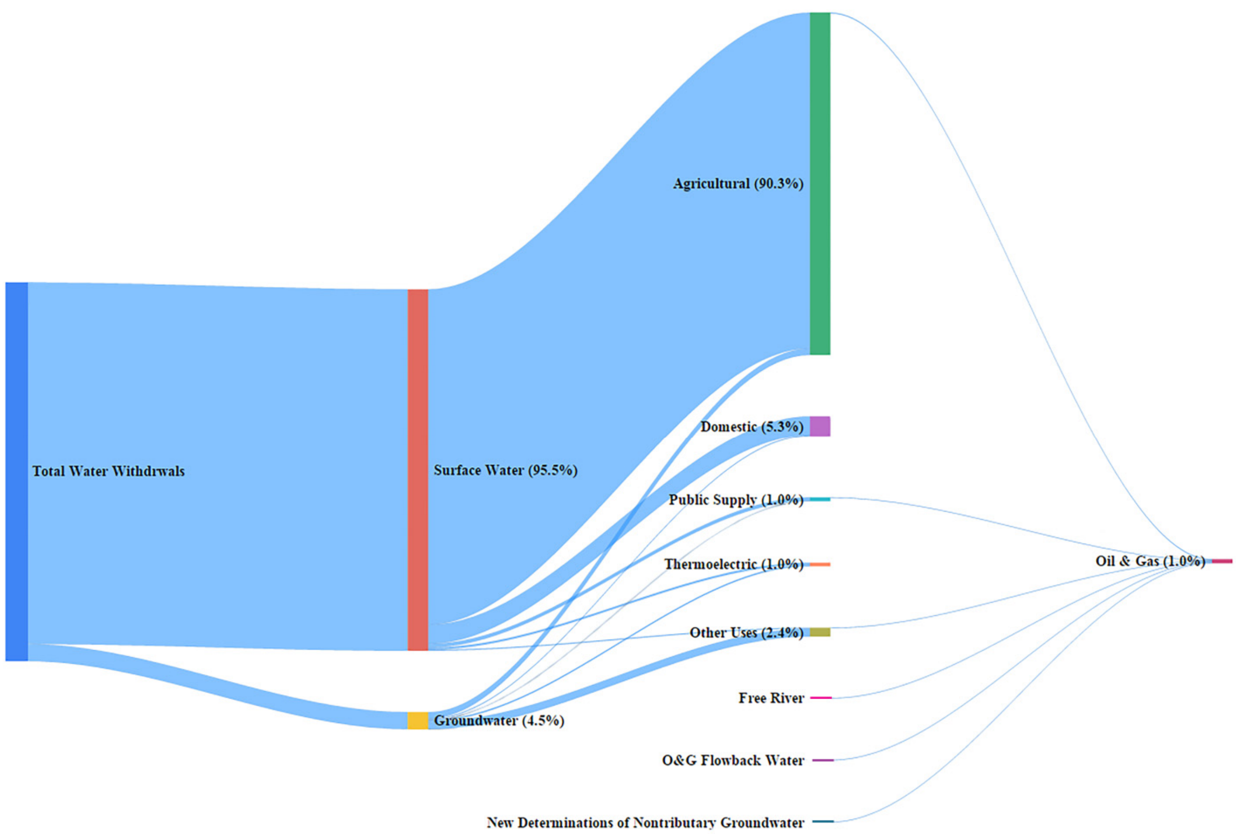


Figure 31: Estimated water withdrawals and their sources in Weld County, along with the sources of water acquisition for oil and gas. The annual average water used for oil and gas development during 2011-2013 is expressed as percentage of total withdrawals [Maupin et al., 2014; FracFocus Chemical Disclosure Registry, 2015].

Figure 31 is a graphical representation of Weld's County water withdrawal estimates based on Maupin et al. [2014] and expressed in percentages of total withdrawals. The last link of the Sankey diagram is portraying the pathways of water acquisition for oil and gas development. According to the FracFocus records for Weld County, the annual average water used for oil and gas for the period of 2011-2013 was 16,240 m^3/day, which, according to the county's total estimated water use [*Maupin et al.*, 2014], is approximately equal to 1.0% of total annual withdrawals. This small percentage of water is not necessarily all from sources within the county. The reason that the links of the last section in Figure 31 are of equal width is due to the fact that it is unknown how much water is transferred from each use.

The picture of water transfers to oil and gas operations in Garfield County is very different from Weld County described in the paragraph above. In general, Garfield County oil and gas companies obtained water rights during the 1950s through the 1970s. The relatively senior date of Garfield County oil and gas companies' water rights gives them some certainty of available water. Some companies lease reservoir units from local and federal water entities. The current water transfer practices scaled to a local level can reveal significant differences from other regions within Colorado. Garfield's County water withdrawal estimates based on Maupin et al. [2014], are expressed in percentages of total withdrawals in Figure 32 by using a Sankey diagram. The last level of connection was added in order to show the pathways of water acquisition for gas development identified and to link water use for oil and gas with other uses. According to the FracFocus records for Garfield County, the annual average water used for oil and gas for the period of 2011-2013 was 19,319 m^3/day, which, according to the county's total estimated water use [*Maupin et al.*, 2014], is approximately equal to 2.0% of total annual withdrawals. This small percentage of water is not necessarily all from sources within the county. Again, the reason that the links of the last section in Figure 32 are of equal width is due to the fact that it is unknown how much water is transferred from each use.

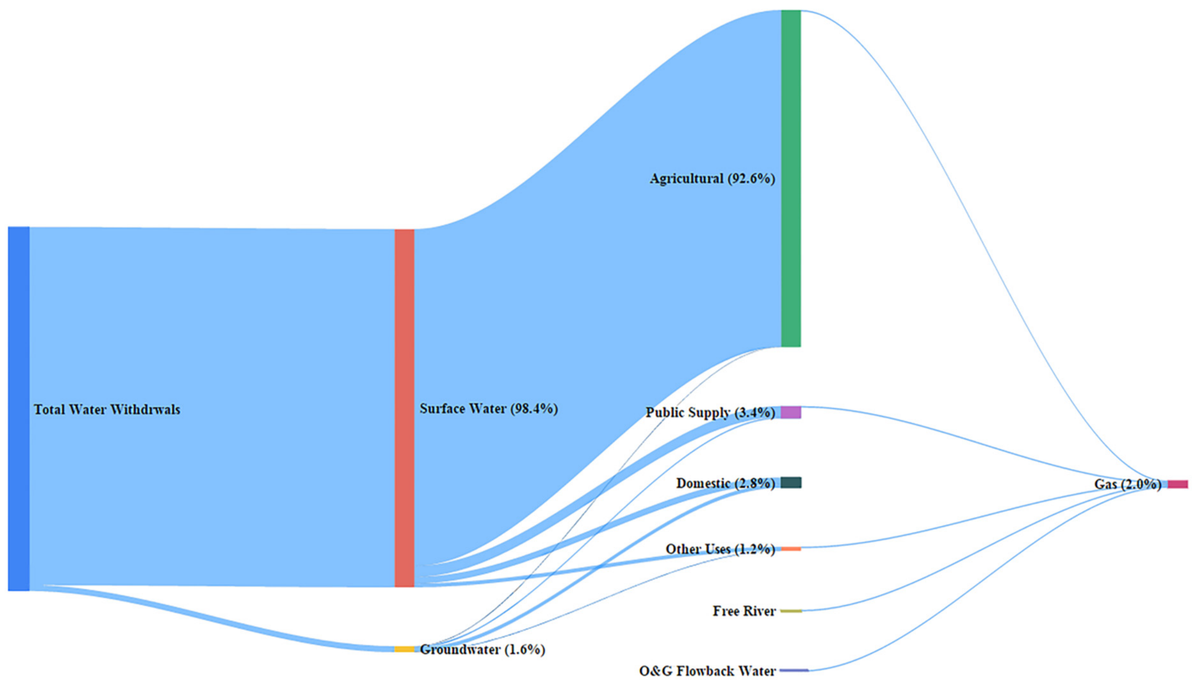


Figure 32: Estimated water withdrawals and their sources in Garfield County, along with the sources of water acquisition for oil and gas. The annual average water used for oil and gas development during 2011-2013 is expressed as percentage of total withdrawals [Maupin et al., 2014; FracFocus Chemical Disclosure Registry, 2015]

3.3.5 Oil and Gas Water Utilization under Extreme Development Scenario in HUC-12 Scale

Figure 33 portrays the estimated water demands for domestic, agricultural, thermoelectric and industrial uses for each HUC-12 that lays within or it is intersected by the Weld's County borders. Agriculture is the largest consumptive sector and apart from the comparisons about water intensity of each use at every HUC-12, it is also clear that in Weld's County southwestern part is experiencing competition for land use. Oil and gas most intense development is taking place in the same area of intense agricultural activity since the Wattenberg field is located at the southwestern part of the County.

According to the extreme water development scenario (Figure 34), water use for oil and gas in most of the HUCs is ranked as the second water user. On the other hand, if oil and gas water demand is presented as percent of the total water demand then in very few HUCs it surpasses 5%. The only exceptions are hydrologic units where agricultural demand is very small or zero, mainly in the north and northeast parts of the County. The lower part of Figure 34 is illustrating the acres per HUC-12 that could be planted with alfalfa or corn and meet net irrigation requirements for a normal water year by using the water used for oil and gas within each HUC-12. The total water amount used for all HUCs it is equivalent to 10,230 acres of alfalfa and 17,218 acres of corn. These numbers of course represent an extreme scenario with the purpose to investigate the tradeoff between water for food and water for unconventional energy. At the same time, it should be reiterated that the period of 2011-2014 is representing a booming period for the oil and gas industry. Under a more realistic scenario, assuming average development in every HUC the acreage becomes 4,237 and 7,131 for alfalfa and corn respectively. In contrast to the 2014 Colorado Agricultural Statistics, a close to normal water year in the South Platte basin, the planted acres for alfalfa were 82,000 and 116,900 for corn in Weld County [*USDA NASS Mountain Regional Office and Colorado State Department of Agriculture*, 2015]. The acreages of the average scenario are representing about 5% more acres of alfalfa or 6% more acres of corn. Despite the fact that it is a relatively small amount of water, it could be a valuable water supplement source during droughts for crops, if the industry moved to reusing and recycling its flowback and produced water.

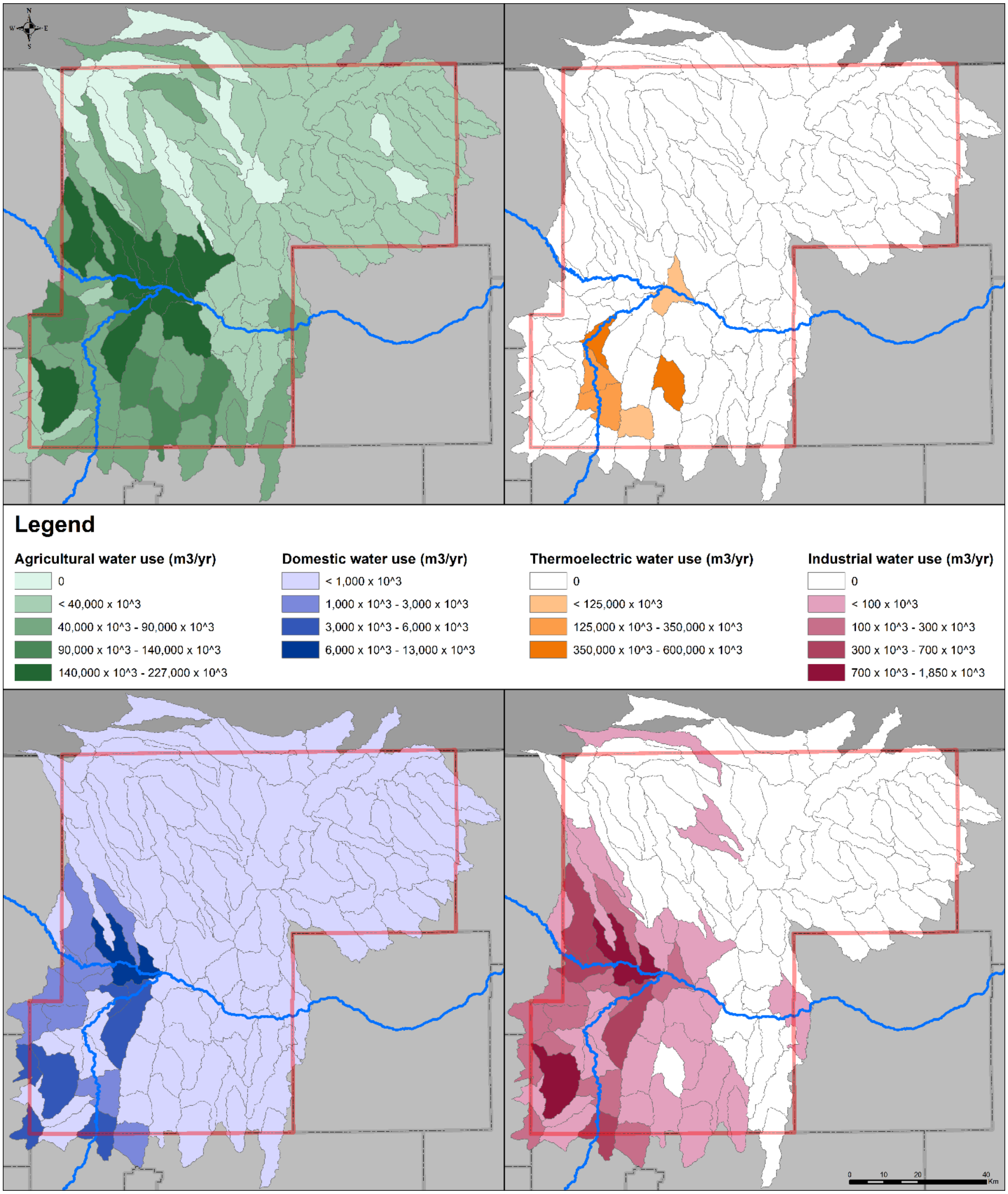


Figure 33: Agricultural, domestic, thermoelectric and industrial water demand per HUC-12 in Weld County, CO [Pickard et al., 2015]

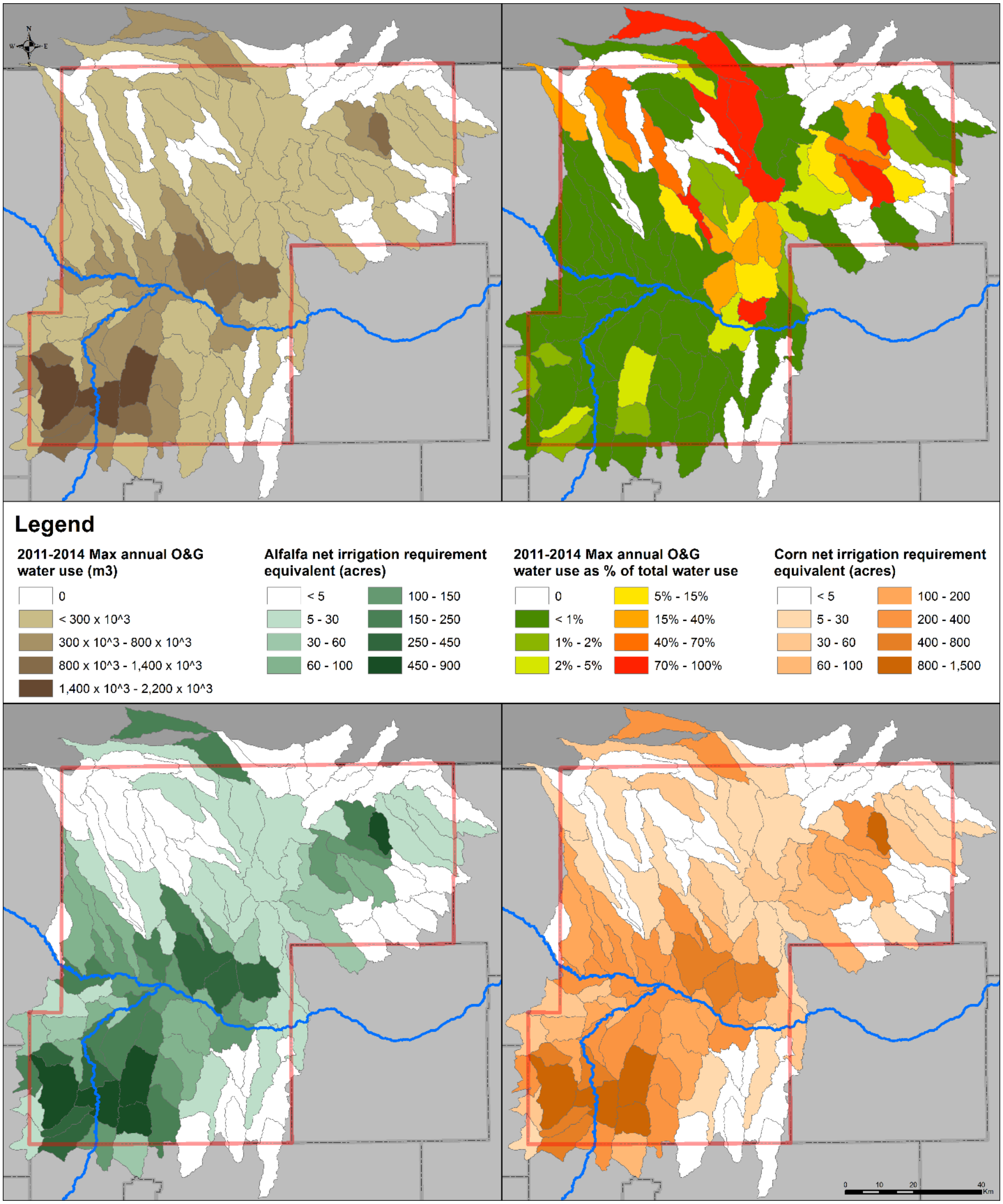


Figure 34: Oil and Gas Maximum Development Scenario in Weld County, CO. Water Use for 2011-2014 per HUC-12; Water Use as % of Total Water use per HUC-12; and the lower section the net irrigation requirement equivelent in acreage for alfalfa and corn

3.4 Conclusions

Colorado is a water-limited region with strict water administration and much attention is paid to the water volumes required for drilling and hydraulic fracturing of oil and gas wells since this water is considered fully consumed as it is rarely returned back to the system. Understanding an emerging water demand, its characteristics and the effects from possible water use shifts within a basin, is important information to communities/regions as they plan for their water future by making decisions about tradeoffs associated with oil and gas development in an integrated management manner. Especially so in water-stressed regions like Colorado, where the water-energy nexus is a topic of heated discussion with conflicting water users and uses. Uncertainties including market fluctuations; advancements in extraction and water recycling technologies; changes in industry regulations; local bans or moratoria; and the availability and price of water each year is making water planning even more challenging.

The current chapter focused on understanding the factors that affect water demand for unconventional oil and gas development. Publicly available water use datasets were analyzed and compared in order to identify well development patterns, normalized water use per well, and flowback water and its disposal practices that are taking place in Garfield and Weld counties. This effort reveals the different oil and gas development practices (types of wells drilled, general pathways of water acquisition, handling of flowback water) and water use intensity in Weld and Garfield counties in Colorado. Furthermore, in order to understand better possible effects of this water demand in different spatial scales, analysis was done at three different spatial resolutions. Such an analysis can offer a complete picture of the water use in oil and gas industry and could be utilized for water resources planning and management in the area by constructing scenarios for this sector. By studying two different formations in Colorado was able to show that the challenges are local and that planning and management of water resources should be tailored for the specific area of interest in order to achieve water sustainability.

More specifically, it was illustrated that in Garfield County there is a decline of new well activity and that the preferred type is primarily multi-stage directional wells. On the east side of the continental divide, in the Denver-Julesburg formation within Weld county, there is a totally different development pattern with a steady, more or less, new well starts and a shift towards multi-stage horizontal. The available data show an increasing length and fracturing stages in horizontal wells. The comparison of single-stage wells and multi-stage wells in Weld County revealed that the normalized per stage water consumption is lower for multi-stage wells.

There is also difference between the two counties regarding the flowback water and its handling. The reported amount of water returning to the surface as percentage of total fluids used is much greater in Garfield County than in Weld. In Garfield County, the industry has adopted a pro-recycle strategy for the flowback water so the water used for the hydraulic fracturing is mainly recycled water. Further research is required to determine if the low re-use practice in Weld County is driven by the limited flowback water quantity or there are water quality issues that dictate such practice. The chemical profile of the flowback water might be such that re-use of treated flowback water might be costlier than its disposal.

Data related to water leases and sales are limited and create challenges when drawing conclusions and comparing findings between the two counties. The main difference in the two counties in terms of water administration is the existence of non-tributary groundwater within Weld County. This gives the opportunity to the industry to have a secure source of water that it is not subject to prior appropriation doctrine. The identification of these wells in Weld County showed that they could potentially satisfy current water demand of the unconventional oil and gas industry. At the same time, another main dissimilarity is that in Garfield County the natural gas companies own the rights of the water they are using and they only lease if the demand is greater that their stock water. In Weld County, the arrangement for water acquisition is decentralized since most oil companies do not hold water rights and they mainly lease the needed water quantities for their operations. The county-level analysis that compared the average amount of water used for shale gas and oil exploration is relatively small if compared with other uses. In Garfield and Weld counties, this use is estimated to represent only 2% and 1% respectively, compared to the total water withdrawn. The analysis on water use for unconventional oil and gas development in HUC-12 level, for the Weld County, revealed that under even a high development scenario, in which the industry's water demand in each HUC-12 would be the highest of the recorded demand between 2011-2014, the total amount of water needed it is a small fraction of other uses.

A major finding of this study is the identification of data gaps and the additional data and metadata needed to be collected in Colorado for such analysis. Water use data for oil and gas development are self-reported by the operators. Different reporting approaches by operators cause inconsistencies and errors that might not always be addressed by the regulating agency unless brought to their attention. Most of the well completion reports used for this study had blank data fields resulting to valuable information to be lost. Some data records were reported multiple times, or in differing units than instructed, which made QA/QC process a cumbersome effort in order to correct or discard unreliable records before further analysis. Having in place an automated reporting system that would ensure compulsory information to be filled and have a filter for potential error entries, would help reduce the loss of information and data errors. Currently, in Colorado, there in not in place a policy that requires the reporting of water acquisition methods and sources, thus missing a significant part of the water life cycle and not allowing water's traceability. Another issue identified in the well completion reports is that flowback water volumes reported are not tied with the number of days flowback water is recovered. Since operators have a 30-day window to submit the well completion report to COGCC, it creates variation on the amount of the water reported. Instead, the reported flowback water should be the amount recovered in a fixed time period in order to have a meaningful unified metric of flowback recovery in every basin. Furthermore, volumes of water used in secondary activities such as dust suppression, drilling mud, and site restoration, although insignificant to the amount used for hydraulic fracturing, should be also included in the report. Finally, data reported to FracFocus combines total water volume per well (freshwater, produced water, or recycled water), which makes it impossible to determine separate volumes for each element.

Analysis of the decrees and lease agreements from the Colorado Division of Water Resources was a very time-intensive process and could be avoided with a different classification, in which separates oil and gas activities from other industrial uses. Moreover, data downloads provided no direct method to distinguish nontributary groundwater sources used for oil and gas operations from other industrial uses. Therefore, additional review of each permit and its supporting documents was required to determine if it was related to oil and gas activities.

Availability of reliable data for the entire water life cycle (source, drilling, fracturing, flowback/produced water recovery, and disposal/reuse) is critical for water resources planning and quantifying tradeoffs associated with oil and gas development. The use of an integrated, standard reporting system would allow for consistency across different scales and geographic regions improving our understanding about water use the sourcing strategies and their effects. This will help reduce conflicting perspectives, which shape the water-energy discussion and allow for water resources planning and management that is environmental and socially acceptable.

References

Burton, T. G., H. S. Rifai, Z. L. Hildenbrand, D. D. Carlton Jr, B. E. Fontenot, and K. A. Schug (2016), Elucidating Hydraulic Fracturing Impacts on Groundwater Quality using a Regional Geospatial Statistical Modeling Approach, *Science of The Total Environment*, *545–546*, 114–126, doi:10.1016/j.scitotenv.2015.12.084.

Chen, H., and K. E. Carter (2016), Water Usage for Natural Gas Production through Hydraulic Fracturing in the United States from 2008 to 2014, *Journal of Environmental Management*, *170*, 152–159, doi:10.1016/j.jenvman.2016.01.023.

Clark, C. E., A. J. Burnham, C. B. Harto, and R. M. Horner (2012), The Technology and Policy of Hydraulic Fracturing and Potential Environmental Impacts of Shale Gas Development, *Environmental Practice*, *14*(04), 249–261, doi:10.1017/S1466046612000415.

Clark, C. E., R. M. Horner, and C. B. Harto (2013), Life Cycle Water Consumption for Shale Gas and Conventional Natural Gas, *Environmental Science & Technology*, *47*(20), 11829–11836, doi:10.1021/es4013855.

COGCC (2015), *Annual Well Starts by County in Oil & Gas Staff Report (September 14, 2015)*, Oil & Gas Staff Report, Colorado Oil and Gas Conservation Commission, Denver, CO.

Colborn, T., K. Schultz, L. Herrick, and C. Kwiatkowski (2014), An Exploratory Study of Air Quality Near Natural Gas Operations, *Human and Ecological Risk Assessment: An International Journal*, *20*(1), 86–105, doi:10.1080/10807039.2012.749447.

Colorado Division of Water Resources (2014), *Sources of Water for Oil and Gas Well Construction in the Denver-Julesburg Basin*, Memo from Colorado Division of Water Resources Deputy Engineer.

Colorado Division of Water Resources (2015), HydroBase Online Tools, Available from: http://water.state.co.us/DataMaps/DataSearch/Pages/DataSearch.aspx#onlinedata (Accessed 10 April 2015)

Colorado Division of Water Resources, Colorado Water Conservation Board, and Colorado Oil and Gas Conservation Commission (2012), *Water Sources and Demand for the Hydraulic Fracturing of Oil and Gas Wells in Colorado from 2010 through 2015*.

Colorado Oil and Gas Conservation Commission (2015), COGCC Data, Available from: http://cogcc.state.co.us/data2.html#/downloads (Accessed 10 August 2015)

Colorado Supreme Court (2009), *Vance Jr., v. Wolfe, Colorado State Engineer*.

Considine, T. J., R. W. Watson, and N. B. Considine (2011), *The Economic Opportunities of Shale Energy Development*, Energy Policy & the Environment Report No. 9, Center for Energy Policy and the Environment at the Manhattan Institute.

FracFocus Chemical Disclosure Registry (2015), FracFocus 2.0, Available from: http://fracfocus.org/ (Accessed 1 August 2015)

Gallegos, T. J., B. A. Varela, S. S. Haines, and M. A. Engle (2015), Hydraulic Fracturing Water Use Variability in the United States and Potential Environmental Implications, *Water Resour. Res.*, n/a-n/a, doi:10.1002/2015WR017278.

Gleick, P. H. (1994), Water and Energy, *Annual Review of Energy and the Environment*, *19*(1), 267–299, doi:10.1146/annurev.eg.19.110194.001411.

Goodman, P. S., F. Galatioto, N. Thorpe, A. K. Namdeo, R. J. Davies, and R. N. Bird (2016), Investigating the Traffic-related Environmental Impacts of Hydraulic-Fracturing (fracking) Operations, *Environment International*, *89–90*, 248–260, doi:10.1016/j.envint.2016.02.002.

Goodwin, S., K. Carlson, B. Bai, L. Rein, K. Knox, and C. Douglas (2013), Improved Water Use Estimates for Drilling and Hydraulic Fracturing in Northeastern Colorado, *Journal of Water Resources and Protection*, *5*, 1262–1267, doi:10.4236/jwarp.2013.512135.

Grubert, E. A., F. C. Beach, and M. E. Webber (2012), Can Switching Fuels Save Water? A life Cycle Quantification of Freshwater Consumption for Texas Coal- and Natural Gas-Fired Electricity, *Environmental Research Letters*, *7*(4), 045801, doi:10.1088/1748-9326/7/4/045801.

Higginbotham, A., A. Pellillo, T. Gurley-Calvez, and T. S. Witt (2010), *The Economic Impact of the Natural Gas Industry and the Marcellus Shale Development in West Virginia in 2009*, College of Business and Economics, West Virginia University.

Hubert, M., and E. Vandervieren (2008), An Adjusted Boxplot for Skewed Distributions, *Computational Statistics & Data Analysis*, *52*(12), 5186–5201, doi:10.1016/j.csda.2007.11.008.

Kondash, A., and A. Vengosh (2015), Water Footprint of Hydraulic Fracturing, *Environmental Science & Technology Letters*, *2*(10), 276–280, doi:10.1021/acs.estlett.5b00211.

Laurenzi, I. J., and G. R. Jersey (2013), Life Cycle Greenhouse Gas Emissions and Freshwater Consumption of Marcellus Shale Gas, *Environ. Sci. Technol.*, *47*(9), 4896–4903, doi:10.1021/es305162w.

Malone, S., M. Kelso, T. Auch, K. Edelstein, K. Ferrar, and K. Jalbert (2015), Data Inconsistencies from States with Unconventional Oil and Gas Activity, *Journal of Environmental Science and Health*, *50*, 501–510, doi:10.1080/10934529.2015.992678.

Maupin, M. A., J. F. Kenny, S. S. Hutson, J. K. Lovelace, N. L. Barber, and K. S. Linsey (2014), Estimated Use of Water in the United States in 2010, *U.S. Geological Survey Circular 1405*, p. *56*, doi:http://dx.doi.org/10.3133/cir1405.

McKenzie, L. M., R. Z. Witter, L. S. Newman, and J. L. Adgate (2012), Human Health Risk Assessment of Air Emissions from Development of Unconventional Natural Gas Resources, *Science of The Total Environment*, *424*, 79–87, doi:10.1016/j.scitotenv.2012.02.018.

McLaughlin, M. C., T. Borch, and J. Blotevogel (2016), Spills of Hydraulic Fracturing Chemicals on Agricultural Topsoil: Biodegradation, Sorption, and Co-contaminant Interactions, *Environ. Sci. Technol.*, doi:10.1021/acs.est.6b00240.

Mehany, M. S. H. M., and A. Guggemos (2015), A Literature Survey of the Fracking Economic and Environmental Implications in the United States, *Procedia Engineering*, *118*, 169–176, doi:10.1016/j.proeng.2015.08.415.

Murray, K. E. (2013), State-Scale Perspective on Water Use and Production Associated with Oil and Gas Operations, Oklahoma, U.S., *Environ. Sci. Technol.*, *47*(9), 4918–4925, doi:10.1021/es4000593.

Nicot, J. P., and B. R. Scanlon (2012), Water Use for Shale-Gas Production in Texas, U.S., *Environmental Science & Technology*, *46*(6), 3580−3586, doi:10.1021/es204602t.

Nicot, J. P., B. R. Scanlon, R. C. Reedy, and R. A. Costley (2014), Source and Fate of Hydraulic Fracturing Water in the Barnett Shale: A Historical Perspective, *Environmental Science & Technology*, *48*(4), 2464–2471, doi:10.1021/es404050r.

Orford, A. (2013), Hydraulic Fracturing & Water: Considering the Water Resources Impacts of Hydraulic Fracturing, *The Water Report*, (110), 1–8.

Osborn, S. G., A. Vengosh, N. R. Warner, and R. B. Jackson (2011), Methane Contamination of Drinking Water accompanying Gas-Well Drilling and Hydraulic Fracturing, *PNAS*, *108*(20), 8172–8176, doi:10.1073/pnas.1100682108.

Pacsi, A. P., N. S. Alhajeri, D. Zavala-Araiza, M. D. Webster, and D. T. Allen (2013), Regional Air Quality Impacts of Increased Natural Gas Production and Use in Texas, *Environ. Sci. Technol.*, *47*(7), 3521–3527, doi:10.1021/es3044714.

Pacsi, A. P., K. T. Sanders, M. E. Webber, and D. T. Allen (2014), Spatial and Temporal Impacts on Water Consumption in Texas from Shale Gas Development and Use, *ACS Sustainable Chem. Eng.*, *2*(8), 2028–2035, doi:10.1021/sc500236g.

Pickard, B. R., J. Daniel, M. Mehaffey, L. E. Jackson, and A. Neale (2015), EnviroAtlas: A New Geospatial Tool to Foster Ecosystem Services Science and Resource Management, *Ecosystem Services*, *14*, 45–55, doi:10.1016/j.ecoser.2015.04.005.

Rahm, B. G., and S. J. Riha (2012), Toward Strategic Management of Shale Gas Development: Regional, Collective Impacts on Water Resources, *Environmental Science & Policy*, *17*, 12–23, doi:10.1016/j.envsci.2011.12.004.

U.S. Department of Agriculture Soil Conservation Service (1988), *Colorado Irrigation Guide: CO210–VI–COIG*.

U.S. Energy Information Administration (2015a), *Natural Gas Gross Withdrawals and Production. Volume in million cubic feet. Release date December 31, 2015*.

U.S. Energy Information Administration (2015b), *Petroleum and other Liquids. Volume in Annual-Thousand Barrels. Crude Oil Production. Release Date: December 31, 2015*.

U.S. Environmental Protection Agency (2012), *Study of the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources: Progress Report*, U.S. EPA Office of Research and Development, Washington, D.C.

U.S. Environmental Protection Agency (2013), Summary of FracFocus 1.0 Hydraulic Fracturing Data, *EPA Analysis of FracFocus 1 Data*. Available from: http://www2.epa.gov/hfstudy/epa-analysis-fracfocus-1-data (Accessed 13 February 2015)

USDA NASS Mountain Regional Office, and Colorado State Department of Agriculture (2015), *Colorado Agricultural Statistics 2015*, Lakewood, Colorado.

Vengosh, A., R. B. Jackson, N. Warner, T. H. Darrah, and A. Kondash (2014), A Critical Review of the Risks to Water Resources from Unconventional Shale Gas Development and Hydraulic Fracturing in the United States, *Environmental Science & Technology*, *48*(15), 8334–8348, doi:10.1021/es40511.

Walton, J., and A. Woocay (2013), Environmental Issues Related to Enhanced Production of Natural Gas by Hydraulic Fracturing, *Journal of Green Building*, *8*(1), 62–71, doi:10.3992/jgb.8.1.62.

4 A Framework for Bridging Data Gaps in Groundwater Level Measurements

4.1 Introduction

Hydrological data are an integral part of informed water resources planning and management for meeting the growing water demands under sustainable conjunctive use of surface and groundwater resources. Environmental datasets are constantly increasing both in spatial and temporal resolution due to technological innovations and advancements that has spurred data collection and reduced collection cost. In the US, there are several national water resources related data repositories from federal organizations (e.g. USGS, USDA, NOAA and EPA). But even in data-rich countries like the US, there are cases where available times series could have missing observations (intermittent or continuous) or be of short span or both. In general, the sources of data fragmentation could be classified into technological (e.g. equipment failure, time off for service), institutional (e.g. not an institutional priority, inadequate funding and/or personnel resources), anthropogenic (e.g. errors in measuring, handling or storing data) and natural (e.g. floods, hurricanes, landslides) causes [*Salas*, 1993].

Long and continuous water resources records are required for hydrologic analysis, as calibration/validation points in physically-based models for understanding and managing the simulated water system and also for forecasting purposes. In many cases though, hydrologists are challenged to work with short and fragmented geophysical time series. Thus, scientific methods were developed to fill in or extend these time series. In relevant literature, different approaches have been suggested to fill in or extend geophysical time-series, such as weighted average methods

[*Pappas et al.*, 2014], regression methods [*Salas et al.*, 1980; *Salas*, 1993; *Makhuvha et al.*, 1997; *Schneider*, 2001; *Koutsoyiannis and Langousis*, 2011], empirical orthogonal functions [*Alvera-Azcárate et al.*, 2007], time series statistical models [*Salas et al.*, 1980; *Salas*, 1993; *Hipel and McLeod*, 1994], interpolation methods [*Ashraf et al.*, 1997; *Teegavarapu and Chandramouli*, 2005; *Teegavarapu*, 2007, 2012], artificial neural networks [*Kuligowski and Barros*, 1998; *Elshorbagy et al.*, 2002], spectral analysis [*Kondrashov and Ghil*, 2006] and others. Filling data gaps specifically in groundwater head time series turns out to be a practically impossible task if the measurements have low frequency and long gaps, due to the intrinsic complexity of geology, subsurface flows, land cover, basin hydrology, and heterogeneity of aquifer parameters, along with weak monitoring initiatives and the anthropogenic interventions (i.e. groundwater pumping, water diversions, recharge structures).

The original Kalman Filter (KF) [*Kalman*, 1960; *Kalman and Bucy*, 1961] and its many variants, where pitfalls of the standard method are overcome and non-linear formulation is possible, have been extensively employed in several environmental fields [*Dee*, 1991; *Walker and Houser*, 2001; *Keppenne and Rienecker*, 2002; *Houtekamer et al.*, 2005; *Aanonsen et al.*, 2009]. There are also many examples in the pertinent literature of the application of the KF technique in hydrology [*Szollosi-Nagy*, 1976; *Li et al.*, 1978; *Rodriguez-Iturbe et al.*, 1978; *Kitanidis and Bras*, 1979, 1980; *Schreider et al.*, 2001; *Zhou et al.*, 2006; *Clark et al.*, 2008; *Xie and Zhang*, 2010; *Wu et al.*, 2013; *Lei et al.*, 2014], hydraulics [*Chiu and Isu*, 1978; *Li et al.*, 1978; *Simons et al.*, 1978; *Sen et al.*, 2004; *Ye and Fenner*, 2011], water quality [*Schrader and Moore*, 1977; *Bowels and Grenney*, 1978; *Pastres et al.*, 2003; *Lee et al.*, 2009; *Javaheri and Babbar-Sebens*, 2014] and water resources systems [*Chan and Loucks*, 1978; *Duong et al.*, 1978; *Okeya et al.*, 2014; *Jung and Lansey*, 2015; *Jung et al.*, 2016].

Early applications [*Kitanidis*, 1976; *McLaughlin*, 1976, 1978; *Wilson et al.*, 1978] examined the opportunities of the KF method in groundwater, despite the computational and methodological constraints present at that time. In groundwater numerical modeling it has been an attractive alternative for parameter estimation in comparison to classical optimization methods. Different KF have been used for solving the inverse problem, in other words, characterizing spatial heterogeneous model parameters (i.e. hydraulic conductivity K, contaminant decay rate, etc.) by assimilating observed data (i.e. piezometric head, contaminant concentration) under different schemes [*Eppstein and Dougherty*, 1996; *Ferraresi et al.*, 1996; *Cahill et al.*, 1999; *Sun et al.*, 2009; *Bailey and Baù*, 2010, 2012; *Zhou et al.*, 2011; *Bailey et al.*, 2012; *Li et al.*, 2012; *Alzraiee et al.*, 2014]. Crestani et al. [2013] compared the ability of the Ensemble Kalman Filter and the Ensemble Smoother (ES) to predict the spatial distribution of hydraulic conductivity in a groundwater flow and transport modeling scheme, showing that EnKF outperforms the smoother. Optimal design of groundwater monitoring networks is a common practical problem where prior information of the hydrogeological parameters is crucial. Studies in this domain employ the KF in different sampling frameworks with noteworthy improvements with main goals to minimize the prediction errors and the cost associated with the number of observation wells [*Herrera and Pinder*, 2005; *Zhang et al.*, 2005; *Kollat et al.*, 2011; *Alzraiee et al.*, 2013].

In the Kalman filtering literature there are very few articles that deal with the problem of filling the gaps in missing data. Bennis et al. [1997] used a multivariable ARIMA model with reference stations and a single-variable linear interpolation, for filling streamflow data, with a Kalman Filter operating both forward and backward in time, succeeding to have better prediction only at the first missing data point and the peak flow, compared with ordinary least squares. In a similar way, in the sense that the Kalman Filter calculates the model parameters and then the smoother is updating the estimates, Alavi [2006] used a regression model of available energy and vapor pressure deficit in order to fill gaps in latent heat flux data, with successful results. Gove and Hollinger [2006] and Jarvis et al. [2004] used Kalman Filters to fill missing data to eddy covariance net carbon fluxes time series. More specifically, Bierkens et al. [1999] calibrated a physically based daily ARIMAX model by embedding it to a simple KF combined with a maximum likelihood criterion. The ARIMAX process used bi-monthly time series of precipitation surplus as the exogenous variable and simulated daily groundwater head levels with the purpose of filling in groundwater head data gaps.

The Autoregressive Integrated Moving Average (ARIMA) models [*Box and Pierce*, 1970] is one of the most common tools for environmental time series analysis and prediction [*Hipel and McLeod*, 1994]. Early on, Salas et al. [1980] classified and reviewed different ARIMA models for hydrologic time series. ARIMA models have been used for analysis and forecasting in a wide range of water studies including floods, drought, water management, water quality [*Irvine and Eberhardt*, 1992; *Toth et al.*, 1999; *Ahmad et al.*, 2001; *Papamichail and Georgiou*, 2001; *Sun and Koch*, 2001; *Durdu*, 2010; *Kim et al.*, 2011; *Wang et al.*, 2014, 2015; *Karavitis et al.*, 2015]. Modeling and forecasting of groundwater levels through ARIMA models with model variations, including seasonality and/or stationarity, is a very common practice [*Changnon et al.*, 1988; *Von Asmuth et al.*, 2008; *von Asmuth et al.*, 2012; *Mirzavand and Ghazavi*, 2014]. *Ahn and Salas* [1997] applied ARIMA models for determining a uniform sampling time in order to represent the groundwater head fluctuations. A different approach for groundwater level forecasting is the use of Artificial Neural Networks [*Daliakopoulos et al.*, 2005; *Uddameri*, 2006; *Shirmohammadi et al.*, 2012; *Chang et al.*, 2016].

Moreover, scholars have proposed methods that combine stochastic time series models of groundwater head with the KF, however, they have focused mainly on inverse applications. Knotters and Bierkens [2000] used the methodology in Bierkens et al. [1999] in order to predict the effect of interventions on groundwater dynamics on a daily time step. The use of KF for parameter estimation for transfer function-noise was the practice of Berendrecht et al. [2003] and Yi et al. [2004]. Bierkens et al. [2001] combined the same framework with auxiliary information (e.g. Digital Elevation Model, soil maps etc.) for calibrating a daily regional ARX model with the groundwater estimated states, which is to be used in applications of optimal prediction, network optimization and conditional simulation. At an earlier study, an empirical KF algorithm incorporating a regional SARIMA model constructed based on monthly groundwater level data from 21 wells, was evaluated for state estimation [*Graham and Tankersley*, 1993]. A scheme for regional spatiotemporal groundwater head prediction that a simple KF and a global optimization algorithm were used for calibrating, at known points, transfer function-noise models with excess precipitation as exogenous input and afterwards the regionalization of the model parameters for ungauged locations via clustering was evaluated by Yuan et al. [2008].

Regional physically based groundwater models are becoming more representative of the aquifer systems due to computational prowess for finer scale representation and improved data availability, offering critical insight to the overall behavior. Despite that, most of the time, their use is limited to the original premises they are built for, resulting in valuable information remaining unutilized. In the current effort, a new methodology for missing groundwater head data estimation is presented. It utilizes information from regional physically based groundwater flow models to build a stochastic linear model, embedding it into an Ensemble Smoother (ES), which uses the measured groundwater head data for its updating scheme. The performed literature review revealed that there is not a similar framework, nor has the Ensemble Smoother (ES) been evaluated for this task. The suggested methodology is an innovative, easy to implement and computationally low cost approach for filling gaps in groundwater head time series. At the same time, it augments the value of readily available regional groundwater models; hence its transferability to any modeled alluvial type aquifer is guaranteed since no aquifer specific parameters are required.

The rest of this chapter consists of the methodology section, which explains the theory of the Ensemble Smoother (ES) and the seasonal ARIMAX model and shows in detail how these two techniques are combined in order to yield the proposed framework of bridging groundwater level data gaps. The next section is divided in two parts: The first is describing the South Platte River Basin and its alluvial aquifer in Colorado, which serves as the area of application, the regional groundwater model employed and the groundwater head measurement data used. As for the second part, the three numerical experiments evaluating the efficacy of the proposed methodology along with the selected well locations are presented. The forth section of this paper includes the results of the numerical experiments and the discussion regarding their performance. The conclusions of this work consist of a brief summary of the proposed framework, underlining its advantages and challenges, and proposing the needed future work that could investigate the effect of different modifications for increasing overall performance.

4.2 Methods

This section gives a brief introduction of the theory of ARIMA and Kalman Filtering methodologies and also introduces the proposed framework for bridging gaps in groundwater head time series.

4.2.1 ARIMA models

The ARIMA models are describing how the present observation of a time series $\{y_t\}$, from a stationary or a non-stationary process, can be determined based on a linear relationship between previous observations ($y_{t-1},\ y_{t-2}, \dots, y_{t-p}$), the current white noise term (ε_t) and ones of previous time steps ($\varepsilon_{t-1},\ \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$) [*Box and Jenkins*, 1976]. The stationary case of an ARIMA model is commonly noted as or ARMA(p, q), and its mathematical formulation is:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \tag{5}$$

where, φ and θ are the autoregressive and moving average coefficients respectively, while p is the order of the autoregressive process, and, q is the order of the moving average process.

Frequently, time series demonstrate seasonal variations, thus the ARIMA model includes additional parameters to account for these linear cyclical relationships. The general notation of such a model is written as ARIMA$(p, d, q)(P, D, Q)s$, where d symbolizes the order of the integrative part (order of differencing), s, the seasonality of the time series, while P, D, Q denote the order of the seasonal autocorrelation, integration, and the moving average processes respectively. Employing the backshift operator B, which is defined as $By_t = y_{t-1}$ and $B^k y_t = y_{t-k}$, $\kappa \epsilon\ |\mathbb{Z}|$, we can define the non-seasonal autoregressive and moving average operators as:

$$\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \tag{6}$$

$$\theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \tag{7}$$

and similarly, the seasonal autoregressive and moving average operators are defined as:

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps} \tag{8}$$

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \dots + \Theta_Q B^{Qs} \tag{9}$$

Time series of dynamic processes exhibiting non-stationarity can be modeled as stationary via differential transformation. In most environmental time series, there is no need of higher order differencing ($d > 2$). The first order differences could be notated as:

$$\nabla y_t = (1 - B)\, y_t = y_t - y_{t-1} \tag{10}$$

In case second order differencing ($d = 2$) is required, then the second order difference is:

$$\begin{aligned} \nabla^2 y_t = (1 - B)^2\, y_t = (1 + B^2 - 2B)\, y_t &= y_t + B^2 y_t - 2B y_t \\ &= y_t + B^2 y_t - 2B y_t = y_t + y_{t-2} - 2y_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = \nabla y_t - \nabla y_{t-1} \end{aligned} \tag{11}$$

Equations 2 - 7 are used to express the general notation of the Seasonal ARIMA (SARIMA) process, which takes the following form:

$$\varphi_p(B)\Phi_P(B^s)\nabla^d \nabla_s^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{12}$$

If exogenous variables $\{x_{tm}\}$ are influencing the stochastic process, in that case, the model is abbreviated often as SARIMAX and then the equation 8 takes the following form:

$$\varphi_p(B)\Phi_P(B^s)\nabla^d \nabla_s^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t + \alpha_m x_{tm} \tag{13}$$

4.2.2 Kalman Filtering Techniques

The Kalman Filter [*Kalman*, 1960] is a sequential filtering technique, which is applied to linear systems whose residuals are following the Gaussian distribution [*Meinhold and Singpurwalla*, 1983]. To break free from the linearity constraints and the difficulty to estimate objectively the error statistics necessary to apply the Kalan Filter, variants of the classical method were developed. One of its advantages is that it combines model and ground truth information of a system into a framework that takes into account uncertainty, both in the model and measurements [*Eigbe et al.*, 1998]. The Extended Kalman Filter (EKF) (see [*Gelb*, 1974]) is suitable for small systems with nonlinear dynamics but there are pitfalls with the linearized transformation [*Julier and Uhlmann*, 2004] and also when vigorous instabilities and strong nonlinearities are present [*Miller et al.*, 1994]. An alternative to the EKF is the Unscented Kalman Filter [*Julier and Uhlmann*, 1997; *Wan and Van der Merwe*, 2001], which does not linearize the models, but instead, uses the actual nonlinear function of the system and propagating a set of points (sigma points) in time.

Another alternative suitable for nonlinear systems is the ensemble Kalman filter (EnKF) [*Evensen*, 1994], which has gained popularity due to its simple formulation, ease of application and also to its minimal computational requirements. It combines a Monte Carlo method for estimating the error statistics and the Kalman filter [*Evensen*, 1994]. Evensen [1997] presented the EnKF successful performance on a chaotic and highly nonlinear system of the Lorenz equations, and there are numerous applications proving its applicability to real case studies [*Evensen*, 2003]. An extension to the EnKF is the Ensemble Kalman Smoother (EnKS) [*Evensen and van Leeuwen*, 2000], which can also be applied in highly nonlinear systems in a sequential way. The main variation from the EnKF is that it updates the state estimates based on the future observations. In general, the filter algorithms are more suitable when forecasting is the primary purpose; while smoothers are more appropriate for process studies [*van Leeuwen and Evensen*, 1996]. On a similar scheme, the Ensemble Smoother (ES) [*van Leeuwen and Evensen*, 1996] is taking advantage of the utilization of the Monte-Carlo technique to represent model's probability densities. The

Ensemble Smoother (ES) does not run in a sequential mode like EnKS and thus, it is suitable for "off-line" applications [*Baù et al.*, 2015]. The main disadvantage of Ensemble Smoother (ES) is its suboptimal performance for nonlinear systems [*van Leeuwen and Evensen*, 1996; *Grønnevik and Evensen*, 2001; *Crestani et al.*, 2013]. In the next subsection, the Ensemble Smoother (ES) will be discussed with more detail.

4.2.3 Ensemble Smoother

The Ensemble Smoother (ES), introduced by *van Leeuwen and Evensen* [1996], is a Monte Carlo-based Kalman method that consists of a forecasting-updating scheme, where the system state is updated in a way that it minimizes the updated state's variance when there are known measurements of the state. The ensemble framework is used for the estimation of the error statistics. Let's define $\boldsymbol{A}^f$ as the ensemble matrix, which is holding a finite set of ensemble members. The ensemble members $\underline{\psi}^i$ are vectors of the forecasted model states, with length n:

$$\boldsymbol{A}^f = \left(\underline{\psi}^1, \underline{\psi}^2, \dots, \underline{\psi}^N\right) \quad \in \Re^{n \times N} \tag{14}$$

The forecast error covariance matrix $\boldsymbol{P}_e^f$ is defined as

$$\boldsymbol{P}_e^f = \frac{(\boldsymbol{A}^f - \overline{\boldsymbol{A}}^\tau)(\boldsymbol{A}^f - \overline{\boldsymbol{A}}^\tau)^T}{N-1} \quad \in \Re^{n \times n} \tag{15}$$

With the assumption that true mean of $\overline{\boldsymbol{A}}^\tau$ is equal to $\overline{\boldsymbol{A}}^f$ for large enough ensemble members the $\boldsymbol{P}_e^f$ can be computed according to the following equation

$$\boldsymbol{P}_e^f = \frac{(\boldsymbol{A}^f - \overline{\boldsymbol{A}}^f)(\boldsymbol{A}^f - \overline{\boldsymbol{A}}^f)^T}{N-1} = \frac{\boldsymbol{A}'^f \boldsymbol{A}'^{f^T}}{N-1} \quad \in \Re^{n\times n} \tag{16}$$

During the update step of the Ensemble Smoother (ES) the matrix $\boldsymbol{A}^f$ is being corrected using m state measurements. The Kalman framework is allowing for incorporating uncertainty not only at the model but also in the measurements. In this regard, the observation vector $\underline{d}^j$, *where j=1,2,…,N*, can be expressed by two parts, the true value vector $\underline{d}$ plus a random noise error vector $\underline{\varepsilon}$,

$$\underline{d}^j = \underline{d} + \underline{\varepsilon}^j \quad \in \Re^m \tag{17}$$

Similar to the forecast ensemble matrix, the measurement vectors can be stored in a matrix

$$\boldsymbol{D} = \left(\underline{d}^1, \underline{d}^2, \dots, \underline{d}^N\right) \quad \in \Re^{m\times N} \tag{18}$$

and the ensemble of error vectors in the $\boldsymbol{\mathcal{E}}$ matrix

$$\boldsymbol{\mathcal{E}} = \left(\underline{\varepsilon}^1, \underline{\varepsilon}^2, \dots, \underline{\varepsilon}^N\right) \quad \in \Re^{m\times N} \tag{19}$$

The error vector $\underline{\varepsilon}$ is assumed to be a random process following a Gaussian distribution with mean equal to zero, thus, the covariance matrix $\boldsymbol{R}_e$ is equal to

$$\boldsymbol{R}_e = \frac{\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{E}}^T}{N-1} \quad \in \Re^{m\times m} \tag{20}$$

The following equation is the update step of the Ensemble Smoother (ES) scheme and it allows to update the estimate of the state of the system, when observations are available.

$$\boldsymbol{A}^u = \boldsymbol{A}^f + \boldsymbol{K}(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{A}^f) \quad \in \Re^{n\times N} \tag{21}$$

Where, $\boldsymbol{H} \in \Re^{m\times N}$ is a binary matrix that maps the locations of the measurements in such way that the rows symbolize the number of the measurement and the columns the time step that was observed. More specifically, if the second measurement is representing the state of the system for $t = 10$, then its location in the $\boldsymbol{H}$ matrix will be $H_{2,10}$.

$$\boldsymbol{K} = \boldsymbol{P}_e^f \boldsymbol{H}^T \left(\boldsymbol{H} \boldsymbol{P}_e^f \boldsymbol{H}^T + \boldsymbol{R}_e \right)^{-1} \quad \in \Re^{n \times m} \tag{22}$$

Lastly, $\boldsymbol{K}$ is the Kalman gain matrix, where the formulation is responsible for minimizing the variance error, while yielding the best linear unbiased estimator. Additionally, the Kalman gain matrix is incorporating both measurement and model covariance matrices $\boldsymbol{R}_e$ and $\boldsymbol{P}_e^f$ into its calculation, signifying that the correction applied to $\boldsymbol{A}^f$ is related to the entrenched uncertainties in these two matrices [*Baù et al.*, 2015].

4.2.4 Framework for Filling Missing Groundwater Head Measurements

Advancements in computational prowess and parallelization schemes that improved efficiency have given the ability to physically model large groundwater systems in complex basins in a more detailed way than it was in the past. Regional models tend to inform about the overall state of the studied groundwater system, thus essentially being a mass balance with little efficacy about local conditions. They lack the ability to fully capture the finer local variations due to system simplifications, imperfect input information, and the scale of model discretization. Hence, even though the model provides the best unbiased estimate for a cell, it cannot be used as a true measure of the system. Groundwater head measurements are still the most reliable information of the systems' state and the fundamental building block of any further analysis. Despite that, in many real cases, the lack of systematic groundwater head measurements in combination with the different sampling timing and frequency has led to groundwater level time-series with gaps and irregular time intervals, thus creating challenges for further analysis.

The proposed framework for inputting missing observations in head levels is intending to bridge the disconnect between regional modeling and localities in system fluctuations, by utilizing all available data, direct (known measurements) and indirect (model information), in an efficient scheme with low computational cost. The framework is illustrated in Figure 35 with its main components listed below:

1. The observation data for the location of interest,
2. The extracted information from the groundwater flow model,
3. Stochastic modeling via Seasonal ARIMA with exogenous inputs, and
4. The Ensemble Smoother (ES) algorithm

First, the observation well of interest is identified and all available groundwater head measurements for that location are recovered from a public repository. The framework does not require the sampling frequency to be consistent across the observation record. The only two prerequisites are to aggregate measurement data to the time step of the groundwater model in case the sampling frequency is higher, and the measurements have been taken within the time span of the simulation of the aquifer. The simulated head time series for the specific well location in the model domain is extracted. Additionally, the time series, representing the average rates for recharge and pumping for several areas with different radii around the well location are retrieved.

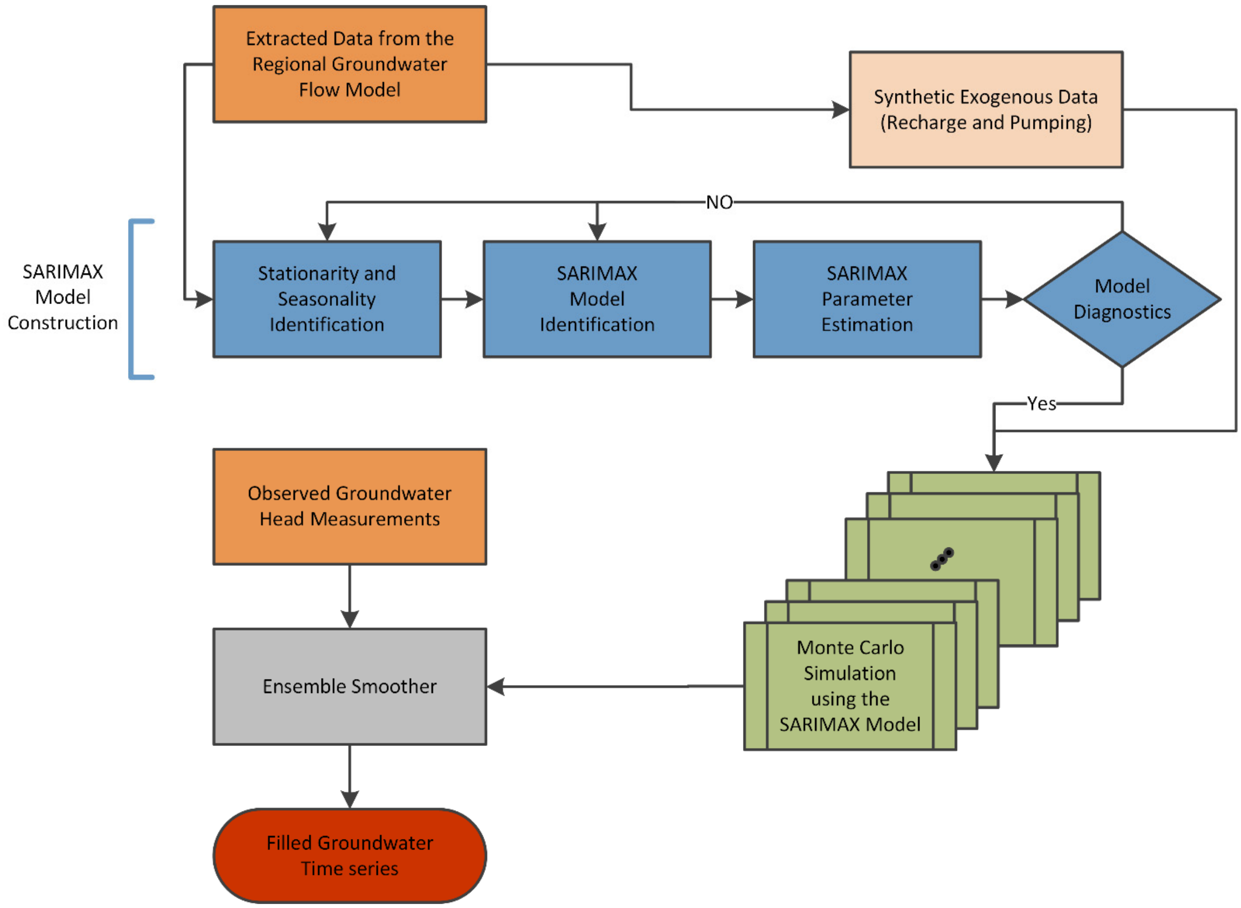


Figure 35: Groundwater Level Measurement Gap-filling Methodology

The next main component of the framework is the construction of a stochastic model representative of the process of the simulated head time series is indicating. The groundwater head level time series is modeled by a seasonal auto-regressive integrated moving average (SARIMAX) process, with exogenous inputs of recharge and pumping in order to give physical meaning to the statistical model. All the different time series of recharge and pumping should be tested during the SARIMA construction process in order to select the model with the best fit. The construction of a SARIMAX model consists of an iterative process of three steps: The identification that includes stationarity and seasonality testing, and, model identification; the parameter estimation; and the last diagnostic checking [*Box and Jenkins*, 1976; *Hipel and McLeod*, 1994]. Two supplementary approaches are used in order to determine the order of integration (d) in the SARIMAX models. The first approach is the use of visual identification with the autocorrelation function (ACF) and partial autocorrelation function (PACF). The second approach runs statistical tests. The augmented Dickey-Fuller (ADF) test was used to determine if a unit root was present in the time series. Additionally, the time series were tested by the Kwiatkowski-Phillips-Schmidt-Shin test, which complements the ADF unit root test, since it has low power against near unit root alternatives [*Karlsson and Löthgren*, 2000]. Seasonality, as well model selection can be identified by looking the ACF and PACF graphs. Table 9 describes the behavior of the ACF and PACF functions for SARIMAX models. For the SARIMAX construction, except from the above approaches, the "forecast" package in R is also involved in order to identify the model automatically, to do parameter selection and to run diagnostics of the SARIMAX models [*Hyndman and Khandakar*, 2008; *Hyndman*, 2016]. The SARIMAX model parameter estimation was done by using maximum likelihood estimation, and also by a t-test to measure statistical significance of each parameter, so not to include spurious terms. The overall fit of model parameters, with the same integration order, was evaluated by the Akaike Information Criterion (AIC) [*Akaike*, 1974]; however, the ACF of the residuals of each fit should also resemble a white noise sequence.

The exogenous variables are assumed to follow in each month a Gaussian distribution. Synthetic exogenous variables are generated based on the original first order statistic of each month and with variances that are a bit larger than those computed. The larger variances were introduced to increase indirectly the ensemble uncertainty in cases where the groundwater flow model was not approximating the behavior of the measured data. The optimal SARIMAX model with the exogenous synthetic time series are used to produce the ensemble of time series needed for the Ensemble Smoother (ES) [*van Leeuwen and Evensen*, 1996]. At the final step, the ES is employed to impute the gaps in the time series, by taking into account also the information provided by the groundwater head observations.

The description above of proposed framework's scheme was designed to have the following key characteristics: Low computational cost, since it is not dependent on the groundwater model for the generating the forecast ensemble for each evaluation point; ease of implementation, the ability to incorporate information from readily available groundwater models irrespective of the coarse scale of application; and the ability of this methodology to be applied in any alluvial aquifer regardless of local conditions.

Table 9: Behavior of identification functions for SARIMA models [Hipel and McLeod, 1994, p.432]

FUNCTION	**TYPES of MODELS**		
	PURE AR $(p,d,0)(P,D,0)s$	**PURE MA** $(0,d,q)(0,D,Q)s$	**MIXED** $(p,d,q)(P,D,Q)s$
ACF	Attenuates	Truncates after lag $q + sQ$	Attenuates
PACF	Truncates after lag $p + sP$	Attenuates	Attenuates

4.3 Materials and Numerical Experiments

4.3.1 Application Area, Available Model and Data

The area selected to evaluate the performance of the proposed computationally low-cost groundwater data-filling framework, is the alluvial aquifer of South Platte River. The South Platte River Basin in northeast Colorado has a drainage area of about 49,000 km^2 [*Dennehy et al.*, 1993] and, it is characterized by a semi-arid climate, with a high dependence on melting snow. The South

Platte River Basin, on the eastern plains, has undergone a huge transformation due to land reclamation, agricultural development and population increase. It is the most populated region of Colorado, since approximately 70% of state's population (more than 3,700,000) is living in this particular area [*Colorado Department of Local Affairs*, 2016]. The South Platte is an over-appropriated basin that has multiple water needs and, apart from the domestic sector, it has to satisfy a large agricultural demand. In 2010, the estimated irrigated parcel area in the basin was 3,426 km^2 (846,634 acres) [*Colorado Decision Support Systems*, 2016]. Industrial activities are also present in the basin, including water demand for unconventional oil and gas development [*Oikonomou et al.*, 2016]. Groundwater plays a substantial role in meeting these water needs, particularly in the agricultural sector.

The physical conditions, anthropogenic stresses and interventions, along with the water legislative environment governing the South Platte basin are adding to water system's complexity. This complexity is also reflected by the different water management approaches over the time to bring groundwater into the prior appropriation system, so as not to harm senior surface water rights [*Waskom*, 2013]. The 2002-2003 drought in Colorado was a catalyst for implementing a stricter water administration, requiring full augmentation for system depletions caused by pumping tributary groundwater. One effect of this water management shift was the curtailment in 2007 of about 5,000 wells that lacked an approved augmentation plan [*Waskom*, 2013]. At the same time, in certain areas of the basin, were reported incidents of adversely impacted properties by high groundwater levels [*Waskom and Oikonomou*, 2014] forcing the State of Colorado to act in this emergency [*Colorado General Assembly*, 2015a].

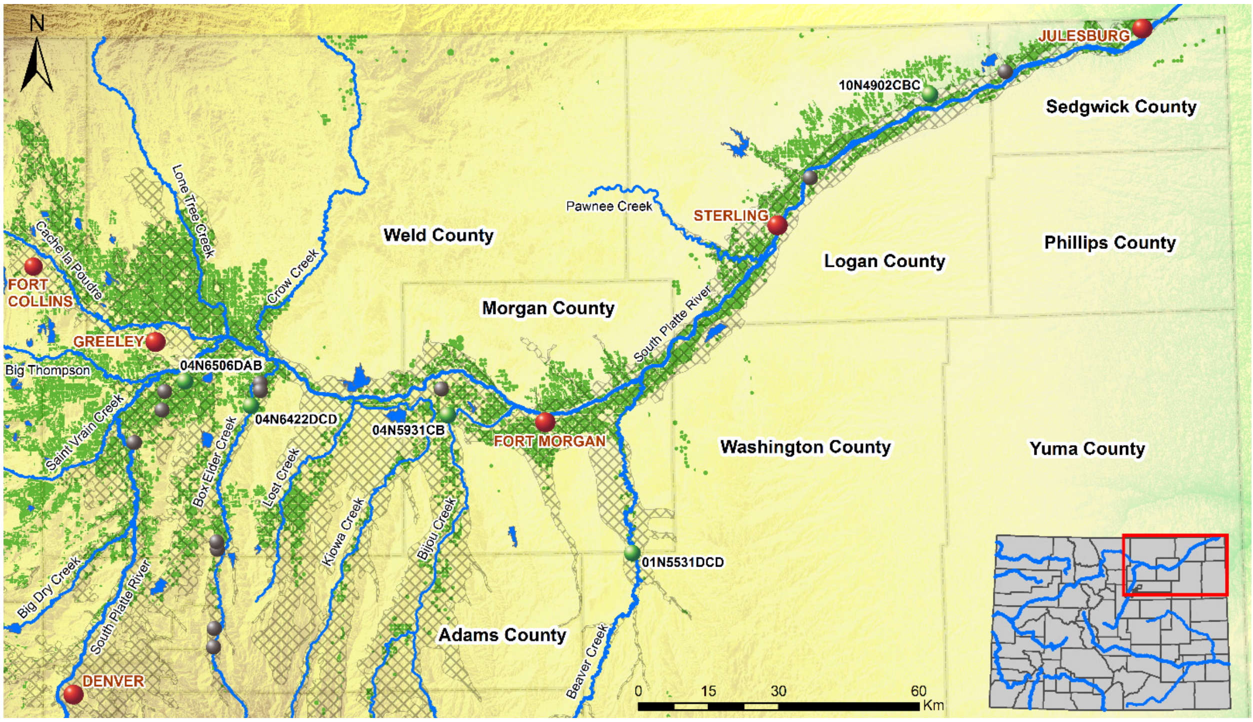


Figure 36: Application Area and location of selected groundwater wells (the mesh represents the alluvium and the irrigated parcels with green color)

The state of Colorado has requested research projects [*Colorado General Assembly*, 2012, 2015b, 2015c], that would assist the administrative authorities to have a better understanding of surface-groundwater interactions and thus, to implement a basin-wide efficient, equitable and integrated water resources planning and management. Through the Colorado Decision Support Systems (CDSS) initiative (www.cdss.state.co.us), a physically-based groundwater flow model was built and tested by CDM-Smith for the South Platte Alluvium (SPDSS-MODFLOW). The model was developed in MODFLOW-2000 [*Harbaugh et al.*, 2000] with a monthly time step from January 1950 to December 2006 [*CDM-Smith*, 2013]. The SPDSS-MODFLOW is covering approximately 6,475 km^2 (2,500 mi^2), and the aquifer is represented as a single layer having a uniform grid of 304.8 m (1,000 ft) which results to 555,440 cells [*CDM-Smith*, 2013]. It was built for regional scale analysis, focusing on aquifer-stream interactions, and thus, it is not able to simulate accurately local groundwater level variations.

In the South Platte, there are few dedicated monitoring wells with continuous data loggers (spanning from 2003 or 2004 until the present), but the majority are irrigation wells measured biannually (Fall and Spring). The water table generally follows a cyclic pattern, with highest observations occurring during the fall period following the irrigation season, and the lowest, in the spring. Some of the observation wells have a longer water level record, going back to the 1930s and 1940s. Of course, there are several years that the irrigation wells have no available groundwater head measurements, in both an intermittent and continuous fashion. It is clear that, through the years, there were periods that the groundwater-monitoring program was not a top priority of the institutions involved in collecting groundwater head measurements.

For the purposes of this study, 18 South Platte Alluvium wells (Figure 36, grey and green spheres), which had the longest available observation record, were selected. These wells are monitored by the Colorado Division of Water Resources (DWR) and the Central Colorado Water Conservancy District (CCWCD). For brevity, the focus is given to only five wells, represented in Figure 36 with green spheres, while the rest are shown in the Appendix G. Their selection was based on their representation of different sections of the alluvium, proximity to the stream, exogenous predictor variables radius of influence, and the modeled stochastic process. The groundwater head observations are publicly available through the State's water database, called "Hydrobase" (http://water.state.co.us/DataMaps/Pages/default.aspx), along with several metadata for each well; such as its position, owner's name, permit number, adjudicated uses, etc. The monthly simulated head at each well's location, the mean recharge and total pumping within different radii from each well were extracted from the SPDSS-MODFLOW. A SARIMAX model for each well was used to model the groundwater head, since all wells exhibited a 12-month cycle. The simulated monthly head for the five selected wells is plotted in Figure 37 with the recorded actual measurements (Appendix H the rest wells). The graphs support the aforementioned statement that the regional model should be used for regional scale analysis. On a local scale, the SPDSS-MODFLOW can only represent a general indication of groundwater behavior (seasonality), without being able to simulate accurately the water table level. At well 01N5531DCD, that is located at the beginning of the alluvium of the Beaver Creek tributary of the South Platte, the model fails to simulate the actual general behavior of the water table at that location. At the same time, it captures the seasonal variability rather well (small head fluctuation from fall to spring season). At the particular cell that the well 04N6422DCD is located at the Box Elder Creek tributary, the SPDSS-MODFLOW is not able to approximate the local water table trend as well; however, at the same time, it is capturing the magnitude of seasonal variability. For the other three wells, the model at these points is representing more appropriately the water table, considering it is a regional model. It should be underlined that, since the groundwater head measurements are taken biannually, it is assumed that those observations are representative of the true mean head for these months. Additionally, all of the tested wells are used for agricultural purposes and the unceasing probability of some outliers to be present due to measurements taken after water was pumped for irrigation.

Evaluating the performance of the proposed methodology for filling gaps in groundwater head time series in a complex hydrologic system with multiple stresses and governed under a strict scheme can be an excellent proof of its robustness. The South Platte basin below Denver and its aquifer, as mentioned above, fulfill all of the criteria that make it an ideal area of study. Overcoming existing data limitations, with methodologies that augment the available water system information, could contribute to new insights that would contribute into taking better decisions towards sustainable groundwater management.

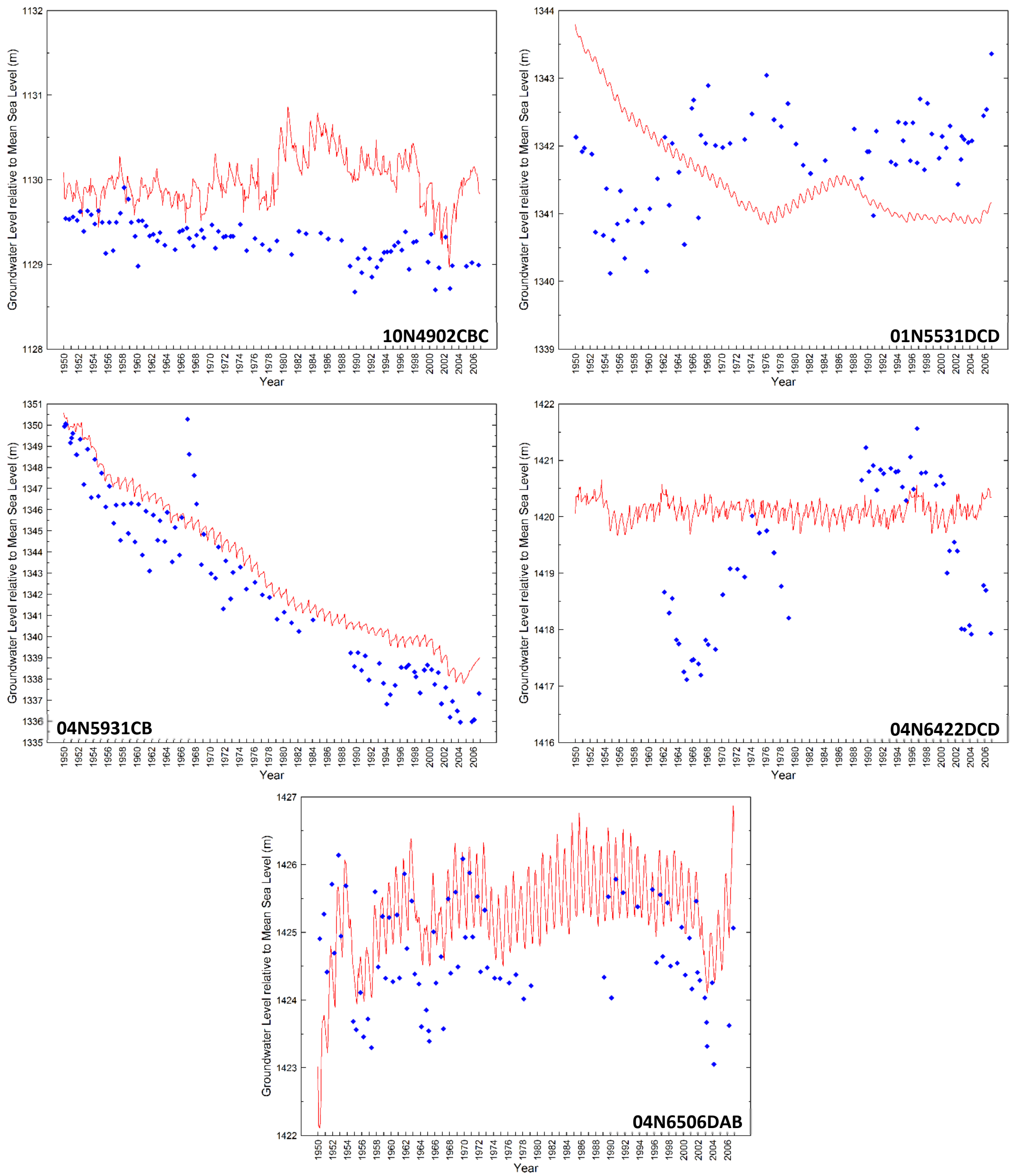


Figure 37: Modeled monthly groundwater level (red line) and biannual observed level (blue dots)

4.3.2 Design and Assumptions of Numerical Experiments

The observed groundwater head over time is the only available measure of evaluating the effectiveness of the methodology. The performance of the proposed methodology for filling data-gaps in groundwater time series has been evaluated by three numerical experiments. The first numerical experiment is assuming that all known water table measurements are available for updating the Ensemble Smoother (ES), except one. The Ensemble Smoother (ES) algorithm is run iteratively, and after each reiteration, another known point is treated as unknown. The average prediction performance is described by calculating dimensioned statistics (in this case in meters), namely: the root mean squared error (RMSE) and the mean absolute error (MAE)

$$RMSE = \left[n^{-1} \sum_{i=1}^{n} e_i^2 \right]^{1/2} \tag{23}$$

$$MAE = n^{-1} \sum_{i=1}^{n} |e_i| \tag{24}$$

where, $e_i = y_i - \hat{y}_i$, y_i is the observed value, $\hat{y}_i$ the estimated value, and n the sample number, of observations. The use of RMSE though, can be misleading many times, due to the fact that the larger the squared error variance is, the larger is the RMSE value (Eq. 23). The MAE (Eq. 24) has been suggested by Willmott and Matsuura [2005] as a more natural measure of average model performance error. In this study, both RMSE and MAE are presented, since they represent the common measures of average prediction performance; however, in the results section, the focus is MAE.

At the second numerical experiment, five random measurements are treated as unknowns, while the RMSE and MAE are calculated and stored. This experiment was selected in order to test the performance of the method to simulate the case of intermittent gaps present in the groundwater head time series. The numerical experiment is repeated for each well with each iteration having a new random set of five values, until there is convergence on the mean and the standard deviation of the sample RMSE. Convergence was assumed when there was no improvement more than 0.1% of both moments for the past 1000 runs. The density histograms of RMSE and MAE were calculated and plotted, because the number of iterations for each well could vary. The density property was expressed in percentages and the entire histogram equals one. The last numerical experiment is constructed to simulate the case of continuous missing values in the groundwater level. A contiguous five-year window is selected randomly and the observations within the window are treated as unknowns. The Ensemble Smoother (ES) algorithm is employed to calculate the RMSE and the MAE of each iteration. Similar to numerical experiment #2, each iteration has a new random five-year window. The numerical experiment for each well is completed with the same convergence assumptions, as in the previous numerical experiment.

As presented in the methodology section, the modeled groundwater head at the well locations was modeled with a SARIMAX model. Different time series of the exogenous predictor variables (recharge and pumping), representing different radii of exogenous influence, were tested in order to represent the stochastic physical process more accurately. For consistency reasons, the set of radii selected to be tested for each well location was 914 m (3,000 ft), 1829 m (6,000 ft), 3,048 m (10,000 ft) and 4,877 m (16,000 ft). The radii originating from each well location that intersected the river, were not taken into account for testing the effect in the SARIMAX fitting. Table 10 shows the best SARIMAX models of each radius of influence and their AIC. In some cases, the model approximation for pumping was not statistically significant in order to be included in the model parameters. The autocorrelation function (ACF) of the residuals of the SARIMAX models

(Figure 38) is evidence of good fit [*Box and Pierce*, 1970; *Box and Jenkins*, 1976]. The autocorrelation plots show that only few autocorrelations are outside the 95% confidence interval around zero. This can be explained by the unconfined aquifer's nonlinear dynamics, since the ARIMA family of models assumes global linearity.

Table 10: SARIMAX models for the tested radii and the AIC

Well ID	ARIMA MODEL	Radii in m (ft)	AIC
10N4902CBC	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	914 (3000)	-194.02
	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	1829 (6000)	-8.37
	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	3,048 (10000)	-29
01N5531DCD	(1,1,2)(1,0,1)[12] with zero mean, xreg = R, P	914 (3000)	-3732.05
	(1,1,2)(1,0,1)[12] with zero mean, xreg = R, P	1829 (6000)	-3765
	(1,1,2)(1,0,1)[12] with zero mean, xreg = R, P	3,048 (10000)	-3828.69
	(1,1,2)(1,0,1)[12] with zero mean, xreg = R, P	4,877 (16000)	-3804.66
04N5931CB	(2,1,0)(1,0,1)[12] with zero mean, xreg = R	914 (3000)	210.02
04N6422DCD	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	914 (3000)	-67.93
	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	1829 (6000)	-84.04
	(2,0,2)(1,0,1)[12] with non-zero mean, xreg = R	3,048 (10000)	-99.85
	(2,0,2)(1,0,1)[12] with non-zero mean, xreg = R	4,877 (16000)	-69.03
04N6506DAB	(1,0,2)(2,0,1)[12] with non-zero mean, xreg = R	914 (3000)	-31.63
	(2,0,1)(2,0,1)[12] with non-zero mean, xreg = R	1829 (6000)	-53.3

The ensemble matrix ($\boldsymbol{A}$) accounting for model uncertainty characterization is populated with 500 time series members, generated from the best SARIMAX model for each well location. The exogenous predictor variables extracted from SPDSS-MODFLOW are assumed to follow a Gaussian distribution for each month. The synthetic exogenous time series used as inputs to SARIMAX have the same monthly mean, with an increased monthly standard deviation. In all 3 numerical experiments, for each well, the same ensemble matrix was used. Furthermore, the statistics of the measurement errors in most practical applications are unknown, so in order to incorporate the measurement uncertainty, the measurement error covariance matrix ($\boldsymbol{R_e}$) was constructed assuming the errors follow a Gaussian distribution with a mean zero and variance of 10.16 cm^2 (1/3 ft^2). This assumption represents the possible human error during a groundwater head measurement. The incorporation of a larger error could also relax the assumption that a single water table observation is assumed to represent the average groundwater head for that specific month.

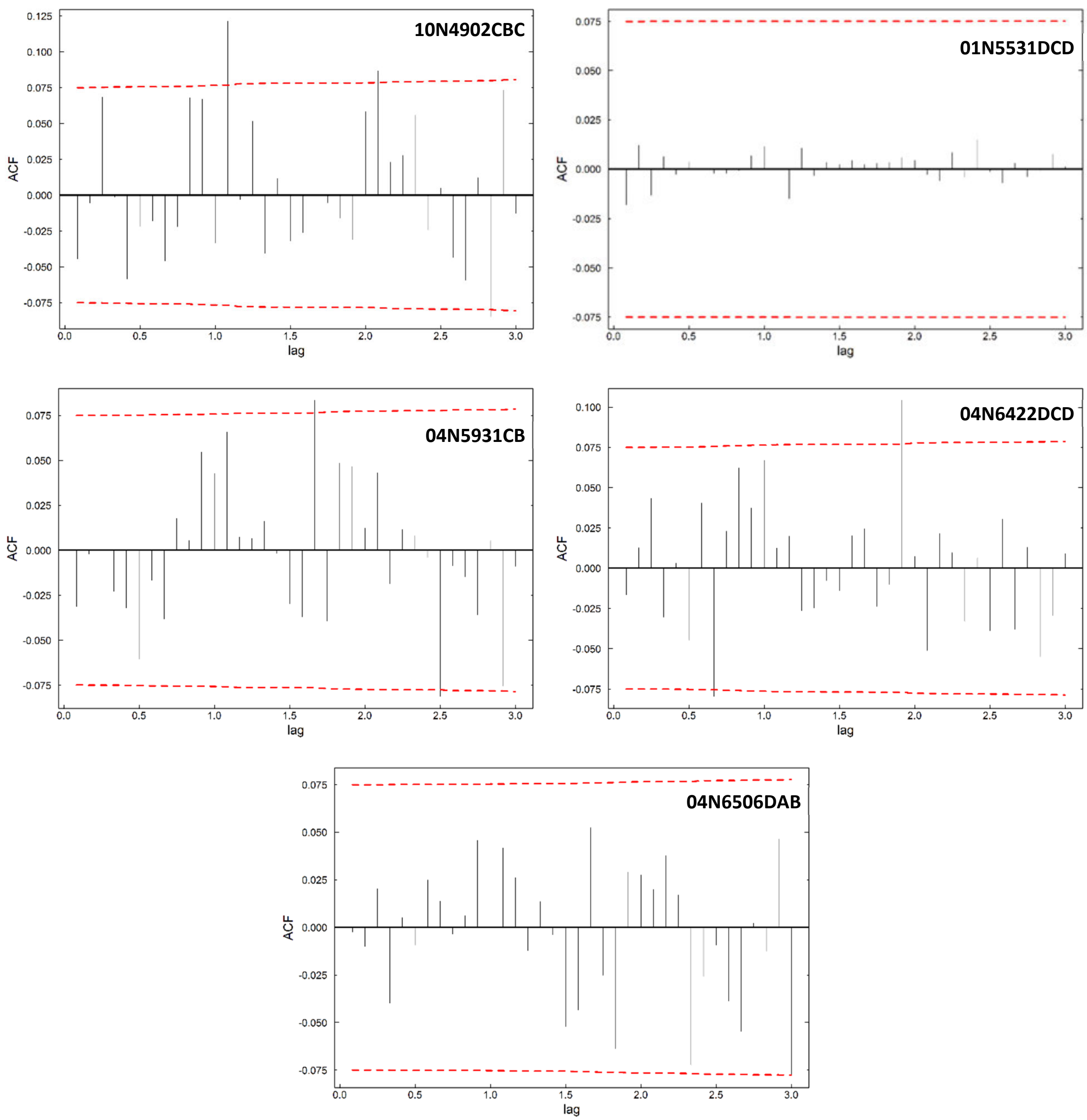


Figure 38: Autocorrelation function of residuals of the SARIMAX models

4.4 Results and Discussion

4.4.1 Numerical Experiment #1

The purpose of this experiment is to illustrate the performance of the suggested framework in the five selected wells where only a single measurement is to be filled in. The scatter plots of the measured versus the predicted value are shown in Figure 39. Each point in the graphs of Figure 39 represents the individual error with its magnitude being the perpendicular distance from the diagonal line. In all scatter plots, the data points are clustered around the diagonal, indicating that the Ensemble Smoother (ES) algorithm has almost no bias. Furthermore, the graphs show a homoscedastic behavior, since in different water table levels of each well there is no difference in the variance of the prediction error. In well 04N6422DCD however, there is a consistent tendency of overestimation in the lower water table values. This bias effect is most probably caused due to the fact that the water table in the period of 1964 to 1970 is at the lowest point and after that, it bounces back.

One key outcome of this numerical experiment was the inability of this methodology to predict piezometric heads when there were no similar measurements in temporal proximity. The water table divergence from temporally adjacent observations could be outliers due to drought/flood conditions, since the South Platte alluvial aquifer responds quickly to surface hydrology variability. The divergence could also be attributed to human errors during measurement or to measurements taken soon after the well was pumped. Low performance is expected when the SARIMAX model fitted on the groundwater levels simulated from the SPDSS-MODFLOW in the specific cell of the aquifer has no similar behavior. A characteristic example of a satisfactory prediction of an observation value, which is significantly different from its immediate temporal neighbors, is the last known point of well 04N6506DAB (Figure 37). Despite the abrupt change of the water table, the proposed methodology managed to give a rather precise estimate. This is because the physical groundwater model of the well 04N6506DAB is able to represent the groundwater upward trend after 2001 and this information is transferred to the SARIMAX model.

Similarly, at temporal points where a change in the water table is taking place, and there are no intermediate observations in between, the predicted value is frequently not accurate. Despite the fact that the Ensemble Smoother (ES) update step is assimilating all available measurements, adjusting the average updated ensemble to follow an abrupt change is not possible when there are no relevant data points and when the SARIMAX model is not representing this change. More specifically, the highest water level observed at well 04N5931CB exhibits a large prediction error (~3m), since there is no observation to inform the update scheme of the Ensemble Smoother (ES) for this change. This specific example is showing how important is to have key measurements that capture accurately the response of the water table in time.

The prediction of an observed "missing" point in general is giving satisfactory results as the MAE is low. In all tested wells– and not only the presented ones– the overall estimation via the Ensemble Smoother (ES) algorithm was closer to the observed value, than filling the "missing" points with the average water table value. The RMSE and MAE values for each well are shown on the top right of each scatter plot of Figure 39. All wells except 04N5931CB have MAE less than 0.3m.

It is worth reiterating that the suggested algorithm was tested on wells measured on a biannual step, which had significant data gaps, ensuring a realistic test on the suggested methodology. The temporal scarcity of observation points in the tested wells in combination with the use of a regional groundwater model, which in many cases is unable to approximate the groundwater fluctuations, are impediments of having even better results.

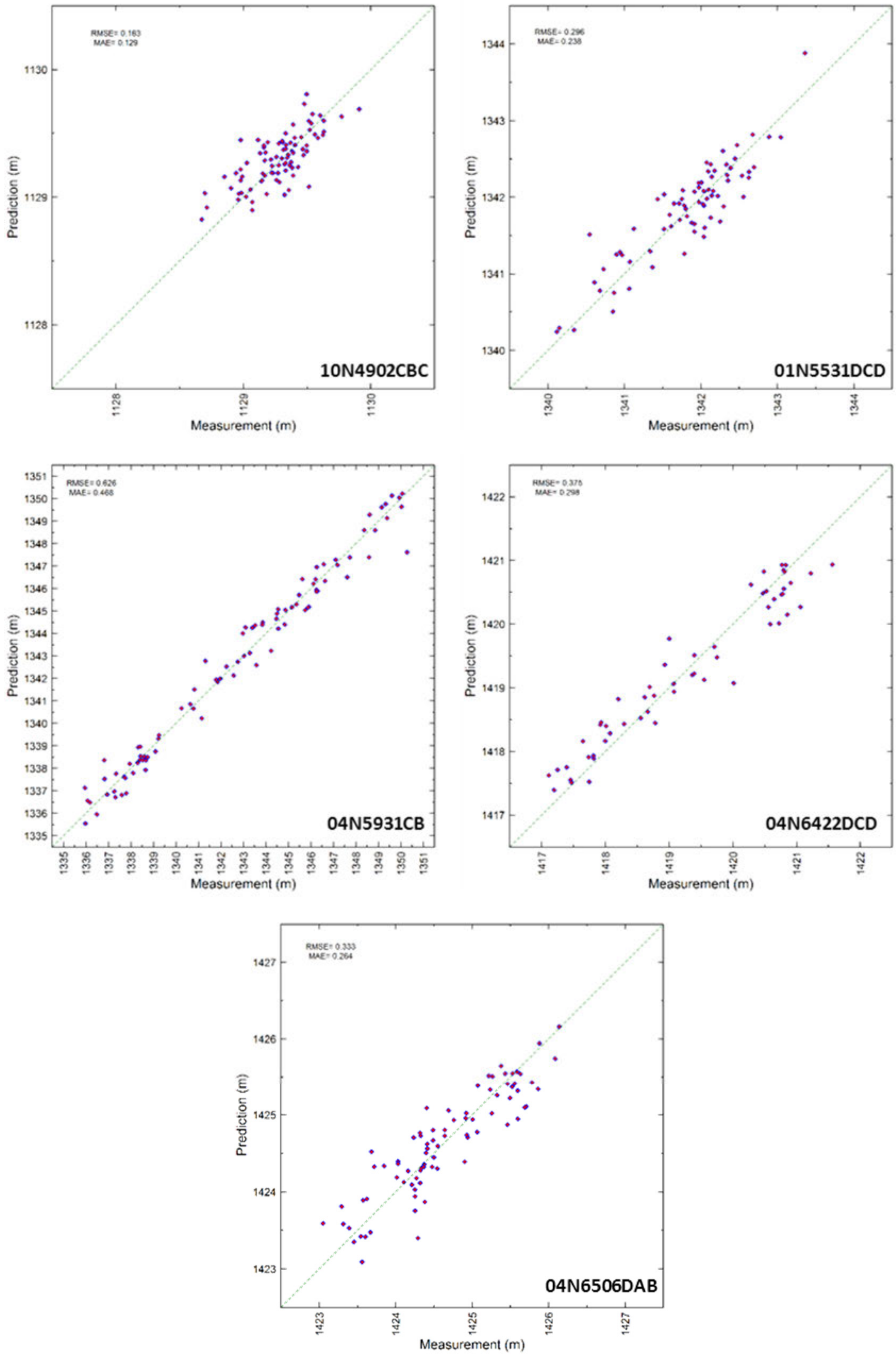


Figure 39: Results of numerical experiment #1

4.4.2 Numerical Experiment #2

Numerical experiment #2 is more challenging than the previous one, since five random observations are treated as unknown. Subsequently, the experiment was repeated enough times in order for the samples to be representative. The probability distribution of two prediction error metrics of the results for each well is presented by density histograms of RMSE and MAE. The iterations needed for the mean and standard deviation of RMSE to converge for each well are not the same and thus frequency histograms are not an appropriate representation. The iterations needed for each well for the first two moments of RMSE to convergence, as the sample size increased, are shown in Table 11.

Table 11: Iterations needed for convergence of the two first moments of RMSE

Well ID	**10N4902CBC**	**01N5531DCD**	**04N5931CB**	**04N6422DCD**	**04N6506DAB**
Iterations	5,157	5,307	8,889	11,433	6,253

The density histograms of all wells are shown in Figure 40 and Figure 41, and the density is expressed in percentage for ease of representation. The histograms appear to be roughly symmetric or slightly right-skewed, since there is the natural limit of zero error, and the temporal points that were larger than the average error in experiment #1 are influencing the average error metrics. The distribution of all wells appear as unimodal with the only exception of the RMSE histogram of well 04N5931CB. The bimodal distribution can be attributed to the increased weight RMSE is giving to larger errors, since the random sample of “unknown” observations could contain values from the cluster of points between 1966-1968, that diverge from well’s general trend (see Figure 37). It is worth mentioning that the bimodality is not present in the histogram of MAE, since the large individual errors are masked.

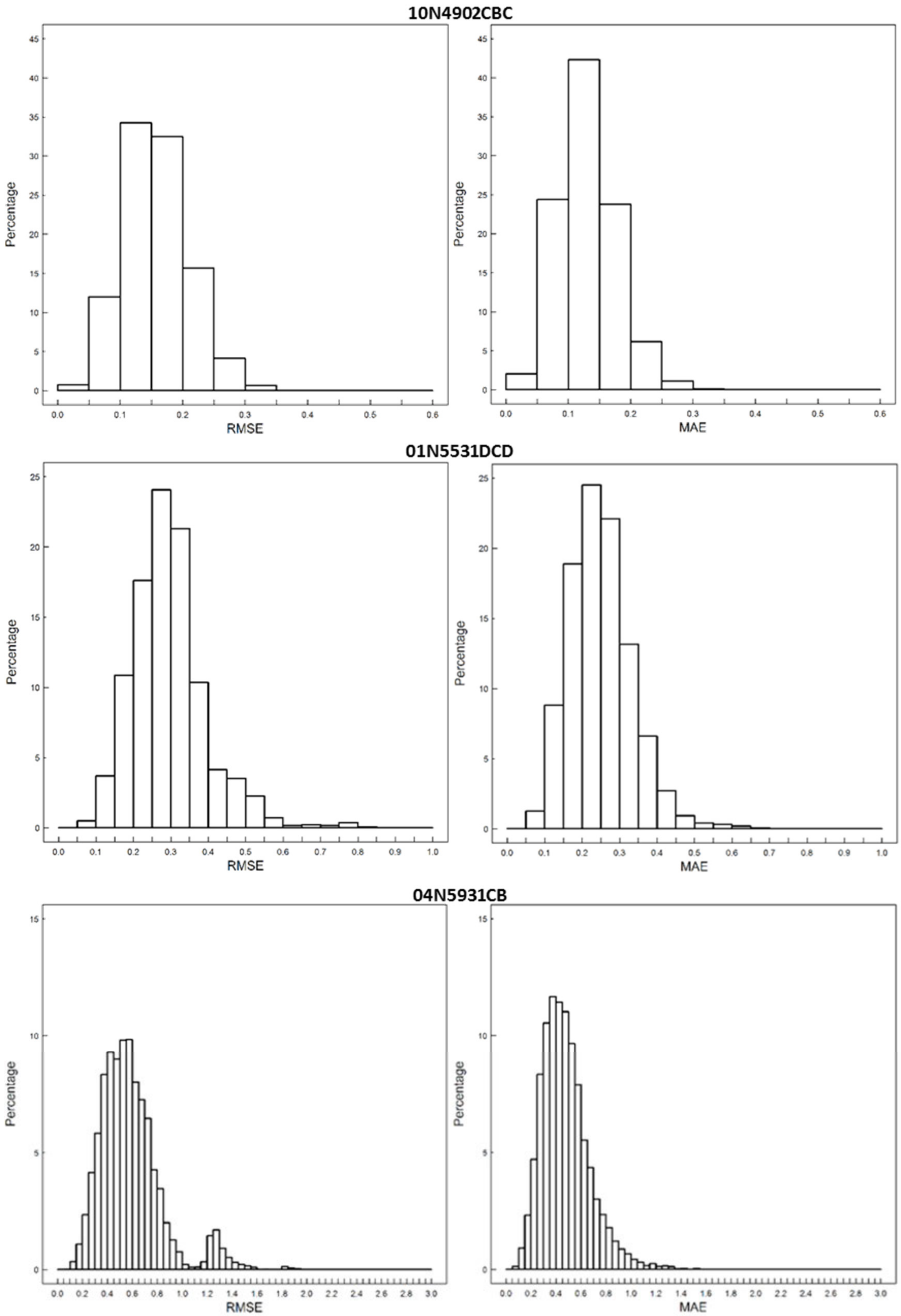


Figure 40: Relative density histograms of RMSE and MAE for wells 10N4902CBC, 01N5531DCD and 04N5931CB

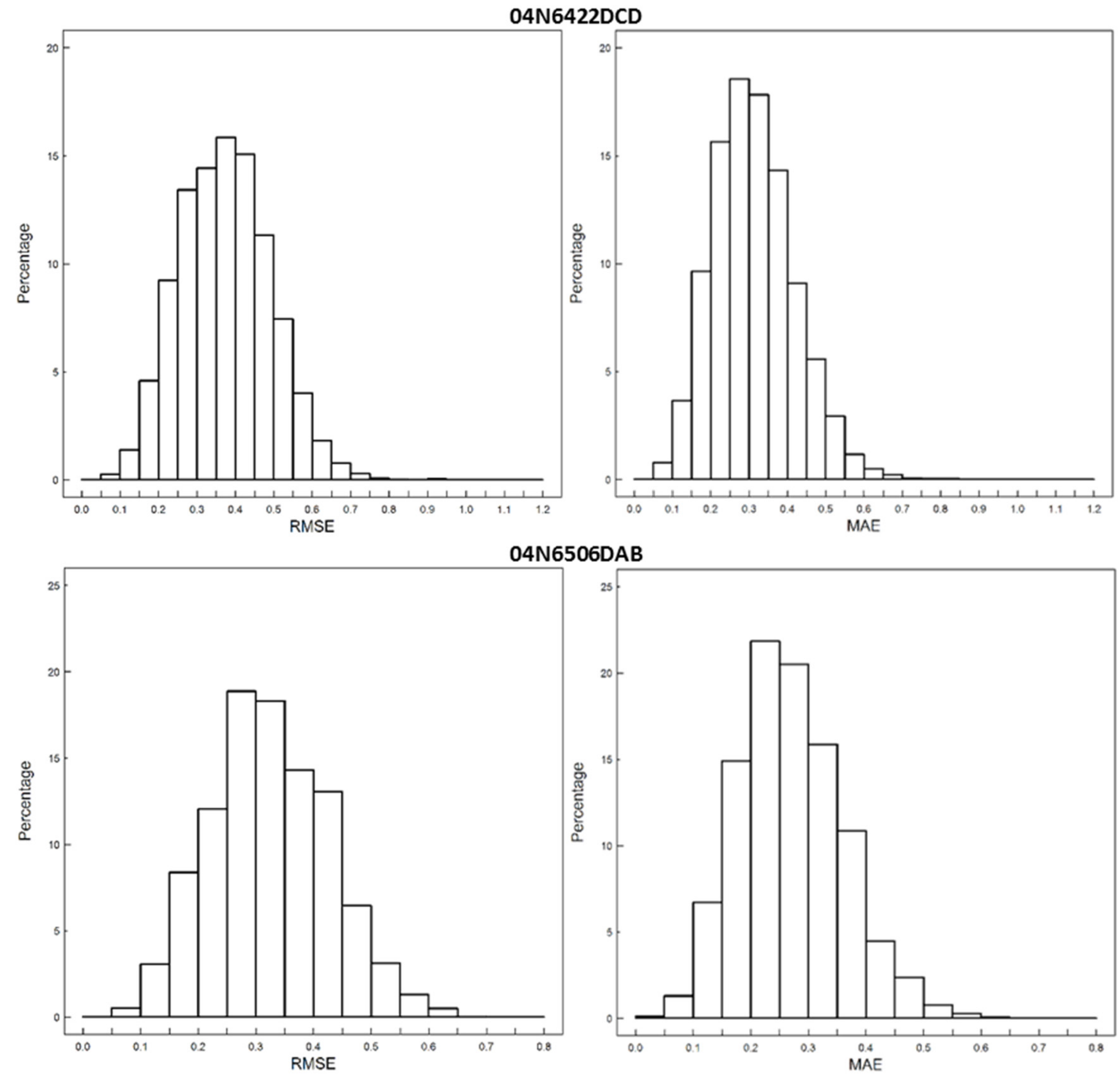


Figure 41: Relative density histograms of RMSE and MAE for wells 04N6422DCD and 04N6506DAB

The proposed methodology for filling the gaps has a very satisfactory performance, since the median MAE is small (less than 0.3 m). Similar to Experiment #1, well 04N5931CB has a larger sample MAE median compared to other wells (just below 0.4 m). It can be considered as a reasonable well prediction, since its “outliers” and non-stationary behavior are challenging factors for accurate prediction. The maximum MAE recorded in these numerical experiments is significant; however, it should be kept in mind that the wells tested have low measurement frequency (biannual) and significant data gaps.

4.4.3 Numerical Experiment #3

Numerical experiment #3 is the most challenging of all numerical experiments, since in a five-year window, the available information of the piezometric level is treated as unknown. Long periods could enclose prolonged drought occurrences and miss critical data points at the update phase of the Ensemble Smoother (ES), resulting to a poor water level prediction. Such information becomes even more important when the subject aquifer is an unconfined alluvial formation when the water table is responding to surface variability. Similar to the previous numerical experiment, the iterative process of treating observations within a five-year period as unknown was terminated when there was convergence in the first two moments of the sample RMSE. The needed iterations for convergence in each well are shown in Table 12. Furthermore, the predication performance of the methodology is presented in an analogous way to the Numerical Experiment #2. Figure 42 and Figure 43 show the density histograms of the RMSE and MAE for all selected random samples.

Table 12: Iterations needed for convergence of the two first moments of RMSE

Well ID	**10N4902CBC**	**01N5531DCD**	**04N5931CB**	**04N6422DCD**	**04N6506DAB**
Iterations	5,114	11,024	9,752	7,631	5,245

It is clear from the presented histograms, that the symmetry on both error metrics is lost with right skewness, and large spread and bi/multi-modality are the key characteristics of the produced error distributions. Only in well 10N4902CBC is the prediction performance within satisfactory levels (maximum MAE is 0.42m) considering the large gap attempted to be filled. In all other wells, the calculated errors in many of the randomly selected samples are large. The spread of the error distributions for well 04N5931CB is 3.5m with many outliers from the main pattern. This is indicative of the poor performance in this specific well; but it also emphasizes the crucial role the known measurements play for the under-prediction time interval.

Due to the gaps existing in the tested time series, especially in wells 04N6422DCD and 04N6506DAB, there were incidents that the randomly selected five-year windows contained only a few observations. The errors from those samples, being the outliers of the general pattern, can be traced in RMSE histograms as the rectangles erecting from low error bins. High error outliers are occurring when the observations under prediction are including possible outliers or these values are too different from the rest of the recorded measurements in the same location. At the same time, this indicates that the methodology is applicable for predicting water table values, while short-term continuous data are missing. Overall, the tested framework should be used with caution particularly in cases which, continuous data gaps are attempted to be filled.

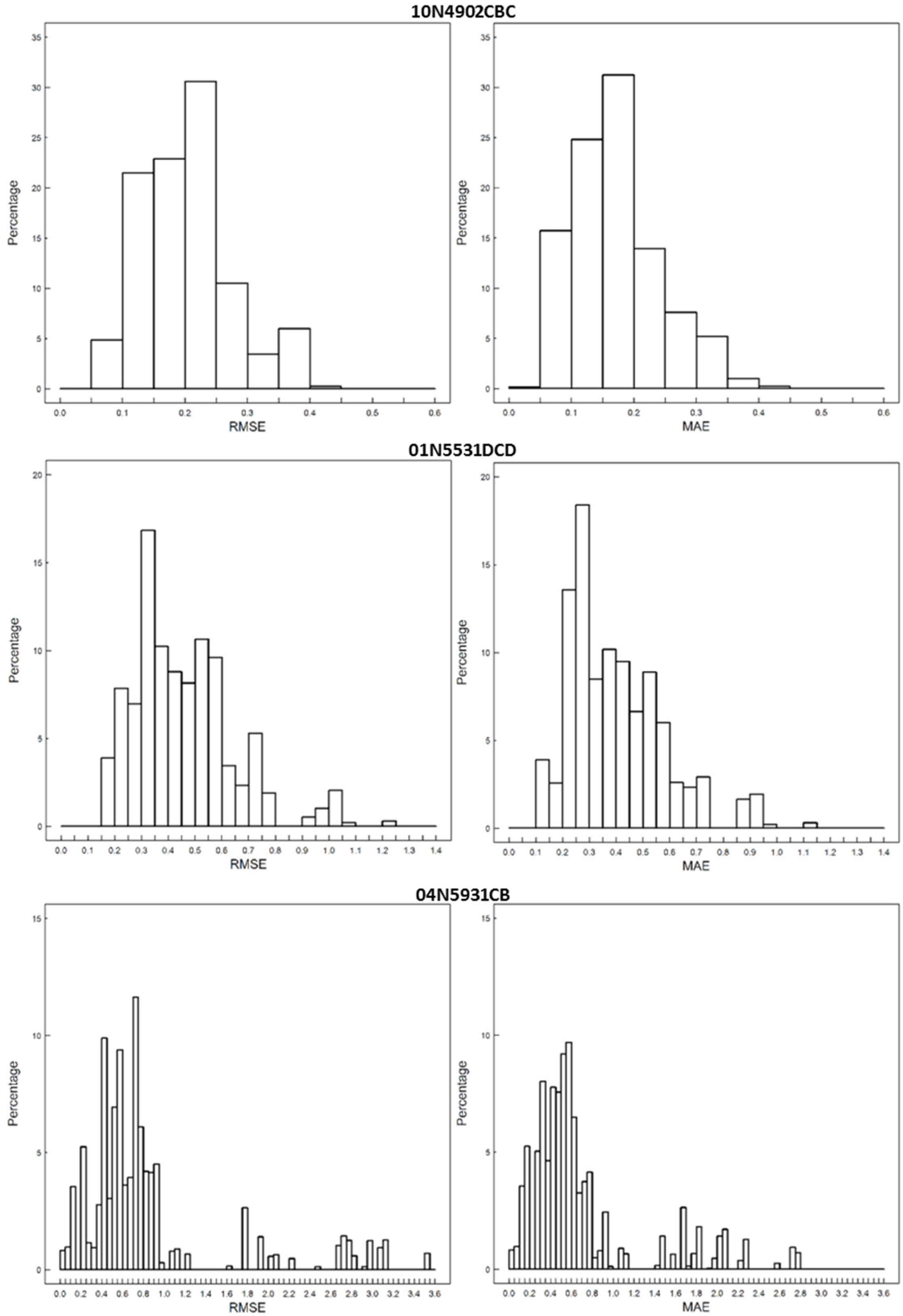


Figure 42: Relative density histograms of RMSE and MAE for wells 10N4902CBC, 01N5531DCD and 04N5931CB

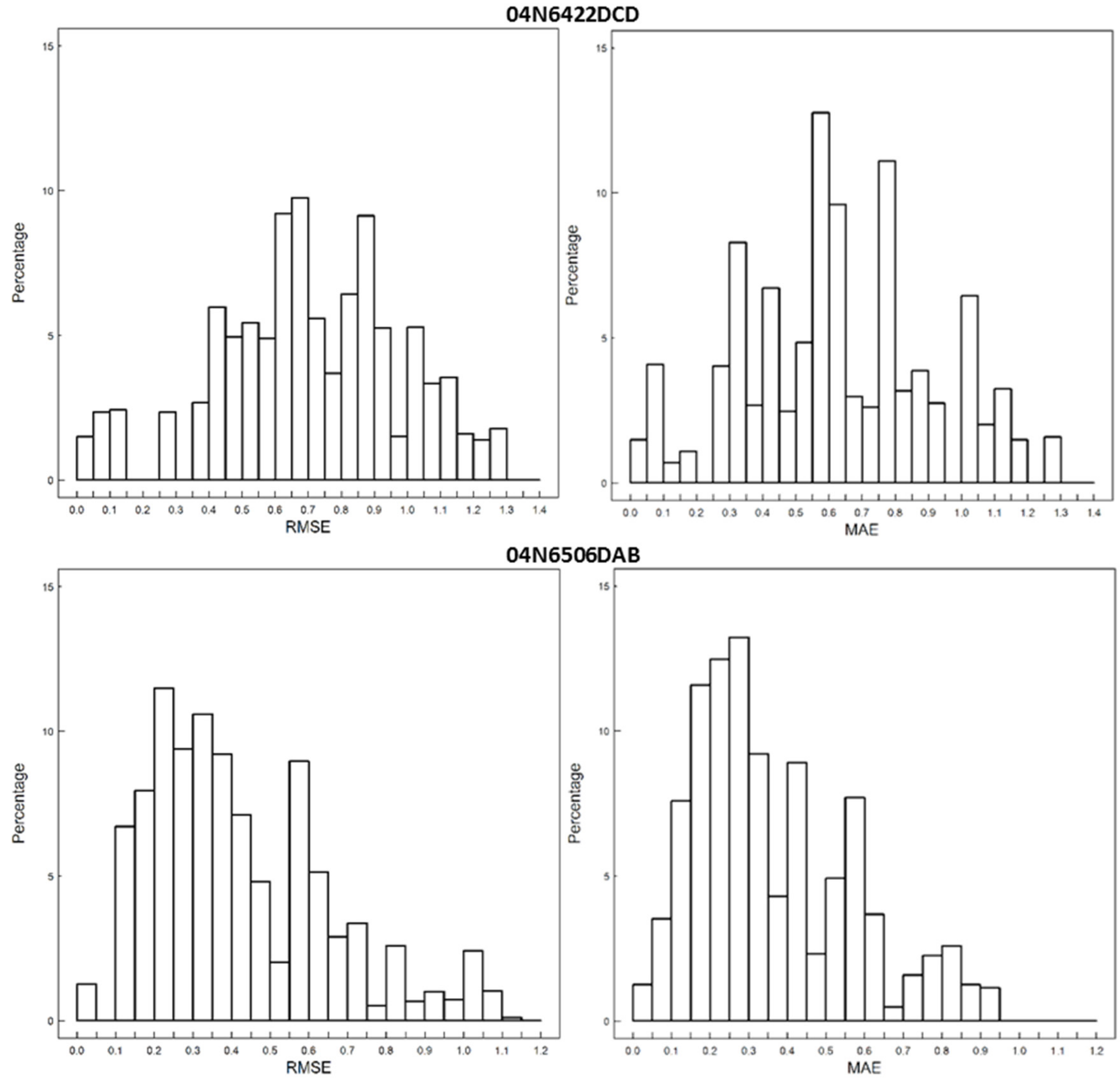


Figure 43: Relative density histograms of RMSE and MAE for wells 04N6422DCD and 04N6506DAB

4.5 Conclusions

Groundwater level data is the building block of data driven analysis and modeling. Groundwater head observations often have missing observations and suffer from inconsistent measurement frequency. In the current effort, a novel framework for filling gaps in groundwater head time series was proposed. It attempts to enhance groundwater head time series by utilizing information of all available sources, direct (known measurements) and indirect (groundwater flow model information), through an efficient scheme with low computational cost. The framework employs an exogenous seasonal autoregressive integrated moving average (SARIMAX) stochastic model to describe the groundwater level fluctuation process and the Ensemble Smoother (ES) for predicting the water table level. The methodology was implemented in South Platter River basin, which has complex hydrology due to its dependence on melting snow with significant hydrology modification. The application of the proposed methodology in such a complex water system adds value to numerical experiments, since it indicates the transferability of the framework to other areas.

Three numerical experiments were designed to evaluate the prediction performance of the proposed framework for 18 well locations with biannual measurements; however, only five are presented for brevity. The scope of the first numerical experiment was to test the performance of the methodology for cases, where a single measurement is missing. Its performance was satisfactory, since the average MAE of all five wells is 0.28m. The algorithm encountered difficulty in predicting a sudden change in head measurement when there were no similar measurements in temporal proximity. Furthermore, if the model was not approximating the general tendency or the seasonality of the measured values, then the prediction error increased. The second numerical experiment's design was an iterative process, in which every run a new set of five randomly chosen observations were treated us unknowns. The experiment was concluded only when there was convergence on the first two moments of the sample RMSE, and the results of all wells are very satisfactory with similar error metrics, as in Experiment #1. On the contrary, Numerical Experiment #3 was the least successful, since this time a random five-year window was selected and the observations that lie within were treated as unknowns. Similar to experiment #2, the iterative process was terminated when there was convergence of the sample RMSE. The histograms produced from the results of Experiment #3 are characterized with right skewness, large spread and bi/multi-modality of the produced error distributions. The importance of measured points to guide the updated step of the Ensemble Smoother (ES) is signified in this experiment. In addition, it should be noted that missing data in sequential fashion does not reveal the presence of extended drought and might overestimate the water table. There was only one well of the five presented where the prediction performance was satisfactory and this can be attributed to the small water fluctuations within the period of measurements. Numerical Experiment #3 proved to be the more challenging and performing poorly in most of the 18 tested wells.

The proposed framework is the first step of a formulation, and there is room for improvement. Future research could evaluate the performance of the framework taking into account that through time, the water table fluctuations could enter different regimes, and thus modeling the process in a piecewise fashion, by different temporal local SARIMAX models, could represent the process more accurately resulting in better performance. Such an approach might produce better results. Another interesting expansion of this work is the use of clustering methods with the purpose of constructing SARIMAX models for the sections of the aquifer with similar processes. The benefit of such an approach would be to prevent the need to construct a new SARIMAX model at every point of interest. Lastly, the performance of a different smoother algorithm, such as the Ensemble Kalman Smoother (EnKS), should be evaluated. With the employment of other Kalman smoothing frameworks, it would be possible to test the incorporation of non-linear stochastic models, which could perhaps result in improving the prediction accuracy. The framework could serve as a valuable tool for enhancing groundwater time series, for both intermittent missing data and of continuous gaps with short span. By advancing, the quality of input data in building physically based groundwater flow models and data driven analysis could gain further insight about aquifer dynamics and improve conjunctive use of surface water and groundwater resources as the integrated water resources management framework is promoting.

References

Aanonsen, S. I., G. Nævdal, D. S. Oliver, A. C. Reynolds, and B. Vallès (2009), The Ensemble Kalman Filter in Reservoir Engineering - A Review, *SPE Journal*, *14*(03), 393–412, doi:10.2118/117274-PA.

Ahmad, S., I. H. Khan, and B. P. Parida (2001), Performance of Stochastic Approaches for Forecasting River Water Quality, *Water Research*, *35*(18), 4261–4266, doi:10.1016/S0043-1354(01)00167-1.

Ahn, H., and J. D. Salas (1997), Groundwater Head Sampling based on Stochastic Analysis, *Water Resour. Res.*, *33*(12), 2769–2780, doi:10.1029/97WR02187.

Akaike, H. (1974), A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, *19*(6), 716–723, doi:10.1109/TAC.1974.1100705.

Alavi, N., J. S. Warland, and A. A. Berg (2006), Filling Gaps in Evapotranspiration Measurements for Water Budget Studies: Evaluation of a Kalman Filtering Approach, *Agricultural and Forest Meteorology*, *141*(1), 57–66, doi:10.1016/j.agrformet.2006.09.011.

Alvera-Azcárate, A., A. Barth, J.-M. Beckers, and R. H. Weisberg (2007), Multivariate Reconstruction of Missing Data in Sea Surface Temperature, Chlorophyll, and Wind Satellite Fields, *J. Geophys. Res.*, *112*(C3), C03008, doi:10.1029/2006JC003660.

Alzraiee, A. H., D. A. Baù, and L. A. Garcia (2013), Multiobjective Design of Aquifer Monitoring Networks for Optimal Spatial Prediction and Geostatistical Parameter Estimation, *Water Resour. Res.*, *49*(6), 3670–3684, doi:10.1002/wrcr.20300.

Ashraf, M., J. C. Loftis, and K. G. Hubbard (1997), Application of Geostatistics to Evaluate Partial Weather Station Networks, *Agricultural and Forest Meteorology*, *84*(3), 255–271, doi:10.1016/S0168-1923(96)02358-1.

von Asmuth, J. R., K. Maas, M. Knotters, M. F. P. Bierkens, M. Bakker, T. N. Olsthoorn, D. G. Cirkel, I. Leunk, F. Schaars, and D. C. von Asmuth (2012), Software for Hydrogeologic Time Series Analysis, Interfacing Data with Physical Insight, *Environmental Modelling & Software*, *38*, 178–190, doi:10.1016/j.envsoft.2012.06.003.

Baù, D., M. Ferronato, G. Gambolati, P. Teatini, and A. Alzraiee (2015), Ensemble Smoothing of Land Subsidence Measurements for Reservoir Geomechanical Characterization, *Int. J. Numer. Anal. Meth. Geomech.*, *39*(2), 207–228, doi:10.1002/nag.2309.

Bennis, S., F. Berrada, and N. Kang (1997), Improving Single-Variable and Multivariable Techniques for Estimating Missing Hydrological Data, *Journal of Hydrology*, *191*(1–4), 87–105, doi:10.1016/S0022-1694(96)03076-4.

Berendrecht, W. L., A. W. Heemink, F. C. van Geer, and J. C. Gehrels (2003), Decoupling of Modeling and Measuring Interval in Groundwater Time Series Analysis based on Response Characteristics, *Journal of Hydrology*, *278*(1–4), 1–16, doi:10.1016/S0022-1694(03)00075-1.

Bierkens, M. F. P., M. Knotters, and F. C. van Geer (1999), Calibration of Transfer Function–Noise Models to Sparsely or Irregularly Observed Time Series, *Water Resour. Res.*, *35*(6), 1741–1750, doi:10.1029/1999WR900083.

Bierkens, M. F. P., M. Knotters, and T. Hoogland (2001), Space-Time Modeling of Water Table Depth using a Regionalized Time Series Model and the Kalman Filter, *Water Resour. Res.*, *37*(5), 1277–1290, doi:10.1029/2000WR900353.

Bowels, D., and W. J. Grenney (1978), Estimation of Diffuse Loading of Water Quality Pollutants by Kalman Filtering, pp. 581–597.

Box, G. E. P., and G. M. Jenkins (1976), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.

Box, G. E. P., and D. A. Pierce (1970), Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, *Journal of the American Statistical Association*, *65*(332), 1509–1526, doi:10.1080/01621459.1970.10481180.

CDM-Smith (2013), *South Platte Decision Support System Alluvial Groundwater Model Report. April 2013. Prepared for the Colorado Water Conservation Board and Colorado Division of Water Resources*, CDM-Smith.

Chan, P. K., and P. D. Loucks (1978), Application of the Stochatic Output Regulator Theory to Optimally Route the Stormwater through a Combined Sewer Network, pp. 745–768.

Chang, F.-J., L.-C. Chang, C.-W. Huang, and I.-F. Kao (2016), Prediction of Monthly Regional groundwater Levels through Hybrid Soft-Computing Techniques, *Journal of Hydrology*, *541*, *Part B*, 965–976, doi:10.1016/j.jhydrol.2016.08.006.

Changnon, S. A., F. A. Huff, and C.-F. Hsu (1988), Relations between Precipitation and Shallow Groundwater in Illinois, *J. Climate*, *1*(12), 1239–1250, doi:10.1175/1520-0442(1988)001<1239:RBPASG>2.0.CO;2.

Chiu, C.-L., and E. O. Isu (1978), Application of Kalman Filter to Open Channel Flow Profile Estimation, pp. 441–458.

Clark, M. P., D. E. Rupp, R. A. Woods, X. Zheng, R. P. Ibbitt, A. G. Slater, J. Schmidt, and M. J. Uddstrom (2008), Hydrological Data Assimilation with the Ensemble Kalman Filter: Use of Streamflow Observations to Update States in a Distributed Hydrological Model, *Advances in Water Resources*, *31*(10), 1309–1324, doi:10.1016/j.advwatres.2008.06.005.

Colorado Decision Support Systems (2016), GIS Data for South Platte, Available from: http://cdss.state.co.us/GIS/Pages/Division1SouthPlatte.aspx (Accessed 5 August 2016)

Colorado Department of Local Affairs (2016), Population Data, Available from: https://demography.dola.colorado.gov/population/

Colorado General Assembly (2012), *South Platte River Alluvial Aquifer Study*.

Colorado General Assembly (2015a), *Emergency Well Pumping Damaging High Groundwater*.

Colorado General Assembly (2015b), *Invasive Phreatophyte Grant Program*.

Colorado General Assembly (2015c), *Pilot Project based on the South Platte Aquifer Study HB 12-1278 Recommendations*.

Crestani, E., M. Camporese, D. A. Baù, and P. Salandin (2013), Ensemble Kalman Filter versus Ensemble Smoother for Assessing Hydraulic Conductivity via Tracer Test Data Assimilation, *Hydrol. Earth Syst. Sci.*, *17*(4), 1517–1531, doi:10.5194/hess-17-1517-2013.

Daliakopoulos, I. N., P. Coulibaly, and I. K. Tsanis (2005), Groundwater Level Forecasting using Artificial Neural Networks, *Journal of Hydrology*, *309*(1–4), 229–240, doi:10.1016/j.jhydrol.2004.12.001.

Dee, D. P. (1991), Simplification of the Kalman Filter for Meteorological Data Assimilation, *Q.J.R. Meteorol. Soc.*, *117*(498), 365–384, doi:10.1002/qj.49711749806.

Dennehy, K. F., D. W. Litke, C. M. Tate, and J. S. Heiny (1993), South Platte River Basin - Colorado, Nebraska, and Wyoming, *JAWRA Journal of the American Water Resources Association*, *29*(4), 647–683, doi:10.1111/j.1752-1688.1993.tb03231.x.

Duong, N., R.-M. Li, and Y. H. Chen (1978), Adaptive Control of Lock and Dam Gate Opening Using a Kalman Filter for Real-Time identification of Upstream-Downstream Stage Relationship, pp. 745–768.

Durdu, Ö. F. (2010), Application of Linear Stochastic Models for Drought Forecasting in the Büyük Menderes River Basin, Western Turkey, *Stoch Environ Res Risk Assess*, *24*(8), 1145–1162, doi:10.1007/s00477-010-0366-3.

Eigbe, U., M. B. Beck, H. S. Wheater, and F. Hirano (1998), Kalman Filtering in Groundwater Flow Modelling: Problems and Prospects, *Stochastic Hydrology and Hydraulics*, *12*(1), 15–32, doi:10.1007/s004770050007.

Elshorbagy, A., S. P. Simonovic, and U. S. Panu (2002), Estimation of Missing Streamflow Data using Principles of Chaos Theory, *Journal of Hydrology*, *255*(1–4), 123–133, doi:10.1016/S0022-1694(01)00513-3.

Evensen, G. (1994), Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model using Monte Carlo Methods to Forecast Error Statistics, *J. Geophys. Res.*, *99*(C5), 10143–10162, doi:10.1029/94JC00572.

Evensen, G. (1997), Advanced Data Assimilation for Strongly Nonlinear Dynamics, *Mon. Wea. Rev.*, *125*(6), 1342–1354, doi:10.1175/1520-0493(1997)125<1342:ADAFSN>2.0.CO;2.

Evensen, G. (2003), The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation, *Ocean Dynamics*, *53*(4), 343–367, doi:10.1007/s10236-003-0036-9.

Evensen, G., and P. J. van Leeuwen (2000), An Ensemble Kalman Smoother for Nonlinear Dynamics, *Mon. Wea. Rev.*, *128*(6), 1852–1867, doi:10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

Gelb, A. (Ed.) (1974), *Applied Optimal Estimation*, MIT Press, Cambridge, Mass.

Gove, J. H., and D. Y. Hollinger (2006), Application of a Dual Unscented Kalman Filter for Simultaneous State and Parameter Estimation in Problems of Surface-Atmosphere Exchange, *J. Geophys. Res.*, *111*(D8), D08S07, doi:10.1029/2005JD006021.

Graham, W. D., and C. D. Tankersley (1993), Forecasting Piezometric Head Levels in the Floridan Aquifer: A Kalman Filtering Approach, *Water Resour. Res.*, *29*(11), 3791–3800, doi:10.1029/93WR01813.

Grønnevik, R., and G. Evensen (2001), Application of Ensemble-based Techniques in Fish Stock Assessment, *Sarsia*, *86*(6), 517–526, doi:10.1080/00364827.2001.10420490.

Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000), *MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model - User guide to Modularization Concepts and the Ground-Water Flow Process*, Open-File Report, U.S. Geological Survey, Reston, Virginia.

Herrera, G. S., and G. F. Pinder (2005), Space-Time Optimization of Groundwater Quality Sampling Networks, *Water Resour. Res.*, *41*(12), W12407, doi:10.1029/2004WR003626.

Hipel, K. W., and A. I. McLeod (1994), *Time Series Modelling of Water Resources and Environmental Systems*, Developments in water science 45, Elsevier, Amsterdam ; New York.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen (2005), Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations, *Monthly Weather Review*, *133*(3), 604–620, doi:10.1175/MWR-2864.1.

Hyndman, R. J. (2016), *forecast: Forecasting Functions for Time Series and Linear Models*.

Hyndman, R. J., and Y. Khandakar (2008), Automatic Time Series Forecasting: The forecast Package for R, *Journal of Statistical Software*, *26*(3), 1–22.

Irvine, K. N., and A. J. Eberhardt (1992), Multiplicative, Seasonal Arima Models for Lake Erie and Lake Ontario Water Levels, *JAWRA Journal of the American Water Resources Association*, *28*(2), 385–396, doi:10.1111/j.1752-1688.1992.tb04004.x.

Jarvis, A. J., V. J. Stauch, K. Schulz, and P. C. Young (2004), The Seasonal Temperature Dependency of Photosynthesis and Respiration in Two Deciduous Forests, *Global Change Biology*, *10*(6), 939–950, doi:10.1111/j.1529-8817.2003.00743.x.

Javaheri, A., and M. Babbar-Sebens (2014), Remote Sensing Data Assimilation In Water Quality Numerical Model Of Eagle Creek Reservoir Using Ensemble Kalman Filter Method, in *11th International Conference on Hydroinformatics, HIC 2014*, New York, NY.

Julier, S. J., and J. K. Uhlmann (1997), New Extension of the Kalman Filter to Nonlinear Systems, in *AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, vol. 3068, pp. 182–193.

Julier, S. J., and J. K. Uhlmann (2004), Unscented Filtering and Nonlinear Estimation, *Proceedings of the IEEE*, *92*(3), 401–422, doi:10.1109/JPROC.2003.823141.

Jung, D., and K. Lansey (2015), Water Distribution System Burst Detection Using a Nonlinear Kalman Filter, *Journal of Water Resources Planning and Management*, *141*(5), 04014070, doi:10.1061/(ASCE)WR.1943-5452.0000464.

Jung, D., Y. H. Choi, and J. H. Kim (2016), Optimal Node Grouping for Water Distribution System Demand Estimation, *Water*, *8*(4), 160, doi:10.3390/w8040160.

Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, *Journal of basic Engineering*, *82*(1), 35–45.

Kalman, R. E., and R. S. Bucy (1961), New Results in Linear Filtering and Prediction Theory, *J. Basic Eng*, *83*(1), 95–108, doi:10.1115/1.3658902.

Karavitis, C. A., C. G. Vasilakou, D. E. Tsesmelis, P. D. Oikonomou, N. A. Skondras, D. Stamatakos, V. Fassouli, and S. Alexandris (2015), Short-Term Drought Forecasting Combining Stochastic and Geo-Statistical Approaches, *European Water*, *49*, 43–63.

Karlsson, S., and M. Löthgren (2000), On the Power and Interpretation of Panel Unit Root Tests, *Economics Letters*, *66*(3), 249–255, doi:10.1016/S0165-1765(99)00237-2.

Keppenne, C. L., and M. M. Rienecker (2002), Initial Testing of a Massively Parallel Ensemble Kalman Filter with the Poseidon Isopycnal Ocean General Circulation Model, *Mon. Wea. Rev.*, *130*(12), 2951–2965, doi:10.1175/1520-0493(2002)130<2951:ITOAMP>2.0.CO;2.

Kim, B. S., S. Z. Hossein, and G. Choi (2011), Evaluation of Temporal-Spatial Precipitation Variability and Prediction using Seasonal ARIMA Model in Mongolia, *KSCE J Civ Eng*, *15*(5), 917–925, doi:10.1007/s12205-011-1097-9.

Kitanidis, P. K. (1976), A Unified Approach to the Parameter Estimation of Groundwater Models, M.Sc. Thesis, Massachusetts Institute of Technology.

Kitanidis, P. K., and R. L. Bras (1979), Collinearity and Stability in the Estimation of Rainfall-Runoff Model Parameters, *Journal of Hydrology*, *42*(1), 91–108, doi:10.1016/0022-1694(79)90008-8.

Kitanidis, P. K., and R. L. Bras (1980), Real-Time Forecasting with a Conceptual Hydrologic Model: 1. Analysis of uncertainty, *Water Resour. Res.*, *16*(6), 1025–1033, doi:10.1029/WR016i006p01025.

Knotters, M., and M. F. P. Bierkens (2000), Physical Basis of Time Series Models for Water Table Depths, *Water Resour. Res.*, *36*(1), 181–188, doi:10.1029/1999WR900288.

Kollat, J. B., P. M. Reed, and R. M. Maxwell (2011), Many-Objective Groundwater Monitoring Network Design using Bias-Aware Ensemble Kalman filtering, Evolutionary Optimization, and Visual Analytics, *Water Resour. Res.*, *47*(2), W02529, doi:10.1029/2010WR009194.

Kondrashov, D., and M. Ghil (2006), Spatio-Temporal Filling of Missing Points in Geophysical Data Sets, *Nonlin. Processes Geophys.*, *13*(2), 151–159, doi:10.5194/npg-13-151-2006.

Koutsoyiannis, D., and A. Langousis (2011), Precipitation, in *Treatise on Water Science*, vol. Volume 2: The Science of Hydrology, edited by P. Wilderer and S. Uhlenbrook, pp. 27–77, Elsevier, Oxford.

Kuligowski, R. J., and A. P. Barros (1998), Using Artificial Neural Networks to Estimate Missing Rainfall Data, *JAWRA Journal of the American Water Resources Association*, *34*(6), 1437–1447, doi:10.1111/j.1752-1688.1998.tb05443.x.

Lee, J. H. W., J. Q. Mao, and K. W. Choi (2009), The Extended Kalman Filter for Short Term Prediction of Algal Bloom Dynamics, in *Advances in Water Resources and Hydraulic Engineering. Proceedings of 16th IAHR-APD Congress and 3rd Symposium of IAHR-ISHS*, edited by C. Zhang and H. Tang, pp. 513–517, Springer Berlin Heidelberg.

van Leeuwen, P. J., and G. Evensen (1996), Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation, *Mon. Wea. Rev.*, *124*(12), 2898–2913, doi:10.1175/1520-0493(1996)124<2898:DAAIMI>2.0.CO;2.

Lei, F., C. Huang, H. Shen, and X. Li (2014), Improving the Estimation of Hydrological States in the SWAT Model via the Ensemble Kalman Smoother: Synthetic Experiments for the Heihe River Basin in northwest China, *Advances in Water Resources*, *67*, 32–45, doi:10.1016/j.advwatres.2014.02.008.

Li, R.-M., Nguyen Duong, and D. M. Simons (1978), Application of the Kalman Filter for Prediction of Stage-Discharge Relationships in Rivers, pp. 459–471.

Makhuvha, T., G. Pegram, R. Sparks, and W. Zucchini (1997), Patching Rainfall Data Using Regression Methods. 2. Comparisons of Accuracy, Bias and Efficiency, *Journal of Hydrology*, *198*(1–4), 308–318, doi:10.1016/S0022-1694(96)03283-0.

McLaughlin, D. (1976), Application of Kalman Filtering to Groundwater Basin Modeling and Prediction, *Real-Time Forecasting/Control of Water Resource Systems, IIASA Proc. Ser*, 109–123.

McLaughlin, D. B. (1978), Potential Applications of Kalman Filtering Concepts to Groundwater Basin Management, pp. 639–655.

Meinhold, R. J., and N. D. Singpurwalla (1983), Understanding the Kalman Filter, *The American Statistician*, *37*(2), 123–127, doi:10.2307/2685871.

Miller, R. N., M. Ghil, and F. Gauthiez (1994), Advanced Data Assimilation in Strongly Nonlinear Dynamical Systems, *J. Atmos. Sci.*, *51*(8), 1037–1056, doi:10.1175/1520-0469(1994)051<1037:ADAISN>2.0.CO;2.

Mirzavand, M., and R. Ghazavi (2014), A Stochastic Modelling Technique for Groundwater Level Forecasting in an Arid Environment Using Time Series Methods, *Water Resour Manage*, *29*(4), 1315–1328, doi:10.1007/s11269-014-0875-9.

Oikonomou, P. D., J. A. Kallenberger, R. M. Waskom, K. K. Boone, E. N. Plombon, and J. N. Ryan (2016), Water Acquisition and Use during Unconventional Oil and Gas Development and the Existing Data Challenges: Weld and Garfield counties, CO, *Journal of Environmental Management*, *181*, 36–47, doi:10.1016/j.jenvman.2016.06.008.

Okeya, I., Z. Kapelan, C. Hutton, and D. Naga (2014), Online Burst Detection in a Water Distribution System Using the Kalman Filter and Hydraulic Modelling, *Procedia Engineering*, *89*, 418–427, doi:10.1016/j.proeng.2014.11.207.

Papamichail, D. M., and P. E. Georgiou (2001), Seasonal Arima Inflow Models for Reservoir Sizing, *JAWRA Journal of the American Water Resources Association*, *37*(4), 877–885, doi:10.1111/j.1752-1688.2001.tb05519.x.

Pappas, C., S. M. Papalexiou, and D. Koutsoyiannis (2014), A Quick Gap Filling of Missing Hydrometeorological Data, *J. Geophys. Res. Atmos.*, *119*(15), 2014JD021633, doi:10.1002/2014JD021633.

Pastres, R., S. Ciavatta, and C. Solidoro (2003), The Extended Kalman Filter (EKF) as a Tool for the Assimilation of High Frequency Water Quality Data, *Ecological Modelling*, *170*(2–3), 227–235, doi:10.1016/S0304-3800(03)00230-8.

Rodriguez-Iturbe, I., J. B. Valdes, and J. M. Velasquez (1978), Applications of Kalman Filter in Rainfall-Runoff Studies, pp. 233–253.

Salas, J. D. (1993), Analysis and Modeling of Hydrologic Time Series, in *Handbook of Hydrology*, edited by D. Maidment, p. 19.1–19.72, McGraw-Hill, New York.

Salas, J. D., J. W. Delleur, V. Yevjevich, and W. L. Lane (1980), *Applied Modeling of Hydrologic Time Series*, Water Resources Publications.

Schneider, T. (2001), Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values, *J. Climate*, *14*(5), 853–871, doi:10.1175/1520-0442(2001)014<0853:AOICDE>2.0.CO;2.

Schrader, B. P., and S. F. Moore (1977), Kalman Filtering in Water Quality Modeling: Theory vs. Practice, in *Proceedings of the 9th Conference on Winter Simulation - Volume 2*, pp. 504–510, Winter Simulation Conference, Gaitersburg, MD.

Schreider, S. Y., P. C. Young, and A. J. Jakeman (2001), An application of the Kalman Filtering Technique for Streamflow Forecasting in the Upper Murray Basin, *Mathematical and Computer Modelling*, *33*(6), 733–743, doi:10.1016/S0895-7177(00)00276-4.

Sen, Z., A. Altunkaynak, and M. Özger (2004), Sediment Concentration and Its Prediction by Perceptron Kalman Filtering Procedure, *Journal of Hydraulic Engineering*, *130*(8), 816–826, doi:10.1061/(ASCE)0733-9429(2004)130:8(816).

Shirmohammadi, B., M. Vafakhah, V. Moosavi, and A. Moghaddamnia (2012), Application of Several Data-Driven Techniques for Predicting Groundwater Level, *Water Resour Manage*, *27*(2), 419–432, doi:10.1007/s11269-012-0194-y.

Simons, D. M., N. Duong, and R.-M. Li (1978), An Approach to Short-Term Water and Sediment Discharge Prediction Using Kalman Filter, pp. 459–471.

Sun, H., and M. Koch (2001), Case Study: Analysis and Forecasting of Salinity in Apalachicola Bay, Florida, Using Box-Jenkins ARIMA Models, *Journal of Hydraulic Engineering*, *127*(9), 718–727, doi:10.1061/(ASCE)0733-9429(2001)127:9(718).

Szollosi-Nagy, A. (1976), An Adaptive Identification and Prediction Algorithm for the Real-Time Forcasting of Hydrological Time Series, *Hydrological Sciences Bulletin*, *21*(1), 163–176, doi:10.1080/02626667609491613.

Teegavarapu, R. S. V. (2007), Use of Universal Function Approximation in Variance-Dependent Surface Interpolation Method: An Application in Hydrology, *Journal of Hydrology*, *332*(1–2), 16–29, doi:10.1016/j.jhydrol.2006.06.017.

Teegavarapu, R. S. V. (2012), Spatial Interpolation using Nonlinear Mathematical Programming Models for Estimation of Missing Precipitation Records, *Hydrological Sciences Journal*, *57*(3), 383–406, doi:10.1080/02626667.2012.665994.

Teegavarapu, R. S. V., and V. Chandramouli (2005), Improved Weighting Methods, Deterministic and Stochastic Data-Driven Models for Estimation of Missing Precipitation Records, *Journal of Hydrology*, *312*(1–4), 191–206, doi:10.1016/j.jhydrol.2005.02.015.

Toth, E., A. Montanari, and A. Brath (1999), Real-Time Flood Forecasting via Combined Use of Conceptual and Stochastic Models, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, *24*(7), 793–798, doi:10.1016/S1464-1909(99)00082-9.

Uddameri, V. (2006), Using Statistical and Artificial Neural Network Models to Forecast Potentiometric Levels at a Deep Well in South Texas, *Environ Geol*, *51*(6), 885–895, doi:10.1007/s00254-006-0452-5.

Von Asmuth, J. R., K. Maas, M. Bakker, and J. Petersen (2008), Modeling Time Series of Ground Water Head Fluctuations Subjected to Multiple Stresses, *Ground Water*, *46*(1), 30–40, doi:10.1111/j.1745-6584.2007.00382.x.

Walker, J. P., and P. R. Houser (2001), A Methodology for Initializing Soil Moisture in a Global Climate Model: Assimilation of Near-Surface Soil Moisture Observations, *Journal of Geophysical Research*, *106*(D11), 11761–11774.

Wan, E. A., and R. Van der Merwe (2001), The Unscencted Kalman Filter, in *Kalman Filtering and Neural Networks*, edited by S. S. Haykin, pp. 221–282, Wiley, New York, NY.

Wang, H. R., C. Wang, X. Lin, and J. Kang (2014), An Improved ARIMA Model for Precipitation Simulations, *Nonlin. Processes Geophys.*, *21*(6), 1159–1168, doi:10.5194/npg-21-1159-2014.

Wang, W., K. Chau, D. Xu, and X.-Y. Chen (2015), Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition, *Water Resour Manage*, *29*(8), 2655–2675, doi:10.1007/s11269-015-0962-6.

Waskom, R. M. (2013), *Report to the Colorado Legislature: HB12-1278 Study of the South Platte River Alluvial Aquifer*, Completion Report No. 226, Colorado Water Institute, Fort Collins, CO.

Waskom, R. M., and P. D. Oikonomou (2014), Improving Groundwater Management in the South Platte Alluvial Aquifer, in *2014 NGWA Groundwater Summit, Denver, CO, May 04 - 07*.

Willmott, C. J., and K. Matsuura (2005), Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance, *Climate research*, *30*(1), 79.

Wilson, J., P. K. Kitanidis, and M. Dettinger (1978), State and Parameter Estimation in Groundwater Models, pp. 657–679.

Wu, X.-L., X.-H. Xiang, C.-H. Wang, X. Chen, C.-Y. Xu, and Z. Yu (2013), Coupled Hydraulic and Kalman Filter Model for Real-Time Correction of Flood Forecast in the Three Gorges Interzone of Yangtze River, China, *Journal of Hydrologic Engineering*, *18*(11), 1416–1425, doi:10.1061/(ASCE)HE.1943-5584.0000473.

Xie, X., and D. Zhang (2010), Data Assimilation for Distributed Hydrological Catchment Modeling via Ensemble Kalman Filter, *Advances in Water Resources*, *33*(6), 678–690, doi:10.1016/j.advwatres.2010.03.012.

Ye, G., and R. A. Fenner (2011), Kalman Filtering of Hydraulic Measurements for Burst Detection in Water Distribution Systems, *Journal of Pipeline Systems Engineering and Practice*, *2*(1), 14–22, doi:10.1061/(ASCE)PS.1949-1204.0000070.

Yi, M.-J., and K.-K. Lee (2004), Transfer Function-Noise Modelling of Irregularly Observed Groundwater Heads using Precipitation Data, *Journal of Hydrology*, *288*(3–4), 272–287, doi:10.1016/j.jhydrol.2003.10.020.

Yuan, X., Z. Xie, and M. Liang (2008), Spatiotemporal Prediction of Shallow Water Table Depths in Continental China, *Water Resour. Res.*, *44*(4), W04414, doi:10.1029/2006WR005453.

Zhang, Y., G. F. Pinder, and G. S. Herrera (2005), Least Cost Design of Groundwater Quality Monitoring Networks, *Water Resour. Res.*, *41*(8), W08412, doi:10.1029/2005WR003936.

Zhou, Y., D. McLaughlin, and D. Entekhabi (2006), Assessing the Performance of the Ensemble Kalman Filter for Land Surface Data Assimilation, *Mon. Wea. Rev.*, *134*(8), 2128–2142, doi:10.1175/MWR3153.1.

5 Conclusions and Recommendations

5.1 Data and Integrated River Basin Management

The methodologies presented in this research cover a range of water resources planning and management activities and could have significant implications in Integrated Water Resources Management efforts. Their unified thread is the common quest for IRBM under changing water resources conditions. The IRBM framework requires high quality information in order to account for the various activities and transformations taking place in a river basin, including their interdependencies. Technology has revolutionized the technical tools of Integrated River Basin Management but in spite of this development there are still challenges to be addressed for improving the implementation of IRBM. One of the main challenges is the development of innovative ways to surpass obstacles imposed by data deficiencies in order to extract the most information possible for assisting decision making.

Even though the South Platte is considered a data rich area with advanced management schemes compared with other parts of the world, the methodological approach in the three case studies reviled areas of concern which are classified in Table 13. South Platte is characterized by a semi-arid climate with limited resources where there is a significant increase of economic activity over the last decades and a constantly expanding population. The water of the basin is considered to be over appropriated as South Platter river is considered to be one of the hardest working rivers of the west. Economic development and the increasing population has put further stress on the already natural water scarcity. As a result, there has been changes in land uses and shifts in water demands trends over the decades. One can argue that these pressures on water resources have triggered the different water administration eras implemented in the basin as a means of problem solving. While the developed infrastructure has served the purposes it was built for, climate uncertainty along with the changing demand setting call for an integrated assessment of future vulnerabilities and alternative management schemes.

Table 13: Categories of concern deriving from the three case studies

River Basin	**Data**	**Transformation to Information**	**Integrated River Basin Management**
Limited water resources, Climatic uncertainty, Over-appropriated water resources, Increasing population, Rapid changes in land uses, Shifting water demands, Infrastructure developed for different pressures	Fragmented, Quality Questions, Non-user friendly access, Non-existing, Inconsistent data collection intervals, Time gaps in monitoring programs, Inappropriate data collection protocols, Not fully cover the spatial extent	No framework to combine data, Complexity, Wrong data used, Models in not appropriate scale for local issues, Lack of multi-scale data analysis approaches	Sub-optimal conjunctive use of surface water and groundwater resources, Non-holistic ways of vulnerability drought assessment, Non-integrated drought impact efforts, Conflicting views derived from limited information, Non-defined water trade-offs for oil & gas development

Under conditions of rapid socioeconomic and natural transformations, IRBM is an adaptive management style where the basin system, consisting of interacting natural sub-systems and anthropogenic sub-systems, is monitored and frequently the system status is reevaluated. Despite having such a need in most river basins in the world, implementation of an adaptive framework is hindered due to data issues and the ways of transforming imperfect data to meaningful information.

Likewise, the cases in this research studied in the South Platte River Basin revealed an array of existing data hindrances including issues of data fragmentation; data quality questions; non-monitored variables; time gaps in monitoring programs; infrequent sampling; limited spatial coverage etc. At the same time, there are concerns about the extracted information from data in order to support the decision-making process under the IRBM framework, since South Platte is a highly complex basin with diverse human activities. Imperfect/incorrect data are often used; the existing models are not scale appropriate to provide solutions in local issues; there is a lack of multi-scale analysis approaches; and a lack of a framework to combine multi-source and diverse data.

Table 14: Realities in South Platte River basin applying IRBM and the premises of an ideal scheme of IRBM.

Practice of IRBM	**Ideal Premises of IRBM**
Non-integrated drought vulnerability practices.	Holistic drought risk assessment and risk management.
Absence of a plan that incorporates fully water for energy (unconventional oil and gas sub-sector not included).	Cross-sectoral integration between water use sub-sectors.
Unconventional oil and gas flowback water has a relative small percentage of reuse.	Integrating water and wastewater management to increase the available resources.
Non-optimal conjunctive use of surface water and groundwater.	Integration of surface water and groundwater management.

In all three case studies, common data hindrances were identified based on the classification of Figure 3 (see Chapter 1). The following Figure 44 illustrates these categories and links the case studies with IRBM technical instruments where the information can be used to improve the implementation closer to the ideal premises of IRBM. Table 14 illustrates the comparison of the main ideal premises of IRBM and the practical issues in the South Platte River Basin as presented in this work. This allows elaboration of the specific contribution areas towards key premises of IRBM.

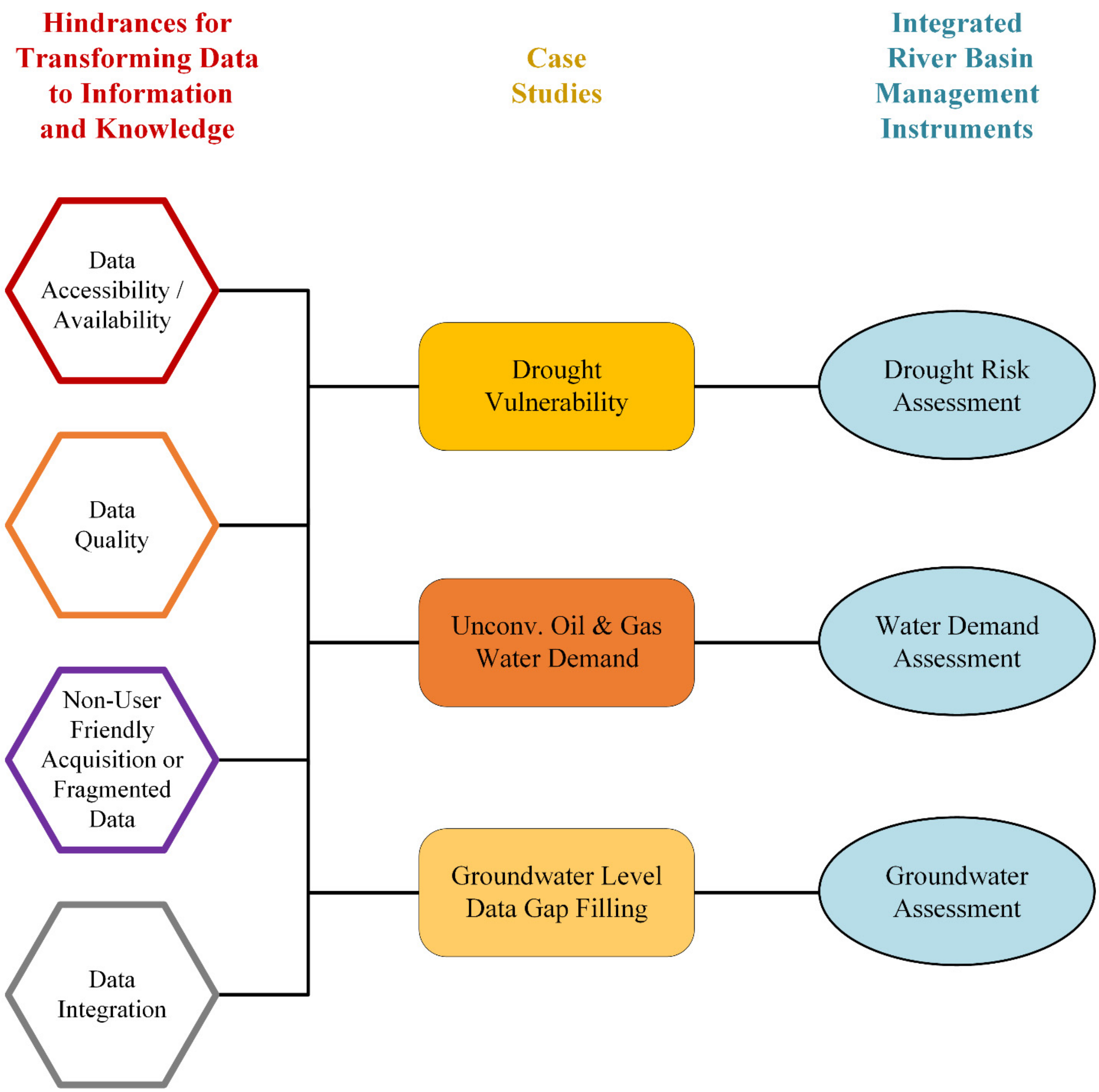


Figure 44: Common data hindrances identified in all three case studies and the direct links of the case studies to technical management instruments/tools in IRBM

5.2 Conclusions

This research effort attempts through the described methodologies to overcome three main challenges, namely: integrated drought vulnerability assessment; unconventional oil and gas demand assessment; and reconstructing groundwater level time series towards IRBM. This was achieved through the illustrated problem solving methodologies integrating data in order to transform them to information and thus to contribute towards bridging the gap among ideal conditions of IRBM and constraints in practice, as portrayed in Figure 45.

Drought events are further challenging the ability of a water resources system to meet demands. The development of a framework that advances the quantification of drought vulnerability based on an integrated approach including aspects of the physical, structural, and socio-economic components was the first objective. The identification of the vulnerability of a system to drought events in a quantitative manner and their delineation was presented. The study evolved the SDVI closer to an operational phase and at the same time tests the index in a geographical local very different than other previous applications. In this context, the objective was achieved.

In a setting with limited resources like the South Platte basin, the finite water resources of river basins are claimed by competing and conflicting uses and users. Understanding shifts in water demands in a basin is essential for updating water management plans and strategies in order to balance water supply and demand. The development of oil and gas shale formations in the US introduced an additional water user which now competes with traditional users for the available water. The characteristics of this emerging water demand for energy development and the factors affecting it are not exactly known since it is a relatively new activity and the reported information is incomplete. Since the second objective was to produce a multi-scale methodology for quantifying and analyzing water demand for unconventional oil and gas development, through the above presented processes this objective was also achieved.

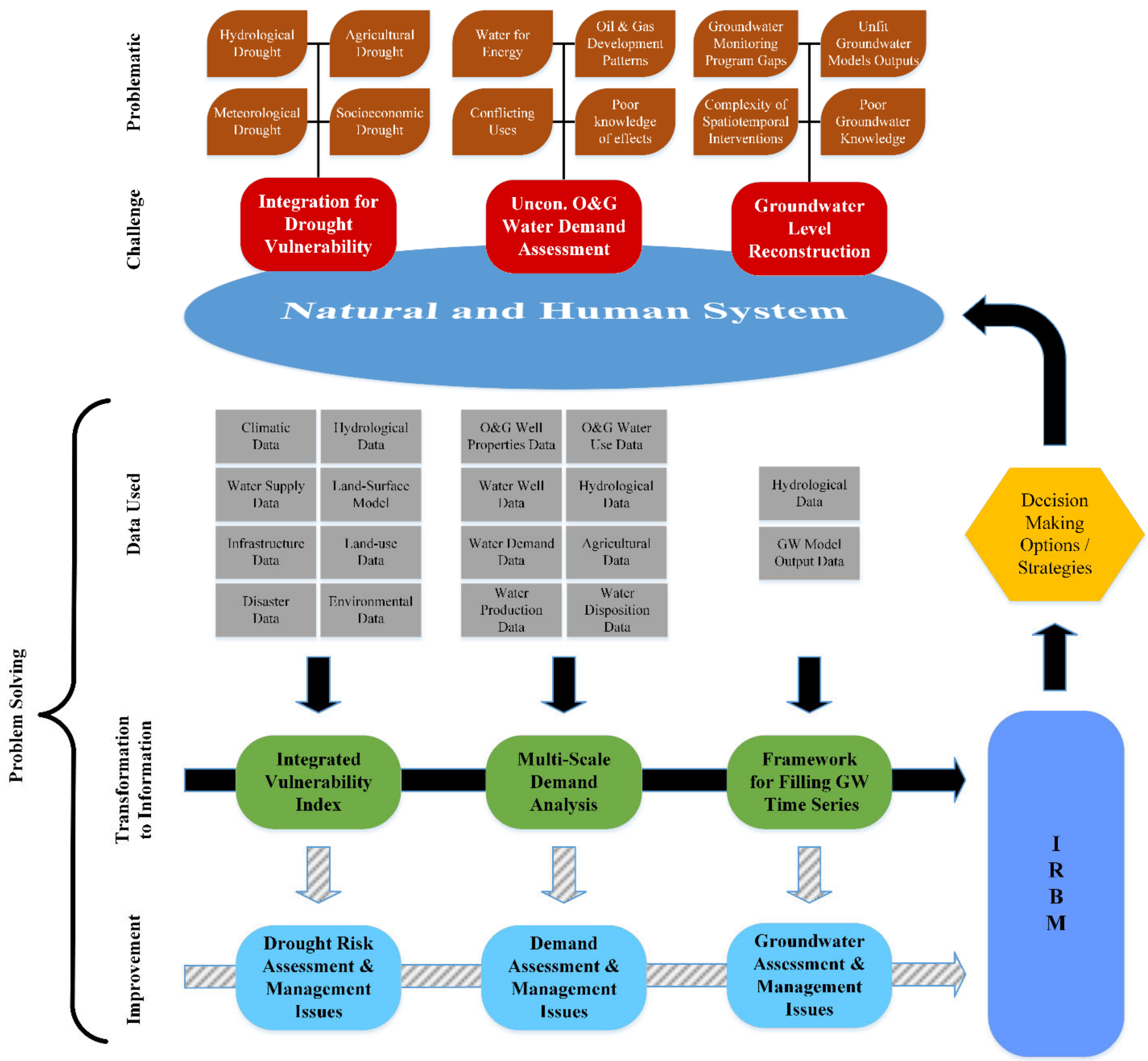


Figure 45: The studied challenges and the problem-solving scheme for integrating data and transform them to information

Groundwater information when compared with surface water information is significantly less available. The difficulties in monitoring groundwater systems has resulted in data gaps. The importance of knowledge of water level variations caused by natural and manmade events is a key for modeling efforts and data driven analysis. Lack of such information leads to sub-optimal use of the groundwater resources. Especially during extreme dry conditions, alluvial groundwater resources are vulnerable but at the same time they can be a great supplement source for mitigating some of the multifaceted drought impacts. The conjunctive use of groundwater and surface water resources is a key premise of IRBM. The last objective of this research was the development and testing of a novel methodology for bridging data gaps in groundwater level measurements in dynamic alluvial formations, by taking advantage of available outputs of regional groundwater physical models in a framework that includes statistical modeling and the Ensemble Smoother. Hence, the presented process shows that this last objective was also achieved.

Overall, a summary of the results and key points from the completed research are presented below:

1. The SDVI calculated with the proposed methodology was proved appropriate for portraying the levels of vulnerability to drought in a fine spatial scale. The drought event of 2012 during the months of July, August and September were used as an example for evaluating the datasets used for the representation of the index's components, which were considered successful since when cross-checking was possible it matched the observations. A contribution of the methodology is an integrated way to portray the spatiotemporal extent and magnitude of the multifaceted drought impacts incorporating a plethora of different data. The integration of the other index components results to delineating vulnerability levels based on societal, physical and structural factors. The SDVI values produced for the South Platte basin seems to offer a deeper understanding of vulnerability of the different system's components.

2. The study of the emergent water demand for unconventional energy reveals that the different oil and gas development practices and water acquisition approaches followed in Weld and Garfield counties in Colorado, is affecting water use intensity. The comparison of the two Colorado counties showed that the challenges are local and thus solutions should be tailored for the specific area of interest in order to achieve sustainable water management. There is an urgent need for more data to be collected in order to understand fully the whole life cycle of this water demand. There were data discrepancies between the available databases. Furthermore, data reported to FracFocus combine total water volume per well (freshwater, produced water, or recycled water), which makes it impossible to determine separate volumes for each element. Finally, volumes of water used in secondary activities such as dust suppression, drilling mud, and site restoration, although insignificant to the amount used for hydraulic fracturing, should be also reported.
3. Three numerical experiments were designed to evaluate the prediction performance of the proposed groundwater data-filling framework for 18 well locations with biannual measurements. The first two numerical experiments were very satisfactory. The algorithm encountered difficulty in predicting a sudden change in head measurement when there were no similar measurements in temporal proximity. Furthermore, if the model was not approximating the general tendency or the seasonality of the measured values, then the prediction error increased. On the contrary, Numerical Experiment #3 was the least successful. It was anticipated to perform less good than the previous two, since in experiment #3, a sequential random five-year window was selected and the observations that lie within that period were treated as unknowns. The framework

should be a valuable tool for enhancing groundwater time series in alluvial formations, for both intermittent missing data and short spanned continuous gaps.

Giving such conclusion in the following section some recommendations are presented towards IRBM implementation.

5.3 Recommendations

While the methodologies proposed have tangible results towards bridging the gap between ideal and real implementation of IRBM under changing conditions, areas of potential recommendations have been identified that could lead in extending further the outcomes of this work.

Regarding the SDVI, the first recommendation is related to the uncertainty in the input values for the index calculation, which is one of the uncertainty sources affecting the results. The difficulty of validating the spatial information and in many cases the limitations of spatial coverage in needed data hinders the use of the index especially in a data scarce area. One promising solution to this problem is using satellite derived information to supplement needed data for the components. Hard measured data are equally important for calculating the index but if there are limited data points available they can be used as validation points. Relying more on satellite data, will lead to a more operational version of the index with the ability to inform about drought vulnerability conditions in near-real time. In addition, the incorporation of demand and supply components for reservoirs would lead to more precise assignment of overall drought vulnerability levels regarding to ecological and societal aspects.

The recommendations based on the research of the Chapter 3 are mainly related to data limitations and quality since they were the main hindrance. Hence, the availability of additional data parameters would greatly improve the understanding of water use for oil and gas development in Colorado. The source of water used in hydraulic fracturing should be included in the reporting as well as the flowback water reported should be associated with a fixed number of days. Suggestions of creating a seamless database with no demarcated responsibilities from the pertinent authorities is needed along with a user friendly way to identify water rights of nontributary water associated with oil and gas activities. Introducing an automated system for reporting with stricter quality criteria accompanied with appropriate training would help minimize future data gaps in future reports. Also, water volumes of water used in secondary activities such as dust suppression, drilling mud, and site restoration, although insignificant to the amount used for hydraulic fracturing, should be also reported. This work could be extended by building a hydrologic models to examine the potential stresses on water resources under drought conditions considering future scenarios for oil and gas industry's water demand, based on the presented findings. Finally, investigating recycle and reuse pathways for flowback and produced water, after appropriate treatment, could lead in cost-effective solutions to reduce industry's water requirements or reveal other beneficial alternative water uses.

The recommendations regarding the proposed groundwater data-filling framework are considering alternative schemes and techniques. The evaluation of the performance of the framework in a piecewise manner by using different temporal local SARIMAX models for different periods. Such an approach might produce better results. Another interesting expansion of this work is the use of clustering methods with the purpose of constructing SARIMAX models for the sections of the aquifer with similar processes. The benefit of such an approach would be to prevent the need to construct a new SARIMAX model at every point of interest. Lastly, the performance of different smoother algorithms, such as the Ensemble Kalman Smoother (EnKS), should be evaluated. With the employment of other Kalman smoothing frameworks, it would be possible to test the incorporation of non-linear stochastic models, which could perhaps result in improving the prediction accuracy.

5.4 Epilogue

All in all, the quest to increase the degree of integration in IRBM efforts through efficient and effective transformation of data to information and knowledge is continuous since the dynamic nature of the study systems call for adaptive approaches in water management. The methodologies developed and presented herein can be part of the technical armory that water resources engineers and managers have towards IRBM. The findings and recommendations of this research effort are subject to localities and data availability constraints. However, the developed methodologies can serve as a guide to other neighboring basins and by extension to other areas of the world facing similar problems, where they may be adapted to each locale.

APPENDICES

Appendix A. Precipitation Stations Metadata

#	Station Name	UTM X	UTM Y	Elev. (m)	Period	Status
1	Akron Washington CO Arpt.	651857.9	4447781	1421.28	1963-2012	Filled
2	Antero Reservoir	422760.9	4316411	2718.82	1963-2012	Complete
3	Bailey	458966.4	4361797	2356.10	1963-2012	Filled
4	Boulder	477231.9	4426892	1671.52	1963-2012	Filled
5	Briggsdale	556945.4	4498458	1473.40	1963-2012	Filled
6	Byers 5 ENE	574758.4	4399287	1554.48	1963-2012	Complete
7	Cheesman	475976.4	4341250	2097.02	1963-2012	Filled
8	Cheyenne Muni. Arpt.	515380.8	4555425	1868.42	1963-2012	Complete
9	Denver Stapleton Int. Arpt.	511182.9	4401498	1611.17	1963-2012	Filled
10	Estes Park	458758.4	4469679	2279.90	1963-2012	Combined
11	Fleming 3 SW	680987.7	4501919	1297.23	1963-2012	Combined + Filled
12	Fort Collins	488893.4	4495995	1525.22	1963-2012	Complete
13	Fort Morgan	600725.2	4457288	1328.62	1963-2012	Filled
14	Greeley UNC	525533.4	4472443	1437.13	1963-2012	Combined
15	Julesburg	729613.9	4540858	1057.35	1963-2012	Filled
16	Kassler	491813.6	4371159	1702.92	1963-2012	Filled
17	Lake George 8 SW	459200.3	4306617	2606.04	1963-2012	Filled
18	Leroy 9 WSW	662532.3	4483878	1386.84	1963-2012	Filled
19	Longmont 2 ESE	494321.8	4444408	1508.76	1963-2012	Filled
20	New Raymer 21 N	595318.5	4531814	1578.86	1963-2012	Combined
21	Sedalia 4 SSE	504115.5	4361567	1821.18	1963-2012	Filled
22	Sedgwick 5 S	709311.4	4526085	1216.15	1963-2012	Filled
23	Sterling	651524.3	4498995	1211.28	1963-2012	Filled
24	Waterdale	482160.7	4475018	1594.10	1963-2012	Filled

Precipitation Time Series Fill-in Details

1. <u>Akron Washington CO Arpt.</u>

   Filled from Akron 4 E station with coefficient of determination 0.824

2. <u>Antero Reservoir</u>

   It had a complete time series

3. <u>Bailey</u>

   Filled from Grant station with coefficient of determination 0.741

4. <u>Boulder</u>

   Filled from Longmont 2 ESE station with coefficient of determination 0.723

5. <u>Briggsdale</u>

   Filled from Fort Morgan, Greeley-Combined and New Raymer-Combined stations with coefficient of determination 0.693

6. <u>Byers 5 ENE</u>

   It had a complete time series

7. <u>Cheesman</u>

   Filled from Bailey and Lake George 8 SWstations with coefficient of determination 0.763

8. <u>Cheyenne Muni. Arpt.</u>

   It had a complete time series

9. <u>Denver Stapleton Int. Arpt.</u>

   Filled from Denver International Airport station with coefficient of determination 0.748

10. <u>Estes Park</u>

    It was combined with Estes Park 1 SSE

11. <u>Fleming 3 SW</u>

    It was combined with Fleming and then filled from Sedgwick 5 S and Sterling stations with coefficient of determination 0.775

12. Fort Collins

    It was complete

13. Fort Morgan

    Filled from Akron Washington Airport, Greeley-Combined, Briggsdale and Sterling station with coefficient of determination 0.737

14. Greeley UNC

    It was combined with Greeley station

15. Julesburg

    Filled from Big Springs (NE) station with coefficient of determination 0.775

16. Kassler

    Filled from Strontia Sp. Dam station with coefficient of determination 0.778

17. Lake George 8 SW

    Filled from Florissant Fossil Beds station with coefficient of determination 0.740

18. Leroy 9 WSW

    Filled from Fleming-Combined and Sterling stations with coefficient of determination 0.797

19. Longmont 2 ESE

    Filled from Boulder, Fort Collins, Greeley-Combined and Waterdale stations with coefficient of determination 0.815

20. New Raymer 21 N

    It was combined with Kauffman 4 SSE

21. Sedalia 4 SSE

    Filled from Castle Rock station with coefficient of determination 0.697

22. Sedgwick 5 S

    Filled from Crook station with coefficient of determination 0.731

23. Sterling

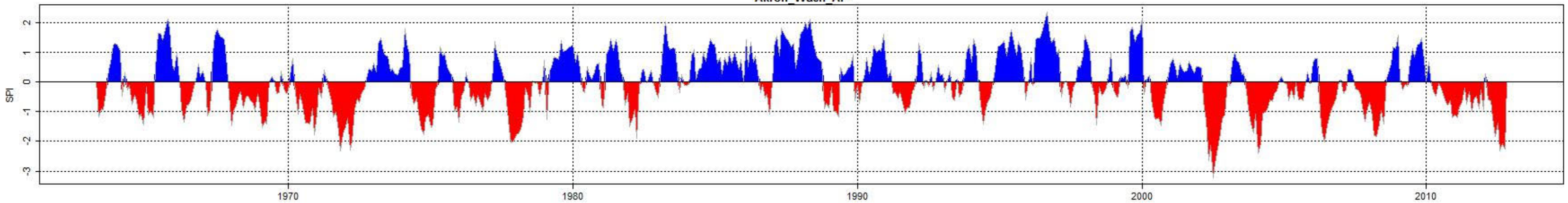
Filled from Fleming-Combined, Leroy 9 WSW and New Raymer-Combined stations with coefficient of determination 0.737

24. Waterdale

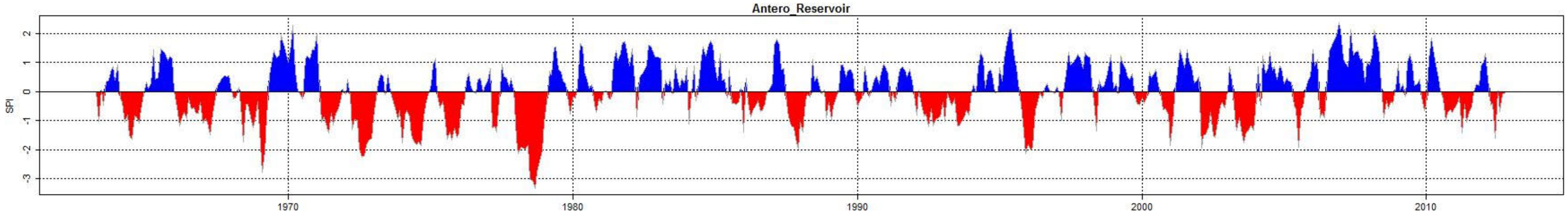
Filled from Flat Iron Reservoir and Fort Collins stations with coefficient of determination 0.776

Appendix B. SPI-6 Graphs

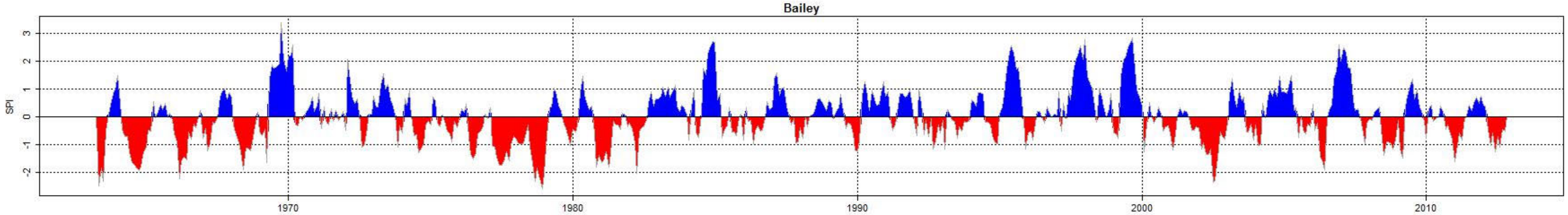
Akron_Wash_AP



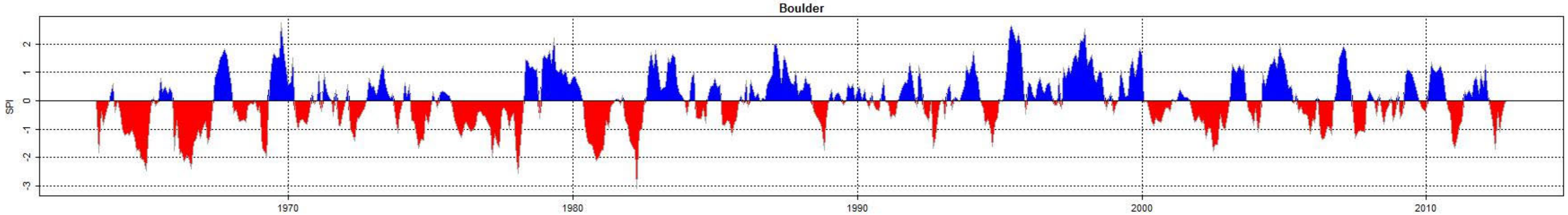
Antero_Reservoir

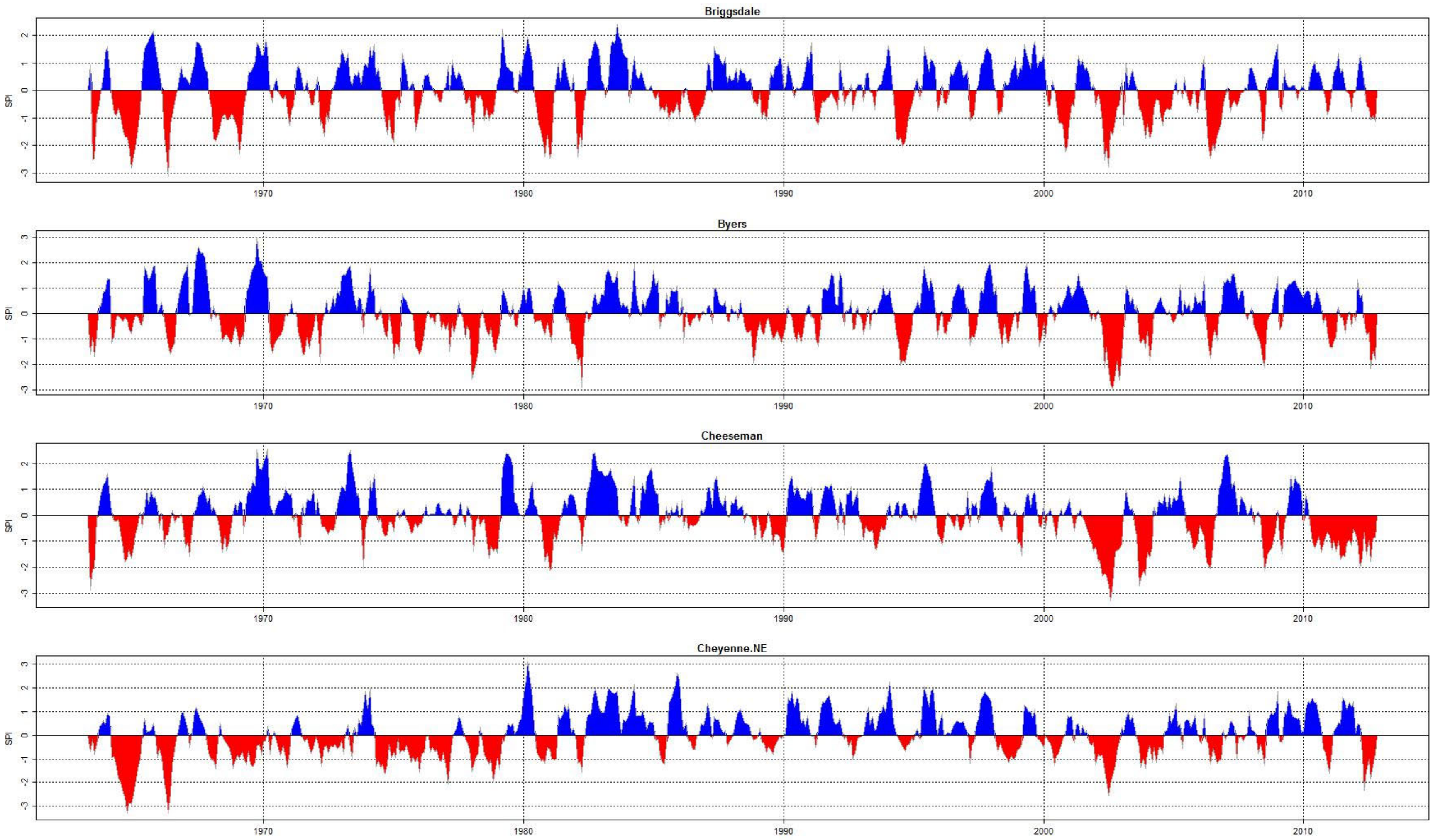


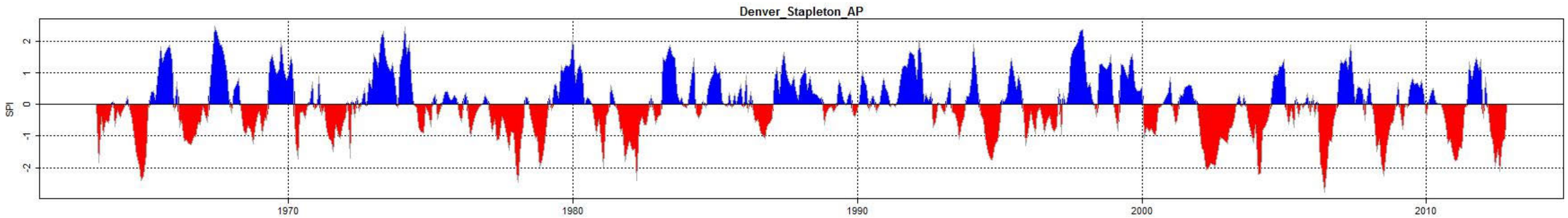
Bailey

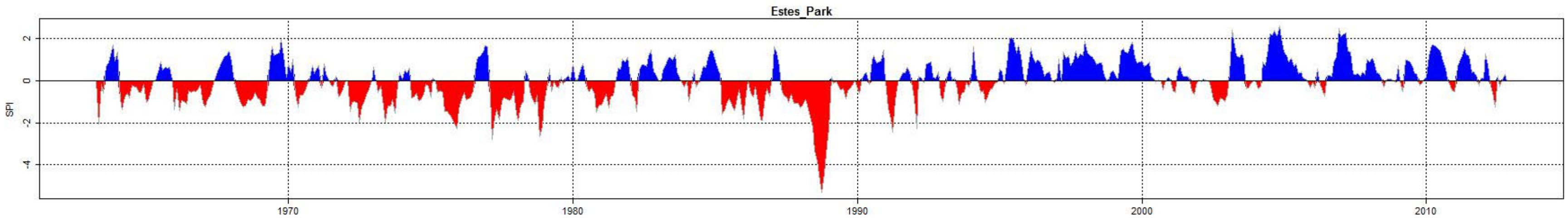


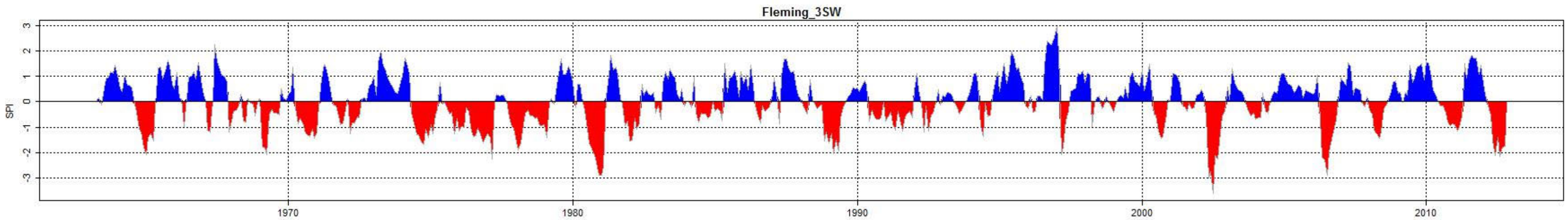
Boulder

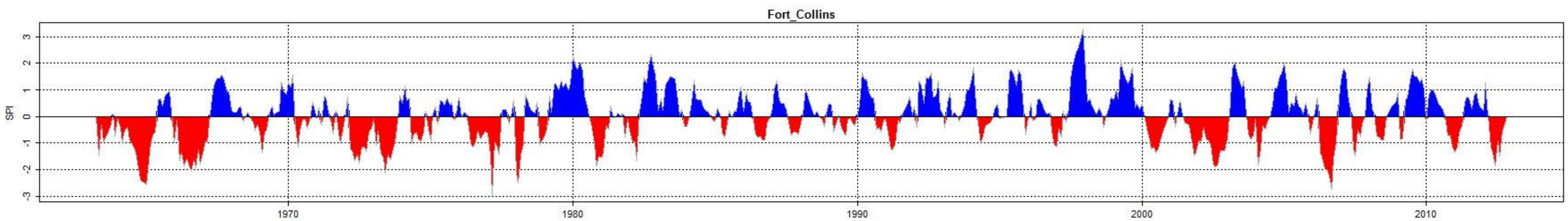


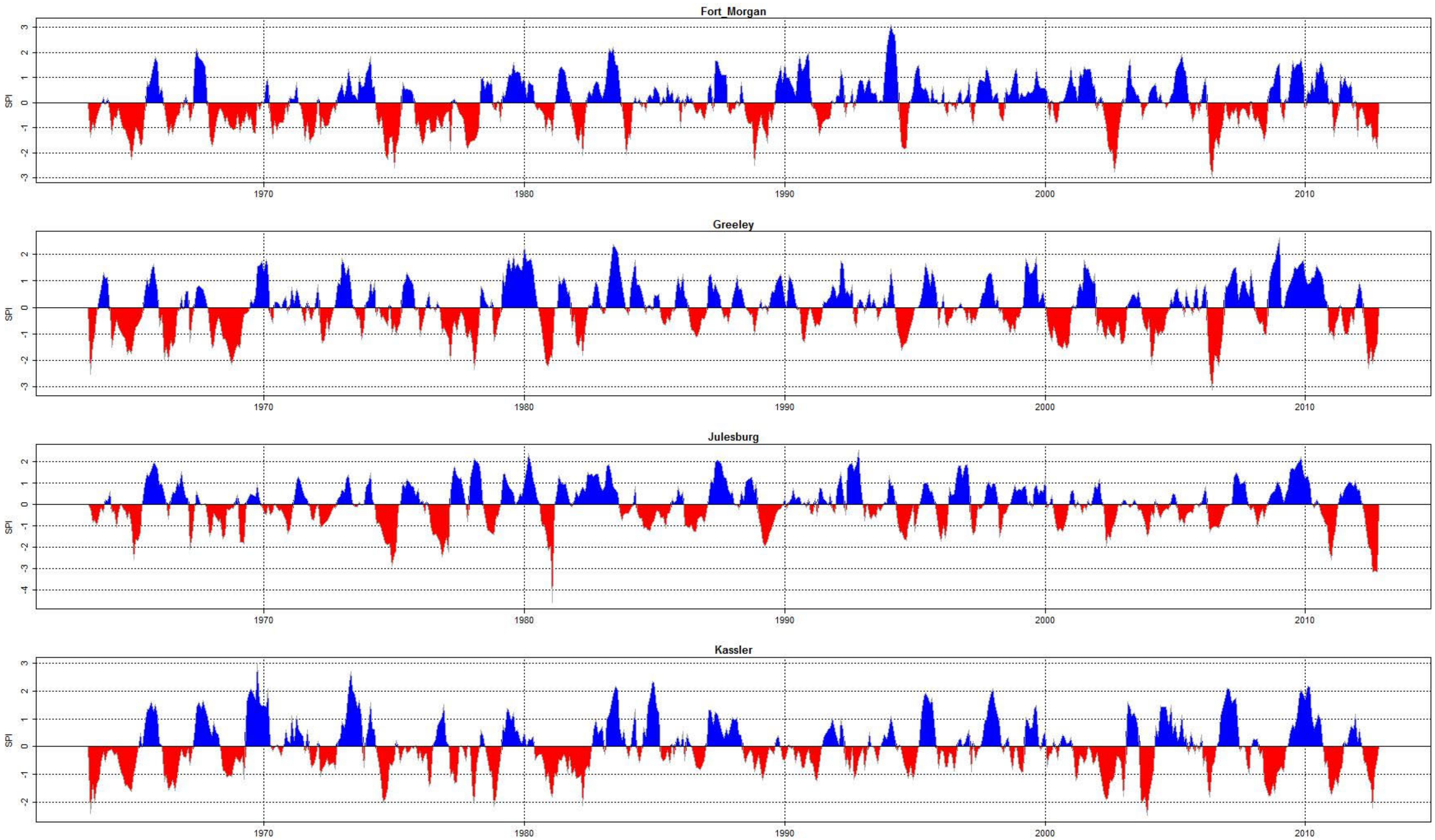

Briggsdale
Byers
Cheeseman
Cheyenne.NE
SPI
1970
1980
1990
2000
2010


Denver_Stapleton_AP
SPI
2
1
0
-1
-2
1970
1980
1990
2000
2010

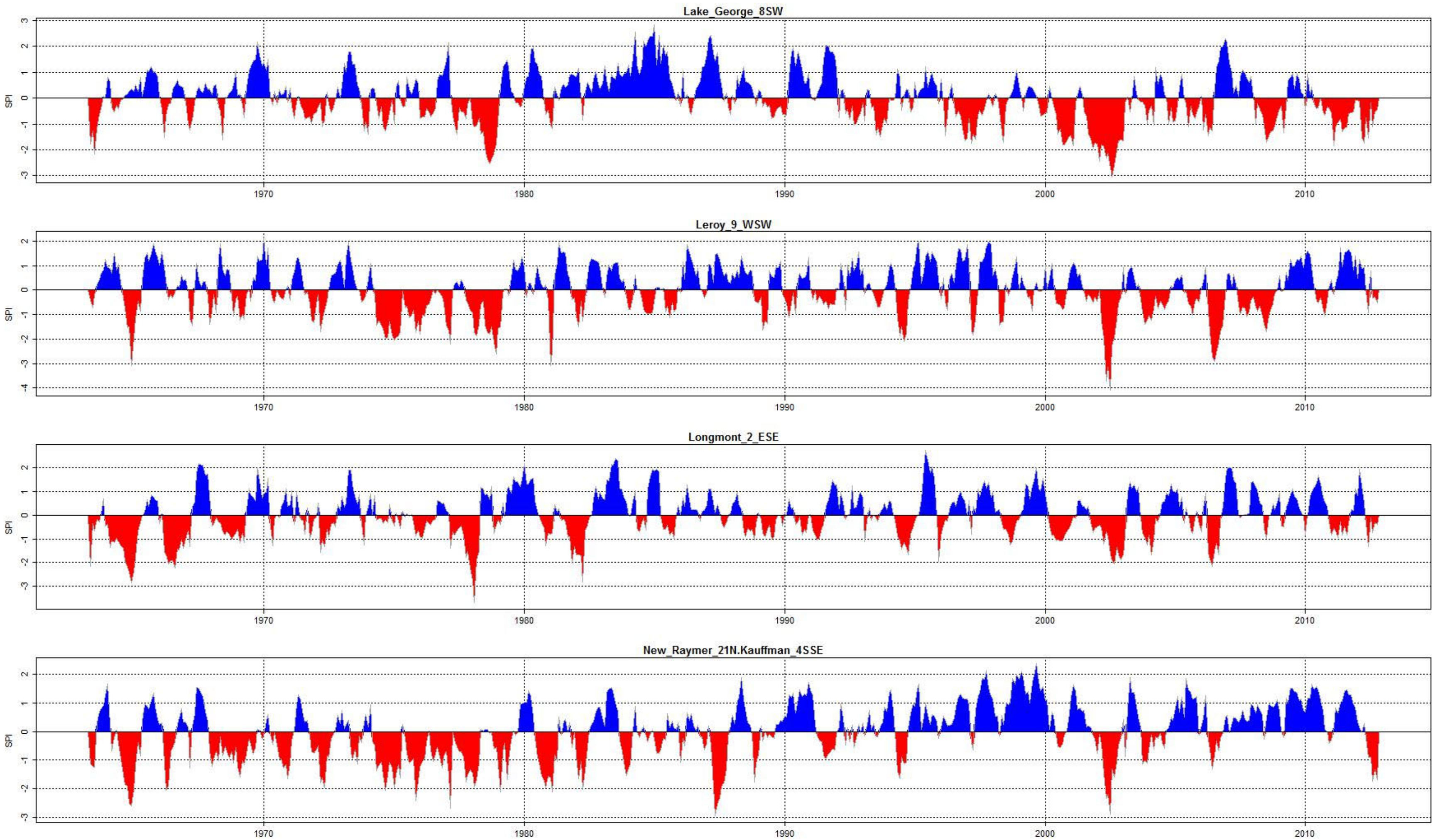

Estes_Park
SPI
2
0
-2
-4
1970
1980
1990
2000
2010


Fleming_3SW
SPI
3
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010

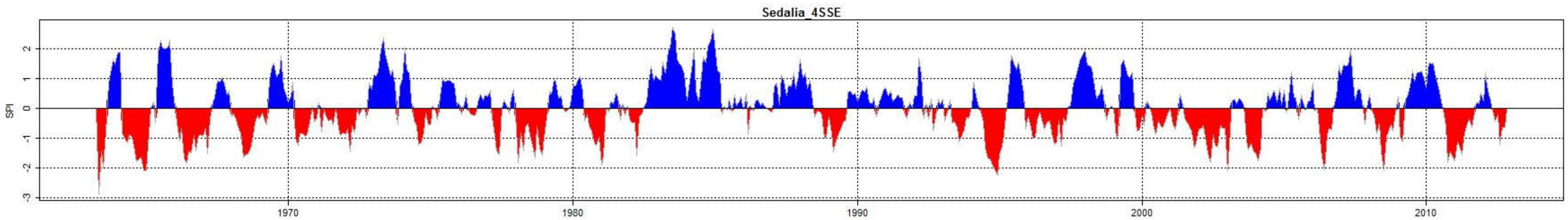

Fort_Collins
SPI
3
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010

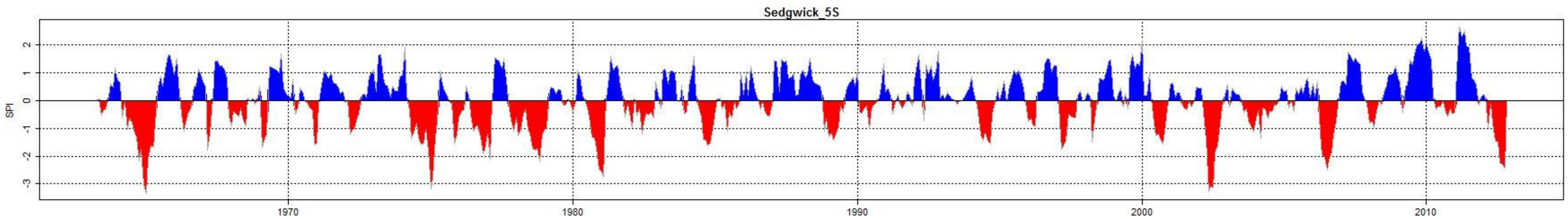


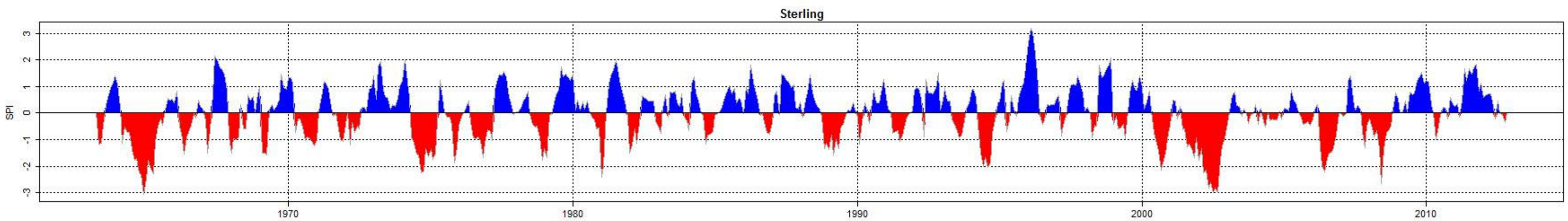
Fort_Morgan
Greeley
Julesburg
Kassler
SPI
1970
1980
1990
2000
2010

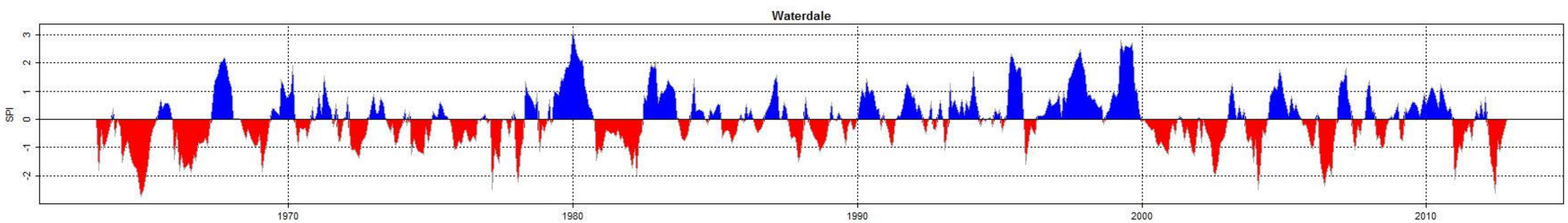


Lake_George_8SW
SPI
3
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010
Leroy_9_WSW
SPI
2
1
0
-1
-2
-3
-4
1970
1980
1990
2000
2010
Longmont_2_ESE
SPI
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010
New_Raymer_21N.Kauffman_4SSE
SPI
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010

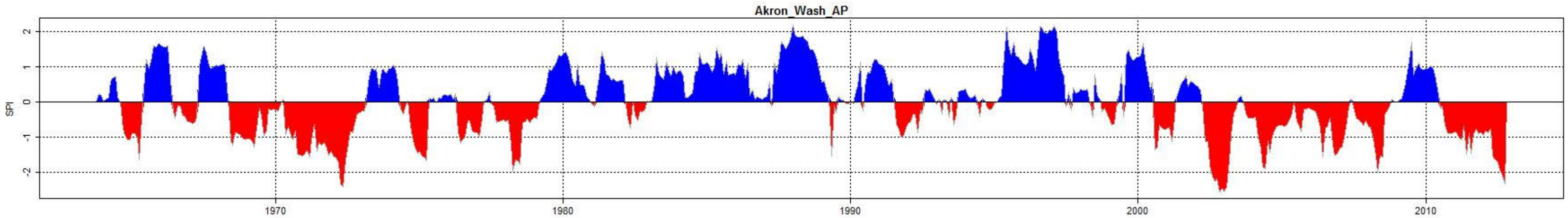

Sedalia_4SSE
SPI
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010

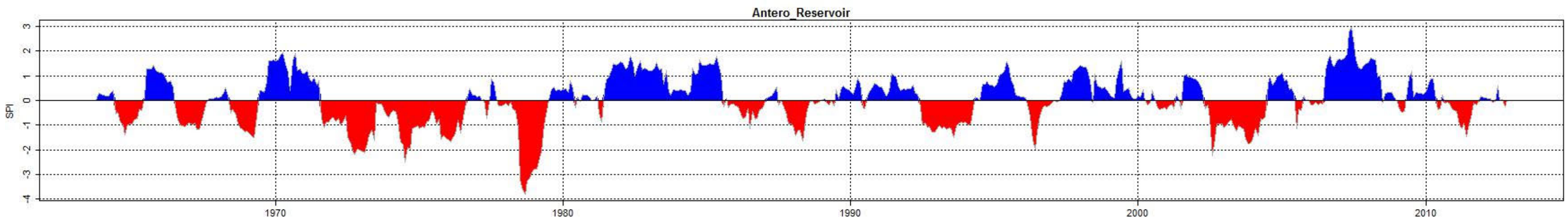

Sedgwick_5S
SPI
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010

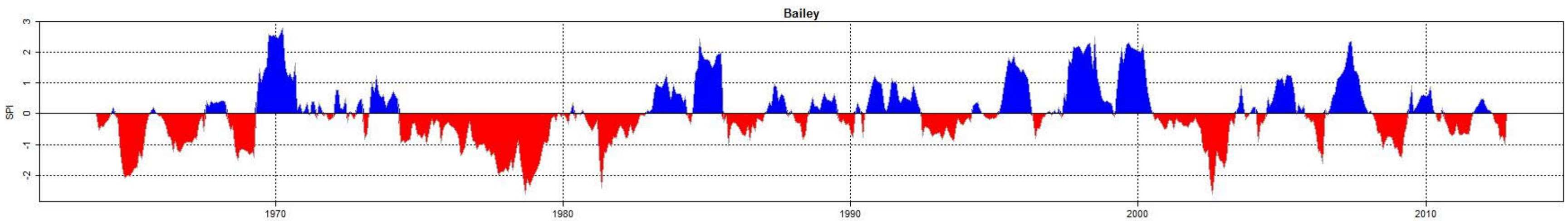

Sterling
SPI
3
2
1
0
-1
-2
-3
1970
1980
1990
2000
2010

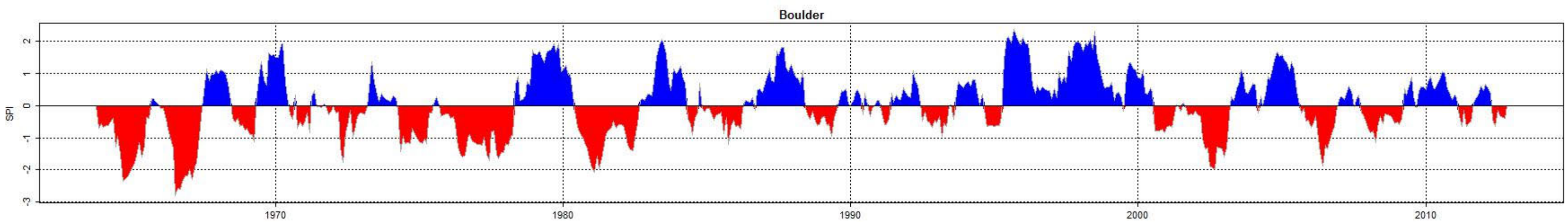

Waterdale
SPI
3
2
1
0
-1
-2
1970
1980
1990
2000
2010

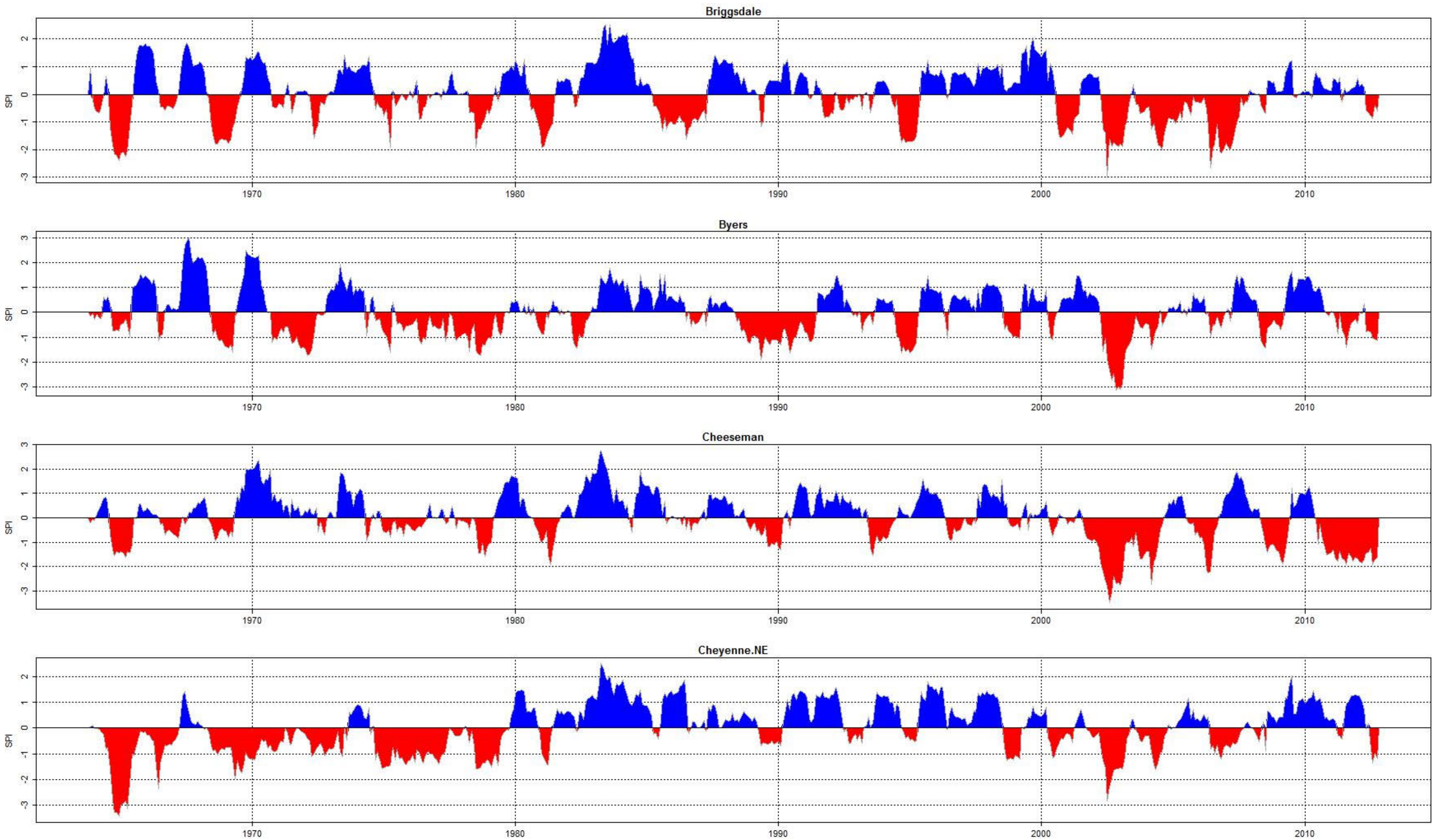
Appendix C. SPI-12 Graphs



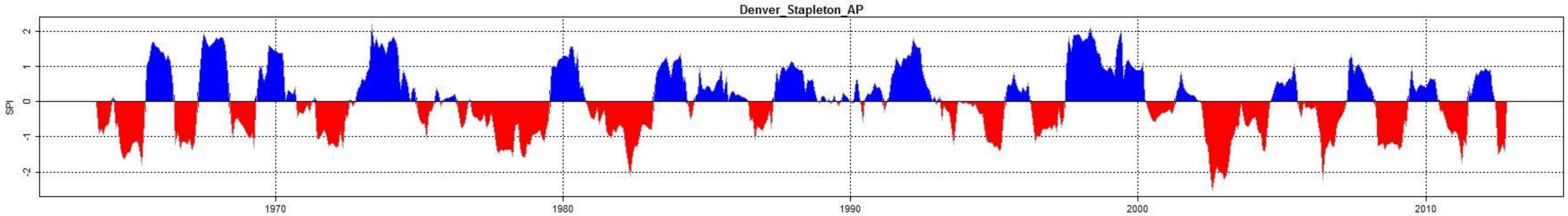


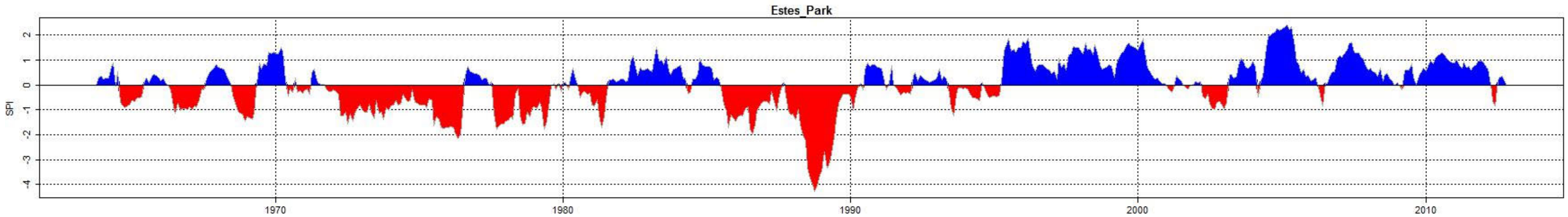


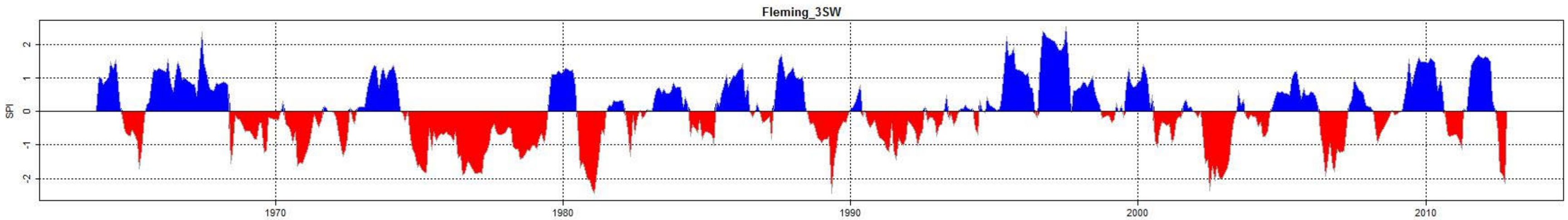


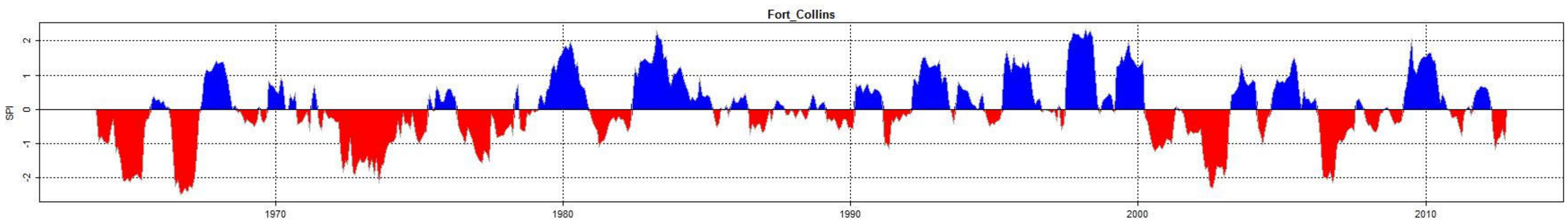


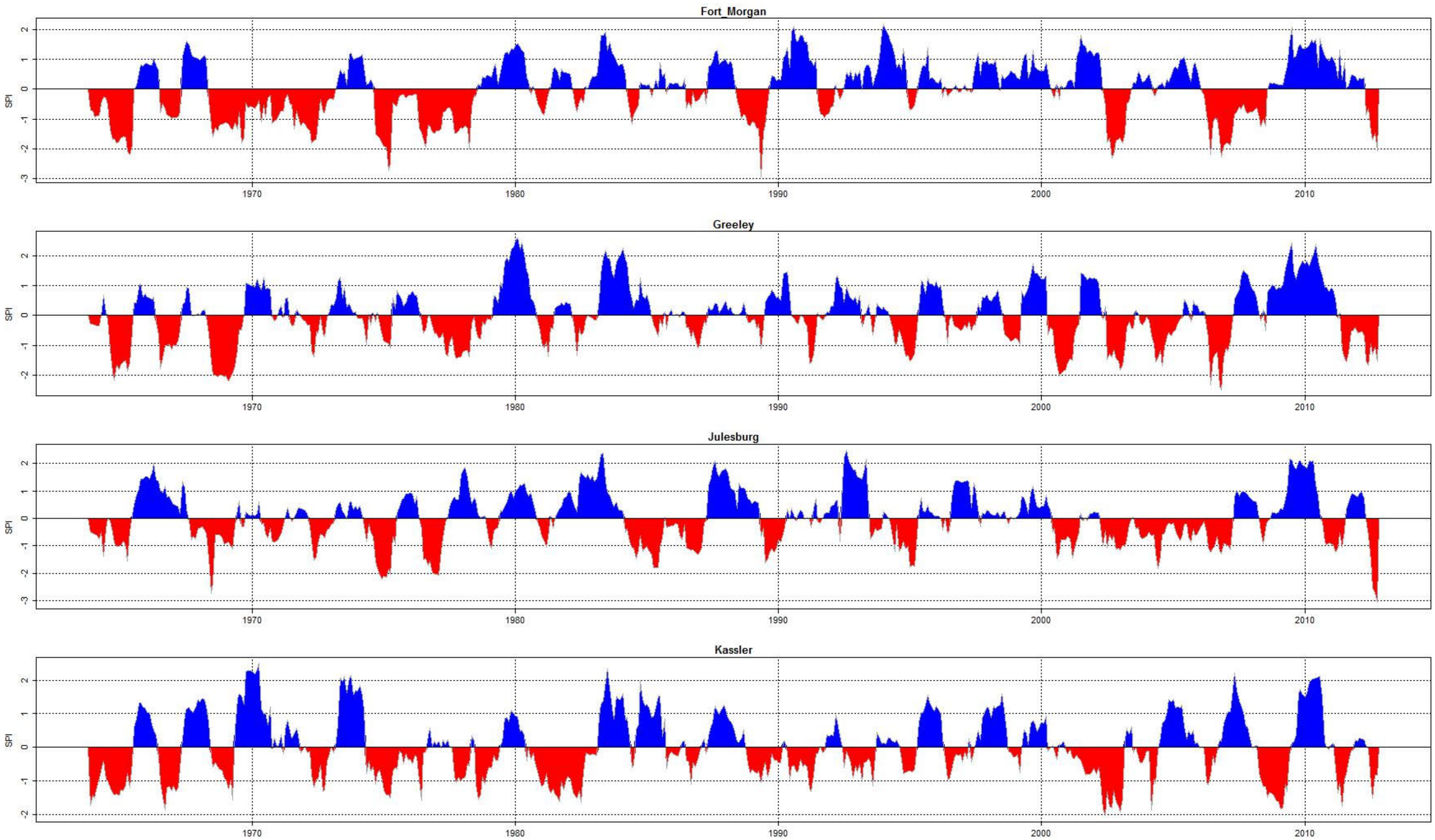
Briggsdale
SPI
1970
1980
1990
2000
2010
Byers
SPI
1970
1980
1990
2000
2010
Cheeseman
SPI
1970
1980
1990
2000
2010
Cheyenne.NE
SPI
1970
1980
1990
2000
2010

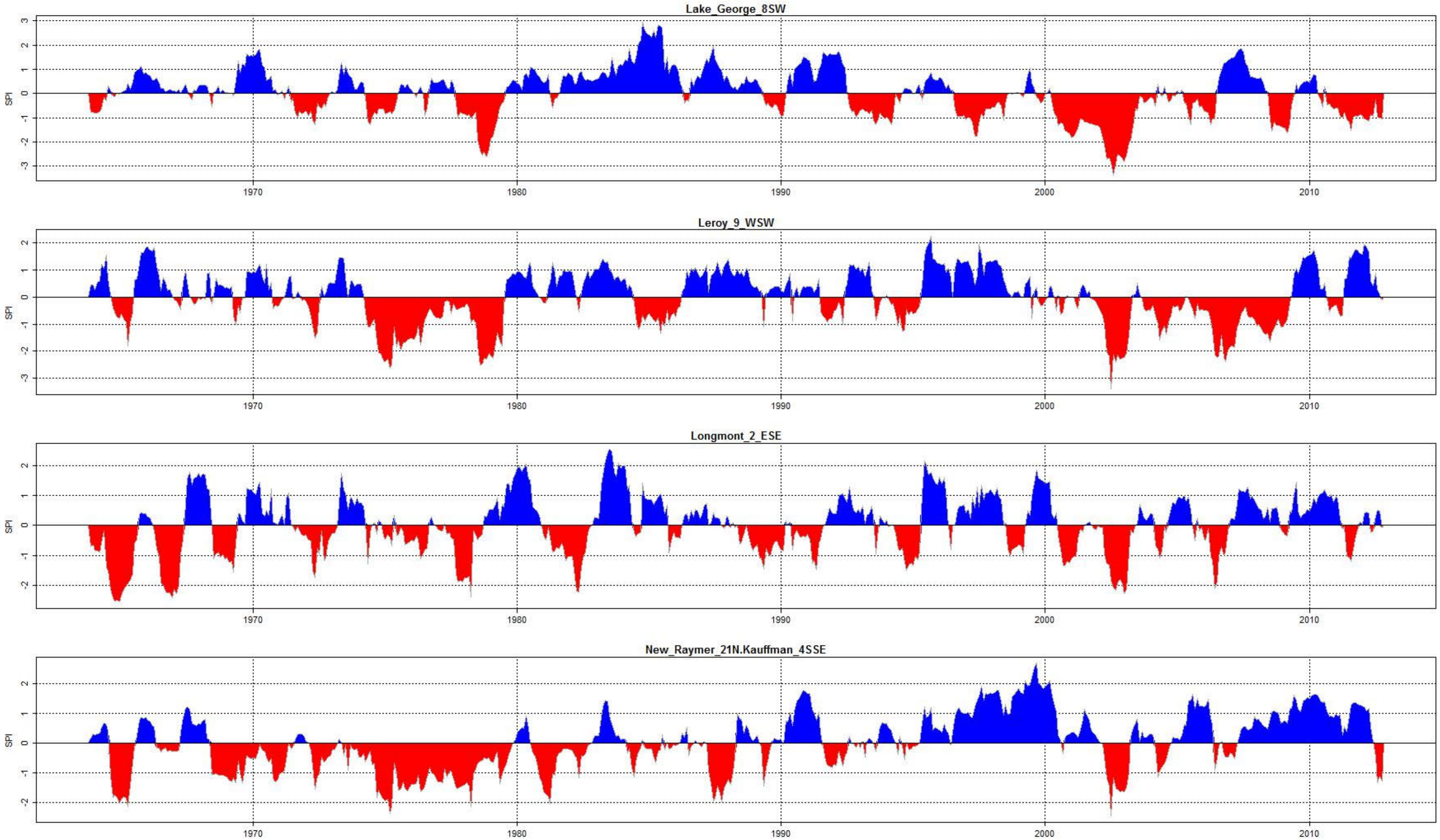

Denver_Stapleton_AP
SPI
2
1
0
-1
-2
1970
1980
1990
2000
2010

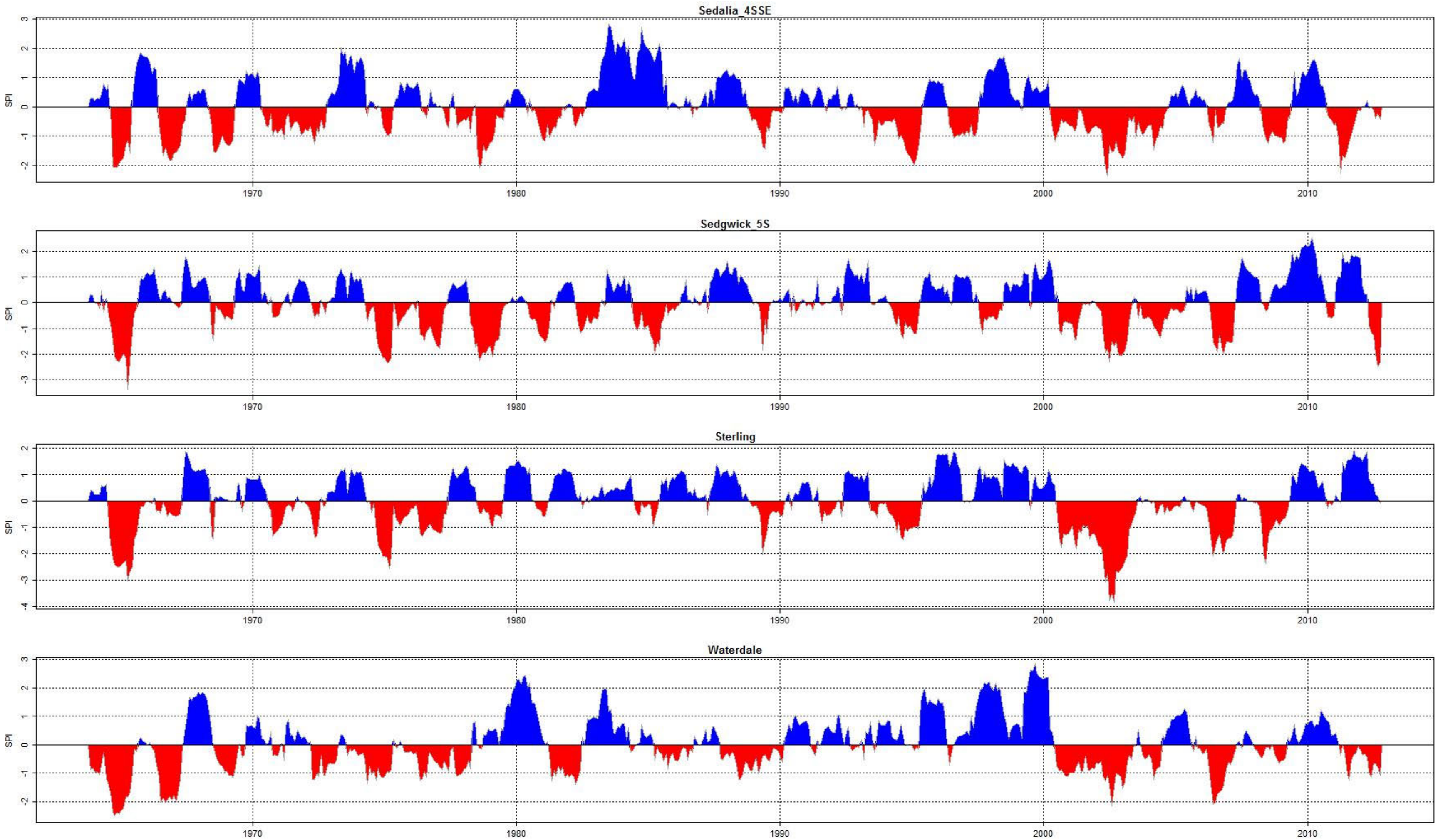

Estes_Park
SPI
2
1
0
-1
-2
-3
-4
1970
1980
1990
2000
2010


Fleming_3SW
SPI
2
1
0
-1
-2
1970
1980
1990
2000
2010


Fort_Collins
SPI
2
1
0
-1
-2
1970
1980
1990
2000
2010


Fort_Morgan
Greeley
Julesburg
Kassler
SPI
1970
1980
1990
2000
2010


Lake_George_8SW
Leroy_9_WSW
Longmont_2_ESE
New_Raymer_21N.Kauffman_4SSE
SPI
1970
1980
1990
2000
2010


Sedalia_4SSE
Sedgwick_5S
Sterling
Waterdale
SPI
1970
1980
1990
2000
2010

Appendix D. Evapotranspiration Station Metadata

#	Station Name	ID	Network	UTM X	UTM Y	Elev. (m)	Period
1	Haxtun	HXT01	CoAgMet	698876.6	4505032.6	1231.39	2000-2016
2	Sterling	ID_108	NCWCD	649217.0	4493283.5	1210.06	2000-2016
3	Brush	ID_107	NCWCD	610990.7	4454763.8	1303.93	2000-2016
4	Eastern Adams County	EAC01	CoAgMet	602923.8	4404663.7	1495.65	2000-2016
5	Wiggins	ID_106	NCWCD	583709.2	4461883.4	1362.76	2000-2016
6	Briggsdale	BRG01	CoAgMet	557622.7	4493989.5	1480.72	2002-2016
7	Kersey 1	KSY01	CoAgMet	539728.0	4469684.9	1409.70	2000-2016
8	Lucerne	LCN01	CoAgMet	524836.0	4480587.7	1447.80	2000-2016
9	Ault	ALT01	CoAgMet	523701.2	4490951.5	1496.57	2000-2016
10	Peckham	PKH01	CoAgMet	523196.6	4462478.4	1432.86	2000-2016
11	Greeley West	ID_224	NCWCD	520601.2	4472638.6	1483.77	2000-2016
12	Eaton	ID_104	NCWCD	519433.6	4491764.8	1510.28	2000-2016
13	Gilcrest	ID_105	NCWCD	513447.2	4456955.7	1455.42	2000-2016
14	Fort Lupton	FTL01	CoAgMet	512889.1	4427890.3	1540.76	2000-2016
15	Fort Collins East	ID_101	NCWCD	503463.1	4496888.2	1571.55	2000-2016
16	CSU - ARDEC	FTC03	CoAgMet	500000.0	4500182.5	1557.53	2000-2016
17	Loveland	ID_102	NCWCD	494149.2	4475394.2	1521.56	2000-2016
18	Longmont South	ID_103	NCWCD	493444.9	4436021.4	1519.12	2000-2016
19	Windy Gap	ID_350	NCWCD	417671.2	4440149.4	2416.76	2000-2016

Appendix E. Ditch Identification numbers

WDID									
0100503	0200885	0400521	0500539	0600516	0600603	0801362	2300894	6400502	8000673
0100507	0200887	0400522	0500542	0600518	0600608	0801412	2300902	6400503	8000674
0100511	0200888	0400523	0500545	0600523	0600610	0801413	2300904	6400504	8000706
0100514	0300905	0400524	0500546	0600525	0600621	0801426	2300922	6400506	8000713
0100515	0300910	0400530	0500547	0600527	0600650	0900535	2300923	6400508	8000729
0100517	0300911	0400532	0500548	0600528	0600735	0900731	2300924	6400511	8000730
0100518	0300912	0400534	0500549	0600532	0700502	0900752	2300926	6400513	8000732
0100519	0300913	0400541	0500550	0600536	0700527	0900767	2300931	6400514	8000759
0100525	0300914	0400543	0500551	0600537	0700540	0900958	2300932	6400515	8000760
0100526	0300915	0400574	0500552	0600538	0700547	2300500	2300933	6400516	8000761
0100687	0300918	0400578	0500553	0600542	0700549	2300502	2300936	6400518	8000773
0100688	0300919	0400582	0500554	0600543	0700551	2300503	2300937	6400520	8000774
0200808	0300922	0400588	0500557	0600551	0700569	2300504	2300948	6400522	8000776
0200809	0300926	0400592	0500558	0600553	0700570	2300505	2300963	6400524	8000777
0200810	0300929	0400599	0500559	0600554	0700597	2300506	2300975	6400525	8000784
0200812	0300930	0400600	0500561	0600560	0700601	2300516	2300977	6400528	8000785
0200813	0300931	0400601	0500563	0600564	0700614	2300564	2300986	6400530	8000792
0200817	0300932	0400602	0500564	0600565	0700632	2300568	2300987	6400531	8000794
0200821	0300934	0400603	0500565	0600566	0700647	2300569	2300991	6400532	8000799
0200822	0300935	0500511	0500568	0600567	0700652	2300573	2300993	6400533	8000800
0200824	0300937	0500523	0500569	0600569	0700698	2300579	2300994	6400535	8000801
0200825	0300994	0500526	0500570	0600570	0700699	2300585	2301003	6403906	8000812
0200826	0301038	0500527	0500571	0600575	0801004	2300586	2301005	8000650	8000827
0200828	0301039	0500528	0500572	0600576	0801124	2300691	2301018	8000651	8000828
0200830	0301041	0500529	0500573	0600580	0801215	2300760	2301020	8000657	8000829
0200834	0400501	0500530	0500574	0600582	0801230	2300763	2301022	8000659	8000831
0200836	0400502	0500534	0500589	0600585	0801237	2300774	2301025	8000660	8000843
0200837	0400503	0500535	0500601	0600586	0801241	2300787	2301075	8000661	8000895
0200872	0400517	0500536	0600501	0600588	0801250	2300789	2301083	8000662	8000896
0200873	0400519	0500537	0600513	0600592	0801264	2300797	2301138	8000667	8000897
0200874	0400520	0500538	0600515	0600593	0801279	2300887	2302910	8000668	8000921

Appendix F. Reservoirs Location and Metadata

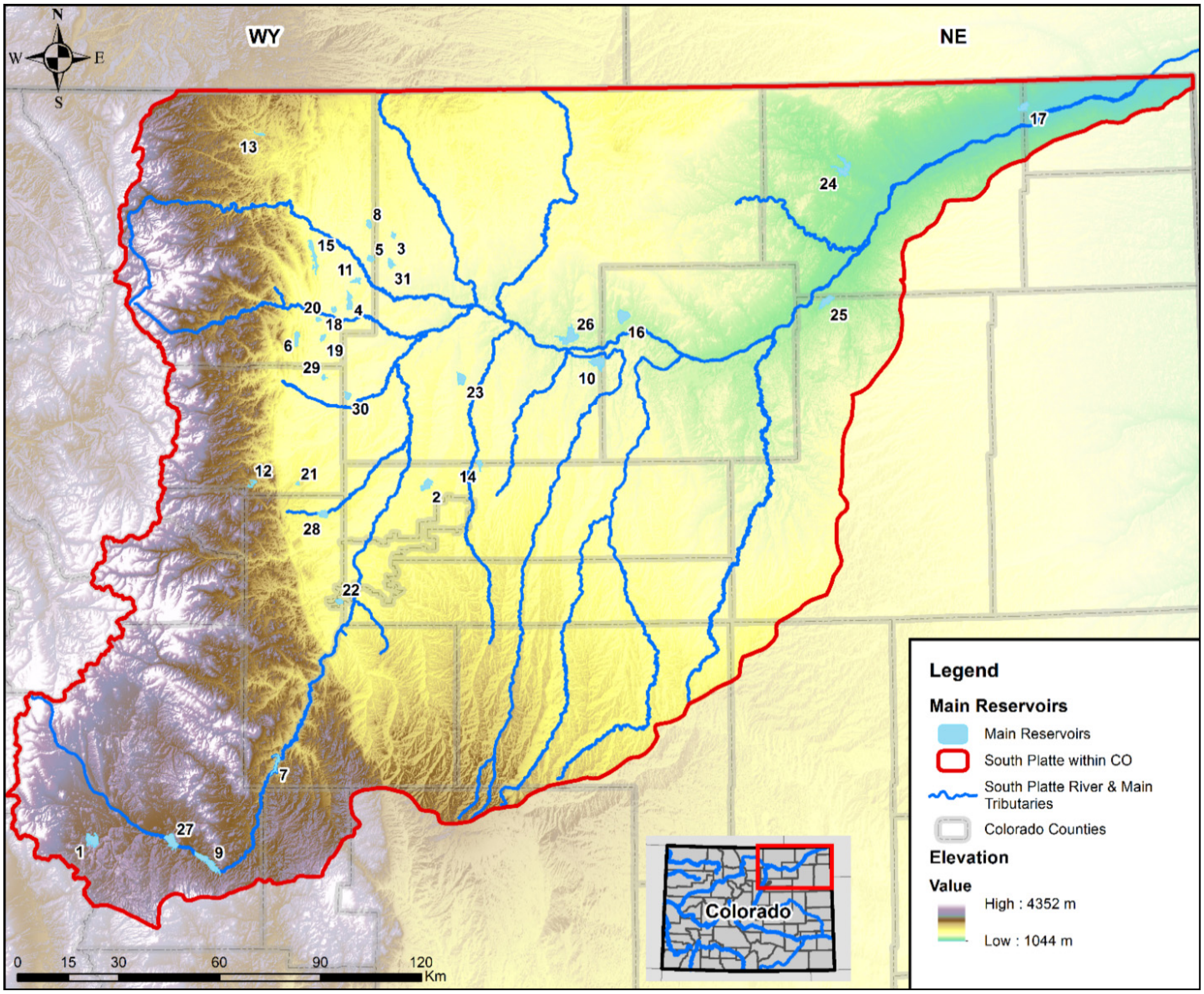


Figure 46: Main South Platte's reservoirs location

Table 15: Main South Platte's reservoirs metadata

#	Reservoir	Station ID	Elev. (m)	County	UTM X	UTM Y
1	Antero Reservoir	6016010	2727.96	Park	422055.78	4316052.29
2	Barr Lake	6016020	1569.72	Adams	519647.05	4422233.14
3	Black Hollow Reservoir	6016030	1548.38	Weld	510149.92	4496581.85
4	Boyd Lake	6016040	1508.76	Larimer	497455.34	4475485.30
5	Cache La Poudre	6016050	1493.52	Larimer	502540.13	4488805.26
6	Carter Lake	6016060	1761.74	Larimer	481308.81	4463298.38
7	Cheesman Lake	6016080	2090.93	Douglas	476689.48	4340115.74
8	Cobb Lake	6016090	1569.72	Larimer	502536.35	4499905.44
9	Elevenmile Canyon Reservoir	6016100	2615.18	Park	458377.41	4305789.18
10	Empire Reservoir	6016110	1371.6	Weld	572269.61	4458072.02
11	Fossil Creek Reservoir	6016120	1481.33	Larimer	499152.54	4482144.86
12	Gross Reservoir	6016130	2209.8	Boulder	469248.07	4422269.87
13	Halligan Reservoir	6016140	1935.48	Larimer	471353.66	4525491.79
14	Horse Creek Reservoir	6016370	1542.29	Adams	535845.90	4428951.70
15	Horsetooth Reservoir	6016150	1661.16	Larimer	485616.65	4494368.78
16	Jackson Lake Reservoir	6016160	1335.02	Morgan	578094.43	4470341.34
17	Julesburg Reservoir	6016170	1127.76	Logan	699509.43	4534802.34
18	Lake Loveland Reservoir	6016180	1530.1	Larimer	493213.25	4474377.96
19	Lone Tree Reservoir	6016190	1569.72	Larimer	489807.82	4465501.99
20	Mariano Reservoir	6016200	1539.24	Larimer	487267.30	4469945.77
21	Marshall Reservoir	6016220	1725.17	Boulder	481207.17	4422230.99
22	Marston Reservoir	6016210	1688.59	Jefferson	493134.51	4386695.29
23	Milton Reservoir	6016230	1456.94	Weld	529775.45	4453344.32

#	Reservoir	Station ID	Elev. (m)	County	UTM X	UTM Y
24	North Sterling Reservoir	6016240	1234.44	Unknown	645981.66	4515775.82
25	Prewitt Reservoir	6016250	1249.68	Washington	637434.73	4475635.34
27	Riverside Reservoir	6016270	1371.6	Weld	562021.01	4463530.91
28	Spinney Mountain Reservoir	16016025	2647.49	Park	446290.228	4313630.375
29	Standley Reservoir	6016280	1676.4	Jefferson	489737.437	4413335.622
30	Terry Reservoir	6016290	1554.48	Boulder	489791.284	4453292.463
31	Union Reservoir	6016300	1508.76	Weld	497445.946	4447736.279
32	Windsor Reservoir	6016310	1456.94	Weld	508475.902	4481039.62

Appendix G. Locations of all studied wells

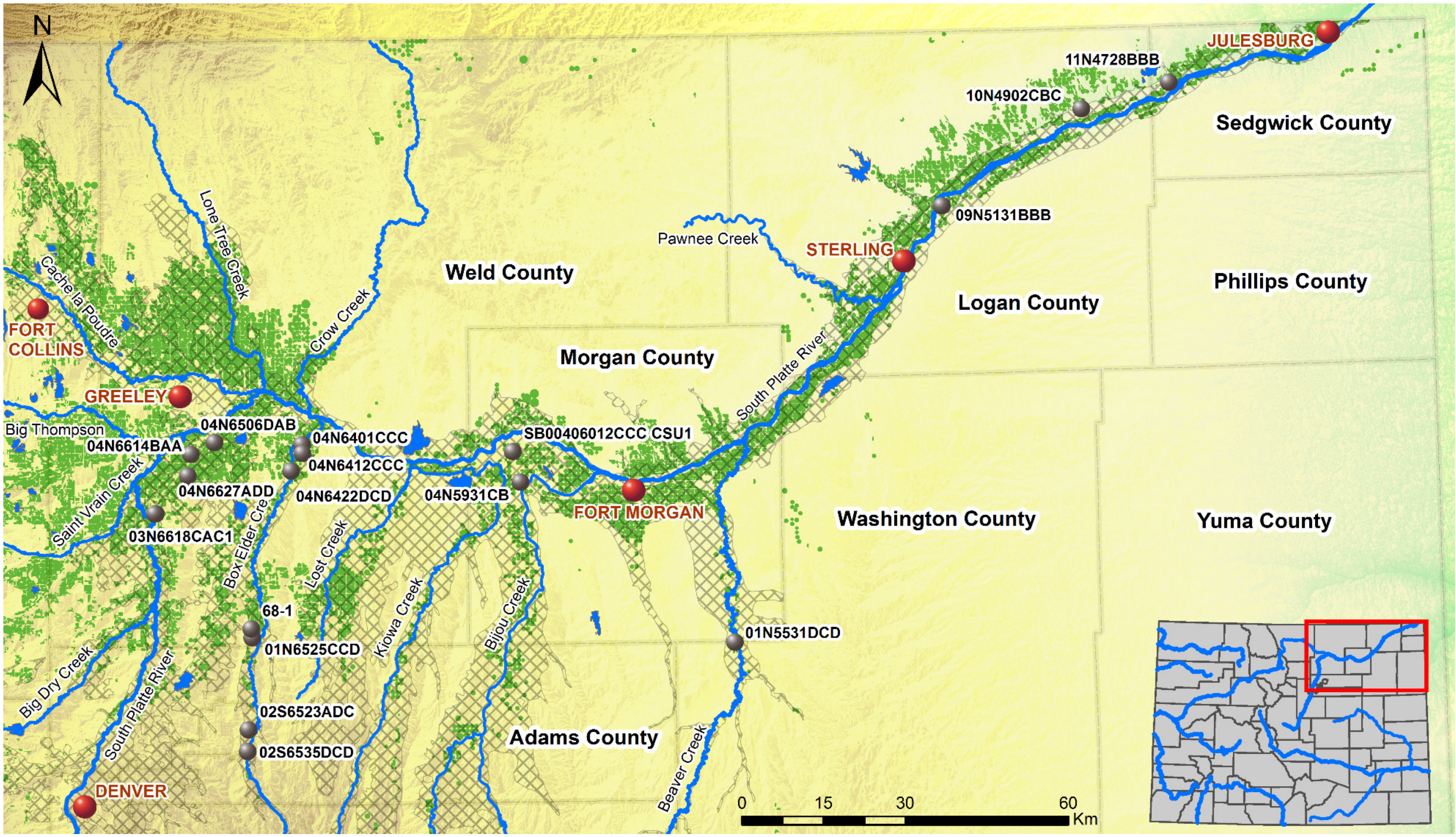
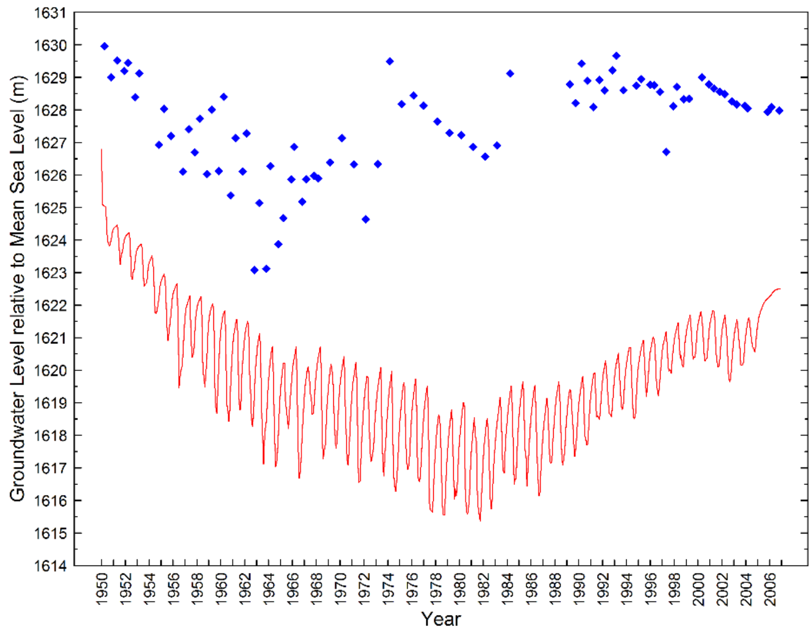


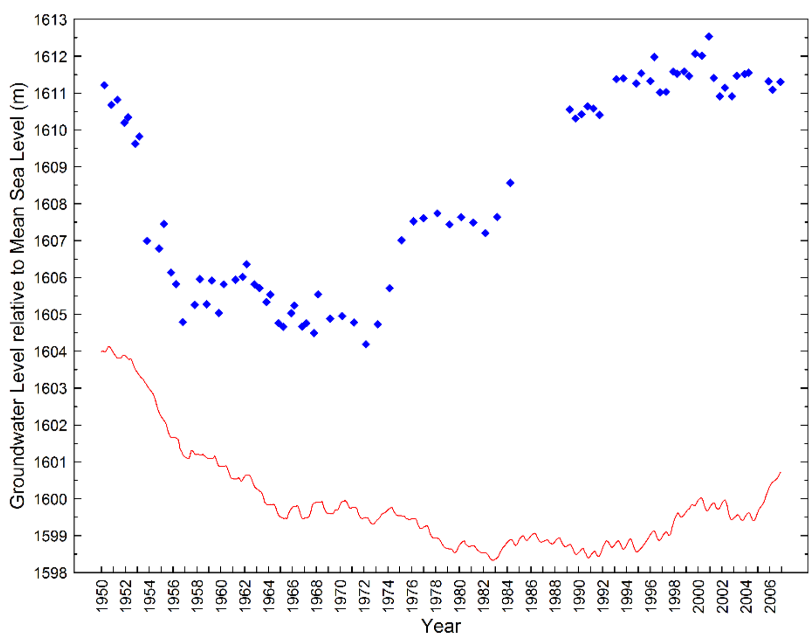
Figure 47: Locations of all studied groundwater wells

Appendix H. Modeled monthly groundwater head & biannual observations

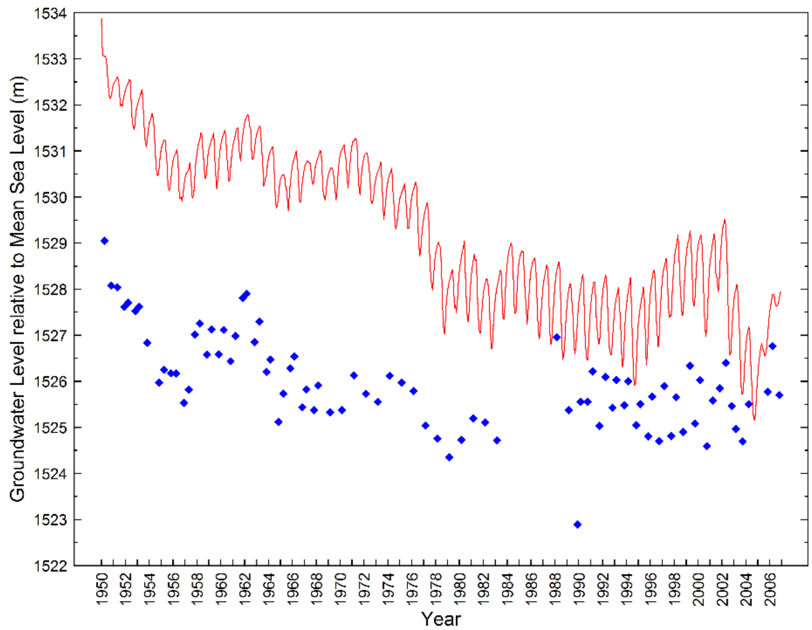
Well ID: 02S6535DCD



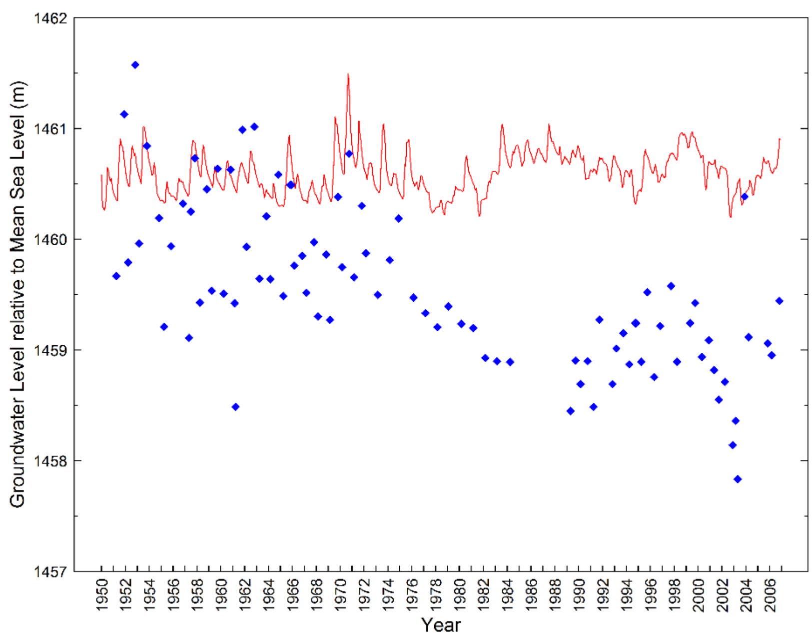
Well ID: 02S6523ADC



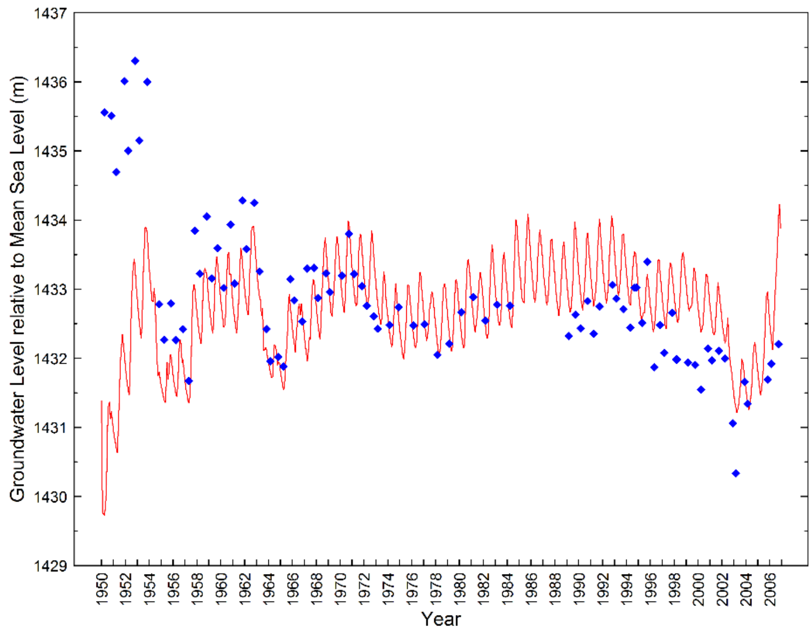
Well ID: 01N6525CCD



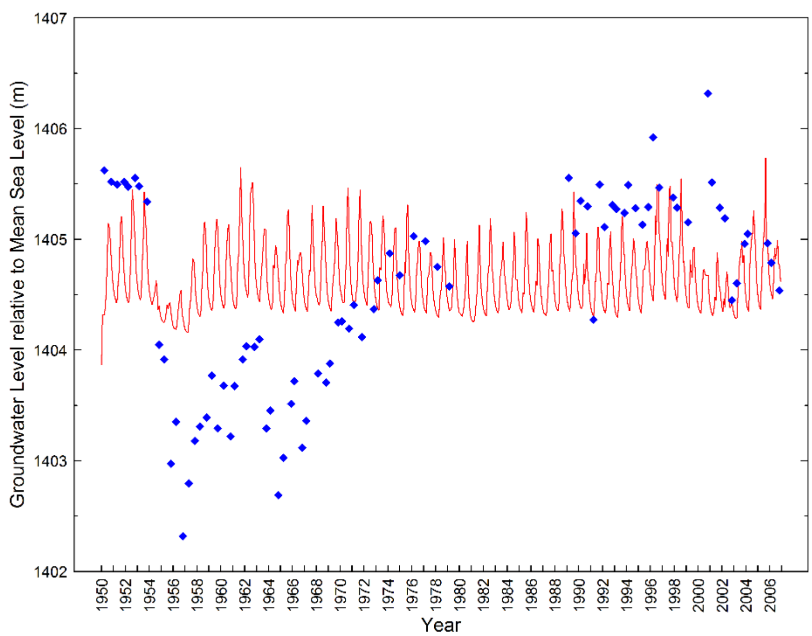
Well ID: 03N6618CAC1



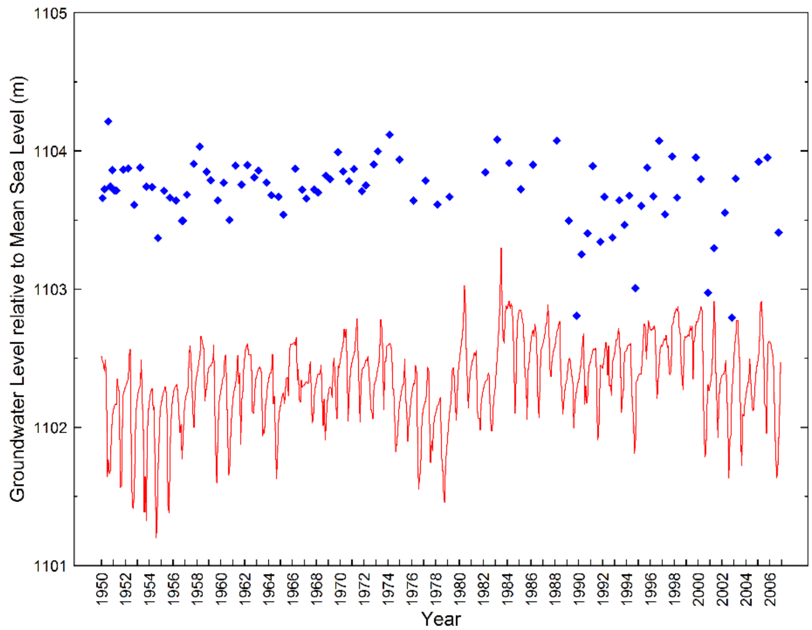
Well ID: 04N6614BAA



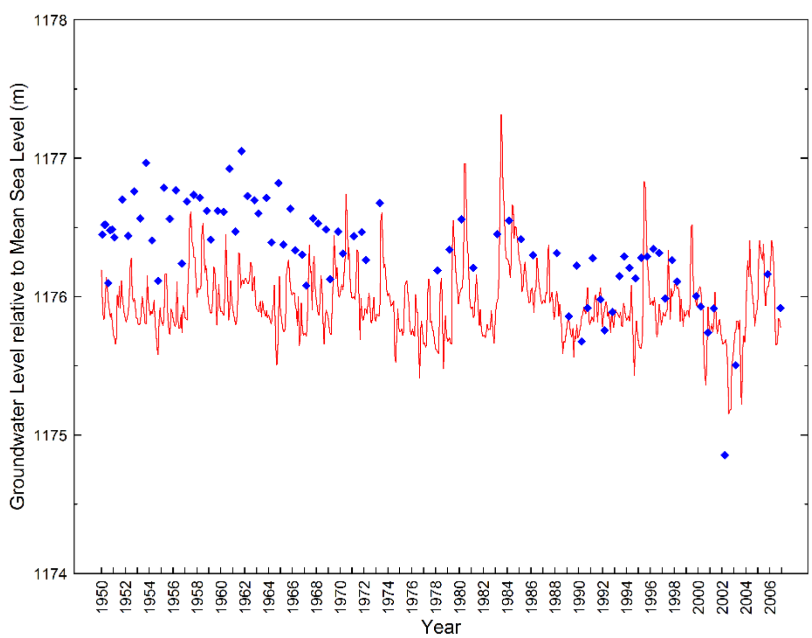
Well ID: 04N6401CCC



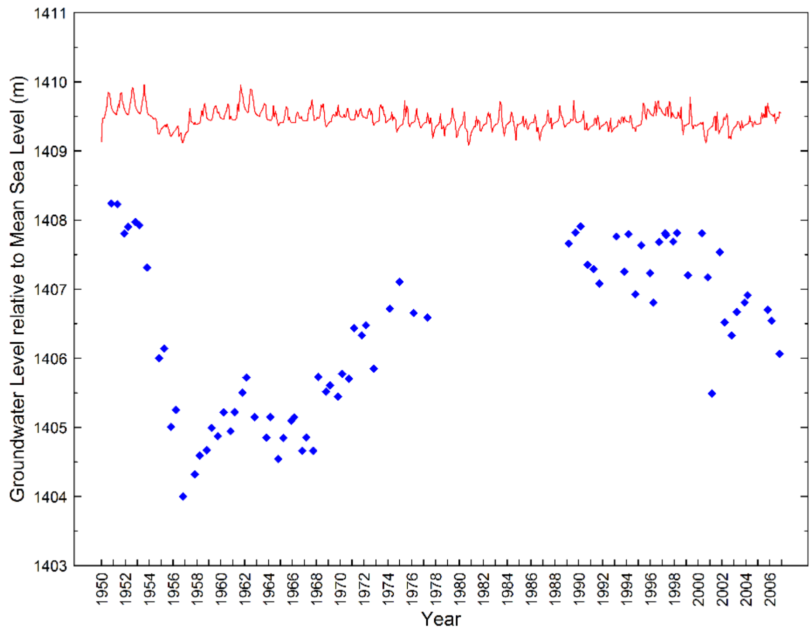
Well ID: 11N4728BBB



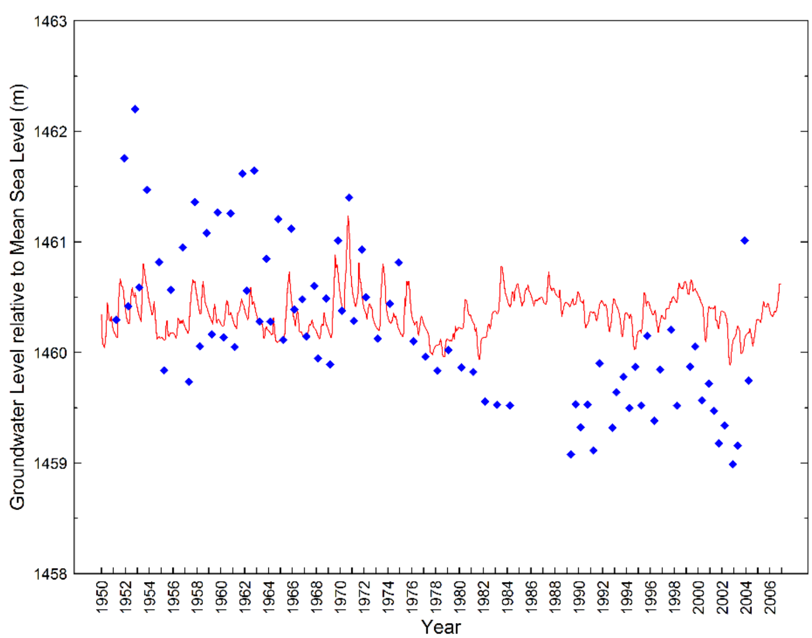
Well ID: 09N5131BBB



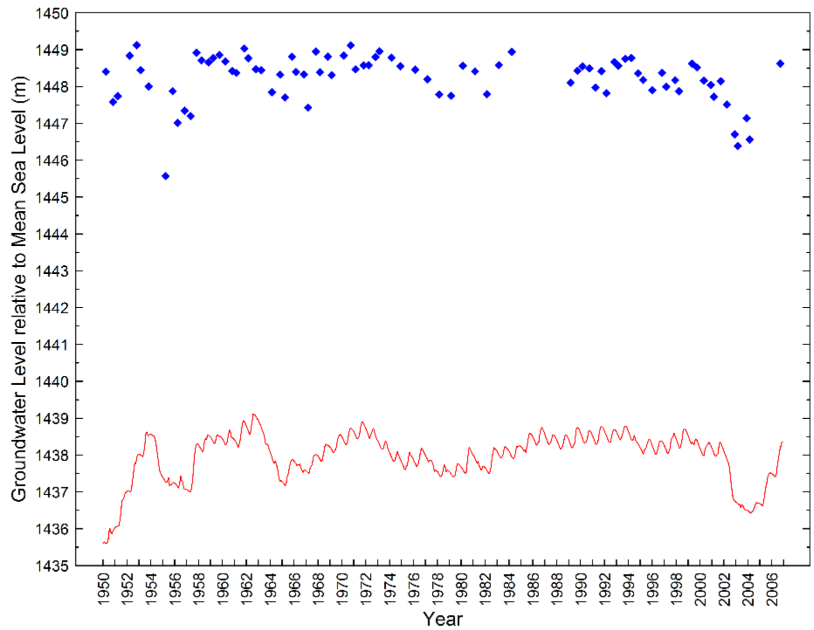
Well ID: 04N6412CCC



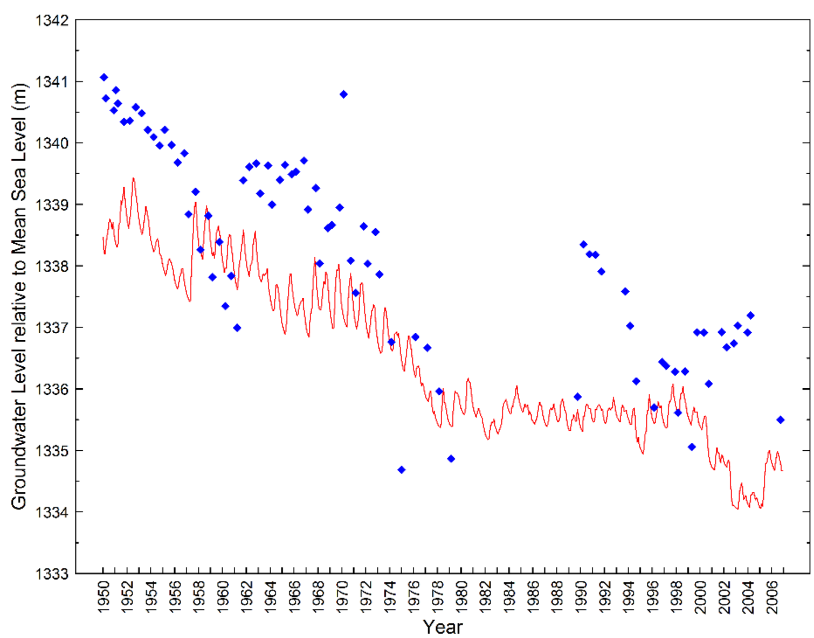
Well ID: SB00306618CAC



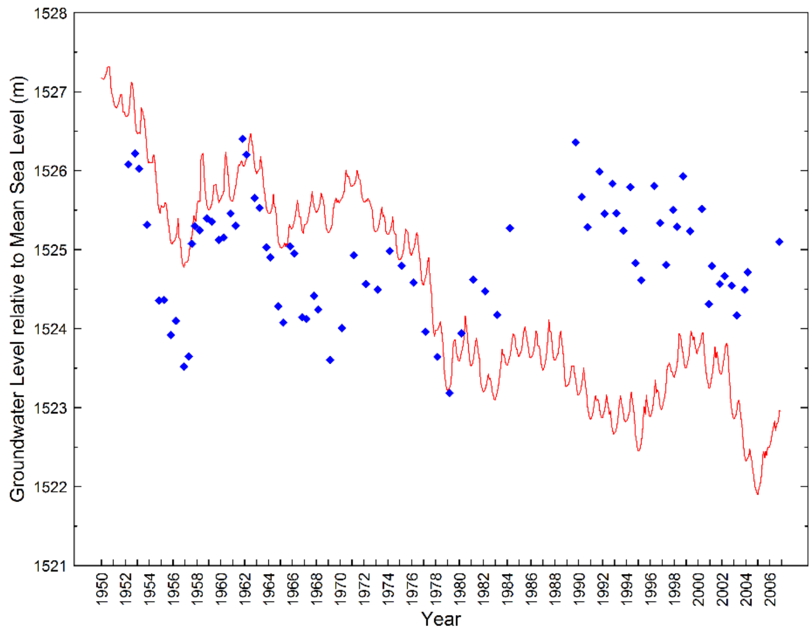
Well ID: 04N6627ADD



Well ID: SB00406012CCC-CSU1



Well ID: 68-1



Groundwater Level relative to Mean Sea Level (m)
1528
1527
1526
1525
1524
1523
1522
1521
1950 1952 1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006
Year

Appendix I. SARIMAX models for the tested radii and the AIC

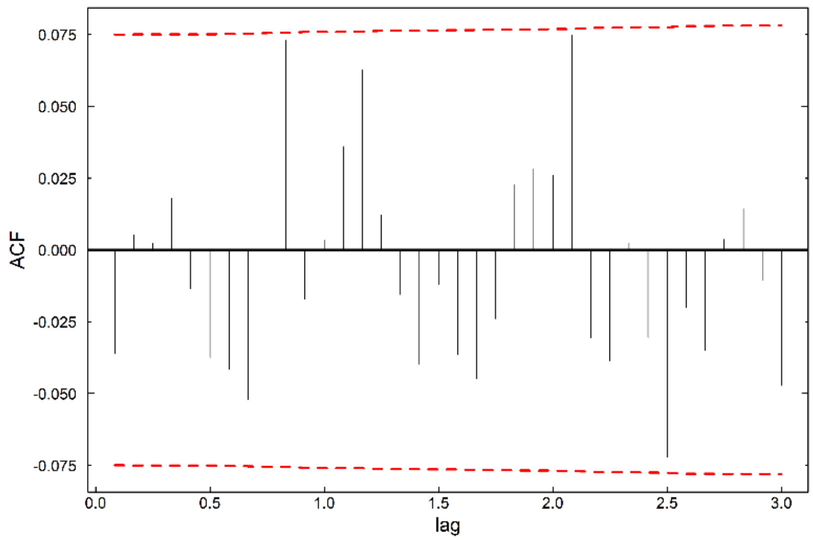
Well ID	ARIMA MODEL	Radii in m (ft)	AIC
02S6535DCD	(1,1,2)(2,0,1)[12] with zero mean, xreg = P	914 (3000)	1764.3
	(1,1,2)(2,0,1)[12] with zero mean, xreg = P	1829 (6000)	1764.32
	(1,1,2)(2,0,1)[12] with zero mean, xreg = R, P	3,048 (10000)	1741.7
	(1,1,2)(2,0,1)[12] with zero mean, xreg = R, P	4,877 (16000)	1734.74
02S6523ADC	(2,1,1)(0,0,3)[12] with zero mean, xreg = R	914 (3000)	-1734.51
	(1,1,2)(2,0,0)[12] with zero mean, xreg = R	1829 (6000)	-1750.94
	(2,1,1)(3,0,1)[12] with zero mean, xreg = R, P	3,048 (10000)	-1744.26
	(2,1,0)(2,0,0)[12] with zero mean, xreg = R, P	4,877 (16000)	-1818.01
01N6525CCD	(2,1,0)(2,0,1)[12] with zero mean, xreg = R	914 (3000)	563.3
	(2,1,1)(2,0,1)[12] with zero mean, xreg = R, P	1829 (6000)	692.09
03N6618CAC1	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	914 (3000)	-668.94
	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	1829 (6000)	-587.59
04N6614BAA	(1,0,2)(3,0,0)[12] with non-zero mean, xreg = R	914 (3000)	533.48
	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	1829 (6000)	544.36
	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	3,048 (10000)	551.35
04N6401CCC	(1,0,1)(3,0,0)[12] with non-zero mean, xreg = R	914 (3000)	475.86
	(1,0,1)(3,0,0)[12] with non-zero mean, xreg = R	1829 (6000)	429.54
	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	3,048 (10000)	222.6

Well ID	ARIMA MODEL	Radii in m (ft)	AIC
	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	4,877 (16000)	246.99
11N4728BBB	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	914 (3000)	565.88
	(1,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	1829 (6000)	574.01
09N5131BBB	(3,0,0)(1,0,1)[12] with non-zero mean, xreg = R, P	914 (3000)	621.35
04N6412CCC	(2,0,1)(1,0,2)[12] with non-zero mean, xreg = R	914 (3000)	-98.92
	(2,0,1)(1,0,2)[12] with non-zero mean, xreg = R	1829 (6000)	-104.16
	(3,0,0)(1,0,2)[12] with non-zero mean, xreg = R, P	3,048 (10000)	-286.79
	(3,0,0)(1,0,2)[12] with non-zero mean, xreg = R, P	4,877 (16000)	-279.02
SB00306618CAC	(1,0,2)(1,0,1)[12] with non-zero mean, xreg = R	914 (3000)	-413.56
04N6627ADD	(3,0,0)(3,0,0)[12] with non-zero mean, xreg = R	914 (3000)	-473.94
	(3,0,0)(3,0,0)[12] with non-zero mean, xreg = R	1829 (6000)	-477.49
	(2,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	3,048 (10000)	-991.87
	(2,0,1)(1,0,1)[12] with non-zero mean, xreg = R, P	4,877 (16000)	-819.56
SB00406012CCC CSU1	(2,1,1)(3,0,0) [12] with zero mean, xreg = R	914 (3000)	-40.99
	(2,1,1)(3,0,0)[12] with zero mean, xreg = R	1829 (6000)	-89.25
	(2,1,2)(2,0,0)[12] with zero mean, xreg = R, P	3,048 (10000)	-109.67

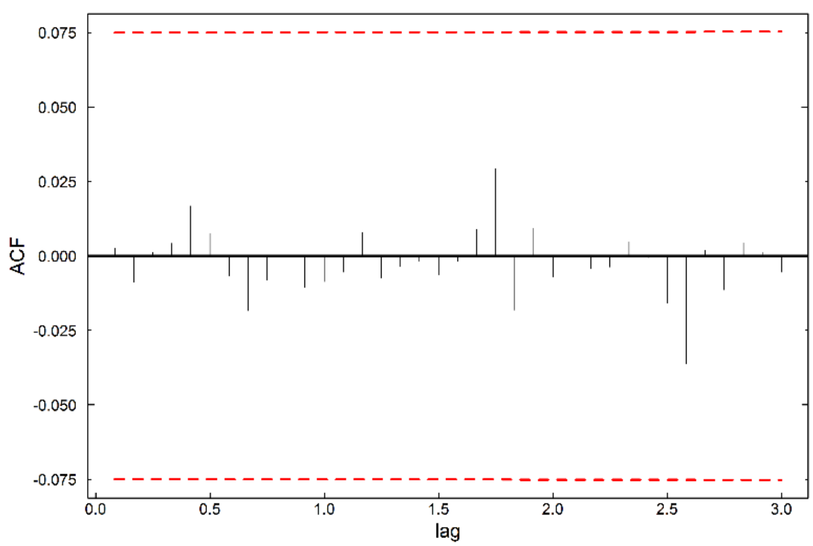
Well ID	ARIMA MODEL	Radii in m (ft)	AIC
68-1	(2,1,2)(1,0,1)[12] with zero mean, xreg = R	914 (3000)	-277.91
	(2,1,2)(1,0,1)[12] with non-zero mean, xreg = R, P	1829 (6000)	-237.9

Appendix J. ARIMA ACF of Residuals

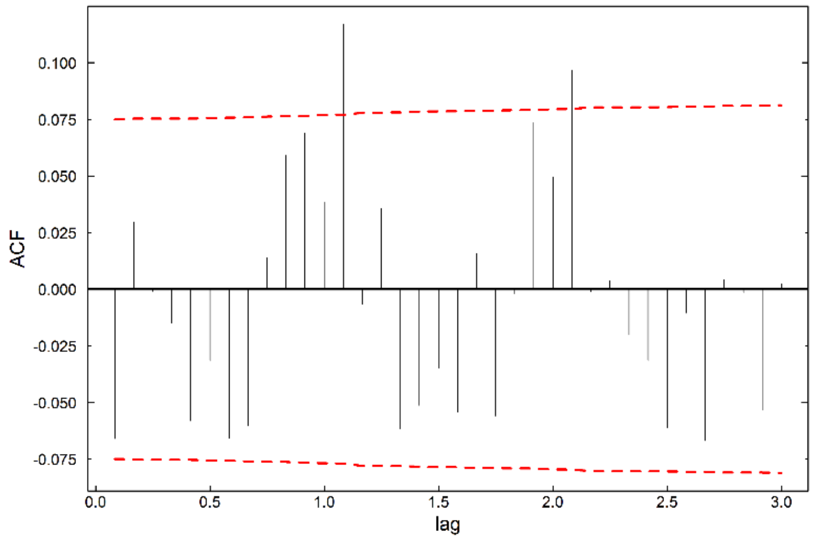
Well ID: 02S6535DCD



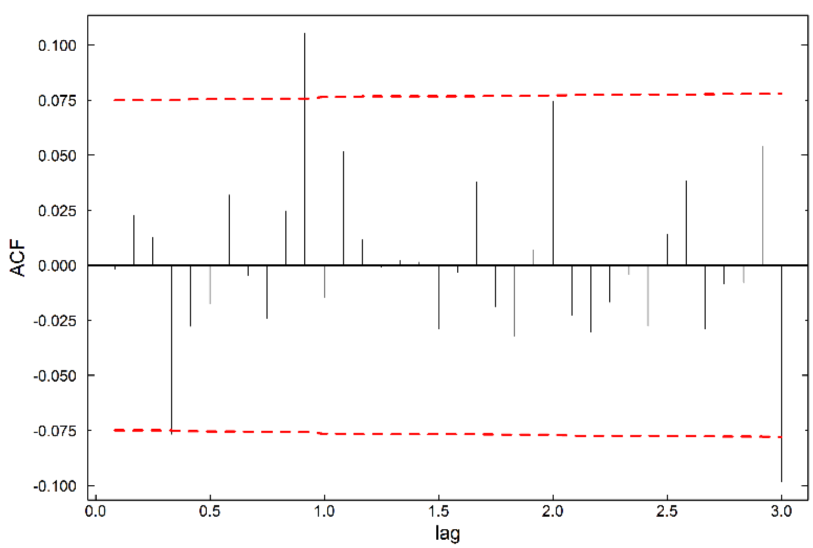
Well ID: 02S6523ADC



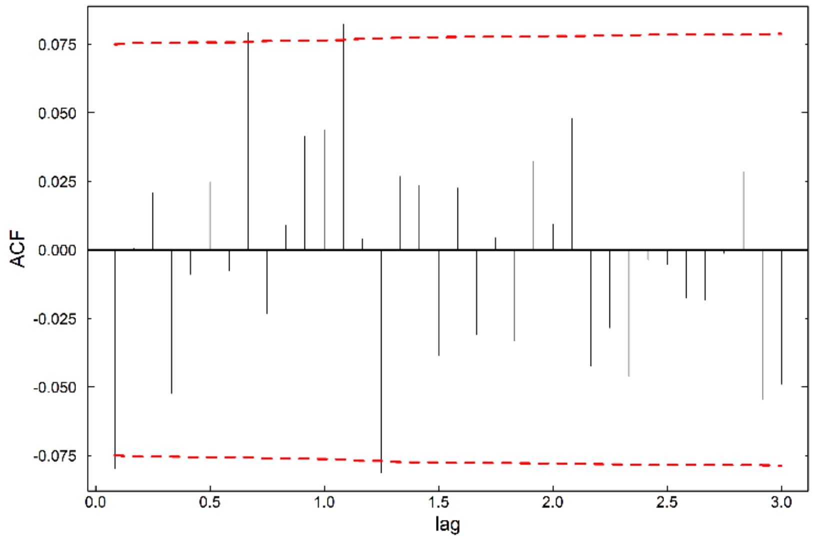
Well ID: 01N6525CCD



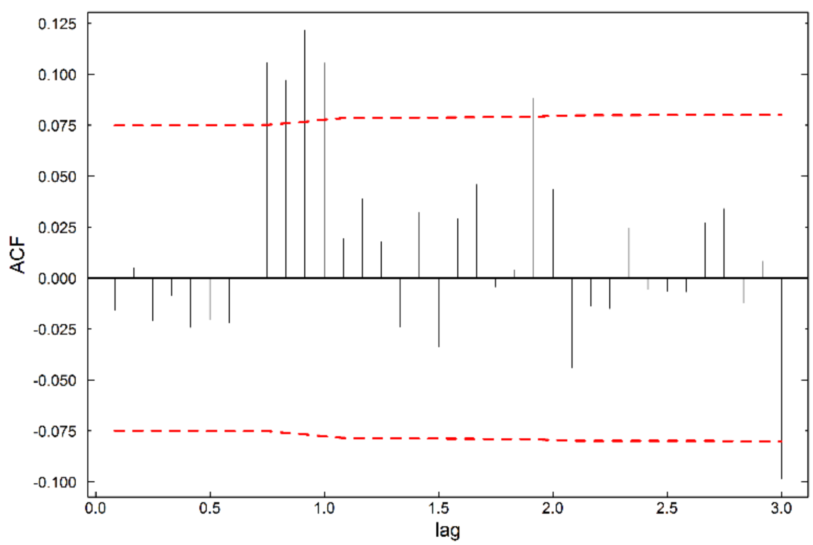
Well ID: 03N6618CAC1



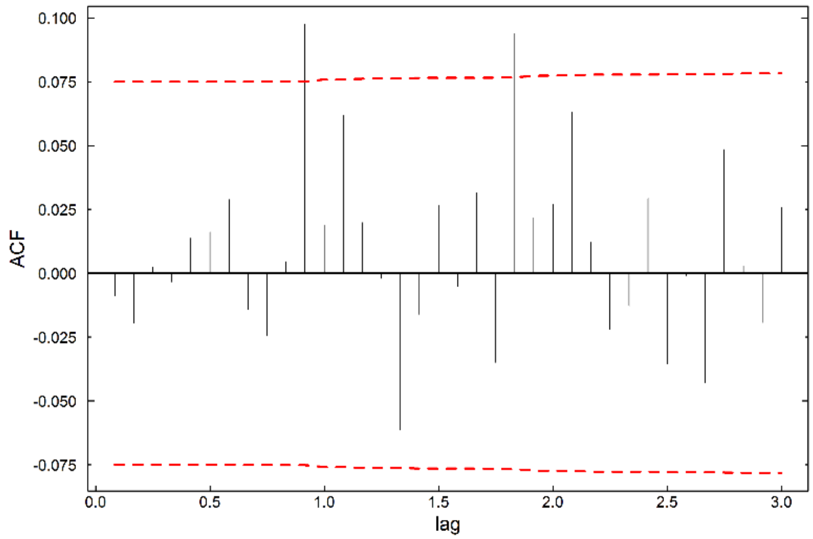
Well ID: 04N6614BAA



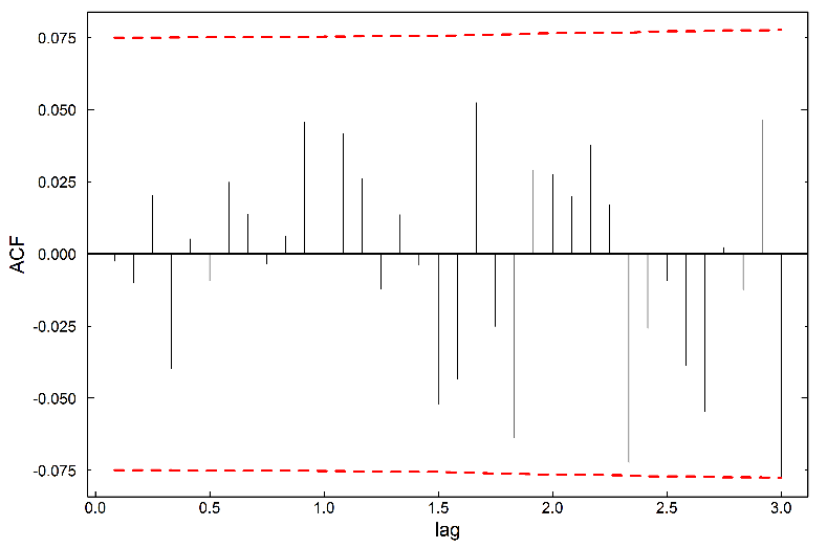
Well ID: 04N6401CCC



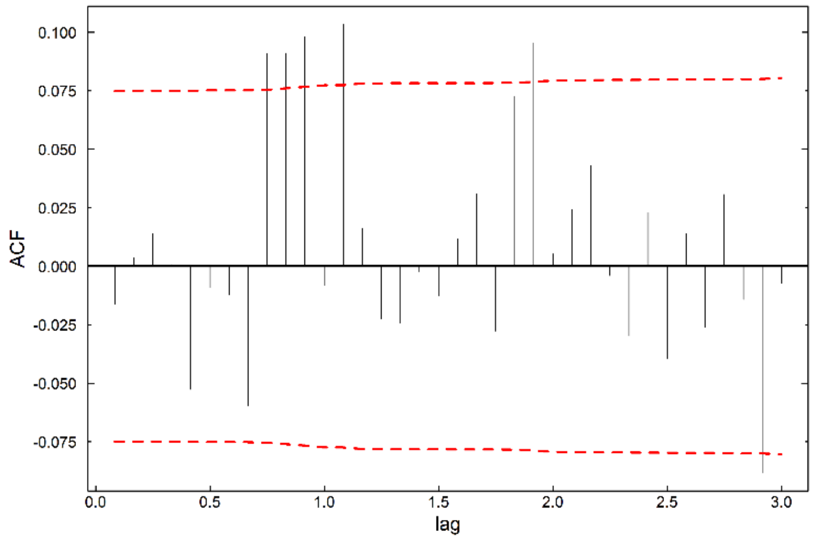
Well ID: 11N4728BBB



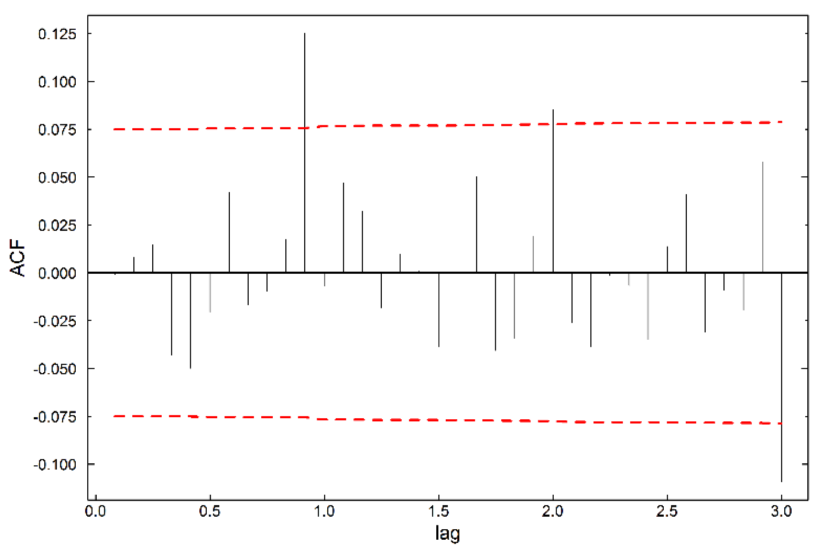
Well ID: 09N5131BBB



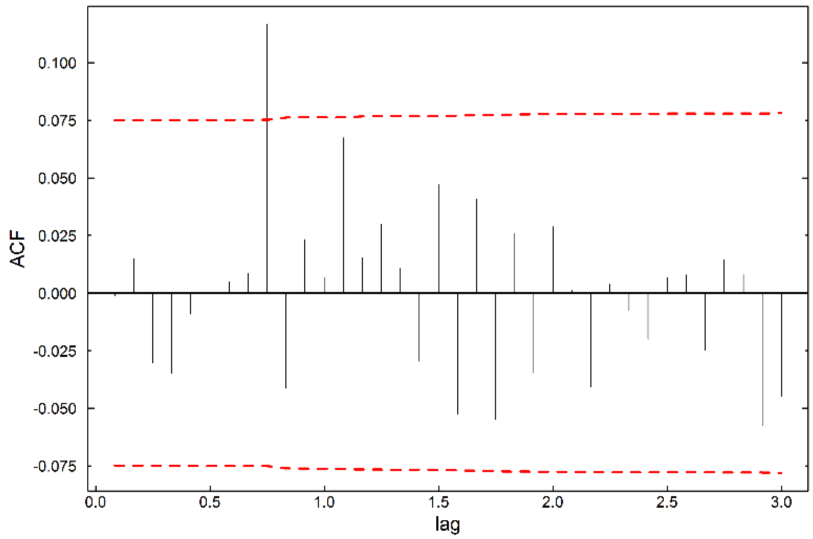
Well ID: 04N6412CCC



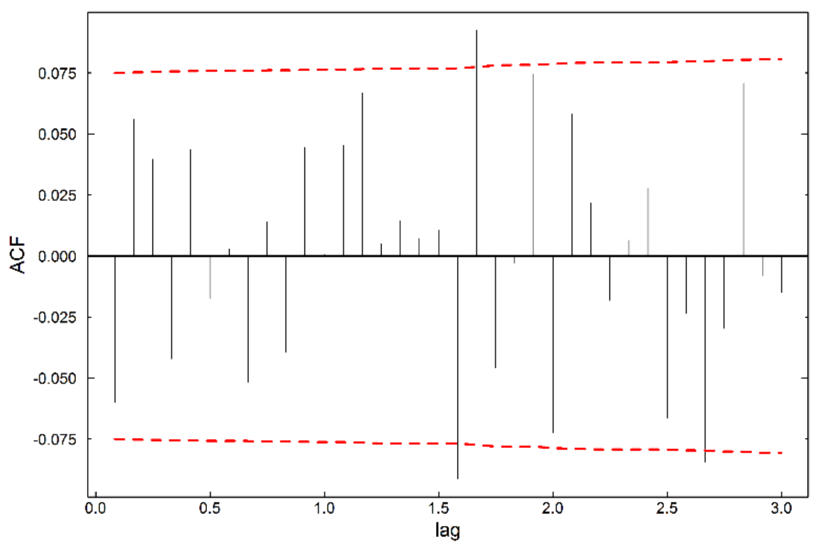
Well ID: SB00306618CAC



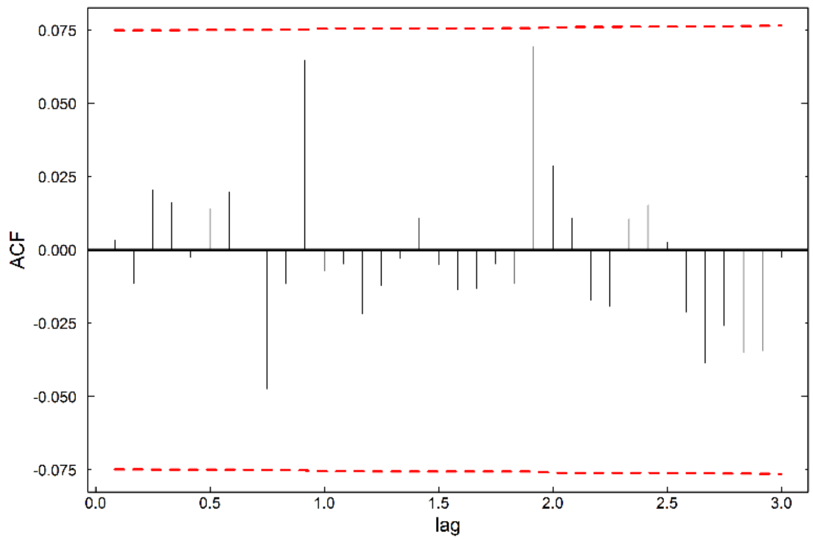
Well ID: 04N6627ADD



Well ID: SB00406012CCC-CSU1

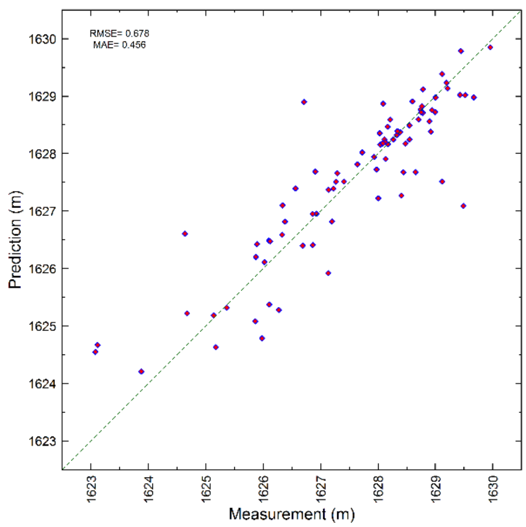


Well ID: 68-1

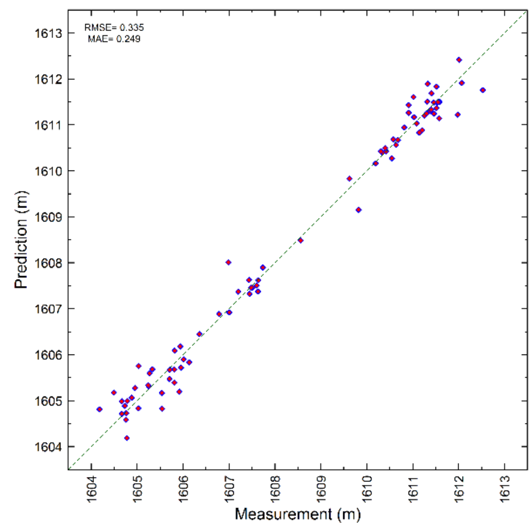


Appendix K. Results of Numerical Experiment #1

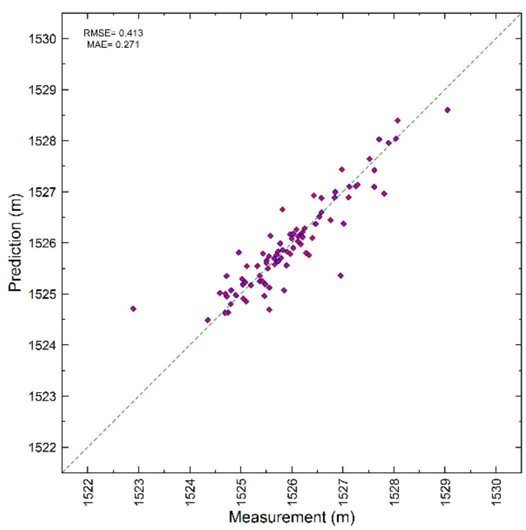
Well ID: 02S6535DCD



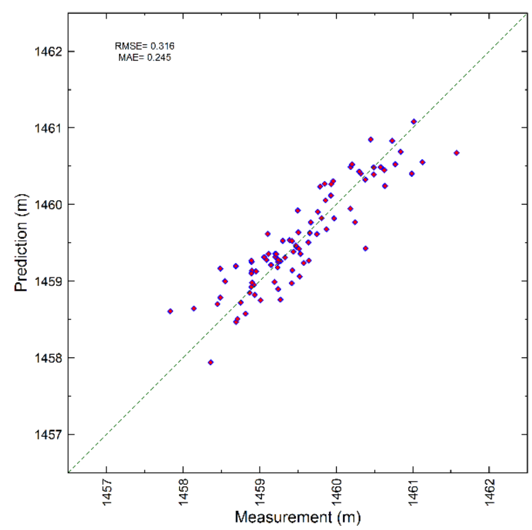
Well ID: 02S6523ADC



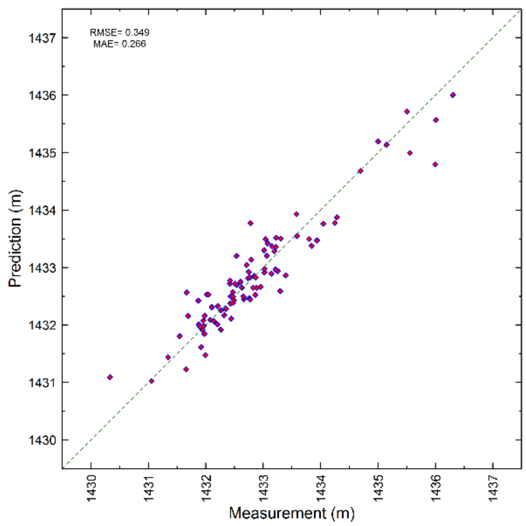
Well ID: 01N6525CCD



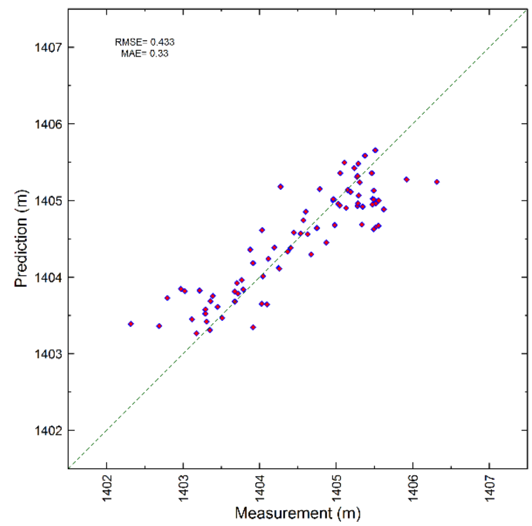
Well ID: 03N6618CAC1



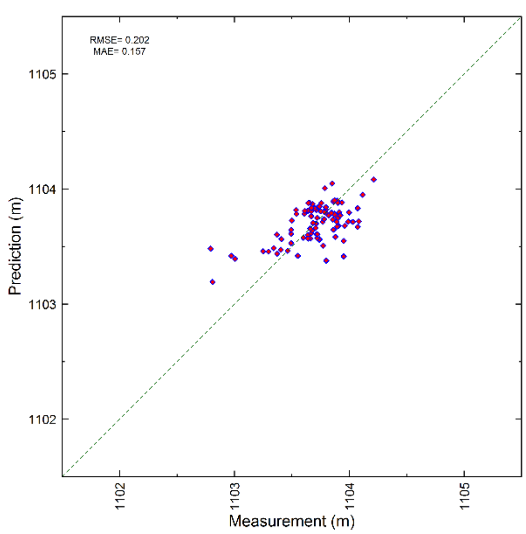
Well ID: 04N6614BAA



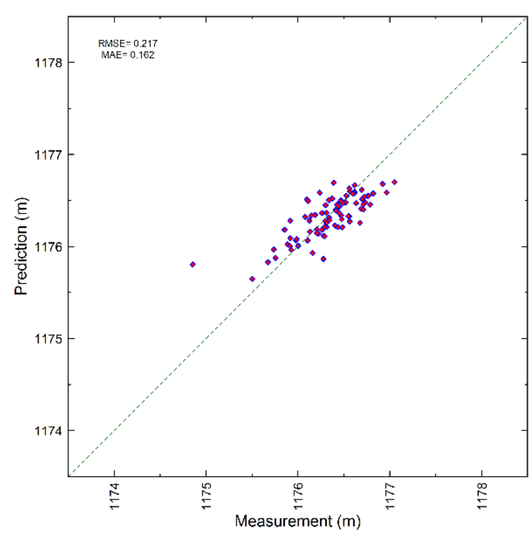
Well ID: 04N6401CCC



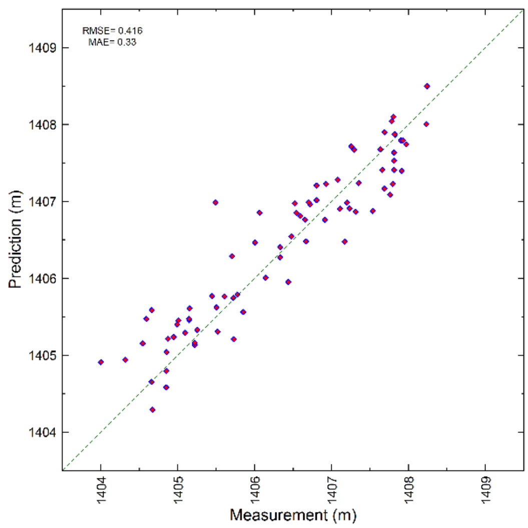
Well ID: 11N4728BBB



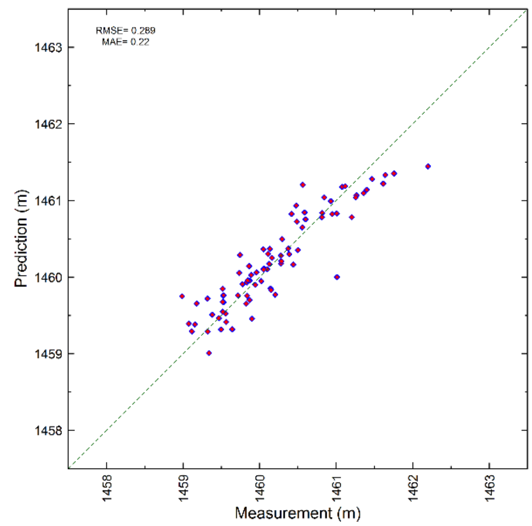
Well ID: 09N5131BBB



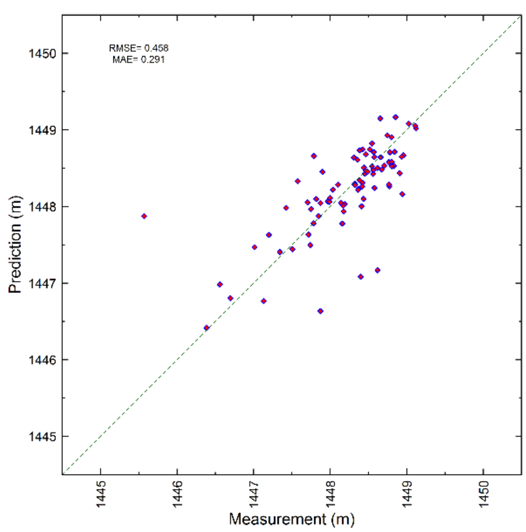
Well ID: 04N6412CCC



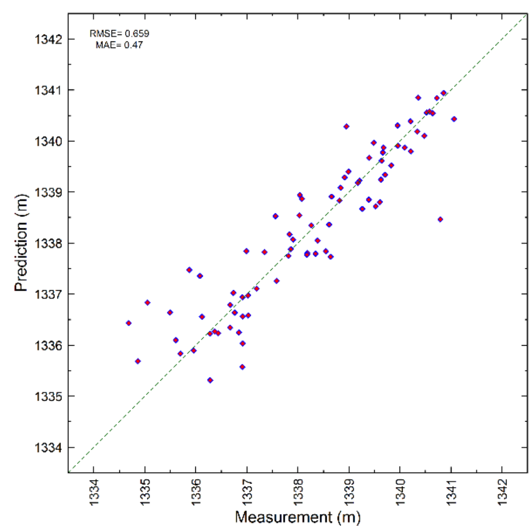
Well ID: SB00306618CAC



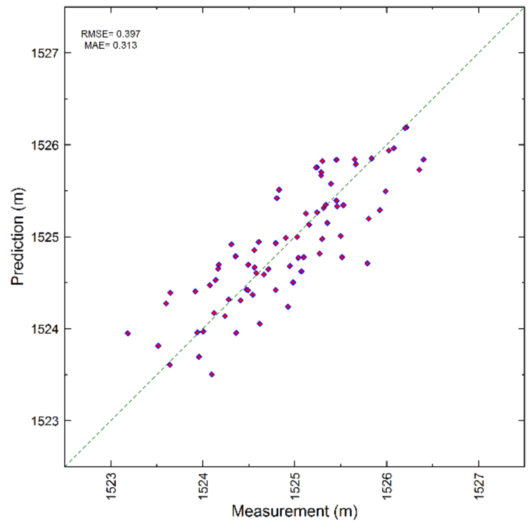
Well ID: 04N6627ADD



Well ID: SB00406012CCC-CSU1



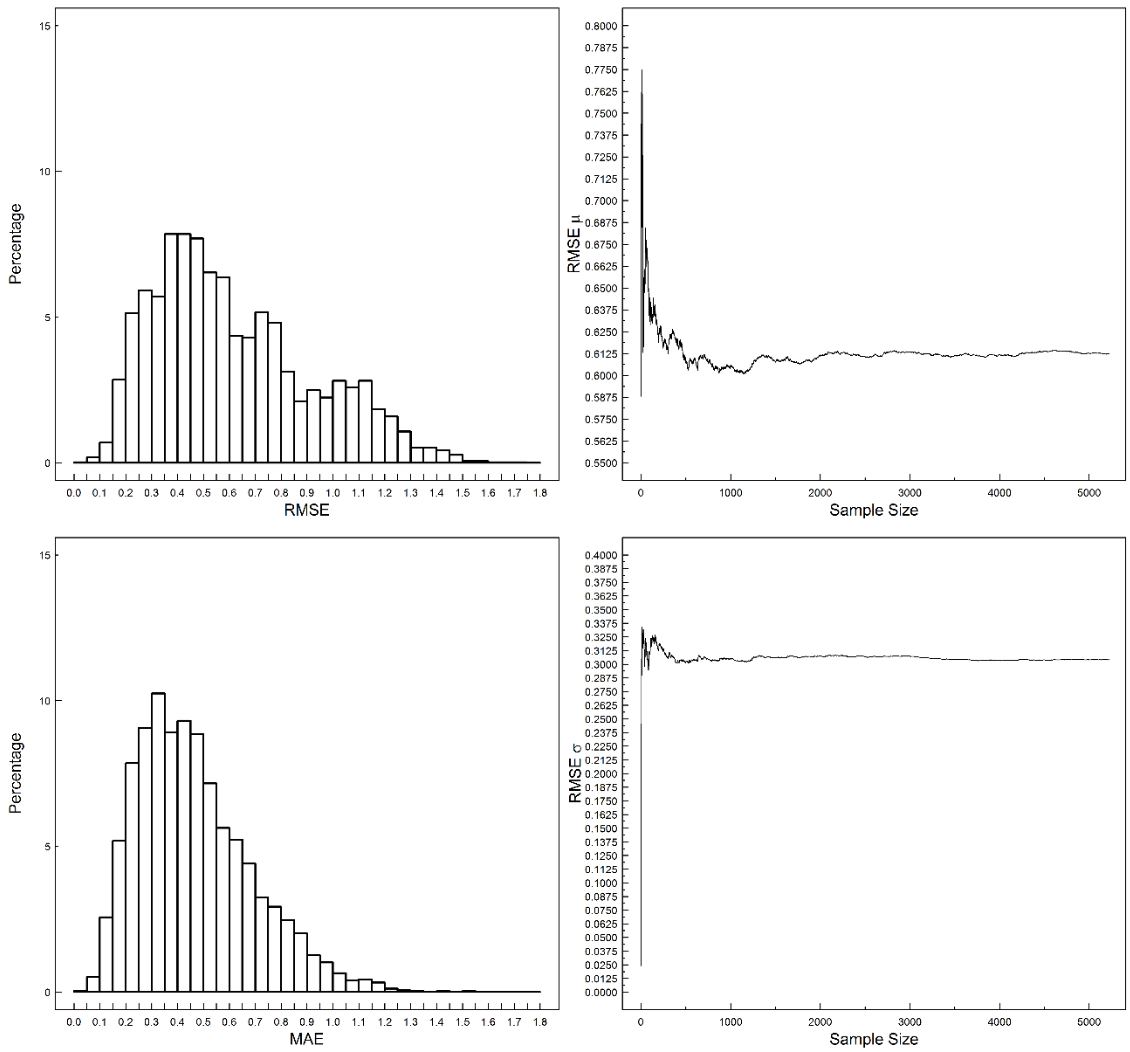
Well ID: 68-1


RMSE= 0.397
MAE= 0.313
Prediction (m)
Measurement (m)
1523
1524
1525
1526
1527

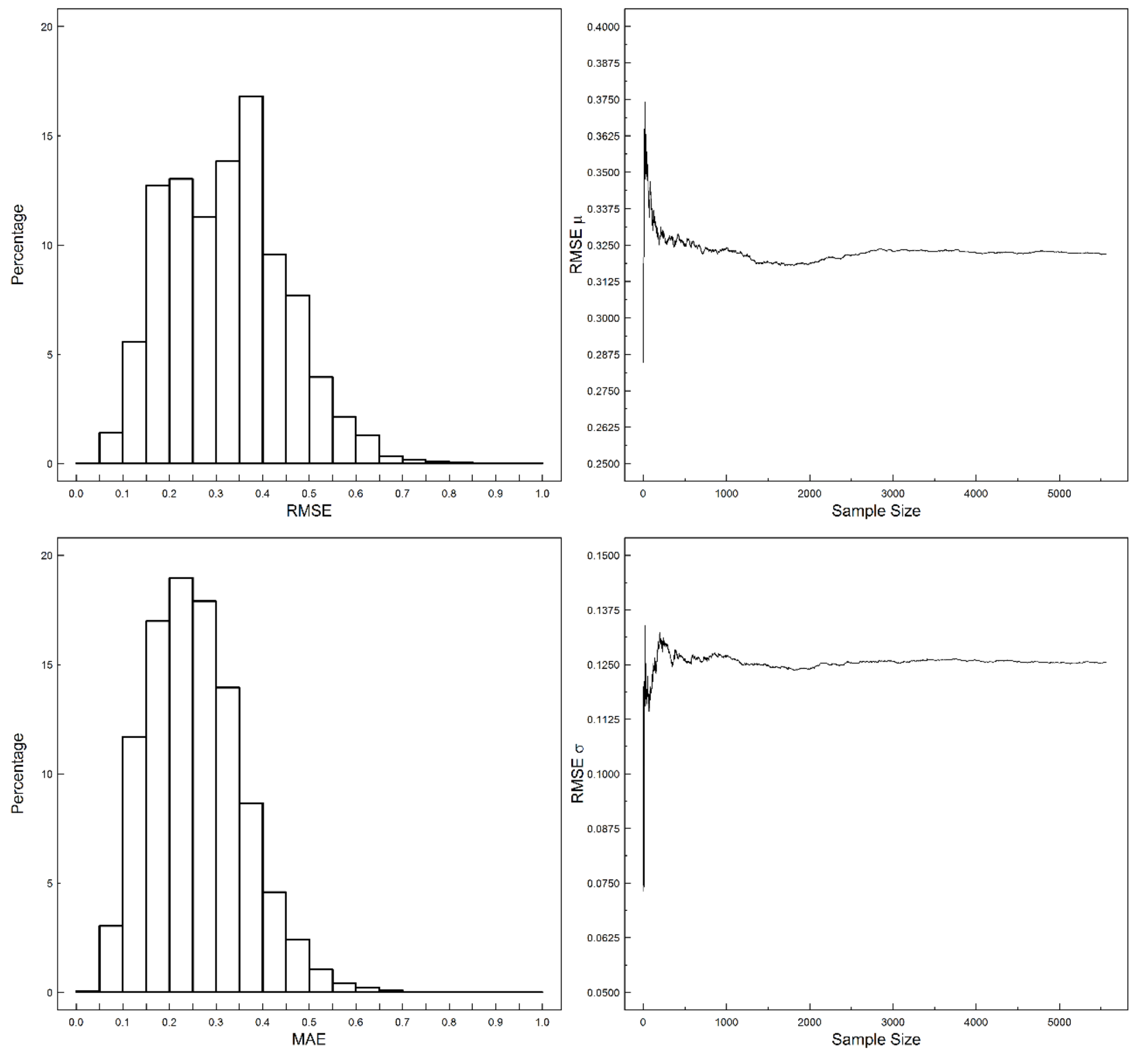
Appendix L. Results of Numerical Experiment #2

Including the five wells presented in the main body of the chapter since their convergence plots was not presented.

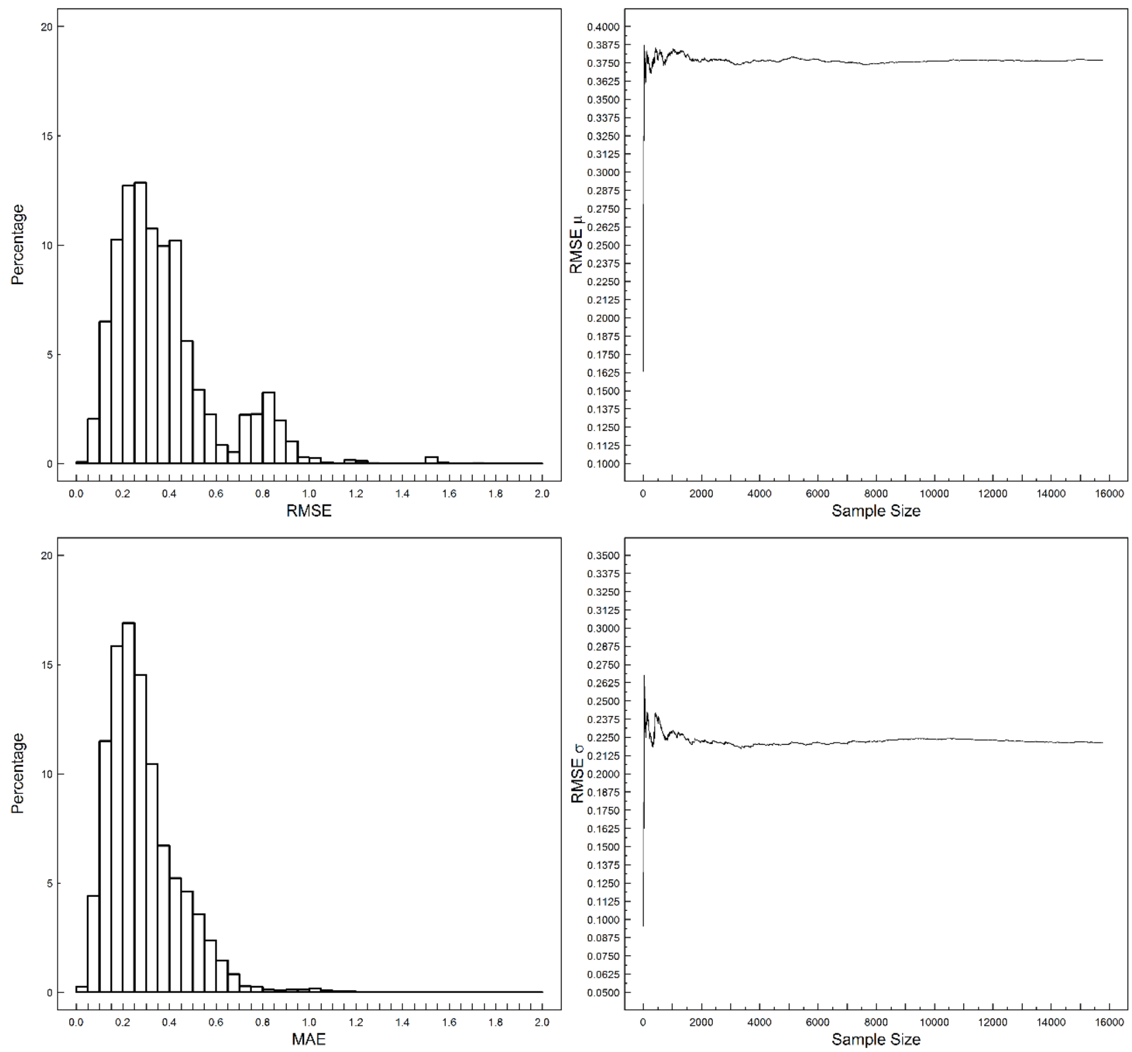
Well ID: 02S6535DCD



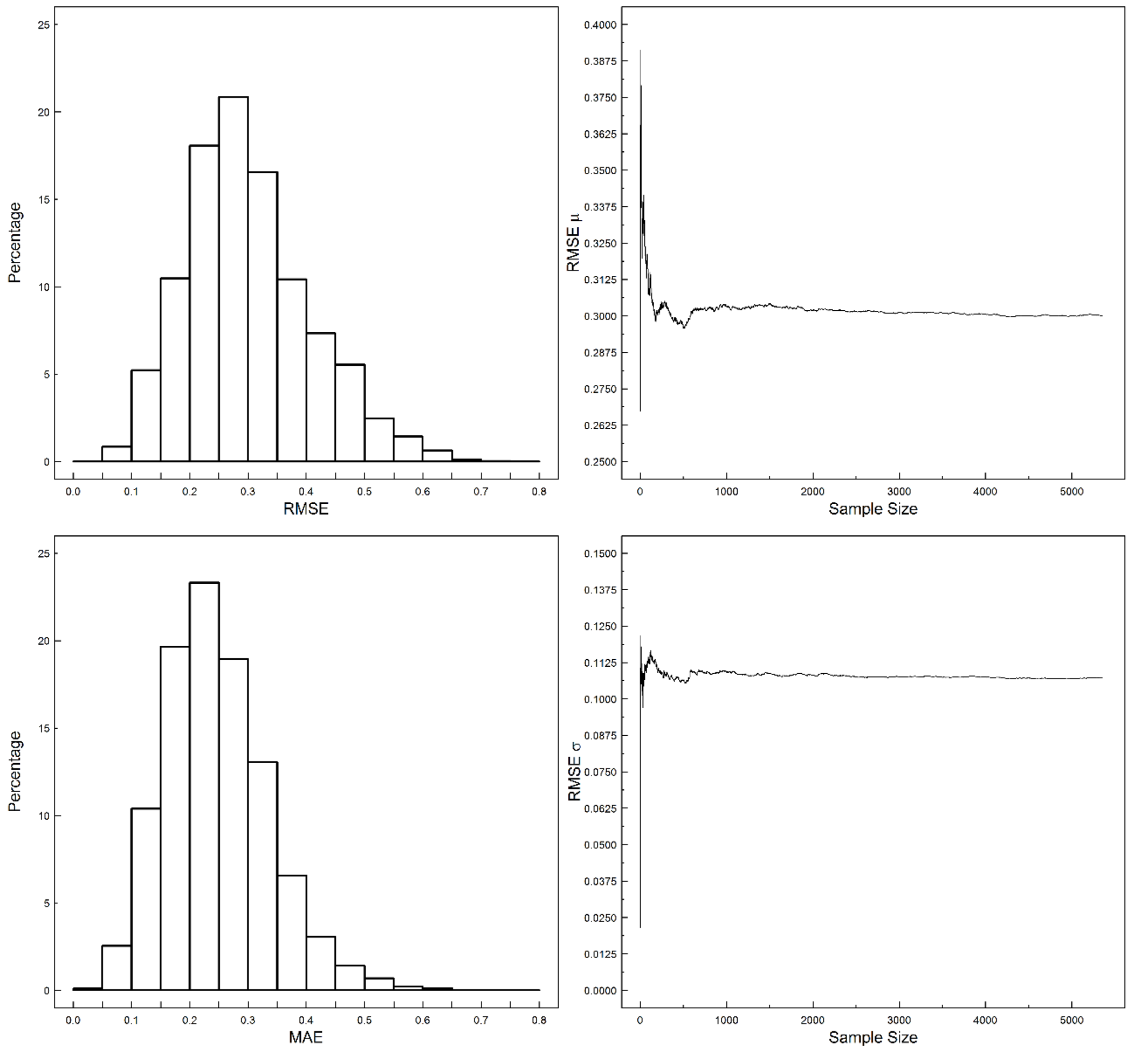
Well ID: 02S6523ADC



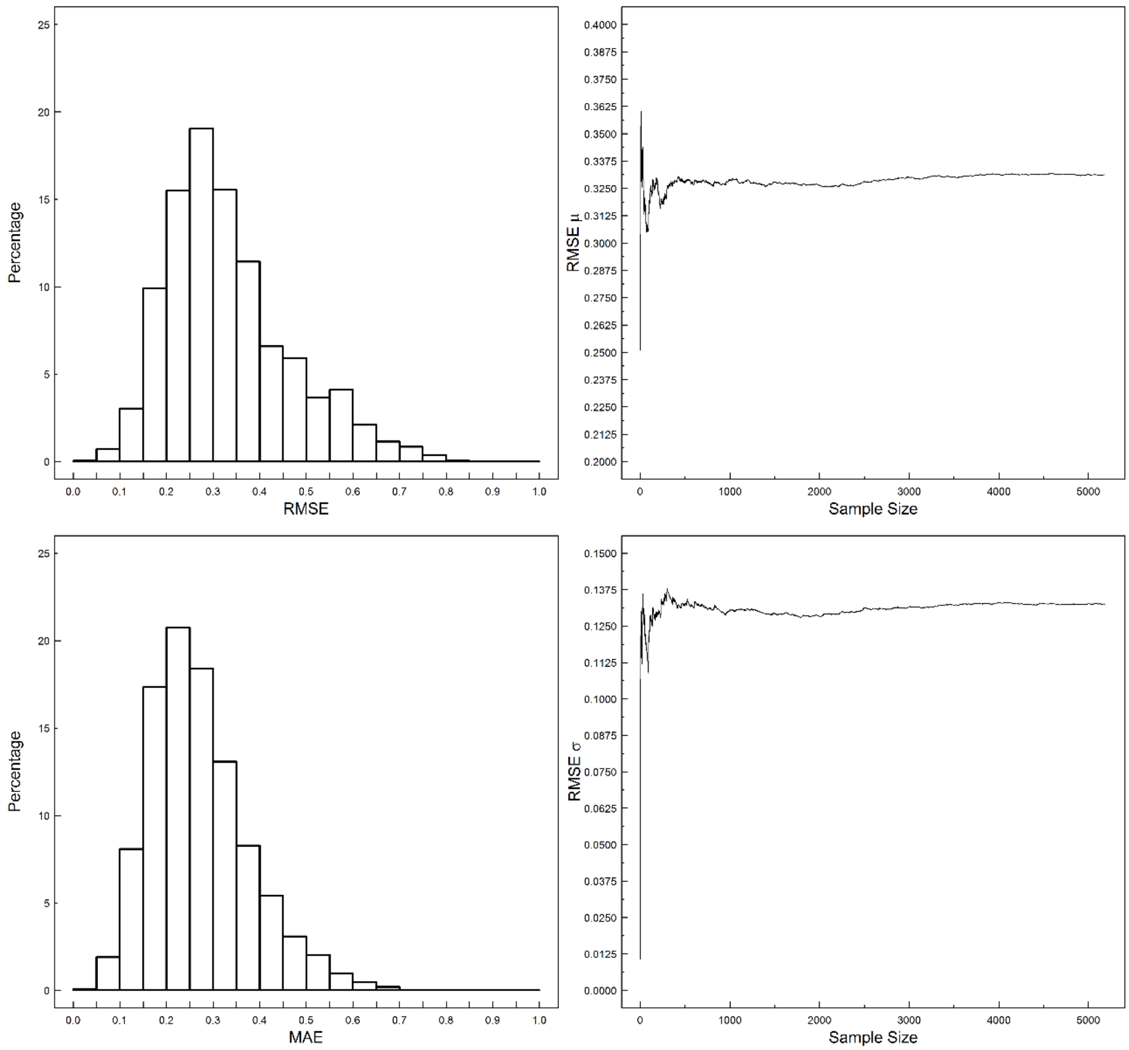
Well ID: 01N6525CCD



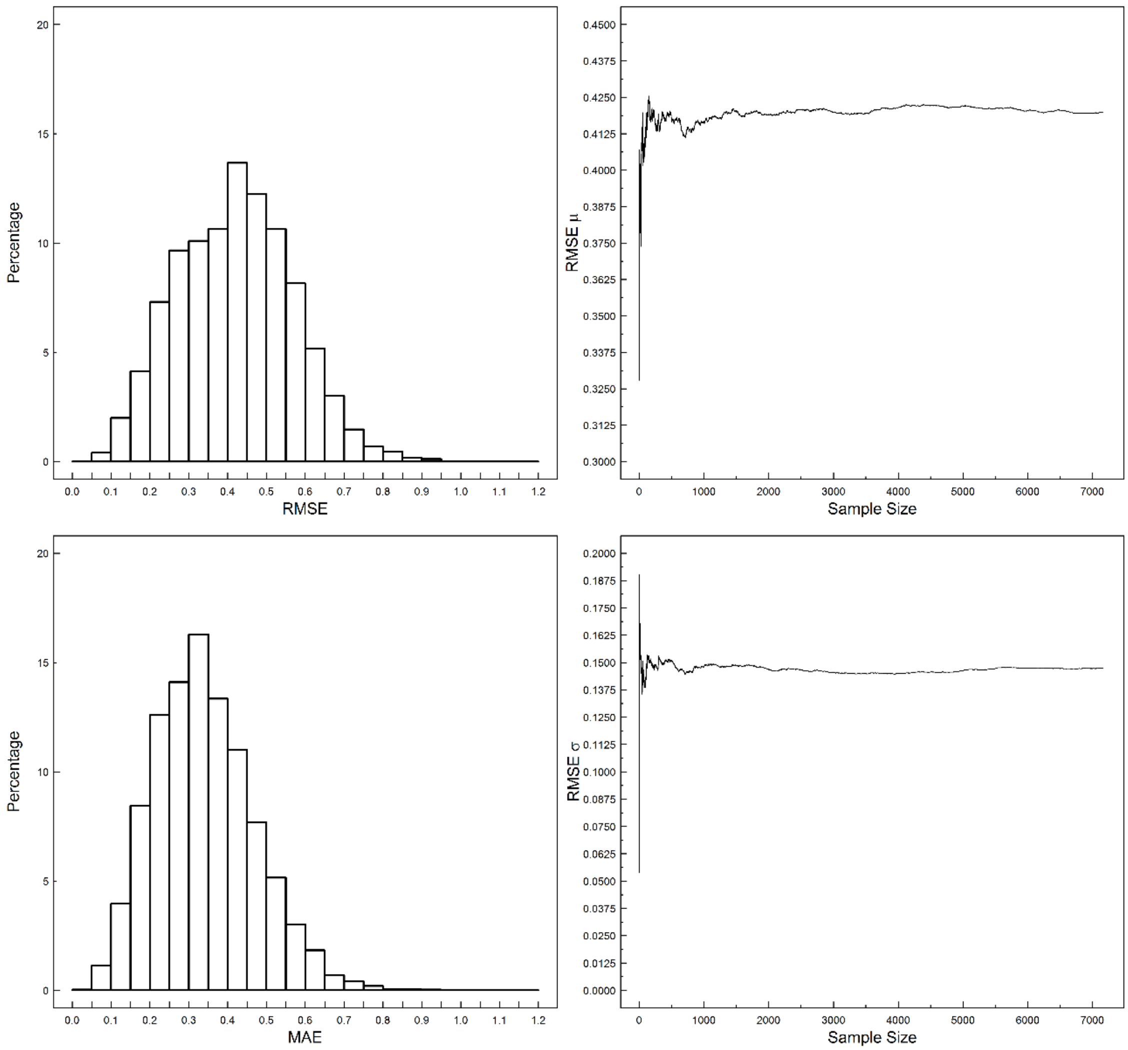
Well ID: 03N6618CAC1



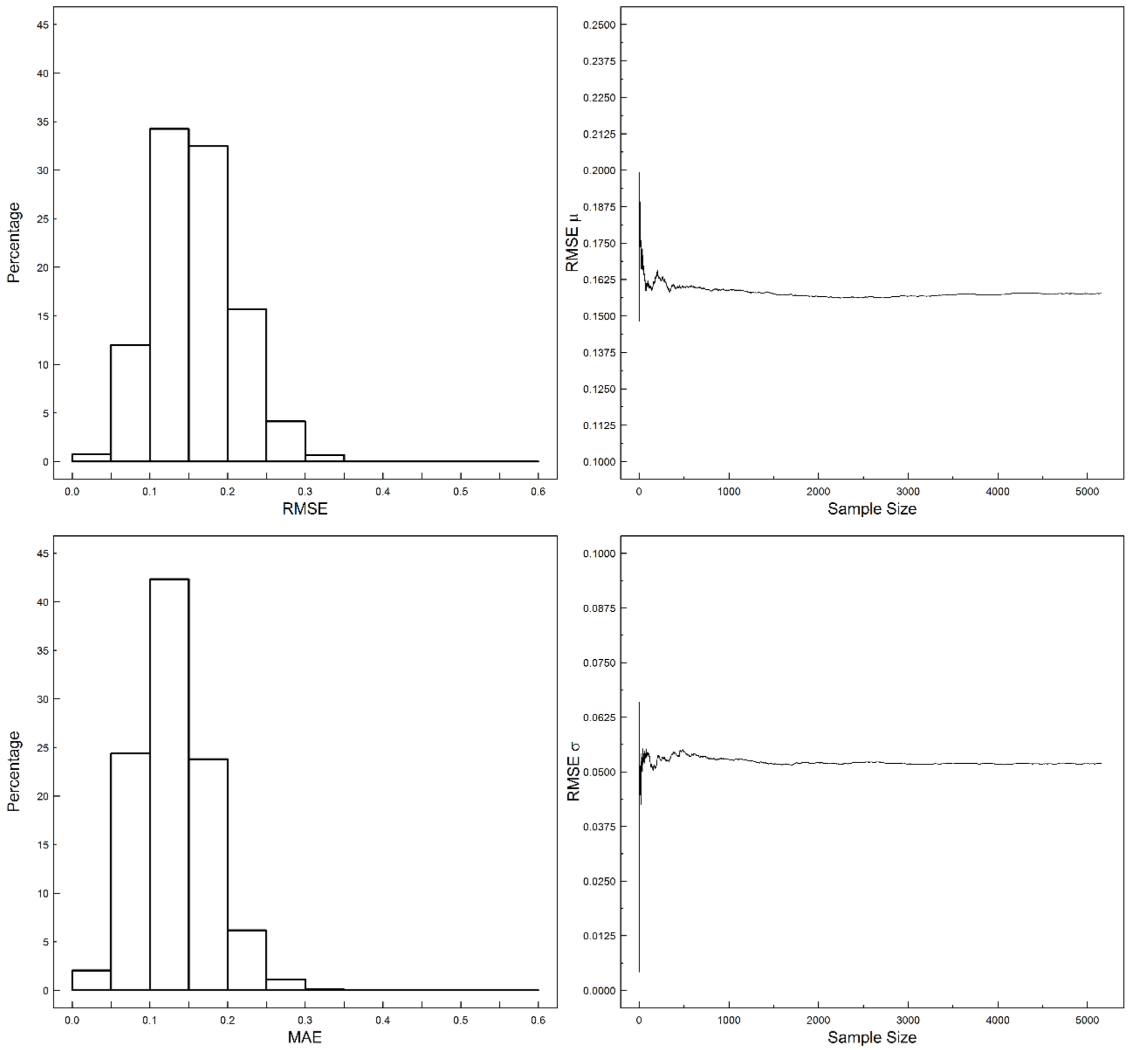
Well ID: 04N6614BAA



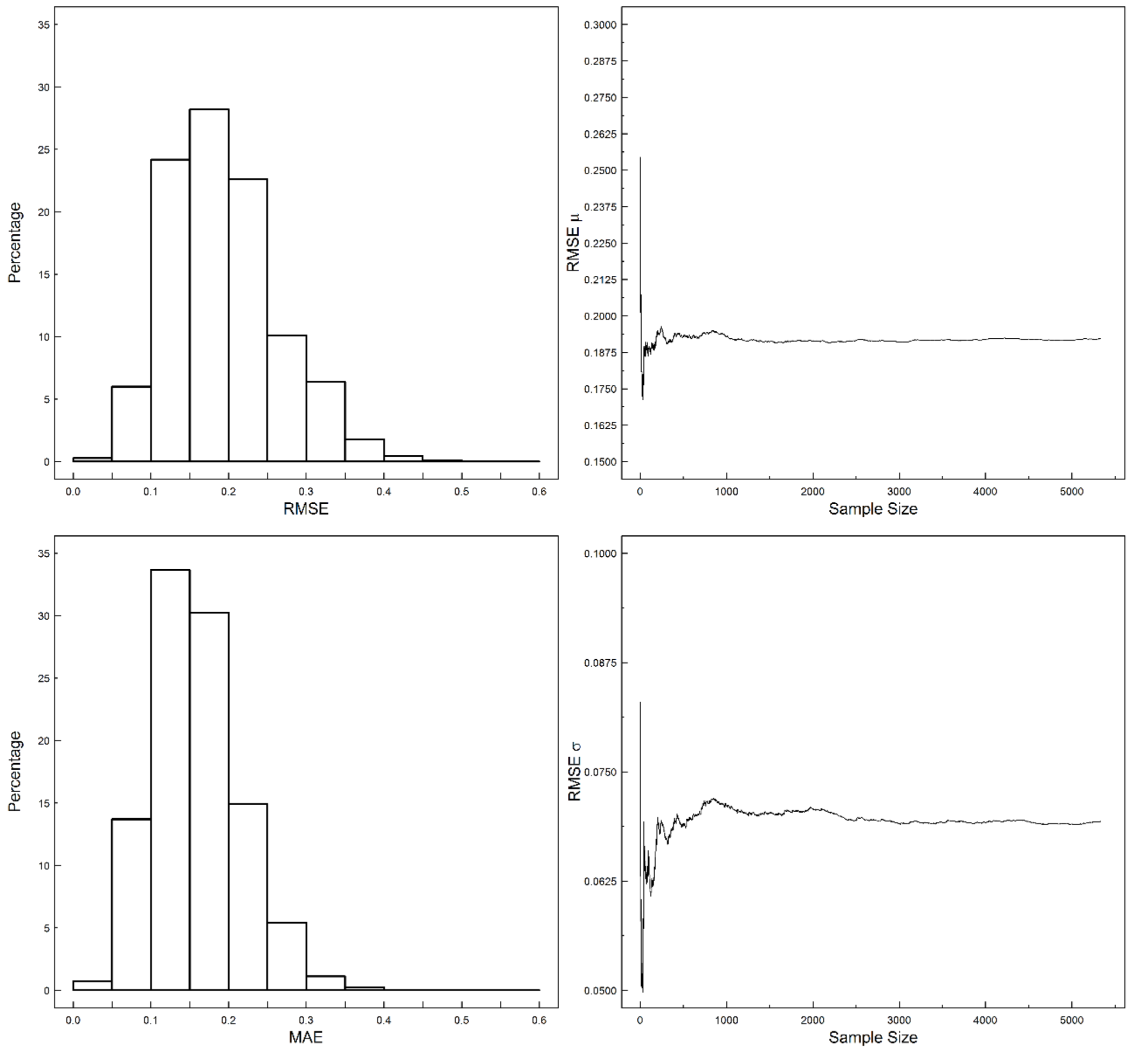
Well ID: 04N6401CCC



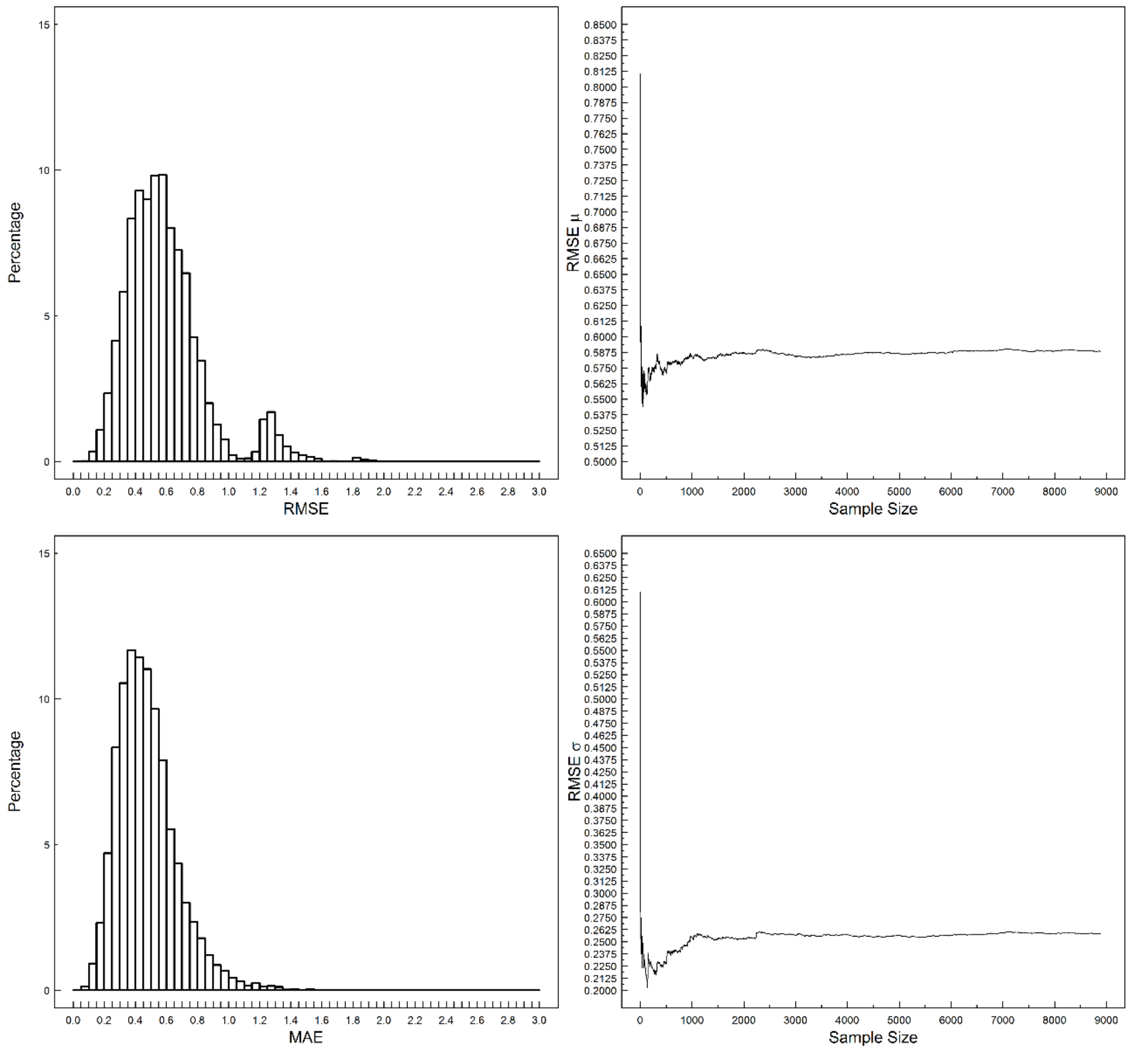
Well ID: 10N4902CBC



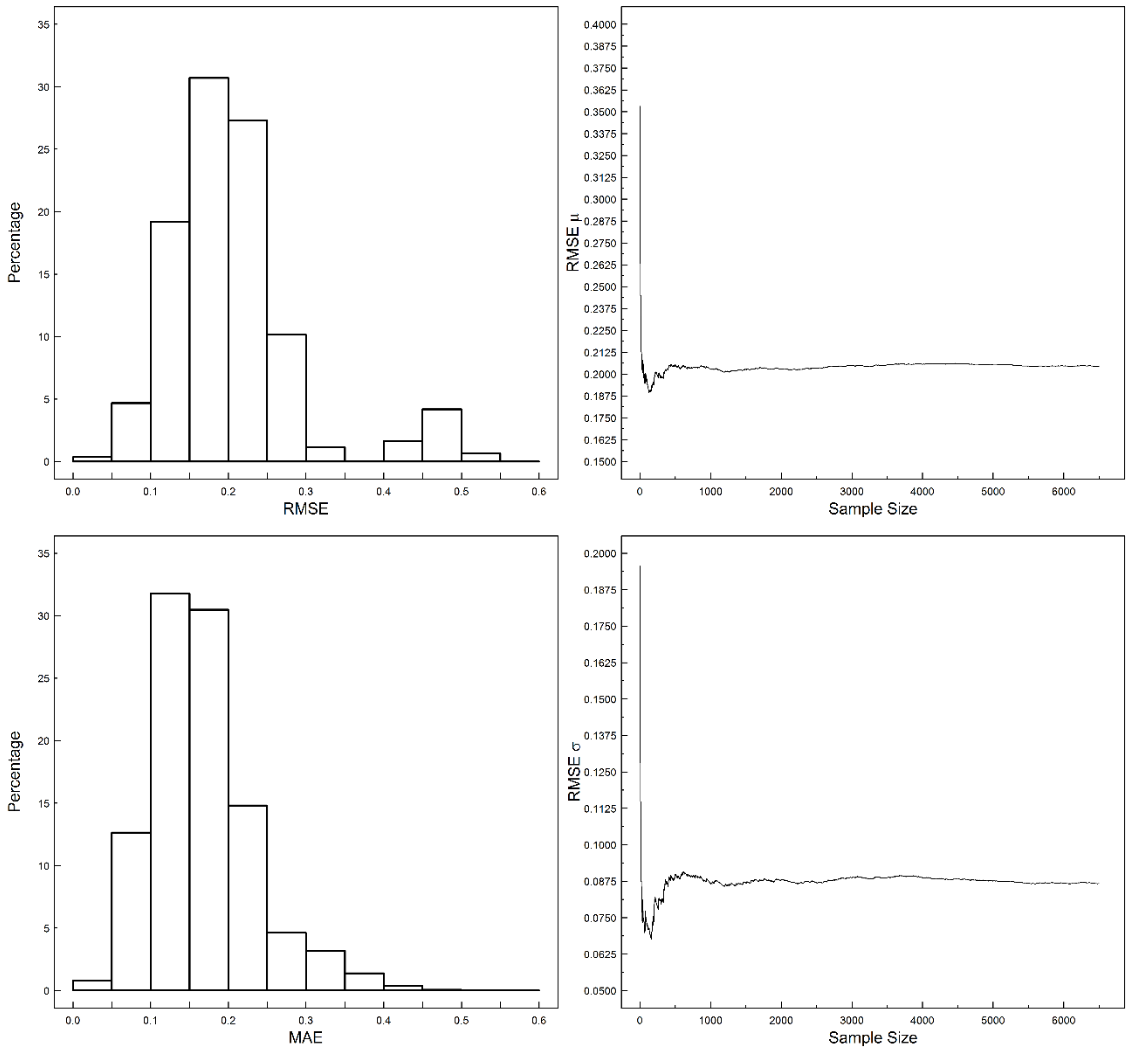
Well ID: 11N4728BBB



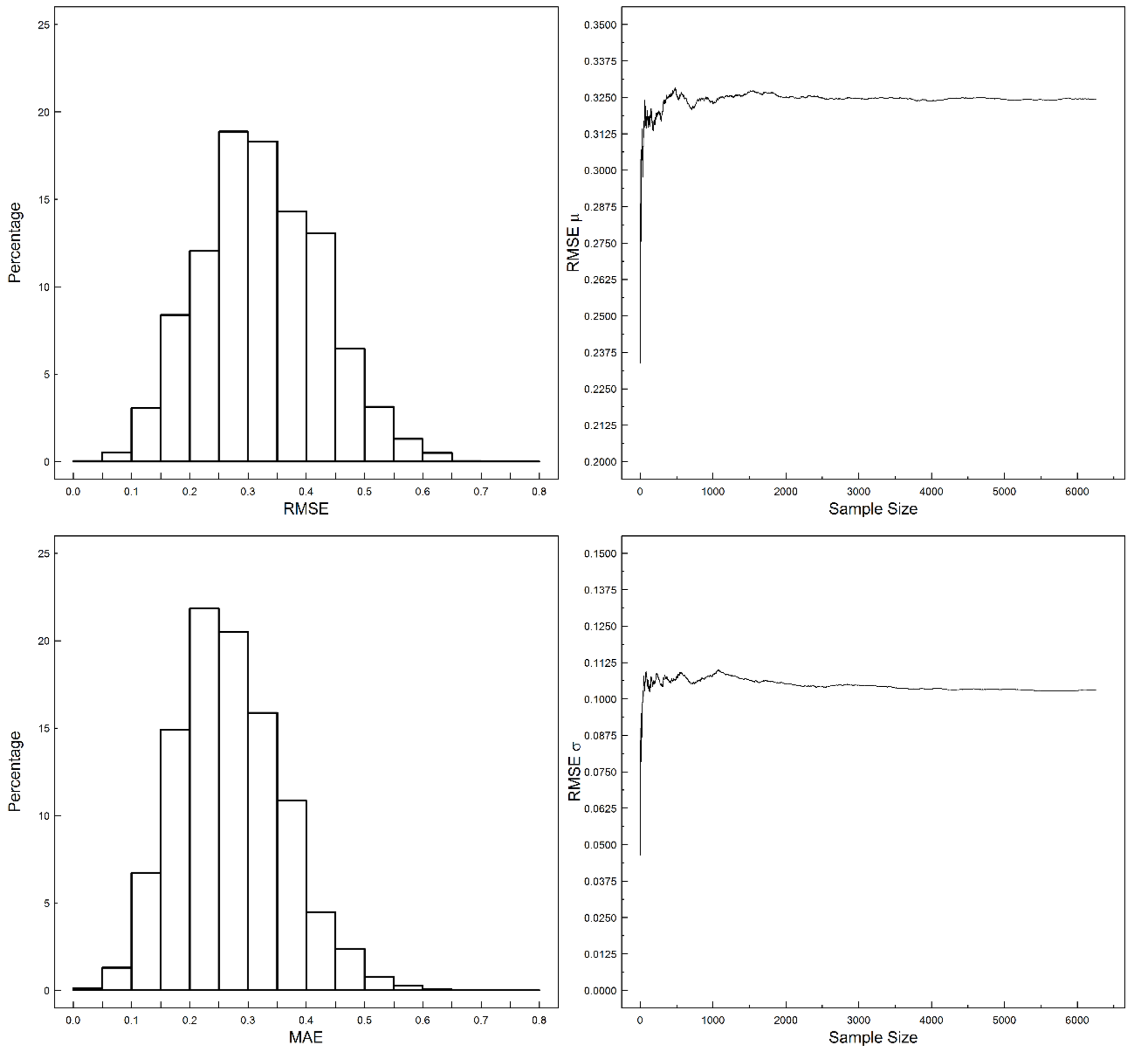
Well ID: 04N5931CB



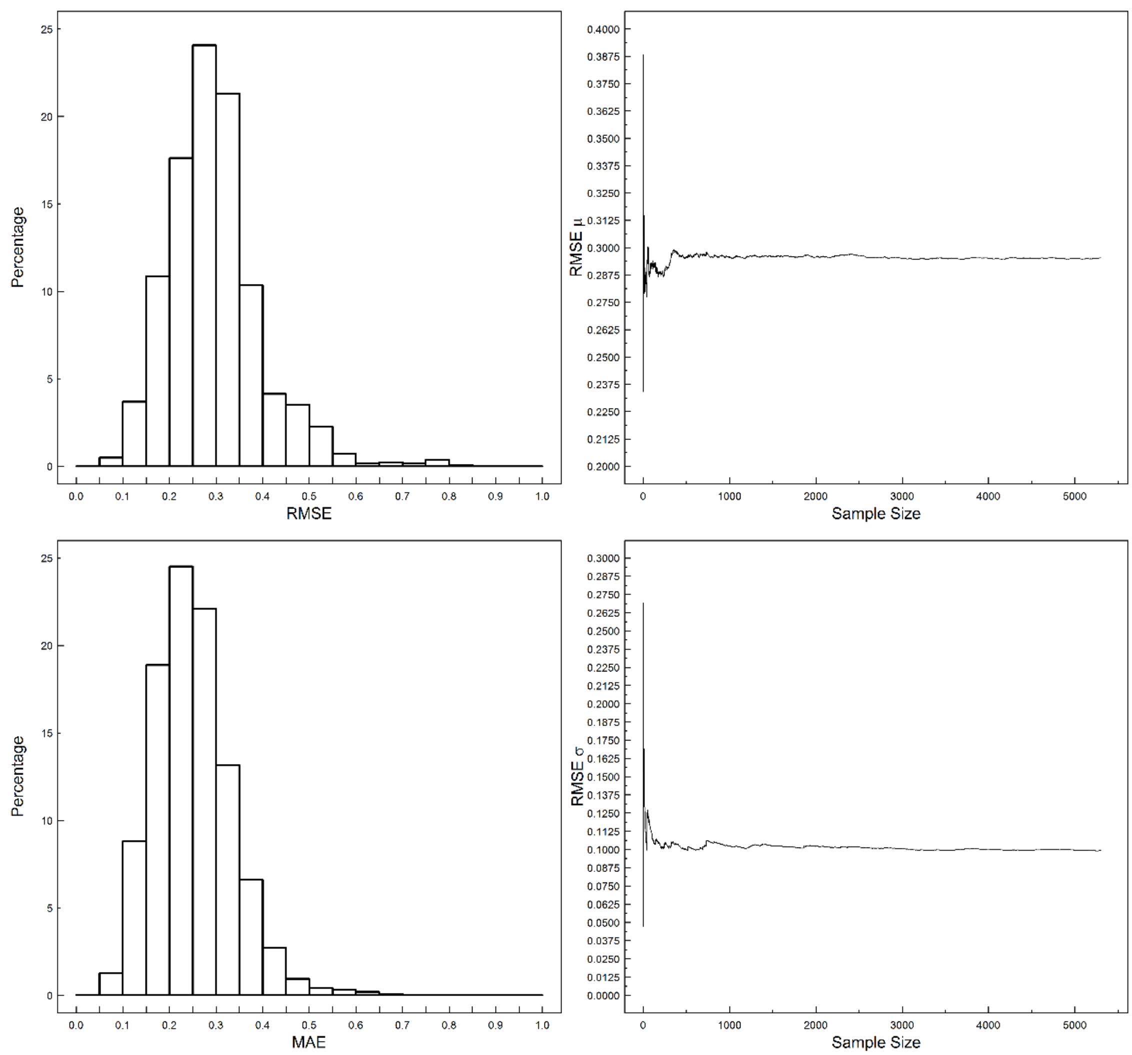
Well ID: 09N5131BBB



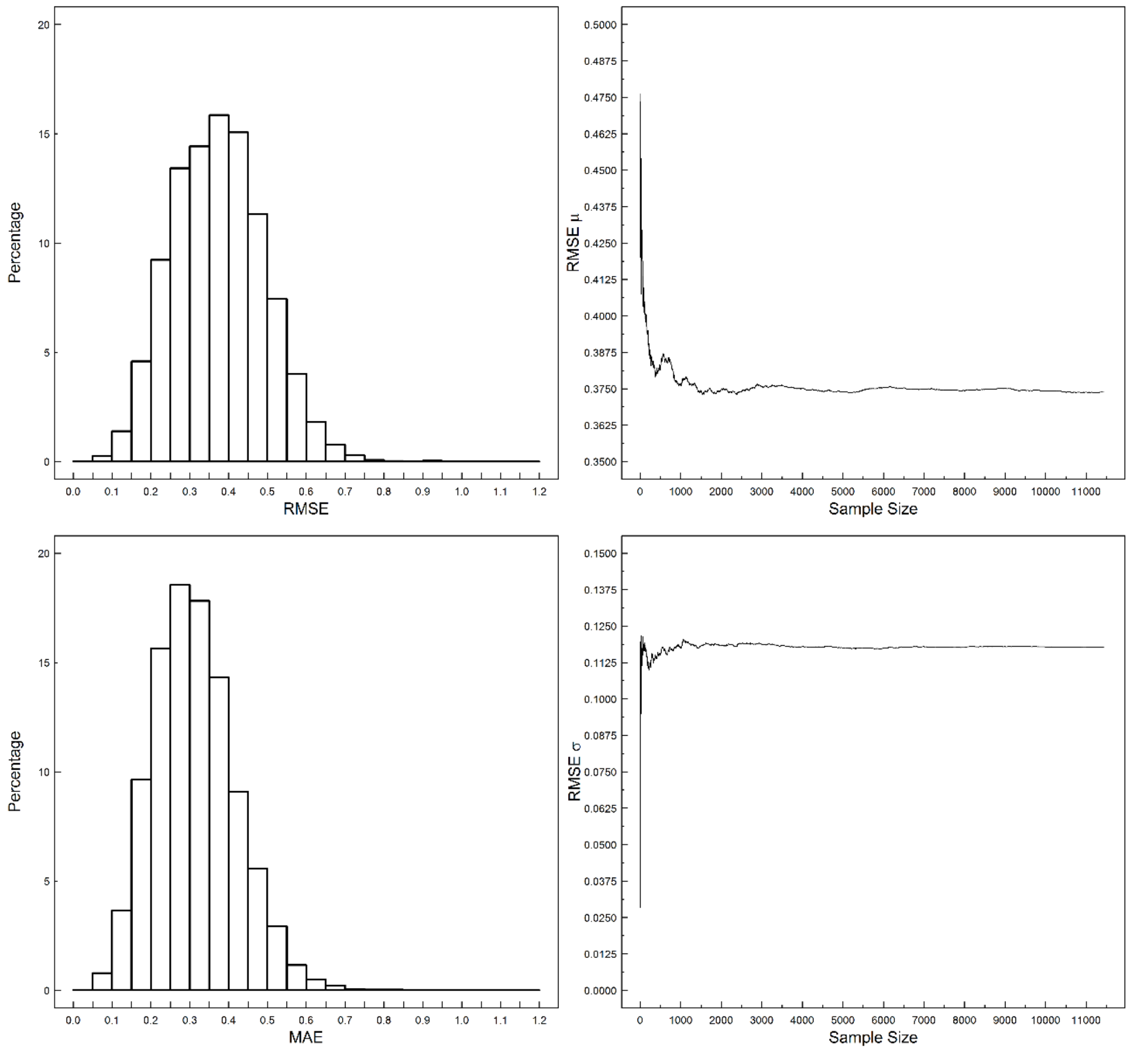
Well ID: 04N6506DAB



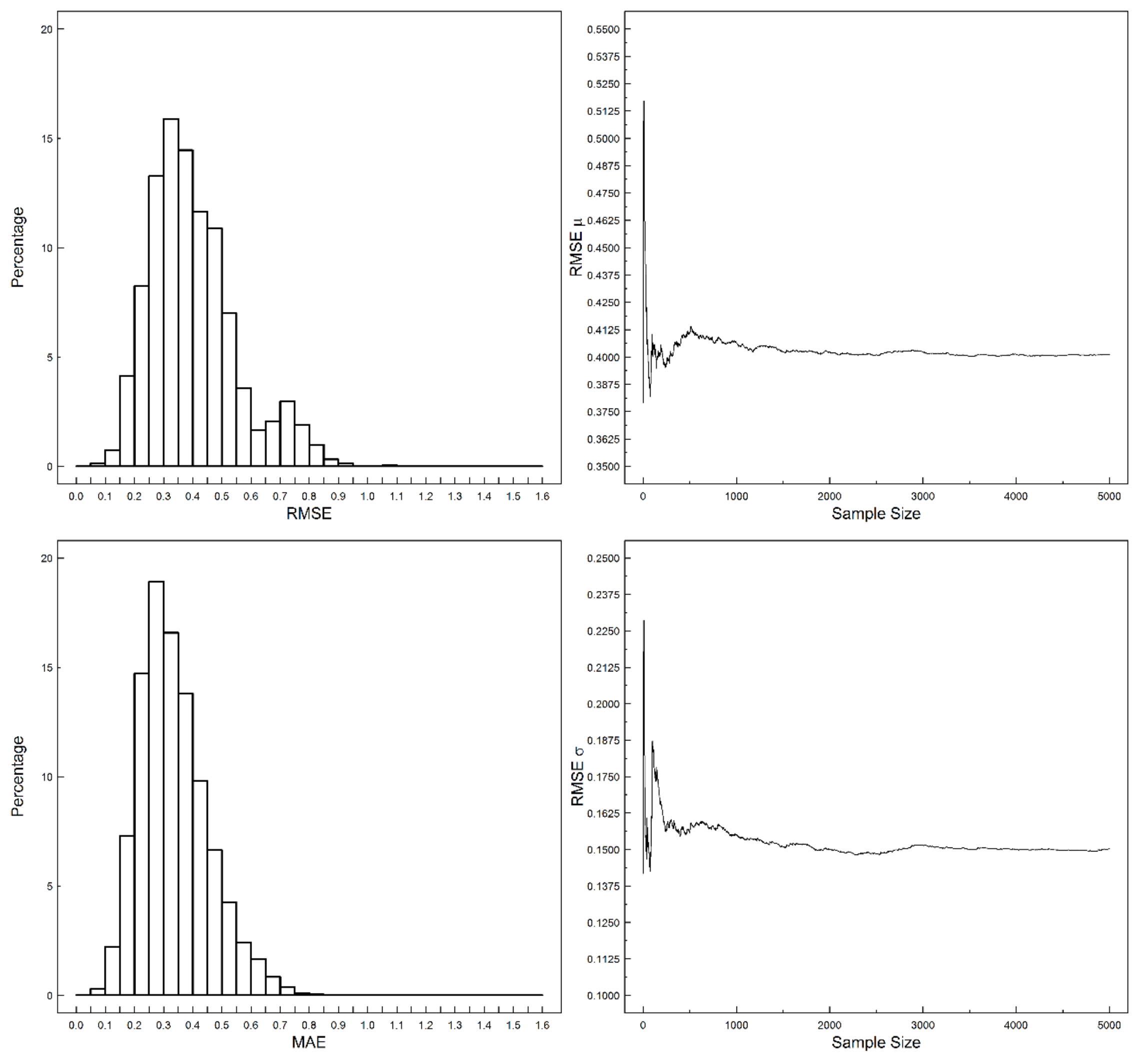
Well ID: 01N5531DCD



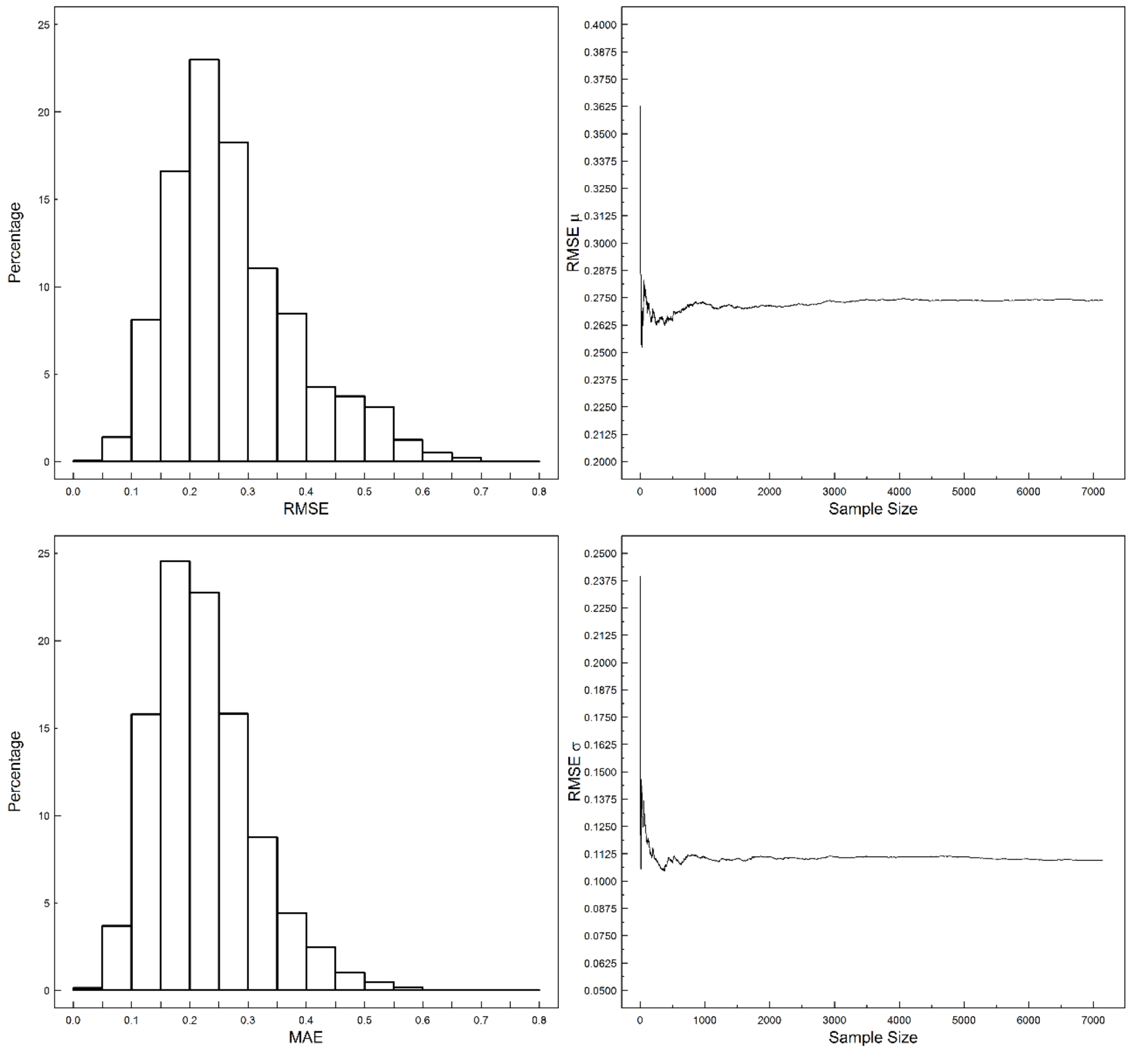
Well ID: 04N6422DCD



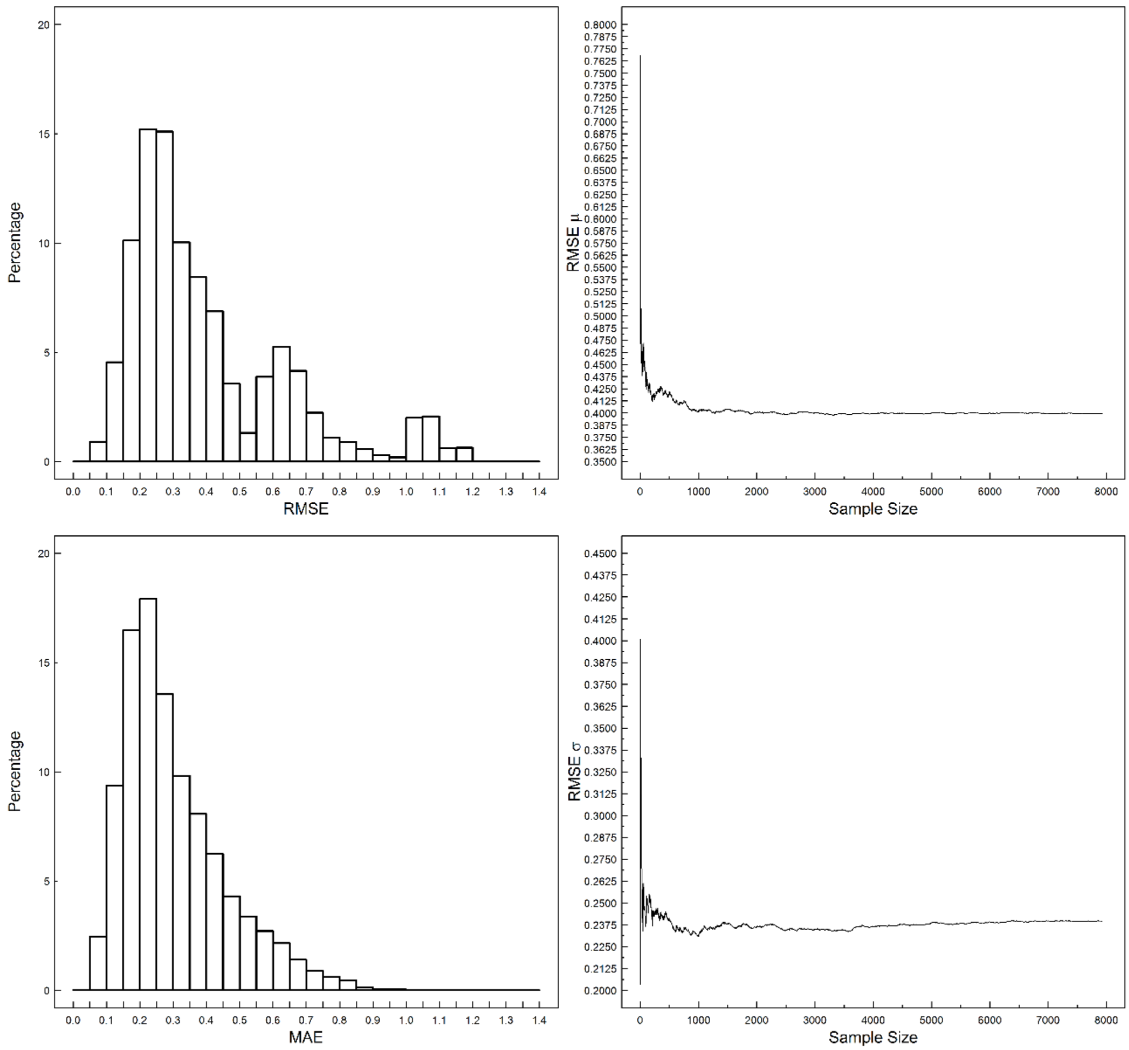
Well ID: 04N6412CCC



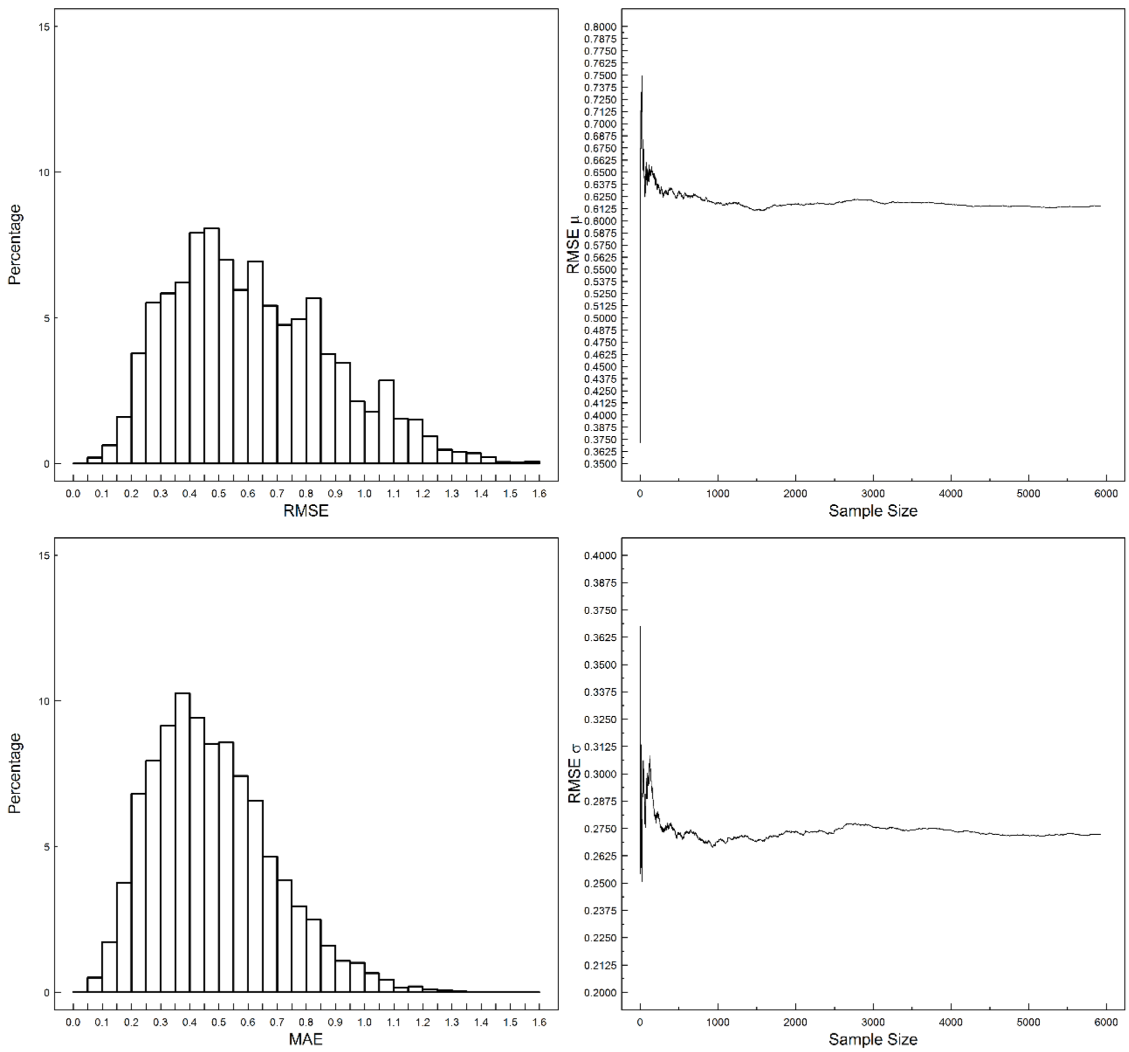
Well ID: SB00306618CAC



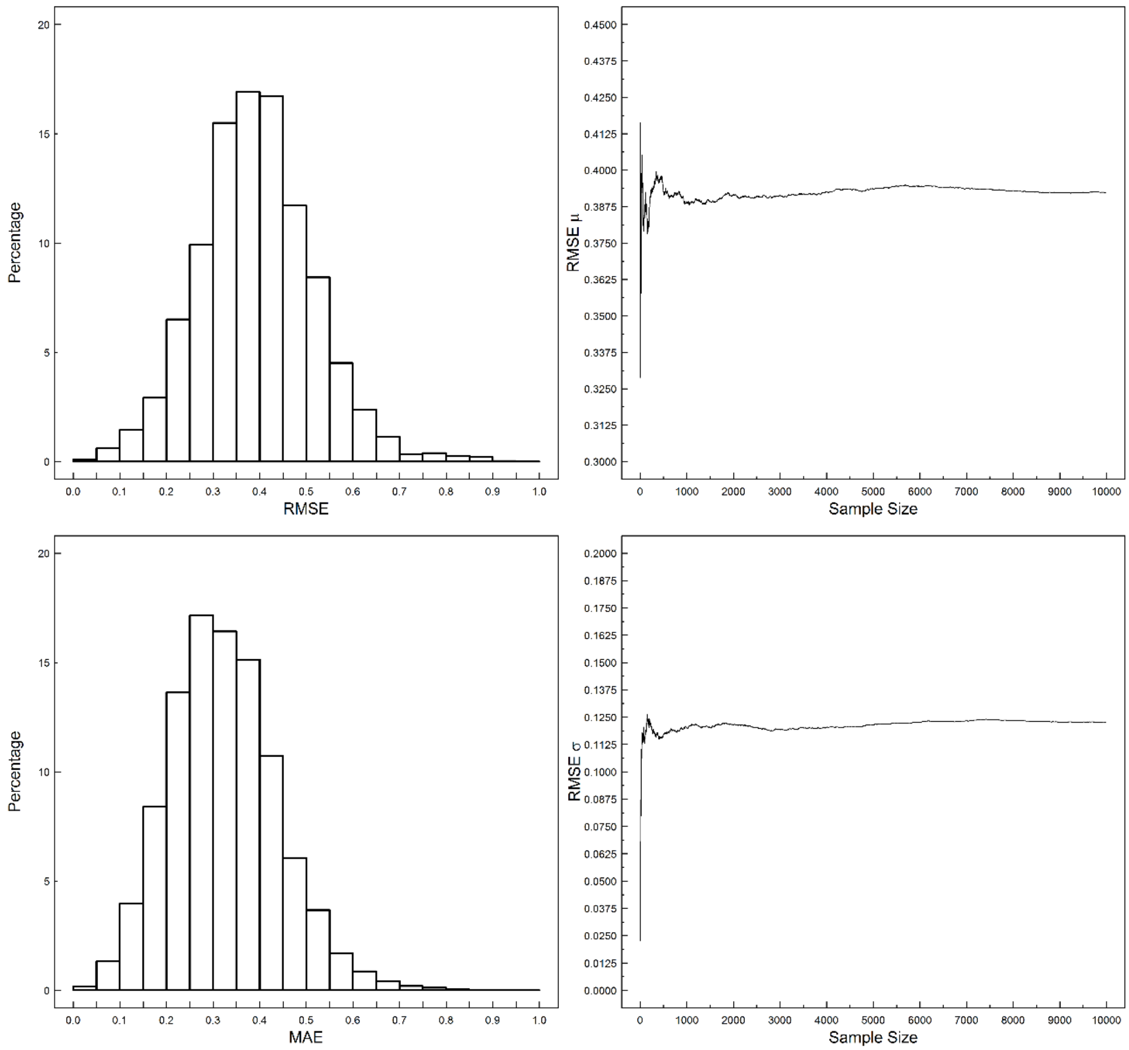
Well ID: 04N6627ADD



Well ID: SB00406012CCC-CSU1



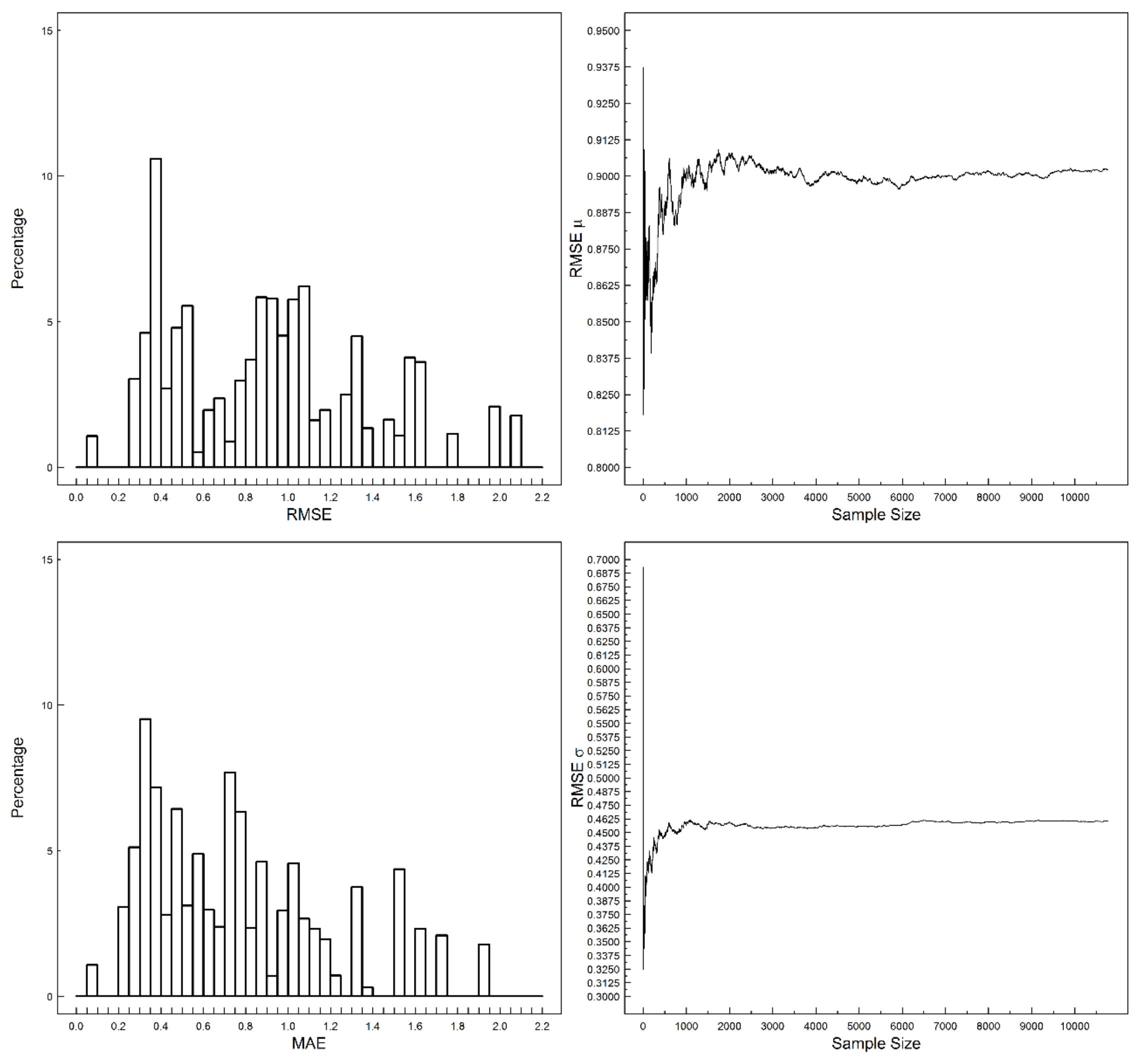
Well ID: 68-1


Percentage
RMSE
RMSE μ
Sample Size
Percentage
MAE
RMSE σ
Sample Size

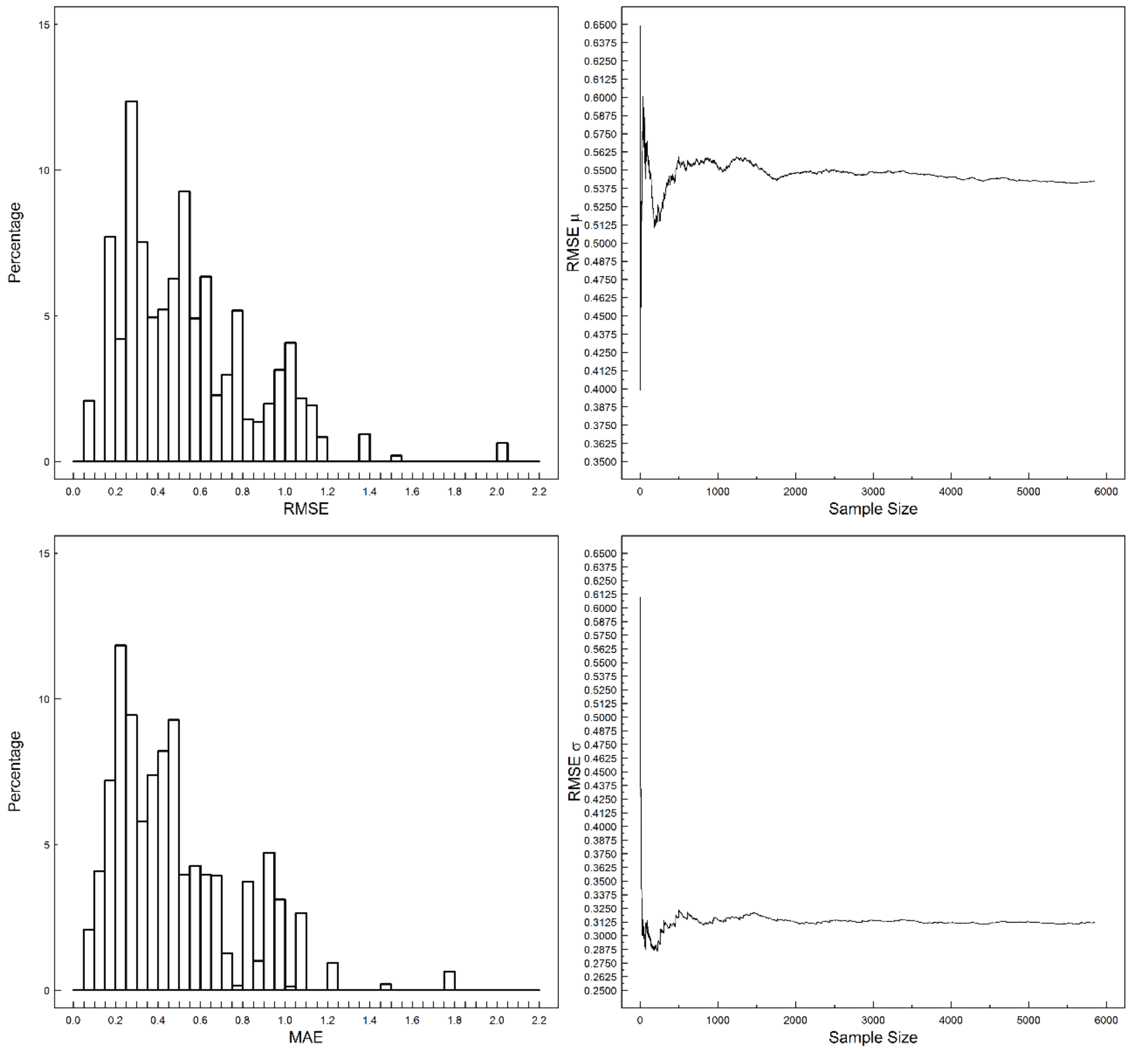
Appendix M. Results of Numerical Experiment #3

Including the five wells presented in the main body of the chapter since their convergence plots was not presented.

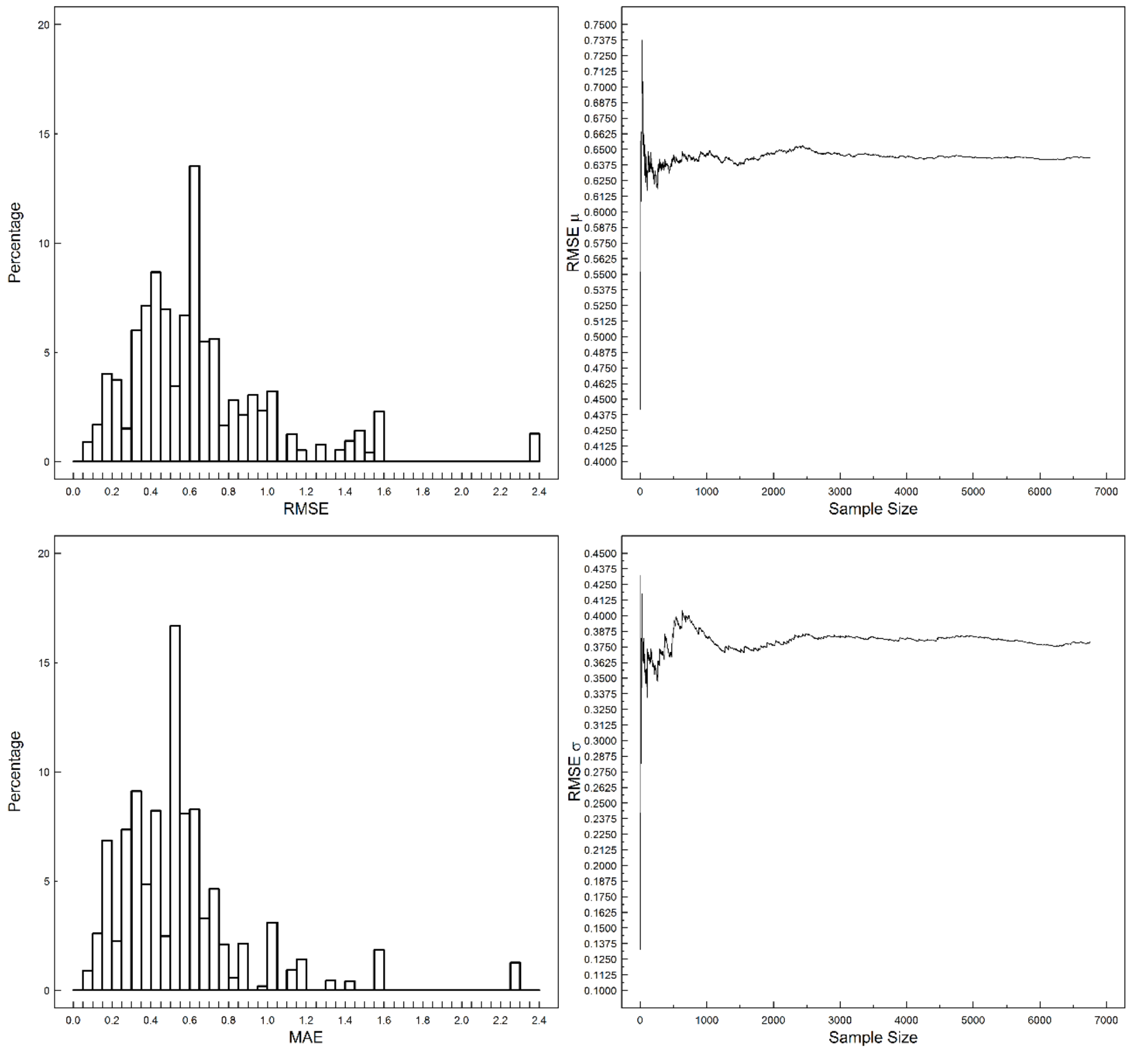
Well ID: 02S6535DCD



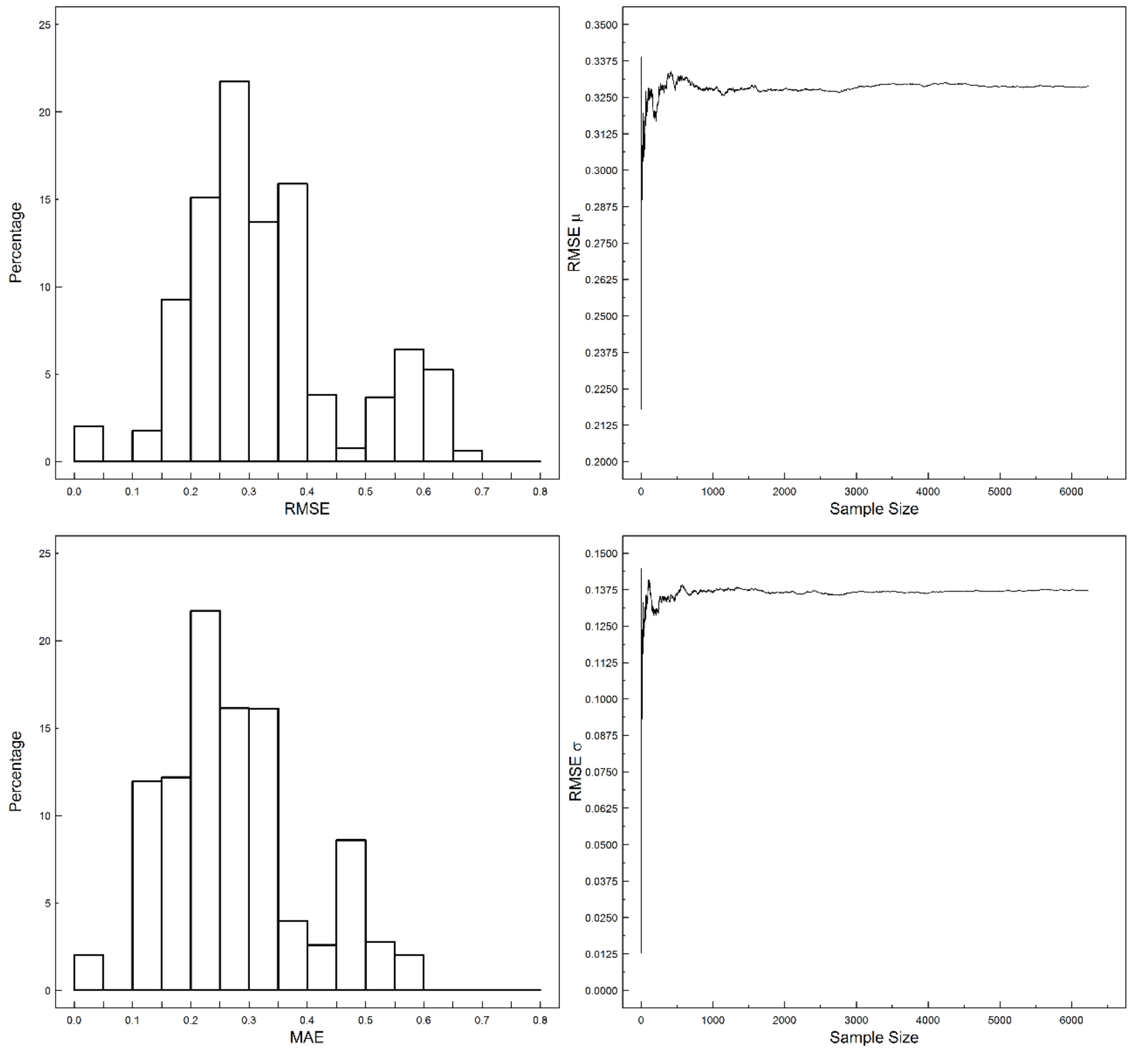
Well ID: 02S6523ADC



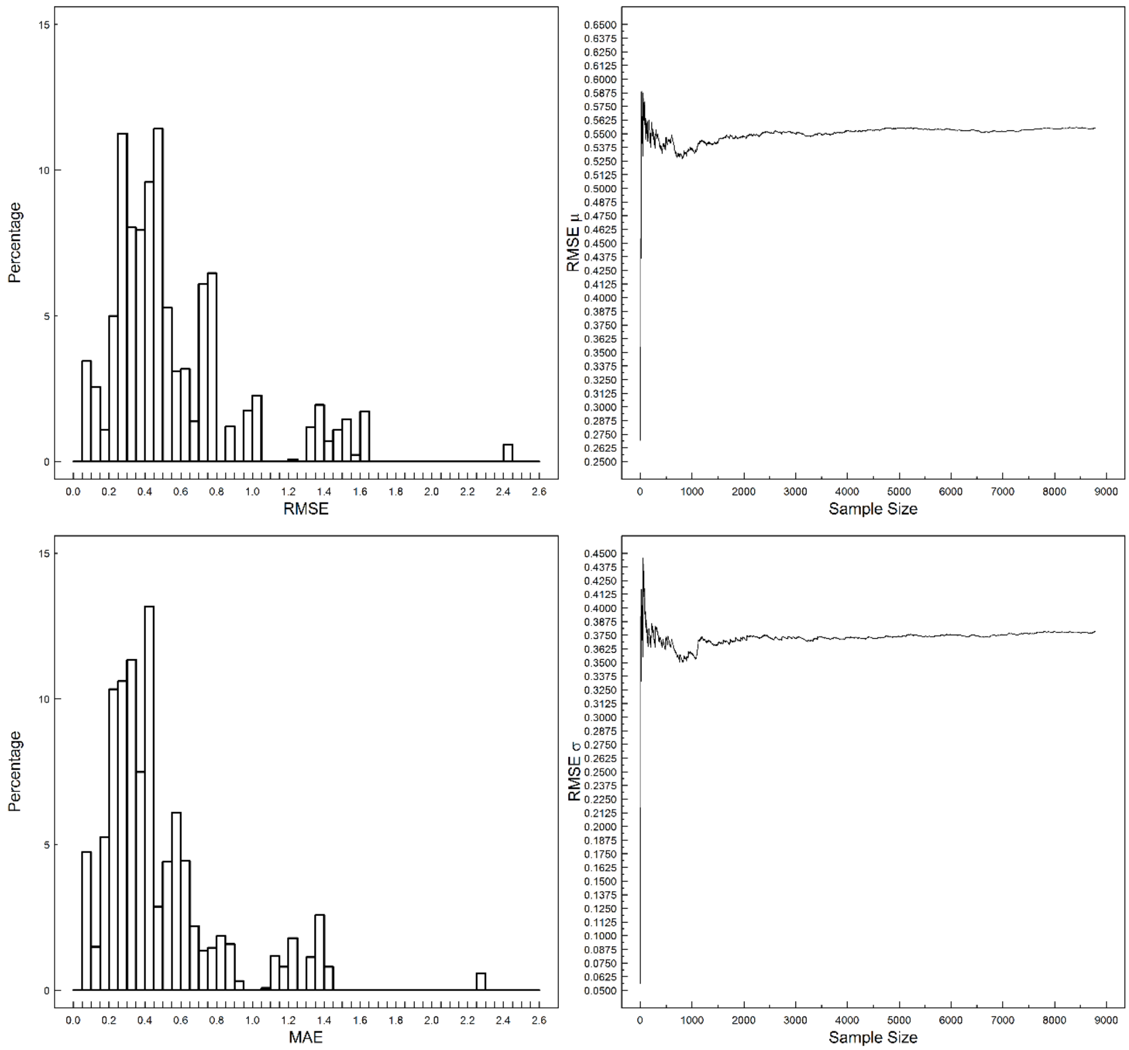
Well ID: 01N6525CCD



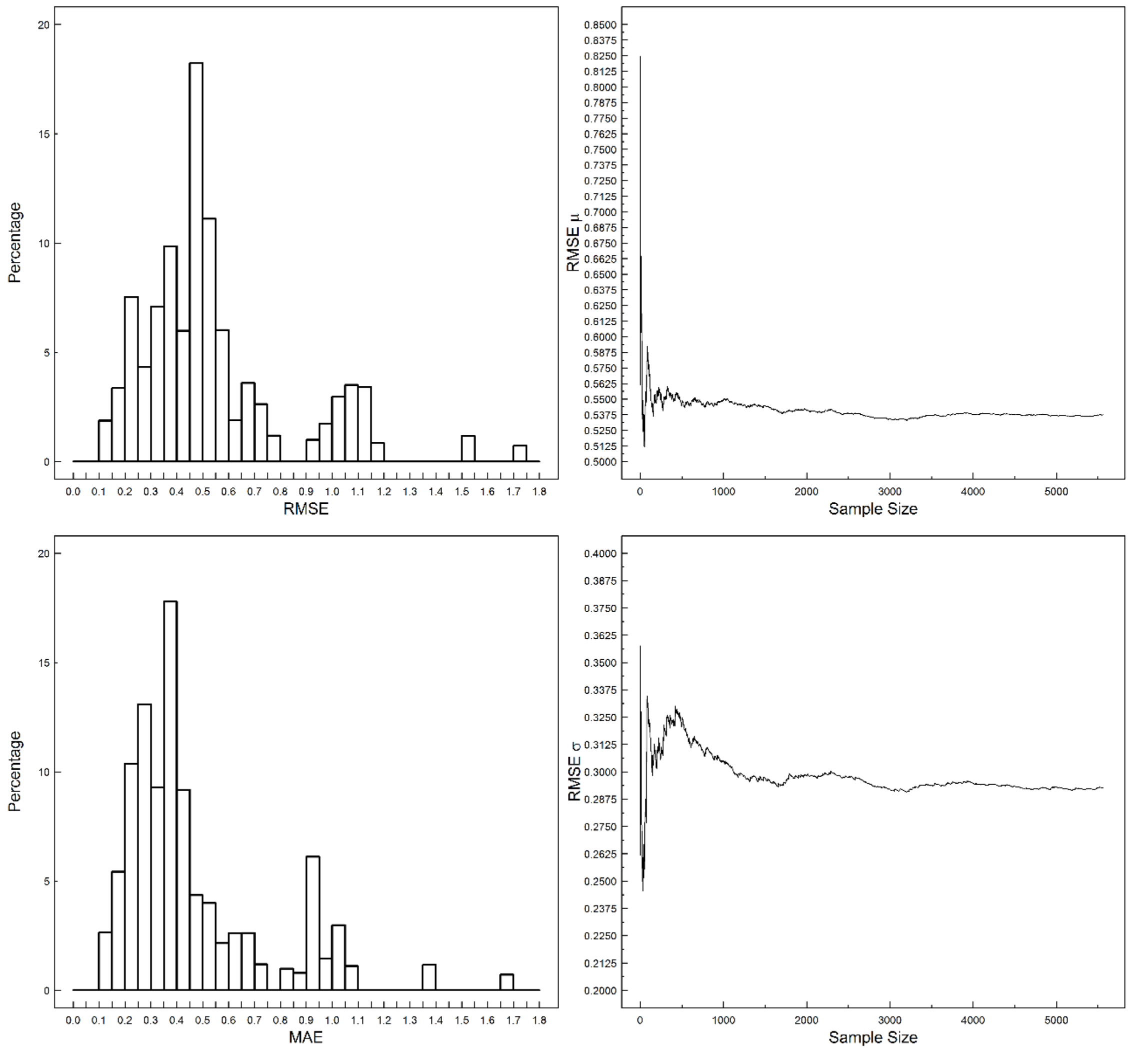
Well ID: 03N6618CAC1



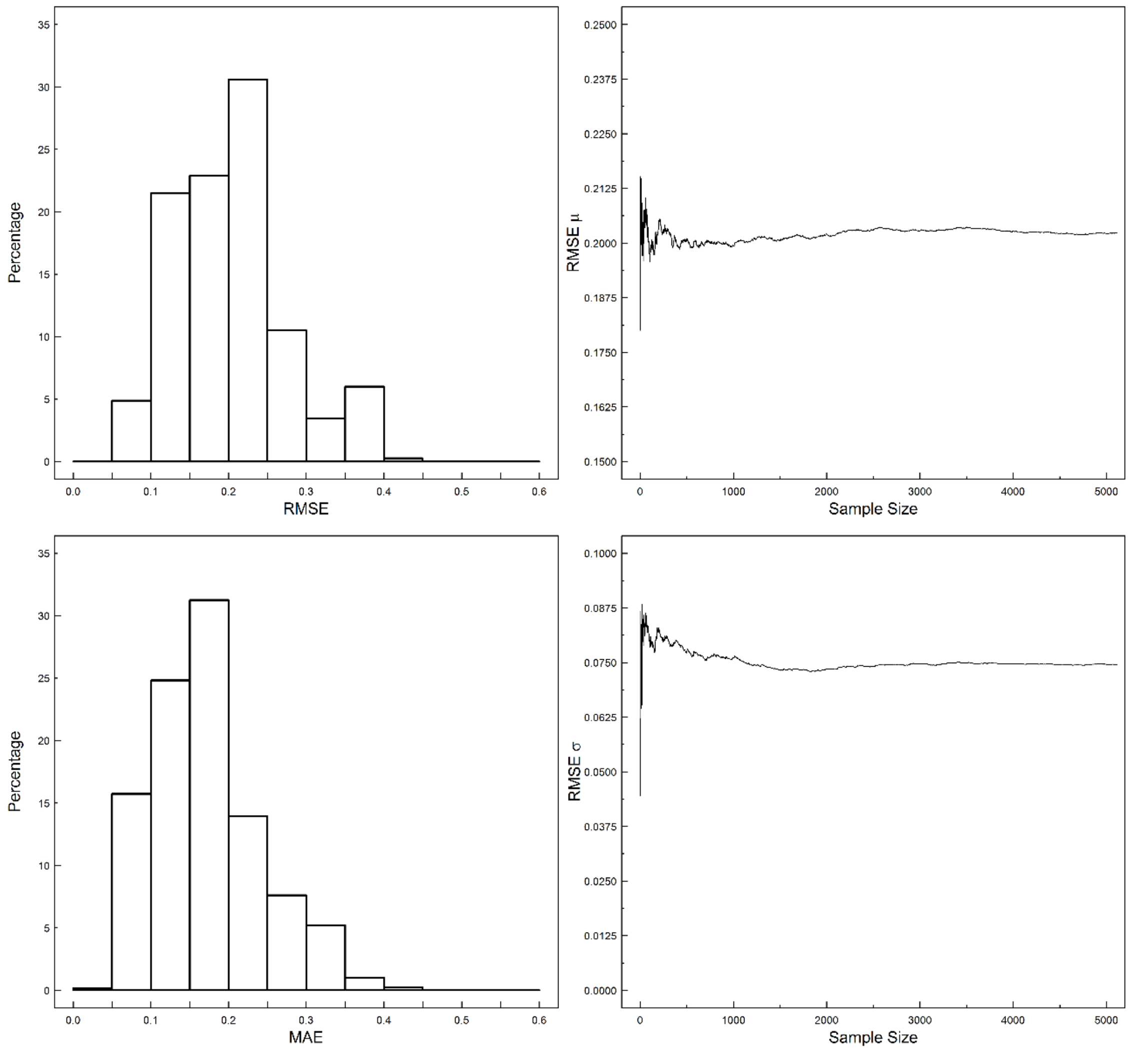
Well ID: 04N6614BAA



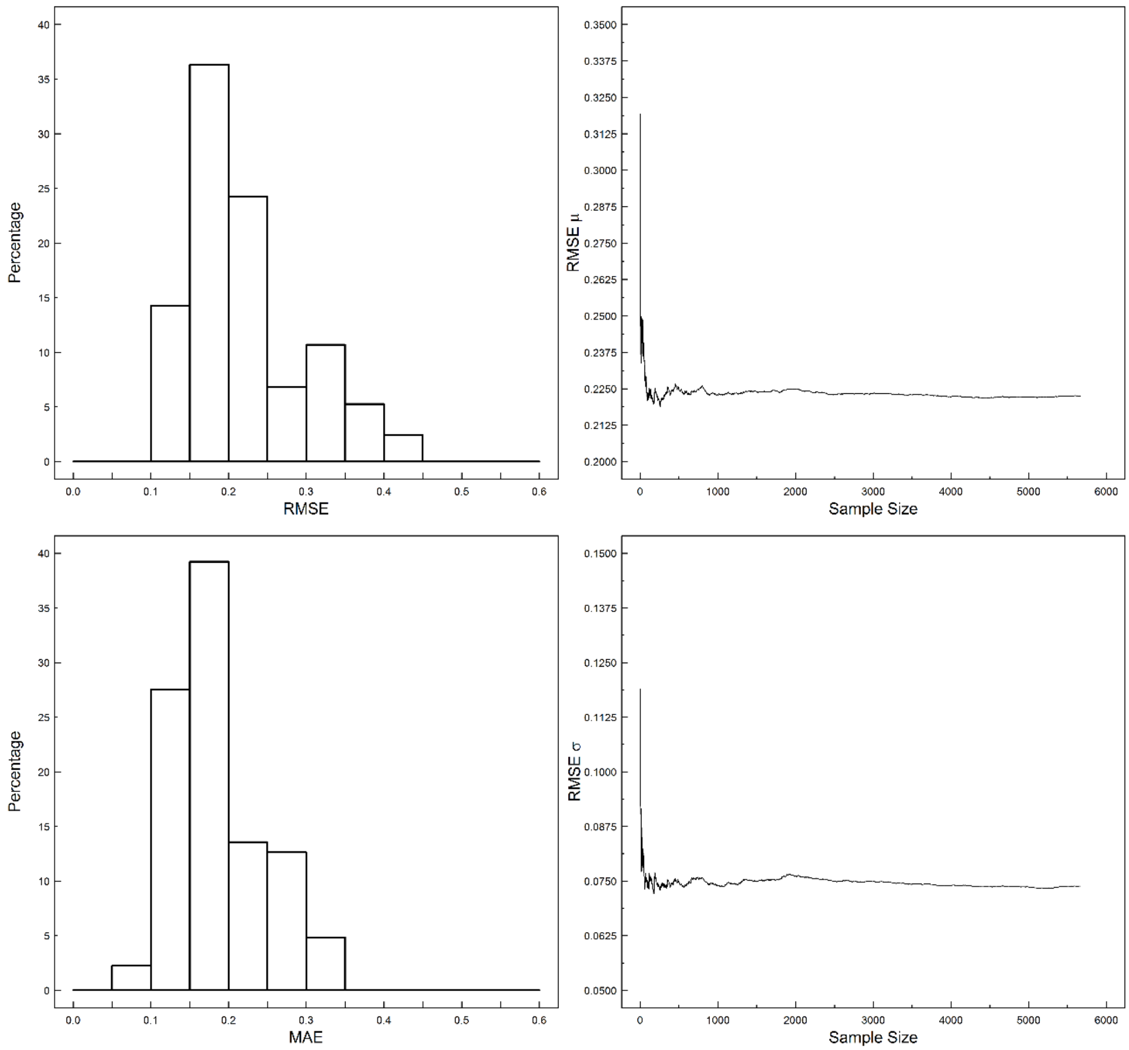
Well ID: 04N6401CCC



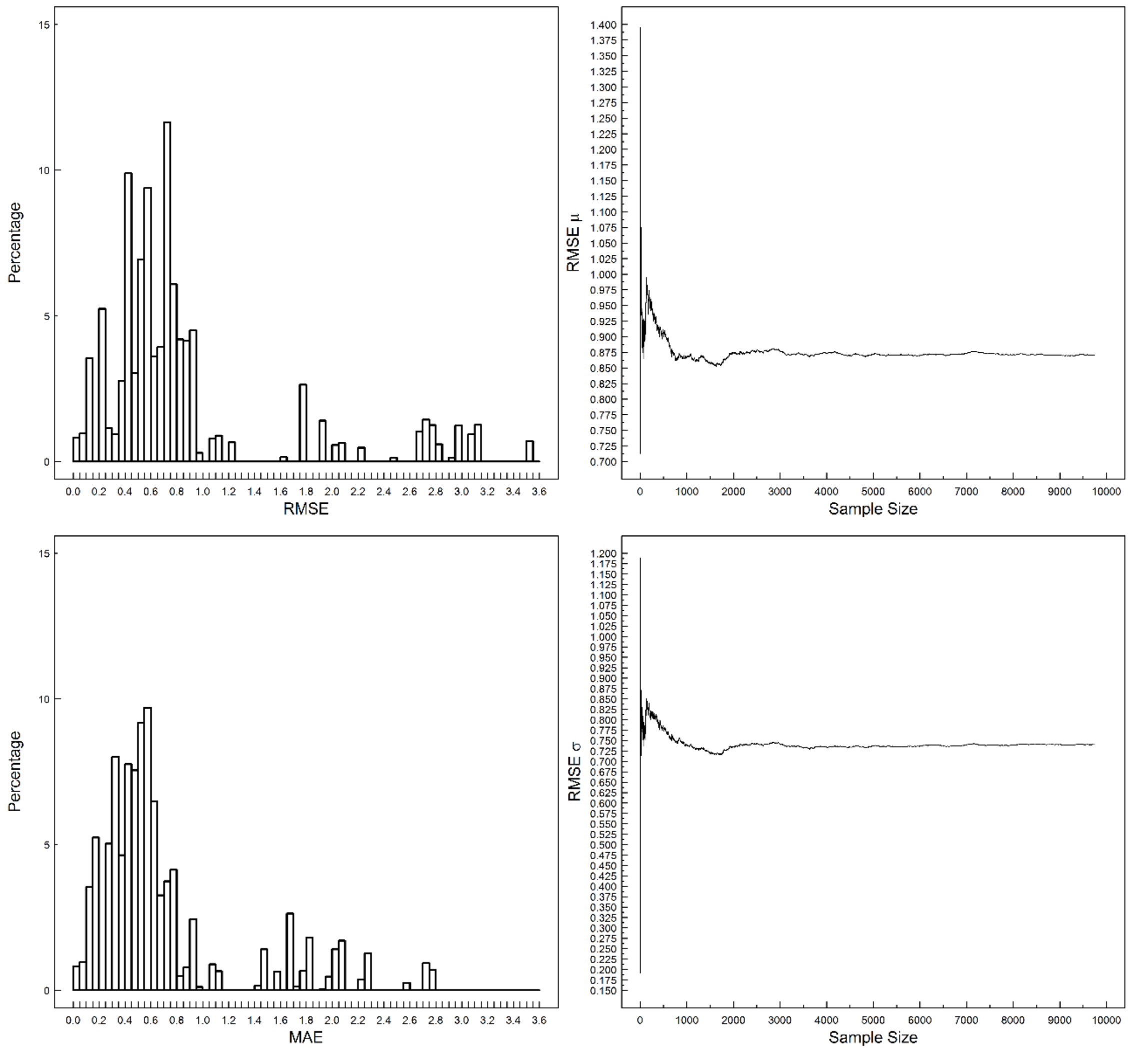
Well ID: 10N4902CBC



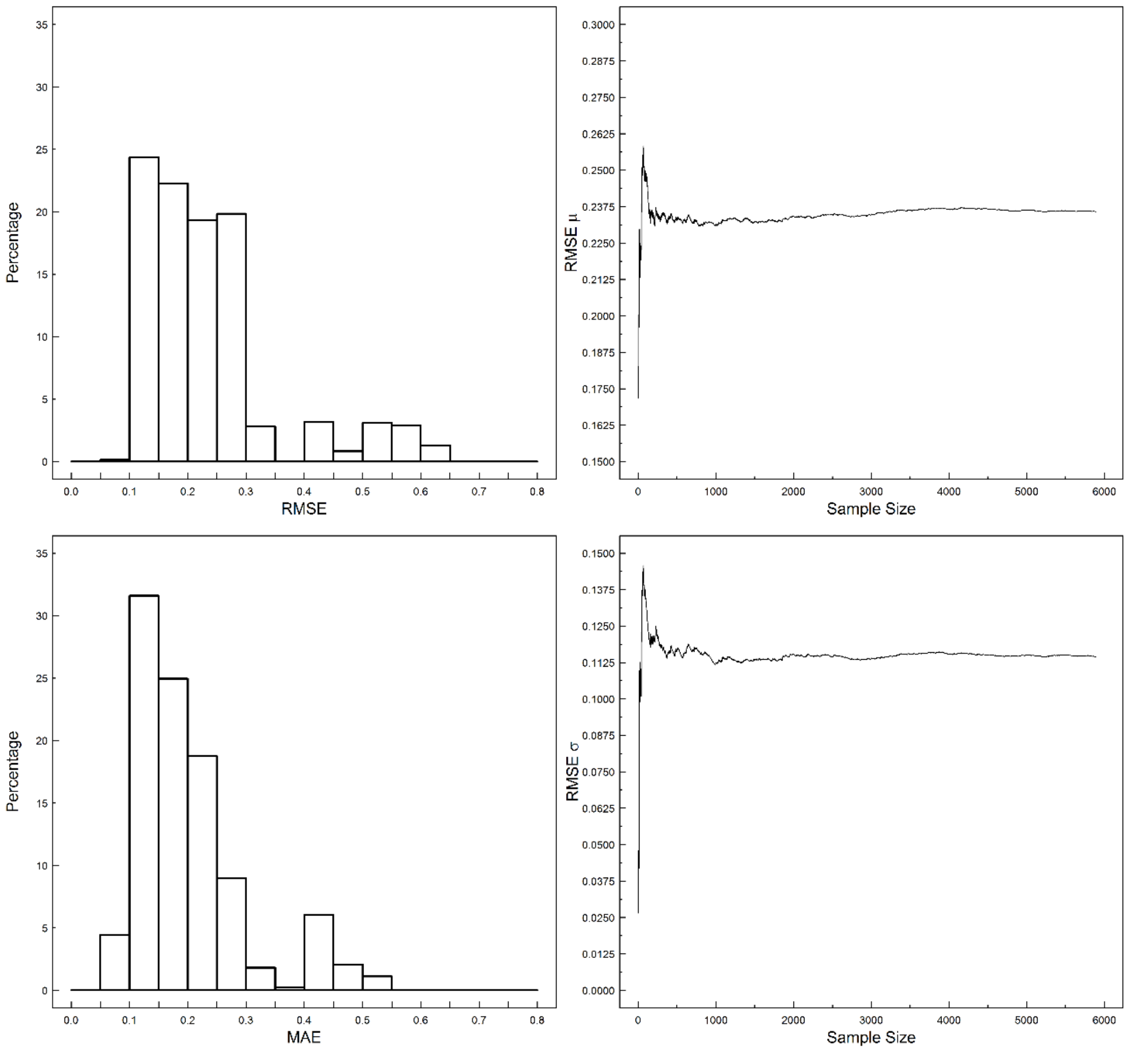
Well ID: 11N4728BBB



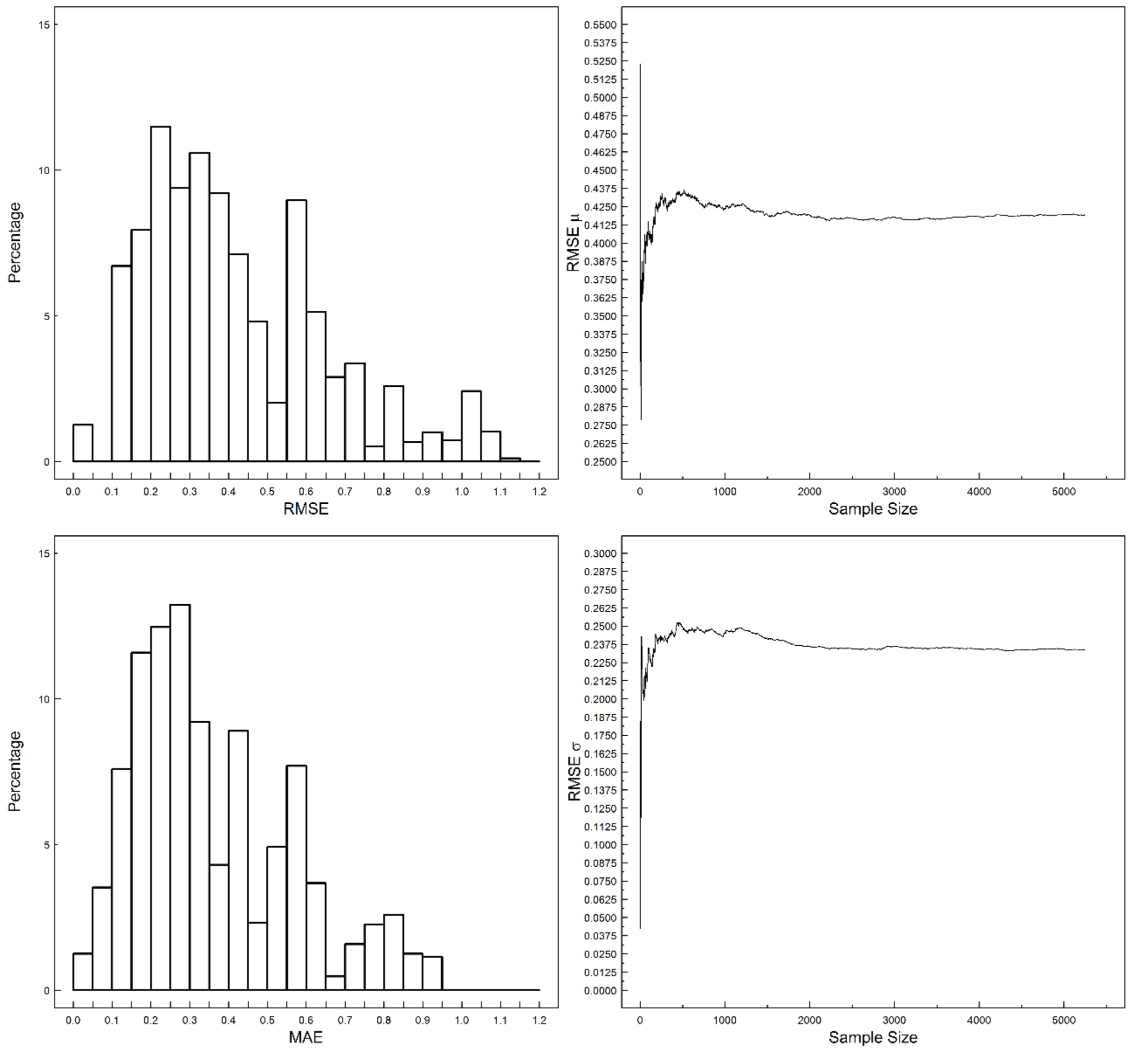
Well ID: 04N5931CB



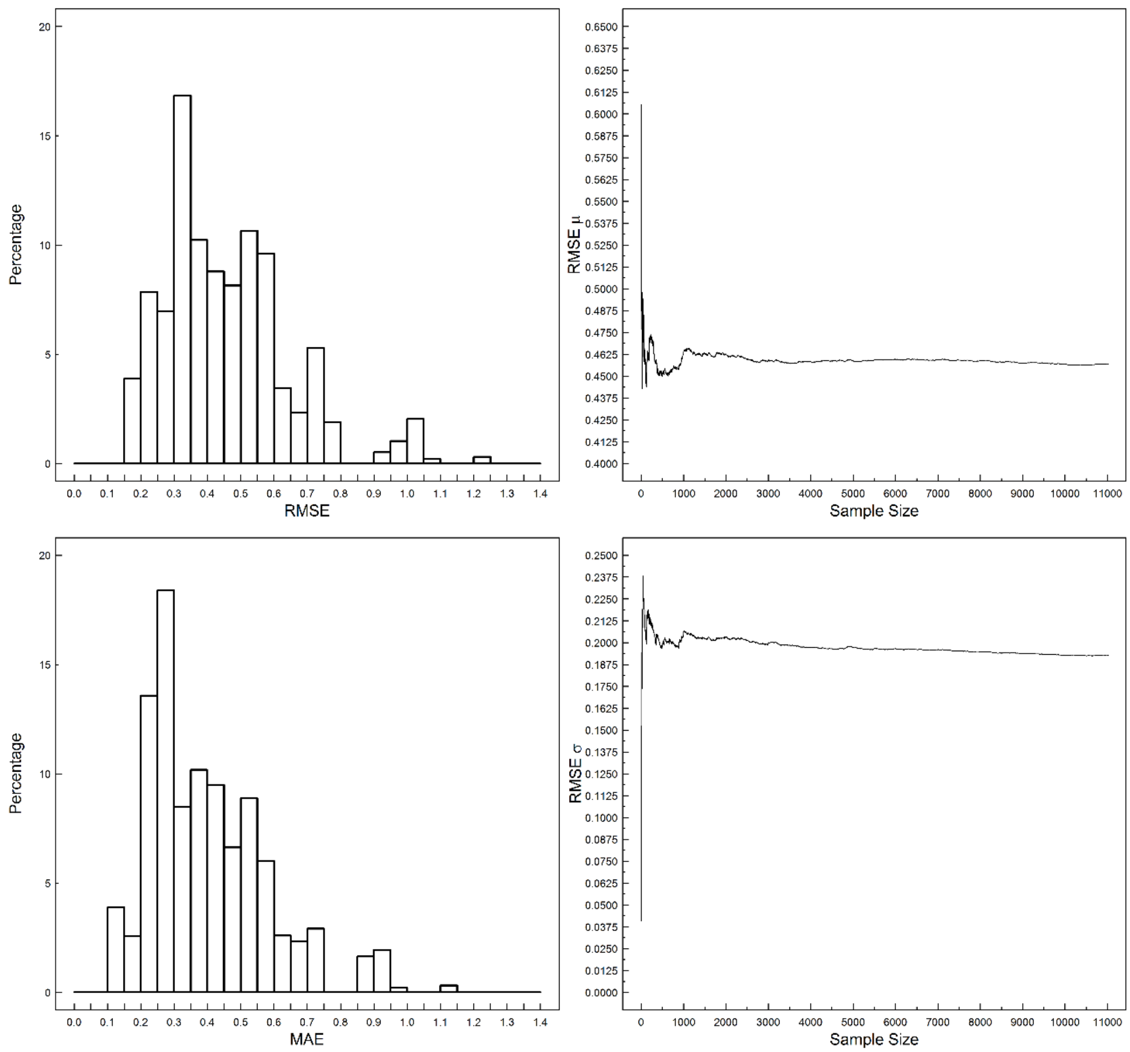
Well ID: 09N5131BBB



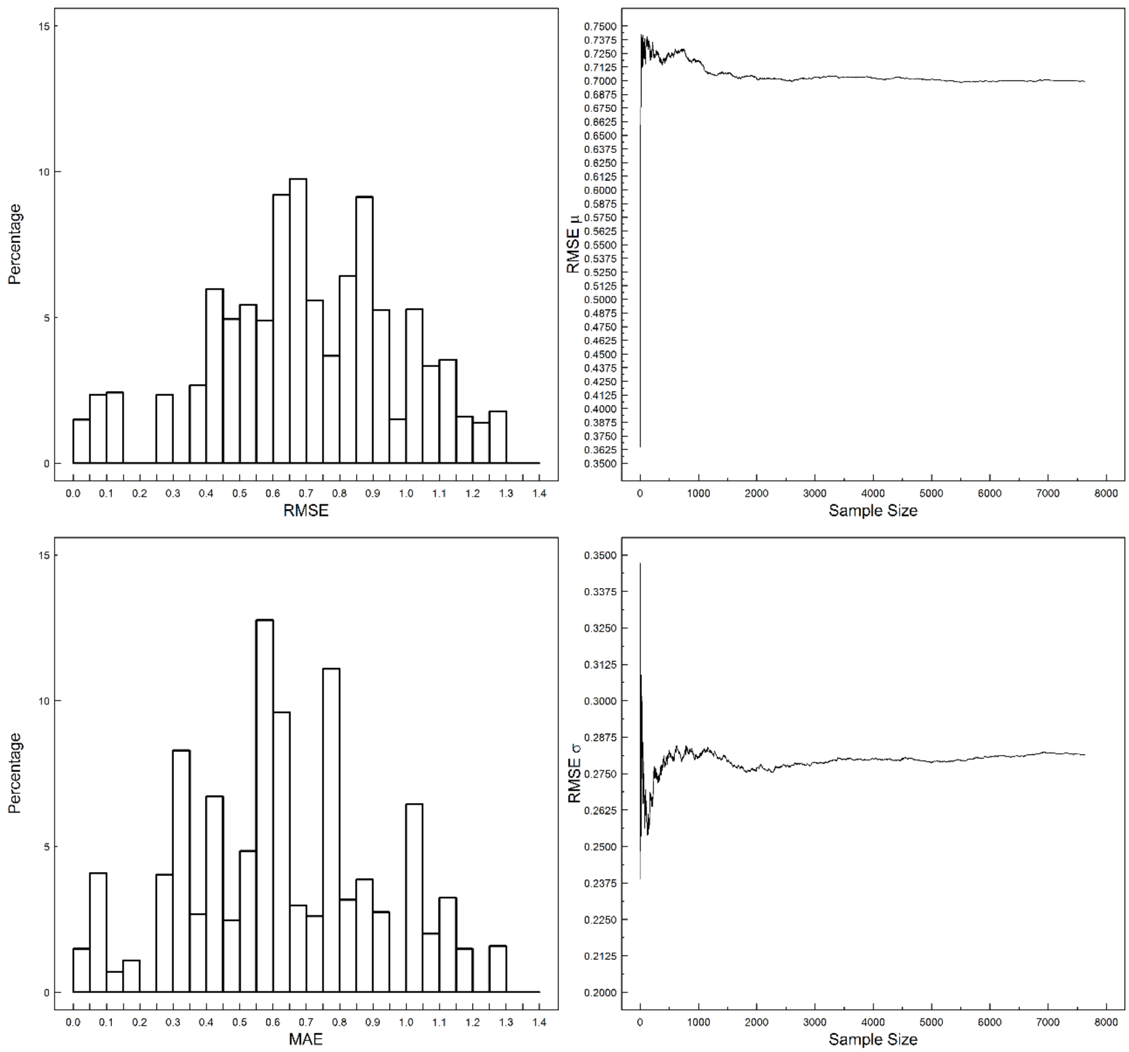
Well ID: 04N6506DAB



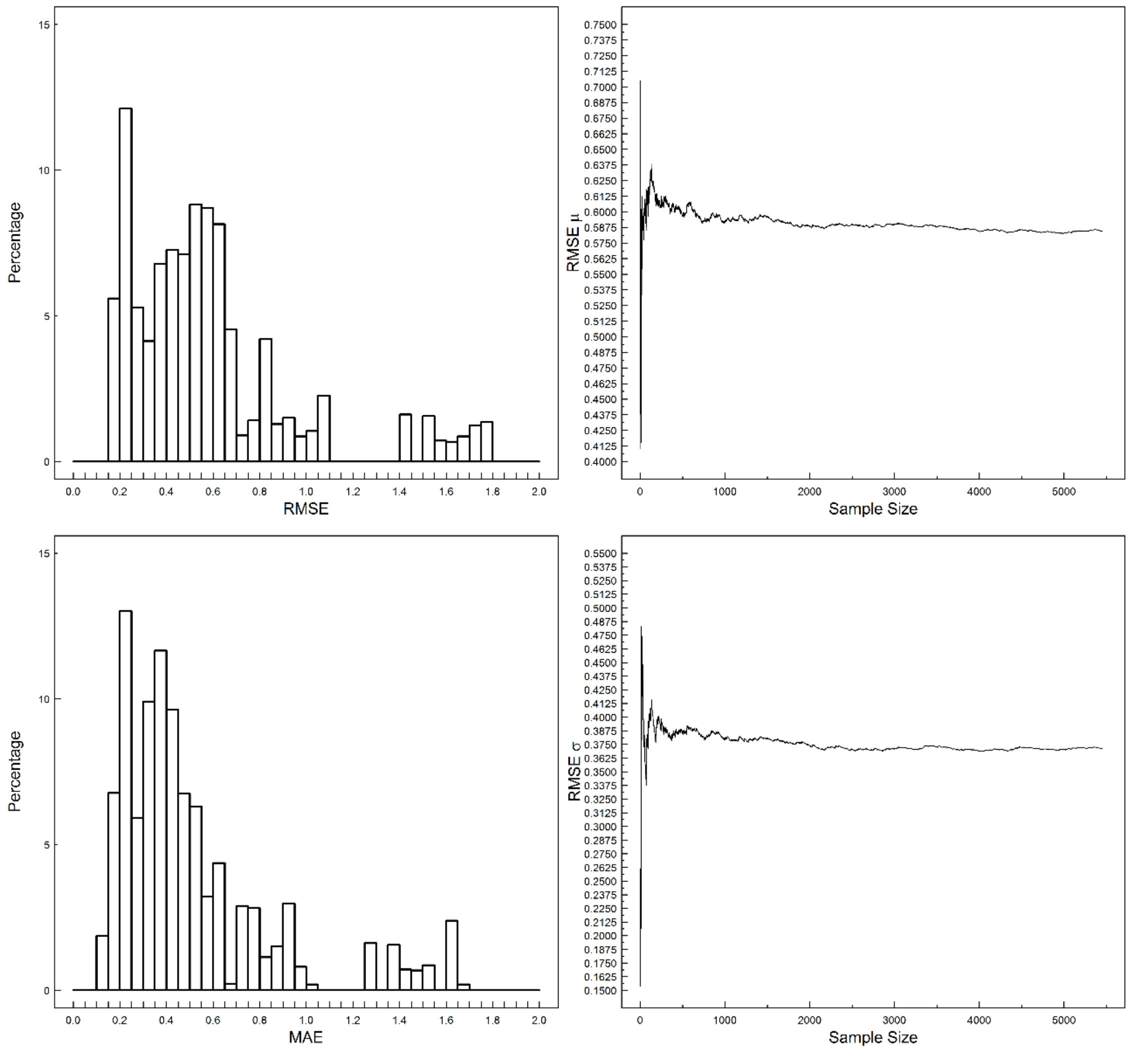
Well ID: 01N5531DCD



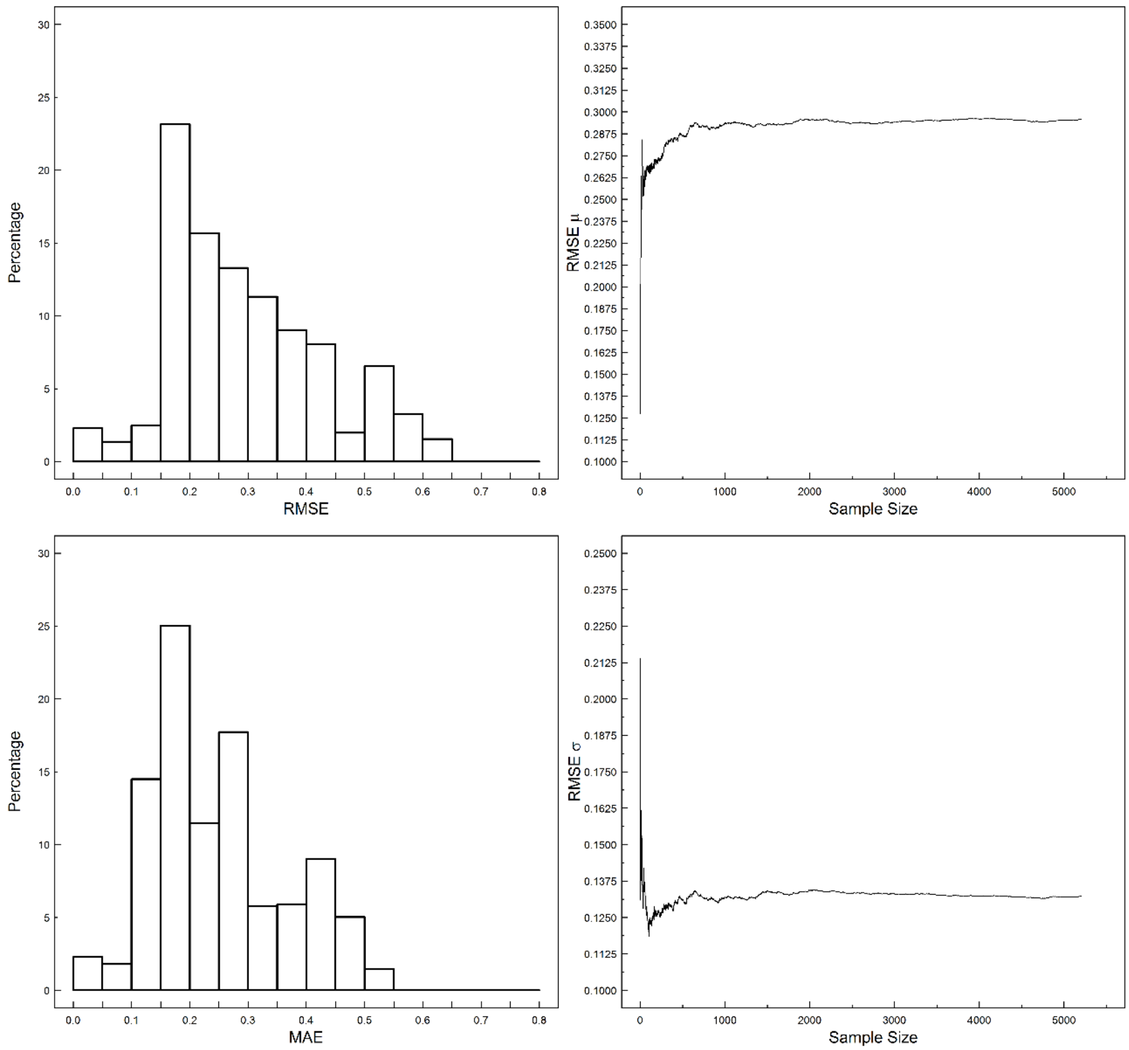
Well ID: 04N6422DCD



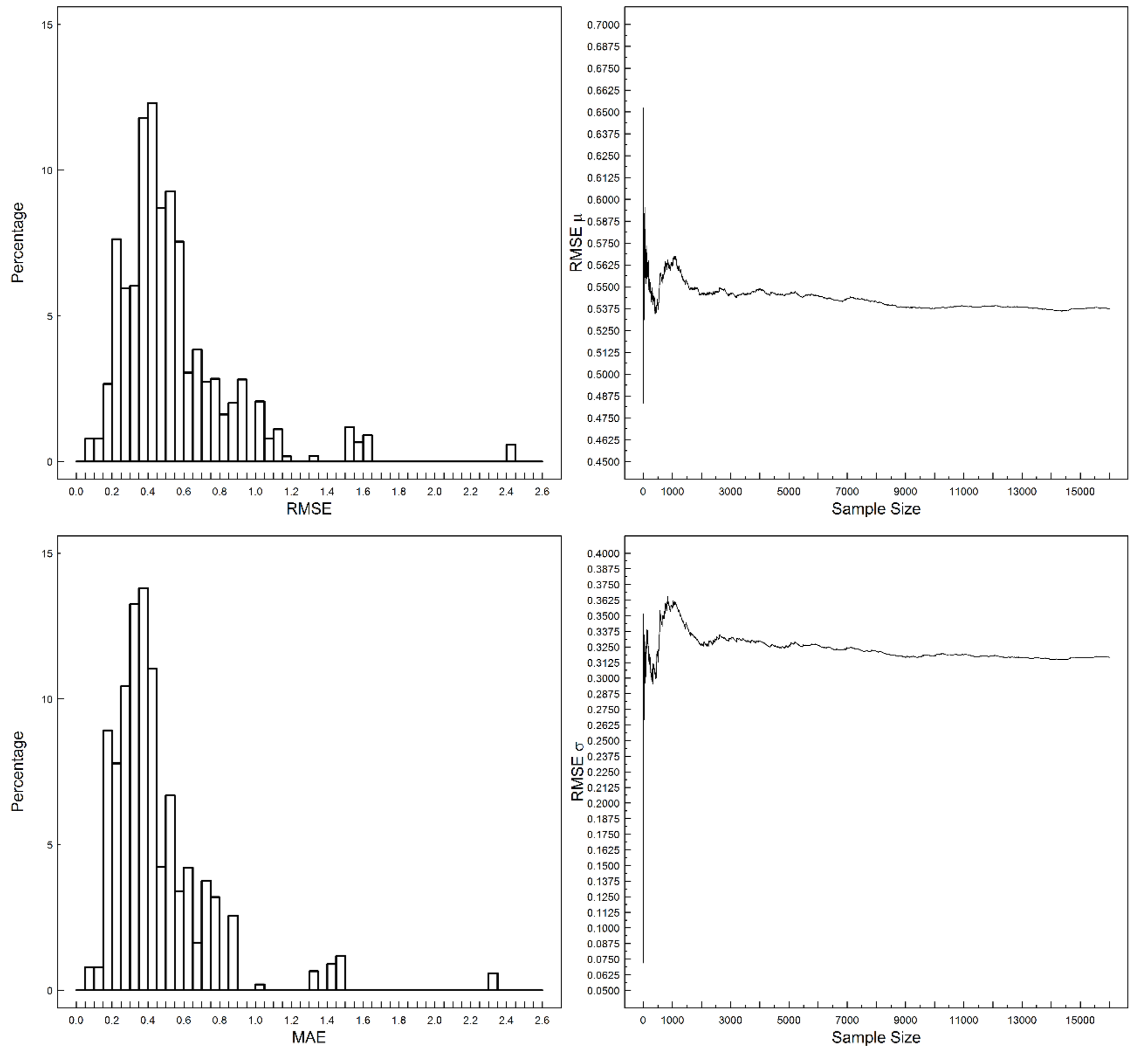
Well ID: 04N6412CCC



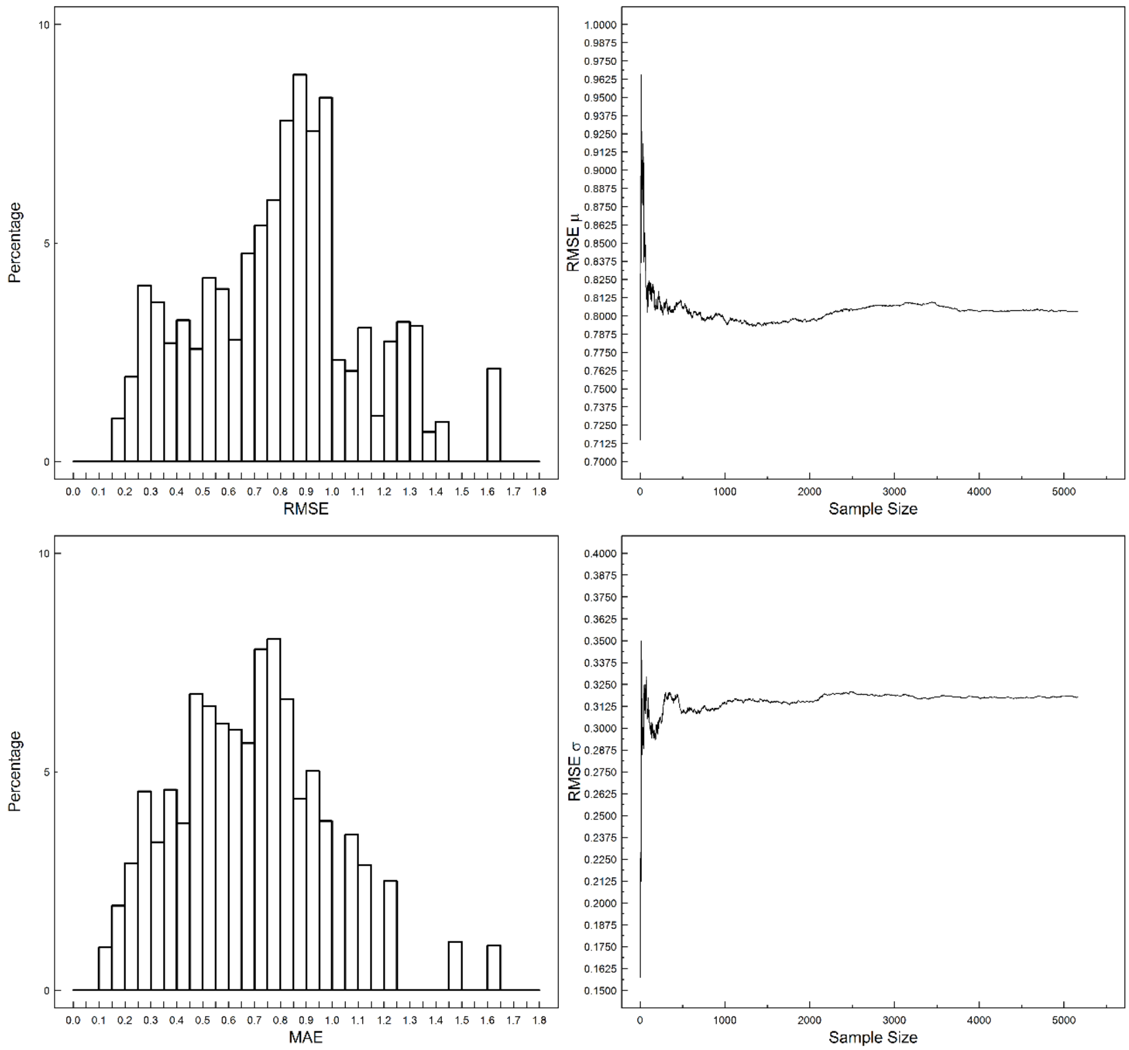
Well ID: SB00306618CAC



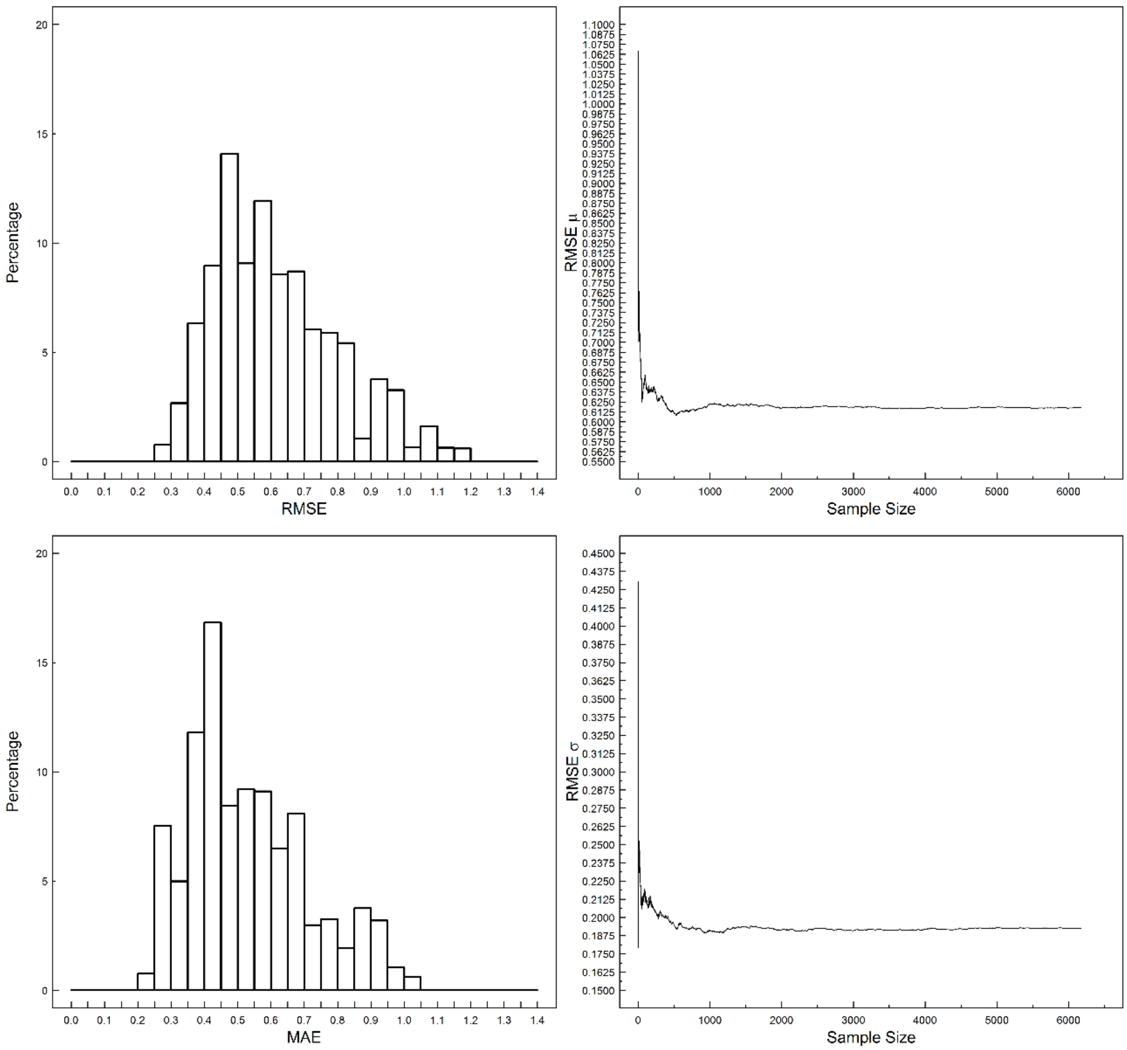
Well ID: 04N6627ADD



Well ID: SB00406012CCC-CSU1



Well ID: 68-1


Percentage
RMSE
RMSE μ
Sample Size
Percentage
MAE
RMSE σ
Sample Size