SPATIO-TEMPORAL ASSESSMENT OF GROUNDWATER RESOURCES IN THE DENVER BASIN AQUIFER SYSTEM

by

Christopher J. Ruybal

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Environmental Science and Engineering).

Golden, Colorado

Date ___________________________

Signed: ________________________________
Christopher J. Ruybal

Signed: ________________________________
Dr. John E. McCray
Thesis Advisor

Signed: ________________________________
Dr. Terri S. Hogue
Thesis Advisor

Golden, Colorado

Date ____________________________

Signed: _________________________
Dr. Terri S. Hogue
Professor and Department Head
Department of Civil and Environmental Engineering

ABSTRACT

Groundwater is an important resource in the Unites States and provides about 40% of the country's public water supply. Withdrawals have dramatically increased in many aquifers, leading to groundwater depletion and questions about future sustainability. In Colorado, the long-term sustainability of the Denver Basin Aquifer System is considered by some as questionable and insufficient to support future demands. Groundwater depletion has been widely documented over the past several decades as groundwater withdrawals have increased and competition for water further stresses supplies. Groundwater monitoring is fundamental to understanding system dynamics, trends in storage, and the long-term sustainability of an aquifer. However, groundwater level data are typically spatially and temporally sparse relative to the data density desired for aquifer-scale analysis. The problems with missing temporal data from water wells in particular has not been addressed in much detail, yet can cause important misinterpretation with regard to groundwater sustainability. This research aims to mitigate some of the problems with current approaches to analyzing water well data by incorporating a new method of spatial-temporal analysis, with particular emphasis on addressing missing temporal data. In addition, we evaluate the ability of the Gravity Recovery and Climate Experiment (GRACE) satellites to improve the temporal sustainability analysis. The methodology is first illustrated using a case study in the Arapahoe Aquifer and is then expanded to all aquifers of the Denver Basin Aquifer System. Remote sensing is utilized from GRACE to provide another perspective on determining groundwater storage changes. Results from this dissertation provide a framework for monitoring and management of groundwater resources along the Colorado Front Range as well as other water-stressed regions of the western U.S.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

AF	Acre-Feet
ASR	Aquifer Storage and Recovery
BM	Biomass
CDWR	Colorado Division of Water Resources
CFRUC	Colorado Front Range Urban Corridor
CWSD	Centennial Water and Sanitation District
DBAS	Denver Basin Aquifer System
ft amsl	Feet Above Mean Sea level
ft	Feet
GRACE	Gravity Recovery and Climate Experiments
GW	Groundwater
km^2	Square kilometers
km^3	Cubic kilometers
Lat	Latitude
Lon	Longitude
M&I	Municipal and Industrial
mi^2	Square miles
PCA	Principal Component Analysis (PCA1, PCA2)
SB	Senate Bill
SPB	South Platte River Basin
SM	Soil moisture

SW	Surface waters
SWE	Snow water equivalent
TWS	Total Water Storage
USGS	U.S. Geological Survey
WISE	Water Infrastructure and Supply Efficiency
(s_i, t_i)	Space-time points
S_s	Specific storage,
S_y	Specific yield
f_i	Known covariates
β_i	Regression coefficients
γ_J	Joint semivariogram
γ_S	Spatial semivariogram
γ_T	Temporal semivariogram
$\hat{\varepsilon}(s_0, t_0)$	Spatio-temporal kriging predictions on the stochastic residual
λ_i	Spatio-temporal kriging weights
σ	Standard deviation
σ^2	Variance of the spatial, temporal, and joint quantities
σ_{nug}^2	Nugget variance
ΔGW	Volumetric change in groundwater storage
ΔH	Groundwater level change
h	Spatial separation distances
A	Area overlying the drained aquifer or area weighting factor
$N(h, u)$	Number of paired observations of z

$Z(s,t)$	Random function model and target variable of groundwater levels
a	Range of the spatial, temporal, and joint quantities
b	Aquifer thickness
$m(s,t)$	Deterministic trend
n	Number of observations limited to the search neighborhood
p	Number of covariates
S	Spatial domain
$s = (x,y)$	Spatial coordinates
T	Time interval
t	Time
u	Temporal separation distances
x	Water level measurement
$\bar{x}$	Groundwater level mean
z	Normalized groundwater level
$z(s_0,t_0)$	Predictions at unmeasured space-time points and variable of interest
$\gamma(h,u)$	Spatio-temporal semivariance
$\varepsilon(s,t)$	Stochastic residual or spatio-temporal residual
κ	Space-time anisotropy ratio that merges spatial distances with temporal distances
μ	Lagrange multiplier

ACKNOWLEDGMENTS

I have had the opportunity to be advised by, Dr. John McCray and Dr. Terri Hogue. I am thankful to them for their mentorship, knowledge, support, and friendship during my time at the Colorado School of Mines. I also want to acknowledge the guidance I received from my Ph.D. committee members, Dr. David Benson, Dr. Kamini Singha, Dr. Wendy Zhou, Dr. Bridget Scanlon, Dr. Dominic DiGiulio, as well as past committee members, Dr. Tissa Illangasekare and Dr. Junko Munakata Marr. Everyone's remarkable expertise and invaluable feedback helped shape my research, inspired me, and helped me through the challenges of Ph.D. research.

This research has been supported by the ConocoPhillips Center for a Sustainable WE^{2}ST, the Department of Civil and Environmental Engineering at the Colorado School of Mines, the Trefny Institute, Colorado Water Institute FY15 Student Research Award 2015CO312B, the Edna Baily Sussman Foundation, and the U.S. EPA's Ground Water and Ecosystems Restoration Division. I am especially grateful for all the funding received.

Thank you to the McCray Research Group, the Hogue Research Group, and the Center for a SustainableWE^{2}ST for enhancing my research through feedback, support, and friendship. A special thanks to Dr. Andrea Blaine for her support and encouragement during my early days in the Center for a Sustainable WE^{2}ST.

I am especially thankful for the incredible support of all my family and friends. Thank you to my parents, Thomas and Jo Ann Ruybal, who have provided unconditional love, support, and encouragement in all my endeavors. They have inspired me to pursue my dreams and have been with me every step of the way. Thank you to my siblings, Jeffrey Ruybal, Kathleen Ruybal, and David Ruybal for putting up with my constant ‘science’ talk and believing in me. Thank you to Ashley Black for her support, encouragement, and love.

I am extremely grateful for such wonderful colleagues, friends, and family, all who have contributed to this dissertation. I look forward to much more friendship and collaboration.

CHAPTER 1

INTRODUCTION

Groundwater is an important resource in the Unites States (U.S.) and provides about 40% of the country's public water supply (Alley et al. 1999; Maupin et al. 2014). It is also the primary source of drinking water for many rural areas and an important source for irrigated agriculture (Scanlon et al. 2012a; Maupin et al. 2014). Withdrawals have dramatically increased since the 20th century leading to groundwater depletion in many aquifer systems and questions about future sustainability (Brown 1999). Since 1990, groundwater volumes have decreased by about 1,000 km^3 across the U.S., highlighting the severity of the problem (Konikow 2015). Some detrimental effects of depletion include reduced well yields, increased pumping cots, the need to drill deeper wells, land subsidence, reduced water quality, and reduced flow to surface waters (Alley et al. 1999; Konikow and Kendy 2005).

The long-term groundwater sustainability of an aquifer system is a complex issue that involves scientific, political, legal, and socioeconomic influences (Alley et al. 1999; Gleeson et al. 2012). Water imbalances in groundwater systems between inflows and outflows can create large deficits, especially when withdrawals are large (Scanlon et al. 2012a). Water level and use data are key for scientific assessments and determination of sustainable withdrawals (Butler et al. 2016). Political influences also help place value on water and promote conservation. For example, in Colorado, the state's Water Plan has been implemented for two years and sets forth a framework at the state, local, and individual level to address future water challenges (Colorado Water Conservation Board 2015). However, the details on how to appropriately quantify sustainability or current groundwater use are not provided. Western states also vary in their legal framework to manage and regulate groundwater withdrawals, but are often focused on short-term management plans that may not sustainable in the long-term (Gleeson et al. 2012; Schlager 2006). A recent study identified socioeconomic variables such as a rural-urban divide, per capita income, education level, pricing structures, and regional location all influence water-use efficiency (Sankarasubramanian et al. 2017). Ensuring adequate water resources to support growing populations and economic development is an ongoing challenged due to the complex overlap of influences.

In Colorado, the long-term sustainability of the Denver Basin Aquifer System (DBAS) is considered by some as questionable and insufficient to support future demands (Everett 2014; CDM 2004). The DBAS is a critical water resource for many uses, including municipal, industrial, agricultural, and domestic use along the Colorado Front Range Urban Corridor. Groundwater depletion has been widely documented over the past several decades as groundwater withdrawals have increased (Topper and Raynolds 2007; Paschke et al. 2011; Bauch et al. 2015). Over the next thirty years, the supply gap is expected to approach 690,000 acre-feet per year (0.85 km^3) for municipal, industrial, and agricultural use (HDR Inc. and Consultants 2015). Population growth is recognized as one of the strongest drivers for increasing future water resource needs in the area (HDR Inc. and Consultants 2015; Colorado Water Conservation Board 2011; Moore et al. 2007). Since 1990, population in the Denver metro area has increased by over 1.8 million people. While much of the water demand will be offset by surface water resources, an increased reliance on groundwater from the DBAS is expected, particularly as other western states with rights to Colorado River water (accounts for 30 – 50% of supply for Front Range in a given year) are likely to demand their full share. Because of this, the DBAS faces future management and resource challenges.

Most aquifers in the arid west, as well as in many other locations across the globe, are facing serious questions about future sustainability (e.g. (Famiglietti 2014; Konikow 2015; Scanlon et al. 2012a). Typically, sustainability is evaluated by analyzing groundwater well records to determine if water levels can be sustained into the future under certain use scenarios. Of course, water levels only tell part of the story, because groundwater availability actually depends on future projects of groundwater storage, which is highly uncertain. In addition, analysis of water well data as it is typically conducted has serious limitations. Data are typically spatially and temporally sparse relative to the data density desired for aquifer-scale analysis. The problems with missing temporal data from water wells in particular has not been addressed in much detail, yet can cause important misinterpretation with regard to groundwater sustainability.

This research aims to mitigate some of the problems with current approaches to analyzing water well data by incorporating a new method of spatial-temporal analysis, with particular emphasis on addressing missing temporal data. In addition to this proposed method, we evaluate the ability of GRACE satellite data to improve the temporal sustainability analysis. The strengths of GRACE data may help minimize the impact of the weaknesses on the water-well data analysis, and vice versa. The research is applied to the DBAS, an important aquifer system in the Colorado Front Range that is expected to be highly stressed with regard to future sustainability.

The DBAS extends from Colorado Springs to Greeley and covers approximately 6,700 mi^2 (~18,000 km^2) (Moore et al. 2004; Paschke et al. 2011). It consists of four layered aquifers in a structural, sedimentary basin containing sandstones and siltstones deposited 65 -70 million years ago (Topper and Raynolds 2007; Robson 1987; Paschke et al. 2011). The DBAS underlies the South Platte Alluvial Aquifer which is another important groundwater system in the area. Groundwater development began in the late 1880s with many wells under artesian pressure, pushing water as high as 100 ft above the ground surface (Topper and Raynolds 2007). However, due to the eventual widespread declines in artesian pressures and as drilling became widespread, legislation of Senate Bills (SB) 73-213 and 85-05 passed prompting how the DBAS aquifers should be managed and groundwater withdrawn (Moore et al. 2004; White 1995). Since 1990, over 23,000 new groundwater wells have been installed, leading to continual water-level declines and in some cases areas are changing from confined to unconfined conditions (Colorado Division of Water Resources 2018; Moore et al. 2004).

There is an ongoing need to continue to monitor and improve understanding of water-level and storage changes in the DBAS due to its high importance. Numerous groundwater models have been developed over the years using various approaches and scales. In 1998, under SB 96-74, a model of the Denver Basin was refined from previous Denver Basin models of Robson (1987), Banta (1989), and the MODFLOW model developed from SB 85-05 (Simpson and Lite 1998). In 1991, the Denver Basin model from Banta (1989) was combined with RIVINT, a river module for MODFLOW, to predict declines in aquifer heads and stream baseflow during times of drought when pumping was occurring. In 2003, MODFLOW-2000 in conjunction with the State's SB-74 model were used to develop a regional Denver Basin groundwater model to estimate water levels in various aquifers under various pumping scenarios (Black & Veatch et al. 2003). In 2011, an extensive model of the DBAS was completed by the U.S. Geological Survey (USGS) to better understand the hydrogeological framework of the system (Paschke et al. 2011). Later this model was then used to understand various pumping schemes to protect the Arapahoe aquifer from depletion of saturated thickness (Banta and Paschke 2012). In 2011, USGS and Rural Water Authority of Douglas County began a study to assess groundwater levels in the DBAS of Douglas County (Everett 2014). Comparison of recent

USGS modeling studies and Colorado Division of Water Resources (CDWR) data do show variable results and different rates of annual depletions across the DBAS aquifers (Paschke et al. 2011; Banta and Paschke 2012; Everett 2014; Flor 2015, 2016).

The DBAS would benefit from additional analysis and interpretation of groundwater level data to better understand stresses in the system. Groundwater monitoring is fundamental to understanding an aquifer, determining direction of flow, and trends in groundwater storage. Knowing were areas are decreasing, increasing, or stable allows managers to appropriately focus their efforts and better manage resources. However, volumetric depletion assessments can be challenging and difficult to assess (Konikow 2015). Many areas lack adequate water-level data, adequate data characterizing aquifer properties, and pumping data. It is necessary to frequently monitor groundwater elevations to capture seasonal fluctuations and long-term trends. Obviously, more data leads to a better understanding of an aquifer system. The DBAS is no exception. All four of the aquifers lack adequate spatial and temporal groundwater level data. Most groundwater wells are irregularly sampled and not every well is sampled every year. This complicates groundwater assessments and determination of trends and stresses.

Groundwater level or storage change evaluations typically employ some sort of interpolation scheme to estimate data at unsampled locations. Point measurements of water table levels are used to generate groundwater surfaces (i.e. potentiometric surfaces) and understand spatial and temporal variations (Sun et al. 2009; Gundogdu and Guney 2007). Numerous spatial interpolation methods exist and have been applied across many disciplines. In general, estimation of data at unsampled locations are represented as a weighted average of surrounding data (Li and Heap 2008). Some common non-geostatistical techniques applied to groundwater mapping include inverse distance weighting (Varouchakis and Hristopulos 2013; Yao et al. 2014), nearest neighbor (Rouhani 1986), splines and trend surfaces (Yao et al. 2014; Xiao et al. 2016), and regression models (Delbari 2014; Ohmer et al. 2017). Kriging is a geostatistical interpolation method and comes in many forms (e.g. Ordinary, Universal, Block, CoKriging.) (Brus and Heuvelink 2007; Hengl et al. 2003; Li and Heap 2008). Kriging is considered the best linear unbiased estimate for the variable in question. The variable is assumed to be a random function with a correlation structure in its spatial distribution that can be explained by a semivariogram. The advantage of kriging compared to other deterministic methods, is that the interpolation error can be estimated.

Groundwater surfaces are generally non-stationary due to a trend in the surface (Maheswaran and Khosa 2013; Gundogdu and Guney 2007). When applying kriging, the equations can be expanded to include a trend or drift in the surface (i.e. universal kriging, kriging with external drift, regression kriging) and account for the drift to achieve stationarity necessary for kriging. Covariates can be used to account for the trend in the data as well as help infer the predictions in un-sampled locations. However, an insufficient number of sample points creating spatial and temporal gasp in the groundwater level record can be problematic. This can potentially lead to anomalies and incorrect representation of the groundwater level or head data. In order to avoid some of these pitfalls, many studies aggregate the data in to bins where the same timestamp is given, even though the data might be from different periods. The likely increases the number of spatial data points to infer the interpolation. However, this assumes that the temporal response of the system response can be represented a sequence of spatial interpolations invoked at user-defined times (Li and Revesz 2004). This presents problems for irregular datasets.

Spatio-temporal approaches offer an alternative to classical geostatic and can overcome some of the limitations of sparse and irregular data. Time is incorporated as another dimension and covariance is modeled as a function of both space and time (Kyriakidis and Journel 1999; Heuvelink and Griffith 2010; Tadic et al. 2017). Spatio-temporal kriging is common to a number of studies that have spatially and temporally mapped precipitation, soil moisture, and satellite observations (Hu et al. 2017; Wang et al. 2015; Zeng et al. 2017). However, despite these applications to environmental research, spatio-temporal kriging applications are lacking in groundwater applications. Spatio-temporal kriging has only been applied to three known studies, including an early groundwater quantity study (Rouhani and Hall 1989), design of groundwater monitoring networks (Júnez-Ferreira and Herrera 2013), and recently groundwater level variations (Varouchakis and Hristopulos 2017). The concept of spatio-temporal kriging in groundwater applications is not a normal concept used in groundwater studies. The recent publication by Júnez-Ferreira and Herrera (2013) and Varouchakis and Hristopulos (2017) may highlight new interest in spatio-temporal methods for groundwater applications.

While these studies introduce spatio-temporal techniques, they are of limited application. The objective of Júnez-Ferreira and Herrera (2013) was to optimize a groundwater head monitoring network in order to minimize the number of the number of samples needed to maintain the current level of uncertainty. Their focus was on a small part of a 379 km^2 unconfined aquifer. Their methodology utilized a Product-Sum model (De Iaco et al. 2001) to model space-time covariance and Kalman filter to estimate the residuals. Varouchakis and Hristopulos (2017) evaluated groundwater levels for wet and dry periods across two separate two year spans. Separable Matérn, diffusive, and Spartan models were used to model space-time trends in variograms for an unconfined aquifer of ~26 km^2 using eleven wells. While these studies have recently introduced space-time geostatistics to the groundwater realm, much validation and assessment is still needed. Further studies are needed to 1) assess the applicability of different space-time variogram models, 2) evaluate larger aquifers, 2) further assess usefulness for confined, semi-confined, and unconfined systems, 3) evaluate longer time periods, 4) evaluate different methods to remove global groundwater trends or drift, and 5) advocate the advantages of spatio-temporal applications for irregular or sparse groundwater datasets. Remote sensing offers unique capabilities in monitoring and understanding spatial and temporal earth system processes.

Lately, the Gravity Recovery and Climate Experiments (GRACE) have been effective in advocating and educating the public on groundwater sustainability. Estimates of gravitational changes can be used to infer time-variable changes in mass above and below the Earth's surface. Most monthly mass changes are due to the redistribution of water within the global water cycle across the oceans, atmosphere, land, and cryosphere. GRACE has been used for many hydrology applications across much of the globe (Seyoum and Milewski 2016; Castle et al. 2014; Sun 2013; Scanlon et al. 2012b; Famiglietti et al. 2011; Strassberg et al. 2009; Rodell et al. 2009; Chen et al. 2014; Tiwari et al. 2009). These studies have looked at droughts (Houborg et al. 2012; Long et al. 2013; Leblanc et al. 2009; Thomas et al. 2014), flooding and surface waters (Swenson and Wahr 2009; Lee et al. 2011; Wang et al. 2011; Zmijewski and Becker 2014; Reager et al. 2014), and groundwater depletion (Rodell et al. 2009; Feng et al. 2013; Scanlon et al. 2012b; Voss et al. 2013)

Studies have shown good agreement between GRACE derived groundwater storage changes and groundwater well observations (Rodell et al. 2007; Chen et al. 2016; Swenson et al. 2006; Castle et al. 2014; Scanlon et al. 2012b; Huang et al. 2016). However, because the

GRACE footprint is ~200,000 km^2, many studies evaluate the use and validity over areas close to or much larger than the footprint. Few studies have evaluated use below the recommended area. In the Central Valley, California, high amounts of pumping for irrigation have allowed storage changes to be detected by GRACE even though the area of interest is ~52,000 km^2 (Scanlon et al. 2012b). In regions such as Mali (~54,971 km^2) with sparse hydrogeological data, GRACE TWS was used to estimate annual recharge (Henry et al. 2011). (Need to add a few more studies).

As with other types of data, GRACE is not without limitations. The gridded GRACE spherical harmonic product is at a course 1° x 1° (~111 km at the equator) resolution while the mascon approach is at a gridded 0.5° x 0.5 ° (~56 km at the equator) resolution. This lower resolution limits the use for smaller scales (Alley and Konikow 2015). Individual grid-cells may also suffer from spatial leakage and measurement errors that are large at small scales (Longuevergne et al. 2010). These errors are assumed to be constant over time, but may actually vary seasonally and annually. GRACE reports changes in water storage in one-dimension. It cannot determine flow velocities, directions of flow, drawdowns, sources of groundwater depletion, differentiate elastic from non-elastic storage and releases, and be used for real-time monitoring due to the few month lag in results being released (Alley and Konikow 2015). Alley and Konikow (2015) suggest that there are more important things to consider like other external effects (e.g. water quality, pumping and lifting costs, land subsidence) that will impact aquifer use long before storage depletions. However, in many areas, there are a lack of data available that limits the understanding of spatial and temporal dynamics (Seyoum and Milewski 2016). This is what makes GRACE appealing, despite the resolution and cautions waring against use at smaller scales. Groundwater data however, is not without limitations and weaknesses. Generally, groundwater data has poor temporal resolution, poor spatial resolution, and does not provide a direct identification of volumetric storage changes without the use of aquifer parameters which are often not well known. GRACE overcomes these limitations and can potentially still add value to systems at the aquifer-scale.

The overarching objective of this dissertation is to mitigate some of the problems with current approaches to analyzing groundwater well data by incorporating a new method of spatial-temporal analysis, with particular emphasis on addressing missing temporal data. In addition to this proposed method, we evaluate the ability of GRACE satellite data to improve the temporal sustainability analysis. The research is applied to the DBAS, an important aquifer system in the Colorado Front Range. This is accomplished through three studies evaluating the advantages of spatio-temporal kriging (Chapter 2), application to the DBAS (Chapter 3), and the applicability of utilizing GRACE to overcome some of the ground-based limitations (Chapter 4).

1.1 Objectives and Research Questions

This section summarizes the objectives for the three main chapters of this dissertation and the efforts taken to address the relevant research questions.

Objective 1: Evaluation of Spatio-Temporal Kriging to Predict Potentiometric Surfaces

The first objective includes evaluation of spatio-temporal kriging as an alternative interpolation method for groundwater data, with a goal of demonstrating the ability to improve potentiometric surface predictions. The following science question was addressed:

Science Question 1: Do estimated potentiometric surfaces using spatio-temporal kriging improve our understanding of groundwater sustainability than traditional methods of groundwater well data analysis (e.g., inverse distance weighting, point to raster) or methods that use only spatial kriging?

Objective 2: Evaluation of Groundwater Storage in the Denver Basin Aquifer System

The second research objective focuses on applying the developed spatio-temporal kriging approach to the DBAS, with the goal of understating storage changes in each aquifer, any spatial patterns, and the influence of the storativity parameter on volumetric estimates. The following science questions were addressed:

Science Question 2: How has groundwater storage changed for aquifers of the DBAS since 1990?

Science Question 3: Utilizing available storativity data, what is the variability in storage change estimates?

Science Question 4: Can we identify the main drivers for groundwater storage changes and how these impact aquifer sustainability?

Objective 3: Evaluation of GRACE at the Aquifer-Scale

The third research objective evaluates the applicability of using the GRACE satellites at the aquifer-scale to help infer groundwater storage changes, with the goal of determining the usefulness and challenges across the DBAS. The following science question was addressed:

Science Question 5: Does analysis of groundwater storage trends from GRACE improve our understanding of groundwater availability/storage at the aquifer-scale compared to those determined from ground-based data for the DBAS?

1.2 Dissertation Organization

This dissertation is organized into five chapters. Chapter 1 provides an overview of the importance of groundwater, the challenges faced in resource evaluations, and outlines the research goals and questions for the material in the main body of the dissertation. Chapter 5 summarizes the conclusions from the dissertation and provides recommendations for future direction of research. The main body of the dissertation (Chapters 2 – 4) describes three research efforts to address the main research objectives and questions. Relevant supporting information for Chapter 3 is provided in Appendix A. The following is a description of the main chapters in the dissertation:

- Chapter 2, “Evaluation of Groundwater Levels in the Arapahoe Aquifer Using Spatio-Temporal Kriging.” The chapter addresses Objective 1. The goal of this chapter is to present and demonstrate a novel spatio-temporal kriging approach used to improve confidence in evaluating groundwater levels across irregular and sparsely gauged aquifer systems. The spatio-temporal approach is defined along with the steps taken to predict groundwater levels at locations in space and time. The Arapahoe Aquifer in the DBAS is used as a case study to further demonstrate the benefits of spatio-temporal approaches in groundwater resource evaluations.
- Chapter 3, “Assessment of Groundwater Depletion and Implications for Management in the Denver Basin Aquifer System.” This chapter addresses Objective 2. Supporting information for this chapter is provided in Appendix A. The goal of this chapter is to extend the spatio-temporal analysis to all four aquifers of the Denver Basin Aquifer System (DBAS) and further evaluate the variability in groundwater depletion and uncertainties in volumetric calculations. Water level changes and storage changes are evaluated from 1990 to 2016. This assessment compliments other groundwater studies in the DBAS and provides a different perspective on analyzing ground-based data.
- Chapter 4, “Monitoring Groundwater Storage Changes Using GRACE at the Aquifer Scale: Accuracy and Validity in the Denver Basin.” This chapter addresses Objective 3. The goal of this chapter is to evaluate the validity of using the GRACE satellites at the

aquifer-scale to better understand groundwater depletion in DBAS. The spatio-temporal analysis presented in Chapters 2 and 3 are used to compare to the GRACE results. GRACE is becoming a popular tool to assess groundwater storage changes. This study highlights the limitations of both GRACE and ground-based data and provides another perspective on evaluating groundwater resources in the DBAS.

Since the framework for the spatio-temporal assessment of the DBAS is setup and introduced in this dissertation, additional data collection or incorporation of other available data can be easily added as it becomes available. This approach enables improvement of predictions, allows for annual updates, and continually helps in water resource management decision-making. Additionally, the methodology described in this dissertation can easily be applied to other groundwater systems.

CHAPTER 2

EVALUATION OF GROUNDWATER LEVELS IN THE ARAPAHOE AQUIFER USING SPATIO-TEMPORAL KRIGING

2.1 Abstract

Groundwater monitoring is fundamental to understanding system dynamics, trends in storage, and the long-term sustainability of an aquifer. Water-level data are the key source of information used to understand the response. However, groundwater level data are often irregularly sampled, leading to temporal gaps in the record, and spatially are not adequately distributed across an aquifer. This presents challenges when spatially interpolating potentiometric surfaces and creating groundwater maps due data availability. We present a spatio-temporal kriging methodology to improve spatial and temporal confidence in groundwater level predictions at unsampled locations. The space-time dataset consists of a trend and residual component modeled with a linear regression and utilize a sum-metric model to represent spatio-temporal covariances. The Arapahoe Aquifer is used as a case study to demonstrate the benefits of spatio-temporal kriging over spatial kriging across a sparsely gauged and irregularly sampled aquifer. The Arapahoe Aquifer is a major source of water for many residents along the Rocky Mountain Front Range in Colorado. The results show superior performance of spatio-temporal kriging to predict groundwater levels over the traditional spatial kriging. Spatio-temporal kriging represents realistic temporal and spatial changes in water levels and avoids some of the problems inherent to spatial kriging. This study demonstrates the power of spatio-temporal kriging to help inform system dynamics in irregularly sampled aquifers.

2.2 Introduction

Groundwater monitoring is essential to groundwater management and provides fundamental information regarding the long-term sustainability and status of an aquifer (Taylor and Alley 2001; Reghunath et al. 2005). However, groundwater level measurements are typically irregularly collected and have many spatial and temporal gaps in the record which presents challenges in understanding spatial and temporal changes and stresses to the system.

A variety of deterministic and geostatistical interpolation methods are often used to spatially interpolate groundwater data to make estimates at unsampled locations. Kriging is a common geostatistical method that takes into account spatial dependence, and unlike deterministic methods (e.g. nearest neighbors, inverse distance weighting), provides a measure of the uncertainty in the predictions. Kriging has been applied to many spatial hydrological analyses, including evaluation of groundwater level fluctuations and groundwater surface map generation (Ahmadi and Sedghamiz 2007, 2008; Ta′any et al. 2009; Sun et al. 2009; Yao et al. 2014; Costelloe et al. 2015). However, most studies only consider spatial relationships when evaluating temporal changes. Data are usually aggregated into temporal bins where the same time stamp is given, even though observations may not be from the same time period (Tadić et al. 2015). This approach assumes that the spatio-temporal response is accurately represented by a sequence of spatial interpolations invoked at several user-defined times. This approach is appropriate if the same locations are sampled in each temporal period (Li and Revesz 2004). However, most datasets are temporally irregular and do not contain consistent spatially located data over all relevant temporal periods. Kriging with such spatio-temporal inconsistent data can lead to anomalies and incorrect representations of groundwater level or head data. An improved representation of temporal trends in a traditionally spatial domain is necessary to avoid temporal anomalies that can lead to an inaccurate spatio-temporal analysis of groundwater data.

One approach for extending spatial geostatistics to space-time geostatistics is to recognize the importance of temporal relationships to help infer predictions. Spatio-temporal approaches incorporate time as a dimension and model covariance/variance as function of both space and time (Kyriakidis and Journel 1999; Heuvelink and Griffith 2010; Tadic et al. 2017). This approach improves the traditional spatial approach by leveraging data in neighboring time periods to compensate for sparseness in temporal data at specific locations. Spatio-temporal kriging has only been applied in three known studies to groundwater data analysis, including an early groundwater quantity study (Rouhani and Hall 1989), design of groundwater monitoring networks (Júnez-Ferreira and Herrera 2013), and recently groundwater level variations (Varouchakis and Hristopulos 2017). Despite the limited applications to groundwater, spatio-temporal kriging is common to a number of other environmental studies. These studies include the analysis and mapping of precipitation (Martínez et al. 2017), MODIS temperature and precipitation (Hengl et al. 2012; Hu et al. 2017), soil moisture, temperature, and electrical conductivity (Wang et al. 2015; Gasch et al. 2015), satellite observed CO_2 (Tadić et al. 2015; Tadic et al. 2017; Zeng et al. 2017), ozone data (Xu and Shu 2015), NO_2 pollutants (De Iaco and Posa 2012; Beauchamp et al. 2017), standardized precipitation index (Bayat et al. 2015), gamma dose rates (Heuvelink and Griffith 2010), solar irradiance forecasting (Aryaputera et al. 2015), and soil heavy metal distribution (Yang et al. 2015).

Current applications of spatio-temporal kriging for groundwater data are of limited application, but demonstrate the benefits of incorporating temporal relationships into the kriging process. Júnez-Ferreira and Herrera (2013) present methods for optimizing a groundwater head monitoring network using spatio-temporal geostatistics. Their focus was to maintain the current monitoring network's level of uncertainty in estimated hydraulic heads while reducing the spatio-temporal redundancy of monitoring data to cut costs for a portion of a 379 km^2 unconfined aquifer. Júnez-Ferreira and Herrera (2013) utilized the product-sum model (De Iaco et al. 2001) to model space-time covariance and Kalman filter to estimate the residuals in the methodology. Varouchakis and Hristopulos (2017) evaluate the performance of three spatio-temporal models in estimating groundwater levels. Separable Matérn, diffusive, and Spartan models were used to model space-time trends in variograms for an unconfined aquifer of ~26 km^2 using eleven wells. While these studies have recently reintroduced space-time geostatistics to the groundwater realm and show promising results, further validation and assessment are still needed. Studies are needed to 1) assess the applicability of additional space-time variogram models to predict groundwater levels, 2) evaluate aquifers on a larger scale, 2) further assess usefulness for confined, semi-confined, and unconfined systems, 3) evaluate longer time periods, 4) evaluate different methods to remove global groundwater trends or drift, and 5) advocate for the advantages of spatio-temporal applications on irregular or sparse groundwater datasets.

This research applies a spatio-temporal kriging approach to groundwater data at scales relevant to regional groundwater resource evaluations. The goal of this study is to improve potentiometric surface estimates at utilizing a regression model to remove the global trend and a sum-metric model to represent spatio-temporal covariances and variances in the sample data. Comparisons against spatio-temporal kriging will be made to traditional spatial kriging, both as monthly estimates and as an annual aggregate. The approach will be demonstrated using the Arapahoe Aquifer, a semi-confined aquifer and important groundwater resource for the greater Denver metro area in Colorado. Additionally, groundwater maps will be generated for the aquifer to understand changes over that past 27 years and identify areas of the aquifer that may be experiencing higher water stress.

2.3 Methods

2.3.1 Spatio-Temporal Random Function Model

Let $z = \{z(s,t)|s \in S, t \ \in T\}$ be a space-time dataset of a variable that is characterized by spatial coordinates $s = (x,y)$ and a time t observed at space-time points $(s_i, t_i) = 1, \dots n$ within spatial domain S and time interval T. Assuming that z is a realization of a spatio-temporal random function model, $Z(s,t)$, predictions of $z(s_0, t_0)$ at unmeasured space-time points can be made (Heuvelink et al. 2015). A basic model to consider for random function Z is (Kyriakidis and Journel 1999),

$$Z(s,t) = m(s,t) + \varepsilon(s,t) \tag{2-1}$$

where $m(s,t)$ is the deterministic trend and $\varepsilon(s,t)$ is the stochastic residual. The trend m can be assumed constant in space and time or allowed to vary as a function of known covariates. This component models the underling variation in the spatio-temporal process. The trend can be constructed by a regression type model that relates the variable of interest, z, to relevant covariates. The residual ε represents the difference between the observations and the predictions from the trend model. Assuming stationarity of spatio-temporal variance and covariance, the stochastic residual ε is assumed to be multivariate normally distributed and depends on spatial and temporal separation distances, h and u. In this study, Z is the target variable of groundwater levels represented as feet above mean sea level.

2.3.2 Trend Component

The trend can be represented as a function of known covariates that span the spatio-temporal domain. Assuming a linear relationship, $m(s,t)$ is represented as (Heuvelink and Griffith 2010),

$$m(s,t) = \sum_{i=0}^{p} \beta_i f_i(s,t) \tag{2-2}$$

where β_i are the regression coefficients (β_0 is the estimated intercept), f_i are the known covariates, and p is the number of covariates. The regression coefficients β_i are estimated from some method, such as ordinary least squares or generalized least squares. Before using a linear regression to determine the trend component for this study, we need to account for the multicollinearity, or overlap in information, in the covariates.

Principal component analysis (PCA) is useful to reduce this overlap. Two covariates used in this study are a digital elevation model and the upper boundary of the Arapahoe Aquifer represented as rasters. Before running the analysis, the two input raster's were linearly stretched to the binary scale 0-255 (Hengl 2009). Upon completion of the PCA, the first component, PCA1, explained 76% of variance and the second, PCA2, 24%. Using PCA1 and PCA2 in place of the digital elevation model and the upper boundary of the aquifer, a regression analysis was performed using various combinations of the predictors, including PCA1, PCA2, latitude (Lat), longitude (Lon), year the measurement was taken, and month the measurement was taken. PCA1 and Lon were determined to be the most important predictors. We assumed a global trend that was temporally constant and only consisted of spatial components. Therefore, the trend $m(s,t)$ for groundwater levels becomes $m(s)$,

$$m(s) = 6.919PCA1 + 0.003509Lon\ + 2504 \quad (2\text{-}3)$$

This regression model explained 90% of the variability in the observed groundwater data. Each variable is statistically significant (p-value <0.001), therefore rejecting the null hypothesis that β = 0. The residuals of the model are normally distributed. The fitted linear regression model is used to remove the global trend in the groundwater level data.

2.3.3 Residual Component Analysis

The estimated spatio-temporal trend $m(s)$ is subtracted from the groundwater level point observations to yield the spatio-temporal residual $\varepsilon(s,t)$. The residual component may have variations and correlations in space and time. The resulting residuals are used to construct the sample semivariogram in space and time.

The sample spatio-temporal semivariogram is calculated as,

$$\gamma(h,u) = \frac{1}{2 \cdot N(h,u)} \sum_{i=1}^{N(h,u)} [\varepsilon(s,t) - \varepsilon(s+h, t+u)]^2 \quad (2\text{-}4)$$

where h is the separation distance for points in space, u is the separation in time, and $N(h,u)$is the number of paired observations of z separted by lag (h,u). The structure of the spatio-temporal semivariogram is similar to that of the purely spatial case in that locations are binned according to their separation distance, or lag. However, the spatio-temporal case uses spatial and temporal distance yielding a semivariogram surface (Heuvelink et al. 2015). The space-time semivariogram solely depends on the lags in space and time. Implementation of spatio-temporal geostatistics was utilized the R software (Team 2014) and gstat and space-time packages (Pebesma 2004; Pebesma 2012; Gräler et al. 2016).

Once the sample spatio-temporal semivariogram has been calculated (Figure 2-3), a model may be fitted. A number of studies have used various models to estimate spatio-temporal covariances and variances, and include the metric model (Dimitrakopoulos and Luo 1994), non-separable models (Cressie and Huang 1999), product sum model (De Iaco et al. 2001; De Iaco and Posa 2012; De Iaco et al. 2012), and the sum-metric model (Myers 2004; Derakhshan and Leuangthong 2006). Further information on space-time models can be found in Kyriakidis and Journel (1999). We utilize the sum-metric model for this study.

The sum-metric model is a combination of the sum and metric models (Rouhani and Hall 1989; Dimitrakopoulos and Luo 1994; Myers 2004; Derakhshan and Leuangthong 2006). The sum model assumes that the sum of the spatial and temporal variogram, or covariance equivalent, represents the spatio-temporal variance. This model treats space and time separately, however it fails the strict definiteness condition for the resulting spatiotemporal variogram and covariance function (Myers and Journel 1990). This problem is resoled when combined with the metric model. The sum-metric model is represented as:

$$\gamma_{ST}(s,t) = \gamma_S(h) + \gamma_T(u) + \gamma_J\left(\sqrt{h^2 + (\kappa \cdot u)^2}\right) \tag{2-5}$$

where κ is the space-time anisotropy ratio that merges spatial distances with temporal distances, and γ_S, γ_T, γ_J, are the spatial, temporal, and joint semivariograms with separate nugget-effects.

Parameters of the spatio-temporal semivariogram (i.e. nugget, range, sill, space-time anisotropy ratio) are estimated by fitting a space-time model to the sample variogram (Heuvelink et al. 2015). We use a Spherical model for the spatial component, a Gaussian model for the temporal component, and a Spherical model for the joint semivariogram. These are represented as,

$$\gamma_S(h)(Spherical) = \begin{cases} \sigma_S^2\left(\frac{3h}{2a_S} - \frac{1}{2}\left(\frac{h}{a_S}\right)^3\right) + \sigma_{nug\_s}^2,\ 0 < h \leq a_S \\ \sigma_S^2 + \sigma_{nug\_s}^2\ ,\ h > a_S \\ 0,\ h = 0 \end{cases} \tag{2-6}$$

$$\gamma_T(u)(Gaussian) = \begin{cases} \sigma_T^2\left(1 - exp\left(-\frac{u^2}{{a_T}^2}\right)\right) + \sigma_{nug\_t}^2\ ,\ u > 0 \\ 0,\ u = 0 \end{cases} \tag{2-7}$$

$$\gamma_J\left(\sqrt{h^2 + (\kappa \cdot u)^2}\right)(Spherical) =$$

$$\begin{cases} \sigma_J^2\left(\frac{3\sqrt{h^2+(\kappa\cdot u)^2}}{2a_J} - \frac{1}{2}\left(\frac{\sqrt{h^2+(\kappa\cdot u)^2}}{a_J}\right)^3\right) + \sigma_{nug\_j}^2,\ 0 < \sqrt{h^2 + (\kappa \cdot u)^2} \leq a_J \\ \sigma_J^2 + \sigma_{nug\_j}^2\ ,\ \sqrt{h^2 + (\kappa \cdot u)^2} > a_J \\ 0,\ \sqrt{h^2 + (\kappa \cdot u)^2} = 0 \end{cases} \tag{2-8}$$

Where σ^2 and a are the variance and range of the spatial, temporal, and joint quantities being mapped, and σ_{nug}^2 is the nugget variance, representative of measurement and data errors in the groundwater level observations. Although other semivariogram models were adequately capable of reproducing the sample semivariogram, the sum-metric model best reproduced the lower semivariances between spatial separation distances from 0 to 25,000 meters and temporal separation from 0 to 100 months, compared to other models (Figure 2-3). Once a model was fit to the sample spatio-temporal semivariogram, kriging was applied to the residuals to predict data at unmeasured locations.

2.3.4 Kriging

Kriging is performed in a similar way as to the purely spatial case. In the case of regression kriging, separate predictions are made for the trend and residual components and then added back together. Regression kriging is mathematically equivalent to kriging with external drift and universal kriging, such that the same input parameters are used (Hengl et al. 2003; Hengl et al. 2004; Hengl et al. 2007). However, the methods differ in how the steps are applied. Spatio-temporal kriging on the stochastic residuals is achieved through ordinary kriging and is the best linear unbiased prediction of $\varepsilon(s_0, t_0)$ as (Hengl et al. 2004; Varouchakis and Hristopulos 2017),

$$\hat{\varepsilon}(s_0, t_0) = \sum_{i=1}^{n} \lambda_i \cdot \varepsilon(s_i, t_i) \tag{2-9}$$

where λ_i are the spatio-temporal kriging weights determined by the spatio-temporal dependence structure of the residuals and $\varepsilon(s_i, t_i)$ are the sample residuals in the neighborhood of the prediction location (s_0, t_0). The optimal kriging weights are obtained by solving (Mateu 2015),

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \cdot \gamma_{ST}\left(s_i - s_j, t_i - t_j\right) + \mu = \gamma_{ST}(s_i - s_0, t_i - t_0)\,, \quad \forall i = 1, \dots, n \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases}, \tag{2-10}$$

where μ is the Lagrange multiplier and n is the number of observations limited to the search neighborhood. Due to the size of the space-time dataset, kriging was limited to the nearest 1000 observations in terms of the degree of correlation. The kriging was also tasked to three clusters, allowing several copies of R to run, which allowed for parallelized computations. These measures decreased computational times for the spatio-temporal kriging and increased the efficiency. The spatio-temporal kriging produced interpolated maps for the entire space-time domain, encompassing 324 months between 1990 and 2016.

The final spatio-temporal estimate of groundwater level at location (s_0, t_0) is given by adding the trend and residual components together,

$$\hat{z}(s_o, t_0) = \widehat{m}(s_0) + \hat{\varepsilon}(s_0, t_0) \tag{2-11}$$

Estimation of the prediction variance is given by (Mateu 2015),

$$Var\big(\hat{\varepsilon}(s_o, t_0) - \varepsilon(s_o, t_0)\big) = \sum_{i=1}^{n} \lambda_i \cdot \gamma_{ST}(s_i - s_0, t_i - t_0) + \mu \tag{2-12}$$

The variance is used to help understand the uncertainty associated with the kriging predictions.

2.3.5 Case Study

The Denver Basin Aquifer System (DBAS) is a major source of water for residents of the Rocky Mountain Front Range and Denver metropolitan area, and underlies an area of about 17,350 km^2 (6,700 mi^2) (Moore et al. 2007). It includes four layered aquifers: the Denver, Dawson, Arapahoe, and Laramie-Fox Hills (Figure 2-1). Each aquifer is located at different depths, and has varying water quality and availability. The aquifer system is recognized as nonrenewable due to it predominately being confined and that it receives little recharge (Graham and Vanslyke 2004).

The Arapahoe Aquifer is the third deepest aquifer of the DBAS and is located within the Arapahoe Formation (Figure 2-1). It has an extent of approximately 11,135 km^2 (4,300 mi^2) and consists of interbedded conglomerate, sandstone, siltstone, and shale (Robson 1987; Robson and Banta 1995). The top of the aquifer is identified as the shale beds of the lower Denver Formation, while the bottom is the top of the shale, coal seams, and beds of sandstone and siltstone from the Laramie Formation shale (Robson 1987; Robson and Banta 1995). The Arapahoe is predominately confined, however crops out along the northern, eastern, and southern edges. The Arapahoe aquifer is desirable for municipal water supply due to its excellent water quality and higher yielding wells (as much as 700 gallons per minute) (Robson and Banta 1995).

Since the 1970s, there has been concern about the sustainability and protection of water resources in the DBAS due to the number of wells extracting water and the increasing population. Despite these legal measures to protect the DBAS, waters users continue to report decreasing water levels and increased pressure to meet increasing populations (Banta 1989; Topper and Raynolds 2007; Moore et al. 2007; Banta and Paschke 2012). Sustainable use of the DBAS requires the ongoing need to improve our understanding of the current state of each aquifer, how each has evolved through time, and an assessment of water level declines and areas experiencing higher water stress. Spatio-temporal techniques offer an opportunity to use existing groundwater level data to better understand the system response, such as water level changes, over the past several decades. Spatio-temporal kriging has yet to be applied to any aquifer in the

DBAS. The results present the most up to date ground-based data study for the Arapahoe Aquifer.

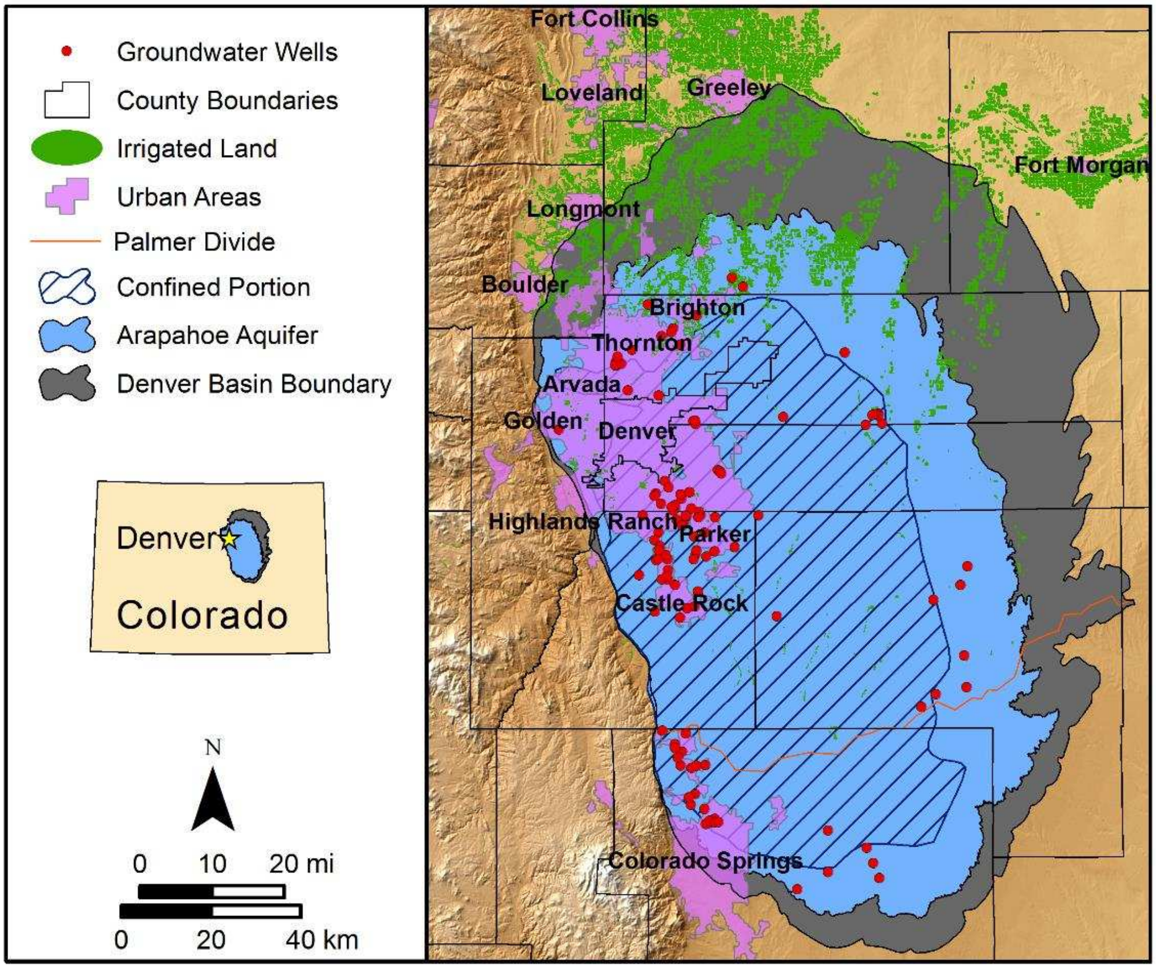

Figure 2-1 Location of the Arapahoe Aquifer in northern Colorado, USA. The diagonal hatch represents the general area where the aquifer is confined. The red points represent available groundwater wells used in the spatio-temporal kriging study. The map also shows locations of urban areas and irrigated land, both large end users of groundwater in the area.

Groundwater level data were obtained through the Colorado Division of Water Resource's Colorado HydroBase (Colorado Division of Water Resources 2017). HydroBase is a central database for water resources data in Colorado and contains historical records from a variety of sources. Many groundwater wells were sampled beginning in the 1980s, with the addition and removal of wells over time. The data were clipped to span the time period between 1990 and 2016 and only wells with a minimum of five observations were included for further analysis. Groundwater wells that did not appear to be at steady state, i.e. pumping wells or under pumping conditions, were also removed. We identified 118 groundwater wells that met these criteria.

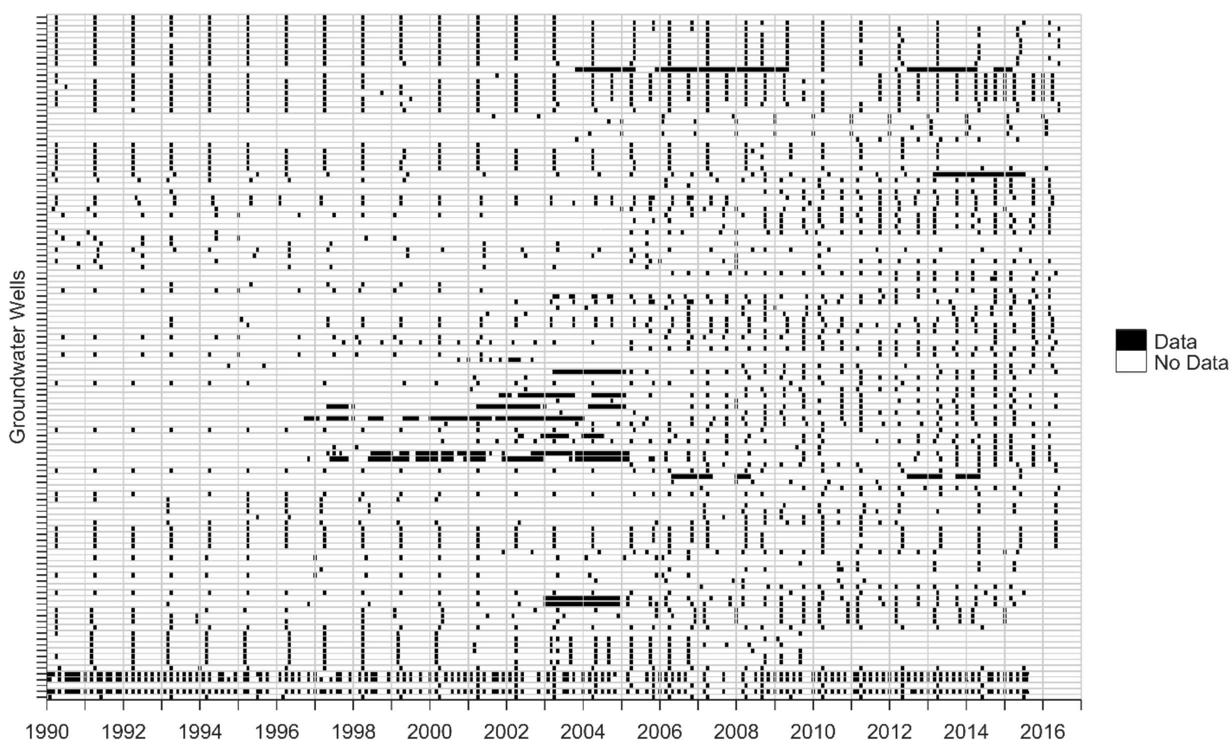

Figure 2-2 Heat map showing available monthly groundwater level data for 118 wells in the Arapahoe Aquifer from 1990 - 2016. Each tick on the y-axis is a unique well. Wells are arranged from increasing latitude (bottom to top) and highlight the irregularity of water level measurements and temporal sparseness of the data.

The distribution of groundwater wells spatially covers all ends of the aquifer, but does highlight areas that lack adequate well coverage, especially in the south-central region (Figure 2-

1). Spatial sparseness of data is a problem common to many aquifer systems. The data were also explored to identify the temporal frequency at which water level measurements were taken (Figure 2-2). Many wells are sampled annually, with a high frequency of these measurements occurring in April. However, many wells are also sampled several times per year, including some that had continuous monthly measurements. Each groundwater well time series was evaluated for outliers that could indicate data measurement or data entry errors. A twelve point moving window was used to identify and remove outliers that were greater that two standard deviations of the moving mean.

2.4 Results and Discussion

2.4.1 Spatio-Temporal Semivariogram

The sample spatio-temporal semivariogram determined from the groundwater level residuals is shown in Figure 2-3. The assumption of stationarity is supported by the spatio-temporal semivariogram of the residuals approaching a constant sill. Semivariance generally increases with increasing time lag, indicating that groundwater level residuals become less similar with increasing temporal distance. A dip in semivariance is observed at separation distances of around 20,000 meters. This would indicate that locations separated by 20,000 meters are more similar than say distances of 10,000 meters. Intuitively this does not make sense, as there are no geological or other system characteristics that could explain this pattern. However, when graphically evaluating these locations, two distinct clusters of data appear across the aquifer due sparseness of spatially distributed data. At this separation distance, these distinct clusters form and lead to the drop in semivariance. It is likely however, that in the absence of the clustering and if there were more spatial and temporal observations available, the true structure would have higher a semivariance at this range. Therefore, the spatio-temporal model was not fit to this dip in semivariance.

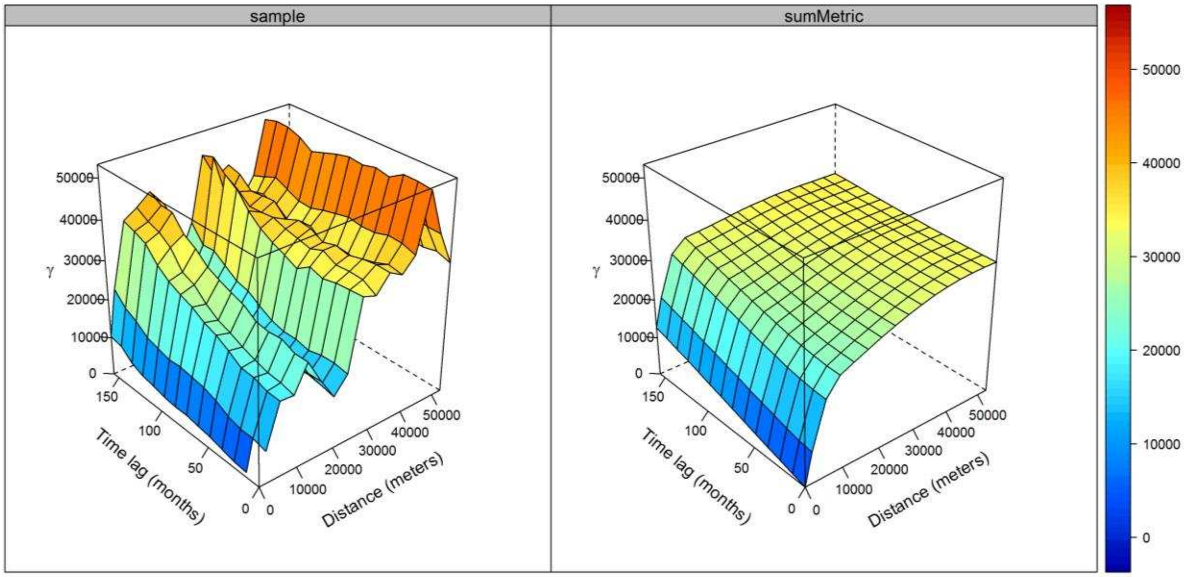

Figure 2-3 Sample (left) and fitted (right) spatio-temporal semivariogram surfaces of the residuals.

The sum-metric model fit to the sample spatio-temporal semivariogram is shown in Figure 2-3. The respective parameters for each of the three variograms used in the model are found in Table 2-1. The model does a good job of fitting the overall structure of the sample spatio-temporal semivariogram. Similar to the sample spatio-temporal semivariogram, the sum-metric model shows increasing semivariance with increasing time lag and spatial distances.

Table 2-1 Parameters for fitted sum-metric model semivariogram.

	Model	pSill	Range	Nugget	stAni	RMSE	MAE
Spatial	Spherical	18341	8121	1150	175	6883	4928
Temporal	Gaussian	7611	290	0			
Joint	Spherical	13841	46711	0			

2.4.2 Cross-Validation

"Leave-one-out" cross-validation was used to assess the kriging model performance that would lead to the best predictions. The cross-validation is applied to the existing groundwater level samples. For comparison to the performance of spatio-temporal kriging, cross-validation was also conducted using traditional spatial kriging both at the monthly interval and as an annual aggregate (Figure 2-4). During cross-validation, each spatio-temporal data location is removed one at a time, while the remaining dataset are used to predict the associated value. Measured and predicted groundwater elevations are then compared for goodness of fit. Due to computational times, the spatio-temporal dataset was limited to the nearest 1,000 observations in terms of the degree of correlation. The leave-one-out approach was repeated for all spatio-temporal locations and all spatial locations for the traditional spatial kriging. This approach provides a means of choosing plausible spatio-temporal semivariogram models.

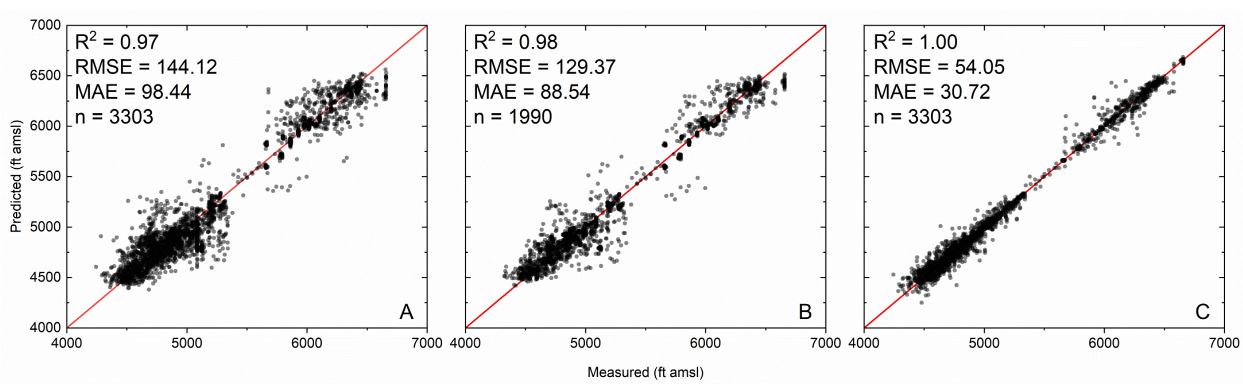

Figure 2-4 Cross-validation scatterplots of potentiometric surface elevations at well observation locations showing the measured surface elevation compared to the predicted elevation using A) spatial kriging with monthly data, B) spatial kriging with annually aggregated data, and C) spatio-temporal kriging.

The cross-validation results show better performance for the spatio-temporal kriging approach than traditional spatial kriging. The spatio-temporal model used in the kriging cross-validation produces points that are roughly symmetrical around the diagonal line (Figure 2-4 C) and indicate that the predictions are unbiased. While the traditional spatial kriging cross validation results also produce points roughly symmetric around the diagonal line, more spreading is evident and show less accurate predictions. In addition to a graphical assessment, four statistics were calculated, the correlation coefficient (R^2), the root mean squared error (RMSE), and mean absolute error (MAE). The high R^2 shows that the kriging predictions are reliable and that there is a high linear correlation between the measured and predicted values. The MAE is similar to RMSE, but is less sensitive to outliers and has a lower value. The spatio-temporal kriging a higher R^2 and lower RMSE and MAE than both spatial kriging variations indicating better prediction of measured groundwater elevations. Spatial kriging as an annual aggregate performs better than spatial kriging at the monthly time interval due to greater spatial data density of data in each time bin to infer the predictions. The results indicate that the spatio-temporal semivariogram from the sum-metric model reasonably represents the spatio-temporal structure of the groundwater level residuals, leading to accurate predictions. Therefore, we can use the spatio-temporal kriging approach to generate potentiometric surfaces for the Arapahoe Aquifer and make estimates at unsampled locations.

2.4.3 Comparison of Spatial to Spatio-Temporal Kriging Maps

Spatial potentiometric surface maps were also created for monthly observed data and data as an annual aggregate using traditional spatial kriging. At each monthly time step, only the wells with observed data in a given month were used to infer the kriging predictions. The number of groundwater wells changes for each time step based on data availability. Differences were calculated between the monthly spatial maps to show monthly changes in the potentiometric surface. For comparison of estimated temporal changes for consecutive months, differences were also calculated for potentiometric surfaces estimated through the spatio-temporal kriging approach. A subset of the maps is shown for April 2004, May 2004, and the monthly difference (Figure 2-5). For the maps created using spatial kriging, changes in groundwater levels primarily caused by the groundwater wells used in the analysis. Additional groundwater wells on the eastern side of the aquifer in May 2004, and variable data availability on the northeastern side for both months, results in large changes between the months. In some places there is up to either a 350 ft decrease in levels or increase up to 150 ft, which are unrealistic changes over a period of one month. For spatio-temporal kriging, the monthly predicted potentiometric surfaces are similar and lend to reasonable monthly changes in levels. Similarly, spatial potentiometric surface maps and difference maps were created for data that are aggregated annually. At each annual time step, wells with multiple measurements during the time period were averaged. The average value was used in the spatial kriging for locations with multiple measurements. A subset of the annual maps is shown for 2005, 2006, and the annual difference (Figure 2-6). Some of the changes in groundwater levels are caused by different available data in each year, but also due to average well data being used in place of actual measurements for different locations. Comparison to the results of spatio-temporal kriging show large differences. Annual changes for spatio-temporal kriging are much smaller than those predicted by spatial kriging of annually aggregated data. The results of the spatial kriging also show unreasonably large areas of increasing or decreasing levels that are likely due to artifacts of aggregating well data and the spatial distribution of available data in each year.

Spatio-temporal kriging produces more realistic monthly and annual changes in groundwater levels. Because spatio-temporal kriging is leveraging data from spatio-temporal neighbors, more information is available to infer the kriging predictions, yields more accurate results, and has smoother transitions. These examples illustrate the problem of using observed data that varies throughout time to create spatial groundwater maps. Changes are then a result of the groundwater wells used in the analysis, rather than actual groundwater level changes.

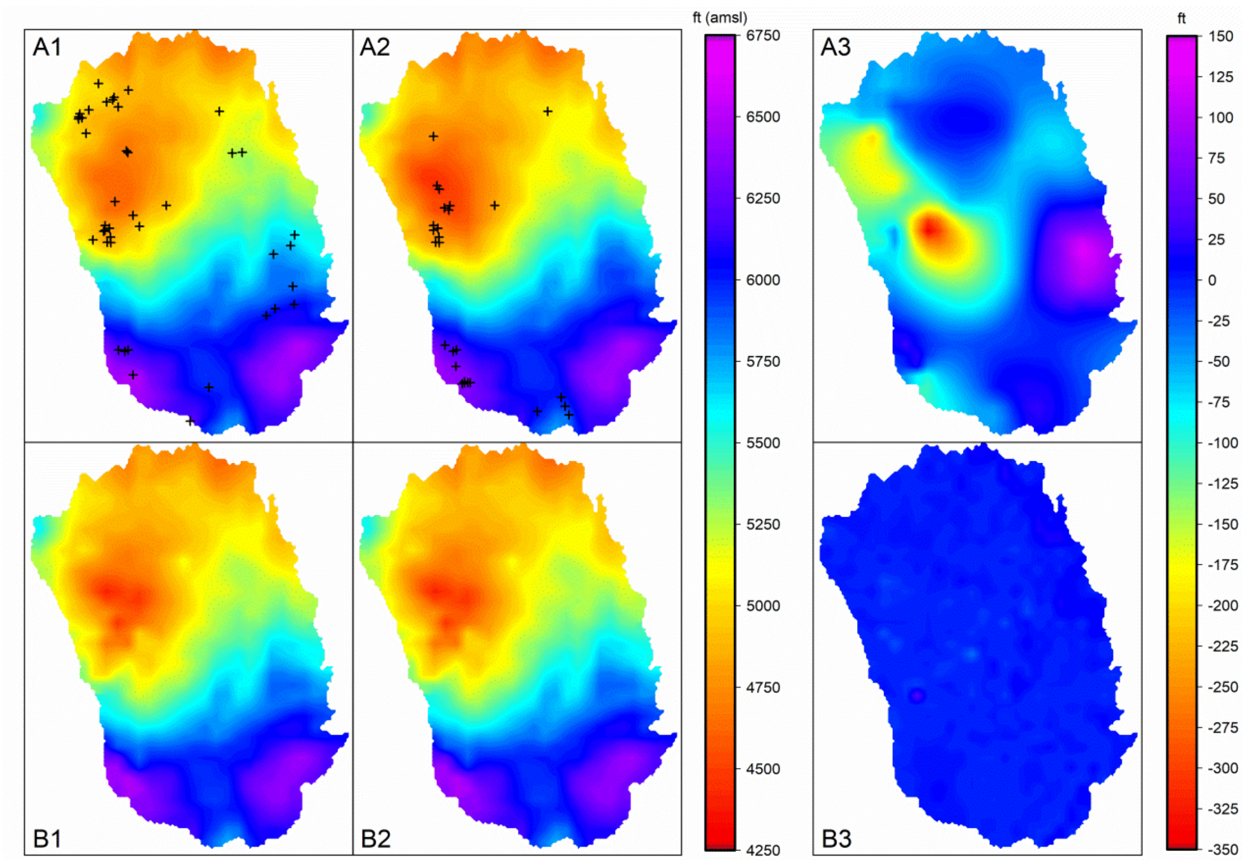

Figure 2-5 Spatial maps using traditional spatial kriging (A) and spatio-temporal kriging (B) for (1) April 2004, (2) May 2004, and (3) the difference between April 2004 and May 2004 potentiometric surfaces. The groundwater wells used in the spatial kriging analysis are shown as '+' in A1 and A2.

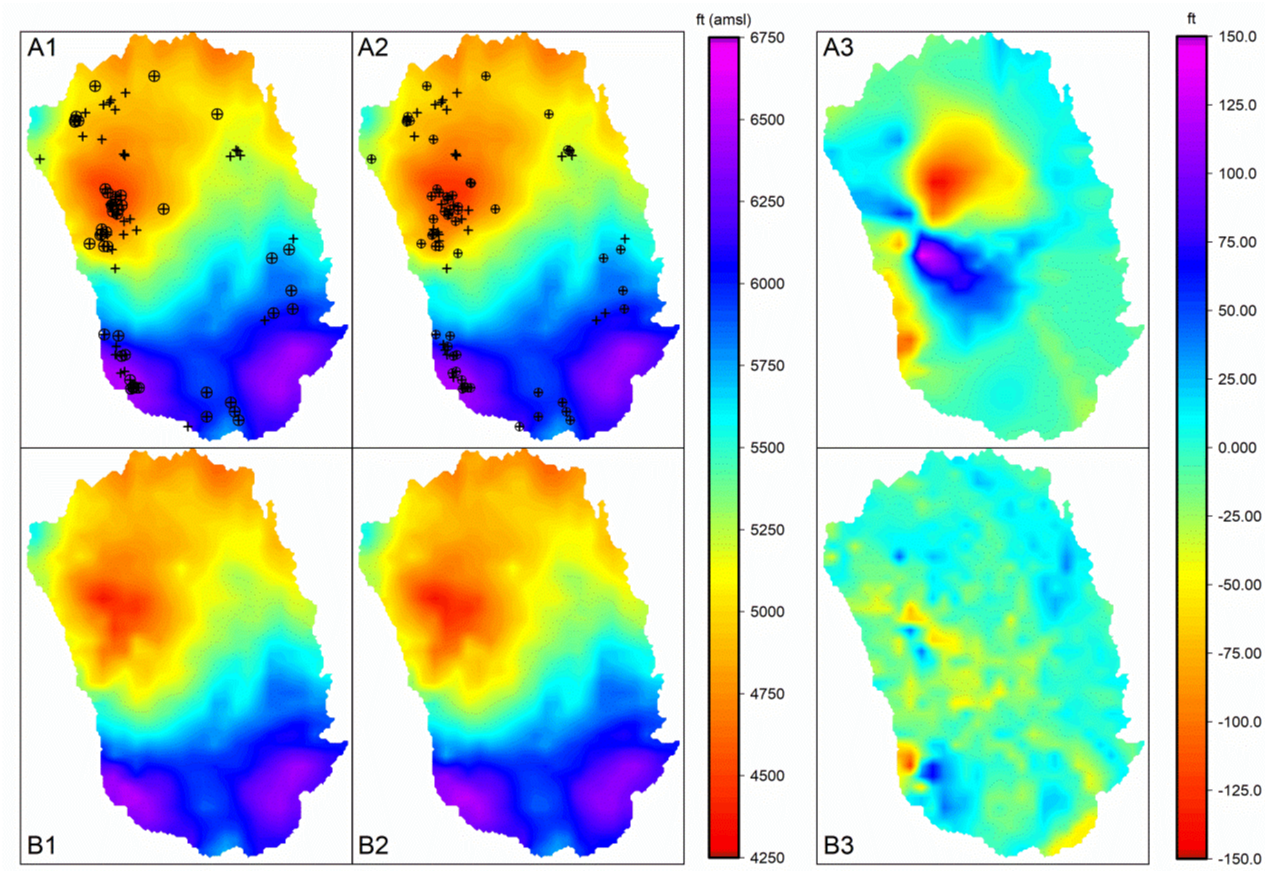

Figure 2-6 Spatial maps using traditional spatial kriging aggregated annually (A) and spatio-temporal kriging (B) for (1) 2005, (2) 2006, and (3) the difference between 2005 and 2006 potentiometric surfaces. The groundwater wells used in the spatial kriging analysis are shown as '+' in A1 and A2. Wells that had multiple measurements and that were aggregated contain a circle around the well.

2.4.4 Spatio-Temporal Kriging Predicted Groundwater Levels

The spatio-temporal kriging approach is applied to the Arapahoe Aquifer from 1990 to 2016 generated 324 monthly groundwater level maps, represented as potentiometric surfaces in feet above mean sea level (ft amsl). A subset of these maps is presented in Figure 2-7 (left side), showing the predicted potentiometric surfaces for the month of April in 1995, 2000, 2005, 2010, and 2015. Each map shows a trended potentiometric surface with higher elevations to the south, reaching 6,750 ft amsl, and lower elevations to the north, reaching 4,250 ft amsl. The spatio-

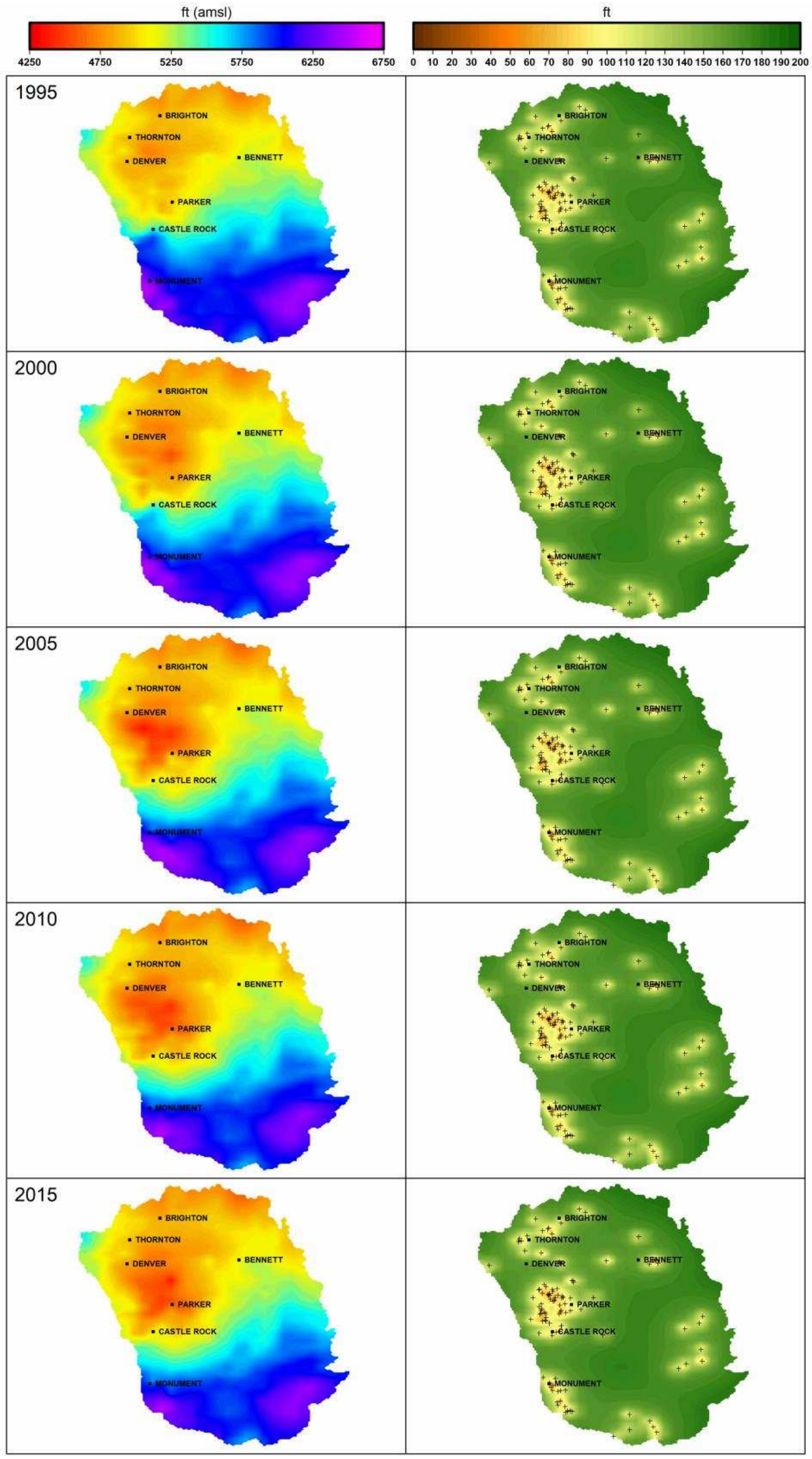

Figure 2-7 Predicted potentiometric surfaces (left column) and standard deviation maps (right column) of the Arapahoe Aquifer. All maps shown are for the month of April in select years. The black '+' on the standard deviation maps show the locations of groundwater wells used in the analysis.

temporal patterns remain similar across the years. However, as shown in Figure 2-7, the potentiometric surface from 1995 to 2000 shows decreased groundwater level elevations near Parker and push south towards Castle Rock. By 2005, the decrease in levels is even larger, indicated by the deeper reds and orange colors on the maps. The light blue color band represent groundwater level elevations around 5,550 ft amsl also appears to get larger and push south of Castle Rock towards Monument. By April 2015, groundwater levels in the aquifer appear to show some recovery and increases in potentiometric surface levels.

The kriging standard deviations, root of the variance, are used to quantify the prediction uncertainty (Figure 2-7 right side). The right side of Figure 2-7 shows the spatial distribution of the standard deviations over the aquifer area as well as the locations of groundwater wells used in the analysis. Each map corresponds to the uncertainty in potentiometric surface predictions for the five April maps on the left. A kriging standard deviations map was produced for each of the 324 months evaluated. As each map shows, lower prediction standard deviations follow the spatial distribution of the groundwater wells used in the predictions. The maps indicate higher confidence in the kriging predictions in areas around the spatial locations of the groundwater wells. Moving to areas where measured data did not exist, i.e. no groundwater wells, standard deviations in the predictions are as high as 200 ft.

Most areas of the Arapahoe Aquifer show decreased groundwater levels from those that existed in 1990 (Figure 2-8). The eastern, shallower portion of the aquifer shows generally less than 50 ft decrease in levels, but less than 25 ft near the groundwater wells where higher confidence exists. A large area of declines greater than 100 ft occurred on the west-central side of the aquifer. The larger decreases occurred under the cities of Parker, Castle Rock, and Monument where levels have dropped by 500 – 600 ft since 1990. These major decreases occur in the confined portion of the aquifer and indicate decreases in the confining pressure. Several areas do show increase levels since 1990, up to 90 ft along a portion of the southwestern edge.

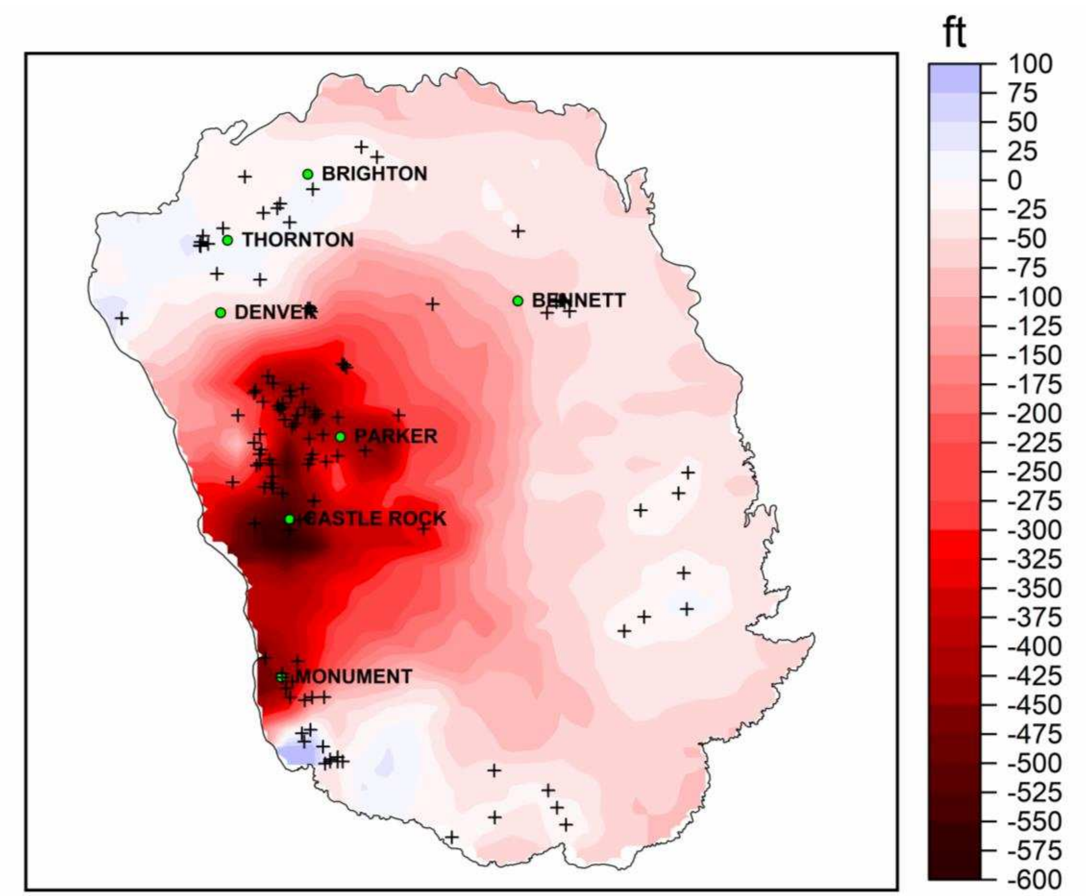

Figure 2-8 Changes in groundwater levels in the Arapahoe Aquifer from 1900 to 2016. The green '+' show the locations of groundwater wells used in the analysis.

2.5 Conclusions

This study demonstrates a methodology for spatio-temporal kriging of groundwater level data and improves predictions of potentiometric surfaces, especially in irregularly sampled and data sparse aquifer systems. Although groundwater level data are often irregularly sampled and spatially sparse, spatio-temporal kriging models the covariance in both space and time and utilizes more information to make estimates at unsampled locations. This approach leverages groundwater level data in adjacent time periods to estimate groundwater levels in the time period of interest and has several advantages over a purely spatial approach to estimating potentiometric surfaces. Spatio-temporal kriging allows for 1) estimation of groundwater levels during times when data is not available, 2) it avoids biases and anomalies caused by kriging with different data available in different time periods, and 3) provides a flexible framework for modeling the spatio-temporal structure.

CHAPTER 3

ASSESSMENT OF GROUNDWATER DEPLETION AND IMPLICATIONS FOR MANAGEMENT IN THE DENVER BASIN AQUIFER SYSTEM

3.1 Abstract

The Denver Basin Aquifer System (DBAS) is a critical groundwater resource for users along the Colorado Front Range Urban Corridor. Groundwater depletion has been documented over the past few decades due to the increased water use among users, presenting long-term sustainability challenges. A spatio-temporal geostatistical analysis is presented using Colorado Division of Water Resources data to estimate potentiometric surfaces and evaluate groundwater storage changes between 1990 and 2016 in each of the four Denver Basin aquifers. The approach utilizes neighboring space-time data to infer the predictions for the timer period of interest. Several key depletion patterns and spatial water level changes emerge in this work. First, hydraulic head changes are the largest in the west-central side of the DBAS and have decreased in some areas by up to 600 ft since 1990. Second, areas to the northwest of Denver show increases in hydraulic head by over 100 ft. Third, the Denver and Arapahoe aquifer show the largest cumulative groundwater depletion volumes, with the highest rates occurring in the 2000s. During this time, parts of the Arapahoe aquifer were declining by up to 55 ft per year. Our results highlight the uncertainty in volumetric predictions under various storage coefficient calculations and emphasize the need for representative aquifer characterization. The observed groundwater storage depletions are due to a combination of factors, which includes population growth

increasing the demand for water, variable precipitation and drought influencing recharge, and increased groundwater pumping.

3.2 Introduction

During the past several decades, competition for water resources has brought significant challenges to Colorado's water supply (Colorado Water Conservation Board 2011). As the population along the Colorado Front Range Urban Corridor (CFRUC) continues to increase, so does the demand on water resources (Moore 2007). It is estimated that in 2016, the Urban Corridor had a population of over 4.5 million, an increase of 73 percent since 1990 (Colorado Department of Local Afairs 2018). During this time period, municipalities such as Parker grew by 856%, Castle Rock by 558%, Monument 511%, and Denver by 48% (U.S. Census Bureau 2005b, 2017b). Strong population growth is recognized as one of the biggest drivers for increasing future municipal and industrial (M&I) water resource needs in the area (HDR Inc. and Consultants 2015; Colorado Water Conservation Board 2011). Because there is little unappropriated surface waters available in the CFRUC, new water demands will require transferring water away from other users and increasing the dependence on groundwater resources (Moore et al. 2007; HDR Inc. and Consultants 2015). This poses future groundwater sustainability challenges in the region, particularly for the Denver Basin Aquifer System (DBAS).

The potential for increased reliance on water resources stresses the importance for improved understanding of groundwater dynamics in the DBAS and the need to proactively manage available resources to ensure long-term sustainability and resilience. Some detrimental effects of groundwater depletion include reduced well yields, increased pumping costs, the need to drill deeper wells, land subsidence, reduced water quality, and reduced flows to surface waters (Alley et al. 1999; Konikow and Kendy 2005).

Numerous groundwater models have been developed and used to analyze groundwater storage changes and the influences of pumping in the DBAS (Robson 1987; Banta 1989; Paschke et al. 2011; Banta and Paschke 2012). Comparison of past modeling studies and Colorado Division of Water Resources (CDWR) data has shown variable results and rates of annual depletions across the different aquifers of the DBAS (Paschke et al. 2011; Banta and Paschke 2012; Everett 2014; Flor 2015, 2016). While models can be advantageous to dynamically estimate many variables, there are disadvantages due to the large amount of spatial and temporal input data needed and they can be time and labor intensive. In addition, groundwater models are also subject to accumulating errors due to the large number of parameters estimated in the modeling process and the associated uncertainty (Xu et al. 2012; Reilly and Harbaugh 2004). Spatial interpolation of groundwater level data can potentially simplify modeling efforts to determine water level and storage changes. However, most groundwater data are irregularly sampled and lack adequate spatial and temporal coverage.

Numerous spatial interpolation schemes exist (e.g. inverse distance weighting, nearest neighbors, kriging) to estimate data at unsampled locations and convert point groundwater level measurements to gridded groundwater surfaces (Li and Heap 2008).Kriging is considered the best linear unbiased estimator and is commonly applied to groundwater problems (Hengl et al. 2003; Maheswaran and Khosa 2013; Gundogdu and Guney 2007). However, data availability in different time periods can create anomalies and inaccurate representations due to the wells used in the analysis and not actual groundwater fluctuations. Spatio-temporal kriging approaches offer an alternative to overcome some of the limitation, especially for sparse and irregular data.

Applications of spatio-temporal kriging in the literature are limited to a few studies (Rouhani and Hall 1989; Júnez-Ferreira and Herrera 2013; Varouchakis and Hristopulos 2017), but common to many other environmental applications (Wang et al. 2015; Tadic et al. 2017; Hengl et al. 2012). The power of spatio-temporal kriging is leveraging neighboring space-time data to help infer the predictions and overcoming some of the sparse data limitations. It has previously been shown that spatio-temporal kriging is advantageous in the DBAS (Chapter 1).

Aside from water level data sparseness, another major uncertainty in regional groundwater storage change assessments is accurately quantifying aquifer storage parameters such as specific yield and specific storage. Changes can be difficult to assess due to sparse and unreliable aquifer characterization data (Gehman et al. 2009; Chen et al. 2003). Specific yield and specific storage can change through groundwater basins due to varying geology and changing depositional environments. Colorado Division of Water Resource's estimates of water storage for the DBAS use a single specific yield that is higher than measured values from cores near Castle Pines and Kiowa (Flor 2016; Woodard et al. 2002). Understanding the uncertainty in storativities is important for assessing the long-term sustainability as well as refining estimates about water availability.

The primary objective of this study is to apply a spatio-temporal kriging approach to the DBAS, with the goal of understanding groundwater storage changes in each aquifer, any spatial patterns, and the uncertainty of the storativity parameter in volumetric estimates. The current study builds upon prior work in the DBAS and presents a historical analysis of groundwater storage and level changes in the DBAS using spatio-temporal geostatistics on CDWR data. Our analysis extends beyond previous studies by including more recent data, accounting for the spatio-temporal correlations in the data, evaluating the uncertainty in groundwater storage calculations, and utilizes a different method to map and understand changes in groundwater levels in the DBAS.

3.3 Study Area

The DBAS is a critical water resource along the CRFUC, in addition to the South Platte River and Alluvial Aquifer (Paschke et al. 2011). It consists of four layered aquifers: from youngest to oldest are the Dawson, Denver, Arapahoe, and Laramie-Fox Hills, respectively (Figure 3-1). It is estimated that the DBAS contains over 200 million acre-feet (AF) of recoverable water (Topper and Raynolds 2007). Robson (1987) provided early estimates of water availability in the DBAS and calculated that of the 467 million AF of water that exists, only 269 million AF is recoverable. While these amounts are quite large, each aquifer yields variable amounts of water and can respond differently to pumping stresses due to the hydrogeological characteristics of the DBAS. As development progresses and demands increase in the CFRUC, these large reserves will be used to supplement water supplies not met by surface waters.

The majority of the DBAS underlies the South Platte River Basin (SPB) which is the most populous basin in the State and produces about 80% of the State's economy and tax base (HDR Inc. and Consultants 2015). Surface water diversions equate to more than 4 million AF per year and groundwater withdrawals exceed 450,000 AF per year (HDR Inc. and Consultants 2015). The SPB also supplements water resources through 400,000 AF per year of trans-basin diversions from the Colorado River Basin and around 100,000 AF per year from the Arkansas, North Platte, and Laramie River Basins (HDR Inc. and Consultants 2015). In 2004, the DBAS provided more than 70% of the south Denver metro areas municipal water supply (CDM 2004). On the eastern side of the DBAS, domestic groundwater use is widespread and many rural areas solely rely on the DBAS for water. In 2007, it was estimated that all four aquifers of the DBAS were permitted to withdrawal over 350,000 AF of water per year (Topper and Raynolds 2007). This has likely increased over the past decade and is projected to continue to increase.

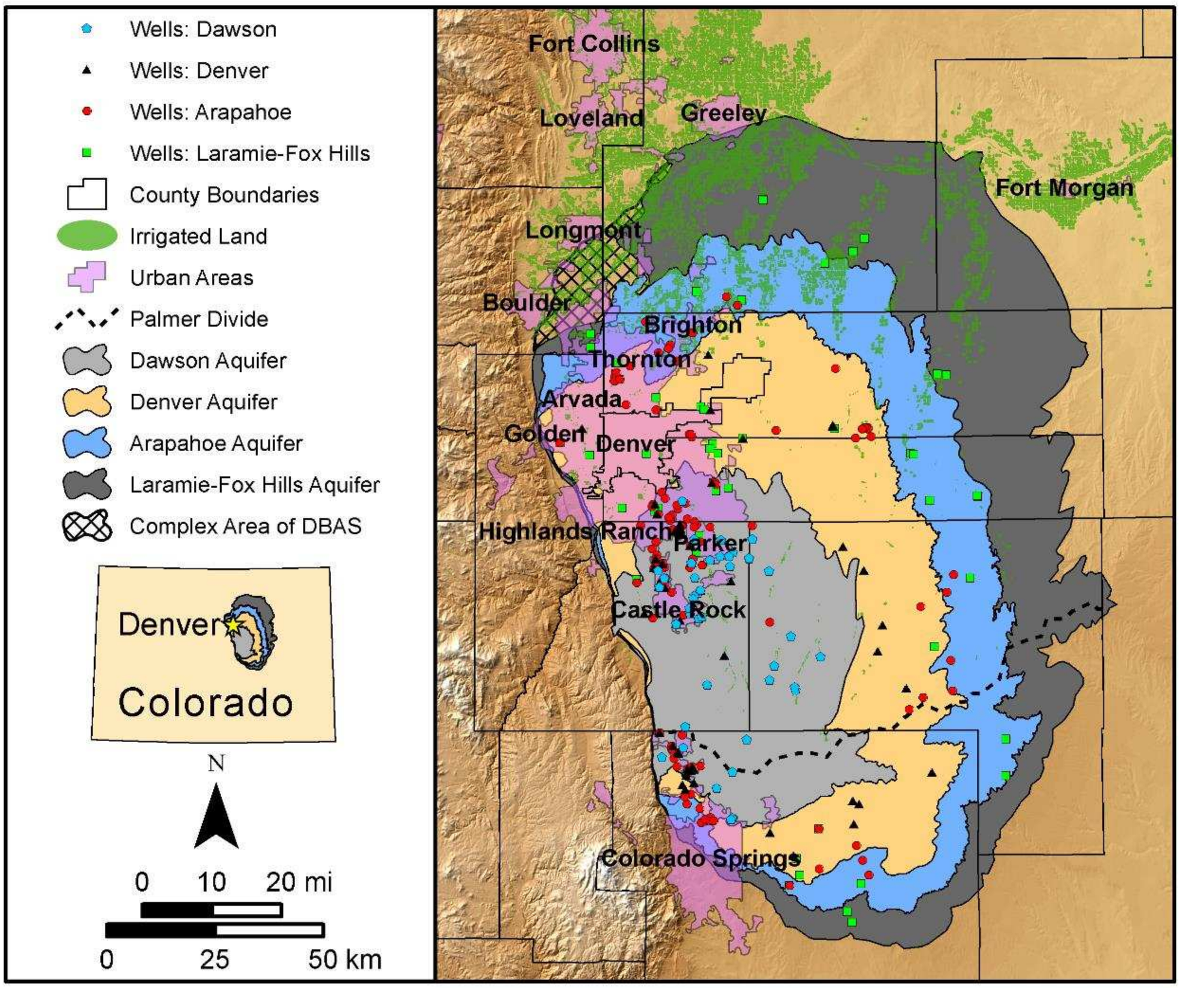

Figure 3-1 Location of the four aquifer of the Denver Basin Aquifer System in northern Colorado, USA. The different points represent available groundwater wells used in the analysis for each aquifer. The map also shows locations of urban areas and irrigated land, both large end users of groundwater in the area. The cross-hatch area represents a complex area within the DBAS and is not considered in this study.

Under medium demand scenarios, by 2050, the SPB M&I supply gap, the difference between existing supplies and future needs, is expected to approach 428,000 AF per year and an agricultural supply gap of 262,000 AF per year (HDR Inc. and Consultants 2015). Municipalities that are newer, at least over the past several decades, rely heavily on DBAS groundwater due to limited surface water available or the rights to use it (Hillier et al. 1978; Moore et al. 2007). The widespread dependence and increasing demand threaten the long-term sustainability and longevity of the DBAS. The effects of decreasing groundwater levels could lead to lower pumping yields for users, increased pumping costs, land subsidence, and reduced flows to streams, springs, and other water bodies (HDR Inc. and Consultants 2015; Van der Gun and Lipponen 2010; Alley et al. 1999; Topper and Raynolds 2007; Konikow and Kendy 2005).

Since the 1970s, there has been concern on the sustainability and protection of water resources in the DBAS (Graham and Vanslyke 2004; Moore et al. 2007). As a result, in 1973, the Colorado legislature passed Senate Bill 213 (SB 73-213) which specified how water pumped from nonrenewable aquifers should be managed (Colorado Session Laws 1973; White 1995). It set criteria for the State Engineer's Office to follow when issuing well permits and it tied pumping to land ownership. It set pumping rates at 1% of the available water per year, ensuring a minimum 100 year life of each aquifer (Jones and Cech 2009).. In 1985, the Colorado General Assembly passed Senate Bill 5 (SB 85-05) which allowed the State Engineer's Office to promulgate rules, known as the Denver Basin Rules, and regulations for groundwater withdrawals (Colorado Session Laws 1985; White 1995). It identified areas of non-tributary groundwater and presumed the 100 year life of the aquifers through pumping rights. However, approximately 47% of the DBAS is located within designated basins, as result of the 1965 Groundwater Management Act. These basins areas are administered by the Colorado Ground Water Commission (CRS 37-90-102) and not the State Engineer's Office, but follow similar practices (Jones and Cech 2009; Topper and Raynolds 2007). In 1996, Senate bill 96-74 (SB96-

74) was passed and commissioned the study of DBAS and South Platte Alluvial Aquiver to help develop decision support system for the SPB (Graham and Vanslyke 2004).

Groundwater depletion in the DBAS has been document over the past several decades and withdrawals have increased (Topper and Raynolds 2007; Colorado Foundation for Water Education 2006; Paschke et al. 2011; Banta and Paschke 2012; Everett 2014; Bauch et al. 2015). The Colorado Division of Water Resources publishes measurements from select wells in the DBAS during various years (Flor 2015, 2016). The data indicate that depending on the aquifer and location, water levels are declining, rising, or relatively stable. Since the 1980s, water level data collection has been irregular and even decreased during some years (Colorado Water Conservation Board 2006).

3.3.1 Groundwater Basins

The Denver Basin Aquifer System (DBAS) extends from Greely to Colorado Springs and from the Rocky Mountain Front Range to the east near Limon (Figure 3-1). It is an asymmetrical bowl shaped basin that dips steeply on the western side, and has lower-angle dipping on the eastern side (Robson 1987; Robson and Banta 1995; Anderman and Ackman 1963). The four layered aquifers have a maximum combined thickness of approximately 3,200 ft. near the Palmer Divide (Paschke et al. 2011). The aquifers are generally unconfined along the outer edges where they meet the surface and confined in the more central areas. Unconsolidated quaternary alluvial aquifers overlie the bedrock aquifers of the DBAS and some groundwater movement occurs between the alluvial and bedrock aquifers (Robson 1996). The DBAS is recognized as a nonrenewable resource because of the slow rate of recharge (Graham and Vanslyke 2004; Banta and Paschke 2012).

The Dawson Aquifer is the uppermost aquifer and the least extensive, covering approximately 1,200 mi^2. It is located within the Dawson Arkose Formation and consists of coarse grained, poorly to well consolidated sandstone interbedded with conglomerate, siltstone, and shale (Robson and Banta 1995; Major et al. 1983; Robson 1987). In the central part, the saturated thickness is between 300 and 400 feet (ft) thick.

The Denver Aquifer is the second uppermost aquifer located within the Denver Formation. It covers approximately 3,000 mi^2 and is in a sequence of moderately consolidated, interbedded shale, claystone, siltstone, and sandstone (Robson and Banta 1995; Major et al. 1983; Robson 1987). This aquifer is considered a leaky confined aquifer to the Dawson above and the Arapahoe Aquifer below (Robson and Banta 1995). Approximately 40% of the Denver Aquifer crops out at the land surface, the largest in the DBAS (Paschke et al. 2011).

The Arapahoe Aquifer is the third aquifer in sequence and is within the Arapahoe Formation. It has an extent of approximately 4,300 mi^2 and consists of interbedded conglomerate, sandstone, siltstone, and shale. The top of the aquifer is identified as the shale beds of the lower Denver Formation, while the bottom is the top of the shale, coal seams, and beds of sandstone and siltstone from the Laramie Formation (Robson and Banta 1995; Major et al. 1983; Robson 1987). This aquifer is the most permeable and highest yielding aquifer within the DBAS. On the northern end, the Arapahoe is divided into an upper and lower aquifer due to a confining shale (Robson 1987). Potential hydraulic connection exists between the Arapahoe and Denver Aquifers on the west-central part of the basin (Paschke et al. 2011).

The Laramie-Fox Hills Aquifer is the deepest in the DBAS and is located within the Fox Hills sandstone and sandstones of the lower Laramie Formation. The aquifer covers an area of approximately 6,700 mi^2 and is the maximum extend of the DBAS. The aquifer consists of fine to very fine grained sandstone or siltstone with interbedded shale and is poorly consolidated (Robson and Banta 1995; Major et al. 1983; Robson 1987). The impermeable Pierre Shale underlies the aquifer and forms the base of the DBAS.

3.4 Methods

3.4.1 Groundwater Level Data

Groundwater data was obtained through the Colorado Division of Water Resource's HydroBase (Colorado Division of Water Resources 2017). HydroBase is a central water resources database for Colorado. Groundwater level data were obtained for all four aquifers of the DBAS for the time period from 1990 and 2016. Groundwater wells with fewer than five observations were not included in the analysis, as well as groundwater wells that appeared to be under pumping conditions and not at steady state. We identified 54 groundwater wells in the Laramie-Fox Hills Aquifer, 118 wells in the Arapahoe Aquifer, 60 wells in the Denver Aquifer, and 44 wells in the Dawson Aquifer that met these criteria. The frequency of groundwater level data collection is highly variable in each aquifer of the DBAS and contains various gaps in the spatial coverage. Many wells are sampled annually in April. However, there are many wells sampled multiple times per year. Each well record was also evaluated for outliers, possibly due to data measurement or data entry errors. A twelve point moving window was used to remove outliers that were greater that two standard deviations from the moving mean.

3.4.2 Spatio-Temporal Kriging

Spatio-temporal kriging was used to make estimates at unsampled locations. A number of advantages exist with spatio-temporal kriging over purely spatial approaches. Spatio-temporal kriging is able to utilize irregular datasets, avoid temporal binning, estimate data during times when data is not available, and leverage data in adjacent time periods to help infer the predictions in the period of interest (Zeng et al. 2014; Wang et al. 2015; Tadic et al. 2017). Implementation of spatio-temporal geostatistics was utilized in the R software (Team 2014) and gstat and space-time packages (Pebesma 2004; Pebesma 2012; Gräler et al. 2016).

We follow the spatio-temporal kriging approach outlined in Chapter 2 and additional information is found in the Supporting Information. Groundwater levels are assumed to follow a random function model that consisted of a trend and residual component that can be modeled separately. The global trend in groundwater levels for each aquifer is assumed constant in time and consists of only spatial components. A regression model is constructed for each aquifer and contains various combinations of predictors (e.g. latitude, longitude, land surface elevation, top boundary of aquifer, principal components 1 and 2 of the land surface and aquifer top elevations). The trend calculated from the regression model is subtracted from the groundwater level observations to yield the residual component. The resulting residual component is used to construct the space-time semivariogram from which a model is fit and used to infer the kriging predictions.

Semivariograms for each aquifer are modeled with a sum-metric model (Rouhani and Hall 1989; Dimitrakopoulos and Luo 1994; Myers 2004; Derakhshan and Leuangthong 2006). The model is represented as,

$$\gamma_{ST}(s,t) = \gamma_S(h) + \gamma_T(u) + \gamma_J\left(\sqrt{h^2 + (\kappa \cdot u)^2}\right) \tag{3-1}$$

where κ is the space-time anisotropy ratio that merges spatial distances with temporal distances, and $\gamma_S, \gamma_T, \gamma_J$, are the spatial, temporal, and joint semivariograms with separate nugget-effects. Various model types (e.g. Gaussian, Spherical, Exponential) were used for the spatial, temporal, and joint semivariograms for each aquifer (Tables A – 1 – 4)

Spatio-temporal ordinary kriging is performed on the residuals. The kriging approach is implemented in a similar fashion as spatial kriging (Mateu 2015). Because space-time datasets can be quite large, the kriging was limited to the nearest 1,000 neighbors and computations were parallelized to decrease computational times. The final spatio-temporal estimate of groundwater levels is given by adding the predicted kriging residuals with the trend component that was originally removed.

3.4.3 Volumetric Groundwater Storage Calculations

Groundwater level changes determined through the spatio-temporal kriging were converted to groundwater storage changes. Calculations are made using changes in water level and an appropriate storativity for unconfined and confined aquifers. The volume of water drained as heads decrease is calculated as (Fetter 2000),

$$\Delta GW_u(unconfined) = \Delta H \cdot A \cdot S_y \tag{3-2a}$$

$$\Delta GW_c(confined) = \Delta H \cdot A \cdot S_s b \tag{3-2b}$$

where ΔGW is the volumetric change in groundwater storage, ΔH is the groundwater level change, A is the area overlying the drained aquifer, S_y is the specific yield, S_s is the specific storage, and b is the aquifer thickness. The aquifers of the DBAS generally contain both unconfined and confined conditions. Each monthly potentiometric surface estimated through spatio-temporal kriging is compared to the upper boundary of the respective aquifer. Potentiometric surface elevations above the upper boundary identify confined conditions whereas elevations below the upper boundary indicate unconfined conditions. The upper boundary for each aquifer were derived from top contour estimates that were interpreted by the CDWR from geophysical logs (Colorado Division of Water Resources 2011).

In the current study, storativities were obtained through CDWR Denver Basin bedrock data (Colorado Division of Water Resources 2011). The available data varied for each aquifer, but highlight the spatial heterogeneities in each one (Figure A – 13 – 16). It is common practice to apply a representative storativity across an entire aquifer when evaluating storage changes. We also apply a single value when determining storage changes across the 27 year period. However, we utilize the mean storativities, median storativities, and combinations of available storativity data for each aquifer to better understand the range of possible storage calculations, depending on the representative storativity values chosen.

3.5 Results and Discussion

The estimated potentiometric surfaces show the distribution of hydraulic heads from 1990 to 2016 (Figures A-9 – 12). Standard deviation maps were also generated to evaluate the reliability and uncertainty in kriging predictions (Figures A-9 – 12). Each aquifer has areas where few data are available to constrain the interpolation. This leads to higher uncertainty in the predictions in data sparse areas. The accuracy of the interpolated potentiometric surfaces also assumes that the groundwater level data are representative of static conditions, the data have minimal measurement errors, the covariates used in the spatio-temporal kriging accurately remove the global trend in the data, the spatio-temporal semivariograms were appropriate, and that the land-surface and aquifer top elevations used in generating potentiometric elevations and identifying confined and unconfined areas are accurate. However, these approximations help understand how much groundwater has been depleted, how water levels have changed spatially, how depletion varies spatially and temporally, and uncertainty in the estimates.

3.5.1 Comparison of Overall Depletion in DBAS Aquifers

Each aquifer of the DBAS varies in magnitude of depletion since 1990 (Figure 3-2). The Dawson aquifer has been depleted by ~ 3 km^3 under average and median storativity conditions. However, calculated depletions range between ~1 to 7 km^3 depending on storativity data determined to be representative of the aquifer. The median scenario depletion represents 1.4% of the groundwater storage available in the DBAS assuming that 200 million AF is recoverable in total from all four aquifers. Depletions in the Denver Aquifer since 1990 equate to ~ 16 – 20 km^3 under median and average storativities. Under these scenarios, the total depletion represents 6.5% of the recoverable groundwater in the DBAS. The range of possible depletion for the Denver Aquifer is between ~ 5 – 55 km^3. The Arapahoe Aquifer is estimated to have depletions of ~ 16 km^3 under median and average storativities. This represents 6.5% of the recoverable groundwater in the DBAS. However, calculated depletions range from ~3 – 34 km^3 depending on the storativity data used in the calculation. The Laramie-Fox Hills Aquifer has been depleted by ~1.6 – 2.1 km^3 under median and average storativities. This represents about 0.6% of the recoverable groundwater in the DBAS. Calculated ranges of depletion for this aquifer are from ~0.4 – 5.2 km^3. Groundwater storage changes estimated under median storativity conditions represents a total of 36.6 km^3 (~29.7 million AF) for all four aquifers combined and accounts for 15% of the total recoverable groundwater in the DBAS. These estimates are variable, however, due to the uncertainty in representative aquifer storativities. The aquifer storativities for each

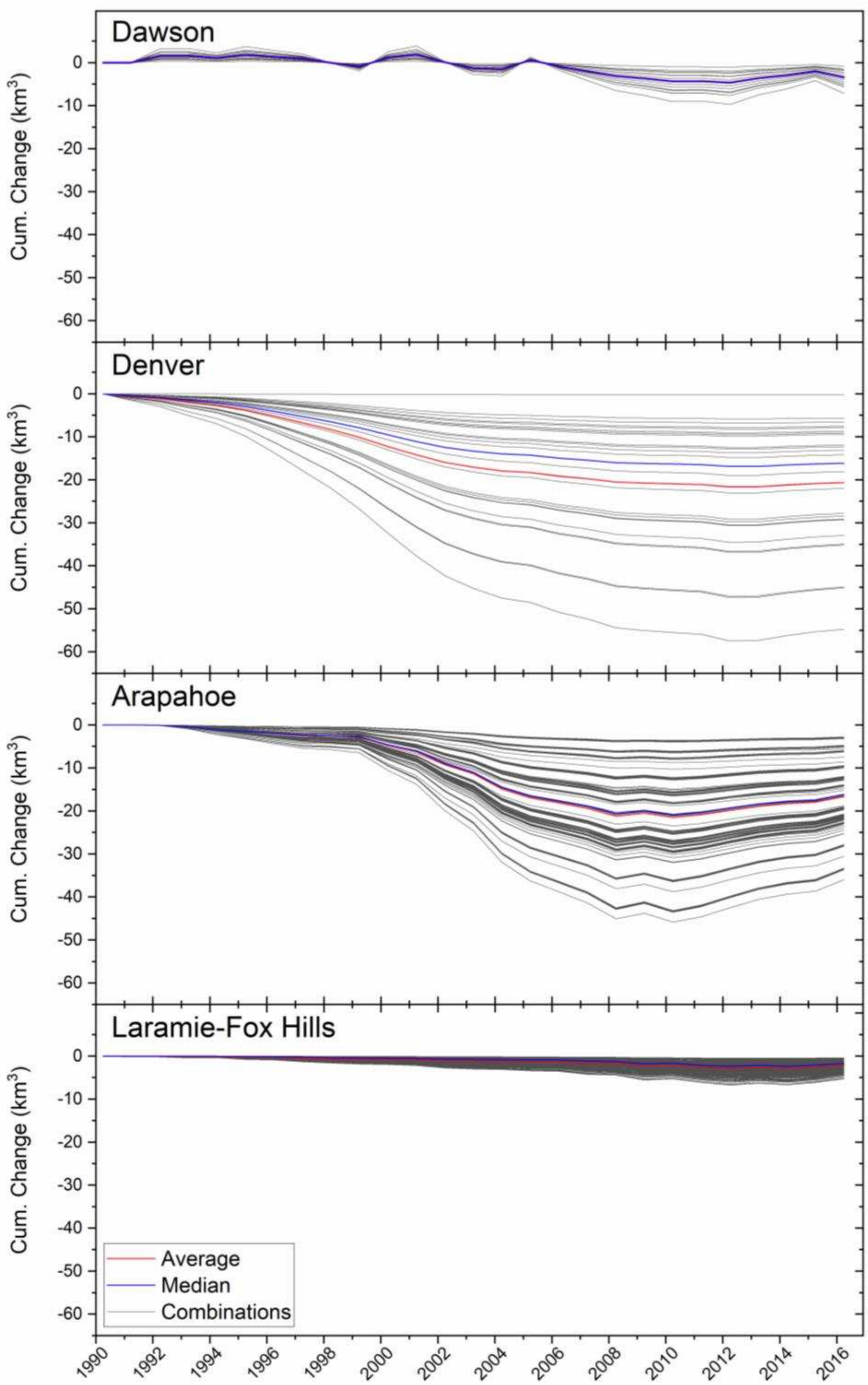

Figure 3-2 Cumulative changes in groundwater storage for the four aquifers of the Denver Basin Aquifer System since 1990. The gray lines represent possible storage change scenarios and are calculated using combinations of reported storativities for each respective aquifer. The red and blue lines are the storage changes calculated from the average and median reported storativities.

DBAS aquifer were not uniformly distributed and were sparse when considering the size of each aquifer (Figures A – 13 – 16). This prevented an area weighted approach in the groundwater storage calculations. In turn, the combinations of storativities highlight the impact and importance of selecting representative parameters.

3.5.2 Spatial Groundwater Changes

Groundwater level changes also vary spatially for each aquifer (Figure 3-3). If depletion were uniform, by 2016 the Dawson Aquifer would have average water table declines of 19 ft, the Denver Aquifer 60 ft, the Arapahoe Aquifer 112 ft, and the Laramie-Fox Hills 118 ft. In the Dawson Aquifer higher depletion was observed on the northern half of the aquifer and little to no depletion in the southern end. The areas near the city of Castle Rock had groundwater level declines of up to 275 ft since 1990. This area also represents the greatest groundwater depletion for the aquifer (Figure 3-4). Areas near the city of Monument and southwest of Castle Rock had estimated increases in groundwater storage up to 0.1 km^3. In the Denver Aquifer, the highest water level changes are localized to the west-central area of the aquifer under the cities of Castle Rock and Parker. Since 1990, water tables have declined by up to 400 ft. Similar to the Dawson Aquifer, most of the decreasing water tables are spatially on the northern half of the aquifer. Volumetric depletions follow similar patterns and range between 0.02 – 0.1 km^3 for the majority of the aquifer. The highest depletions were located under the city of Parker and to the northwest of Castle Rock. As previously reported in Chapter 2, the Arapahoe Aquifer is estimated to have water table declines by up to 600 ft in the west-central areas of the aquifer. The highest depletion extends from the Highlands Ranch area, between Denver and Parker, down to the city of Monument. From this central area, water table declines radially decrease outwards to the outer edge of the aquifer. Areas to the northwest of Denver and south of Monument towards Colorado

Springs do show estimated increases in groundwater levels since 1990. Volumetric storage depletions for much of the aquifer are less than 0.04 km^3. The highest depletion occurs west of Castle Rock and has reached up to 0.4 km^3. The Laramie-Fox Hills Aquifer is estimated to have experienced water table declines by up to 600 ft in the areas underlying Parker and the town of Bennett. With the exception for the town of Bennett, this spatial pattern and total water table changes are similar to those of the Arapahoe Aquifer. Much of the aquifer shows decreased water table levels since 1990, but to the northeast of Denver, levels have increased. The majority of the aquifer has depletions less than 0.01 km^3. The highest depletion is predicted along the southeastern edges of the aquifer and reaches 0.25 km^3.

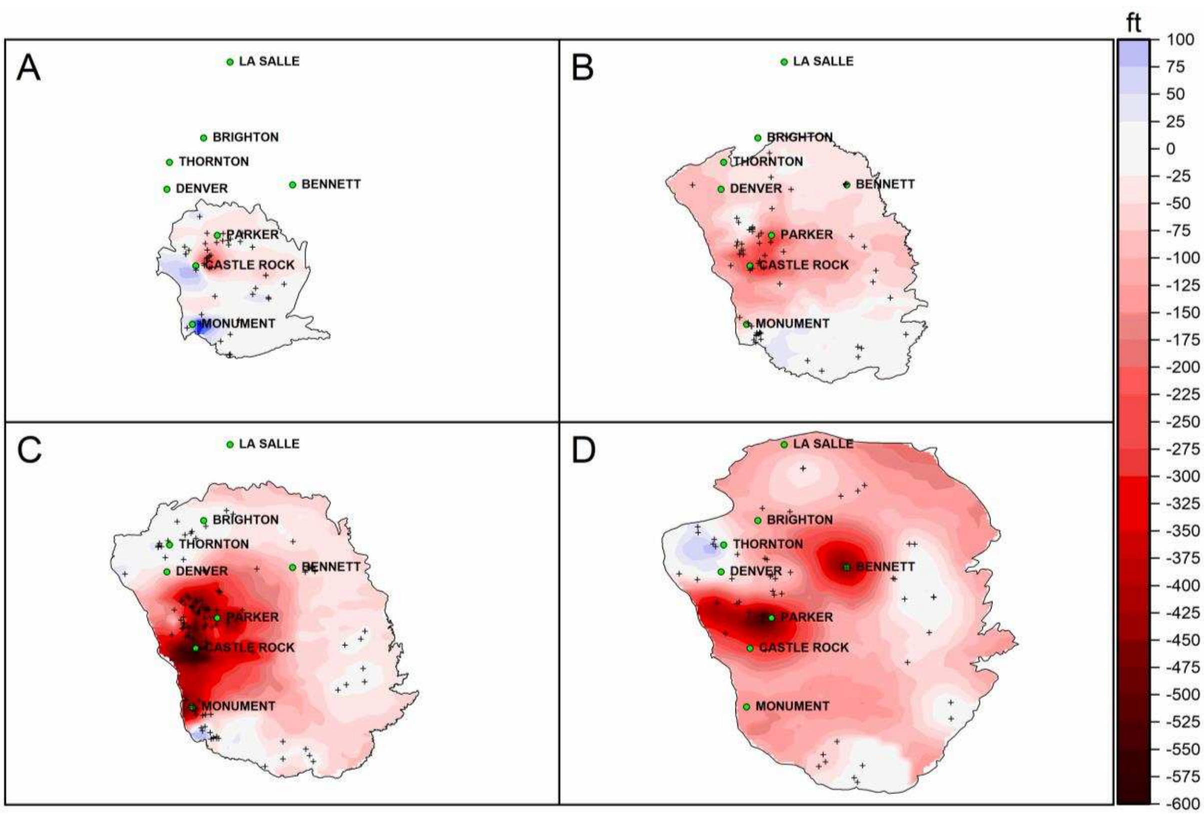

Figure 3-3 Changes in groundwater levels in the A) Dawson, B) Denver, C) Arapahoe, and D) Laramie-Fox Hills Aquifers from 1990 – 2016. The black '+' show the locations of groundwater wells used in the analysis.

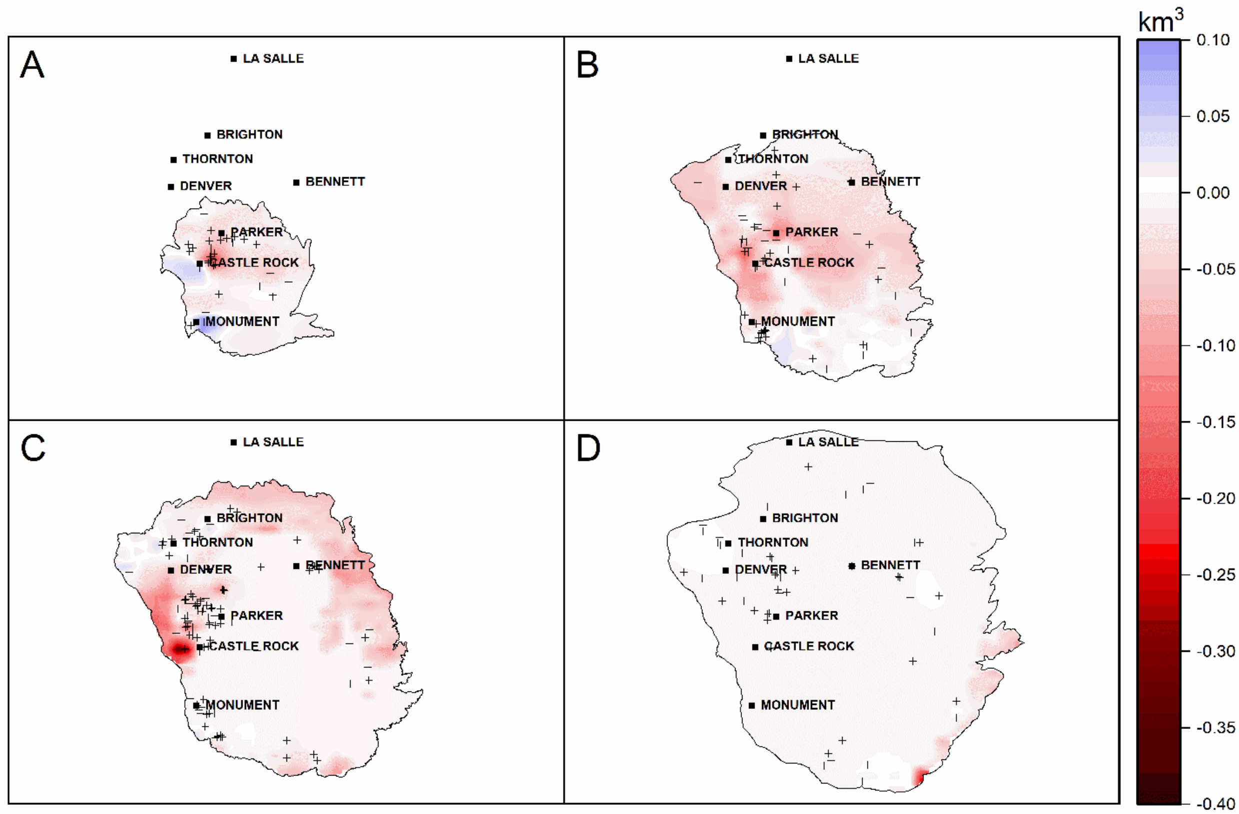

Figure 3-4 Changes in groundwater storage in km^3 for the A) Dawson, B) Denver, C) Arapahoe, and D) Laramie-Fox Hills Aquifers from 1990 – 2016. The black '+' show the locations of groundwater wells used in the analysis.

Spatial depletion patterns for each of the aquifers might not completely match the water table change patterns and extremes. The Denver, Arapahoe, and Laramie-Fox Hills Aquifers are predominantly under confined conditions, but do become unconfined along the edges where each aquifer crops out. In the confining portions, large decreases in groundwater levels are due to a reduction in aquifer pressures and do not always translate to large volumes of water. Water is released by compression of the aquifer and expansion of water. However, when conditions become unconfined, the storage coefficient is represented by the specific yield and can be orders of magnitude larger than the specific storage. The water released is due to the drainable porosity by gravity drainage and yields larger volumes. The western edge of the DBAS dips more steeply

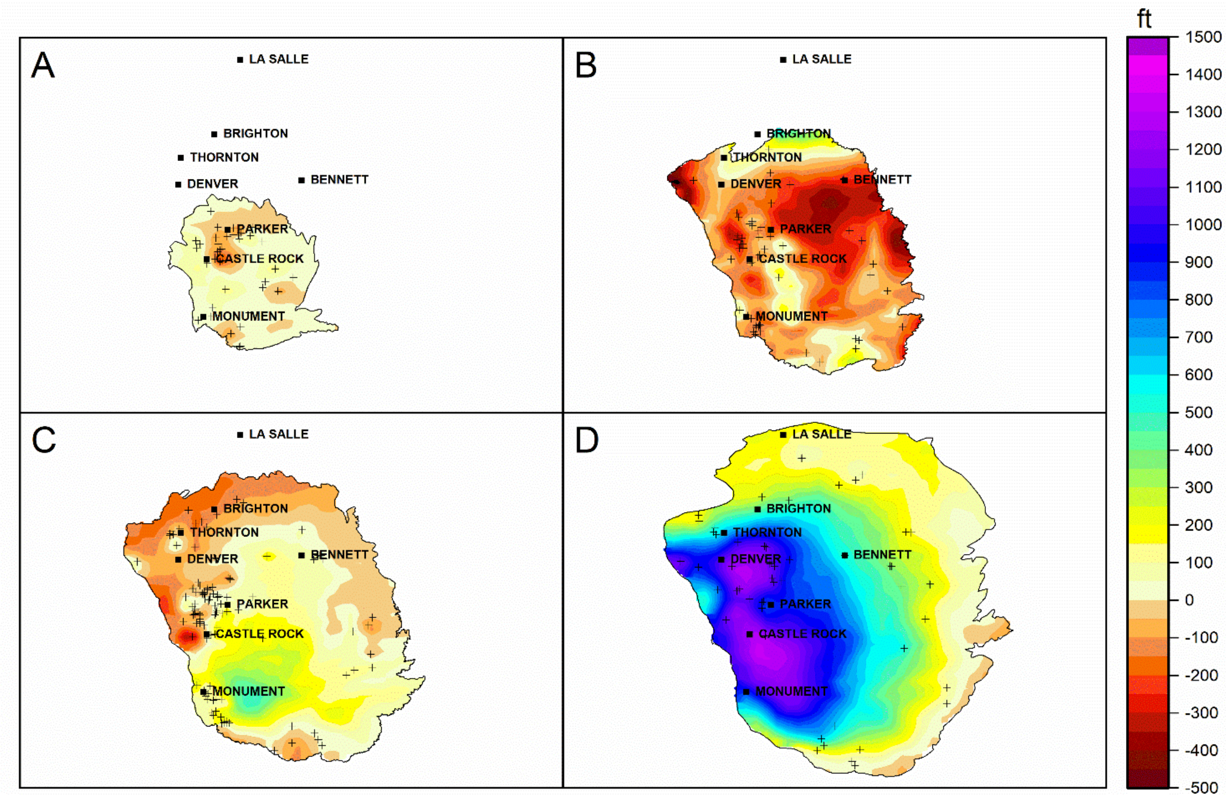

Figure 3-5 Height of the potentiometric surface in relation to the top designation of each aquifer unit for the A) Dawson, B) Denver, C) Arapahoe, and D) Laramie-Fox Hills by the end of 2016. Positive values indicate water levels above the top of each aquifer and confining conditions. Negative values indicate water levels below the top designation of each aquifer and are considered unconfined conditions.

than the central and eastern part due to the asymmetrical, bowl shape of system. Therefore, large changes in water levels can cause the potentiometric surface to drop below the upper confining unit more easily. Looking at differences between the potentiometric surface elevations and the top unit of each aquifer help identify the areas experiencing higher depletion. Areas with potentiometric surfaces below the aquifer top elevations are subject unconfined conditions and therefore larger calculated depletion volumes for the same water level change. In the Denver Aquifer, the higher depletion areas are also the areas where the water level have changed the largest and caused the confined portion to become unconfined (Figure 3-5). In the Arapahoe

Aquifer, the areas with the highest volumetric depletion are to the west of Caste Rock and show where the confined aquifer has become unconfined. In 2004, Raynolds (2004) predicted that over the next several decades some areas of the Arapahoe Aquifer would transfer from confined to unconfined conditions. In the Laramie-Fox Hills aquifer, the potentiometric surface is between 400 and 1500 ft above the upper confining unit. The Dawson Aquifer is an entirely unconfined aquifer even though estimated water tables are above the designated aquifer top.

3.5.3 Temporal Groundwater Changes

The rates of change in groundwater levels vary across each aquifer and have changed over the past 27 years. Linear decadal rates are shown in Figure 3-6 for three time periods, 1990 to 1999, 2000 to 2009, and 2010 to 2016. The prior two decades experienced higher depletion rates in the west-central areas underling Monument and spanning to Denver. The Dawson Aquifer has generally experienced increasing and decreasing water levels of up to 5 ft per year. During the 2000s decade, levels near castle rocks experienced between 5 and 20 ft per year declines. Since 2010, the majority of the aquifer shows increasing water levels. The Dawson Aquifer displays increasing and decreasing storage between 1990 and 2005. During this period depletion does appear to follow drought and precipitation patterns as can be seen in 2001 which began as moderate drought conditions and led to extreme and exceptional drought until 2004. The Dawson Aquifer is the uppermost and an unconfined aquifer that would be influenced by variable precipitation and recharge. The Denver Aquifer experienced higher rates of depletion across a larger area in the 1990s than in the 2000s. The highest depletion was up to 25 ft per year and underlies areas of rapid population growth during the time. The southern areas of the aquifer show up to 10 ft per year increases in hydraulic head, but slow down during the 2000s. The Arapahoe Aquifer shows much higher rates of depletion in the west-central areas of the aquifer

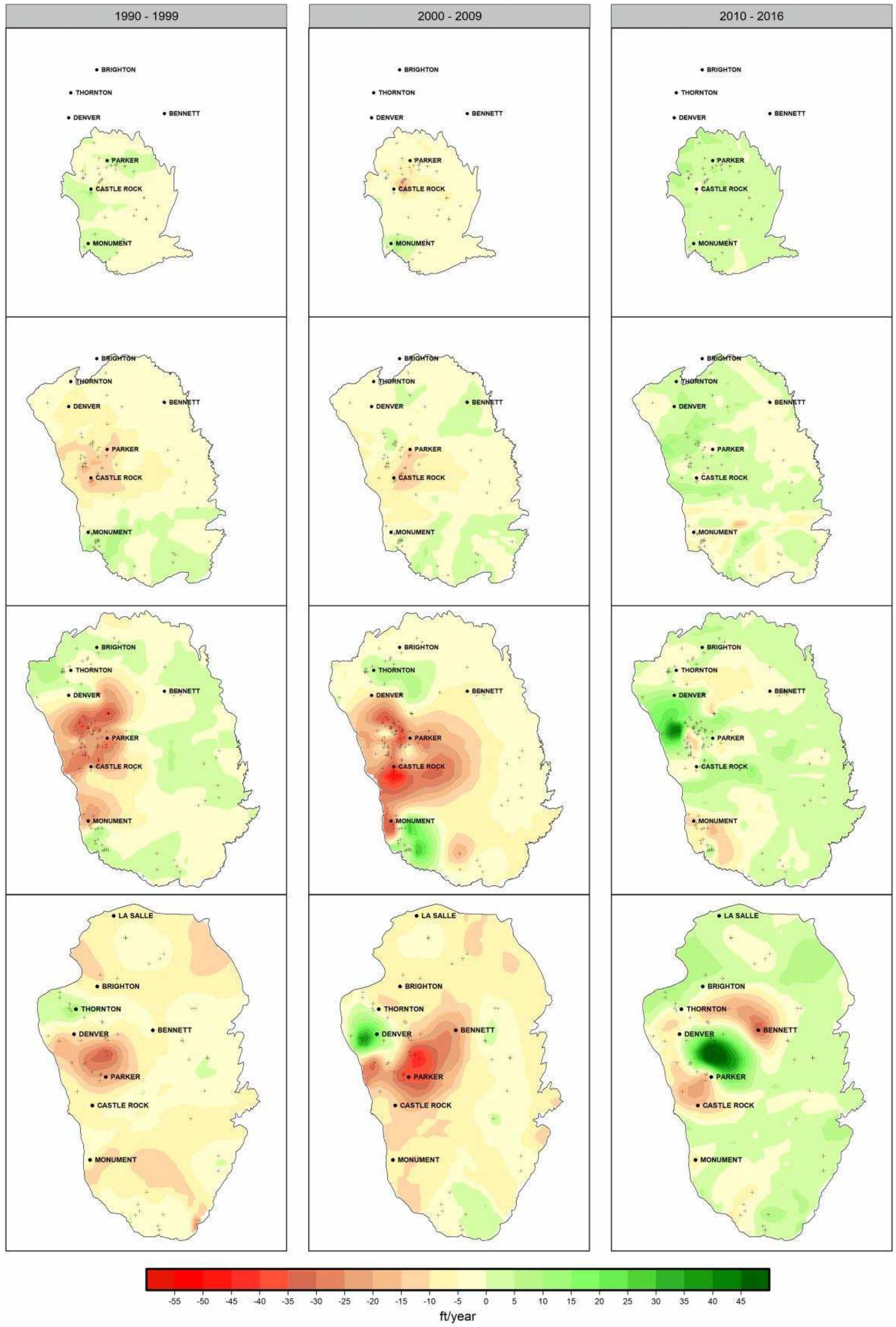

Figure 3-6 Linear decadal rates of groundwater level changes across each of the four Denver Basin Aquifers. The aquifers are ordered from shallowest to deepest (top to bottom) and from earlier decades to more recent (left to right). The black '+' show the locations of groundwater wells used in generating potentiometric surfaces for the analysis.

ranging from 15 to 40 ft per year. During the 2000s, the area experiencing high groundwater level declining grew in size and experienced decreases in hydraulic head of up to 55 ft per year near Castle Rock. Since 2010, rates of water level change are drastically different and show much of the aquifer increasing in water levels. It is estimated that Castle Rock now experiences about 5 ft per year declines and matches to what the city reports as average annual declines (Town of Castle Rock Colorado 2018). The Laramie-Fox Hills Aquifer started to see declines between 10 and 35 ft per year under the same areas that the other aquifers were experiencing high declines. During the 2000s, rates in this area reached up to 50 ft per year and shifted northeast by the town of Bennett. Since 2010, water level rates appear to have reversed and many areas are experiencing increasing levels, even as high as 45 ft per year north of Parker. Assessment of these trends highlights areas were management is critical.

3.5.4 Drivers of Groundwater Change

3.5.4.1 Population

Groundwater depletion for aquifers of the DBAS is likely a result of a combination of many factors including population growth, reduced precipitation and recharge, and groundwater well pumping, each adding additional stress. Population growth appears to have a large influence on the demand for water as is evident in the spatial patterns for water level change. Municipalities such as Parker and Castle Rock have long been dependent on groundwater resources of the DBAS and also overlie areas that have experienced up to 600 ft of water level declines. The population of these two municipalities grew steadily from 1990 to 1998. However, after 1998 the rate of population growth increased for the next decade before slowing down. A plot of population growth for these to municipalities (Figure 3-7) shows two pivot points that

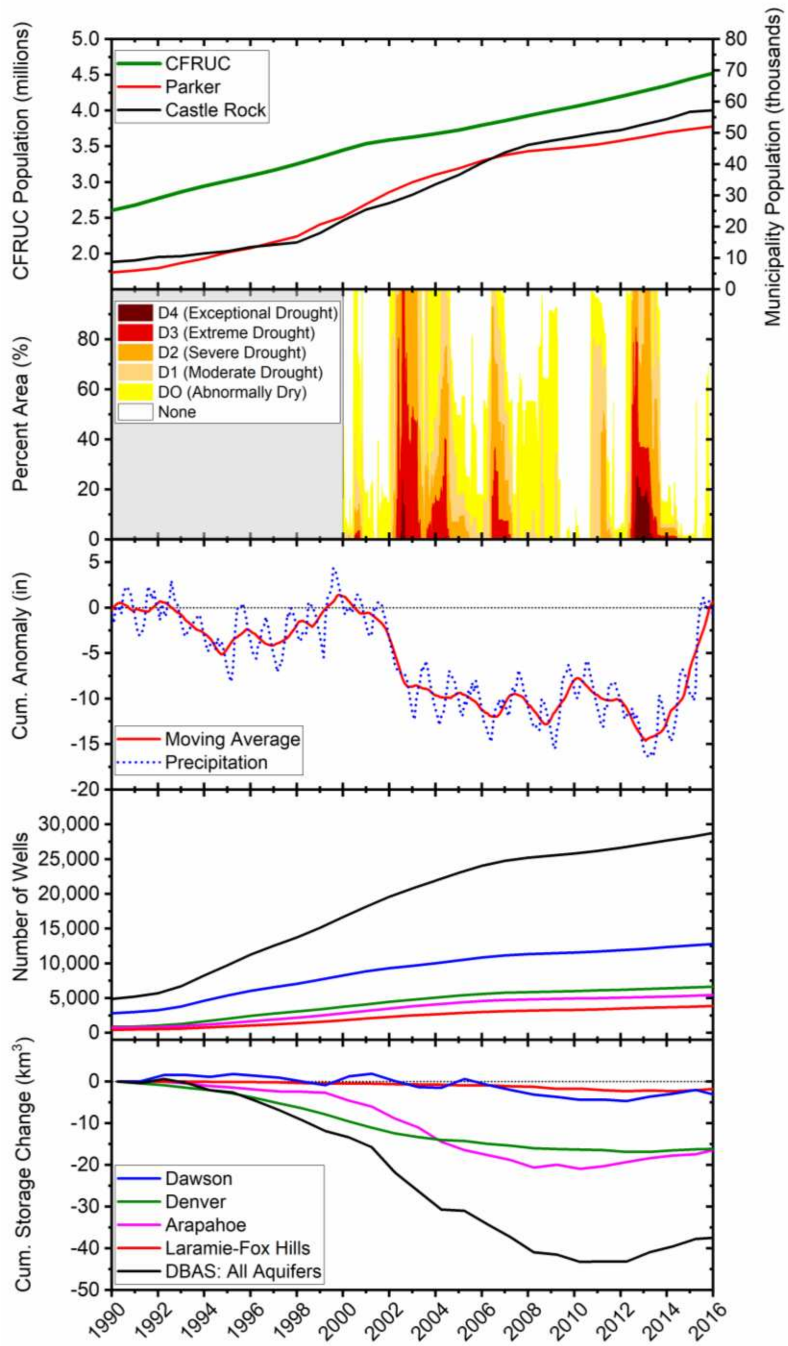

Figure 3-7 Comparison of possible influences on groundwater storage changes in the DBAS from 1990 – 2016. The top panel shows population growth for the Colorado Front Range Urban Corridor (CFRUC) and two municipalities that have observed some of the largest water level declines over the 27 year period. The second panel from the top shows the percent area of the basin experiencing variable drought conditions since 2000. The middle panel shows precipitation patterns represented as a cumulative anomaly and departure from the long term mean. The next panel shows the number of groundwater wells through time in each aquifer of the DBAS. The bottom panel shows groundwater storage changes for each aquifer under median storativity conditions.

also closely resemble the visual trends in groundwater storage depletion. . The percent population change in these cities helps explain some of the decadal rates of change and the large decreases in groundwater levels (Table 1).

Table 3-1 Percent population change for select cities across the DBAS during three time periods. Data sources: (U.S. Census Bureau 2005a; Colorado Department of Local Afairs 2013; U.S. Census Bureau 2017a)

	1990 - 2000	2000 - 2010	2010 - 2016
Denver	18.6%	8.2%	15.5%
Littleton	19.5%	3.5%	11.0%
Greenwood Village	40.2%	26.2%	13.1%
Parker	314.2%	92.3%	13.0%
Castle Rock	135.2%	138.5%	19.6%
Monument	116.4%	180.6%	18.6%

3.5.4.2 Variable Precipitation

The 30-yr normal annual precipitation from 1981 – 2010 across the DBAS varies spatially and ranges from 12 to 20 inches (PRISM Climate Group 2016). Evaluating the departure from the long term mean as an anomaly can help understand when precipitation was above or below average. An alternative of this is to evaluate the impact of precipitation declines is through the cumulative anomaly, which can help provide a measure of how much precipitation the areas has missed out on over the past 27 years. This is shown in the third panel of Figure 3-7. It is possible that changes in precipitation patterns, which would influence the amount of surface recharge, have contributed to the overall depleting groundwater storage in the DBAS. Prior to 2000, variability in precipitation led to decreased precipitation between 1992 and 1995, but eventual recovery of precipitation by 2000. The most impactful turning point in annual precipitation occurred after 2000. The cumulative anomaly continued to become more negative and signify reduced precipitation compared to the 1990s. By 2014, precipitation recovered to average conditions. Coincidently, the large period of below average annual precipitation also tracks with the larger groundwater depletions of the Denver and Arapahoe Aquifers suggesting some influence to the system by reduced recharge. Comparison of the cumulative anomaly to the U.S. Drought Monitor also show this time period experiencing two exceptional drought periods (Figure 3-7).

3.5.4.3 Increased Groundwater Development

The amount of water pumped from the DBAS is not well monitored. Following the methods of Paschke et al. (2011), the CDWR well-permit database was used to determine the number of wells constructed from 1990 to 2016 and as a proxy to pumping (Colorado Division of Water Resources 2018). The database was filtered to only include groundwater wells with a well construction date and a specific aquifer distinction within the DBAS. Wells screened across multiple aquifers were excluded. Approximately 13,050 wells were completed in the Dawson Aquifer, 6,751 wells in the Denver Aquifer, 5,535 wells in the Arapahoe Aquifer, and 3,948 wells in the Laramie-Fox Hills Aquifer. In total, approximately 29,284 wells existed in the DBAS by 2016 given our criteria. The wells include domestic, monitoring, municipal, stock, commercial, and irrigation uses. Figure 3-7 displays the cumulative number of wells constructed over time. Three distinct periods of well construction appear and also have some resemblance to the cumulative storage change patterns. The fastest rate of wells constructed in the DBAS was between 1993 and 2007 and increased by over 18,000 wells. During this period, the Denver and Arapahoe Aquifers also had the largest storage declines (Figure 3-7) suggesting that as the number of wells increased, so did the amount of pumping. However, after 2007 the number of new wells constructed each year decreased, possibly due reduced number of permits being issued by the State Engineer's Office, alternative sources of water being developed (e.g. Rueter-Hess Reservoir), or conservation measures.

3.5.5 Concerns for Land Subsidence

Bedrock formations of the DBAS have elastic properties and are compressible (Robson 1987). Groundwater withdrawals leading to large groundwater level declines could potentially stress the system, leading to land subsidence. Little attention however has been given to potential subsidence across the DBAS as only several studies make any mention. Galloway et al. (1999) report subsidence in the Denver area as a result of groundwater abstractions, but don't provide any additional details. Using rock-compressibility data for other areas, Robson (1987) estimated that in the Arapahoe Aquifer, land subsidence due to elastic compression could range from 0.2 to 2 inches per 100 ft of water level decline. In the Laramie-Fox Hills Aquifer it is estimated to range from 0.07 to 0.7 inches per 100 ft of water level declines. Using these estimates, water levels declines since 1990 in the Arapahoe Aquifer below the Denver metro reach around 75 ft, which would indicate a potential of 0.15 to 1.2 inches land subsidence could have occurred due to water declines in the Arapahoe Aquifer. Below the city of Castle Rock, up to 600 ft water level declines since 1990 could have led to 1.5 to 12 inches land subsidence. The subsidence estimates could potentially be larger, depending on water level declines in the three other aquifers of the DBAS, with a contribution from each aquifer. Robson (1987) estimated that water level declines in the Arapahoe Aquifer from 1883 to 1960 led to 0.8 to 8 inches of subsidence across the Denver metro. Because the DBAS dips to the east, it is estimated that regional compaction would begin on the western side of the aquifer (Morgan 2007).

3.5.6 Future Planning

Long-term depletion can be driven by unsustainable withdrawals and the short-term variability of precipitation and recharge (Konikow 2015). To combat withdrawals and store supplies for future use, Aquifer Storage and Recovery (ASR) is gaining speed in the DBAS. Centennial Water and Sanitation District (CWSD) has successfully implemented ASR in the DBAS since the mid-1990s (Centennial Water and Sanitation District 2017). CWSD continues to expand its infrastructure and more than 4.5 billion gallons of water have been injected into over 25 wells permitted in the Denver, Arapahoe, and Laramie-Fox Hills Aquifers. Areas near Highlands Ranch show increasing hydraulic heads in the Arapahoe and Laramie-Fox Hills Aquifer (Figure 3-7) which is also where CWSD operates its ASR field. It's possible that the observed increasing hydraulic heads are due to ASR and reduced groundwater pumping. Other water providers in the areas such as Denver Water, South Metro Water Supply Authority, East Cherry Creek Valley Water and Sanitation District, the Town of Castle Rock, Rangeview Metropolitan District, and Pinery Water and Wastewater District are all exploring ASR to secure future water supplies (Denver Water 2018; Hecox 2015; Colorado Water Insitiute 2017).

Many other projects are being are also being implemented to prepare for future demands. A regional project known as WISE (Water Infrastructure and Supply Efficiency) recently became operational in late 2017 and is a partnership meant to share existing water resources and secure a renewable water supply to South Metro communities (South Metro Water Supply Authority 2017). In this project, water used by Denver and Aurora is recaptured, and when available, is able to be shared with participating partners (e.g. Highlands Ranch, Parker, and Castle Rock) to help reduce reliance on DBAS groundwater.

Until 2012, Parker solely relied on groundwater to provide drinking water to its residents. In 2012, the Rueter-Hess Reservoir was completed and expanded to 72,000 acre feet. Combined groundwater and surface water resources are now used (Parker Water and Sanitation District 2016). Castle Rock gets most of its water from the DBAS with approximately 85% being groundwater and only 15% coming from renewable supplies along Plum Creek. The Plum Creek Water Purification facility has been treating and delivering renewable water since 2013. Castle Rock is making plans to convert its water supplies to be more than 75% renewable surface water by 2050. The town has also purchased water storage in Rueter-Hess Reservoir and is participating in a project to capture water supplies in Chatfield Reservoir. Aquifer storage and recovery also is a component of Castle Rock's long-term storage plan (Town of Castle Rock Colorado 2018). Over the years, about 80% of Highlands Ranch water supply has been surface water derived, but since 2011 it has become 99%. Transitioning to more renewable water supplies will be important to help extend the life of the DBAS and reduce depletion rates.

3.5.7 Comparison to USGS Denver Basin Model

The USGS Denver Basin model by Paschke et al. (2011) provides a good means to compare some of our results, although some time periods don't overlap. Calibration of the USGS Denver Basin Model did not utilize data from the CDWR due to the limited number of annual water level measurements that prevented adequate spatial coverage to generate potentiometric maps. However simulated heads were compared to select CDWR wells. The USGS model provided variable results and had areas of better agreement than others to CDWR data (Paschke et al. 2011). Since this spatio-temporal kriging analysis of groundwater levels in the DBAS solely relied on CDWR data, the results too are variable to those of the USGS model. The spatio-temporal kriging approach applied in this study leveraged data in adjacent time periods to help infer the kriging predictions at the time of interest. This increased the available data used to generate potentiometric surface maps and possibly overcomes some of the limitation of irregular and sparse data. Comparing the USGS 2003 simulated change in hydraulic head for the Dawson Aquifer to those of this study (Figure 3-3) show some similarities and differences. Both show decreasing hydraulic head around Castle Rock and Parker that propagates outward and both show increases levels along the northwestern edge that is attributed to urban recharge. However, one major difference is that the spatio-temporal kriging results show increasing hydraulic heads underlying Monument while the USGS Denver Basin model shows decreasing levels. In the Denver Aquifer, the USGS model simulated decreasing hydraulic heads in the west-central portions under Castle Rock and Parker reaching 250 ft declines since 1880. This study also shows larger decreases in this area up to 350ft, though this study is 13 years longer. Differences in patterns for head change are predominately in the northern half of the aquifer. This study shows overall decreasing levels (Figure 3-3), while the USGS model shows the area underlying Denver to be increasing in water level. Storage change patterns in the Arapahoe Aquifer are similar between this study and the USGS Denver Basin Model. The areas underlying Monument, Castle Rock, Parker, and Highlands Ranch are all experiencing the largest declining storage changes. In the UGSG study, by 2003 hydraulic heads had decreased by around 400 ft in this area. Since 1990, this study estimates up to 600 ft decreases in hydraulic head. Both this study and the USGS study show increasing heads in the northwestern area as well as south of Monument. However, this study does not show increasing levels along the eastern edges due to a lack of data to infer the kriging predictions. Agreement between the USGS model and this study for the Laramie-Fox Hills Aquifer is highest in the areas underlying Castle Rock, Parker, and Denver. The USGS model shows some depletion around Thornton, while this analysis shows hydraulic head increases by up to 80 ft. Hydraulic head near the town of Bennett show decreases up to 600 ft since 1990 and a cone of depression that radially propagates outward. The USGS model does show up to 200 ft decreases in hydraulic head in the area since 1880, but does not appear to have a similar cone of depression. One possibility for the differences could be the lack of nearby well data to Bennett to help with the kriging predictions. Overall, comparison to the USGS Denver Basin model study show some similarities and agreement to hydraulic head changes, but also some noticeable differences in data sparse areas.

Simulated cumulative change in groundwater storage for the DBAS by the USGS model shows considerable differences than this study. By 2003, the total cumulative depletion was predicted to be 1.00 km^3 and was expected to reach 1.30 km^3 by 2008 (Paschke et al. 2011; Konikow 2013). In this study, applying median storativities calculated from CDWR data (Figures S13 – 16) yielded 24 km^3 depletion in the bedrock aquifers by 2003. By the end of 2016, total cumulative depletion is predicted to have reached 36 km^3 (Figure7). However, as noticed in Figure 3-2, the calculated depletion is heavily dependent on the storativity values chosen. This study highlights the uncertainty that results from either applying one representative storage coefficient across the entire aquifer or the uncertainty due to aquifer property heterogeneities which is common to many groundwater quantity studies. It is plausible that selecting different storativities could lead to much lower cumulative depletion predictions. Paschke et al. (2011) multiplied specific yield values by the sild-sand fraction of model cells to get an adjusted specific yield. Future work could include applying a similar procedure to area weight storativities to better predict volumetric storage changes in the DBAS.

As in all cases, future estimates would benefit from additional data and monitoring wells across each aquifer. Filling in the spatial gaps could help improve potentiometric surface predictions and provide additional space-time points to leverage. While this study only evaluated CDWR data that resides in Colorado's HydroBase, additional groundwater level data from other sources would greatly improve future predictions. This study also points to the importance of representative aquifer storage parameters when calculating volumetric storage changes. Additional work is needed to compile more storativity data and a better distribution across each aquifer.

3.6 Conclusions

The DBAS is a critical resource for many users along the CFRUC, but faces long-term sustainability challenges due to increasing demands for water in the region. Spatio-temporal kriging techniques were used to estimate potentiometric surfaces and evaluate groundwater depletion in the DBAS. Spatio-temporal kriging provides an effective means for interpreting sparse and irregular groundwater data. There is an ongoing need to continue to analyze and interpret DBAS groundwater data to better understand the resources available. This study complements the USGS Denver Basin modeling study and provides a different approach to assessing depletion and status of the aquifers by integration space-time approaches.

A wide range of calculated storage depletions are possible depending on the storativities chosen to be representative of the aquifer. The Denver and Arapahoe Aquifer exhibited the largest range of storage depletions with varying storativities. Using the median storativities from CDWR data, it is estimated that since 1990, the DBAS has been depleted by 36.6 km^3 (~29.7 million AF) and accounts for ~15% of the total estimated recoverable groundwater in the DBAS. However, the results show that it is possible for storage depletion to be much less or much more, depending on the storativities deemed representative. Groundwater level declines since 1990 are more severe in the west-central areas of the Denver, Arapahoe, and Laramie-Fox Hills Aquifers.

Spatial patterns of water level changes and depletion are likely a combination of population growth in municipalities that rely heavily on groundwater, variable precipitation and drought influencing surficial recharge, and an increase in the number of wells pumping groundwater.

Large variability in cumulative depletions should be considered in management decisions. Volumetric groundwater assessments can be difficult and often the uncertainty in estimates is unexplored. Because the storativities have a large influence on this estimated amount of depletion, so too does their influence on estimated amount of recoverable water in the DBAS. Future work would benefit from refining the spatial distribution of storativities in each aquifer and incorporating additional data or monitoring networks into the analysis.

CHAPTER 4

MONITORING GROUNDWATER STORAGE CHANGES USING GRACE AT THE AQUIFER SCALE: ACCURACY AND VALIDITY IN THE DENVER BASIN

4.1 Abstract

Few studies have evaluated the validity of using GRACE over smaller basins which are less than the recommended footprint of ~200,000 km^2, despite the interest in using remote sensing as another tool to asses groundwater depletion and storage trends in areas lacking monitoring data. Given the lack of adequate temporal and spatial groundwater data in many water-limited systems and the growing need for water resources, it is important to determine whether GRACE can still provide useful information at smaller scales typical of many aquifer systems to help inform and understand system dynamics. Temporal data sparseness is a significant problem common in groundwater well data analysis (see Chapter 2), and spatial water-well data scarcity is also common. Thus, it is logical that GRACE data, which is provided monthly and averaged over aquifer-scale pixels, might improve storage analyses compared to using water-well data alone, even at smaller scales. The strengths of GRACE data may help minimize the impact of the weaknesses of the water-well data analysis, and vice versa. The purpose of this study is to assess the accuracy and validity of applying GRACE at the aquifer-scale across the Denver Basin Aquifer System (DBAS) to understand groundwater storage changes and provide another tool for assessing groundwater storage. GRACE derived groundwater storage changes are compared to ground-based groundwater-well data for the 4

aquifers of the DBAS and the overlying South Platte Alluvial Aquifer. Under the current processing, the ground-based results do not show good agreement to GRACE derived groundwater anomalies. This work highlights some uncertainties in appropriately processing GRACE for groundwater applications and further work is still need to assess the usefulness across the DBAS.

4.2 Introduction

Remote sensing has become a powerful tool for assessing environmental processes and is an important source of data to improve hydrologic modeling. Groundwater studies have benefited from the addition of the Gravity Recovery and Climate Experiment (GRACE) satellites as one of the latest remote sensing tools. Estimates of gravitational changes through GRACE can be used to infer time variable changes in mass which are attributed to the redistribution of water both above and below the Earth's surface. GRACE has been used to look at droughts (Houborg et al. 2012; Long et al. 2013; Leblanc et al. 2009; Thomas et al. 2014), flooding and surface waters (Swenson and Wahr 2009; Lee et al. 2011; Wang et al. 2011; Zmijewski and Becker 2014; Reager et al. 2014), groundwater depletion (Rodell et al. 2009; Feng et al. 2013; Scanlon et al. 2012b; Voss et al. 2013), to infer aquifer storage parameters (Sun et al. 2010), and used to enhance groundwater models and global hydrologic models (Li et al. 2012; Güntner 2008).

GRACE has provided useful and unique hydrological information globally and at regional scales. However, there is much debate on the usefulness and applicability of GRACE at scales smaller than 200,000 km^2, which is a common scale for aquifers. Alley and Konikow (2015) recently called to attention some of the limitations and misperceptions of GRACE due to the popularity and gaining interest of utilizing GRACE for hydrologic investigations. Some of the limitations identified include lower resolution that limits use at smaller scales, spatial leakage and measurement errors that can be large, and the inability to show direction of water flow or sources of depletion because changes are represented in one-dimension (Longuevergne et al. 2010; Alley and Konikow 2015). However, despite these limitations, there is still much interest in using GRACE due to lack of adequate ground-based data in many places. Indeed, the relatively sparse spatial and temporal data availability can negatively impact aquifer storage and sustainability analyses, even for highly utilized aquifers with the typical availability and quantity of data (see Chapters 2 and 3).

Several studies have evaluated the use of GRACE below the recommended 200,000 km^2 footprint. In the Central Valley, California, high amounts of pumping for irrigation have allowed storage changes to be detected by GRACE even though the area of interest is ~52,000 km^2 (Scanlon et al. 2012b). In regions such as Mali (~54,971 km^2) with sparse hydrogeological data, GRACE was used to estimate annual recharge (Henry et al. 2011). In three locations across the U.S., an artificial neural network (ANN) was used to statistically downscale GRACE to use as a predictor of water levels for various wells within a single grid-pixel (Sun 2013); these authors determined that GRACE can be used to infer changes in water levels when an ANN was applied. Katpatal et al. (2017) characterize the performance of four GRACE pixels across aquifer systems with varying lithology in India. Hachborn et al. (2017) evaluate the sensitivity of GRACE derived groundwater to represent ground-based data across a study area of 45,000 km^2 in southern Ontario using four pixels. It is possible that GRACE can serve as another tool in the "toolbox" and help provide a big picture overview in data sparse areas. The strength of GRACE is realized when combined with local data to assess system dynamics (Richey et al. 2016; Wouters et al. 2014).

In many places, there is a lack of in situ groundwater well data with adequate temporal and spatial coverage to validate GRACE observations (Chen et al. 2016; Sun et al. 2010; Wouters et al. 2014; Sun 2013). In situ observations help understand the dynamics of groundwater systems and provide a means by which to compare the accuracy and limitations GRACE. However, proper comparisons can be limited by uncertainties in specific yield, specific storage, and representative groundwater wells. Given these uncertainties, many studies still have shown good agreement between GRACE derived groundwater storage changes and groundwater well observations (Rodell et al. 2007; Chen et al. 2016; Swenson et al. 2006; Castle et al. 2014; Scanlon et al. 2012b; Huang et al. 2016). Temporal data sparseness is a significant problem common in groundwater well data analysis (see Chapters 2 and 3), and spatial water-well data scarcity is also common. Thus, it is logical that GRACE data, which is available at the monthly scale and averaged over aquifer-scale pixels, might improve storage analyses compared to using water-well data alone, even at smaller scales. The strengths of GRACE data may help minimize the impact of the weaknesses of the water-well data analysis, and vice versa. The primary objective of this study is to assess the viability of using GRACE to evaluate groundwater storage changes in the multilayered Denver Basin Aquifer System.

4.3 Study Area

The Denver Basin Aquifer System (DBAS) is an important water resource for many users along the Colorado Front Range (Paschke et al. 2011). It is an asymmetrical, bowl shaped basin that extends from Greely to Colorado Springs and from the Front Range to Limon (Figure 4-1). It consists of four layered aquifers and from youngest to oldest are the Dawson, Denver, Arapahoe, and Laramie-Fox Hills. The aquifers are unconfined along the outer edges where they meet the surface, and confined in the more central areas. The DBAS is recognized as a nonrenewable resource because of the slow rate of recharge (Graham and Vanslyke 2004; Banta and Paschke 2012). Additional details on the study area are found in Chapter 3.

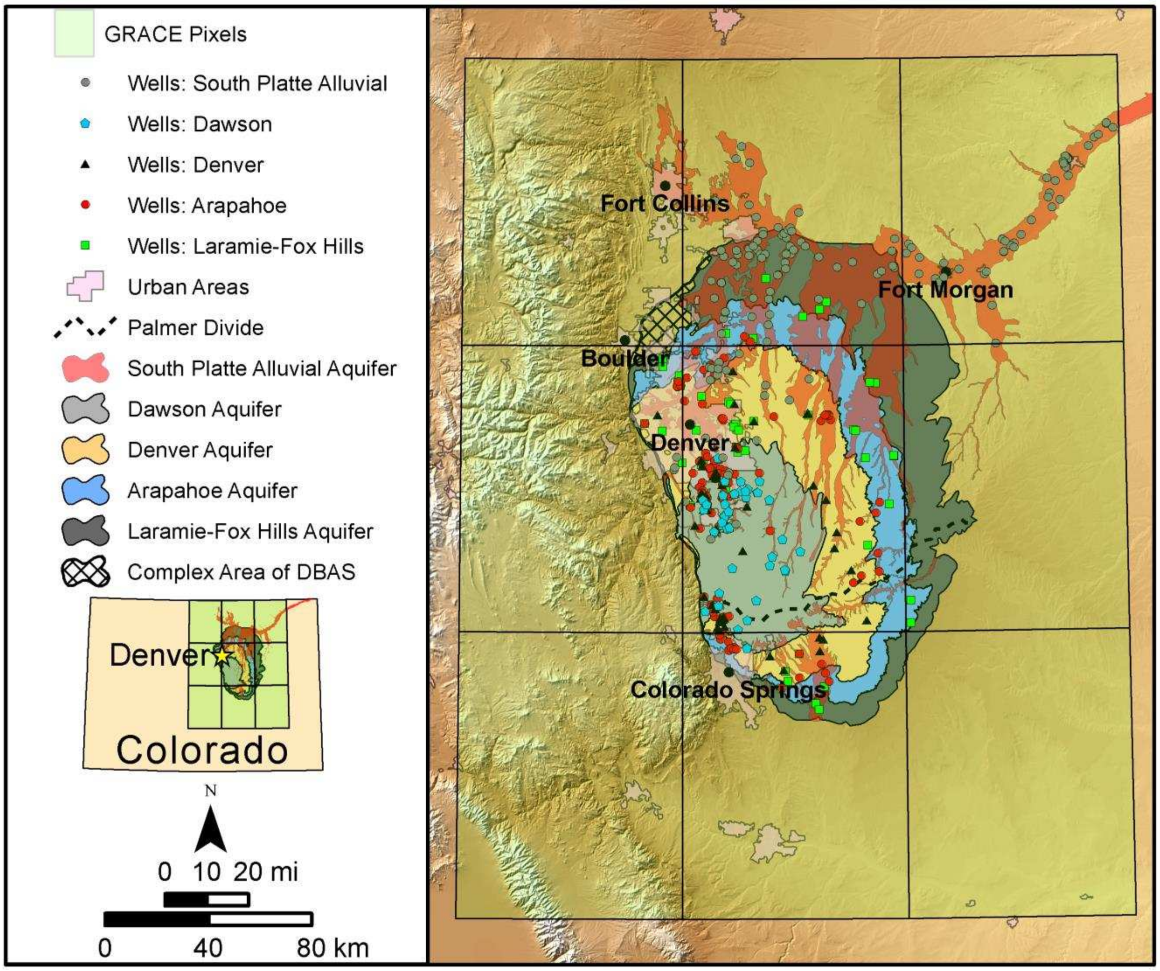

Figure 4-1 Study area showing the location of the GRACE pixels overlying the four aquifers of the Denver Basin Aquifer System in northern Colorado, USA. The available groundwater wells used in the analysis and for comparison to GRACE are shown.

Groundwater depletion has occurred in the DBAS over the past several decades as withdrawals have increased (Topper and Raynolds 2007; Paschke et al. 2011). Since the 1980s, groundwater level measurements in the DBAS have been irregular and contain many spatial and temporal gaps. Many wells are sampled annually, with a high frequency during the month of April. Other wells are sampled either multiple times a year or every couple of years.

4.4 Methods

4.4.1 Ground-Based Groundwater Data

Groundwater data for all four aquifers of the DBAS were obtained through the Colorado Division of Water Resource's HydroBase, a central water resources database for Colorado (Colorado Division of Water Resources 2017). Groundwater data were filtered to only include wells with greater than five observations and that appeared to be under static conditions. We identified 54 groundwater wells in the Laramie-Fox Hills Aquifer, 118 wells in the Arapahoe Aquifer, 60 wells in the Denver Aquifer, and 44 wells in the Dawson Aquifer that met these criteria. A twelve point moving window was used to remove outliers that were greater that two standard deviations from mean, which are likely due to data entry or measurement errors.

4.4.2 Normalized Groundwater Method

Normalized groundwater level anomalies are compared to the GRACE derived groundwater signal as an initial analysis of the potential to utilize GRACE at the aquifer-scale. A normalized approach was selected as an initial approach to avoid the uncertainties in applying aquifer storativities. Anomalies are calculated as,

$$z = \frac{x - \bar{x}}{\sigma} \tag{4-1}$$

where z is the normalized groundwater level (-), x is the water level measurement, $\bar{x}$ is the groundwater level mean relative to the 2004 – 2009 time mean baseline and then to the 2002 – 2015 time mean for the study period. Anomalies are generated for each well within the five aquifers in the study area. Representative signals from each aquifer are then calculated as an average of the well observation anomalies. The cumulative groundwater signal, which is comparable to GRACE, is then calculated by adding all of the representative aquifer anomalies together.

4.4.3 Spatio-Temporal Method

Spatio-temporal kriging was also used to fill temporal data gaps of groundwater levels for wells in the DBAS to and obtain representative groundwater level changes for comparison to GRACE. This approach was conducted after the initial normalized groundwater level analysis previously discussed. We utilize the data and techniques outlined in Chapters 2 and 3. Spatio-temporal kriging leverages data in adjacent space-time periods to help infer the predictions for the time period of interest. This helps overcome some of the limitations in irregularly sampled and data sparse aquifers. In general, groundwater levels are assumed to follow a random function model that consists of a trend and residual component that can be modeled separately. A global trend is removed from each aquifer and the residuals are spatio-temporally kriged. The final estimate of groundwater levels is taken by adding the predicted residuals back to the global trend component. We utilized the sum-metric model for modeling the sample semivariograms (Rouhani and Hall 1989; Dimitrakopoulos and Luo 1994; Myers 2004; Derakhshan and Leuangthong 2006).

Groundwater level changes for individual wells were converted to representative groundwater anomalies for comparison to GRACE. Anomalies are first computed relative to the

2004 – 2009 time mean baseline and then to the 2002 – 2015 time mean baseline to represent the baseline for this study period. The time mean for each groundwater location is determined from 2004 – 2009 and then subtracted from each groundwater measurement. Then the mean from 2002 – 2016 is determined and subtracted from each anomaly value determined in the first time mean removal. Anomalies are represented as a change in water table height relative to baseline mean. Calculations are then made using changes in the anomaly and an appropriate storativity for unconfined and confined aquifers. The volume of water drained as heads decrease is calculated as (Fetter 2000),

$$\Delta GW_u(unconfined) = \Delta H \cdot A \cdot S_y \qquad (4\text{-}2a)$$

$$\Delta GW_c(confined) = \Delta H \cdot A \cdot S_s b \qquad (4\text{-}2b)$$

where ΔGW is the adjusted groundwater anomaly, ΔH is the monthly anomaly represented as a change in water table height relative to the baseline mean, A is an area weighting factor, S_y is the specific yield, S_s is the specific storage, and b is the aquifer thickness. The aquifers of the DBAS generally contain both unconfined and confined conditions. Uncertainties exist in determining a representative storativity due to spatial heterogeneities. We utilize a combination of available storativity data to highlight the uncertainty this parameter has on estimated storage changes. Rodell et al. (2007) identify that mischaracterization of Sy can lead to a mismatch when compared to GRACE.

4.4.4 GRACE Data and Processing

GRACE data are processed at three main centers, the Center for Space Research (CSR) at the University of Texas at Austin, the Jet Propulsion Laboratory (JPL), and the German Research Center for Geoscience (GFZ). These centers differ by approaches used to process the data, such as background models, how the orbits are integrated, and the maximum degree of estimated spherical harmonics (Wouters et al. 2014). The gravity field data are represented through Stokes coefficients of spherical harmonic functions. Initial processing removes systematic errors, random errors, and reduces noise at the higher frequencies (Wouters et al. 2014; Wahr et al. 2006; Swenson and Wahr 2006). Several filtering approaches exist (e.g. 300 km Gaussian filter), but can lead to signal loss. Bias and leakage corrections are then used to restore the signal (Swenson and Wahr 2002). Using an ensemble of the three processing centers (calculated as an arithmetic mean) has been shown to reduce noise, scatter, and is more reliable than the product of a standalone center (Sakumura et al. 2014). In this study, an ensemble of the three processing centers is used.

Data from the three processing centers are presented as monthly solutions and represent averages in time and space (Landerer and Swenson 2012). Gravitational anomalies are converted to equivalent water heights and are often referred to as terrestrial water storage or total water storage (TWS) anomalies (Wahr et al. 1998). Over land, TWS is a measure of relative storage changes on the surface (i.e. lakes, reservoirs, rivers, snow, and ice) and the subsurface (i.e. soil moisture and groundwater). In this study, mass changes are based on Release 5 (RL05) spherical harmonics and datasets spans from April 2002 to March 2016, giving 152 monthly datasets with 13 missing months (due to battery maintenance of satellites). The datasets are a 1° x 1° gridded product (~111 km at the equator). This is provided as a Level-3 product, as post-processing and geophysical corrections have been applied. GRACE land data are available at http://grace.jpl.nasa.gov, supported by the NASA MEaSUREs Program.

Signal attenuation during post-processing of GRACE data necessitates the need to apply gain factors to restore the lost signal (Landerer and Swenson 2012). Signals not restored limit the accurate representation of a closed water balance. Generally three methods are applied and include scale factors (Long et al. 2015; Landerer and Swenson 2012; Sun 2013; Rodell et al. 2009; Chen et al. 2014; Feng et al. 2013), additive approaches (Long et al. 2015; Longuevergne et al. 2010), and multiplicative approaches (Long et al. 2015; Swenson and Wahr 2007; Longuevergne et al. 2010). Scale factors are one of the most widely used approaches (Long et al. 2015; Sun 2013). Factors are calculated using a least squares fit between filtered and unfiltered TWS from land surface models (LSM) (Long et al. 2015). The LSM used to determine the factors potentially bias TWS for smaller spatial scales and can lead to larger errors (Landerer and Swenson 2012). At the basin scale, the temporal variability in scaling factors is said to be lower than at the grid cell (Long et al. 2015). Gridded scale factors are provided with the GRACE data and are derived from the Community Land Model 4.0 (CLM4.0) of the National Center for Atmospheric Research (NCAR) (Landerer and Swenson 2012). The same grid of scale factors is used across the entire time span of the GRACE data. It is suggested though, that these factors may vary temporally (Rodell et al. 2009). To use a scale factor, GRACE TWS anomalies are simply corrected by multiplying the scaling factors to the corresponding GRACE data grid cells. To be consistent with the baseline period for this study, GRACE TWS anomalies are adjusted by computing the average over 2002 – 2016 for each grid point and subsequently subtracting this value from all time steps.

A more recent basis function used to fit the inter-satellite observations are mass concentration blocks (mascons) (Watkins et al. 2015). An advantage of mascons over the spherical harmonic coefficient approach is that the noise filtering is applied at the Level-2 step, rather than during the post-processing of Level-2 data. Mascons have less reliance on empirical scale factors to accurately represent mass estimates and have improved spatial resolution for smaller scale basins (Watkins et al. 2015). The solution solves for 4,551 equal area 3° spherical cap masons globally. The data are represented as 0.5 degree grid cells (~56 km at the equator), but represent 3-degree equal area caps. The mascon solutions are also utilized for comparison to the spherical harmonic solutions over the study area.

GRACE TWS is often disaggregated to isolate different components that make up the total signal. TWS represents a vertically integrated product that over land includes surface waters (SW), snow water equivalent (SWE), biomass (BM), soil moisture (SM), and groundwater (GW) (Strassberg et al. 2009; Scanlon et al. 2012b). These are represented in equation form as,

$$\Delta TWS = \Delta SW + \Delta SWE + \Delta BM + \Delta SM + \Delta GW \tag{4-3}$$

where Δ represents the change (i.e. monthly). If auxiliary data sources are used (e.g. LSM, gauges), using the GRACE TWS, equation 1 can be rearranged to back-calculate the change in groundwater storage (ΔGW) as,

$$\Delta GW = \Delta TWS - \Delta SW - \Delta SWE - \Delta BM - \Delta SM \tag{4-4}$$

The calculated changes can then be compared to groundwater well data or used to estimate long-term or short-term ground water storage changes. However, depending on the location of the area of interest, some of the components that comprise TWS can be ignored. Over the Norther High Plains Aquifer, changes in surface water storage and plant canopy storage (biomass) were insignificant to changes in TWS at the annual time scale (Seyoum and Milewski 2016; Strassberg et al. 2009). Instead, TWS was dominated by groundwater storage with the exception of a few years where the unsaturated zone (i.e. soil moisture) dominated the TWS response. Similarly, it was observed that groundwater storage changes dominated the TWS observations in a global study (Richey et al. 2015). However, contrary to this, during the 2011 drought in Texas, it was shown that changes in groundwater storage had a low contribution to the TWS change (Long et al. 2013). In other places, such as over Alberta Canada, accounting for soil moisture, snow, and ice improved the correlation between GRACE derived groundwater storage and groundwater observations (Huang et al. 2016). In this study, we assume soil moisture, snow water equivalent, surface water bodies (e.g. reservoirs), and groundwater account for the majority of the TWS signal.

Land surface models (LSMs) such as Global Land Data Assimilation System (GLDAS) and the North American Land Data Assimilation System (NLDAS) are commonly used to remove components of the TWS signal due to lack of measured in situ data and/or good spatial and temporal coverage of these components. However, several studies have indicated that soil moisture problems can exist from LSM's causing unrealistic storage changes for groundwater (Huang et al. 2012; Long et al. 2013; Huang et al. 2016; Breña Naranjo et al. 2014). Most LSM's don't account for deep soil moisture (greater than 2m) and ground water storage changes (Rodell et al. 2009). In the High Plains Aquifer, it was found that soil moisture variability was limited to the upper 2 m of the soil profile (Strassberg et al. 2009). Even in areas that were heavily irrigated, it was found that the upper 2 m accounts for 76 – 98% of the variability in soil moisture. In this study soil moisture data were obtained from NLDAS using an ensemble of the Noah and Mosaic models. This product is a gridded 12.5 km resolution.

Soil moisture output from NLDAS was modified to account for increased soil moisture due to irrigation. NLDAS does not account for added soil moisture due to irrigation (Xia et al. 2012).Estimates of actual evapotranspiration (Eta) from the operational version of the Simple Surface Energy Balance model (SSEBop) were used as a proxy to the total applied irrigation water (Savoca et al. 2013). ETa is used to represent the associated consumptive use and a portion of the applied irrigation water. An irrigation land cover map was used to identify irrigated land in the study area and the ETa was masked to include only the regions that were considered to be irrigated land. The irrigation season in the area spans from April 1 to October 15. The added soil moisture from irrigation was assumed to only occur during this period. It was determined that plant canopy storage output within NLDAS was considerably smaller than the GRACE TWS output. Thus, the biomass term (BM) in equations 2 and 3 is neglected in deriving the groundwater storage signal. Surface water influences to the GRACE signal were attributed to reservoir storage. Approximately 18 reservoirs were identified and the corresponding storage records were obtained.

Snow and ice are represented as snow water equivalent (SWE) in equation 4-2 and 4-3. Many studies use GLDAS for SWE input (e.g. (Huang et al. 2016; Chen et al. 2016; Henry et al. 2011; Yang and Chen 2015) and others use NLDAS (e.g. (Wouters et al. 2014; Breña Naranjo et al. 2014; Long et al. 2013; Scanlon et al. 2012b). Comparison of the 12.5 km cell size of NLDAS with Snow Telemetry (SNOTEL) sites in the Rocky Mountains, which are point measurement, indicates NLDAS under predicts SWE. In this study, the Snow Data Assimilation System (SNODAS) data product is chosen to represent SWE. SNODAS is a 1 km gridded snow product that is used instead of NLDAS due to better agreement with SNOTEL sites. Because SNODAS's period of record begins in October of 2003, SWE values from NLDAS from April 2002 – September 2003 were multiplied by a calculated scaling factor (for better agreement with SNOTEL sites) and used for the missing record in SNODAS. This was done to have SWE data corresponding to the entire GRACE record, which began in April 2002. SNODAS is then upscaled to match the GRACE resolution. Soil moisture, snow water equivalent, and reservoir storage are all converted to anomalies to be comparable to GRACE. Both the 2004 – 2009 and 2002 – 2013 averages are subtracted from values for each component.

4.5 Results and Discussion

4.5.1 GRACE Total Water Storage

The GRACE Total Water Storage (TWS) signal (Figure 4-2) shows annual cycles with highs and lows following the winter and summer months, respectively. The TWS is represented as an average of the nine pixel study area (Figure 4-1). The TWS fluctuated around baseline from 2002 – 2008 before overall increasing until 2011. After 2011, TWS decreases to be below baseline for the rest of the study period. Of all the water storage components considered in Equation 4-3, the dominant signal comes from soil moisture (Figure 4-2). This is possibly due to large evapotranspiration (ET) losses across the study area. ET fluxes commonly exceed precipitation, leading to greater fluctuations in soil moisture. The soil moisture anomaly closely tracks the GRACE TWS seasonal signal and position relative to baseline conditions. Several of the annual amplitudes show a slight lag in the soil moisture anomaly relative to GRACE TWS.

Soil moisture from NLDAS is adjusted to account for additional soil moisture from irrigation that is not accounted for in the model. The other signals (i.e. reservoir storage, snow water equivalent) make up only a small component of the GRACE TWS signal. *GRACE*

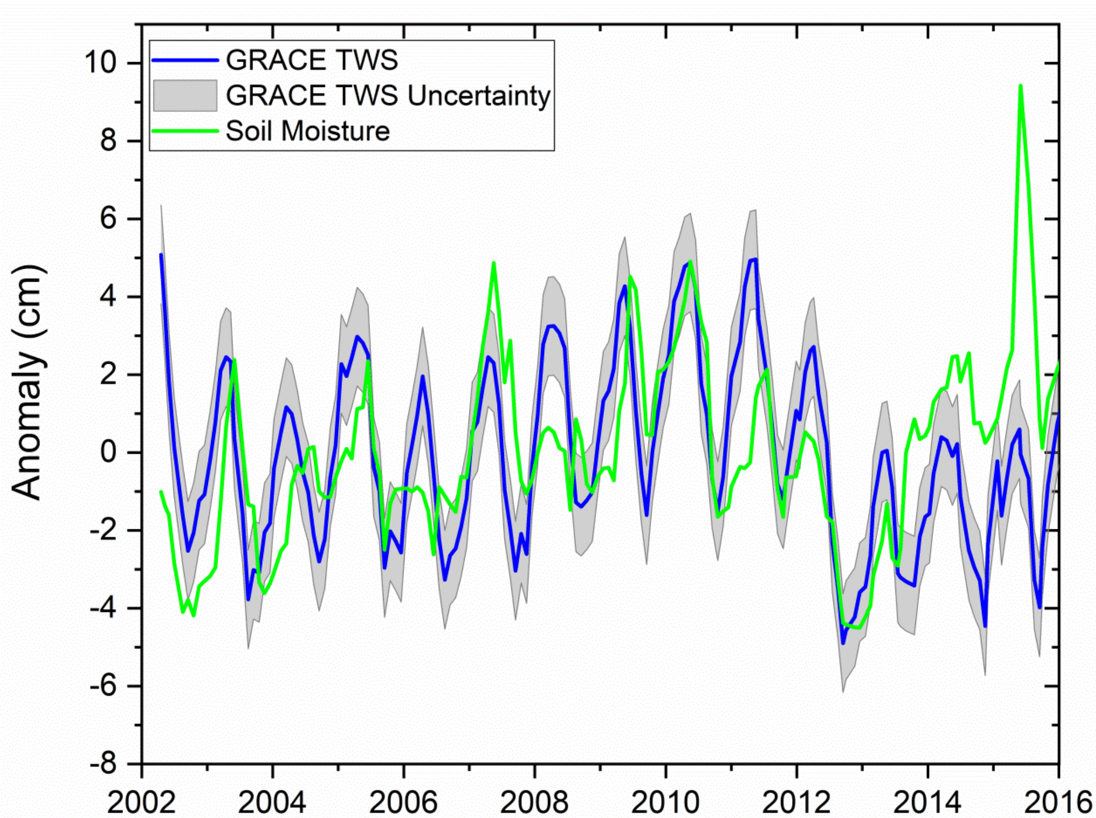

Figure 4-2 Monthly GRACE TWS changes (in cm) and monthly soil moisture anomaly (in cm) across the study area.

4.5.2 Groundwater Storage

The monthly changes in GRACE derived groundwater storage are shown in Figure 4-3 (in cm) and represents the average groundwater signal across the nine-pixel study area. The results generally show higher rates of depletion, or anomaly decreases, during times of extreme and exceptional drought. In the case of the 2012 – 2013 drought period, groundwater anomalies continue to decrease until early 2015. The anomaly fluctuates around baseline for much of the period of record, except for after mid-2013. The overall linear trend from 2002 – 2016 shows 0.3 cm/year decrease in levels. However, prior to 2014, the long-term rate of depletion was 0.1 cm/year. During the extended drought from 2002 – 2008, linear rates of groundwater depletion amounted to about 0.5 cm/year. From a volumetric perspective, across the DBAS, the trend from 2002 – 2016 equates to ~0.07 km^3/year depletion. The long-term depletion rate prior to 2014 was ~0.03 km^3/year.

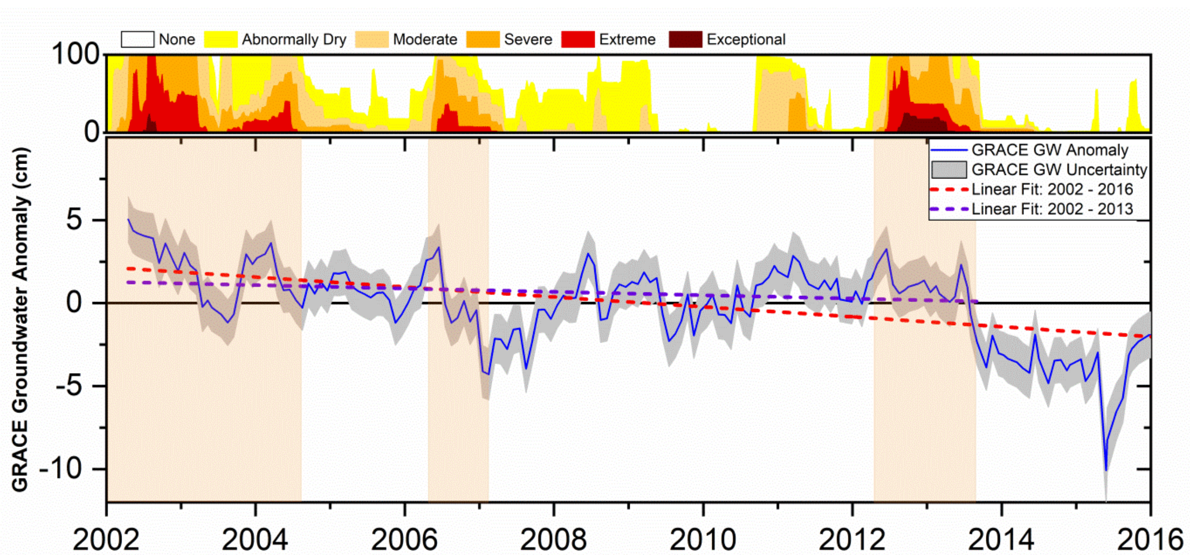

Figure 4-3 Monthly GRACE derived groundwater (GW) changes (in cm) shown in the bottom plot are compared to the U.S. Drought Monitor (top) across the study area. The U.S. Drought monitor represents the percent of the study area that is within a given drought index.

4.5.3 Comparison to Groundwater Well Observations

Groundwater wells were first compared as a normalized groundwater levels (Figure 4-4). Comparisons of the water-levels for the cumulative well observations for all aquifers are consistent with the increasing and decreasing patterns of the GRACE derived groundwater signal. Individually, the normalized levels do not appear to match the GRACE signal for a given aquifer, but do have periods were the trends match. The depletion trends in the Denver Aquifer after 2012 appear to strongly influence the declines that the GRACE derived groundwater signal also detects. Because the cumulative normalized well observations show similar trends as the GRACE derived groundwater signal, further analysis is conducted on using storativities and confined versus unconfined wells to better understand the potential for GRACE to help infer temporal changes in levels.

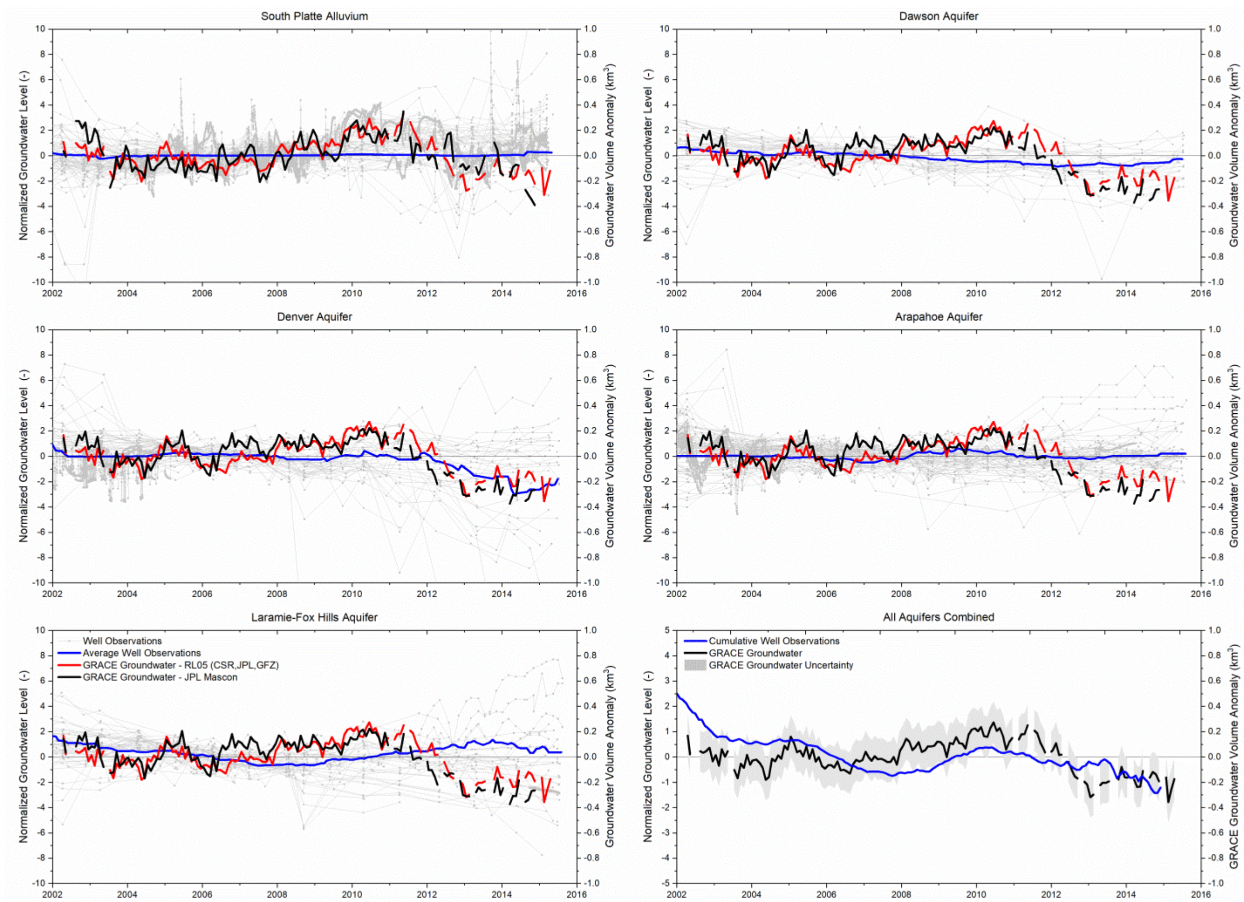

Figure 4-4 Normalized groundwater levels for each of the five aquifers in the study area compared to GRACE derived groundwater.

Groundwater wells with temporal gaps filled using spatio-temporal kriging were used next better understand the contribution of each aquifer on the total signal represented by the

GRACE (Figure 4-5). Wells in both unconfined and confined areas were included in the analysis and a range of storativities were used in the calculations to better understand the influence of the storativity parameter on the observed trends, considering that storativity equals specific yield for unconfined aquifers and equals specific storage multiplied by aquifer thickness for confined aquifers. The average anomaly produced from the individual well anomalies in a given aquifer is used to represent the aquifer specific signal. Results comparing the well observations with GRACE do not show similar trends and variations at the individual aquifer level (Figure 4-5). There are occasions though where the two time series being compared for a given aquifer portray similar increasing or decreasing signals. The groundwater anomalies using observational wells exhibit the largest variability for the Laramie-Fox Hills, Arapahoe, and Dawson aquifers. Groundwater fluctuations in these aquifers highlight the influence storativity has on the anomaly shape and magnitude of variations. These three aquifers also show a shift toward negative anomalies after around 2008. The groundwater anomalies for the Denver and South Platte Alluvial aquifer had small variations in time.

The five aquifer anomalies were combined into one total anomaly that would theoretically represent the GRACE groundwater derived anomaly. The areas of each aquifer areas were used to weight each respective anomaly and the contribution towards the combined signal. The combined anomaly shows close resemblance to the anomaly of the Laramie-Fox Hills Aquifer, due to the Laramie-Fox Hills anomaly receiving the largest weight based on area. The combined anomaly bears some resemblance of the GRACE groundwater signal. Both show decreasing levels from 2002 – 2007. After 2007, the GRACE groundwater anomaly generally increases and hovers around baseline. The combined groundwater anomaly continues to decrease until about 2009 before increasing. At 2012, both show shift towards decreasing levels until

2015. However, the combined anomaly has poorer agreement than did the combined normalized groundwater levels. This suggests possible misrepresentation through using both confined and unconfined aquifers, incorrectly applying area weights based on aquifer size, or improper storativity values.

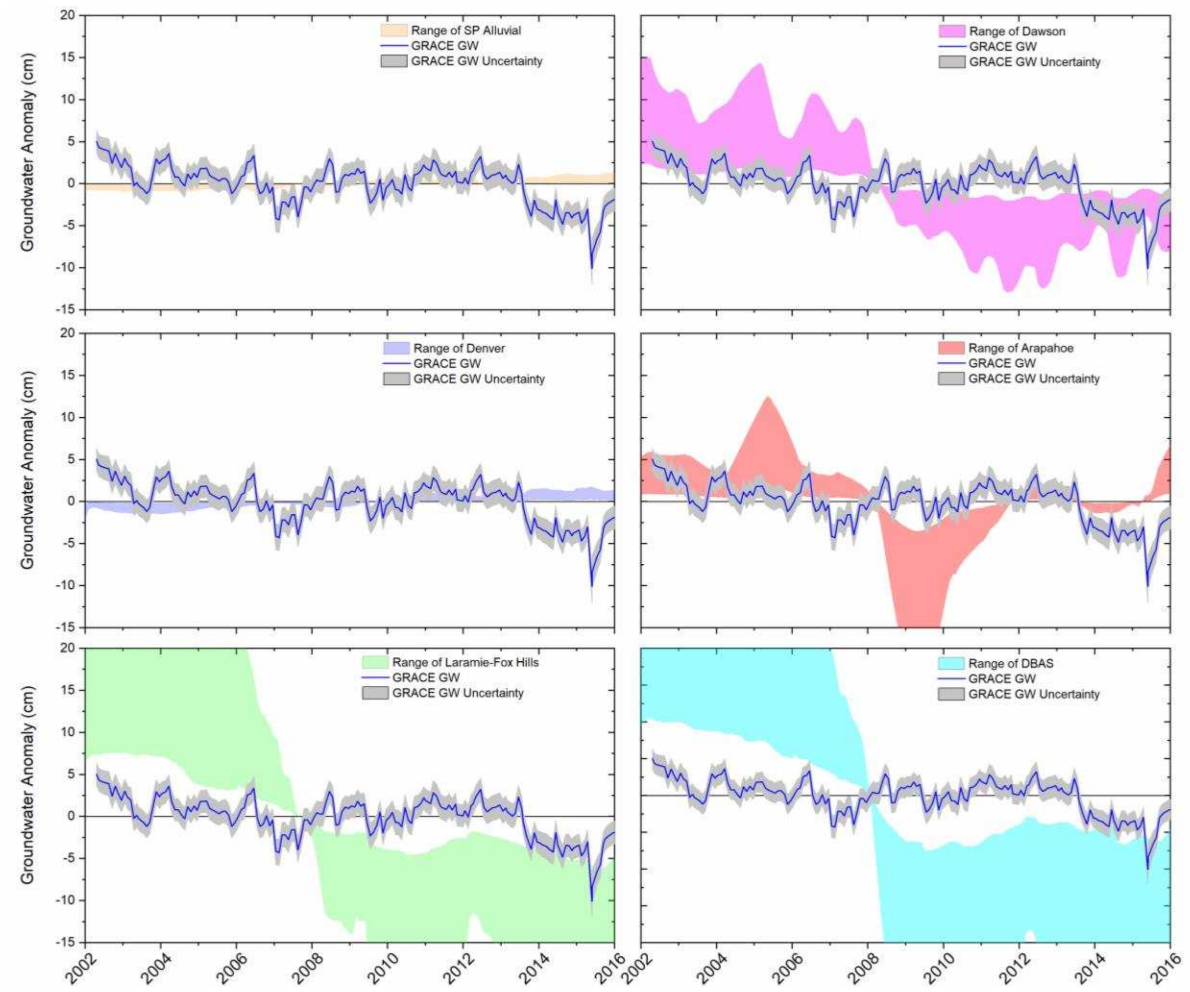

Figure 4-5 Monthly GRACE derived groundwater anomaly compared to aquifer specific anomalies for the four DBAS aquifers, the South Platte Alluvial Aquifer, and the DBAS combined. The color range of each aquifer represents the range of possible anomalies that can be calculated which are depended on the storativity values chosen to represent the groundwater well data.

The range of anomalies were calculated with specific yields between 0.04 – 0.25 and confined storativities from 0.00002 – 0.008. In each aquifer, the average anomaly was dominated by wells in unconfined portions of the aquifer due to the specific yields being used instead of confined storativities in the calculations. Use of smaller specific yields leads to anomalies that are closer in magnitude to the GRACE derived groundwater anomaly.

Each aquifer contained groundwater-well water level observations that were increasing, deceasing, or relatively stable. Of these wells, several had anomalies that closely resembled the

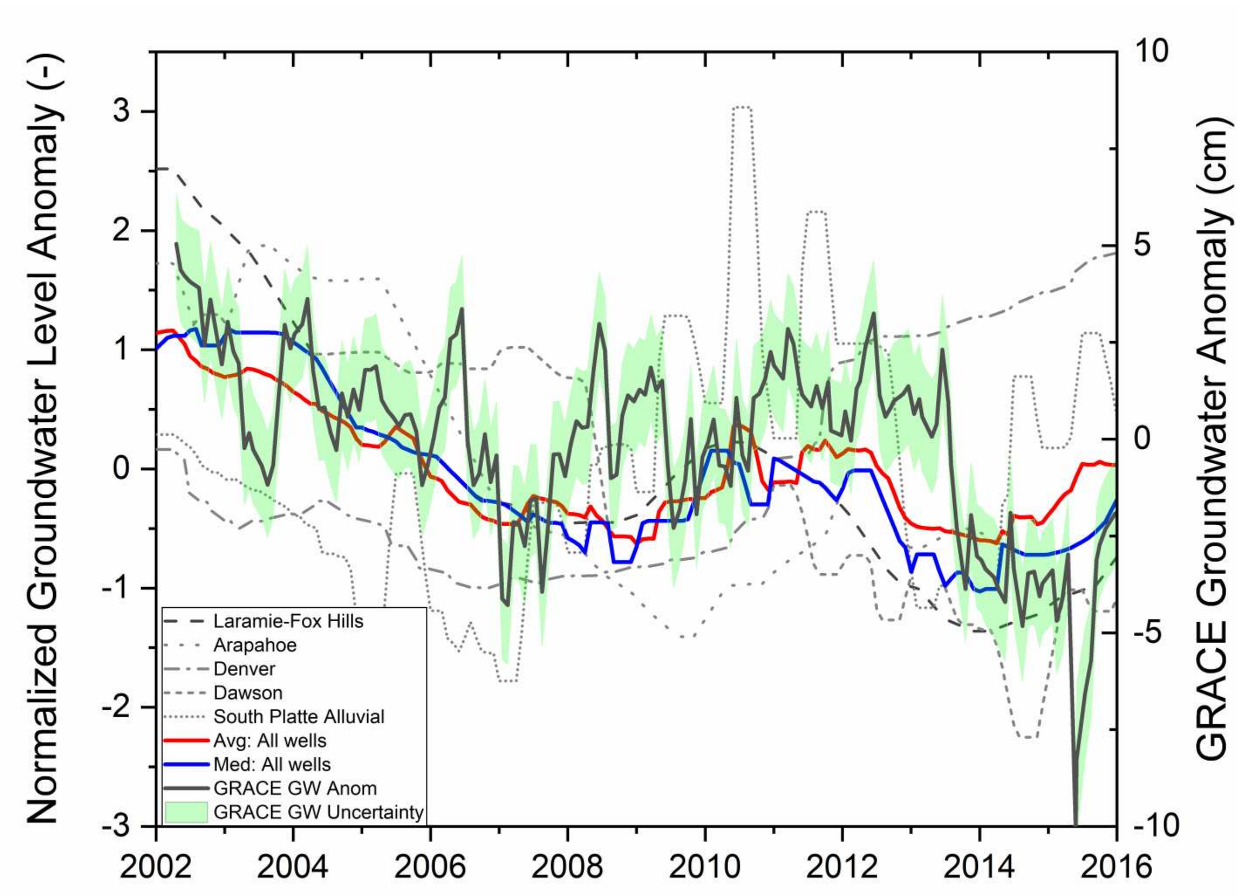

Figure 4-6 GRACE derived groundwater anomaly compared to select groundwater well time series from each aquifer in study area, represented as a normalized groundwater level anomaly (i.e. long-term mean removed and divided by standard deviation). The combined average and median anomaly of the select wells in each aquifer are also shown.

GRACE derived groundwater anomaly. One well from each aquifer was selected and normalized by the standard deviations of the time series. Each well was selected based on visual inspection of the trends. This was simply done as an exploratory analysis to see if well selection played a role on agreement or disagreement to GRACE derived groundwater anomalies. The average and median signal from the group of wells was calculated and plotted for comparison to the GRACE groundwater anomaly (Figure 4-6). The average and median time series portrays a similar signal to the GRACE groundwater anomaly. There is a decreasing trend from 2002 – 2008, an increasing trend from 2008 – 2012, and then a decreasing trend from 2012 – 2014. This highlights the high influence well selection has on agreement or disagreement to the GRACE groundwater anomaly.

4.5.4 Limitations and Uncertainty

There are several limitations and uncertainties inherent in the analysis of GRACE for applications at the aquifer-scale. Comparison for the accuracy and validity of GRACE was restricted solely to time series analysis of groundwater-level observation wells. Many GRACE studies that compare to observation well data utilize well time series. Due to the sparse nature of groundwater wells in each aquifer and irregular temporal measurements, we decided to not compare GRACE to estimated changes in an aquifer wide potentiometric surface at this time.

The evaluation of GRACE groundwater estimates in this study utilized the common GRACE Level 3 post-processed outputs that are scaled using CLM4-baed gain factors (Landerer and Swenson 2012). Scaling factors are used to restore signal attenuation and bias that occurs during postprocessing techniques that result from spatial leakage (Swenson and Wahr 2002; Landerer and Swenson 2012). The scaling factors are generally applied across all temporal periods to restore lost signal. However, it is possible that using the same scaling factor across all temporal periods is inappropriate for our study area. These factors may also not be appropriate if there is no significant spatial correlation of hydrological signals that can be corrected by a simple factor (Nanteza et al. 2016).

In this study, no additional postprocessing was conducted to either GRACE data or data from the other hydrologic components. Some studies apply similar processing techniques that are used for GRACE (i.e. spherical harmonics truncated at degree 60, decorrelation filter to destrip, and 200 – 3—km Gaussian filter to smooth) to the other hydrologic components (Longuevergne et al. 2010; Nanteza et al. 2016; Döll et al. 2014). Other studies incorporate a priori information of point mass changes inside and outside of a basin that may cause leakage and influence the GRACE results (Longuevergne et al. 2013). Some error in this study could be due to the lack of accounting of outside influences and not incorporating a priori spatial information for point masses (e.g. reservoirs) as suggested by (Longuevergne et al. 2013).

One of the goals of this study was to evaluate the application of GRACE at the aquifer–scale. Despite numerous publications in the literature for applying GRACE below 200,000 km^2, it is generally recommended to not use GRACE at smaller scales. However, it is possible that some useful information about water storage can be inferred. The nine pixels in used in the study have an area ~86,000 km^2. Disagreement with aggregated observational well data and utilizing storativities could be due to several reasons. First, the GRACE TWS signal represented over the nine pixel area might be more representative of the larger regional signal and to some extent attenuated from the larger decreases in the five aquifers. Therefore when the other hydrologic components were removed, the derived groundwater signal could be biased low if the localized signal were in actuality much higher than the regional signal. The second reason for disagreement might have to do with the distribution of observational well data. Because the dynamics of individual wells in the DBAS and South Platte Alluvial Aquifer are quite variable, it is possible that we didn't have a large enough dataset from which to generate representative aquifer changes, that would correspond to average conditions for the areal size of a GRACE pixel, especially when separating confined from unconfined wells as was done in our analysis. Confined aquifer levels may not be represented by GRACE because even under depletion, the aquifer remains saturated, albeit at lower pressure and doesn't constitute enough mass redistribution to be detected (Alley 2016). However, in each aquifer, there existed groundwater wells that had similar water-level patterns as GRACE derived groundwater. When plotting these wells versus GRACE, there was much better agreement. This suggests that when making comparisons to GRACE, not all groundwater well are appropriate to use for comparison. Aquifer wide variabilities in storativity (Longuevergne et al. 2013) are another influence on the disagreement. One storativity value is attributed to each well within a designated aquifer due to lack of spatial storativity data. Having well specific storativities might improve determining accurate anomaly changes for each well by allowing an area-weighted analysis of spatially variable storativity values. Different aquifer weightings schemes were also used when determining the combined groundwater anomaly signal for the ground-based data. Weighting the contribution by aquifer size led to the results presented. Finally, it was determined that soil moisture was the largest contributor to the GRACE TWS signal. We applied a correction the NLDAS soil moisture to account for irrigated agriculture between the irrigation season between April 1 and October 15. We assumed that actual evapotranspiration (ETa) could be used as a proxy to represent a portion of the applied irrigation water that would remain as added soil moisture. We did not apply any scaling to reduce the ETa so that less added soil moisture would be considered. Our current approach assumed a shallow soil moisture profile (200 cm) is representative of the entire vadose zone. This might be true for areas with deep vadose zones that are well drained, but for shallower vadose zones, this may not be true. The vadose zones across the DBAS and South Platte Alluvial Aquifer are quite variable and may be between 5 to 100 ft. Accounting for monthly soil moisture fluctuations below 200 cm may be important for accurately representing GRACE derived groundwater changes.

4.6 Conclusions

Over the past decades, groundwater depletion in the Denver Basin Aquifers has increased as population rapidly grown and water withdrawals increased. Groundwater well data across the DBAS is irregular with many spatial and temporal gaps and presents challenges for assessing the dynamics and long-term sustainability of groundwater resources in the area. The GRACE satellites offer a unique opportunity to assess water resources due to the high temporal frequency of data collection and big picture view. Although GRACE does have limitations, some are different than those of using groundwater level and aquifer characterization data to assess groundwater storage changes in an aquifer. In this study, the accuracy and validity of GRACE over the Denver Basin Aquifer System is assessed from 2002 – 2016.

At this point it is uncertain whether or not GRACE can directly be applied to understand groundwater resources in the DBAS. A large number of uncertainties exist in processing GRACE data and correctly removing the various hydrologic storage components. While reasonable agreement was shown using normalized groundwater levels, much poorer agreement to GRACE was shown when separating unconfined from confined wells and applying various storativities. Further analysis should be done in determining aquifer wide storativities and improving the representative time series for each aquifer. GRACE still may useful if the other uncertainties and proper aquifer representation can be determined for more accurate comparisons.

CHAPTER 5

CONCLUSIONS

Most aquifers in the arid west, as well as in many other locations across the globe, are facing serious questions about future sustainability (e.g. (Famiglietti 2014; Konikow 2015; Scanlon et al. 2012a). Typically, sustainability is evaluated by analyzing groundwater well records to determine if water levels can be sustained into the future under certain use scenarios. Of course, water levels only tell part of the story, because groundwater availability actually depends on future projects of groundwater storage, which is highly uncertain. In addition, analysis of water well data as it is typically conducted has serious limitations. Data are typically spatially and temporally sparse relative to the data density desired for aquifer-scale analysis. Another issue revolves around aquifer storativities, which are fundamental to understanding the response of an aquifer to water-level and storage changes. However the availability and reliability of storativity data leads to further uncertainties in groundwater storage and sustainability calculations. The problems with missing temporal data from water wells in particular has not been addressed in much detail, yet can cause important misinterpretation with regard to groundwater sustainability.

This research aims to mitigate some of the problems with current approaches to analyzing water well data by incorporating a new method of spatial-temporal analysis, with particular emphasis on addressing missing temporal data. In addition to this proposed method, we evaluate the ability of GRACE satellite data to improve the temporal sustainability analysis. Continual monitoring and assessment of aquifer dynamics is essential for long-term management and sustainability of groundwater resources. Assessing aquifer wide water-level and storage changes using sparse datasets is not trivial because of the limited information being used to infer changes in a heterogeneous system without introducing unnecessary bias.

In this dissertation a spatio-temporal kriging methodology is introduced which leverages data in space and time to minimize the temporal sparseness of data and understand changes in groundwater systems. The methodology is first illustrated using a case study in the Arapahoe Aquifer and is then expanded to all aquifers of the Denver Basin Aquifer System (DBAS). Remote sensing is utilized from the Gravity Recovery and Climate Experiment (GRACE) satellites to provide another perspective on determining groundwater storage changes. Results from this work provide a framework for monitoring and management of groundwater resources along the Colorado Front Range Urban Corridor as well as other water-stressed regions of the western U.S.

5.1 Summary of Findings

The following summarizes each objective of this dissertation, which are presented as a discussion of how the findings relate to the science questions.

<u>Objective 1: Evaluation of Spatio-Temporal Kriging to Predict Potentiometric Surfaces</u>

Science Question 1: Do estimated potentiometric surfaces using spatio-temporal kriging improve our understanding of groundwater sustainability than traditional methods of groundwater well data analysis (e.g., inverse distance weighting, point to raster) or methods that use only spatial kriging?

Chapter 2 evaluates spatio-temporal kriging as an alternative interpolation method with the goal of improving the accuracy of potentiometric surface predictions. This study is only one of three other known studies identified in the literature to apply spatio-temporal kriging specifically to a groundwater system. Our approach is different than the other studies in that we consider a different space-time model, apply the spatio-temporal approach at a much larger aquifer-scale, consider confined and unconfined areas, and utilize a linear regression to remove the trend component from the data. It is also the first known study of this type in the DBAS.

The generated groundwater maps provide a visual and quantitative assessment of the uncertainty in predictions for groundwater levels for both spatial and spatio-temporal approaches. Potentiometric surfaces created through spatial kriging are subject to unrealistic changes in hydraulic head and can lead to an erroneous analysis of spatial ground water conditions for a specific time, and in particular to errors in a temporal analysis, problems that are common in sparsely monitored areas. Two likely causes exist for erroneous potentiometric surfaces created through traditional spatial kriging. First, the differences in spatial maps are a result of the wells used in the analysis, rather than the true differences in groundwater level data. This leads to problems when trying to understand changes because the fluctuations are attributed to the presence or absence of well data. Second, because temporal aggregation of data is sometimes used before traditional spatial kriging, well data with multiple measurements during the year are averaged and can lead to artificial levels being compared. Differences would then represent changes from an average condition, making interpretation of the change difficult to assess. Spatio-temporal kriging allows for smoother and more realistic groundwater changes to be represented. We assume the strength of the temporal and spatial dependence represented in the semivariogram and shape are the same for all locations over the study period. Spatio-temporal kriging helps overcome some limitation of temporally sparse data, however, it cannot increase the spatial density beyond the original well locations used in the analysis. To increase the prediction accuracy for some areas, more spatial monitoring locations are needed. The cross-validation for the spatio-temporal approached predicted measured water level elevations much better than did the spatial kriging approach. Using neighboring temporal and spatial data improved the potentiometric surface predictions because more information was available to help infer an appropriate estimate.

The Arapahoe Aquifer in the DBAS is used as a case study to highlight the benefits of spatio-temporal kriging. While rates of change have been variable over the 27 year study period, rates were the highest during the 2000s with areas of the aquifer experiencing up to 55 ft per year declines in levels. The west-central side of the aquifer shows the largest overall declines, with some declines up to 600 ft under the municipalities of Parker and Castle Rock.

<u>Objective 2: Evaluation of Groundwater Storage in the Denver Basin Aquifer System</u>

Chapter 3 focuses on applying the developed spatio-temporal kriging approach to understanding groundwater level and storage changes in the DBAS, any spatial patters, and the influence of the storativity parameter on volumetric estimates. This study presents the most current information and results for groundwater level and storage changes in DBAS.

Science Question 2: How has groundwater storage changed for aquifers of the DBAS since 1990?

Each aquifer in the DBAS varies in overall depletion over the 27 year study period. Higher rates of depletion and larger storage declines occurred in the 1990s and 2000s. The cumulative groundwater depletions are highest for the Denver and Arapahoe Aquifers. Due to the larger declines in hydraulic head in these two aquifers, areas transitioned from confined to unconfined conditions, leading to higher calculated depletion volumes than would otherwise be calculated assuming confined conditions. Historically, these areas have been assumed to be confined. The higher rates of depletion began in the early 1990s for the Denver Aquifer, while the Arapahoe Aquifer depletion rates increased in the 2000s. The Dawson and Laramie-Fox Hills Aquifers experience much less depletion. The Dawson Aquifer displays increasing and deceasing storage patters that follow drought and precipitation patterns. The Laramie-Fox Hills Aquifer displays steady declines in storage over the study period. However, the cumulative depletion volumes for the Laramie-Fox Hills Aquifer are likely larger due to predicted potentiometric surfaces being higher than the top boundary of the aquifer in areas thought to be unconfined, leading to a specific storage being erroneously used in the calculations instead of a specific yield. Due to the spatial sparseness in the data, few groundwater wells are located near the edges of the aquifer to help infer the predictions for this aquifer.

Some areas of the DBAS aquifers do show increasing groundwater storage. In the Dawson Aquifer, the increases are located near Monument and west of Castle Rock and reach up to 0.1 km^3 in areas. The Denver Aquifer also shows increases in storage south of Monument up to 0.03 km^3. The Arapahoe and Laramie-Fox Hills Aquifers show increases near Highlands Ranch and to the northwest of Denver. The Highlands Ranch area is also the location of aquifer storage and recovery wells operated by Centennial Water and Sanitation District.

Potentiometric surfaces by the end of 2016 indicate areas of the Denver and Arapahoe Aquifer are transitioning from confined to unconfined conditions. Results indicate that much of the Denver Aquifer's eastern and western portions are unconfined and only a small area remains confined in the central region. In the Arapahoe Aquifer, the western side is starting to become unconfined and is also where the largest storage changes are occurring. The Laramie-Fox Hills Aquifer is still dominated by confining conditions. The west-central region is estimated to have up to 1,500 ft of hydraulic head above the upper confining unit. The actual unconfined nature of the aquifers must be carefully considered for future groundwater storage and sustainability evaluations because quantifications using specific storage are likely to be much different than those that use specific yield.

Science Question 3: Utilizing available storativity data, what is the variability in storage change estimates?

Using median reported storativities from CDWR data, the estimated total DBAS depletion since 1990 is 36.6 km^3 (~29.7 million AF) and represents ~15% of the total estimated recoverable groundwater. However, the variability of estimated change is quite large depending on the storativity value selected for each aquifer. The Dawson aquifer varies between ~1 - 7 km^3 cumulative depletion from 1990 to 2016. The Denver Aquifer ranges from ~5 – 55 km^3, the Arapahoe Aquifer from ~3 – 34 km^3, and the Laramie-Fox Hills aquifer from ~0.4 – 5.2 km^3. These results highlight the importance of selecting a representative storativity when calculating volumetric storage changes. This study applied a single storativity, different for each aquifer, when calculating volumetric changes. This is common in many groundwater assessments. However, this research considered the uncertainty in storativity values, an uncommon practice in water resources evaluations. The results highlight the importance of proper characterization of storativity and should be of high interest and importance for water managers. Clearly, an assessment of volumetric storage changes in the DBAS would benefit from better characterization of storativities and possibly even applying an area weighted approach. Comparison of estimated groundwater storage changes were variable when compared to results from a previous USGS Denver Basin modeling study (Paschke et al. 2011). The most noticeable differences occurred in data sparse areas. The highest agreement generally occurred for areas in the west-central portion of each aquifer under the municipalities of Castle Rock, Parker, and Denver and is likely due to a higher density of data. By 2003, it is estimated that a total of 24

km^3 had been depleted from the DBAS aquifers, which is 24 times larger than the cumulative depletions predicted by the USGS model.

Science Question 4: Can we identify the main drivers for groundwater storage changes and how these impact aquifer sustainability?

A combination of population growth, reduced precipitation and recharge, and groundwater pumping are likely the largest factors contributing to groundwater depletion and reduced storage. The population across the Colorado Front Range Urban Corridor has increased by 73% since 1990, and now over 4.5 million people reside in the area. Municipalities with limited to no surface water access (e.g. Parker, Castle Rock) in the 1990s and 2000s relied heavily on groundwater from the DBAS and are also the areas overlying the largest water-level declines. Population is expected to continue to grow, further adding stress on limited resources. Precipitation departures from the area normal show reduced levels in the 2000s with recovery by 2014. Severe to exceptional drought was common in the early 2000s and 2012, likely leading to increased groundwater reliance because of reduced surface water access. Groundwater well development was also large during the study period. Over 18,000 wells were constructed between 1993 and 2007 and by 2016, at least 30,000 wells withdrew water from the DBAS. It is difficult to identify the relative contributions for each of these factors. However, obtaining better information and data on pumping volumes is necessary to better refine these contributions.

Objective 3: Evaluation of GRACE at the Aquifer-Scale

Science Question 5: Does analysis of groundwater storage trends from GRACE improve our understanding of groundwater availability/storage at the aquifer-scale compared to those determined from ground-based data for the DBAS?

Chapter 4 evaluates the applicability and usefulness of using the GRACE satellites at the aquifer-scale. Based on the altitude of the satellites (~450 km), the footprint of GRACE is estimated to be ~200,000 km^2 and have an accuracy of 1.5 cm for larger footprints (Famiglietti and Rodell 2013). Over the past several years, there has been much debate on the usefulness of GRACE at scales relevant for water managers, as well as using GRACE below the recommended footprint (Alley and Konikow 2015).

At this point it is uncertain whether or not GRACE improves our understanding of groundwater availability in the DBAS. While the general trends from an initial normalized groundwater level approach closely matched the observed trends from the GRACE derived groundwater anomaly, an approach that applied spatio-temporal kriging to fill temporal gaps in the data and utilize aquifer specific storativities had much poorer performance. A large number of uncertainties were identified and possible a source of error. It is possible that the GRACE TWS used to derive the groundwater anomalies represents a regionally attenuated signal that does not fully capture the changes across the DBAS. Soil moisture appeared to heavily influence GRACE TWS. The variability is soil moisture was only accounted for 200 cm in depth from the surface and actual evapotranspiration was used as a proxy for added soil moisture due to irrigation. Further work should focus on refining soil moisture estimates. In addition, further analysis should focus on determining aquifer wide storativities and determining a better method to combine the representative time series for each aquifer into a composite signal. GRACE still may useful if the uncertainties can be corrected and proper methods for combining the composite aquifer signal can be determined.

5.2 Concluding Remarks

The work presented in this dissertation highlights the benefits and advantages that spatio-temporal kriging and remote sensing products can contribute to groundwater sustainability assessments. For spatio-temporal kriging, future work should focus on the impacts of model selection, proper grid cell size for predictions, inclusion of spatio-temporal semivariograms that account for spatial anisotropy, identifying additional predictors of groundwater levels, and trend removal techniques on predicted potentiometric surfaces. In the DBAS, groundwater wells should be added in areas that exhibit the highest variance to optimize the monitoring network and provide more certainty of hydraulic heads in data sparse areas. Additionally, more data is needed on the spatial distribution of storativities. While it is unknown at this point whether or not GRACE helps improve our understanding of groundwater availability at the aquifer-scale, future work should focus on addressing the uncertainties and sources of error in comparing GRACE to ground-based data. The results of this work fortuitously highlight the possibilities of exploring land subsidence across the DBAS, which potentially could be attributed to groundwater withdrawals, especially near Castle Rock and Parker. Multiple imputation techniques could be added to the spatio-temporal kriging approach to explore estimates of missing data to help capture uncertainty, increase data density, and further improve confidence in estimates. Finally, future work could explore possible influences of climate signals such as El Niño Southern Oscillation (ENSO) or Pacific Decadal Oscillation (PDO) on long-term groundwater availability and sustainability in the DBAS.

REFERENCES CITED

Ahmadi, S.H., and A. Sedghamiz. 2007. Geostatistical analysis of spatial and temporal variations of groundwater level. *Environmental monitoring and assessment* 129 no. 1-3: 277.

Ahmadi, S.H., and A. Sedghamiz. 2008. Application and evaluation of kriging and cokriging methods on groundwater depth mapping. *Environmental Monitoring and Assessment* 138 no. 1-3: 357-368.

Alley, W., T. Reilly, and O. Franke. 1999. Sustainability of ground-water resources: US Geological Survey Circular 1186, 79 p.

Alley, W.M. 2016. AGI's 2016 Critical Issues Forum on groundwater. Golden, CO.

Alley, W.M., and L.F. Konikow. 2015. Bringing GRACE down to earth. *Groundwater*.

Anderman, G., and E. Ackman. 1963. Structure of the Denver-Julesburg Basin and surrounding areas: in Rooky Mtn. *Assoc. Geologists Guidebook to the Geology of the northern Denver Basin and adjacent uplifts*.

Aryaputera, A.W., D. Yang, L. Zhao, and W.M. Walsh. 2015. Very short-term irradiance forecasting at unobserved locations using spatio-temporal kriging. *Solar Energy* 122: 1266-1278.

Banta, E.R. 1989. Hydrologic effects of pumpage from the Denver basin bedrock aquifers of northern El Paso County, Colorado. US Geological Survey.

Banta, E.R., and S.S. Paschke. 2012. Demonstration optimization analyses of pumping from selected Arapahoe aquifer municipal wells in the west-central Denver Basin, Colorado, 2010-2109. *USA: US Geological Survey Scientific Investigations Report* 5140: 2012.

Bauch, N.J., M. Musgrove, B.J. Mahler, and S. Paschke. 2015. The quality of our Nation's waters: water quality in the Denver Basin aquifer system, Colorado, 2003-05. US Geological Survey.

Bayat, B., M. Nasseri, and B. Zahraie. 2015. Identification of long-term annual pattern of meteorological drought based on spatiotemporal methods: evaluation of different geostatistical approaches. *Natural Hazards* 76 no. 1: 515-541.

Beauchamp, M., C. de Fouquet, and L. Malherbe. 2017. Dealing with non-stationarity through explanatory variables in kriging-based air quality maps. *Spatial Statistics* 22: 18-46.

Black & Veatch, Rick Giardina & Associates, HRS Water Consultants, Hydrosphere Resource Consultants, and Mulhern MRE. 2003. South Metro Water Supply Study - Executive Summary. The South Metro Water Supply Study Board.

Breña Naranjo, J.A., A.D. Kendall, and D.W. Hyndman. 2014. Improved methods for satellite based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations. *Geophysical Research Letters* 41 no. 17: 6167-6173.

Brown, T.C. 1999. Past and future freshwater use in the United States. *Gen. Tech. Rep. RMRS-GTR-39. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 47 p.* 39.

Brus, D.J., and G.B. Heuvelink. 2007. Optimization of sample patterns for universal kriging of environmental variables. *Geoderma* 138 no. 1: 86-95.

Butler, J.J., D.O. Whittemore, B.B. Wilson, and G.C. Bohling. 2016. A new approach for assessing the future of aquifers supporting irrigated agriculture. *Geophysical Research Letters* 43 no. 5: 2004-2010.

Castle, S.L., B.F. Thomas, J.T. Reager, M. Rodell, S.C. Swenson, and J.S. Famiglietti. 2014. Groundwater depletion during drought threatens future water security of the Colorado River Basin. *Geophysical research letters* 41 no. 16: 5904-5911.

CDM. 2004. Statewide Water Supply Initiative.

Centennial Water and Sanitation District. 2017. Highlands Ranch Water Supply.

Chen, J., J. Li, Z. Zhang, and S. Ni. 2014. Long-term groundwater variations in Northwest India from satellite gravity measurements. *Global and Planetary Change* 116: 130-138.

Chen, J.L., C.R. Wilson, B.D. Tapley, B. Scanlon, and A. Güntner. 2016. Long-term groundwater storage change in Victoria, Australia from satellite gravity and in situ observations. *Global and Planetary Change* 139: 56-65.

Chen, X., J. Goeke, J.F. Ayers, and S. Summerside. 2003. Observation Wll Network Design for Pumping Tests in Unconfined Aquifers. *JAWRA Journal of the American Water Resources Association* 39 no. 1: 17-32.

Colorado Department of Local Afairs. 2013. Total Population Counts and Change from 2000 to 2010: Cities and Counties.

Colorado Department of Local Afairs. 2018. County and Municipal Population Timeseries.

Colorado Division of Water Resources. 2011. GIS Data Library: Denver Basin Bedrock Data.

Colorado Division of Water Resources. 2017. HydroBase, ed. Colorado Division of Water Resources. Denver, CO.

Colorado Division of Water Resources. 2018. DWR Well Application Permit, ed. Colorado Division of Water Resources. Colorado Information Marketplace.

Colorado Foundation for Water Education. 2006. The Groundwater Puzzle. In *Headwaters*, vol. Summer 2006. Denver, CO: Colorado Foundation for Water Education.

Colorado Session Laws. 1973. Concerning Underground Water. Chapter 441: Water Rights and Irrigation.

Colorado Session Laws. 1985. Concerning Ground Water and Making an Appropriation in Connection Therewith. Chapter 285: Water and Irrigation.

Colorado Water Conservation Board. 2006. SPDSS Phase 3 Task 44.2 Denver Basin Region Water Level Technical Memorandum FINAL 65.

Colorado Water Conservation Board. 2011. Statewide water supply initiative 2010.

Colorado Water Conservation Board. 2015. Colorado's Water Plan.

Colorado Water Insitiute. 2017. Aquifer Storage and Recovery. In *Colorado Water*, vol. 34 Issue 4.

Costelloe, J., T. Peterson, K. Halbert, A. Western, and J. McDonnell. 2015. Groundwater surface mapping informs sources of catchment baseflow. *Hydrology and Earth System Sciences* 19 no. 4: 1599-1613.

Cressie, N., and H.-C. Huang. 1999. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association* 94 no. 448: 1330-1339.

De Iaco, S., S. Maggio, M. Palma, and D. Posa. 2012. Towards an automatic procedure for modeling multivariate space–time data. *Computers & Geosciences* 41: 1-11.

De Iaco, S., D.E. Myers, and D. Posa. 2001. Space–time analysis using a general product–sum model. *Statistics & Probability Letters* 52 no. 1: 21-28.

De Iaco, S., and D. Posa. 2012. Predicting spatio-temporal random fields: Some computational aspects. *Computers & Geosciences* 41: 12-24.

Delbari, M. 2014. Accounting for exhaustive secondary data into the mapping of water table elevation. *Arabian Journal of Geosciences* 7 no. 10: 4221-4233.

Denver Water. 2018. Aquifer Storage and Recovery Study, ed. D. Water.

Derakhshan, H., and O. Leuangthong. 2006. A Review of Separable Spatiotemporal Models of Regionalization.

Dimitrakopoulos, R., and X. Luo. 1994. Spatiotemporal modelling: covariances and ordinary kriging systems. In *Geostatistics for the next century*, 88-93. Springer.

Döll, P., H. Müller Schmied, C. Schuh, F.T. Portmann, and A. Eicker. 2014. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resources Research* 50 no. 7: 5698-5720.

Everett, R. 2014. Groundwater levels in the Denver Basin bedrock aquifers of Douglas County, Colorado, 2011-2013 Scientific Investigations Report 56.

Famiglietti, J.S. 2014. The global groundwater crisis. *Nature Climate Change* 4 no. 11: 945.

Famiglietti, J.S., M. Lo, S.L. Ho, J. Bethune, K.J. Anderson, T.H. Syed, S.C. Swenson, C.R. de Linage, and M. Rodell. 2011. Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters* 38 no. 3: n/a-n/a.

Famiglietti, J.S., and M. Rodell. 2013. Water in the balance. *Science* 340 no. 6138: 1300-1301.

Feng, W., M. Zhong, J.-M. Lemoine, R. Biancale, H.-T. Hsu, and J. Xia. 2013. Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. *Water Resources Research* 49 no. 4: 2110-2118.

Fetter, C.W. 2000. *Applied hydrogeology*: Prentice hall.

Flor, A. 2015. Groundwater Levels in the Denver Basin Bedrock Aquifers, ed. C. D. o. W. Resources.

Flor, A. 2016. Groundwater Levelsin the Denver Basin Bedrock Aquifers, ed. Colorado Division of Water Resources.

Galloway, D.L., D.R. Jones, and S.E. Ingebritsen. 1999. *Land subsidence in the United States*: US Geological Survey.

Gasch, C.K., T. Hengl, B. Gräler, H. Meyer, T.S. Magney, and D.J. Brown. 2015. Spatio-temporal interpolation of soil water, temperature, and electrical conductivity in 3D + T: The Cook Agronomy Farm data set. *Spatial Statistics* 14: 70-90.

Gehman, C.L., D.L. Harry, W.E. Sanford, J.D. Stednick, and N.A. Beckman. 2009. Estimating specific yield and storage change in an unconfined aquifer using temporal gravity surveys. *Water Resources Research* 45 no. 4.

Gleeson, T., W.M. Alley, D.M. Allen, M.A. Sophocleous, Y. Zhou, M. Taniguchi, and J. VanderSteen. 2012. Towards sustainable groundwater use: setting long-term goals, backcasting, and managing adaptively. *Groundwater* 50 no. 1: 19-26.

Graham, G., and G. Vanslyke. 2004. Development of the Regulatory Framework for Denver Basin Aquifers. *The Mountain Geologist*.

Gräler, B., E. Pebesma, and G. Heuvelink. 2016. Spatio-Temporal Interpolation using gstat.

Gundogdu, K.S., and I. Guney. 2007. Spatial analyses of groundwater levels using universal kriging. *Journal of Earth System Science* 116 no. 1: 49-55.

Güntner, A. 2008. Improvement of Global Hydrological Models Using GRACE Data. *Surveys in Geophysics* 29 no. 4: 375-397.

Hachborn, E., A. Berg, J. Levison, and J.T. Ambadan. 2017. Sensitivity of GRACE-derived estimates of groundwater-level changes in southern Ontario, Canada. *Hydrogeology Journal* 25 no. 8: 2391-2402.

HDR Inc., and W.S.W. Consultants. 2015. South Platte Basin Implementation Plan: South Platte Basin Roundtable/Metro Basin Roundtable.

Hecox, E. 2015. Storing Water Underground Holds Promise for South Metro, ed. South Metro Water Supply Authority.

Hengl, T. 2009. *A practical guide to geostatistical mapping*: Hengl.

Hengl, T., G.B. Heuvelink, and A. Stein. 2003. Comparison of kriging with external drift and regression kriging. ITC.

Hengl, T., G.B.M. Heuvelink, M. Perčec Tadić, and E.J. Pebesma. 2012. Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theoretical and Applied Climatology* 107 no. 1: 265-277.

Hengl, T., G.B.M. Heuvelink, and D.G. Rossiter. 2007. About regression-kriging: From equations to case studies. *Computers & Geosciences* 33 no. 10: 1301-1315.

Hengl, T., G.B.M. Heuvelink, and A. Stein. 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* 120 no. 1: 75-93.

Henry, C.M., D.M. Allen, and J. Huang. 2011. Groundwater storage variability and annual recharge using well-hydrograph and GRACE satellite data. *Hydrogeology Journal* 19 no. 4: 741-755.

Heuvelink, G., and D.A. Griffith. 2010. Space–Time Geostatistics for Geography: A Case Study of Radiation Monitoring Across Parts of Germany. 地理学的时空地统计学： 横跨德国部分区域的辐射监测的案例研究. *Geographical Analysis* 42 no. 2: 161-179.

Heuvelink, G.B., E. Pebesma, and B. Gräler. 2015. Space-time geostatistics. *Encyclopedia of GIS*: 1-7.

Hillier, D.E., R.E. Brogden, and P.A. Schneider Jr. 1978. Hydrology of the Arapahoe Aquifer in the Englewood-Castle Rock area south of Denver, Denver Basin, Colorado IMAP.

Houborg, R., M. Rodell, B. Li, R. Reichle, and B.F. Zaitchik. 2012. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research* 48 no. 7.

Hu, D., H. Shu, H. Hu, and J. Xu. 2017. Spatiotemporal regression Kriging to predict precipitation using time-series MODIS data. *Cluster Computing* 20 no. 1: 347-357.

Huang, J., J. Halpenny, W. van der Wal, C. Klatt, T.S. James, and A. Rivera. 2012. Detectability of groundwater storage change within the Great Lakes Water Basin using GRACE. *Journal of Geophysical Research: Solid Earth* 117 no. B8: n/a-n/a.

Huang, J., G. Pavlic, A. Rivera, D. Palombi, and B. Smerdon. 2016. Mapping groundwater storage variations with GRACE: a case study in Alberta, Canada. *Hydrogeology Journal*: 1-18.

Jones, P.A., and T. Cech. 2009. *Colorado water law for non-lawyers*: University Press of Colorado.

Júnez-Ferreira, H.E., and G.S. Herrera. 2013. A geostatistical methodology for the optimal design of space–time hydraulic head monitoring networks and its application to the Valle de Querétaro aquifer. *Environmental Monitoring and Assessment* 185 no. 4: 3527-3549.

Katpatal, Y.B., C. Rishma, and C.K. Singh. 2017. Sensitivity of the Gravity Recovery and Climate Experiment (GRACE) to the complexity of aquifer systems for monitoring of groundwater. *Hydrogeology Journal*: 1-11.

Konikow, L.F. 2013. Groundwater depletion in the United States (1900− 2008). US Geological Survey.

Konikow, L.F. 2015. Long-Term Groundwater Depletion in the United States. *Groundwater* 53 no. 1: 2-9.

Konikow, L.F., and E. Kendy. 2005. Groundwater depletion: A global problem. *Hydrogeology Journal* 13 no. 1: 317-320.

Kyriakidis, P.C., and A.G. Journel. 1999. Geostatistical space–time models: a review. *Mathematical geology* 31 no. 6: 651-684.

Landerer, F., and S. Swenson. 2012. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research* 48 no. 4.

Leblanc, M.J., P. Tregoning, G. Ramillien, S.O. Tweed, and A. Fakes. 2009. Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia. *Water Resources Research* 45 no. 4: n/a-n/a.

Lee, H., R.E. Beighley, D. Alsdorf, H.C. Jung, C.K. Shum, J. Duan, J. Guo, D. Yamazaki, and K. Andreadis. 2011. Characterization of terrestrial water dynamics in the Congo Basin using GRACE and satellite radar altimetry. *Remote Sensing of Environment* 115 no. 12: 3530-3538.

Li, B., M. Rodell, B.F. Zaitchik, R.H. Reichle, R.D. Koster, and T.M. van Dam. 2012. Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology* 446: 103-115.

Li, J., and A.D. Heap. 2008. A review of spatial interpolation methods for environmental scientists.

Li, L., and P. Revesz. 2004. Interpolation methods for spatio-temporal geographic data. *Computers, Environment and Urban Systems* 28 no. 3: 201-227.

Long, D., L. Longuevergne, and B.R. Scanlon. 2015. Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resources Research* 51 no. 4: 2574-2594.

Long, D., B.R. Scanlon, L. Longuevergne, A.Y. Sun, D.N. Fernando, and H. Save. 2013. GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. *Geophysical Research Letters* 40 no. 13: 3395-3401.

Longuevergne, L., B.R. Scanlon, and C.R. Wilson. 2010. GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA. *Water Resources Research* 46 no. 11: n/a-n/a.

Longuevergne, L., C. Wilson, B. Scanlon, and J. Crétaux. 2013. GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage. *Hydrology and Earth System Sciences* 17 no. 12: 4817-4830.

Maheswaran, R., and R. Khosa. 2013. Long term forecasting of groundwater levels with evidence of non-stationary and nonlinear characteristics. *Computers & Geosciences* 52: 422-436.

Major, T.J., S.G. Robson, J.C. Romero, and S. Zawistowski. 1983. Hydrogeologic data from parts of the Denver basin, Colorado. *United States Geological Survey Open File Report*: 83-274.

Martínez, W.A., C.E. Melo, and O.O. Melo. 2017. Median Polish Kriging for space–time analysis of precipitation. *Spatial Statistics* 19: 1-20.

Mateu, J. 2015. *Spatial and spatio-temporal geostatistical modeling and kriging*: John Wiley & Sons.

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. US Geological Survey.

Moore, J., R. Raynolds, and P. Barkmann. 2004. Groundwater mining of bedrock aquifers in the Denver Basin–past, present, and future. *Environmental Geology* 47 no. 1: 63-68.

Moore, J.E., R.G. Raynolds, and M. Dechesne. 2007. Bedrock aquifers and population growth in the Denver Basin, Colorado, USA. *Episodes* 30 no. 2: 115.

Morgan, P. 2007. Potential for Subsidence Associated with Groundwater Withdrawal in the Denver Basin. In *2007 GSA Denver Annual Meeting*.

Myers, D.E. 2004. Estimating and modeling space-time variograms. In *Proceedings of the joint meeting of TIES-2004 and ACCURACY-2004*.

Myers, D.E., and A. Journel. 1990. Variograms with zonal anisotropies and noninvertible kriging systems. *Mathematical Geology* 22 no. 7: 779-785.

Nanteza, J., C. de Linage, B. Thomas, and J. Famiglietti. 2016. Monitoring groundwater storage changes in complex basement aquifers: An evaluation of the GRACE satellites over East Africa. *Water Resources Research* 52 no. 12: 9542-9564.

Ohmer, M., T. Liesch, N. Goeppert, and N. Goldscheider. 2017. On the optimal selection of interpolation methods for groundwater contouring: An example of propagation of uncertainty regarding inter-aquifer exchange. *Advances in Water Resources* 109: 121-132.

Parker Water and Sanitation District. 2016. Rueter-Hess Reservoir.

Paschke, S.S., E.R. Banta, J.A. Dupree, J.P. Capesius, and D.W. Litke. 2011. Groundwater availability of the Denver Basin aquifer system, Colorado. US Geological Survey USGS Professional Paper 1770 274.

Pebesma, E. 2012. spacetime: Spatio-temporal data in r. *Journal of Statistical Software* 51 no. 7: 1-30.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30 no. 7: 683-691.

PRISM Climate Group. 2016. 30-Year Normal Precipitation: Annual, ed. Oregon State University.

Raynolds, R.G. 2004. Stratigraphy and water levels in the Arapahoe aquifer, Douglas County area, Denver Basin, Colorado. *The Mountain Geologist*.

Reager, J.T., B.F. Thomas, and J.S. Famiglietti. 2014. River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nature Geosci* 7 no. 8: 588-592.

Reghunath, R., T.S. Murthy, and B. Raghavan. 2005. Time series analysis to monitor and assess water resources: A moving average approach. *Environmental Monitoring and Assessment* 109 no. 1-3: 65-72.

Reilly, T.E., and A.W. Harbaugh. 2004. *Guidelines for evaluating ground-water flow models*: US Department of the Interior, US Geological Survey.

Richey, A.S., B.F. Thomas, M.-H. Lo, J.T. Reager, J.S. Famiglietti, K. Voss, S. Swenson, and M. Rodell. 2016. Reply to comment by Sahoo et al. on "Quantifying renewable groundwater stress with GRACE". *Water Resources Research*: n/a-n/a.

Richey, A.S., B.F. Thomas, M.H. Lo, J.T. Reager, J.S. Famiglietti, K. Voss, S. Swenson, and M. Rodell. 2015. Quantifying renewable groundwater stress with GRACE. *Water Resources Research*.

Robson, S.G. 1987. Bedrock aquifers in the Denver Basin, Colorado; a quantitative water-resources appraisal. US Geological Survey.

Robson, S.G., and E.R. Banta. 1995. Ground water atlas of the United States: Segment 2, Arizona, Colorado, New Mexico, Utah. US Geological Survey, .

Rodell, M., J. Chen, H. Kato, J.S. Famiglietti, J. Nigro, and C.R. Wilson. 2007. Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal* 15 no. 1: 159-166.

Rodell, M., I. Velicogna, and J.S. Famiglietti. 2009. Satellite-based estimates of groundwater depletion in India. *Nature* 460 no. 7258: 999-1002.

Rouhani, S. 1986. Comparative Study of Ground-Water Mapping Techniques. *Ground Water* 24 no. 2: 207-216.

Rouhani, S., and T.J. Hall. 1989. Space-Time Kriging of Groundwater Data. In *Geostatistics: Proceedings of the Third International Geostatistics Congress September 5–9, 1988, Avignon, France*, ed. M. Armstrong, 639-650. Dordrecht: Springer Netherlands.

Sakumura, C., S. Bettadpur, and S. Bruinsma. 2014. Ensemble prediction and intercomparison analysis of GRACE time variable gravity field models. *Geophysical Research Letters* 41 no. 5: 1389-1397.

Sankarasubramanian, A., J.L. Sabo, K.L. Larson, S.B. Seo, T. Sinha, R. Bhowmik, A.R. Vidal, K. Kunkel, G. Mahinthakumar, E.Z. Berglund, and J. Kominoski. 2017. Synthesis of public water supply use in the United States: Spatio-temporal patterns and socio-economic controls. *Earth's Future* 5 no. 7: 771-788.

Savoca, M.E., G.B. Senay, M.A. Maupin, J.F. Kenny, and C.A. Perry. 2013. Actual evapotranspiration modeling using the operational Simplified Surface Energy Balance (SSEBop) approach. US Geological Survey.

Scanlon, B.R., C.C. Faunt, L. Longuevergne, R.C. Reedy, W.M. Alley, V.L. McGuire, and P.B. McMahon. 2012a. Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proceedings of the national academy of sciences* 109 no. 24: 9320-9325.

Scanlon, B.R., L. Longuevergne, and D. Long. 2012b. Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *Water Resources Research* 48 no. 4: n/a-n/a.

Schlager, E. 2006. Challenges of governing groundwater in US western states. *Hydrogeology journal* 14 no. 3: 350-360.

Seyoum, W.M., and A.M. Milewski. 2016. Monitoring and comparison of terrestrial water storage changes in the northern high plains using GRACE and in-situ based integrated hydrologic model estimates. *Advances in Water Resources* 94: 31-44.

Simpson, H.D., and C. Lite. 1998. Denver Basin and South Platte River Basin Technical Study Senate Bill 96-074. Special Water Committee.

South Metro Water Supply Authority. 2017. WISE Partnership delivers water, marks new era of cooperation.

Strassberg, G., B.R. Scanlon, and D. Chambers. 2009. Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States. *Water Resources Research* 45 no. 5.

Sun, A.Y. 2013. Predicting groundwater level changes using GRACE data. *Water Resources Research* 49 no. 9: 5900-5912.

Sun, A.Y., R. Green, M. Rodell, and S. Swenson. 2010. Inferring aquifer storage parameters using satellite and in situ measurements: Estimation under uncertainty. *Geophysical Research Letters* 37 no. 10.

Sun, Y., S. Kang, F. Li, and L. Zhang. 2009. Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. *Environmental Modelling & Software* 24 no. 10: 1163-1170.

Swenson, S., and J. Wahr. 2002. Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity. *Journal of Geophysical Research: Solid Earth (1978–2012)* 107 no. B9: ETG 3-1-ETG 3-13.

Swenson, S., and J. Wahr. 2006. Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters* 33 no. 8.

Swenson, S., and J. Wahr. 2007. Multi-sensor analysis of water storage variations of the Caspian Sea. *Geophysical Research Letters* 34 no. 16: n/a-n/a.

Swenson, S., and J. Wahr. 2009. Monitoring the water balance of Lake Victoria, East Africa, from space. *Journal of Hydrology* 370 no. 1–4: 163-176.

Swenson, S., P.J.F. Yeh, J. Wahr, and J. Famiglietti. 2006. A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois. *Geophysical Research Letters* 33 no. 16: n/a-n/a.

Ta′any, R.A., A.B. Tahboub, and G.A. Saffarini. 2009. Geostatistical analysis of spatiotemporal variability of groundwater level fluctuations in Amman–Zarqa basin, Jordan: a case study. *Environmental Geology* 57 no. 3: 525-535.

Tadic, J.M., X. Qiu, S. Miller, and A.M. Michalak. 2017. Spatio-temporal approach to moving window block kriging of satellite data v1. 0. *Geoscientific Model Development* 10 no. 2: 709.

Tadić, J.M., X. Qiu, V. Yadav, and A.M. Michalak. 2015. Mapping of satellite Earth observations using moving window block kriging. *Geoscientific Model Development* 8 no. 10: 3311-3319.

Taylor, C.J., and W.M. Alley. 2001. *Ground-water-level monitoring and the importance of long-term water-level data*: Geological Survey (USGS).

Team, R.C. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2012) ISBN 3-900051-07-0. *www. R-project. org*.

Thomas, A.C., J.T. Reager, J.S. Famiglietti, and M. Rodell. 2014. A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophysical Research Letters* 41 no. 5: 1537-1545.

Tiwari, V., J. Wahr, and S. Swenson. 2009. Dwindling groundwater resources in northern India, from satellite gravity observations. *Geophysical Research Letters* 36 no. 18.

Topper, R.E., and B. Raynolds. 2007. *Denver Basin Groundwater*: Colorado Foundation for Water Education.

Town of Castle Rock Colorado. 2018. Long-Term Water.

U.S. Census Bureau, P.D. 2005a. City and Town Postcensal Tables: 1990-2000, Population Estimates for States, Counties, Places and Minor Civil Divisions: Annual Time Series, April 1, 1990 Census to July 1, 2000 Estimate

U.S. Census Bureau, P.D. 2005b. City and Town Postcensal Tables: 1990-2000.

U.S. Census Bureau, P.D. 2017a. Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2016 Release Dates: For the United States, regions, divisions, states, and Puerto Rico Commonwealth, December 2016. For counties, municipios, metropolitan statistical areas, micropolitan statistical areas, metropolitan divisions, and combined statistical areas, March 2017. For cities and towns (incorporated places and minor civil divisions), May 2017.

U.S. Census Bureau, P.D. 2017b. Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2016.

Van der Gun, J., and A. Lipponen. 2010. Reconciling groundwater storage depletion due to pumping with sustainability. *Sustainability* 2 no. 11: 3418-3435.

Varouchakis, E.A., and D.T. Hristopulos. 2017. Comparison of spatiotemporal variogram functions based on a sparse dataset of groundwater level variations. *Spatial Statistics*.

Varouchakis, E., and D. Hristopulos. 2013. Comparison of stochastic and deterministic methods for mapping groundwater level spatial variability in sparsely monitored basins. *Environmental monitoring and assessment* 185 no. 1: 1-19.

Voss, K.A., J.S. Famiglietti, M. Lo, C. de Linage, M. Rodell, and S.C. Swenson. 2013. Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water Resources Research* 49 no. 2: 904-914.

Wahr, J., M. Molenaar, and F. Bryan. 1998. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth (1978–2012)* 103 no. B12: 30205-30229.

Wahr, J., S. Swenson, and I. Velicogna. 2006. Accuracy of GRACE mass estimates. *Geophysical Research Letters* 33 no. 6: n/a-n/a.

Wang, J., Y. Ge, G. Heuvelink, and C. Zhou. 2015. Upscaling In Situ Soil Moisture Observations to Pixel Averages with Spatio-Temporal Geostatistics. *Remote Sensing* 7 no. 9: 11372.

Wang, X., C. de Linage, J. Famiglietti, and C.S. Zender. 2011. Gravity Recovery and Climate Experiment (GRACE) detection of water storage changes in the Three Gorges Reservoir of China and comparison with in situ measurements. *Water Resources Research* 47 no. 12.

Watkins, M.M., D.N. Wiese, D.N. Yuan, C. Boening, and F.W. Landerer. 2015. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth* 120 no. 4: 2648-2671.

White, C.B. 1995. Sustainable Use of the Denver Basin. In *Sustainable Use of the West's Water (Summer Conference, June 12-14)*.

Woodard, L.L., W. Sanford, and R.G. Raynolds. 2002. Stratigraphic variability of specific yield within bedrock aquifers of the Denver Basin, Colorado. *Rocky Mountain Geology* 37 no. 2: 229-236.

Wouters, B., J. Bonin, D. Chambers, R. Riva, I. Sasgen, and J. Wahr. 2014. GRACE, time-varying gravity, Earth system dynamics and climate change. *Reports on Progress in Physics* 77 no. 11: 116801.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, and J. Meng. 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. *Journal of Geophysical Research: Atmospheres* 117 no. D3.

Xiao, Y., X. Gu, S. Yin, J. Shao, Y. Cui, Q. Zhang, and Y. Niu. 2016. Geostatistical interpolation model selection based on ArcGIS and spatio-temporal variability analysis of groundwater level in piedmont plains, northwest China. *SpringerPlus* 5 no. 1: 425.

Xu, J., and H. Shu. 2015. Spatio-temporal kriging based on the product-sum model: some computational aspects. *Earth Science Informatics* 8 no. 3: 639-648.

Xu, T., A.J. Valocchi, J. Choi, and E. Amir. 2012. Improving groundwater flow model prediction using complementary data-driven models. In *XIX International Conference on Computational Methods in Water Resources, Univ. of Ill., Urbana-Champaign, Ill.*

Yang, P., and Y. Chen. 2015. An analysis of terrestrial water storage variations from GRACE and GLDAS: The Tianshan Mountains and its adjacent areas, central Asia. *Quaternary International* 358: 106-112.

Yang, Y., J. Wu, and G. Christakos. 2015. Prediction of soil heavy metal distribution using Spatiotemporal Kriging with trend model. *Ecological Indicators* 56: 125-133.

Yao, L., Z. Huo, S. Feng, X. Mao, S. Kang, J. Chen, J. Xu, and T.S. Steenhuis. 2014. Evaluation of spatial interpolation methods for groundwater level in an arid inland oasis, northwest China. *Environmental earth sciences* 71 no. 4: 1911-1924.

Zeng, Z.-C., L. Lei, K. Strong, D.B. Jones, L. Guo, M. Liu, F. Deng, N.M. Deutscher, M.K. Dubey, and D.W. Griffith. 2017. Global land mapping of satellite-observed CO2 total columns using spatio-temporal geostatistics. *International Journal of Digital Earth* 10 no. 4: 426-456.

Zeng, Z., L. Lei, S. Hou, F. Ru, X. Guan, and B. Zhang. 2014. A Regional Gap-Filling Method Based on Spatiotemporal Variogram Model of CO2 Columns. *IEEE Transactions on Geoscience and Remote Sensing* 52 no. 6: 3594-3603.

Zmijewski, K., and R. Becker. 2014. Estimating the Effects of Anthropogenic Modification on Water Balance in the Aral Sea Watershed Using GRACE: 2003–12. *Earth Interactions* 18 no. 3: 1-16.

APPENDIX A

SUPPORTING INFORMATION FOR CHAPTER 3

A.1 Spatio-Temporal Semivariogram Plots and Tables

Groundwater levels are assumed to follow a random function model, consisting of a trend and residual component that can be modeled separately. The global trend in groundwater levels is assumed constant in time and consists of only spatial components. A regression model is constructed for each aquifer and contains various combinations of predictors. The trend calculated from the regression model is subtracted from the groundwater level observations to yield the residual component. The resulting residual component is used to construct the spatio-temporal semivariogram from which a model is fit and used to infer the kriging. The predicted groundwater levels are taken as the sum of the trend component and the kriged residuals.

The sample spatio-temporal semivariograms determined from the groundwater level residuals and the sum-metric model fit to the residuals are shown in the following plots for each aquifer. The parameters for each fitted sum-metric model are shown in the following tables.

A.1.1 Dawson Aquifer

The groundwater level trend (m) is calculated as,

$$m(s) = 0.2438AquiferTop + 0.01403Lon - 10770Lat + 878.2Lat^2 + 3037Lat^3 - 2246Lat^4 - 2638 \quad \text{(A-1)}$$

where AquiferTop is the top elevation contour of the aquifer at the longitude (Lon) and latitude (Lat) of spatial location s.

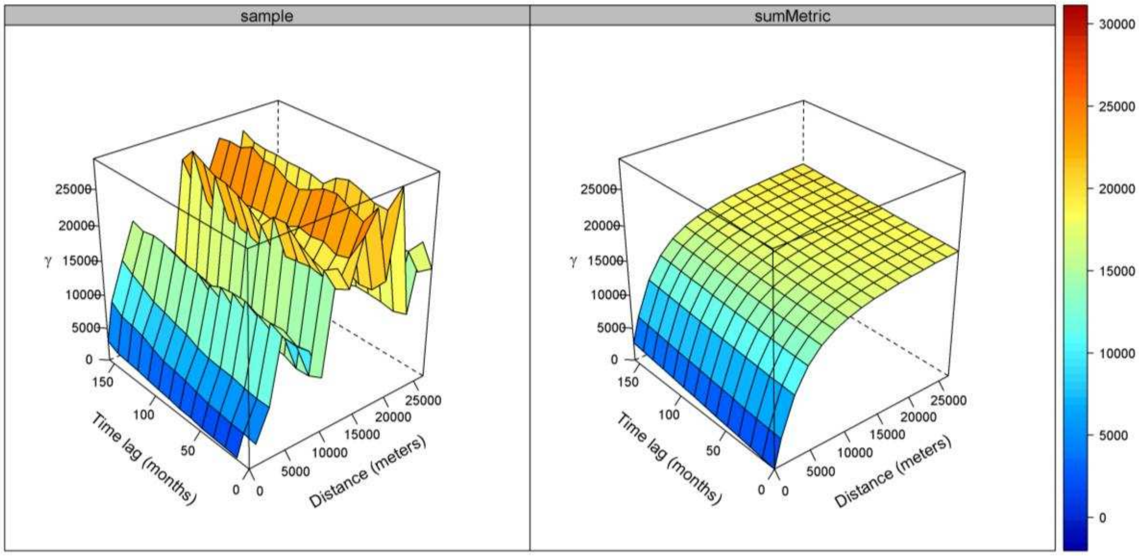

Figure A-1 Sample (left) and fitted (right) spatio-temporal semivariogram surfaces of the residuals for the Dawson Aquifer.

Table A-1 Parameters for fitted sum-metric model semivariogram for the Dawson Aquifer.

	Model	pSill	Range	Nugget	stAni	RMSE	MAE
Spatial	**Exponential**	14127	3185	0	29	4811	3815
Temporal	**Linear**	1314	156	0			
Joint	**Exponential**	4994	13435	0			

A.1.2 Denver Aquifer

The groundwater level trend (m) is calculated as,

$$m(s) = 4169.70PCA2 - 397.94PCA2^2 - 3956.72PCA2^3 - 3556.01Lon - 1626.14Lon^2 - 17900.46Lat - 99.85Lat^2 + 3947.86Lat^3 - 5905.77 \quad \text{(A-2)}$$

where PCA2 is Principal Component 2 from a Principal Component Analysis using land surface and aquifer top elevations at the longitude (Lon) and latitude (Lat) of spatial location s.

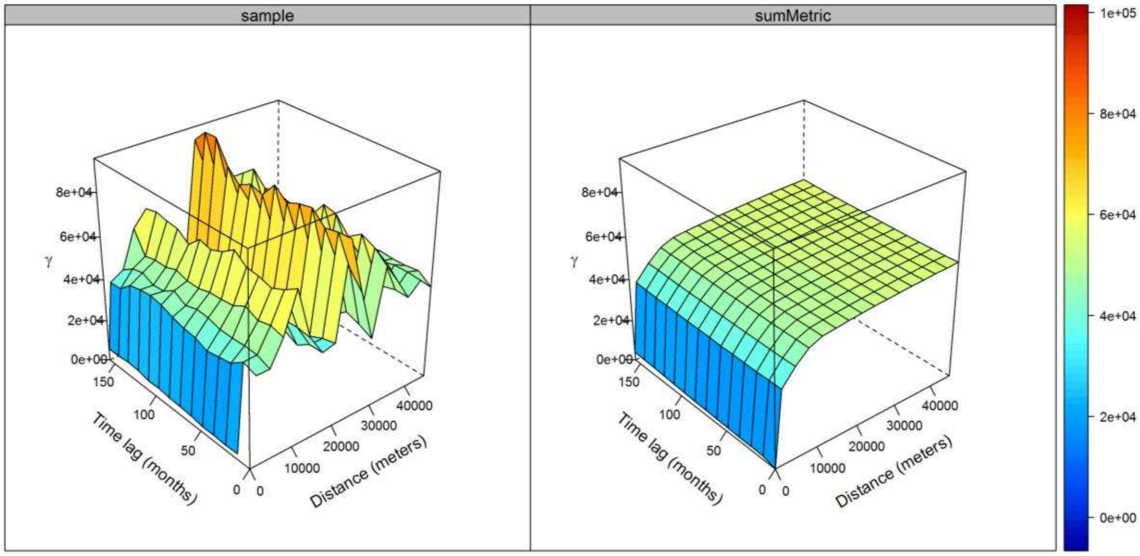

Figure A-2 Sample (left) and fitted (right) spatio-temporal semivariogram surfaces of the residuals for the Denver Aquifer.

Table A-2 Parameters for fitted sum-metric model semivariogram for the Denver Aquifer.

	Model	pSill	Range	Nugget	stAni	RMSE	MAE
Spatial	**Spherical**	15341	10731	29720	16	13273	10258
Temporal	**Gaussian**	4931	244	0			
Joint	**Exponential**	9821	25091	1600			

A.1.3 Arapahoe Aquifer

The groundwater level trend (m) is calculated as,

$$m(s) = 6.919PCA1 + 0.003509Lon\ + 2504 \qquad \text{(A-3)}$$

where PCA1 is Principal Component 1 from a Principal Component Analysis using land surface and aquifer top elevations at the longitude (Lon) of spatial location s.

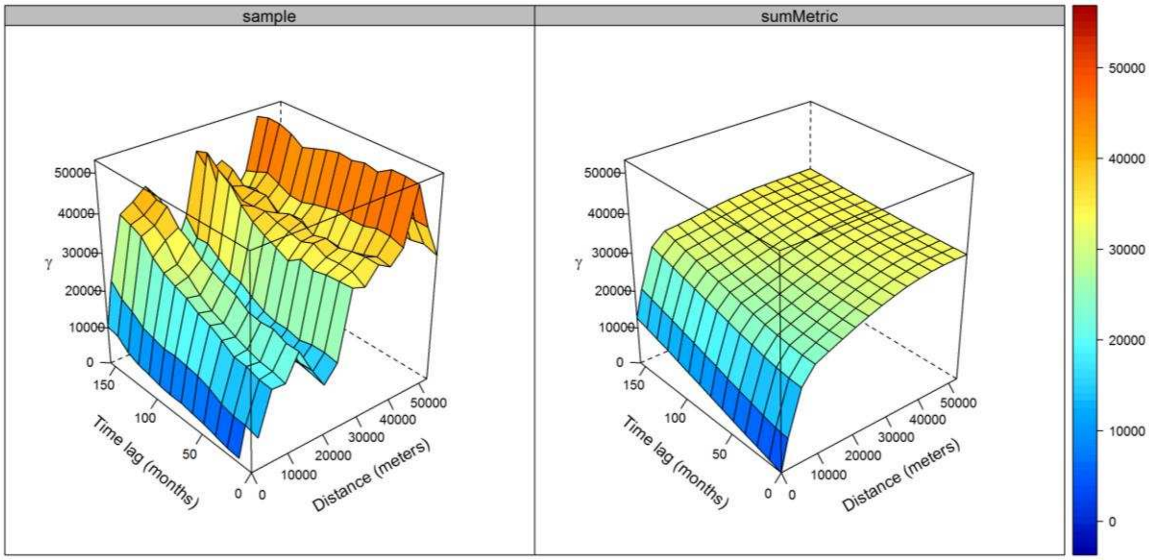

Figure A-3 Sample (left) and fitted (right) spatio-temporal semivariogram surfaces of the residuals for the Arapahoe Aquifer

Table A-3 Parameters for fitted sum-metric model semivariogram for the Arapahoe Aquifer

	Model	pSill	Range	Nugget	stAni	RMSE	MAE
Spatial	**Spherical**	18341	8121	1150	175	6883	4928
Temporal	**Gaussian**	7611	290	0			
Joint	**Spherical**	13841	46711	0			

A.1.4 Laramie-Fox Hills Aquifer

The groundwater level trend is calculated as,

$$m(s) = 6.156PCA1 - 0.001826Lon\ + 5330 \quad \text{(A-4)}$$

where PCA1 is Principal Component 1 from a Principal Component Analysis using land surface and aquifer top elevations at the longitude (Lon) of spatial location s.

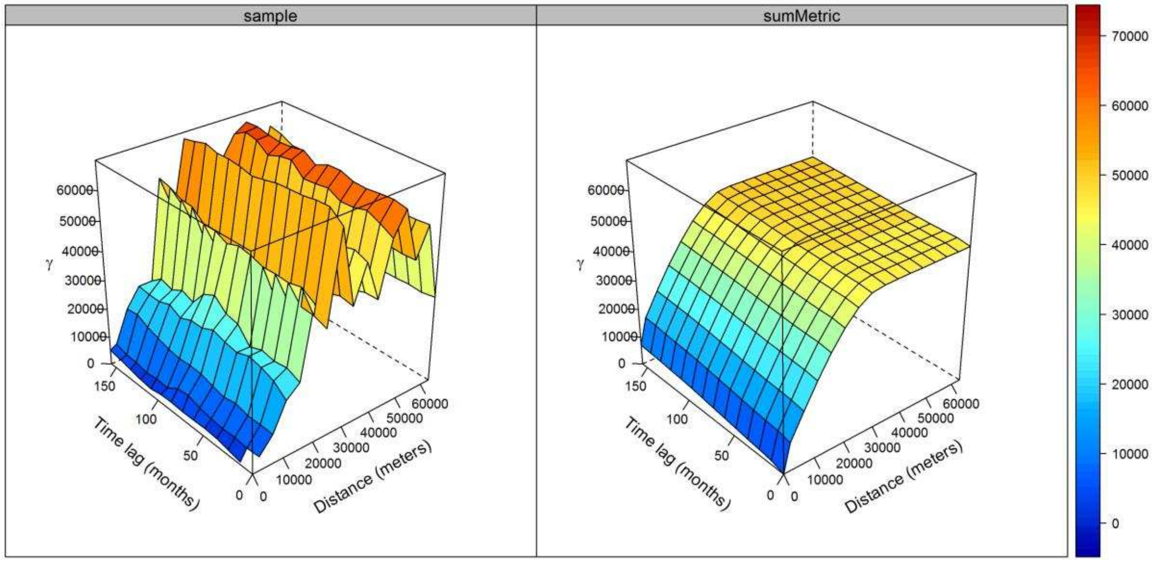

Figure A-4 Sample (left) and fitted (right) spatio-temporal semivariogram surfaces of the residuals for the Laramie-Fox Hills Aquifer.

Table A-4 Parameters for fitted sum-metric model semivariogram for the Laramie-Fox Hills Aquifer.

	Model	pSill	Range	Nugget	stAni	RMSE	MAE
Spatial	**Spherical**	38281	31781	5020	280	10821	9250
Temporal	**Gaussian**	4821	145	640			
Joint	**Spherical**	2661	34271	470			

A.2 Kriging Cross-Validation Plots

"Leave-one-out" cross-validation was used to assess the kriging model performance that would lead to the best predictions. The cross-validation is applied to the existing groundwater level samples. During cross-validation, each spatio-temporal data location is removed one at a time, while the remaining dataset are used to predict the associated value. Measured and predicted groundwater elevations are then compared for goodness of fit.

A.2.1 Dawson Aquifer

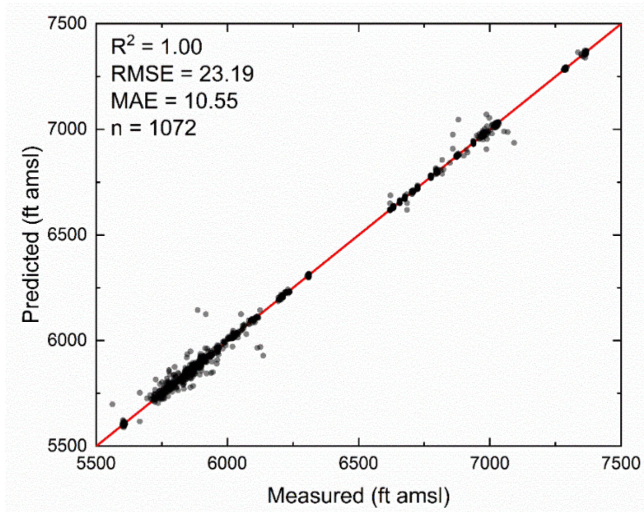

Figure A-5 Dawson Aquifer cross-validation scatterplot of potentiometric surface elevations at well observation locations showing the measured surface elevation compared to the predicted elevation using spatio-temporal kriging.

A.2.2 Denver Aquifer

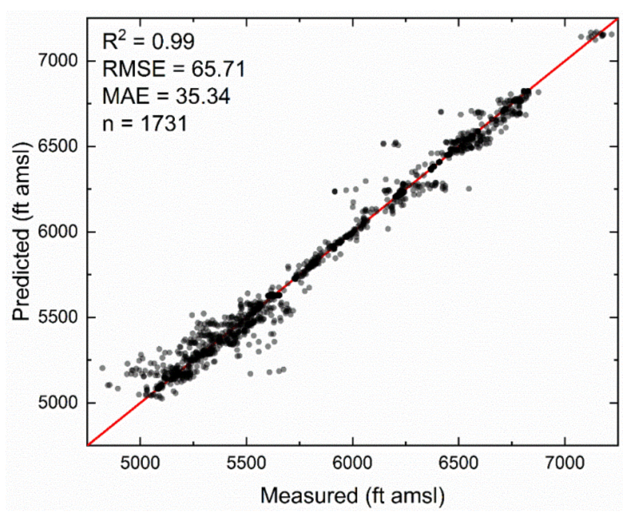

Figure A-6 Denver Aquifer cross-validation scatterplot of potentiometric surface elevations at well observation locations showing the measured surface elevation compared to the predicted elevation using spatio-temporal kriging.

A.2.3 Arapahoe Aquifer

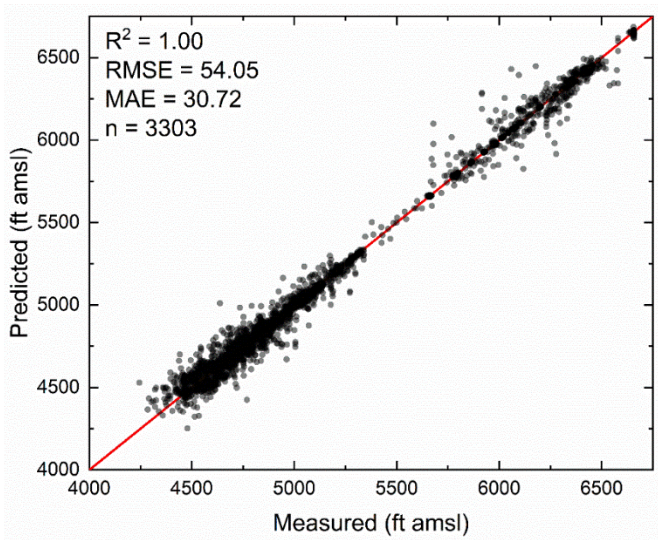

Figure A-7 Arapahoe Aquifer cross-validation scatterplot of potentiometric surface elevations at well observation locations showing the measured surface elevation compared to the predicted elevation using spatio-temporal kriging.

A.2.4 Laramie-Fox Hills Aquifer

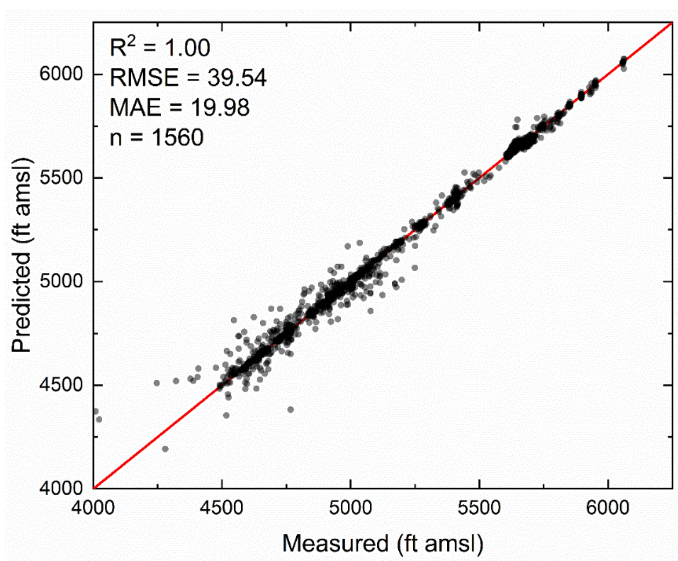

Figure A-8 Laramie-Fox Hills Aquifer cross-validation scatterplot of potentiometric surface elevations at well observation locations showing the measured surface elevation compared to the predicted elevation using spatio-temporal kriging.

A.3 Potentiometric Surfaces and Standard Deviations for Select Years

A.3.1 Dawson Aquifer

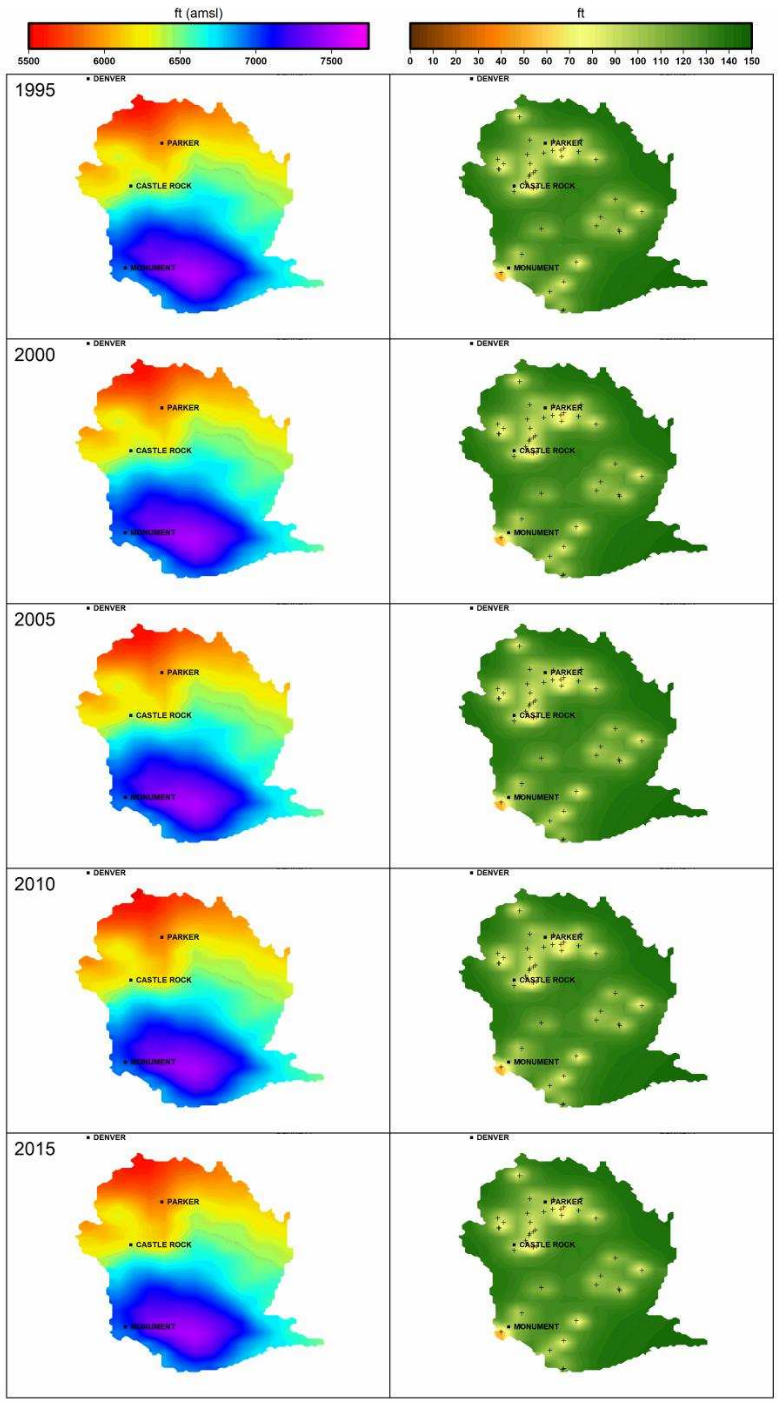

Figure A-9 Predicted potentiometric surfaces (left column) and standard deviation maps (right column) for the Dawson Aquifer. All maps shown are for the month of April in select years. The black '+' on the standard deviation maps show the locations of groundwater wells used in the analysis.

A.3.2 Denver Aquifer

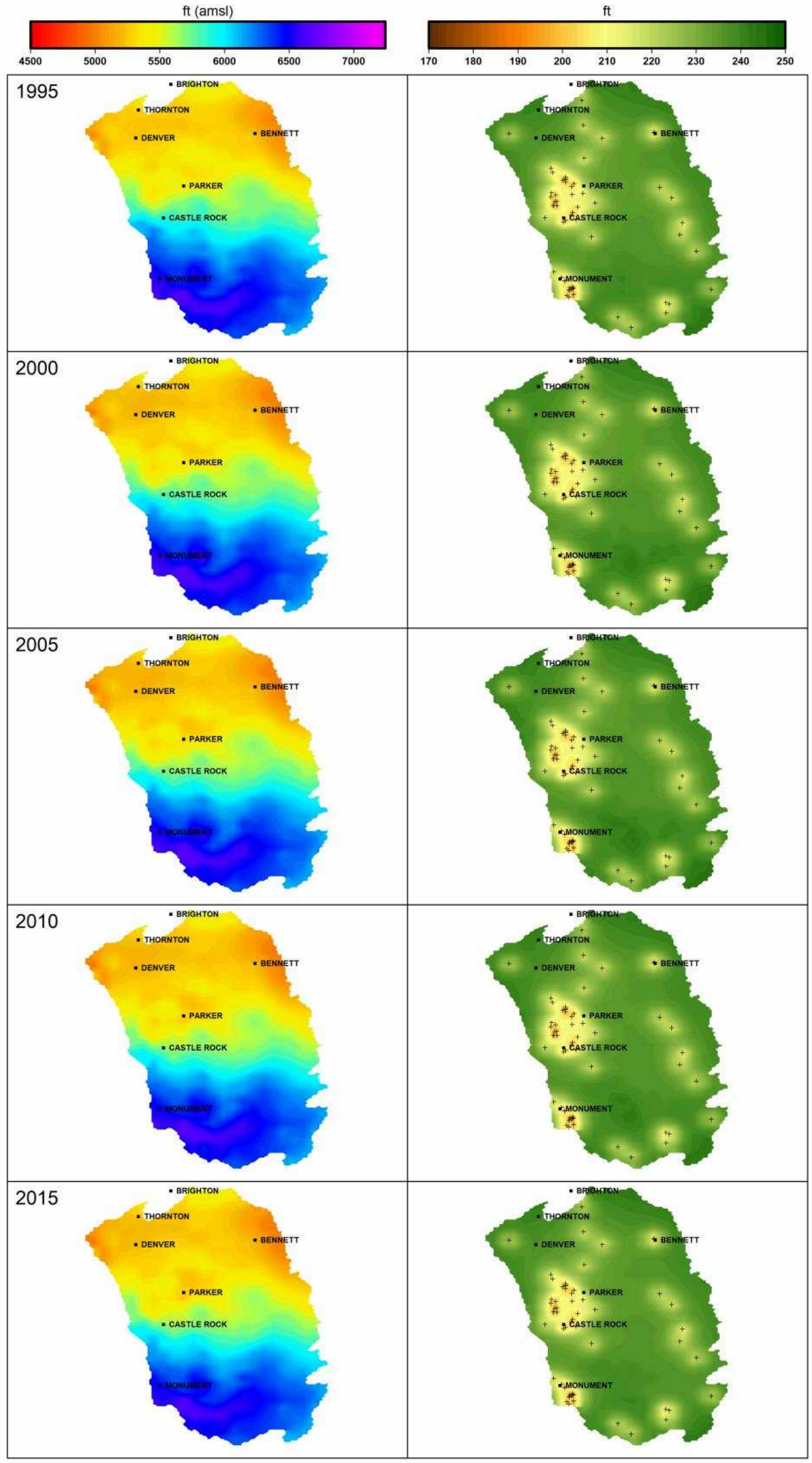

Figure A-10 Predicted potentiometric surfaces (left column) and standard deviation maps (right column) for the Denver Aquifer. All maps shown are for the month of April in select years. The black '+' on the standard deviation maps show the locations of groundwater wells used in the analysis.

A.3.3 Arapahoe Aquifer

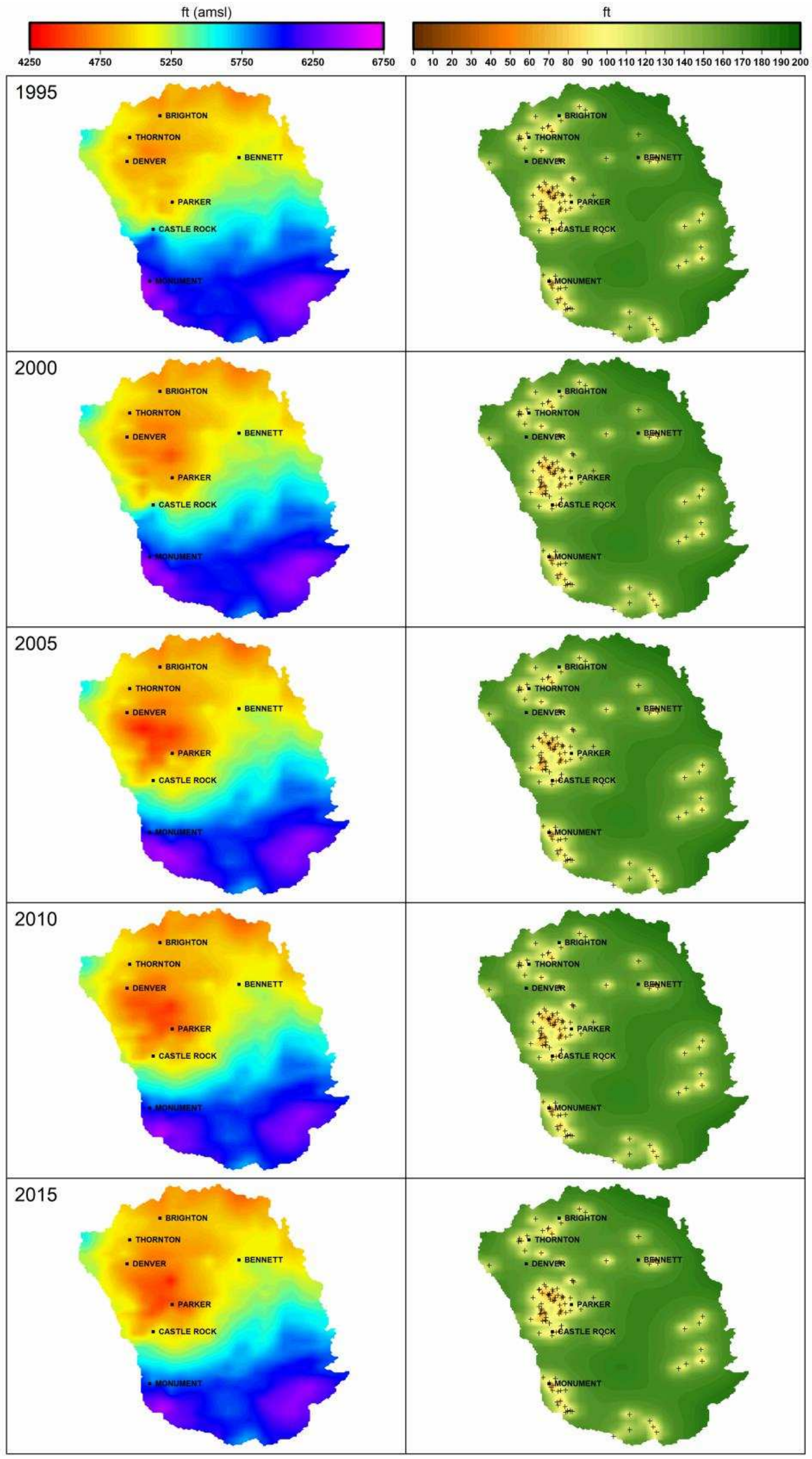

Figure A-11 Predicted potentiometric surfaces (left column) and standard deviation maps (right column) for the Arapahoe Aquifer. All maps shown are for the month of April in select years. The black '+' on the standard deviation maps show the locations of groundwater wells used in the analysis.

A.3.4 Laramie-Fox Hills Aquifer

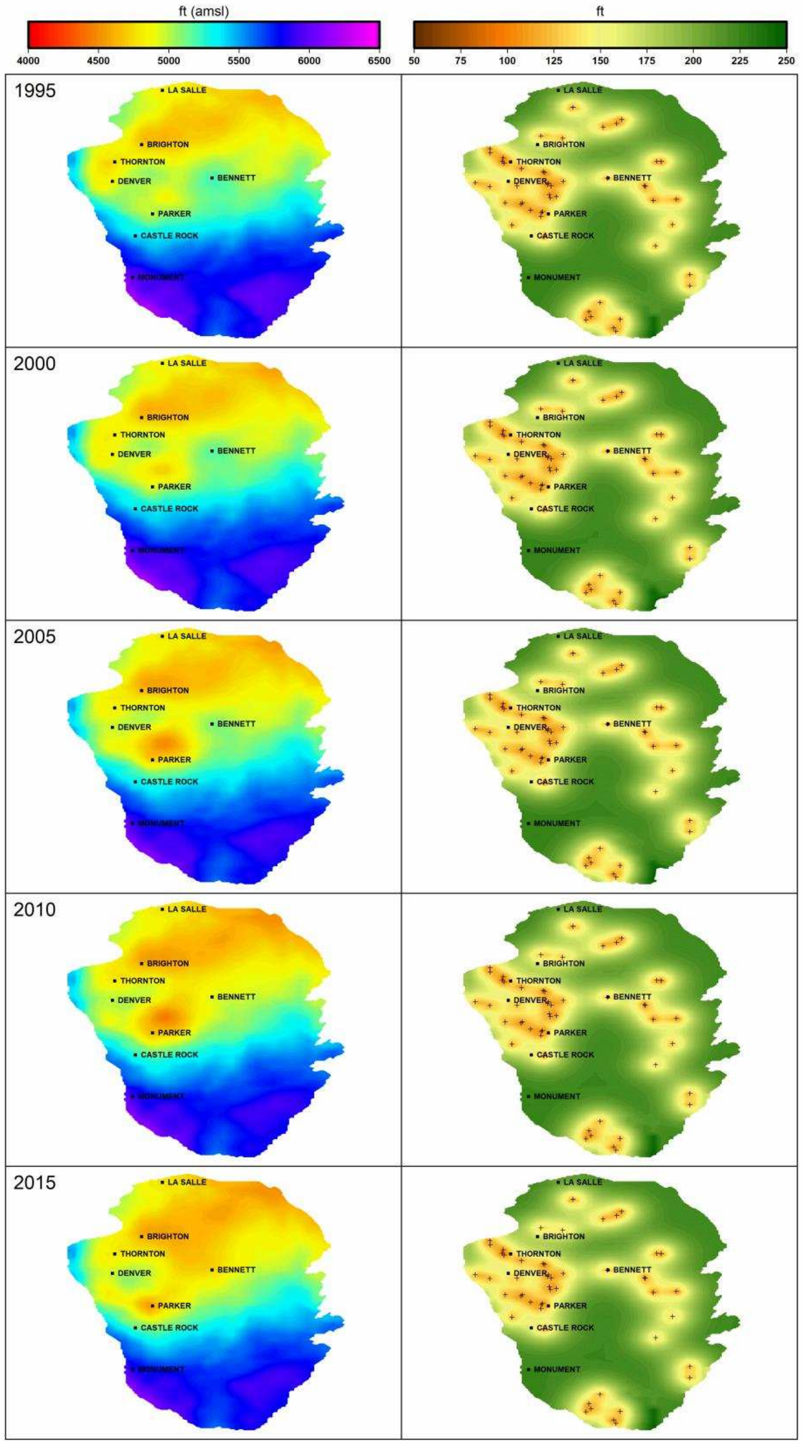

Figure A-12 Predicted potentiometric surfaces (left column) and standard deviation maps (right column) for the Laramie-Fox Hills Aquifer. All maps shown are for the month of April in select years. The black '+' on the standard deviation maps show the locations of groundwater wells used in the analysis.

A.4 Aquifer Properties (Sy and Ssb) Locations

A.4.1 Dawson Aquifer

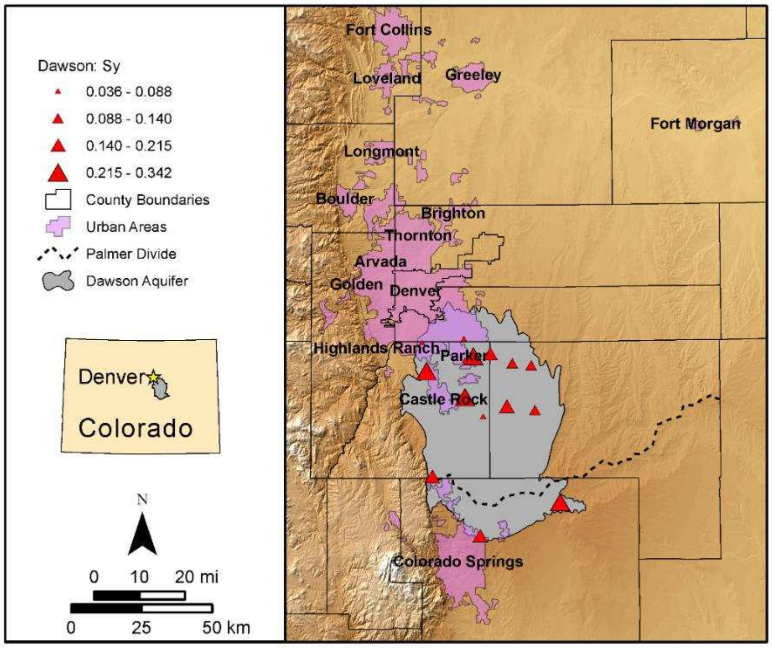

Figure A-13 Location of aquifer data used in storage calculations for the Dawson Aquifer.

A.4.2 Denver Aquifer

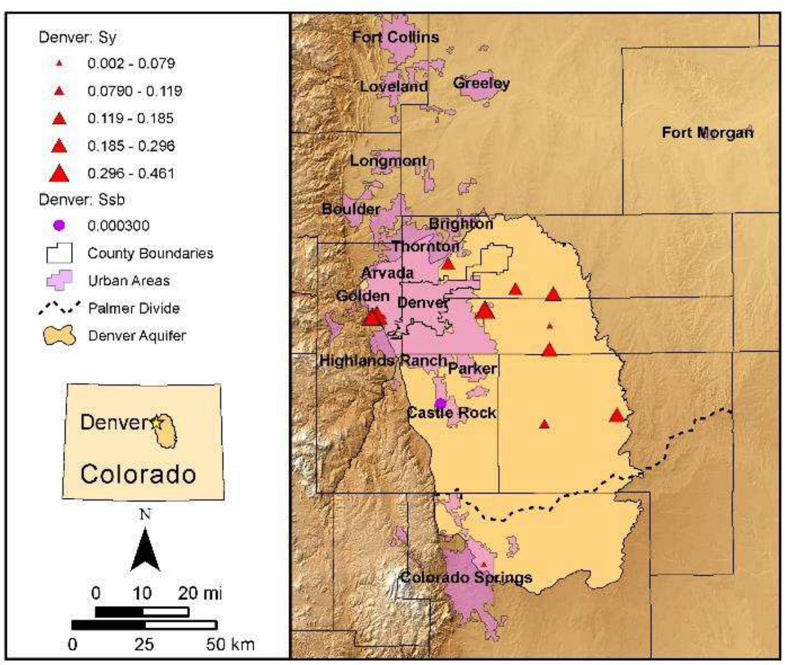

Figure A-14 Location of aquifer data used in storage calculations for the Denver Aquifer.

A.4.3 Arapahoe Aquifer

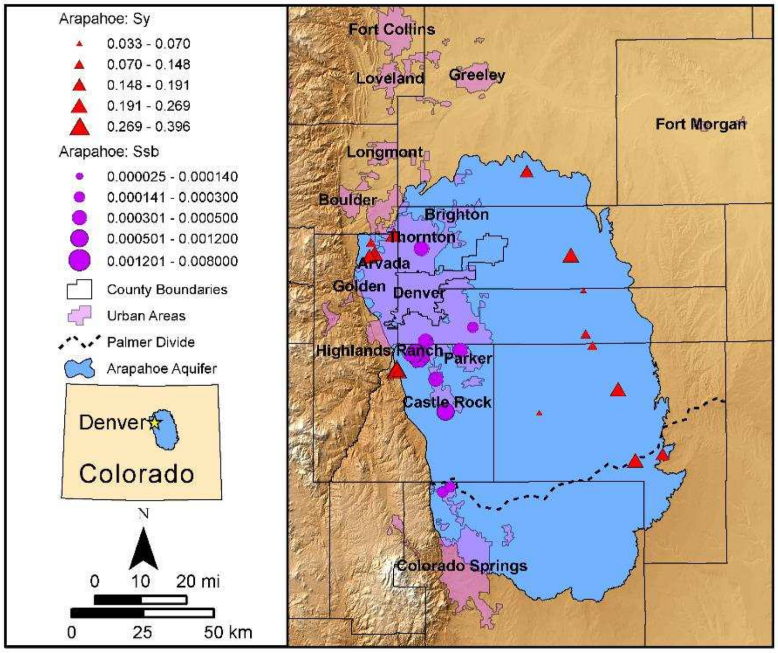

Figure A-15 Location of aquifer data used in storage calculations for the Arapahoe Aquifer.

A.4.4 Laramie-Fox Hills Aquifer

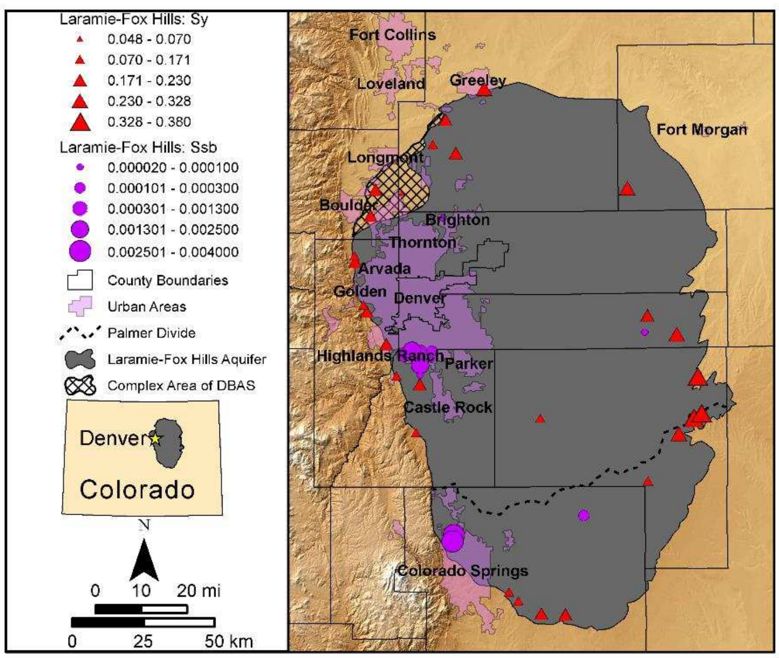

Figure A-16 Location of aquifer data used in storage calculations for the Laramie-Fox Hills Aquifer.