UCLA
UCLA Electronic Theses and Dissertations

Title

Science-based approaches to water resources management: Studies in remote sensing, groundwater and California's Central Valley

Permalink
https://escholarship.org/uc/item/8f14x1fb

Author
Miro, Michelle

Publication Date
2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Science-based approaches to water resources management:

Studies in remote sensing, groundwater and California's Central Valley

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Civil Engineering

by

Michelle Elizabeth Miro

2017

© Copyright by

Michelle Elizabeth Miro

2017

ABSTRACT OF THE DISSERTATION

Science-based approaches to water resources management:

Studies in remote sensing, groundwater and California's Central Valley

by

Michelle Elizabeth Miro

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 2017

Professor Steven Adam Margulis, Chair

This dissertation is motivated by the principle that data availability and scientific analysis are fundamental for effective natural resources management. The research in this thesis presents approaches that can enhance water management institutions' ability to more comprehensively measure and manage groundwater resources. This research draws from a diverse scientific body of work in numerical modeling, remote sensing science, hydrology and public policy. A robust, artificial neural network model is presented that downscales GRACE gridded land datasets (~150,000km^2) to higher-resolution (~16km^2) groundwater storage change estimates, a 100-fold higher resolution. This modeling approach uses minimum input data - five key data sets and minimally processed GRACE data - and thus has applicability to data scarce regions. For California's Central Valley, downscaled groundwater storage change maps can be used to inform groundwater management as they point to specific sub-regional patterns in groundwater storage change. This dissertation also presents a framework intended to strengthen the scientific underpinnings of groundwater management in California. A methodology is developed to calculate sustainable yield under California's new Sustainable Groundwater Management Act (SGMA) that is flexible to varying input data as well as to a given region's local socio-economic and environmental dynamics. The long-term implications of three different determinations of sustainable yield are assessed through an empirical groundwater balance that is projected to 2040. The results of these three scenarios show that there are tradeoffs to be had between groundwater availability, future climate uncertainty and socio-economic preferences that must be carefully weighed. Finally, research is presented that addresses the future of remote sensing. A novel approach to quantify the value of geospatial data for decision makers is presented, along with an unbiased assessment of a rapidly developing branch of remote sensing – privately-owned and privately-funded small satellites. Overall, groundwater management in California is a critical example of the need for robust management strategies in the face of increasing resource scarcity and rising climate variability. California is not alone in this task. The lessons and takeaways presented in this dissertation can be applied to address similar natural resource management challenges across the globe.

The dissertation of Michelle Elizabeth Miro is approved.

James Famiglietti

William Yeh

Alex Hall

Mekonnen Gebremichael

Steven Adam Margulis, Committee Chair

University of California, Los Angeles

2017

TABLE OF CONTENTS

1. Introduction **1**
1.1 Research questions 2
1.2 Organization 4
1.3 References 5

2. Downscaling GRACE remote sensing datasets to high-resolution groundwater maps of California's Central Valley **6**
2.1 Introduction and background 6
2.1.1 Downscaling GRACE data 8
2.1.2 Goals and objectives 10
2.2 Study region: California's Central Valley 11
2.3 The neural network approach and input data 13
2.4 Results and analysis 18
2.4.1 Approach 1: In situ point data for calibration 20
2.4.2 Approach 2: Kriged groundwater surface for calibration 20
2.4.3 Approach 3: Kriged groundwater surface for calibration (2002-2006) with validation over entire surface (2007-2010) 22
2.4.4 Finalized neural network model results 23
2.5 Conclusions 28
2.6 References 30

3. A hydrologically-based framework for calculating sustainable yield under California's Sustainable Groundwater Management Act (SGMA) **39**
3.1 Introduction 39
3.2 The challenge of defining sustainability for groundwater management 42
3.3 Case study – the South San Joaquin Irrigation District 45
3.4 A novel framework to calculate sustainable yield 47
3.4.1 Step 1: Quantifying a baseline sustainable yield 47
3.4.2 Step 2: Determining constraints to the baseline sustainable yield 51
3.4.3 Step 3: Projecting basin response over the GSP management horizon 56
3.5 Conclusions 62
3.6 References 65

4. Diagnosing the health of the commercial remote sensing market - How developing value-added solutions can foster industry growth **72**
4.1 Introduction 72
4.2 Overview of the business to business market for commercial remote sensing 73
4.3 Strategies to promote growth in the commercial remote sensing industry 78
4.4 A case study of Walmart's operations and supply chain management 82
4.5 Remote sensing applications for Walmart 85
4.6 Valuing geospatial data within Walmart's sales forecasting capacity 87
4.7 Conclusions 93
4.8 References 95

5. Conclusions **100**
5.1 Contributions 102
5.2 Future Work 103
5.3 References 108

6. Appendix A – Basics of Neural Networks **110**
6.1 References 114

LIST OF FIGURES

Figure 2.1 Map of California groundwater basins **12**

Figure 2.2 Maps of annual change in groundwater storage (m). **18**

Figure 2.3 Groundwater storage change for 2010 (m), Absolute errors for 2010 (m); Relative errors for 2010 (%). **22**

Figure 2.4 Neural network model downscaled groundwater storage change (m) maps from 2002-2010. **24**

Figure 2.5 Spatial resolution of remote sensing and in situ water storage change data over study region for 2010 (m). **25**

Figure 2.6 Percent influence of input variables on neural network model output. **26**

Figure 2.7 Cumulative annual groundwater storage change (km^3) for the San Joaquin groundwater basin. **28**

Figure 3.1 Sustainable yield management scenarios. **45**

Figure 3.2 Hill's method - Estimating sustainable yield from changes in groundwater levels. **48**

Figure 3.3 Modified Hill's Method for Central Valley's wet and dry year climatology. **50**

Figure 3.4 Sustainable yield calculation for SSJID (1994-2014). **51**

Figure 3.5 Undesirable results variables and relationship to groundwater level. **54**

Figure 3.6 Recalculating sustainable yield under a streamflow constraint. **55**

Figure 3.7 Sustainable yield calculation for Scenario 2 – managed depletion, recovery. **60**

Figure 3.8 Projected average annual groundwater levels for SSJID from 2017 to 2040. **61**

Figure 4.1 Relationship between data, information and intelligence. **77**

Figure 4.2 The spatial and temporal resolution of existing remote sensing data sources. **81**

Figure A1 General structure of a two-layer feedforward neural network. **111**

LIST OF TABLES

Table 2.1 Uninterpolated points (50% calibration, 50% validation) **20**

Table 2.2 Kriged groundwater surface (50% calibration, 50% validation) **21**

Table 2.3 Kriged groundwater surface **23**

Table 3.1 Undesirable Results thresholds, recommended references for GSAs **52**

Table 3.2 Best-fit lines between groundwater level and each undesirable result **54**

Table 3.3 Sustainable yield values for management scenarios in SSJID **59**

Table 4.1 Types of remote sensing data available **75**

Table 4.2 Walmart's supply chain key strategies and challenges **85**

Table 4.3 Applications of remote sensing data to Walmart's supply chain **87**

Table 4.4 Time series of independent variables, weekly (2/2010-7/2013) **89**

Table 4.5 Statistical test results for multivariate regression of Models 1-3 **92**

ACKNOWLEDGMENTS

This dissertation is the compilation of work under the guidance and leadership of Dr. James Famiglietti. Thank you for your mentorship, for your devotion to science communication and for the example you have set for me and for so many scientists to come. You have set the bar high.

The research in Chapters 2 and 3 is co-authored with Dr. Famiglietti. Chapter 2 is a version of:

> Miro, M. and J. Famiglietti (2017). “Downscaling GRACE remote sensing datasets to high-resolution groundwater maps of California’s Central Valley”. Water Resources Research. *Under review.*

Chapters 2 and 3 were supported by grants from the NSF Graduate Research Fellowship Program, the University of California Office of the President, Multi-campus Research Programs and Initiatives, and the NASA GRACE Science Team.

Chapter 4 is the product of a research project with Geoffrey Torrington at the RAND Corporation, who co-authored the paper. It is a version of:

> Miro, M. and G. Torrington (2017). "Diagnosing the health of the commercial remote sensing market - How developing value-added solutions can foster industry growth.” Remote Sensing Applications: Society and Environment. *Under Review.*

This work was supported by the U.S. Department of Defense and the internal support of the RAND Corporation, a nonprofit, nonpartisan, research institution.

I would also like to acknowledge the members of my committee for their time, perspectives and insight. Finally, thank you to Dr. Mark Gold – a true visionary and leader in the field. I am grateful to have learned from you.

VITA

EDUCATION

Ph.D. Civil Engineering 2017 *expected*
Hydrology and Water Resources - University of California, Los Angeles
M.S. Civil Engineering 2012
Hydrology and Hydraulic Engineering - University of Illinois Urbana-Champaign
B.A. Humanities 2010
Health and Policy - Special Honors, High Honors – Univ. of Texas at Austin

EXPERIENCE

- *Graduate Researcher* 2014 – present
  UCLA - Los Angeles, CA
- *Summer Associate and Adjunct Researcher* 2016 – present
  RAND Corporation - Washington, D.C.
- *Graduate Researcher* 2012 – 2014
  University of California Center for Hydrologic Modeling - Irvine, CA
- *Research Assistant* 2010 – 2012
  University of Illinois Urbana-Champaign - Champaign, IL
- *Technical Editor* 2014 – 2016
  American Journal Experts - Nationwide, USA
- *Industrial Attachment* 2010
  Ghana Water Company, Technical Services Division - Ho, Volta Region, Ghana

TEACHING EXPERIENCE

- *Teaching Fellow* 2015 – 2016
  UCLA Institute of the Environment and Sustainability - Los Angeles, CA
  ENV166 Leadership in Water Management
- *Teaching Assistant* 2013
  University of California, Irvine - Irvine, CA
  CEE 20 Engineering Problem Solving
- *Teaching Assistant* 2011 – 2012
  University of Illinois Urbana-Champaign - Champaign, IL
  CEE 201 Civil and Environmental Systems Engineering
  CEE 434 Environmental Systems I

PUBLICATIONS

Miro, M. and J. Famiglietti (2017). "Downscaling GRACE remote sensing datasets to high-resolution groundwater maps of California's Central Valley". Water Resources Research. *Under review.*

Famiglietti, J. and **M. Miro** (2017). "Our wild, wet winter doesn't change this reality — California will be short of water forever". Los Angeles Times Op-Ed. 7 March 2017.

Miro, M. and G. Torrington (2017). "Diagnosing the health of the commercial remote sensing market - How developing value-added solutions can foster industry growth." Remote Sensing Applications: Society and Environment. *Under Review.*

Massoud, E., Purdy, A.J., **Miro, M.**, Hallerback, S., Famiglietti, J. and J. Vrugt (2017). "Quantifying groundwater sustainability in California's Central Valley - An empirical method to estimate and project groundwater depletion and recharge". *In Preparation.*

Cai, X and **Miro, M.** (2014). "Sustainable Water Resources Management in China: Critical Issues and Opportunities for Policy Reforms". Oxford Companion to Economics of China.

Miro, M. and Cai, X (2012). "Integrating institutional and local decision making with emergent environmental phenomena: the case of the Republican River Basin". University of Illinois.

Im, EO, *et al.* (2010). "'A waste of time': Hispanic women's attitudes toward physical activity". Women & Health. 50(6):563-79.

CONFERENCE PRESENTATIONS

"A Hydrologically-based Method for Calculating Sustainable Yield under California's Sustainable Groundwater Management Act". American Geophysical Union Fall Meeting. December 2016. Oral presentation.

"Downscaling GRACE for groundwater management in California's Central Valley". National Ground Water Association Groundwater Summit. April 2016. Oral presentation.

"Integrating institutional and local decision making with emergent environmental phenomena: the case of the Republican River Basin". EWRI World Environmental and Water Resources Congress. May 2012. Oral presentation.

"Water resources and systems in rural Tamil Nadu, India". National Ground Water Association Groundwater Summit. April 2010. Poster. Award received for top student research.

CERTIFICATES

Leader in Sustainability - UCLA Institute of the Environment and Sustainability	2016
Mediation - certified by State of California per Dispute Resolution Program Act 1986	2014
Non-profit Management - University of Texas at Austin	2010

AWARDS

National Science Foundation Graduate Research Fellowship	2012 – 2016
1st and 2nd Place, Net Impact Consulting Challenge, UCLA Anderson School	2015 – 2016
Borlaug Institute for Global Food Security Summer Fellow	2013
National Ground Water Association Farvolden Scholarship	2010
Undergraduate Endowed Presidential Research Fellowships	2008 – 2010
Phi Beta Kappa	2010

1. Introduction

Groundwater is the most heavily extracted raw material worldwide [*NGWA*, 2013]. For many communities, it is the primary source of water, accounting for 40 percent of the global drinking water supply and 38 percent of the water used for irrigation [*NGWA*, 2013]. During times of drought and in water-stressed, arid regions, these percentages increase dramatically as the availability of surface water diminishes, and communities become even more dependent on groundwater to meet their needs. As a result, many aquifers have become severely, and even irreversibly, depleted. In California's Central Valley, which accounts for one-fifth of the nationwide demand for groundwater, water tables have been declining over the past few decades [*Faunt et al.*, 2009]. This intensive overuse of groundwater resources has resulted in land subsidence, degraded water quality and increasing costs of extraction as water tables decline. Government officials and scientists have extensively reported these impacts, drawing from monitoring data, hydrologic models, first-hand accounts and satellite observations. A recent study has shown that the Central Valley lost 20.3 cubic kilometers of water between 2003 and 2010 [*Famiglietti et al.*, 2011]. Much of this loss was due to extensive groundwater pumping to support irrigated agriculture [*Famiglietti et al.*, 2011]. Moreover, significant drought conditions that began in 2011 caused additional declines in the Central Valley of approximately 10 cubic kilometers of freshwater annually in 2012 and 2013 [*Famiglietti,* 2014]. The dichotomy of rapid depletion and increasing dependence on groundwater was highlighted by drought in California between 2011-2015, gaining the eye of both the media and politicians. In September of 2014, California Governor Jerry Brown signed landmark groundwater management legislation [*CA DWR,* 2015]. This series of bills, known collectively as the Sustainable Groundwater Management Act (SGMA), are targeted at providing a statewide framework for local groundwater management and imposing regulations on groundwater extraction [*Water Education Foundation*, 2015]. As a result of SGMA, groundwater, formerly unregulated in most of the state, will now be subject to mandatory sustainability plans and increasing government intervention and oversight.

In California, traditional water management paradigms have focused heavily on surface water and therefore lack the institutional structures, mandates and informational frameworks necessary for effective groundwater management. Thus, the implementation of SGMA presents a new opportunity for water management agencies to manage groundwater resources in a more judicious and sustainable way. To do so, water managers and institutions need to learn to bridge the gap between hydrologic science and policy - to develop the tools necessary to craft and administer institutionally appropriate and physically relevant groundwater management plans.

The overall motivation of this dissertation is to develop science-based tools that provide utility to water management and is based on the premise that data availability and scientific analysis are fundamental for effective management. This dissertation is thus targeted at enhancing institutions' ability to measure groundwater resources and to manage groundwater extractions by fusing hydrologically-based methods, numerical models and remote sensing.

1.1 Research questions

Groundwater management in California serves as a critical example of a resource that requires data availability and scientific tools as it begins adopting new institutions and regulations under conditions of dramatic and ongoing change. In this context, the following research questions guided the work that is detailed in subsequent chapters.

I. Can we better characterize the spatial and temporal patterns in California groundwater?

To address this question in Chapter 2, research is presented that compares three different sources of information on California groundwater in a portion of the Central Valley. These three sources are i) in situ groundwater well data from publicly available monitoring networks, ii) satellite-based remote sensing of water storage and iii) downscaled remote sensing data obtained empirically by combining data sources i) and ii).

II. How can groundwater water agencies methodologically implement the concept of sustainable yield into groundwater management plans?

In Chapter 3 of this dissertation, a framework for calculating sustainable yield according to California's new groundwater management paradigm, SGMA, is detailed. This framework is both flexible and systematic, and includes modifications to traditional methodologies that make them appropriate for California's hydrology and established approaches to water management.

III. How should groundwater sustainability be defined and assessed for California?

Chapter 3 of this dissertation also addresses this question. The work synthesizes a large body of literature on sustainability in groundwater management and applies it to California. It also examines the long-term impacts of three management strategies on groundwater levels in a sub-region of California's Central Valley.

IV. Is the privately-owned, commercial remote sensing industry robust enough to provide information in the future?

Chapter 4 assesses the overall health of the private remote sensing market. The first section of this chapter integrates expert interviews and existing customer data to show that the private remote sensing industry has stagnated. Three strategies are then presented as a means to generate the industry growth necessary to sustain it into the future. These strategies also provide insight into the future potential of privately-owned remote sensing data.

V. What is the value of remote sensing based information?

Chapter 4 also addresses this question. It presents a case study that values remote sensing based temperature data for a private company. Using publicly available sales data, this chapter details a model that quantifies the value-added of remote sensing data to store-level revenue.

1.2 Organization

The following three chapters, which together make up this dissertation are each complete with a review of relevant literature and descriptions of the study region. They will also answer the presented research questions and detail how specific scientific objectives were addressed. Chapter 2 presents an overview of groundwater monitoring in California, a theoretical approach to applying neural networks to downscale remote sensing data and three modeling approaches for a neural network downscaling model in California's Central Valley. Chapter 3 describes a three-step framework for calculating sustainable yield under SGMA. Chapter 4 details a two-part study that investigates the health of commercial remote sensing and then quantifies the value-added of remote sensing data for decision makers. The main contributions of each chapter and future work are presented in the Conclusion.

1.3 References

California Department of Water Resources (2015), Legislation, Sustainable Groundwater Management Act (SGMA), California Department of Water Resources. http://www.water.ca.gov/cagroundwater/legislation.cfm

Famiglietti, J. (2014), Epic California Drought and Groundwater: Where Do We Go From Here? National Geographic Water Currents. http://voices.nationalgeographic.com/2014/02/04/epic-california-drought-and-groundwater-where-do-we-go-from-here/

Famiglietti, J. S., Lo, M. Ho, S.L., Bethune, J., Anderson, K., Syed, T. Swenson, S. de Linage, C. and M. Rodell (2011), Satellites measure recent rates of groundwater depletion in California's Central Valley, Geophysical Research Letters 38, L03403. doi:10.1029/2010GL046442

Faunt, C. (2009), Groundwater Availability of the Central Valley Aquifer, California, U.S. Geological Survey. Professional Paper 1766. http://pubs.usgs.gov/pp/1766/

National Ground Water Association (NGWA) (2013), Facts about global groundwater usage, Web. Sept 2013. http://www.ngwa.org/Fundamentals/Documents/global-groundwater-use-fact-sheet.pdf

Water Education Foundation (2015). "The 2014 Sustainable Groundwater Management Act: A Handbook to Understanding and Implementing the Law". Water Education Foundation. Sacramento, CA. http://www.watereducation.org/sites/main/files/file-attachments/groundwatermgthandbook_oct2015.pdf

2. Downscaling GRACE remote sensing datasets to high-resolution groundwater maps of California's Central Valley

2.1 Introduction and background

Groundwater monitoring has historically relied on a network of local observations of well levels, or in situ measurements. In the United States, the United States Geological Survey (USGS) maintains a network of 850,000 active monitoring wells that provides fundamental data on groundwater quantity and quality [*Taylor and Alley*, 2001] and enables essential regional studies [*Faunt,* 2009; *Faunt and Sneed*, 2015]. In many other parts of the world, however, groundwater observation networks often lack adequate spatial and temporal coverage, they are often underfunded, and therefore they may well be unreliable [*Mogheir et al.,* 2005; *Shah et al.*, 2000]. Even in the United States, where a relative abundance of well data provides information to water managers on short- and long-term water level trends at specific locations, more monitoring sites are needed to better understand the groundwater surface and the spatial distribution of pumping patterns [*Taylor and Alley*, 2001].

To overcome the shortcomings from sparse observation networks and insufficient in situ data, significant progress has been made in the way the groundwater surface and its behavior are represented. These advances have come from the fields of groundwater modeling [*Harbaugh et al.*, 2000; *Faunt*, 2009; *Dagan*, 1982; *Hendricks Franssen and Kinzelbach*, 2008], monitoring network design [*Reed et al.*, 2000] and geostatistical analysis of groundwater data [*Hughes and Lettenmaier*, 1981; *Kitanidis and Vomvoris*, 1983; *Sun et al.*, 2009]. While this research has made huge strides in characterizing groundwater from limited data, many of these studies focused on small, sub-basin scales and failed to capture wider spatial trends in groundwater.

Satellite remote sensing can complement existing monitoring networks and modeling studies and can help compensate for gaps in spatial and temporal coverage. In particular, following the work of *Rodell and Famiglietti*, [2002], *Yeh et al*. [2009] and *Rodell et al.,* [2007], several authors have now demonstrated that NASA's Gravity Recovery and Climate Experiment (GRACE) can reliably measure monthly groundwater storage changes in the large aquifer systems of the world. Some examples include the Ogallala aquifer [*Strassberg et al.*, 2007], northwestern India [*Rodell et al.,* 2009], California's Central Valley [*Famiglietti et al.,* 2011], South America's Guarani aquifer [*Munier et al.,* 2012], the Middle East [*Voss et al.,* 2013], the North China Plain [*Feng et al.,* 2013] and several others [*Famiglietti,* 2014; *Richey et al.,* 2015 a,b].

Despite these studies, the ability of GRACE to monitor changes at finer scales, which could directly benefit local water management authorities, is limited. This is largely due to the low spatial resolution of its observations (~200,000 km^2), and researchers and hydrogeologists have noted these drawbacks [*Famiglietti and Rodell,* 2013; *Alley and Konikow,* 2015]. The lack of ground truthing and the potential errors in retrieval algorithms are also cited as deficiencies in remotely sensed data [*Fekete et al.*, 2015]. A higher-resolution GRACE data product would significantly improve information availability for local-scale decision makers, as well as offer novel data for regions that do not have adequate in situ monitoring networks. To supplement the shortage in regional in situ data and improve upon the resolution of GRACE data, this study downscales GRACE and creates a hybrid product that utilizes available local observations along with GRACE estimates of changes in total water storage to accurately characterize local changes in groundwater availability. Our approach has potential for use in data scarce regions worldwide, as it requires only minimal hydrologic data and GRACE estimates of changes in total water storage to simulate groundwater storage change in a complex aquifer system.

2.1.1 Downscaling GRACE data

The majority of research approaches for downscaling remote satellite data originate in the climate modeling literature, owing to the need to better understand the regional impacts of global change. Two approaches, dynamical and statistical, are the most common [*Wilby and Wigley*, 1997]. Dynamical methods typically utilize a higher-resolution, physically-based model using low-resolution data, such as those from a global climate model or general circulation model, as the lateral boundary conditions. The GRACE data assimilation approach of *Zaitchik et al.* [2008] was effectively a form of dynamical downscaling. Monthly GRACE observations of terrestrial water storage change (i.e. the change in the sum of snow, surface water, soil moisture and groundwater) were assimilated into a physically-based land surface model at the scale of the major watersheds of the Mississippi River Basin [*Zaitchik et al*., 2008]. The output of the physical models – the higher resolution, modeled water storage changes within the major watersheds – was forced to sum to the lower-resolution assimilated constraint from GRACE. The physics of the land surface modeling and atmospheric forcings were used to distribute the GRACE data to finer scales. Data assimilation methods have the strong advantage of being physically consistent but, at the same time, require significant computational time, limiting their applicability [*Schoof*, 2013].

Statistical downscaling methods, instead, draw upon relationships between coarser-scale input data and finer-resolution target data. [*Wilby et al.*, 1998]. A variety of statistical techniques have been applied and studied in the downscaling literature, including classification-based methods, regression models, Markov chains and stochastic models [*Wilby et al.,* 2004]. The advantages of these methods are that they are relatively flexible to various data types and spatial and temporal scales, they can generate uncertainty estimates of parameters and model output and they are generally easy to apply [*Wilby et al.,* 2004]. Statistical methods are, however, based on the assumptions that the input and target data fully capture the dynamics of the system under study and that these dynamics are valid even outside of the observation period [*Wilby et al.,* 2004]. Studies that have compared dynamical and statistical downscaling approaches have revealed relatively similar results between the two types of methods [*Schoof,* 2013].

Some researchers have also opted for a hybrid option. *Huang et al.* [2015] developed a physically-based statistical modeling approach that combines both methods, promoting the decrease in computational time as one of the method's advantages. Purely statistical methods, though, offer ease of use, even lower computational requirements and simplicity, which has been shown to be an advantageous attribute for downscaling hydrologic data [*Chiew et al.,* 2010; *Fasbender and Ouarda*, 2010; *Frost et al.,* 2011].

Given these benefits, we have adopted a tried and true statistical downscaling approach that will be novel in its application to downscale GRACE data. It will rely on derived relationships from local observations instead of on equations based on physical processes. Of the possible numerical methods available for this approach, artificial neural networks were selected. This technique has been widely used for statistical downscaling in the hydrosciences [*Schoof and Pryor*, 2001; *Dibike and Coulibaly*, 2006; *Fowler et al.*, 2007], in spatial data analysis [*French et al.*, 1992; *Zhu,* 2000; *Jin et al.*, 2006], in studies for groundwater management [*Chu and Chang*, 2009], for predicting groundwater levels [*Daliakopoulos et al.*, 2005; *Krishna et al.,* 2008; *Yang et al.*, 2009; *Yoon et al.*, 2011; *Taormina et al.,* 2012], as well as for predicting groundwater levels with GRACE data [*Sun,* 2013]. Neural network studies have also illustrated the method's ability to simulate complex hydrological characteristics across various regions and time periods [*Hsu et al.*, 1997; *Hsu et al.*, 1995]. In addition, artificial neural networks are highly capable of processing different types of data efficiently [*Turban et al.,* 2008]. This allows the proposed model to use input data (GRACE, meteorological forcings and soil types) similar to previous, well-established data assimilation studies, yet depart from these physically-based methods conceptually and offer a quicker computational time. The neural network model also has the flexibility to incorporate alternative data sets and future GRACE data releases, such as GRACE RL05M Mascon Solutions and future GRACE Follow-On (GRACE-FO) data [*Wiese*, 2015; *JPL,* 2016].

2.1.2 Goals and objectives

Here we present a neural network model to spatially downscale GRACE data from ~200,000 km^2 to ~16 km^2, as well as to vertically isolate the groundwater component from GRACE estimates of total water storage. We apply the downscaling model to the time period 2002-2010 in order to generate a series of annual, high-resolution maps of changes in groundwater storage over a portion of California's Central Valley. In doing so, we seek to determine whether or not a neural network, numerical downscaling approach is appropriate for downscaling remote sensing data. We also examine the optimal spatial and temporal characteristics of the calibration dataset that would inform future work and a more widespread application of downscaled remote sensing data to groundwater management. We assess this by testing the type of groundwater information (i.e., point data or interpolated surfaces) that improves the neural network's estimate of groundwater change in a given year. Finally, our modeling approach also investigates the best way to calibrate and validate the model in time and if the method has the capability to project forward in time.

2.2 Study region: California's Central Valley

In California's Central Valley (see Figure 2.1), where agricultural water use accounts for one-fifth of the nationwide demand for groundwater, groundwater levels have been declining dramatically over the past few decades [*Faunt*, 2009]. Dependence on groundwater resources is even more pronounced during times of drought as communities and farmers have few alternatives to meet their water needs [*Howitt et al.,* 2015]. In many areas of the Central Valley, the intensive overuse of and reliance on groundwater resources has resulted in land subsidence, degraded water quality and increasing costs of extraction due to deepening water tables [*Howitt et al.*, 2015; *Faunt and Sneed*, 2015]. *Faunt and Sneed* [2015] show the region has lost 100 km^3 of groundwater since the 1960s. Between 2003 and 2010, GRACE satellite observations showed that the Central Valley lost 20.3 km^3 of groundwater - primarily due to extensive groundwater pumping to support agriculture [*Famiglietti et al.*, 2011]. Dire drought conditions that began in 2011 have caused additional water losses of approximately 10 km^3 of freshwater between 2012 and 2013 [*Famiglietti*, 2014]. From 2012 to present, land subsidence within the Central Valley, which is the result of water loss and compacting sediments within an aquifer, reached up to 280mm in some places [*Faunt and Sneed*, 2015]. Another estimate shows that peak rates of subsidence – 500mm/year – occurred during 2014 [*Farr et al.*, 2015; *Faunt and Sneed*, 2015].

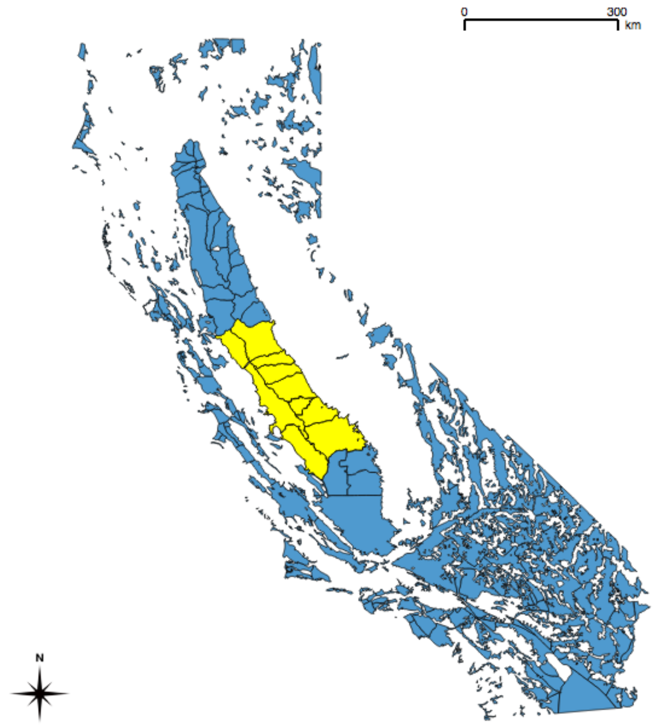

Figure 2.1 Map of California groundwater basins [*CA DWR*, 2016]. The study region – the San Joaquin Valley – is highlighted in yellow.

Despite these dramatic physical impacts to Central Valley aquifers, a comprehensive assessment of California groundwater basins has not been performed since 1980 [*RMC*, 2014]. The California Department of Water Resources (DWR) cites the lack of information to properly quantify groundwater overdraft as the main reason for this gap in analysis [*RMC*, 2014]. This is not unique to California. Calls by scientists, engineers and water managers for more extensive monitoring networks that provide better information for water management have been commonplace throughout the 20th century and still continue to be so today [*Taylor and Alley*, 2001].

Ongoing drought conditions, continued groundwater losses and dramatic rates of land subsidence all point to the need for effective management and heightened monitoring of California's groundwater resources [*Howitt et al.*, 2015]. GRACE observations provide comprehensive information on drought impacts, climate change and groundwater and have proven to be a powerful tool for understanding regional water resources behavior [*Famiglietti and Rodell*, 2015]. Results from GRACE-based studies have already been used to inform decision-making processes in California's Central Valley, Texas, and the American Southeast [*Lo and Famiglietti*, 2013]. Refining GRACE to higher resolution estimates of groundwater changes would provide a significant value-added to groundwater management efforts and upcoming implementation of California's Sustainable Groundwater and Management Act (SGMA) [*CA DWR*, 2015].

2.3 The neural network approach and input data

The GRACE downscaling model employs an artificial neural network (ANN) to combine low-resolution GRACE data with higher resolution hydrologic variables in order to predict changes in local groundwater storage and to, in effect, vertically isolate the groundwater component of the GRACE signal. ANNs are particularly useful for this task as they are often employed in spatial data analysis and offer the ability to efficiently and comprehensively handle large, diverse and noisy spatial datasets [*Turban et al*., 2008]. Because they are not yet widely used in numerical downscaling of remote sensing data, this research also represents a novel application of ANNs.

In essence, the ANN derives non-linear, empirical relationships between GRACE, the input hydrologic datasets to the downscaling model and the output variable – groundwater storage change. These relationships are represented numerically by a network of empirical equations that are fit during the network learning, or calibration, process. Our downscaling model employs a two-layer feedforward neural network, which was calibrated with a Bayesian regularization backpropagation learning algorithm [*MathWorks*, 2016; *Turban et al.*, 2008]. A more complete discussion on the training algorithm and neural network architecture can be found in *MacKay* [1992] and *Turban et al.* [2008], respectively, as well as in Appendix A.

Because neural networks are data-driven models, the quality and nature of the data used as inputs are of critical importance. Previous studies that employed neural networks to predict groundwater levels utilized both environmental and hydrologic variables and included distinct combinations of: precipitation, temperature, surface discharge (in riparian groundwater systems), tidal levels (in coastal aquifers), and potential evapotranspiration [*Lallahem et al.*, 2005; *Krishna et al.*, 2008; *Yoon et al.*, 2011; *Taormina et al.,* 2012]. This study extends this work to focus on storage change rather than groundwater levels. To do so, we use precipitation and temperature data along with GRACE and other key local hydrogeologic datasets (soil type and slope) that are shown in the literature to be significant predictors of terrestrial water storage change [*Reager and Famiglietti*, 2013]. Together, these variables – GRACE observations of terrestrial water storage, slope, soil type, precipitation and temperature – serve as the hydrologic input data to the neural network model, which is calibrated to changes in in situ groundwater storage. Once calibrated, the downscaling model utilizes the fit empirical relationships between these datasets to generate new estimates of changes in aquifer storage from an alternate set of hydrologic input data from either a new region or from a different point in time. Because the model is calibrated to changes in storage from groundwater alone, the downscaling model also vertically isolates the groundwater component of the GRACE input data.

Hydrologic datasets used as model inputs were obtained from the 2002-2010 period over California's Central Valley. They include: 2.5min (~4km) resolution precipitation and mean temperature data from PRISM [*PRISM Climate Group*, 2014]; 10-meter DEM [*USGS*, 2009], processed in ArcGIS for slope; 10-meter NRCS soil maps from the Gridded Soil Survey

Geographic (gSSURGO) Database [*NRCS*, 2014]. All input data were discretized to the 4km by 4km target spatial resolution.

GRACE Release 5 (R05) data compiled and processed by the Center for Space Research (CSR) for the period 2002-2010 were used as model inputs and can be found at http://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/. These data consist of monthly measurements across the land surface at a 1-degree by 1-degree resolution. Each grid was multiplied by its scale factor, as provided by GRCTellus, in order to adjust for attenuation of the signal during smoothing and destriping. This procedure is outlined in *Landerer and Swenson* [2012]. Next, each GRACE grid cell in the study region was discretized and spatially interpolated to the target 4km resolution. To do this, the original GRACE grid cell value was treated as the centroid of the new 4km discretized grid cell and was interpolated to the centroid of each of the neighboring GRACE grid cell values. A linear interpolation was performed to fill in the rest of the discretized grid between the centroid GRACE value and the values at the corner points. In this way, GRACE data were treated as a surface, taking into account not only a single grid cell but also its neighbors. This allowed the model to incorporate more information about the spatial distribution of groundwater change, rather than just considering a single magnitude. To annualize the GRACE data and make it comparable to the in situ groundwater data, the storage change for twelve months of each year, starting in February, were summed to obtain an annual storage change value.

The groundwater data that serves as the calibration and validation datasets for the neural network model were taken from 2,189 wells across San Joaquin County [*California Water Data Library*, 2015]. This dataset can be accessed at http://www.water.ca.gov/waterdatalibrary/. It is important to note that region covered by the model domain, as shown in Figure 2.1, includes the eleven groundwater subbasins of the San Joaquin Valley Groundwater Basin and encompasses an area of 15,100 km^2 [*CA DWR*, 2003]. In general, the basin's hydrogeology is characterized by unconsolidated alluvium and consolidated rocks and includes both confined and unconfined aquifers [*CA DWR,* 2003]. The presence of a Corcoran Clay confining layer in most of the Central Valley indicates the transition from unconfined to confined aquifers [*Faunt*, 2009]. Across the study region, hydrogeological studies have shown that most confined aquifers begin at a depth of over 60 meters [*Burow et al.*, 2004; *Provost and Pritchard Consulting Group*, 2014]. The depth of most of the wells in the dataset ranged between 10 and 25 meters, with very few wells at a depth of over 45 meters and less than 4% over 60 meters. Because nearly all wells tapped unconfined aquifers, groundwater storage change for each well was calculated using a specific yield of 10%, which was reported as the average value for the San Joaquin Valley by the USGS [*Bertoldi et al.*, 1991]. The specific yield was multiplied by annual groundwater level change, calculated as the difference in well levels in the winter (December - February) months from one year to the next. This assured the capture of the full irrigation season in a given year. A more specific description on the complex hydrogeology of the individual subbasins can be found in California's Department of Water Resources Bulletin 118 basin descriptions: http://www.water.ca.gov/groundwater/bulletin118/.

The neural network downscaling approach, reduces some of the hydrogeological complexity found in the natural system, as we do not directly include any hydrogeological data on these multiple aquifer systems. The model, instead, relies on empirical relationships derived between groundwater change behavior and the input datasets (GRACE, precipitation, temperature, slope and soil type). While these empirical relationships, which vary across space and time, reflect not only temporal and spatial trend in extraction, they also represent the physical characteristics of the multiple aquifer systems.

The calibration and validation data were formatted in two ways for the neural network – distributed point data and spatially interpolated maps of the study region, as shown Figure 2.2. The distributed point dataset was obtained by directly applying the groundwater storage change estimates to a 4 km grid of the region based on each wells latitude and longitude. If more than one well fell within a given grid cell, the average of all wells was used. The spatially interpolated groundwater storage change maps were created by kriging the individual groundwater storage change estimates, as this approach has been found to best approximate groundwater spatially [*Zimmerman et al.*, 1999; *Sun et al.*, 2009]. A more complete discussion on the kriging methodology can be found in *Delhomme* [1978], and its applicability to represent characteristics of multiple aquifer systems can be found in *Martin and Frind* [1998]. More specifically, ordinary kriging was applied to groundwater well data points using an empirically fit spherical semivariogram with a 300-meter nugget, a similar value to regional kriging approaches employed in the San Joaquin basin [*Faunt*, 2009]. Spherical semivariograms are commonly used with ordinary kriging, and the choice of semivariogram is often determined empirically. In this case, a spherical semivariogram resulted in the lowest mean error, average standard error and root mean square error when compared to pentaspherical and exponential semivariograms. We followed the procedures outlined in *Arétouyap et al.* [2016] and *Nikroo et al.* [2009] for semivariogram selection. One kriged map was created for each year. The annual change in groundwater storage, rather than the water levels themselves, was used to create a comparable dataset to the GRACE data, which reflects variations in water storage. In this way, both GRACE and the groundwater data both represented a monthly height difference in water storage.

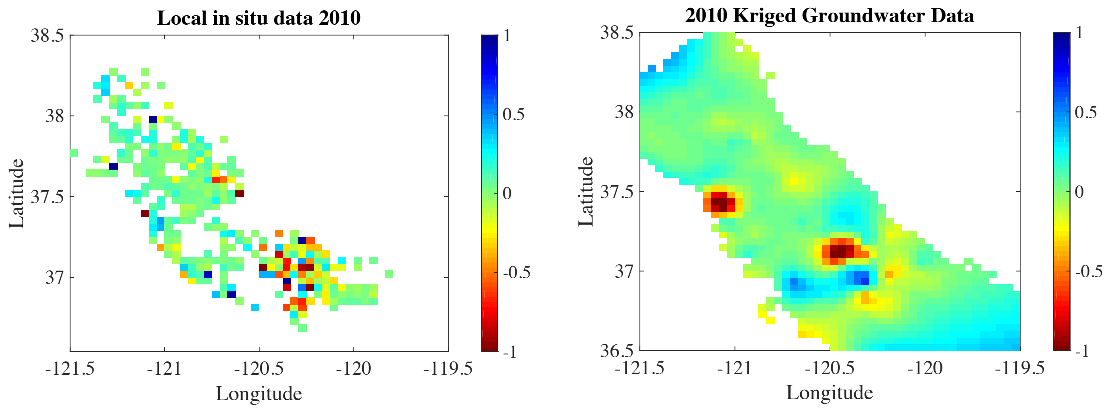

Figure 2.2 Left: Map of annual change in groundwater storage (m) from available in situ well levels for 2010. Right: Map of annual change in groundwater storage (m) from kriged in situ well levels for 2010.

2.4 Results and analysis

The results and analysis are divided according to the calibration data type and validation approach used. Calibration and validation of the neural network model were performed with two data types – distributed point data and spatially interpolated maps of the study region. The model was then validated either on a subset of the data within a given year (50% of the original dataset, randomly selected and kriged separately) or on the entire dataset in a series of different years (2007-2010). More specifically, the first calibration method uses annual groundwater storage changes from individual wells in each year (2002-2010) for calibration and validates the model over a subset of groundwater storage changes in each year. The second method uses spatially interpolated groundwater storage changes, e.g. groundwater maps, in each year (2002-2010) for calibration and validates the model over a subset of the map within each year. This subset is kriged separately following the same procedures to insure data independence. The final method uses spatially interpolated groundwater maps for calibration for 2002-2006 and validates the model over the years 2008-2010. By analyzing these approaches we are able to determine what type of calibration and validation data best informs the network and improves its performance. The performance of each approach was assessed through the use of various model evaluation statistics (Nash-Sutcliffe model efficiency coefficient, root mean squared error, correlation coefficient) and measures of the spatial distribution of model error (relative error, absolute error).

In order to quantify the relative contribution of each input dataset onto the neural network model output, we applied a method proposed by *Garson* [1991]. Garson's method is based on the weights of the calibrated neural network and has been widely cited and is widely used in neural network studies [*Song et al.*, 2013; *Brosse et al.*, 1999; *Gevrey et al.*, 2003; *Olden and Jackson*, 2002]. This method identifies the percentages of influence, $Q_{ik}(\%)$, of each of the input variables on the model's prediction of the output variable [*Song et al.*, 2013; *Garson*, 1991]. It is defined by the following equation:

$$Q_{ik}(\%) = \frac{\sum_{j=1}^{n}\left(\frac{|w_{ij}|}{\sum_{i=1}^{m}|w_{ij}|}|v_{jk}|\right)}{\sum_{i=1}^{m}\left(\sum_{j=1}^{n}\left(\frac{|w_{ij}|}{\sum_{i=1}^{m}|w_{ij}|}|v_{jk}|\right)\right)}\times 100 \qquad (2.1)$$

Where w_{ij} represents the weights between the input variables (neurons), $i = 1,2, \ldots m$, and each of the two hidden layers, $j = 1,2 \ldots n$; v_{jk} represents the weights between the hidden layers and the output variable (neuron), $k = 1,2, \ldots l$; and the number of input neurons, hidden layers and output neurons were $m = 5$, $n = 2$, $l = 1$, respectively. The results of this method are shown in Figure 2.5 and discussed below in Section 2.4.4.

2.4.1 Approach 1: In situ point data for calibration

The first approach calibrates the model with annual groundwater storage changes from each available grid cell in each year (2002-2010). In this approach, 50 percent of the available well information in a given year was selected randomly from the dataset and used for calibration, and the remaining 50 percent was set aside for validation of the model. Results are shown below in Table 2.1. From Table 2.1, we can see that NSE validation values mostly fall within the acceptable range (0-1), but the correlation between simulated and observed values is fairly low. Visual inspection of the spatial distribution of simulated groundwater data also performed poorly, as little to no heterogeneity in the spatial pattern was visible. For these reasons, this approach does not fully capture groundwater behavior across space and during the time period of study.

Table 2.1 Uninterpolated points (50% calibration, 50% validation)

	Calibration Results			Validation results		
Year	NSE	Corr. Coeff.	RMSE (m)	NSE	Corr. Coeff.	RMSE (m)
2002	0.5185	0.1665	0.0512	0.1435	0.1222	0.0586
2003	0.8731	0.2543	0.1210	0.4831	0.3098	0.1061
2004	0.3555	0.3578	0.1036	0.1569	0.3397	0.0845
2005	0.3603	0.2745	0.0814	0.0967	0.2397	0.0941
2006	0.2683	0.1566	0.0861	0.0683	0.1770	0.1200
2007	0.1580	0.4180	0.5608	0.5851	0.2489	0.1044
2008	0.8732	0.2189	0.1215	0.2977	0.2211	0.1263
2009	0.8152	0.2340	0.1159	0.0773	0.1426	0.1466
2010	0.0448	0.1749	0.0853	0.1676	0.0818	0.1099

2.4.2 Approach 2: Kriged groundwater surface for calibration

The second approach to the neural network validation and calibration used a spatially interpolated (kriged) groundwater dataset. Similar to the first approach, 50 percent of the kriged groundwater data was used for calibration, and the remaining portion of the dataset (50 percent) was used to validate the model. By calibrating the model to a best guess of the groundwater surface in the region as opposed to sparse point data, we provided more information to the neural network during the calibration process. Table 2.2 shows error indicators of model results. This approach produced acceptable NSE values (ranging from 0.2445 to 0.9577) for calibration and validation (ranging from 0.0391 to 0.7511), indicating that the model's simulated values are better predictors than observed values alone [*Moriasi et al.*, 2007]. From Table 2.2, it is also evident that the model results have a high degree of correlation to the calibration and validation datasets.

Table 2.2 Kriged groundwater surface (50% calibration, 50% validation)

	Calibration Results			Validation results		
Year	NSE	Corr. Coeff.	RMSE (m)	NSE	Corr. Coeff.	RMSE (m)
2002	0.8364	0.9146	0.0266	0.3981	0.6359	0.0610
2003	0.9431	0.9717	0.0800	0.7511	0.8907	0.2390
2004	0.5624	0.7502	0.0754	0.0692	0.5227	0.3698
2005	0.6976	0.8393	0.0414	0.3185	0.5798	0.1326
2006	0.5799	0.7604	0.0511	0.0453	0.1602	0.0818
2007	0.6111	0.7826	0.3772	0.2096	0.3102	0.6236
2008	0.9577	0.9787	0.0690	0.3285	0.7219	0.2114
2009	0.8721	0.9365	0.1236	0.0391	0.7560	0.1924
2010	0.2445	0.4966	0.0541	0.2547	0.4843	0.0519

Figure 2.3 illustrates the modeled spatial distribution of groundwater change in 2010 along with the absolute and relative errors between groundwater calibration data and model outputs. The error maps show the majority of the absolute and relative error is close to zero (shown in green). In addition, there is little spatial bias in model error; however, most of the error does correspond to areas of more extreme values, indicating that the model's ability to predict extrema (peak and troughs) may be limited.

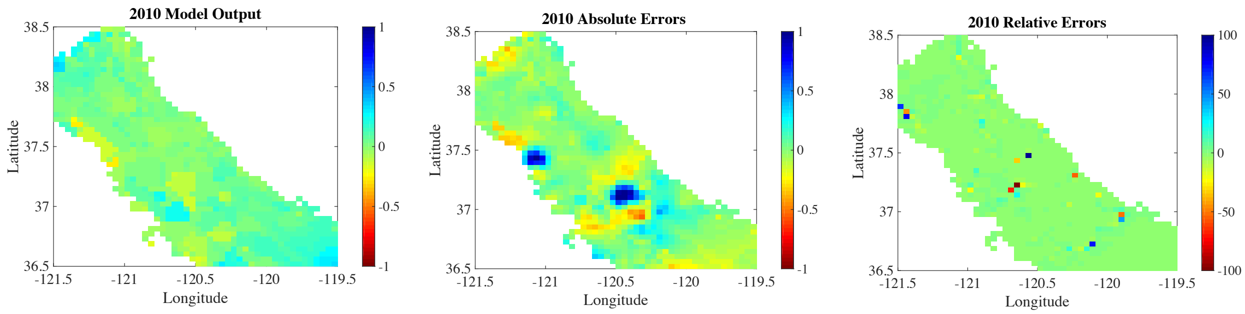

Figure 2.3 Left: Simulated groundwater storage change for 2010 (m); Center: Absolute errors for 2010 (m); Right: Relative errors for 2010 (%).

2.4.3 Approach 3: Kriged groundwater surface for calibration (2002-2006) with validation over entire surface (2007-2010)

The final approach for calibration and validation of the neural network model utilized the same spatially interpolated (kriged) dataset from the second approach but validated the model over four annual time periods (2007-2010) rather than on a portion of the data within each year. In this case, calibration of the neural network model was performed over the first set of years (2002-2006). Table 2.3 shows the overall performance and error indicators of model output. We can see that again NSE values fall in the acceptable range for calibration but are outside of this range for the validation time periods. Further visual inspection of the spatial distribution of both absolute error and relative error for this modeling approach shows significantly more error than in the second approach.

Table 2.3 Kriged groundwater surface

Calibration (2002-2006), Validation (2007-2010)			
Year	NSE	Corr. Coeff.	RMSE (m)
2002	0.5509	0.7429	0.0240
2003	0.8752	0.9355	0.0673
2004	0.6887	0.8302	0.0582
2005	0.8360	0.9143	0.0471
2006	0.6839	0.8270	0.0479
2007	-0.1029	0.1772	0.5246
2008	-3.7980	0.1965	0.6708
2009	-0.3598	0.2301	0.1605
2010	-0.0029	0.3048	0.0642

2.4.4 Finalized neural network model results

The output data of the neural network model contained the least error and highest correlations when using the second approach. Because the model was unsuccessful in simulating groundwater change in new time periods, it is clear that the model requires some groundwater information as an input. This also highlights one of the limitations of an empirically-based downscaling model – once calibrated to a particular period of time or location in space, the model may not accurately represent the groundwater changes in a new region or time period. However, following the second approach, which calibrates each year to a widely available interpolated set of groundwater storage change, the proposed model can acceptably simulate the groundwater surface and downscale GRACE data to the 4km resolution. The maps shown below in Figure 2.4 are the final output of the model. Error data for these maps can be found in Table 2.2.

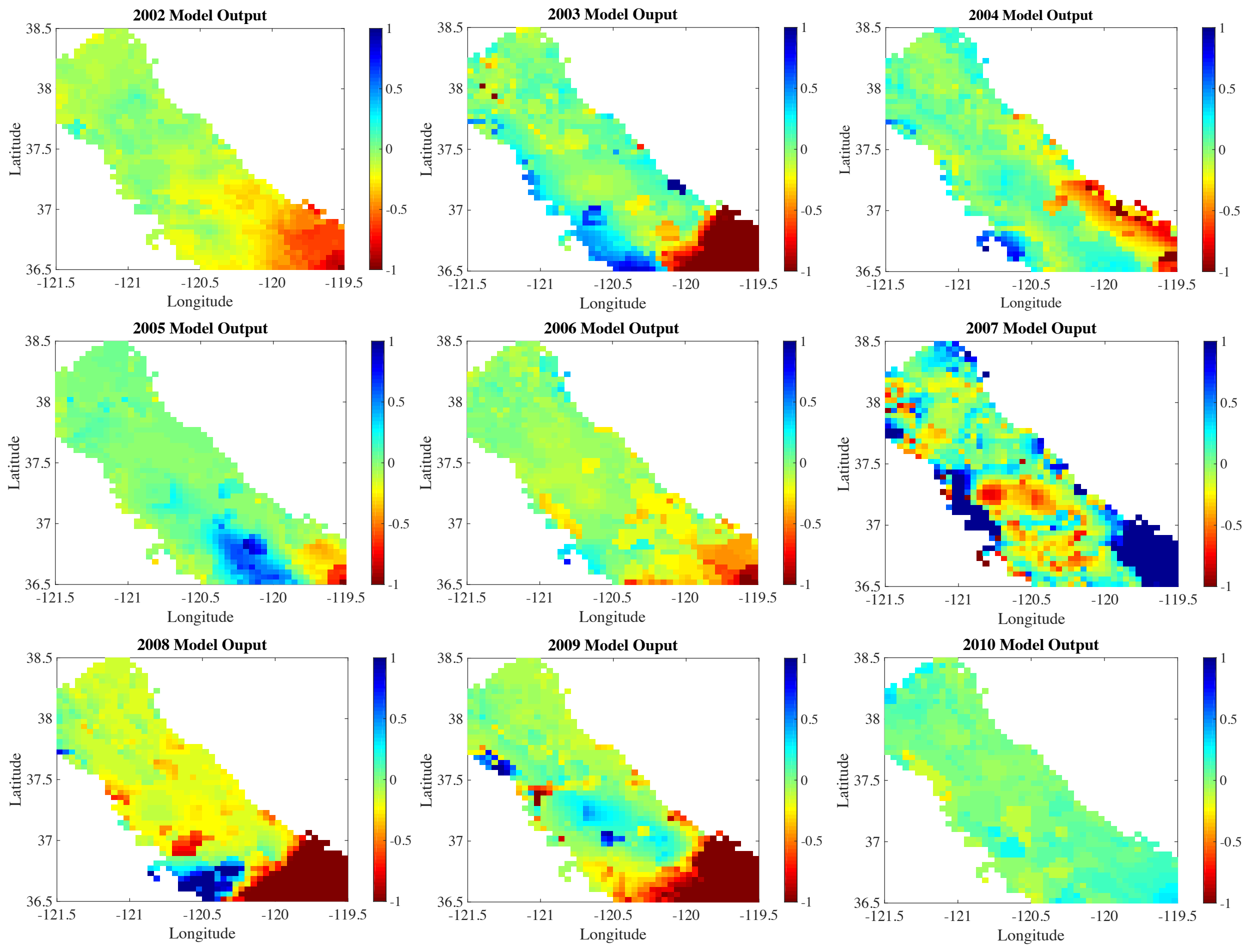

Figure 2.4 Neural network model downscaled groundwater storage change (m) maps from 2002-2010 of San Joaquin, Central Valley, CA.

From Figure 2.4, we can see that the large majority of the groundwater declines (shown in red) during the 2002-2010-time period in California's Central Valley occurred in the eastern and southern portions of the southern Central Valley. These hotspots of groundwater depletion show up to 1 meter of storage loss per unit area in some locations of the southern and eastern portions of the study region. Merced, for example, is located at 37.3022° N, 120.4830° W (in the center of the eastern half of the map) and shows between 0.25 and 1 m of groundwater storage loss in all years except 2007. Other areas, shown in greens and blues, experienced relative increases in groundwater storage. These locations varied from year to year. The central portions of the southern end of the map had recurring increases in the groundwater table (see years 2003, 2005 and 2009). Further study of these blue and red regions could help elucidate why and how certain regions may be losing or gaining groundwater.

Overall, the model output maps point to a high degree of heterogeneity in groundwater behavior compared to GRACE data. As such, it is critical to keep in mind the increase in resolution these maps provide. Figure 2.5 below shows the resolution of the GRACE estimates of total water storage and currently available groundwater well data for this region in 2010. Looking at the GRACE data in Figure 2.5 we can see a slight average regional increase in groundwater storage. However, the model output from this study shows that at a more local level groundwater may be exhibiting more dramatic increases and decreases. These extrema are more or less averaged out at such a low spatial resolution seen in GRACE data. The in situ groundwater data shown in Figure 2.5 does capture some of these highs and lows but fails to provide adequate spatial coverage.

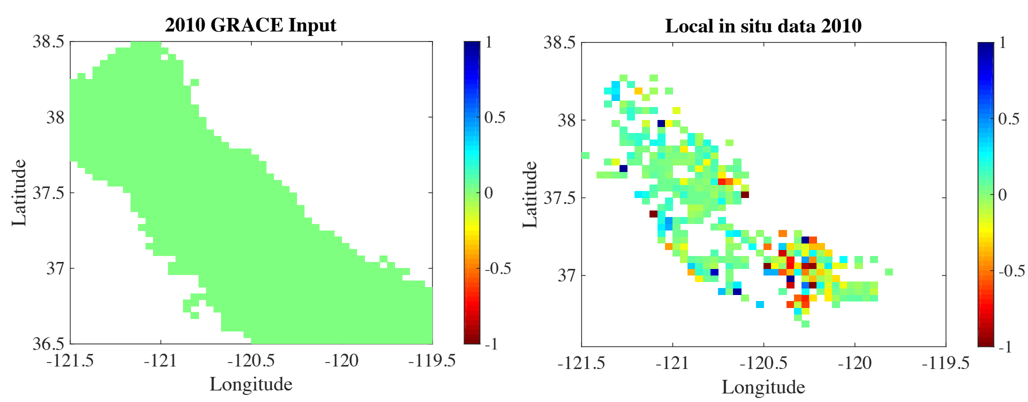

Figure 2.5 Spatial resolution of available remote sensing and in situ water storage change data over study region for 2010 (m).

Figure 2.6 shows the percent by which each input variable influences the model output, as calculated from Equation 2.1, in our final neural network model. GRACE has the highest percent influence (PI) on model output, 38.76%, followed by precipitation, 21.99%, temperature, 15.54%, slope, 12.41%, and soil type, 11.30%. These values illustrate that GRACE is able to explain a significant portion of groundwater storage change in the San Joaquin portion of the Central Valley. Because GRACE only provides low spatial resolution information, the PI values also show that the remaining input variables are necessary to achieve model output at our target, higher spatial resolution.

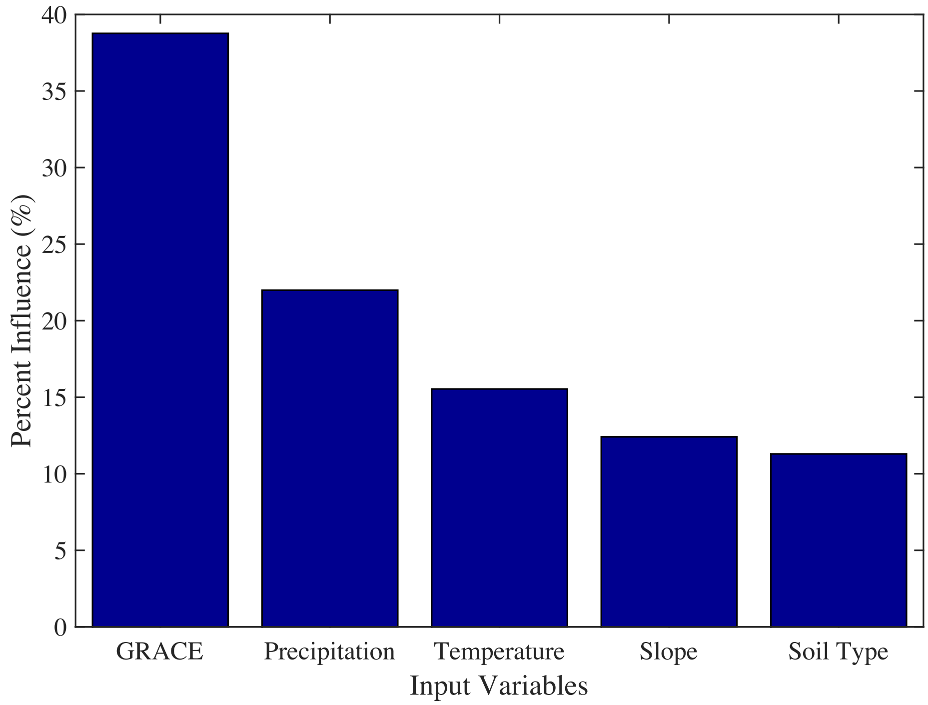

Figure 2.6 Percent influence of input variables on neural network model output.

A time series of cumulative groundwater storage change over the 2002-2010 study period is shown in Figure 2.7 as estimated from the three model approaches, from in situ data and from GRACE estimates of changes in total water storage. The comparison of the two time series show that in some years GRACE appears to be overestimating annual storage loss (2002-2004) and gain (2006) in this region. This may be due to gaps in the spatial coverage of well data, where significant groundwater pumping may be occuring. It could also be the result of surface water storage dynamics, as GRACE also detects changes in surface water. In addition, the model output may be understimating the groundwater change due to errors in the spatial interpolation methodology. The actual annual change is most likely somewhere between the two lines.

Overall, the model output demonstrates that when remote sensing and monitoring data are used together, as in our neural network model, they are able to provide a clearer picture of local and regional groundwater patterns than the use of each data type in isolation (shown in Figure 2.5). The comparison of the high-resolution maps generated in this study to regional groundwater models could further confirm or improve the effectiveness of this method. In addition, a deeper analysis of the implications of the findings in this study to local groundwater management would be highly beneficial for groundwater managers.

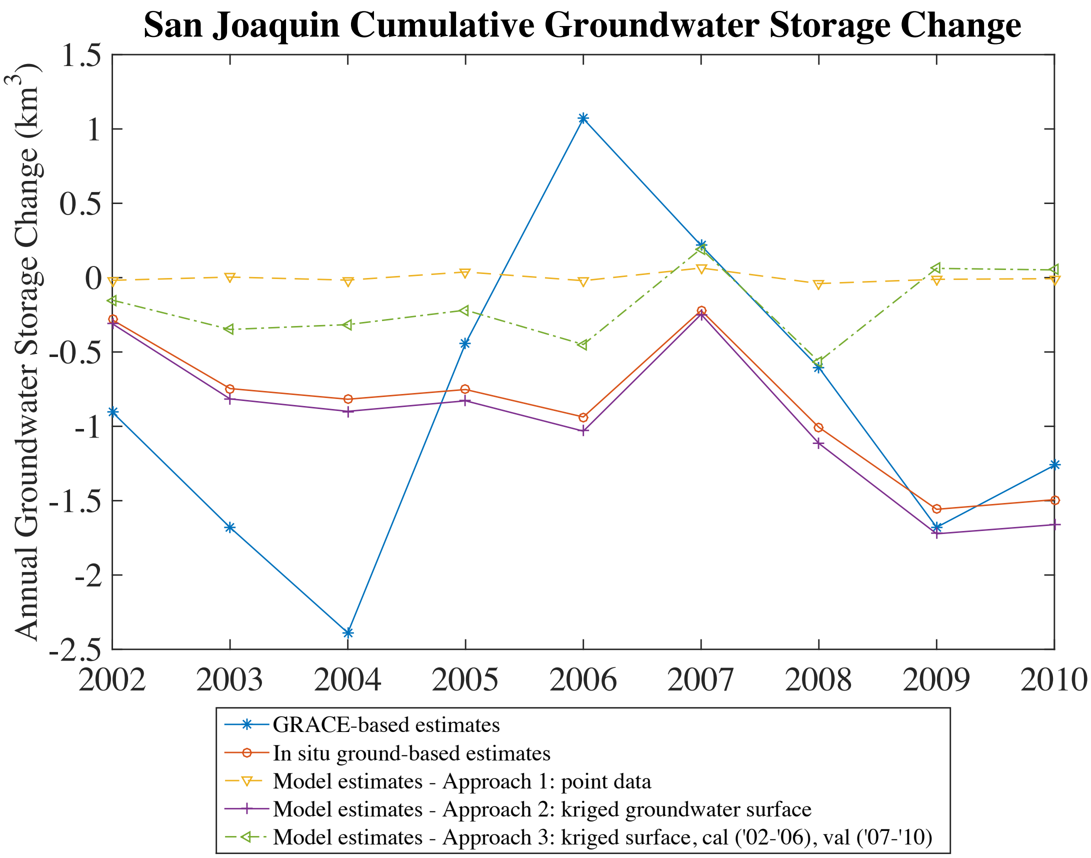

Figure 2.7 Cumulative annual groundwater storage change (km^3) for the San Joaquin groundwater basin, as estimated by GRACE water storage changes (blue), by ground-based in situ groundwater data (red) and by the three downscaling models (yellow, purple, green).

2.5 Conclusions

Sustainable planning and management of groundwater resources requires accurate information about trends in groundwater availability. GRACE has already proven to be a powerful data source for regional groundwater assessments in many areas around the world, yet its applicability to more localized studies and its utility to water management authorities has been constrained by its limited spatial resolution (~200,000 km^2) [*Famiglietti and Rodell*, 2013; *Rowlands et al.*, 2005]. We developed a robust, artificial neural network model that downscales GRACE gridded land datasets (~1 degree) to higher-resolution (~4km) groundwater storage change estimates. The model utilized GRACE estimates of variations in total water storage and a series of widely available hydrologic variables (PRISM precipitation and temperature data, DEM-derived slope and NRCS soil type) to derive spatial patterns in groundwater behavior. The neural network downscaling model was able to effectively simulate groundwater storage change over the central and southern portions of the Central Valley with NSE values ranging from 0.0391 to 0.7511. This study also showed that the model required richer estimations of groundwater data (kriged datasets) for improved calibration and validation performance. The results of the downscaling model – high-resolution maps of groundwater storage change – illustrated the high heterogeneity in groundwater behavior and the tendency for more dramatic declines in the groundwater table to occur in the southern and western portions of the San Joaquin Valley and Tulare Groundwater Basins.

Overall, the extension of GRACE data by means of numerical downscaling represents a unique contribution to the scientific remote sensing community and advances the state of current remote sensing-based hydrologic science. This approach departs from data assimilation methods in that it is model-independent and thus offers more flexibility in data scarce environments and with changing input data products (i.e. new data releases or alternate remote sensing products). This has implications for world-wide applicability in developing regions, where models and dense monitoring networks may not be freely available. This neural network model also constitutes an alternative, numerical approach to improving the resolution of remote sensing products and offers a hybrid solution between low-resolution GRACE data and sparse groundwater monitoring networks.

2.6 References

Arétouyap, Z., Njandjock Nouck, P., Nouayou, R., Ghomsi Kemgang, F.E., Piépi Toko, A.D., and J. Asfahani (2016), Lessening the adverse effect of the semivariogram model selection on an interpolative survey using kriging technique, SpringerPlus, 5(1), 549. http://doi.org/10.1186/s40064-016-2142-4

Bertoldi, G., Johnston, R. and K.D. Evenson (1991), Groundwater in the Central Valley, California – A Summary Report, U.S. Geological Survey Professional Paper 1401-A. https://pubs.er.usgs.gov/publication/pp1401A

Brosse, S., Guegan, J.F., Tourenq, J.N, and S. Lek (1999), The use of artificial neural networks to assess fish abundance and spatial occupancy in the littoral zone of a mesotrophic lake, Ecological Modelling, 120, 2–3, http://dx.doi.org/10.1016/S0304-3800(99)00110-6

Burow, K.R., Shelton, J.L., Hevesi, J.A., and G.S. Weissmann (2004), Hydrogeologic Characterization of the Modesto Area, San Joaquin Valley, California: U.S. Geological Survey Scientific Investigations Report, https://pubs.usgs.gov/sir/2004/5232/sir_2004-5232.pdf

California Department of Water Resources (2016), California's Groundwater: Bulletin 118, California Department of Water Resources. http://www.water.ca.gov/groundwater/bulletin118/gwbasins.cfm

California Department of Water Resources (2015), Legislation, Sustainable Groundwater Management Act (SGMA), California Department of Water Resources. http://www.water.ca.gov/cagroundwater/legislation.cfm

California Department of Water Resources (2003), Califorina's Groundwater: Bulletin 118 – Update 2003, California Department of Water Resources. http://www.water.ca.gov/groundwater/bulletin118/docs/Bulletin_118_Update_2003.pdf

California Water Data Library (2013), Contour Data Report, California Department of Water Resources. http://www.water.ca.gov/waterdatalibrary/

Chiew, F., Kirono, D., Kent, D., Frost, A., Charles, Timbal, S., Nguyen, K. and G. Fu (2010), Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates, Journal of Hydrology, 387: 1–2, http://dx.doi.org/10.1016/j.jhydrol.2010.03.025

Chu, H., and L. Chang (2009). "Optimal control algorithm and neural network for dynamic groundwater management." Hydrological Processes 23: 2765-2773. doi: 10.1155/2013/128393

Dagan, G. (1982), Stochastic modeling of groundwater flow by unconditional and conditional probabilities: 1. Conditional simulation and the direct problem, Water Resources Research, 18(4), 813–833, doi:10.1029/WR018i004p00813

Daliakopoulos, I., Coulibaly, P. and I. Tsanis (2005), Groundwater level forecasting using artificial neural networks, Journal of Hydrology. 309:1-4. doi: 10.1016/j.jhydrol.2004.12.001

Delhomme, J.P. (1978), Kriging in the hydrosciences, Advances in Water Resources, 1(5), 251-266, http://dx.doi.org/10.1016/0309-1708(78)90039-8

Dibike, Y. and P. Coulibaly (2006), Temporal neural networks for downscaling climate variability and extremes, Neural Networks, 19 (2), 135-144, http://dx.doi.org/10.1016/j.neunet.2006.01.003

Famiglietti, J. and M. Rodell (2014), Water in the balance, Science 340 (6138), 1300-1301. doi: 10.1126/science.1236460

Famiglietti, J.S. (2014), The global groundwater crisis, Nature Climate Change, 4(11), pp.945-948. doi:10.1038/nclimate2425

Famiglietti, J. S., Lo, M. Ho, S.L., Bethune, J., Anderson, K., Syed, T. Swenson, S. de Linage, C. and M. Rodell (2011), Satellites measure recent rates of groundwater depletion in California's Central Valley, Geophysical Research Letters 38, L03403. doi:10.1029/2010GL046442

Farr, T., Jones, C. and Z. Lui (2015), Progress report: subsidence in the Central Valley, California, Sacramento (CA): California Department of Water Resources. http://water.ca.gov/groundwater/docs/NASA_ REPORT.pdf

Fasbender, D. and T. Ouarda (2010), Spatial Bayesian Model for Statistical Downscaling of AOGCM to Minimum and Maximum Daily Temperatures, Journal of Climate. 23. doi: 10.1175/2010JCLI3415.1

Faunt, C. and M. Sneed (2015), Water Availability and Subsidence in California's Central Valley, San Francisco Estuary and Watershed Science, 13(3). Retrieved from: http://escholarship.org/uc/item/6qn2711x

Faunt, C. (2009), Groundwater Availability of the Central Valley Aquifer, California, U.S. Geological Survey. Professional Paper 1766. http://pubs.usgs.gov/pp/1766/

Fekete, B.M., Robarts, R.D., Kumagai, M., Nachtnebel, H.P., Odada, E. and A.V. Zhulidov (2015), Time for in situ renaissance, Science, 349, 6249: 685–686. doi:10.1126/science.aac7358

Feng, W., Zhong, M., Lemoine, J.M., Biancale, R., Hsu, H.T. 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(4), pp. 2110-2118. doi:10.1002/wrcr.20192

Fowler, H.J., Blenkinsop, S. and C. Tebaldi (2007), Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. International Journal of Climatology, 27, 1547–1578. doi:10.1002/joc.1556

French, M., W. Krajewski, and R. Cuykendall (1992), Rainfall forecasting in space and time using a neural network, Journal of Hydrology 137. doi: 10.1016/0022-1694(92)90046-X

Frost, A., Charles, S., Timbal, B., Chiew, F., Mehrotra, R., Nguyen, K., Chandler, R., McGregor, J., Fu, G., Kirono, D., Fernandez, E. and D. Kent (2011), A comparison of multi-site rainfall downscaling techniques under Australian conditions, Journal of Hydrology. 408. http://dx.doi.org/10.1016/j.jhydrol.2011.06.021

Garson, G. (1991), Interpreting neural network connection weights. Artificial Intelligence Expert 6, 47-51. http://dl.acm.org/citation.cfm?id=129449.129452

Gevrey, M. Dimopoulos, I., and S. Lek (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models, Ecological Modelling, 160 (3), 249-264, http://dx.doi.org/10.1016/S0304-3800(02)00257-0

Harbaugh, A., Banta, E., Hill, M. and M. McDonald (2000), MODFLOW-2000, The U.S. Geological Survey Modular Ground-water Model – User guide to moduralization concepts and the ground-water flow process, U.S. Geological Survey. Open-File Report 00-92. https://pubs.usgs.gov/of/2000/0092/report.pdf

Hendricks Franssen, H. J., and W. Kinzelbach (2008), Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, Water Resources Research, 44, W09408, doi:10.1029/2007WR006505

Howitt, R. MacEwan, D., Medellín-Azuara, J., Lund, J. and D. Sumner (2015), Economic Analysis of the 2015 Drought for California Agriculture, UC Davis Center for Watershed Science. https://watershed.ucdavis.edu/files/biblio/Final_Drought%20Report_08182015_Full_Report_WithAppendices.pdf

Hsu, K., Gupta, H.V. and S. Sorooshian (1995), Artificial Neural Network Modeling of the Rainfall-Runoff Process, Water Resources Research, 31(10), 2517–2530, doi:10.1029/95WR01955.

Hsu, K., Goa, X. Sorooshian, S. and H. Gupta (1997), Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks, J. Appl. Meteor. 36, 1176–1190. http://dx.doi.org/10.1175/1520-0450(1997)036%3C1176:PEFRSI%3E2.0.CO;2

Huang, H.Y., Capps, S., Huang, S.C. and A. Hall (2015), Downscaling near-surface wind over complex terrain using a physically-based statistical modeling approach, Clim Dyn. 44: 529. doi: 10.1007/s00382-014-2137-1

Hughes, J. and D. Lettenmaier (1981), Data requirements for kriging: Estimation and network design, Water Resources Research. doi: 10.1029/WR017i006p01641

Jet Propulsion Laboratory (2016), GRACE Follow-On, California Institute of Technology. NASA. http://gracefo.jpl.nasa.gov/

Jin, L., Lin, J. and K. Lin (2006), Precipitation Prediction Modeling Using Neural Network and Empirical Orthogonal Function Base on Numerical Weather Forecast Production, Proceedings of the 6th World Congress on Intelligent Control and Automation. Dailan, China. doi: 10.1109/WCICA.2006.1712859

Kitanidis, P. and E. Vomvoris (1983), A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations, Water Resources Research. 19:3. doi:10.1029/WR019i003p00677

Krishna, B., Satyaji Rao, Y. R. and T. Vijaya (2008), Modelling groundwater levels in an urban coastal aquifer using artificial neural networks, Hydrological Processes. 22. doi:10.1002/hyp.6686

Lallahem, S., Mania, J., Hani, A. and Y. Najjar (2005), On the use of neural networks to evaluate groundwater levels in fractured media, Journal of Hydrology. 307:1-4. http://dx.doi.org/10.1016/j.jhydrol.2004.10.005

Landerer, F. and S. Swenson (2012), Accuracy of Scaled GRACE Terrestrial Water Storage estimates, Water Resources Research. doi:10.1029/2011WR011453

Lo, M. and J. Famiglietti (2013), Irrigation in California's Central Valley strengthens the southwestern U.S. water cycle, Geophysical Research Letters 40: 301-306. doi:10.1002/grl.50108

MacKay, D. (1992), Bayesian Interpolation, Neural Computation, Vol. 4, No. 3, pp. 415–447. doi: 10.1007/978-94-017-2219-3_3

Martin, P.J. and E.G. Frind (1998), Modeling a Complex Multi-Aquifer System: The Waterloo Moraine. Ground Water, 36: 679–690. doi:10.1111/j.1745-6584.1998.tb02843.x

Mathworks (2016), Neural Network Toolbox, MATLAB. http://www.mathworks.com/products/neural-network/

Mogheir, Y., De Lima, J.L.M.P. and V.P. Singh (2005), Assessment of Informativeness of Groundwater Monitoring in Developing Regions (Gaza Strip Case Study), Water Resources Management 19: 737. doi:10.1007/s11269-005-6107-6

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., and T. Veith (2007), Model evaluation guidelines for systemic quantification of accuracy in watershed simulations, Transactions of the ASABE. 50(3):885-900. doi: 10.13031/2013.23153

Munier, S., Becker, M., Maisongrande, P. and A. Cazenave (2012), Using GRACE to detect Groundwater Storage variations: the cases of Canning Basin and Guarani Aquifer System, International Water Technology Journal, 2, pp.2-13. http://iwtj.info/wp-content/uploads/2012/11/V2-N1-P1.pdf

Nash, J. E., and J.V. Sutcliffe (1970), River flow forecasting through conceptual models: Part 1. A discussion of principles, J. Hydrology 10(3): 282-290. http://dx.doi.org/10.1016/0022-1694(70)90255-6

Nikroo, L., Kompani-Zare, M., Sepaskhah, A., Rashid, S. and F. Shamsi (2010), Groundwater depth and elevation interpolation by kriging methods in Mohr basin of Fars Province in Iran, Environmental Monitoring and Assessment, 166 , 387–407 http://dx.doi.org/10.1007/s10661-009-1010-x

NRCS (2014), Gridded Soil Survey Geographic (gSSURGO) Database, USDA Natural Resources Conservation Service. https://gdg.sc.egov.usda.gov/

Olden, J. and D. Jackson (2002), Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks, Ecological Modelling, 154 (1–2), 135-150, http://dx.doi.org/10.1016/S0304-3800(02)00064-9

PRISM Climate Group (2014), PRISM Climate Data, Northwest Alliance for Computational Science and Engineering. Oregon State University. http://www.prism.oregonstate.edu.

Provost and Pritchard Consulting Group (2014), Madera Regional Groundwater Management Plan. sgma.water.ca.gov/basinmod/docs/download/824

Reager, J.T., and J. Famiglietti (2013), Characteristic mega-basin water storage behavior using GRACE, Water Resources Research: 3314-3329. doi:10.1002/wrcr.20264

Reed, P., Minsker, R. and A. Valocchi (2000), Cost-effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation, Water Resources Research. 36:12. doi: 10.1029/2000WR900232

Richey, A.S., Thomas, B.F., Lo, M.H., Reager, J.T., Famiglietti, J.S., Voss, K., Swenson, S. and M. Rodell (2015a), Quantifying renewable groundwater stress with GRACE, Water Resources Research, 51(7). doi:10.1002/2015WR017349

Richey, A.S., Thomas, B.F., Lo, M.H., Famiglietti, J.S., Swenson, S. and M. Rodell (2015b), Uncertainty in global groundwater storage estimates in a Total Groundwater Stress framework, Water Resources Research, 51(7). doi:10.1002/2015WR017351

Rowlands, D., Luthcke, S., Klosko, S., Lemoine, F., Chinn, D., McCarthy, J., Cox, C. and O. Anderson (2005), Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, Geophys. Res. Lett. doi:10.1029/2004GL021908

RMC (2014), An Evaluation of California Groundwater Management Planning, California Water Foundation. Sacramento, CA. http://www.californiawaterfoundation.org/uploads/1405009350-GMPReport2014(00256304xA1C15).pdf

Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J. and C.R. Wilson (2007), Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE, Hydrogeology Journal, 15(1). doi:10.1007/s10040-006-0103-7

Rodell, M. and J.S. Famiglietti (2002), The potential for satellite-based monitoring of groundwater storage changes using GRACE: the High Plains aquifer, Central US, Journal of Hydrology, 263(1). http://dx.doi.org/10.1016/S0022-1694(02)00060-4

Rodell, M., Velicogna, I. and J.S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, Nature, 460(7258). doi:10.1038/nature08238

Schoof, J. and S.C. Pryor (2001), Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. International Journal of Climatology, 21, 773–790. doi:10.1002/joc.655

Schoof, J. (2013), Statistical Downscaling in Climatology, Geography Compass. 7:4. doi: 10.1111/gec3.12036

Shah T., Molden, D. Sakthivadivel, R. and D. Seckler (2000), The global groundwater situation: Overview of opportunities and challenges, Colombo, Sri Lanka. International Water Management Institute. http://publications.iwmi.org/pdf/H025885.pdf

Sneed, M., Brandt, J. and M. Solt (2013), Land subsidence along the Delta-Mendota Canal in the northern part of the San Joaquin Valley, California, 2003–10, U.S. Geological Survey Scientific Investigations. http://pubs.usgs.gov/sir/2013/5142/

Song, K., Park, Y.S., Zheng, F. and H. Kang (2013), The application of Artificial Neural Network (ANN) model to the simulation of denitrification rates in mesocosm-scale wetlands, Ecological Informatics, 16, 10-16, http://dx.doi.org/10/1016/j.ecoinf.2013.04.002

Strassberg, G., Scanlon, B. and M. Rodell (2007), Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA)", Geophys. Res. Lett., 34, 14402, doi:10.1029/2007GL030139.

Sun, A. (2013), Predicting groundwater level changes using GRACE data, Water Resources Research 49: 5900-5912. doi:10.1002/wrcr.20421

Sun, Y., Kang, S., Li, F. 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 and Software. 24:10. http://dx.doi.org/10.1016/j.envsoft.2009.03.009

Taormina, R., Chau, K. and R. Sethi (2012), Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon, Engineering Applications of Artificial Intelligence. 25:8. doi:10.1016/j.engappai.2012.02.009

Taylor, C. and W. Alley (2001), Ground-water-level monitoring and the importance of long-term water-level data, U.S. Geological Survey Circular 1217. USGS. Denver, CO. http://pubs.usgs.gov/circ/circ1217/pdf/circular1217.pdf

Turban, E., Sharda, R., Aronson, J. E., and D.N. King (2008), Business Intelligence: A Managerial Approach, Pearson Prentice Hall. Upper Saddle River, N.J.

USGS (2009), "1/3-Arc Second National Elevation Dataset", U.S. Geological Survey. http://nationalmap.gov

Voss, K., Famiglietti, J.S., Lo, M., de Linage, C., Rodell, M. 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 Resour. Res., 49, doi:10.1002/wrcr.20078.

Wiese, D.N. (2015), GRACE monthly global water mass grids, NETCDF RELEASE 5.0. Ver. 5.0. PO.DAAC, CA, USA. http://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/

Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P. and L.O. Mearns (2004), Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods, Task Group of Data and Scenario Support for Impacts and Climate Analysis (TGICA). Intergovernmental Panel on Climate Change.

Wilby, R.L., and T.M.L. Wigley (1997), Downscaling general circulation model output: a review of methods and limitations, Progress in Physical Geography: 530-548. doi: 10.1177/030913339702100403

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and D.S. Wilks (1998), Statistical downscaling of general circulation model output: A comparison of methods, Water Resour. Res., 34(11), 2995–3008, doi:10.1029/98WR02577.

Yang, Z., Lu, W., Long, Y., and P. Li (2009), Application and comparison of two prediction models for groundwater levels: A case study in Western Jilin Province, China, Journal of Arid Environments. 73:4-5. http://dx.doi.org/10.1016/j.jaridenv.2008.11.008

Yeh, P., Swenson, S., Famiglietti, J. and M. Rodell (2009), Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), Water Resources Research. 42:12. doi:10.1029/2006WR005374

Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O. and K.K. Lee (2011), A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer, Journal of Hydrology. 396:1-2. http://dx.doi.org/10.1016/j.jhydrol.2010.11.002

Zaitchik, B., M. Rodell and R. Reichle (2008), Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin, J. Hydrometeor, 9, 535–548. http://dx.doi.org/10.1175/2007JHM951.1

Zhu, A. (2000), Mapping soil landscape as spatial continua: The Neural Network Approach, Water Resources Research: 663. doi:10.1029/1999WR900315

Zimmerman, D., Pavlik, C., Ruggles, A. and M. Armstrong (1999), An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting, Mathematical Geology. 31:4. doi:10.1023/A:1007586507433

3. A hydrologically-based framework for calculating sustainable yield under California's Sustainable Groundwater Management Act (SGMA)

3.1 Introduction

Over the past 100 years of agricultural development in California, the state has lacked an established framework for groundwater management [*Grantham and Viers,* 2014]. This lack of oversight has led to dramatic rates of groundwater depletion throughout California's Central Valley, a region with high agricultural demand for water but variable climatology [*Faunt,* 2009]. The Central Valley is semi-arid and drought-prone, with highly variable wet and dry years [*Faunt et al.*, 2016]. In dry years, when surface water supplies diminish, farmers depend heavily on groundwater resources to meet their water needs [*Faunt*, 2009]. Since the 1960's, recharge during wet periods has failed to replenish the high rates of depletion during dry periods, and the Central Valley has lost nearly 100 million acre-feet of groundwater from storage [*Faunt et al.*, 2016]. On average, this amount of water would be enough to supply the state's urban areas for 12 years [*Pacific Institute*, 2014].

Beyond reduction in groundwater storage, the region has seen significant impacts to both groundwater quality and the land surface. As groundwater has been removed from the system, the geology of the Central Valley's subsurface has caused significant land subsidence, or compaction of the land surface, at rates of nearly 20 inches per year in some locations [*Faunt et al.,* 2016; *Farr et al.*, 2017]. Increasing nitrate concentrations in groundwater due to extensive agriculture, particularly in the eastern portion of the Central Valley, has also jeopardized groundwater availability [*Burow et al.*, 2013; *Dzurella et al.*, 2015]. This is problematic for both agricultural and urban users, as migration of nitrate contaminant plumes has been identified between subbasins [*Burow et al.*, 2013]. Agricultural sustainability in the Central Valley is also threatened by soil salinization and saline intrusion into groundwater basins, which are secondary and tertiary effects to increased removal of groundwater from storage [*Schoups et al*., 2005].

Over time, the accumulation of these physical impacts and the growing recognition of the importance of groundwater as a source of water for both agricultural and urban regions led to the creation and passage of the 2014 Sustainable Groundwater Management Act (SGMA). SGMA is a complete overhaul for groundwater management across the state. It requires the formation of Groundwater Sustainability Agencies (GSAs), who must define their boundaries and develop and implement Groundwater Sustainability Plans (GSPs) [*Sustainable Groundwater Management Act,* 2014]. This is no easy task. Not only are most of these agencies brand new to direct groundwater management, but much of the Central Valley also lacks adequate monitoring and information for management [*Croyle*, 2014].

GSAs must complete their GSPs by 2020 - for critically-overdrafted groundwater basins - and by 2022 for all other non-adjudicated basins [*CA DWR*, 2016]. GSAs will then have an additional 20 years to fully implement these plans and achieve sustainability [*CA DWR*, 2016]. While GSAs retain local authority over much of SGMA implementation, the state maintains a regulatory "backstop" in their ability to review GSPs and intervene if they are deemed ineffective [*State Water Resources Control Board*, 2016]. A more lengthy description of the requirements, timelines and best management practices for SGMA implementation can be found by accessing California's Department of Water Resources (DWR) SGMA Toolbox [*CA DWR*, 2017].

This chapter will focus on clarifying and strengthening one component of GSPs – the concept of sustainable yield – a key prerequisite to introducing measurements to curb depletion rates, subsidence and water quality degradation. SGMA mandates that GSAs manage annual groundwater pumping (also termed extractions) according to the sustainable yield of their basin. The language of SGMA defines sustainable yield as the "maximum quantity of water… that can be withdrawn annually from a groundwater supply without causing an undesirable result" [*Sustainable Groundwater Management Act*, 2014]. Undesirable results include 1) "depletion of supply from chronic lowering of groundwater levels"; 2) "reduction of groundwater storage"; 3) "seawater intrusion"; 4) "water quality degradation"; 5) "land subsidence"; and 6) "adverse impacts from the depletion of interconnected surface waters" [*Sustainable Groundwater Management Act*, 2014]. The concept of sustainable yield is based on the notion that by limiting extractions to the sustainable yield of the system, basin managers will not only be preventing the undesirable results listed above, but they will also be prolonging the useful life of the aquifer. Because an important focus of SGMA is its strong emphasis on local management, how to approach the calculation of sustainable yield is therefore left to the discretion of each newly formed GSA [*Joseph,* 2016].

Lessons from California's own history of implementing sustainable yield, and its precursor safe yield, in adjudicated basins point to the need for a consistent approach. *Langridge et al.* [2016] cite the use of six different definitions of safe yield statewide, each with highly variable methods for calculation. *Rudestam and Langridge* [2014] also detail California water agency representatives' lack of clear understanding of and confidence in the concept of sustainable yield. It is also important to note that the majority of the State's experiences with the concept of sustainable yield come from adjudicated basins - those in which parties have previously filed suit for and were granted groundwater rights. These basins, most of them located outside of the Central Valley in urban and semi-urban areas throughout Southern California, are essentially exempt from SGMA implementation [*Langridge et al.,* 2016; *State of California,* 2017].

We contend that SGMA's current definition for sustainable yield is not only ambiguous but also lacks grounding in physics. Instead, we believe that integrating hydrologically-sound methods into the concept of sustainable yield offers a more robust framework for groundwater resource management. This paper suggests one approach to calculating sustainable yield that is based on a synthesis of scientific inquiry and analysis. Rather than focusing on providing yet another definition for sustainable yield, we introduce a flexible framework that basin managers can rely on to quantify sustainable yield values and analyze their impacts of over the management horizon. We apply this three-step framework in the context of a case study of the South San Joaquin Irrigation District (SSJID) in California's Central Valley. In doing so, we assess the long-term applicability of the sustainable yield framework by performing a groundwater balance through the planning horizon of SGMA, to 2040.

3.2 The challenge of defining sustainability for groundwater management

The idea of sustainable yield or its predecessor, safe yield, is not new. In the 1970s, the "modern" approach for a hydrologist to determine safe yield consisted of creating a complete and detailed model of the groundwater system that could characterize the response of both groundwater and surface water to various stresses [*Lohman*, 1972]. With this model, the hydrologist and the groundwater manager would then work together to determine the most equitable distribution of water [*Lohman*, 1972]. However, even with this, ambiguity remained on the actual definition. *Lohman* [1972] states, "the term 'safe yield' has about as many definitions as the number of people who have defined it". *Thomas* [1951] writes that safe yield is an "Alice-in-Wonderland term which means whatever its user chooses".

While debate continues, a general consensus in the community has emerged. Many hydrologists that have written extensively on the concept of safe and sustainable yield have highlighted the importance of physically based, numerical modeling approaches [*Bredehoeft et al.*, 1982; *Zhou*, 2009]. In one such debate, researchers have shown that sustainable yield values can be calculated based on mass balance approaches [*Sophocleous*, 2000; *Kalf and Woolley*, 2005; *Zhou*, 2009] as long as the sustainable yield values (i.e. pumping) are significantly less than recharge [*Sophocleous*, 2000]. If, instead, the sustainable yield is allowed to equal the natural recharge rate, undesired "capture" of natural outflow to surface water bodies can occur [*Lohman*, 1972; *Balleau and Mayer,* 1988; *Bredehoeft*, 2002; *Devlin and Sophocleous*, 2006]. In this case, groundwater abstractions are supplied by both surface and groundwater and may cause reductions in surface water availability, complicating surface water right allocations [*Theis,* 1940*; Alley et al.,* 1999*; Seward et al.,* 2006]. Ultimately, without proper accounting or accurate numerical models, groundwater managers cannot ensure the sustainability of local groundwater and interconnected surface water systems.

However, groundwater sustainability is relevant to more than just the physical system. In the end, "groundwater has value only by virtue of its use", and sustainable yield is the volume of groundwater that will supply an overlaying socio-economic system [*Freeze and Cherry*, 1979]. Hydrologists have discussed the interconnection of physical and human system in the context of groundwater management, highlighting importance of social, economic and environmental constraints on groundwater resources [*Alley et al.,* 1999; *Alley and Leake*, 2004; *Maimone,* 2004; *Devlin and Sophocleous*, 2006; *Zhou,* 2009]. The quantification of these non-physical constraints, though, depends largely on the local socio-economic environment surrounding groundwater use [*Rudestam and Langridge*, 2014]. From this perspective, sustainability is therefore best defined locally based on an analysis of the “social acceptability of impacts” [*Herczeg and Leaney*, 2002; *Rudestam and Langridge*, 2014].

Under SGMA, GSAs must determine what impacts to both the local physical groundwater system and the overlaying socio-economic system are bearable. This is written into SGMA through the prevention of undesirable results. Recall SGMA’s definition of sustainable yield: “maximum quantity of water… that can be withdrawn annually from a groundwater supply without causing an undesirable result”, in which undesirable results can be viewed as six indicator variables that help GSAs determine and maintain sustainability [*Sustainable Groundwater Management Act*, 2014]. GSAs must quantify acceptable thresholds for groundwater depletion, quality degradation, land subsidence and streamflow reduction and not surpass them – underscoring the importance of correctly identifying the “social acceptability of impacts” when assessing long-term sustainability [*Herczeg and Leaney*, 2002]. For example, in the context of groundwater quantity, a GSA may decide to pursue no net groundwater depletion in their basin or may restrict pumping altogether (see Figure 3.1). These decisions would in turn govern how much water is extracted and recharged annually and how groundwater levels recover over time.

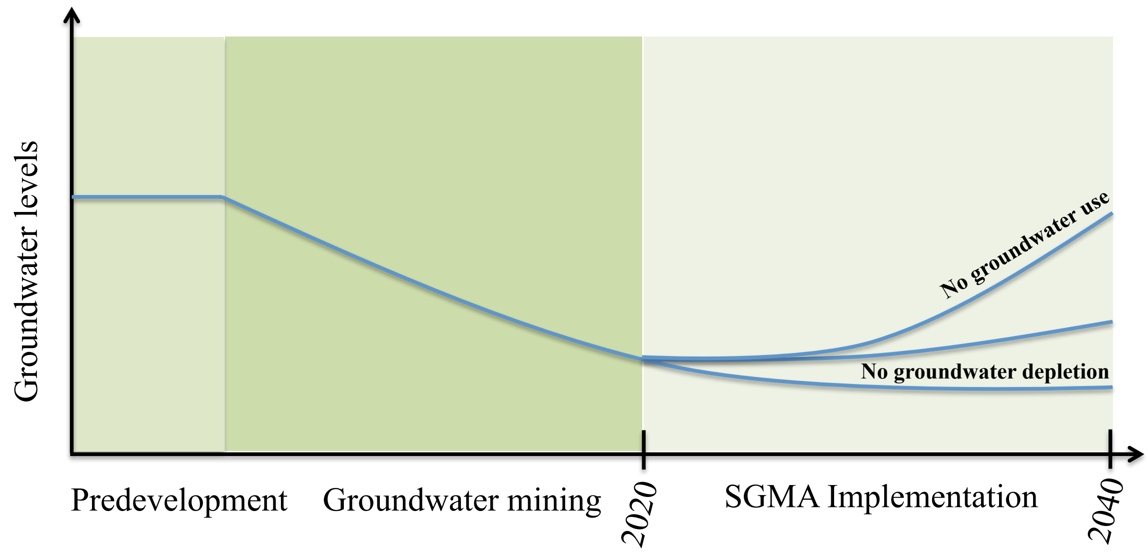

Figure 3.1 Sustainable yield management scenarios, adapted from *U.S. Department of Interior* [2005].

3.3 Case study – the South San Joaquin Irrigation District

The South San Joaquin Irrigation District (SSJID) covers 72,000 acres of the Eastern San Joaquin Groundwater Basin in the northeastern San Joaquin Valley [*Nakagawa,* 2004]. The district was formed in 1909 and holds pre-1914 water rights to 600,000 acre-feet of the Stanislaus River, which it shares equally with a neighboring district – the Oakdale Irrigation District [*Nakagawa,* 2004; *Davids Engineering*, 2015]. Most of the water available in the SSJID is used to irrigate approximately 54,000 acres of cultivated land [*Davids Engineering*, 2015]. However, agricultural water demands are not fully met by available surface water, and groundwater serves as an important resource, especially during dry periods [*Davids Engineering*, 2015]. As a result, the SSJID has experienced significant groundwater declines of 30-40 feet over the past several decades, an average of 1.7 feet per year [*Nakagawa,* 2004; *CA DWR,* 2006].

The estimated safe yield for the district is 72,000 acre-feet per year [*Nakagawa,* 2004]. During the irrigation season, average recharge is estimated at 97,000 acre-feet and average pumping at 40,000 acre-feet. During dry years, average pumped volumes surpass 65,000 acre-feet. Within the boundaries of the SSJID are also three urban areas whose potable water supply is fully dependent on groundwater – the cities of Escalon, Ripon and Manteca [*Nakagawa,* 2004; *Black Water Consulting Engineers*, 2016]. Nitrate contamination and declining groundwater levels are impacting the availability of groundwater for domestic use [*Nakagawa,* 2004; *Black Water Consulting Engineers*, 2016].

The district manages 28 groundwater wells, from which groundwater is pumped into open channels for distribution; well depths range from 80 to 800 feet (350 feet on average) [*Davids Engineering*, 2015]. Wells penetrate into the top two layers of four permeable water-bearing formations that are comprised of sands, gravels, silts and clays [*Davids Engineering*, 2015]. The presence of finer sediments has contributed to land subsidence in the region [*CA DWR,* 2006]. More details on the geology of the subsurface, specific yield values and groundwater storage capacity can be found in Basin 5-22.01 of California's Bulletin 118 [*CA DWR*, 2006].

The California Department of Water Resources has identified the Eastern San Joaquin Groundwater Basin as one of the state's critically-overdrafted basins [*CA DWR*, 2016]. This means that GSPs must be developed by 2020 and fully implemented to "sustainability" by 2040 [*CA DWR*, 2016]. SSJID has applied to be a GSA, along with 17 other existing water management agencies in the Eastern San Joaquin subbasin [*CA DWR*, 2017]. If all of the current GSA proposals are accepted, the SSJID will have to coordinate with seven neighboring GSAs that either fully or partially overlap with the SSJID's boundaries [*CA DWR*, 2017].

3.4 A novel framework to calculate sustainable yield

Under SGMA, the process of quantifying sustainable yield can be divided into three steps – 1) Quantification of a baseline sustainable yield value; 2) Identification of potential constraints (from undesirable results) to the baseline sustainable yield from Step 1 and determination of a constraint-adjusted sustainable yield; and 3) Projection of basin response to use of sustainable yield-based strategies over the management horizon.

3.4.1 Step 1: Quantifying a baseline sustainable yield

We take a graphical approach to calculating sustainable yield that is based on a basin's response to pumping levels under both wet and dry hydrologic conditions. This approach, often termed Hill's Method, is a widely used empirical method that relates basin-wide change in groundwater levels to groundwater extraction (pumping) [*Conklin,* 1946; *Butler et al.*, 2016; *Loáiciga*, 2016; *Whittemore et al.,* 2016]. According to this method, which is shown graphically in Figure 3.2, the baseline sustainable yield corresponds to the average level of extraction (measured along the x-axis) that causes zero average groundwater level change (measured along the y-axis). In practice, a best-fit line is used to calculate the sustainable yield value from a plot of average annual groundwater levels, calculated from Equation 3.1, and annual groundwater extraction, as reported or estimated by local basin managers [*Loáiciga*, 2016].

$$\Delta GWL = \frac{1}{Area} \sum_{k=1}^{K} \Delta GWL_k * Area_k \qquad (3.1)$$

Where, ΔGWL_k is the average annual change in groundwater level for a given well, k, in a groundwater basin (ft); $Area_k$ is the representative spatial extent of well k in the basin (ft^2).

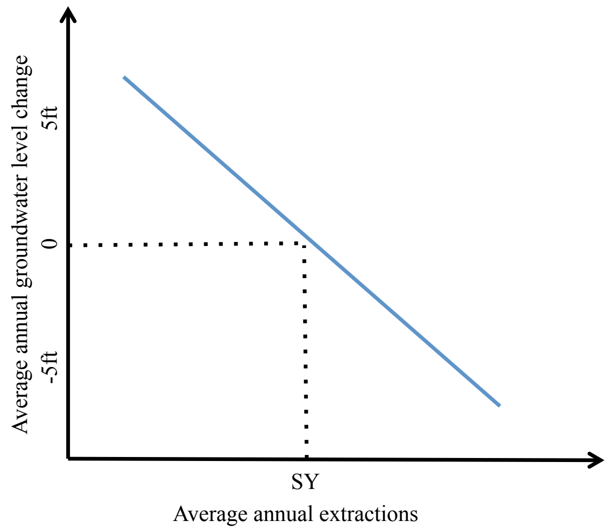

Figure 3.2 Hill's method - Estimating sustainable yield from changes in groundwater levels

Hill's method offers a number of advantages, including simplicity [*Hill*, 2006; *Butler et al.*, 2016], flexibility to varying types and amounts of data [*Loáiciga*, 2016] and ability to represent a variety of spatial scales, or potential management units. Notably, it is appropriate for areas where monitoring data may be limited and where only basin-wide extraction values and groundwater hydrographs from a few sample wells are available [*Loáiciga*, 2016]. This is the case in much of the Central Valley, particularly if we consider the narrow spatial scale of management proposed by many potential GSAs. In many locations groundwater fluxes (lateral flows, etc.) are also often poorly understood, and those estimates that do exist often do not vary with time. Thus, at this stage, little data is available to complete a full groundwater balance. Groundwater models are good tools to substitute direct readings of groundwater levels from wells, and are encouraged by DWR as a best management practice for SGMA [*CA DWR*, 2016]. As these are developed, model output for groundwater level change can be incorporated into this framework. Improvements to monitoring networks and availability of groundwater models, both of which may be an outgrowth of SGMA implementation, will provide greater spatial coverage, enhancing the accuracy of Hill's method.

Hill's method does, however, have a few limitations. First, it represents the groundwater level response to pumping and therefore cannot fully capture the dynamics of confined aquifers. Second, the method assumes constant recharge to the basin and is thus limited in its ability to accurately estimate sustainable yield under highly variable climatology. As a result of these constraints, the method is most applicable to basins that are both predominantly unconfined and receive relatively consistent inter-annual rainfall. For basins that do not fall in that category, we propose two modifications to this approach so that it remains applicable in the context of GSPs.

The first modification is relevant to the case of a confined system. Annual change in groundwater storage – calculated from known storativity values – can be used in place of groundwater levels on the y-axis (see Figure 3.2).

The second modification to Hill's method mirrors existing water management frameworks and will allow basin managers to account for the impact of higher or lower recharge amounts. This modification consists of determining a separate sustainable yield for dry and wet years. Each year, beginning in February and based on river flow in the Sacramento and San Joaquin regions, the California's State Water Resources Control Board determines whether the incoming hydrologic year will be a wet or a dry year [*CA DWR*, 2013]. Groundwater basin managers can use this information to predict both regional surface water allocations and potential recharge amounts, as well as to calculate a separate sustainable yield for each hydrologic year type. Figure 3.3 illustrates this modification. Average groundwater level change (from Equation 3.1) and the corresponding extraction volume are plotted for each year but are divided by hydrologic year type. A best-fit line is then used to calculate a separate sustainable yield value for each hydrologic year type. As expected, the sustainable yield is lower in a dry year than in a wet year.

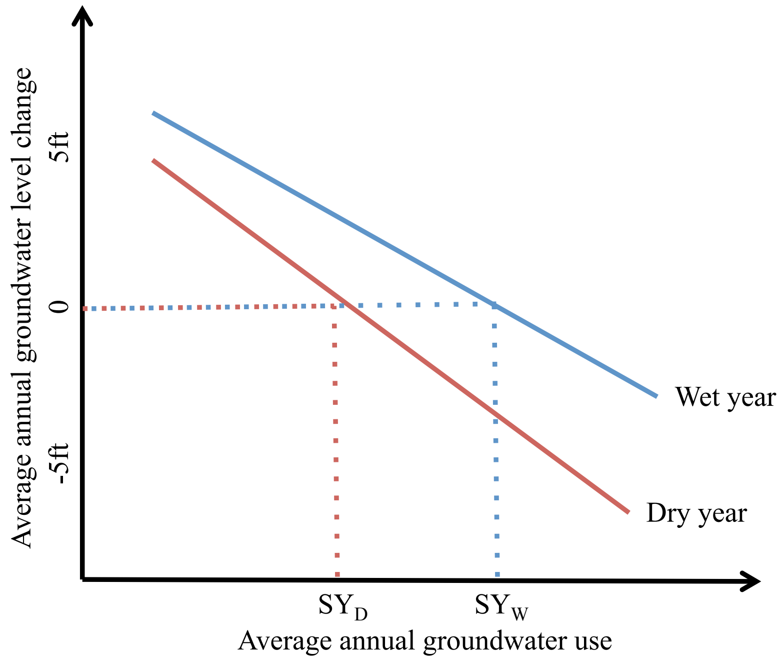

Figure 3.3 Modified Hill's Method to capture Central Valley's wet and dry year climatology

We apply the second modification to Hill's method and calculate a baseline wet and dry year sustainable yield for the SSJID. We use six sample wells equally distributed across the district with temporally continuous and error-free well logs. We use this data to calculate the average annual groundwater change from Equation 3.1. Annual extraction volumes were available from the SSJID Agricultural Management Plan from 1994-2014 [*Davids Engineering*, 2015]. From Figure 3.4, the wet year sustainable yield is estimated at 73,608 acre-feet, and the dry year sustainable yield at 20,686 acre-feet. These yields correspond to the baseline sustainable yields that will be used as inputs for Step 2 of our framework.

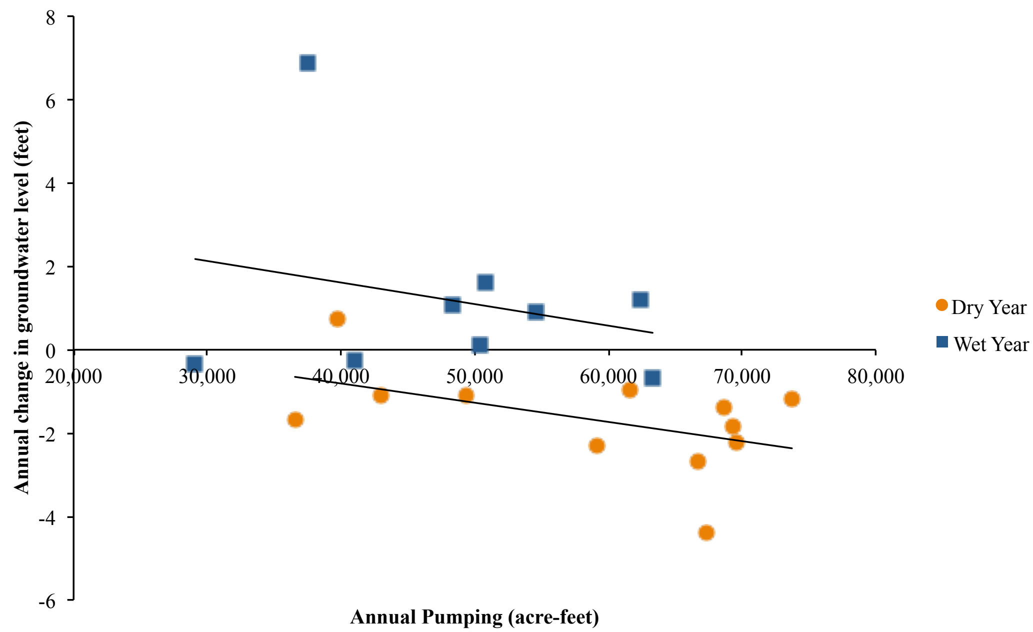

Figure 3.4 Sustainable yield calculation for SSJID (1994-2014)

3.4.2 Step 2: Determining constraints to the baseline sustainable yield

Step 2 characterizes the relationship between pumping, groundwater depletion and the different undesirable results to determine if a constraint exists to the baseline sustainable yield. This step is based on the premise that a certain level of groundwater depletion or pumping may trigger the occurrence of an undesirable result. If this depletion level or pumping volume is known, the sustainable yield value may need to be reduced to ensure avoidance of the undesirable result.

SGMA mandates consideration of six undesirable results but does not specifically quantify what constitutes an "undesirable" level, or threshold, for each. Instead, quantification of these thresholds is left to each GSA. Some standards already exist to guide GSAs in establishing threshold values for each undesirable result [*Sustainable Groundwater Management Act*, 2014]. Table 3.1 contains potential threshold values that GSAs can refer to, based on established maximum values from federal or state regulations. Where definitive numerical thresholds do not exist in the literature, documents that offer relevant guidance are included as a reference. The thresholds shown in Table 3.1 are only guidelines values. Each threshold should be extensively studied and quantified based on the local properties of each management basin, including its subsurface, land use and stream networks, as well as reflective of local socio-economic preferences. *Christian-Smith and Abhold* [2015] also offers recommendations for quantifying thresholds.

Table 3.1 Undesirable Results thresholds and/or recommended references for GSAs

Undesirable Result	Thresholds - Reference texts
Depletion of supply • Chronic lowering of groundwater levels • Reduction in groundwater storage	*Sustainable yield management, including:* • Aquifer sustainability [*Custodio*, 2002; *van der Gun and Lipponen*, 2010; *Richey et al.*, 2015] • Socio-economic dimensions [*Foster and Loucks, 2006*; *MacEwan et al.*, 2017]
Saline intrusion	*Total dissolved solids (TDS): 500 mg/L* • SWRCB recommendation for drinking water [*SWRCB*, 2016] • Negative impacts on crops [*SWRCB*, 2016]
Water quality degradation	*Nitrate concentrations: 3 mg/L (recommended); 10 mg/L (maximum)* • Concentrations over 10 mg/L are unsafe for human consumption [*US EPA,* 2016] • 20-40 mg/L impact livestock [*Crowley et al.*, 1974]
Land subsidence	*Rates over 1 foot/year* • 0.5 - 1 feet/year are historically high rates [*Faunt et al.,* 2016] • Billions in estimated costly impacts to infrastructure [*Borchers and Carpenter*, 2014]
Surface water depletion	*Avoidance of baseflow reduction* • Surface water rights fully allocated [*Grantham and Viers*, 2014] • Groundwater tables depend on recharge from surface water [*Scanlon et al.*, 2012] • Numerical modeling of groundwater-surface water interaction is critical [*Sophocleous*, 2002; *Alley and Leake*, 2004; *Faunt*, 2009]

In the previous section, we introduced a method that derives linear relationships between groundwater levels (or storage) and groundwater extraction to calculate baseline sustainable yield values for wet and dry years. Similarly, we can form linear relationships between each undesirable result in Table 3.1 and groundwater levels. Because the quantification of sustainable yield in Step 1 is based on a threshold for groundwater level change, Step 2 is focused on the groundwater quality, land subsidence and surface water undesirable results. In addition, the wet year-dry year distinction becomes less discernible in the relationship between groundwater level and nitrate concentration, for example, so the aggregate data is used.

To apply this step to the SSJID, we gathered data on each of the undesirable results. Land subsidence was obtained data from the nearest GPS monitoring station outside Salida, CA, which began recording in 2008 [*UNAVCO*, 2017]. Groundwater quality and salinity indicators – dissolved nitrogen and TDS – were obtained from four available monitoring wells from the California Water Data Library [*CA DWR*, 2017]. Streamflow data was obtained from U.S. Geological Survey (USGS) stream gauge #11303000 located on the Stanislaus River near Ripon, CA [*USGS*, 2017]. Figures 3.5a-3.5d show the derived relationships between groundwater levels and each of the four undesirable results variables.

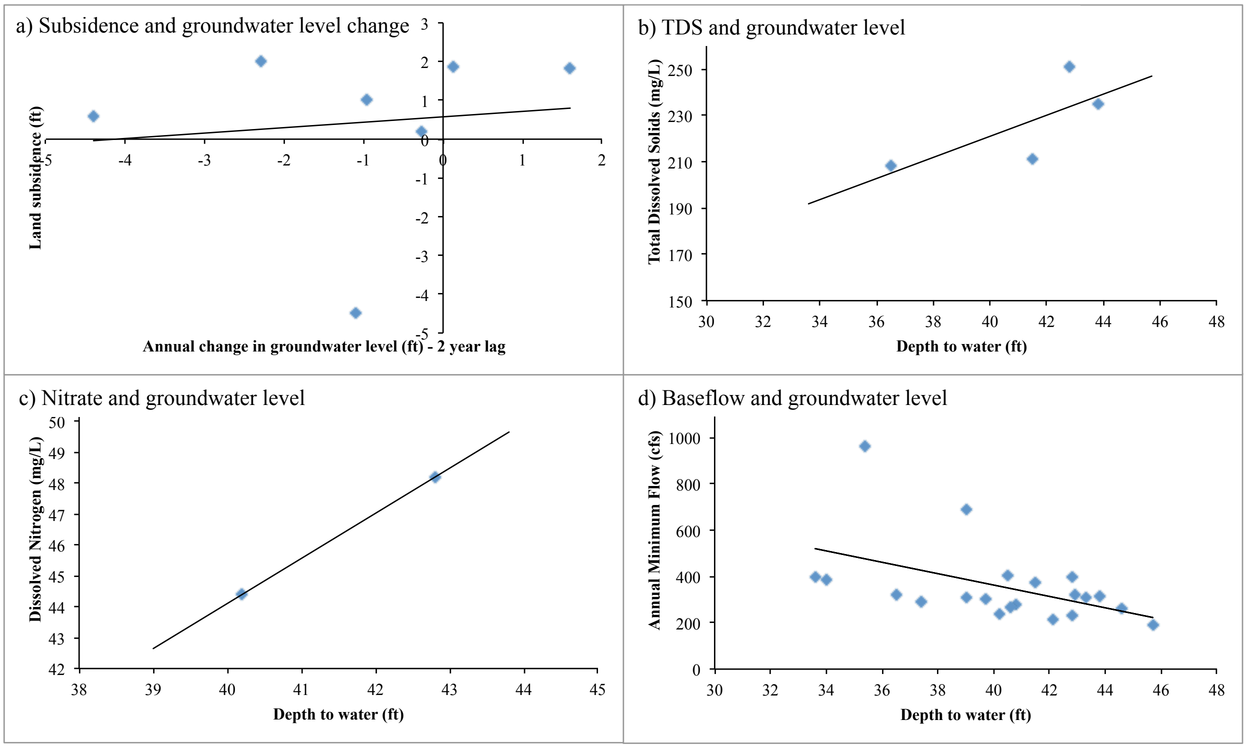

Figure 3.5 Undesirable results variables and relationship to groundwater level

Using the best-fit lines from the relationships derived in Figure 3.5, groundwater level thresholds for each undesirable result can be calculated as shown in Table 3.2.

Table 3.2 Best-fit lines between groundwater level and each undesirable result

Variable	Equation	R^2	Threshold	GWL Threshold*
Subsidence	*Subsidence = 0.14*ΔGWL + 0.58*	0.014	-1 ft/year	-11.27 feet/year**
Salinity	*TDS = 4.56*GWL + 38.66*	0.522	500 mg/L	101.2 feet
Nitrate	*Nitrate = 1.46*GWL - 14.35*	1	40 mg/L	37.19 feet
Streamflow	*Streamflow = -24.59*GWL + 1346.5*	0.231	200 cfs	46.63 feet

*Note: *GWL Thresholds indicate the depth to water surface at which undesirable results thresholds would be surpassed. **Subsidence was not found to have a strong linear relationship with groundwater levels, and as a result annual change in groundwater level with a two-year lag was used.*

To connect Step 1 to Step 2, we need to relate the groundwater level thresholds in Table 3.2 to the baseline sustainable yield values. To do so, we start with the known depth to the water surface in a given year. If, for example, the current average annual groundwater level was measured at a depth to water of 50 feet, and using streamflow as the constraint variable, Table

3.2 would indicate that groundwater levels would need to be raised by 3.37 feet to either reverse or prevent reduction in streamflow. In this case, the sustainable yield during a wet year would be calculated from Hill's Method, shown in Figure 3.6, using 3.37 feet as the y-intercept rather than a 0 feet change in groundwater level. This is a hypothetical example. In the case of the SSJID, the current sustainable yield values are constrained by groundwater level and storage change and the baseline sustainable yield values calculated in Step 1 hold. If groundwater levels drop in the future, a surface water constraint to sustainable yield, similar to the hypothetical example, may need to be implemented.

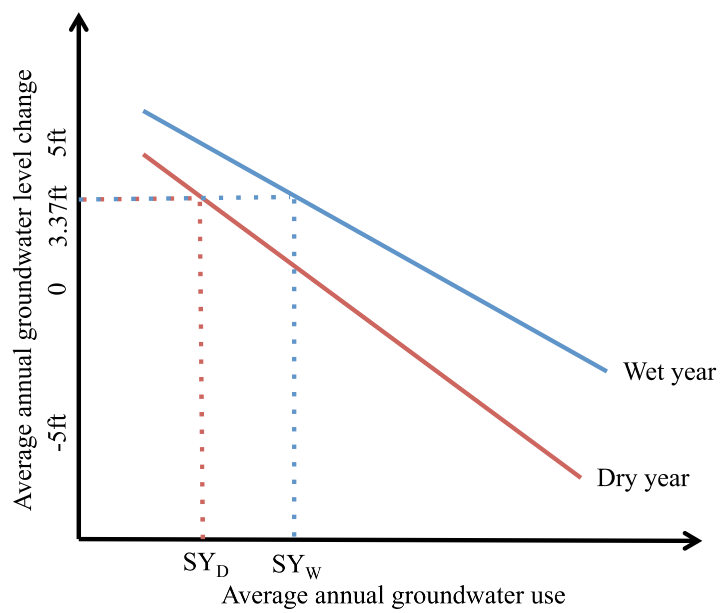

Figure 3.6 Recalculating sustainable yield under a streamflow constraint

The relationships in Figure 3.5 and Table 3.2 assume a causal and linear relationship between each variable and groundwater levels, but in reality this may not be the case for all variables. Depending on the land use or geology in a given GSA, not every undesirable result in SGMA will be tied to groundwater extraction. In Figure 3.5c, for example, a linear relationship is evident, but research shows increasing nitrate concentrations are likely driven by agricultural practices, not necessarily groundwater [*Burow et al.*, 2013]. In addition, the lack of data for nitrate concentrations in this region leaves any empirical relationship unconvincing. Here, sustainable yield should not be constrained because of increasing nitrate concentrations. Instead, GSAs will need to approach groundwater management in a holistic way and incorporate other dimensions of water management to curb undesirable results like nitrate contamination.

3.4.3 Step 3: Projecting basin response over the GSP management horizon

Under SGMA, GSPs must show how the basin will reach sustainability over the planning horizon using sustainable yield-based strategies. Managers will need to project a basin's response to either 2040 or 2042, depending on the basin classification, and assess whether this response is sustainable. There are a number of methods currently available to managers to achieve this, including applying future climate data to a groundwater balance approach or to a numerical groundwater model [*Scibek and Allen*, 2006; *Maxwell and Kollet*, 2008]. Our framework follows an empirical groundwater balance approach that is similar in concept to *Whittemore et al.* [2016] and *Butler et al.* [2016]. They demonstrate the utility of using simple, linear regression equations to relate pumping to groundwater level change [*Whittemore et al.*, 2016]. We modify their approach slightly to include recharge as a separate term, rather than as a part of the *Net inflow* term in Equation 3.2. The *Net Inflow* term is intended to capture all fluxes in and out of the groundwater system, including all lateral and vertical inflows and outflows. In this approach, this term is derived empirically and is part of the *b* variable in Equation 3.4.

We perform a water balance based on the following system of equations, modified from *Butler et al.* [2016].

$$\Delta Storage = \Delta GWL \times Area \times S_s = (Net\ inflow + Recharge) - Pumping \tag{3.2}$$

$$\Delta GWL = \frac{Net\ inflow}{Area \times S_s} + \frac{Recharge - Pumping}{Area \times S_s} \tag{3.3}$$

$$\Delta GWL = b + a*(Recharge - Pumping) \tag{3.4}$$

Where, for a specific time period and groundwater basin, $\Delta Storage$ is the change in storage (acre feet); ΔGWL is the change in groundwater level (feet); *Area* is the spatial extent of the basin (feet squared); S_s is the specific yield (for an unconfined aquifer) or storativity (confined); *Net inflow* is the net inflow to the system, accounting for lateral inflow and outflow (acre-feet); *Recharge* is recharge from interconnected surface water and percolation from the land surface (acre-feet); and *Pumping* is the groundwater extraction volume (acre-feet). Coefficients *b* and *a* are constants that are fit using linear least squares regression [*Whittemore et al.*, 2016]. Simplification from Equation 3.2 to Equation 3.4 is possible under the assumption that S_s and *Net inflow*, largely made up of lateral flows across the aquifer, do not vary in time [*Butler et al.* 2016].

To fit the parameters *a* and *b* in Equation 3.4 for the SSJID, we use data on annual groundwater level change (as described in Step 1) and on recharge and pumping values available in the SSJID Agricultural Management Plan for 1994-2014 [*Davids Engineering,* 2015]. The fit empirical model is shown in Equation 3.5, and has an $R^2 = 0.49$.

$$\Delta GWL = -6.23 + 0.00006*(Recharge - Pumping) \tag{3.5}$$

Equation 3.5 is used to project the impact of sustainable yield management strategies on groundwater levels in the future. To do so, we obtain data on precipitation and recharge from a downscaled climate change model – the California Basin Characterization Model – and apply a mask to the boundaries of the SSJID [*Flint et al.*, 2013]. The California Basin Characterization

Model includes 18 future precipitation and recharge projections from five CMIP-3 and nine CMIP-5 General Circulation Models under three emission scenarios and four different representative concentration pathways (RCPs) [*Flint et al.*, 2013]. With the precipitation data, we classify a given future year as wet or dry based on annual precipitation, where a wet year receives greater than 13 inches of rainfall [*Davids Engineering,* 2015]. The *Recharge* term is broken down into three components – 1) percolation from precipitation, estimated from the downscaled GCM data; 2) return flows from applied water, based on historical averages for wet years and dry years; 3) seepage from surface water, also based on historical averages for wet years and dry years [*Davids Engineering,* 2015].

The *Pumping* term in Equation 3.5, the annual volume of water extracted from a groundwater basin, will vary depending on the hydrological year type and the corresponding sustainable yield value. In wet years, *Pumping* is defined as the basin's wet year sustainable yield and in dry years, the dry year sustainable yield. At this point in the framework, basin managers can elect to vary the *Pumping* term and assess the impacts of various definitions of sustainable yield on future groundwater levels. This may be a common approach for confined aquifers, where only managed depletion is possible. In unconfined systems, a GSA may elect to allow depletion due to socio-economic reasons. In such situations, a management goal may not be zero groundwater level change in a given year. These situations can also be easily incorporated into our framework.

We illustrate this point by applying three different sustainable yield-based scenarios to the *Pumping* term in Equation 3.5. In Scenario 1, each wet year and dry year sustainable yield are calculated according to the SSJID case study shown in Step 1 and Step 2. In this scenario, the baseline sustainable yield is determined under an annual groundwater level change of 0 feet and is not constrained by any undesirable results, as was shown in Step 2 for the SSJID. Figure 3.3 illustrates the sustainable yield calculation for Scenario 1. In Scenario 2, we present a case in which there is some gradation in how a GSA defines their threshold for groundwater level change. In this scenario, a GSA opts to allow depletion in dry years and recovers the basin groundwater level in wet years. Figure 3.7 further illustrates this idea – the appropriate pumping level for a two-foot drop in groundwater levels can be calculated during dry years from our modified Hill's Method. We can then calculate the corresponding pumping level to allow for a two-foot increase in groundwater levels during a wet year. These wet year and dry year sustainable yield values may be more socio-economically appropriate for the region. Finally, we compare these two scenarios to a third, in which the basin elects to make no changes to groundwater management and continues "business-as-usual" pumping during wet and dry years. The three scenarios wet year and dry year sustainable yield values for our case study of the SSJID are shown in Table 3.3.

Table 3.3 Sustainable yield values for management scenarios in SSJID

	Scenario 1 Baseline	Scenario 2 Recovery-Depletion	Scenario 3 Business-as-usual
Wet year	73,608 acre-feet	33,608 acre-feet	48,606 acre-feet
Dry year	20,686 acre-feet	60,686 acre-feet	58,712 acre-feet

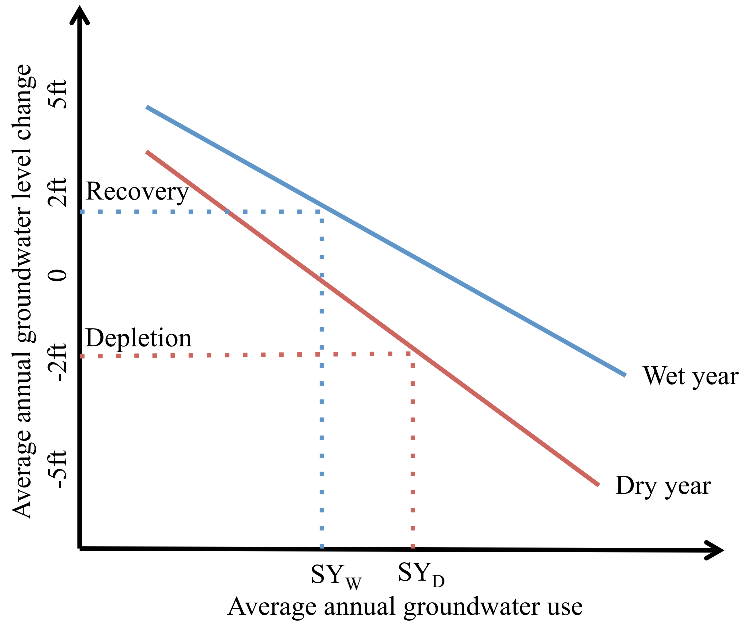

Figure 3.7 Sustainable yield calculation for Scenario 2 – managed depletion, recovery

We then run our empirical model (Equation 3.5) with these three sustainable yield pumping scenarios under the 18 different climate model datasets. The results are shown in Figures 3.8a-c. In all three figures, the groundwater levels calculated from the 18 climate simulations are designated by the grey lines, and the mean of the simulations by the bolded black line. In Figure 3.8a, the baseline sustainable yield scenario, groundwater levels are shown to stabilize over time. A similar trend is also evident for the mean of the climate change projections in Scenario 2, Figure 3.8b. Both Scenario 1 and 2 avoid the dramatic depletion of the business-as-usual Scenario 3. However, there is more uncertainty in the depletion-recovery Scenario 2; the spread of the possible groundwater levels due to future climate is much wider than those in Scenario 1. A management regime like Scenario 2, which may be more practical for many GSAs, must be undertaken with more care. The occurrence of possible undesirable results, such as streamflow depletion, is more likely under this scenario. For the SSJID, the groundwater level threshold calculated in Step 2 at which surface water baseflow depletion may occur is 46.63 feet.

Future year sustainable yield values in Scenario 2 may be constrained by the streamflow undesirable result if recovery periods do not make up for managed depletion years.

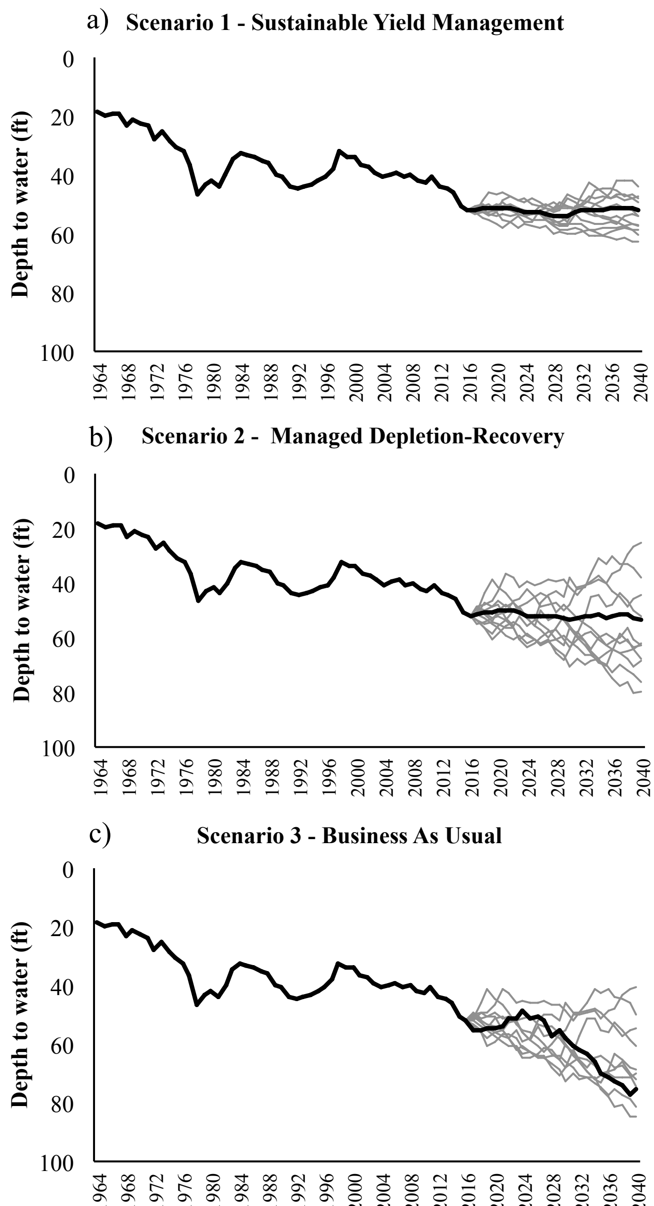

Figure 3.8 Projected average annual groundwater levels for SSJID from 2017 to 2040, under 18 climate change models (grey lines) and three management scenarios. Mean of all climate projections shown in black. Historical data shown from 1964 to 2016 [*CA DWR*, 2017].

3.5 Conclusions

The implementation of SGMA gives California a new opportunity to incorporate science-based approaches into continually evolving groundwater management schemes. To do so, water managers must bridge the gap between hydrologic science and policy and employ the necessary tools to craft and administer institutionally appropriate and physically relevant groundwater management plans. This study examines the concept of sustainable yield, in the context of SGMA. We do not offer a new definition for groundwater sustainability but rather provide a framework that GSAs can use to assess the long-term impacts of sustainable yield-based management scenarios that take into account both future groundwater availability and potential undesirable effects.

We present a three-step framework to quantify sustainable yield and illustrate this approach through a case study of the SSJID. In Step 1, we introduce a modified Hill's method that is more appropriate for California's highly variable hydrology. Basin managers can use this method to derive sustainable yield values – where sustainability can take the classic definition of zero annual average groundwater level change or represent any other groundwater level change (depletion) that basin managers consider tolerable. In Step 2, we present an approach to integrating the groundwater quality, land subsidence and surface water undesirable results into the quantification of sustainable yield. This method is flexible to varying types input data and can incorporate more complex relationships between the undesirable results and groundwater levels, such as models, as they are developed. In Step 3, the overall sustainability of three different definitions for sustainable yield is assessed through an empirical groundwater balance that is projected to 2040. The results of the three scenarios show that there are tradeoffs between groundwater availability, future uncertainty and socio-economic preferences that must be carefully weighed. This step will allow basin managers to do that.

There are a number of limitations to both the present study as well as local groundwater management that will be addressed as models are developed for the region and more monitoring data becomes available. This study employs available datasets of groundwater use, groundwater level change, local streamflow, etc., but there is a significant lack of information on other key variables, particularly for water quality and land subsidence. The development of more spatially distributed and temporally consistent monitoring data should be a priority for GSAs. Information is a prerequisite to effective management. The incorporation of results from new groundwater models into the proposed framework will also allow groundwater managers to better account for the effects of pumping on interconnected surface water bodies. Numerical aquifer and streamflow models that are run over the long term are one of the best management practices necessary to achieve a comprehensive understanding groundwater behavior, and should therefore be a requirement for determining streamflow thresholds under SGMA [*Alley and Leake*, 2004].

Finally, the concept of sustainable yield cannot be static, and water managers will need to adapt and change – recalculating annual sustainable yield values – as environmental and socio-economic conditions shift in time [*Meyer,* 1993; *Sophocleous,* 2000; *Alley and Leake*, 2004]. While SGMA mandate five year reporting periods, sustainable yield values should be determined annually, incorporating all new monitoring data as it becomes available.

Practical, day-to-day groundwater management, while based on hydrologic and scientific principles, cannot ignore the realities of each local economy and its inhabitants. In some parts of the Central Valley, groundwater sustainability will likely mean full depletion over time. We hope that this framework will allow GSAs to define sustainability as they see fit while at the same time help them identify and quantify any potential physical consequences that come with it.

3.6 References

Alley, W.M., Reilly, T.E., and Franke, O.L. (1999), Sustainability of ground-water resources: U.S. Geological Survey Circular 1186, 79 p. Available at: https://pubs.usgs.gov/circ/circ1186/pdf/circ1186.pdf

Alley, W.M. and S.A. Leake (2004), The Journey from Safe Yield to Sustainability, Ground Water, 42(1), DOI: 10.1111/j.1745-6584.2004.tb02446.x

Balleau, W.P., Mayer, A.B. (1988), The transition from ground- water mining to induced recharge in generalized hydrogeologic systems, Proceedings of FOCUS conference on Southwestern Ground Water Issues, National Water Well Association, Dublin, Ohio, pp. 81–103.

Black Water Consulting Engineers (2016), City of Escalon Water System Evaluation Report, City of Escalon, http://cityofescalon.org/e-documents/public_works/water/103_water_eval.pdf

Borchers, J. and M. Carpenter (2014), Land Subsidence from Groundwater Use in California, Luhdorff & Scalmanini Consulting Engineers, California Water Foundation. http://californiawaterfoundation.org/wp-content/uploads/PDF/1397858208-SUBSIDENCEFULLREPORT_FINAL.pdf

Bredehoeft, J.D., Papadopoulos, S. and H. Cooper (1982), Groundwater: The Water-Budget Myth, Scientific Basis for Water-Resource Management, Studies in Geophysics, Washington, DC. National Academy Press.

Bredehoeft, J.D. (2002), The Water Budget Myth Revisited: Why Hydrogeologists Model, Ground Water, 40(4). DOI: 10.1111/j.1745-6584.2002.tb02511.x

Burow, K., Jurgens, B., Belitz, K. and N. Dubrovsky (2013), Assessment of regional change in nitrate concentrations in groundwater in the Central Valley, California, USA, 1950s–2000s, Environmental Earth Sciences, 69: 2609. doi:10.1007/s12665-012-2082-4

Butler, J., Whittemore, D., Wilson, B. and G. Bohling (2016), A new approach for assessing the future of aquifers supporting irrigated agriculture, Geophysical Research Letters, 43, 2004–2010, doi:10.1002/ 2016GL067879.

California Department of Water Resources (2006), San Joaquin Valley Groundwater Basin Eastern San Joaquin Subbasin, California's Groundwater Bulletin 118. http://www.water.ca.gov/groundwater/bulletin118/basindescriptions/5-22.01.pdf

California Department of Water Resources (2013), California River Indices, CA Water Plan Update 2013, http://www.water.ca.gov/waterplan/docs/cwpu2013/Final/vol4/hydrology/01RiverIndex_Update2013_Final.pdf

California Department of Water Resources (2016), Sustainable Groundwater Management Program – DRAFT Groundwater Sustainability Plan (GSP) Emergency Regulations Guide, State of California Department of Water Resources. http://water.ca.gov/groundwater/sgm/pdfs/GSP_Regulations_Guide.pdf

California Department of Water Resources (2016), Critically Overdrafted Basins, State of California Department of Water Resources. http://www.water.ca.gov/groundwater/sgm/pdfs/COD_BasinsTable.pdf

California Department of Water Resources (2016), Best Management Practices for the Sustainable Management of Groundwater – Modeling BMP, State of California Department of Water Resources. http://www.water.ca.gov/groundwater/sgm/pdfs/BMP_Modeling_Final_2016-12-23.pdf

California Department of Water Resources (2017), SGMA Portal – GSA Formation Notification System, State of California Department of Water Resources. http://sgma.water.ca.gov/portal/#gsa

California Department of Water Resources (2017), SGMA Toolbox, State of California Department of Water Resources. http://www.water.ca.gov/cagroundwater/sgmatoolbox.cfm.

California Department of Water Resources (2017), Water Data Library, http://www.water.ca.gov/waterdatalibrary/index.cfm

Christian-Smith, J. and K. Abhold (2015), Measuring What Matters – Setting Measurable Objectives for Achieving Groundwater Management in California, Union of Concerned Scientists. www.ucsusa.org/measuringwhatmatters

Conklin, H. (1946), Utilization of ground-water storage in stream system development, Transactions of the American Society of Civil Engineer 111.

Crowley, J., Jorgensen, N., Kahler, L., Satter, L., Tyler, W. and M.F. Finner (1974), Effect of Nitrate in Drinking Water on Reproductive and Productive Efficiency of Dairy Cattle, Water Resources Center, University of Wisconsin Technical Report WIS WRC 74-06.

Croyle, W. (2014), Public Update for the Drought Response - Groundwater Basins with Potential Water Shortages and Gaps in Groundwater Monitoring, California Department of Water Resources, Drought Management Team, Drought Management.

Custodio, E. (2002), Aquifer overexploitation: what does it mean?, Hydrogeology Journal, 10, doi: 10.1007/s10040-002-0288-6

Davids Engineering (2015), 2015 Agricultural Water Management Plan, South San Joaquin Irrigation District, Available at: http://www.ssjid.com/assets/pdf/2015-Ag-Water-Management-Plan.pdf

Devlin, J.F. and M. Sophocleous (2006), Hydrogeology Journal, The persistence of the water budget myth and its relationship to sustainability, 14: 267. doi:10.1007/s10040-005-0485-y

Dzurella, K., Pettygrove, G., Fryjoff-Hung, A., Hollander, A. and T. Harter (2015), Potential to assess nitrate leaching vulnerability of irrigated cropland, Journal of Soil and Water Conservation, 70 (1), doi:10.2489/jswc.70.1.63

Farr, T.G., C. Jones, Z. Liu (2017), Progress Report: Subsidence in California, March 2015 – September 2016, Submitted to California Department of Water Resources. Available at: http://www.water.ca.gov/waterconditions/docs/2017/JPL%20subsidence%20report%20final%20for%20public%20dec%202016.pdf

Faunt, C.C. ed. (2009), Groundwater Availability of the Central Valley Aquifer: U.S. Geological Survey Professional Paper 1766, 225 p. Available at https://pubs.usgs.gov/pp/1766/

Faunt, C.C., Sneed, M., Traum, J. and J. Brandt (2016), Water availability and land subsidence in the Central Valley, California, USA, Hydrogeology Journal 24: 675. doi:10.1007/s10040-015-1339-x

Flint, L. E., A. L. Flint, J. H. Thorne, and R. Boynton (2013), Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. Ecological Processes 2:25 http://dx.doi.org/10.1186/2192-1709-2-25

Foster, S. and D. Loucks (2006), Non-renewable groundwater resources – A guidebook on socially-sustainable management for water-policy makers, United Nations Educational, Scientific and Cultural Organization. http://www.ircwash.org/sites/default/files/Foster-2006-Nonrenewable.pdf

Freeze, R.A., and Cherry, J.A. (1979), Groundwater, Englewood Cliffs, NJ, Prentice-Hall, 604 p.

Grantham, T. and J. Viers (2014), 100 years of California's water rights system: patterns, trends and uncertainty" Environmental Research Letters, doi.org/10.1088/1748-9326/9/8/084012

Herczeg, A., Leaney, F., Onkaparinga Catchment Water Management Board (S.A.). (2002), Groundwater flow systems and recharge in the McLaren Vale prescribed wells area. Canberra: CSIRO Land and Water, http://nla.gov.au/nla.arc-13482

Hill, M. C. (2006), The Practical Use of Simplicity in Developing Ground Water Models. Ground Water, 44: 775–781. doi:10.1111/j.1745-6584.2006.00227.x

Joseph, T. (2016), Best Management Practices for the Sustainable Management of Groundwater: Monitoring Networks and the Identification of Data Gaps, California Department of Water Resources, Sustainable Groundwater Management Program. http://www.water.ca.gov/groundwater/sgm/pdfs/BMP_Monitoring_Networks_Final_2016-12-23.pdf

Kalf, F. and D. Woolley (2005), Applicability and methodology of determining sustainable yield in groundwater systems, Hydrogeology Journal, 13. doi: 10.1007/s10040-004-0401-x

Konikow, L. F., and S.A. Leake (2014), Depletion and capture: revisiting "The source of water derived from wells". Groundwater 52.S1, doi: 10.1111/gwat.12204

Langidge, R., Brown, A., Rudestam, K., and E. Conrad (2016), An Evaluation of California's Adjudicated Groundwater Basins, University of California, Santa Cruz. http://www.water.ca.gov/cagroundwater/docs/UCSC_Adjudicated_Groundwater_Basins_Report_FINAL.pdf

Loáiciga, H.A. (2016), The Safe Yield and Climate Variability: Implications for Groundwater Management, Groundwater. doi: 10.111/gwat.12481

Lohman, S. (1972), Ground-water hydraulics, Geological Survey Professional Paper 708, US Government Printing Office. https://pubs.usgs.gov/pp/0708/report.pdf

MacEwan, D., Cayar, M., Taghavi, A., Mitchell, D., Hatchett, S. and R. Howitt (2017), Hydroeconomic modeling of sustainable groundwater management, Water Resour. Res., 53, doi:10.1002/2016WR019639

Maimone, M. (2004), Defining and Managing Sustainable Yield. Ground Water, 42: 809–814. doi:10.1111/j.1745-6584.2004.tb02739.x

Maxwell, R.M., and S.J. Kollet (2008), Interdependence of groundwater dynamics and land-energy feedbacks under climate change. Nature Geoscience, 1(10), 665-669. doi:10.1038/ngeo315

Meyer, J. (1993), Changing concepts of system management. Sustaining Our Water Resources, 78, The National Academies Press, https://doi.org/10.17226/2217

Nakagawa, B. (2004), Eastern San Joaquin Basin Groundwater Management Plan, Northeastern San Joaquin County Groundwater Banking Authority, San Joaquin County Department of Public Works. Available at: http://www.gbawater.org/Portals/0/assets/docs/IRWMP-2014/Groundwater-Management-Plan-Final.pdf

Pacific Institute (2014), Issue Brief: Urban Water Conservation and Efficiency Potential in California, Available at: http://pacinst.org/app/uploads/2014/06/ca-water-urban.pdf

Richey, A., Thomas, B. Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., Swenson, S. and M. Rodell (2015), Quantifying renewable groundwater stress with GRACE, Water Resources Research, 51, 5217–5238, doi:10.1002/2015WR017349

Schoups, G., Hopmans, J., Young, C., Vrugt, J., Wallender, W., Tanji, K., and S. Panday (2005), Sustainability of irrigated agriculture in the San Joaquin Valley, California, Proceedings of the National Academy of Sciences, 102(43), doi: 10.1073/pnas.0507723102

Scibek, J., and D. M. Allen (2006), Modeled impacts of predicted climate change on recharge and groundwater levels, Water Resources Research, 42, doi:10.1029/2005WR004742.

Seward, P., Xu, Y., and L. Brendonck, (2006), Sustainable groundwater use, the capture principle, and adaptive management. Water SA, 32(4). http://dx.doi.org/10.4314/wsa.v32i4.5287

Sophocleous, M. (1997), Managing water resources systems: why “safe yield” is not sustainable, Ground water, 35(4), 561-561, doi: 10.1007/s10040-001-0170-8

Sophocleous, M. (2000), From safe yield to sustainable development of water resources – the Kansas experience. Journal of Hydrology, 235, doi: 10.1016/S0022-1694(00)00263-8

Sophocleous, M. (2002), Interactions between groundwater and surface water: the state of the science, Hydrogeology Journal, 10: 52. doi:10.1007/s10040-001-0170-8

State of California (2017), Legislation - More about SGMA, State of California, California Groundwater. http://groundwater.ca.gov/moresgma.cfm

State Water Resources Control Board (2016), Groundwater Information Sheet – Salinity, State Water Resources Control Board, Division of Water Quality, GAMA Program. http://www.waterboards.ca.gov/gama/docs/coc_salinity.pdf

State Water Resources Control Board (2016), Triggering State Intervention – Sustainable Groundwater Management Act (SGMA), California Environmental Protection Agency. Available at: http://www.waterboards.ca.gov/water_issues/programs/gmp/sgma.shtml#materials

Sustainable Groundwater Management Act (2014), Sustainable Groundwater Management Act [And Related Statutory Provisions from SB1168 (Pavley), AB1739 (Dickinson), and SB1319 (Pavley) as Chaptered], Available at: https://www.opr.ca.gov/docs/2014_Sustainable_Groundwater_Management_Legislation_092914.pdf

Theis, C.V. (1940), The source of water derived from wells, Civil Engineering, v. 10, no. 5, p. 277-280. Available at: https://water.usgs.gov/ogw/pubs/Theis-1940.pdf

Thomas, H. (1951), The conservation of ground water. New York, Toronto, London. McGraw Hill Book Co.

Todd, D. (1959), Ground-water hydrology, New York, John Wiley and Sons.

UNAVCO (2017), P781 – Overview, UNAVCO Plate Boundary Observatory (PBO), https://www.unavco.org/instrumentation/networks/status/pbo/overview/P781

U.S. Department of Interior (2005), Water Management of the Regional Aquifer in the Sierra Vista Subwatershed, Arizona—2004, Report to Congress, Available at: https://water.usgs.gov/Section321.2004_050705.pdf

U.S. Environmental Protection Agency (2016), Estimated Nitrate Concentrations in Groundwater Used for Drinking, https://www.epa.gov/nutrient-policy-data/estimated-nitrate-concentrations-groundwater-used-drinking

U.S. Geological Survey (2017), USGS 11303000 Stanislaus R at Ripon, CA, National Water Information System. https://waterdata.usgs.gov/nwis/uv?PARAmeter_cd=00095,00010,00300,00400&site_no=11303000

Whittemore, D. O., Butler Jr, J. J., & Wilson, B. B. (2016), Assessing the major drivers of water-level declines: New insights into the future of heavily stressed aquifers, Hydrological Sciences Journal, 61(1), 134-145. http://dx.doi.org/10.1080/02626667.2014.959958

Van der Gun, J. and A. Lipponen (2010), Reconciling Groundwater Storage Depletion Due to Pumping with Sustainability, Sustainability, 2(11). doi:10.3390/su2113418

Zhou, Y. (2009), A critical review of groundwater budget myth, safe yield and sustainability, Journal of Hydrology, 370. doi : 10.1016/j.jhydrol.2009.03.009

4. Diagnosing the health of the commercial remote sensing market - How developing value-added solutions can foster industry growth

4.1 Introduction

Remote sensing refers to the collection of any information about the observable Earth from a distance, such as from a satellite. The uses of remote sensing data are many and varied, with notable applications in the government sector, including land use mapping, ocean temperature measurement and monitoring, and tracking floods and natural disasters; and in the private sector, including crop yield optimization, insurance assessments, and spatial measurement tools and investment data [*NOAA,* 2015; *Kearns,* 2015; *Thenkabail et al.*, 2004]. In this paper, we focus on satellite-based remote sensing at a global scale, which includes various forms of imagery, measurements of temperature, wind, chemical and biological properties and altimetry, among others [*Pohl and Van Ganderen*, 1998]. Satellite-based remote sensing data offer a relatively unbiased data source and provides an almost endless supply of spatially referenced information on anything that varies by space and in time.

The commercial, or private, remote sensing industry has grown rapidly over the last decade [*BBC Research,* 2016, *Lal et al.*, 2015]. Venture capital firms have invested an unprecedented amount in new start-up companies – nearly $1.8 billion in the first half of 2015 alone – and government agencies, such as the National Oceanic and Atmospheric Administration (NOAA), are considering shifting to the private sector for their satellite-based remote sensing needs [*Gissot et al.*, 2015; *Shepherd,* 2016]. However, little work has been done to verify whether the speculation within this industry is justifiable, that is, whether the industry has the diverse revenue streams and robustness needed to sustain this growth over time.

While previous studies have focused on remote sensing suppliers, an understanding of the health of the remote sensing market requires equal attention to existing and potential demand [*BBC Research*, 2016]. In particular, a more detailed investigation of the business-to-business (B2B) market is needed. A healthy B2B market signals that the industry is large enough to support growth, and that potential customers can depend on the capabilities they need to be available into the future. Current and potential customers for remote sensors include government entities looking to privatize space-based earth observation, investment firms, insurance companies, natural resource agencies, and commercial firms [*GeoConnections*, 2016; *Hope*, 2016]. A complete understanding of future growth paths for commercial remote sensing requires an analysis of potential new customers and their data needs as well as a means of quantifying the value-added of remote sensing data to meet those needs.

In this chapter we take a holistic look at current and potential markets for the commercial remote sensing industry and identify opportunities where remote sensing data could uniquely benefit firms' operations or increase profits. The specific objectives of this chapter are to: (1) provide an overview of the current B2B market for remote sensing; (2) identify strategies to make commercial remote sensing firms more responsive and robust to the needs of the market, especially new sectors; (3) use a case study to examine potential demand for commercial remote sensing in supply chain and business operations; (4) develop a model to quantify the value-added of remote sensing information.

4.2 Overview of the business to business market for commercial remote sensing

We conducted a literature review to examine trends in supply and demand in the remote sensing market. On the supply side, the commercial remote sensing industry has changed dramatically during the last decade [*BBC Research*, 2016]. From the 1990s to early 2010s, the industry grew by only a few firms and satellites annually [*Navulur et al.*, 2013]. Since 2010, however, the race for dominance in commercial remote sensing has accelerated, and many more companies are now entering the industry, with at least a few dozen satellites going up annually [*Foust,* 2013]. Global revenues from the remote sensing satellite industry have grown by about 16% annually since 2010 and reached nearly $1.8 billion USD in 2015 [*The Tauri Group*, 2016].

The rapid industry shift that has occurred since 2010 has been the result of technological advancements, novel business models, and decreased costs. In the past, conventional remote sensing firms, such as DigitalGlobe, spent millions of dollars designing, building, launching, and supporting a single satellite [*Dalby,* 2015]. Over time, the industry matured by improving spatial and spectral resolution, power and robustness, which, in turn, led to larger and larger satellites [*Dalby*, 2015]. In comparison, recent upstart entrants like Planet Labs and Terra Bella have developed and proliferated small satellites ("smallsats"), many of which are no larger than a loaf of bread [*Lal et al.*, 2015; *Dillow,* 2015]. These newer satellite firms have reversed the traditional model of providing increasingly higher resolution, power, and operational lifetime by instead providing "decent" resolution for short missions at increasingly cheaper costs. While payloads like DigitalGlobe's WorldView-3 are over 2,800kg [*Dalby,* 2015; *Kramer,* 2016], smallsats typically weigh between 5kg to 500kg [*Lal et al.*, 2015; *Dillow*, 2015].

A comparison of the cost of Planet Labs' Flock 1 to that of DigitalGlobe's upcoming WorldView-4 provides another clear example of the shift towards cheaper, nimbler satellites. Planet Labs' Flock 1, which cost an estimated $2.8 million for 28 satellites [*Dalby,* 2015], has a much lower price tag than WorldView-4, with a cost of $750 million. Planet Labs' lower costs have been made possible thanks to the use of commercial off-the-shelf (COTS) sensors rather than conventional, custom-built optical sensors [*Lal et al.*, 2015]. Smallsat owners and operators like Planet Labs have been able to maintain their cost advantage while staying on the forefront of new technology in part due to the rapid technological advancements in COTS sensors themselves [*Lal et al.*, 2015].

While remote sensing suppliers may differ in the number of payloads and the structure of their business model, the key differentiators in their product offerings are the spatial resolution and spectral bands or sensor type. Further, Table 4.1 illustrates that many of the newer market entrants – such as UrtheCaste, Planet Labs and BlackSky – provide information at a medium spatial (1-5 m) and temporal (daily or sub-daily) resolution.

Table 4.1 Types of remote sensing data available

Company	Spectral bands/sensor type	Spatial resolution	Revisit rate	Number of satellites
Terra Bella	Panchromatic, R, G, B, NIR	0.9 – 2 m	3x daily	3, 18*
UrtheCast	Panchromatic, R, G, B, NIR, X-band and L-band SAR*	1 – 4 m	<Daily	24*
Planet Labs	R, G, B, NIR**	2.7 – 4.9 m	<Daily	127, >150*
DigitalGlobe	Panchromatic, multispectral, SWIR	0.31 – 30 m	<Daily	4, 1*
BlackSky*	Multispectral	1 m	<Daily	60*
Aquila Space*	R, G, B, NIR, Red edge	2.5 – 22 m	>Daily	30*
NorStar*	Hyperspectral, IR	n/a	n/a	40*

**Only planned satellites; **Experimental use only*
Sources: *Krebs,* 2016; *eoPortal*, 2016.

In contrast to the dynamism of the supply side, the demand side of the remote sensor industry is much more static and driven by a few, mostly long-time, customers. The primary sectors buying and using remote sensing data are defense (59% of global demand), natural resources (10%), energy (9%) and infrastructure (8%) [*Keith*, 2015]. The government—at multiple levels (state, federal, local)—remains the dominant user [*de Selding,* 2016]. A few private industry verticals - insurance, agriculture, and mining - are also notable customer segments [*de Leeuw et al.*, 2014; *Hope*, 2016; *Darrow*, 2015; *Gissot,* 2015].

The rapid growth of commercial remote sensing capacity has created significant opportunities for firms to develop novel remote sensing data services and market themselves to a wider variety of potential clients. Some companies, such as Spaceknow, Orbital Insight, RS Metrics, Descartes Labs, and Ursa Space Systems, are responding to a growing demand for geospatial data-based analytics, numerical methods that refine geospatial data to provide more targeted information [*Lal et al.*, 2015]. Most offer a suite of data products that are tailor-made according to their client base and internal expertise. Examples of such data products include parking lot counts, oil storage tank measurements, commodity crop forecasts, and estimates of surface water availability [*RS Metrics,* 2016; *Ursa Space Systems*, 2016; *Spaceknow,* 2016; *Orbital Insight*, 2016]. However, even though significant opportunities have emerged for firms to create these novel remote sensing data services, commercial firms are still not fully capitalizing on available prospects. The majority of remote sensing data continues to be sold to traditional users – state and local governments and the agriculture and insurance industries [*Private Conversation*, 2016]. Investment firms, which are traditionally adept at utilizing diverse data sets, are a notable emerging consumer [*Hope*, 2016].

Our research suggests that a key reason for the lag in the growth and diversification of the consumer base for remote sensing data is the disconnect between the business intelligence needs of potential clients and the services that remote sensing companies are providing. We illustrate this point through Figure 4.1, which shows a U.S. Department of Defense (DoD) view of the way in which data are collected through remote sensing and distilled into intelligence that fulfills a mission or business need. Collection, processing, and analysis are identified as the three processes by which data refinement occurs, but there are barriers than can impede a firm's movement from one phase to the next.

Commercial remote sensing providers – large and small, old and new – have solved or are solving the data collection problem, as shown by the smallsat leaping over the first barrier in Figure 4.1. However, would-be business consumers require more than the data alone: they need data to be translated into useful information and ultimately into products that solve business intelligence problems. Successfully overcoming the barriers to progress and transforming remote sensing data into effective business solutions requires insight into business-specific operating challenges. It is likely that the development of tailored solutions for one client cannot typically be leveraged for other customers.

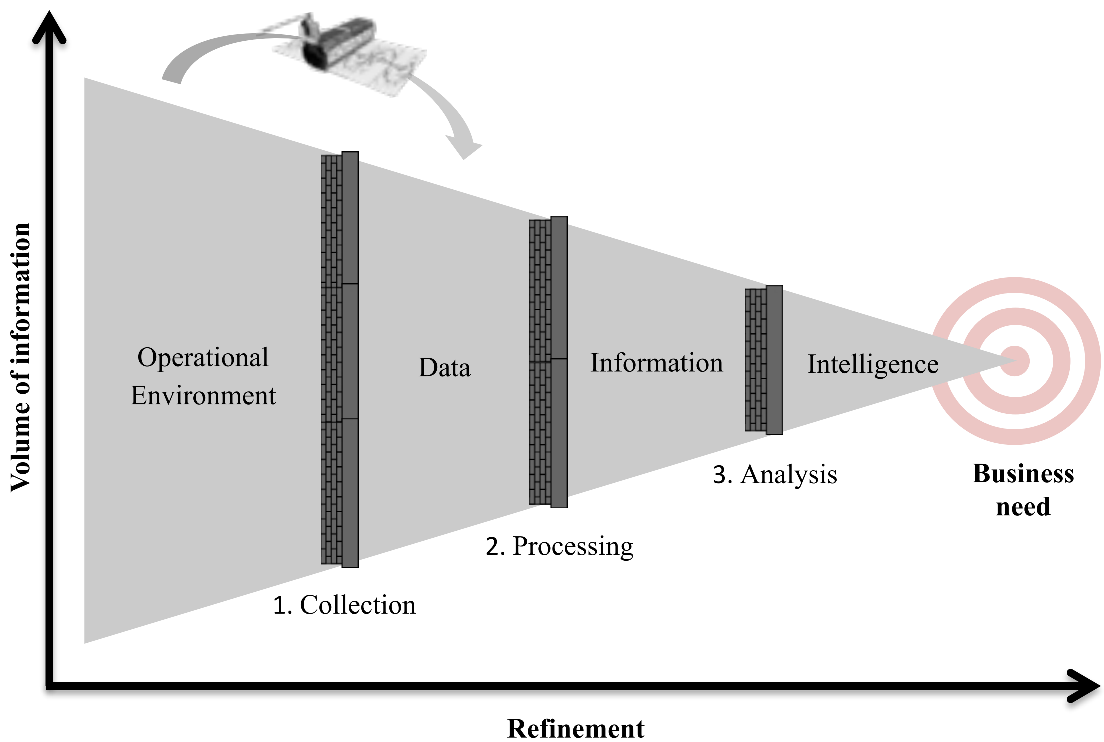

Figure 4.1 Relationship between data, information and intelligence, adapted from *Joint Chiefs of Staff*, 2013

Although a growing number of companies are likely interested in incorporating remote sensing data into their decision processes, many do not have the capacity or organizational structure to process data into actionable intelligence on their own, and they have been largely unwilling to invest internally in the expertise required [*Private Conversation*, 2016]. Thus, a gap exists between the continued emphasis among commercial remote sensing suppliers on providing *data* as their key product, and business users' need to have that data transformed into actionable *intelligence*. We contend that this gap has contributed to the inability of the commercial remote sensing industry to cultivate new users on a large scale—a situation that has led some studies to posit whether the recent rapid growth in commercial space is sustainable. If the key product offered by remote sensing providers is data alone, how can a private company with an inexpensive, lightweight imager compete with free U.S. government-provided Landsat data of comparable quality? [*Gissot et al*., 2015; *Private Conversation*, 2016].

4.3 Strategies to promote growth in the commercial remote sensing industry

As with any emerging market, the way in which customers and companies will adapt to changing market conditions is difficult to predict [*Burgel and Murray*, 2000; *Private Conversation*, 2016]. However, our research has identified three strategies to assist commercial remote sensing companies in monetizing the value of their data and services in a way that will sustain and grow the industry. First, gaining a better understanding of would-be customers' data needs, especially in the area of business intelligence, is critical to identifying opportunities for generating value-added products to meet those customers' needs. In particular, remote sensing firms might focus on identifying key entry points within potential clients' business processes where commercial remote sensing offerings could provide a decision advantage for them. Second, our research suggests that products based upon medium spatial and temporal resolution offer opportunities for providing value-added products. Higher spatial resolution leads to larger (and more expensive) spacecraft, and higher temporal resolution is generally unavailable in the free data sets. Third, we believe that data refinement services (often loosely termed "analytics") offer another key avenue for growth, by providing tailored business intelligence to would-be customers. We discuss each of these ideas in more detail below.

Understanding the customer is crucial for generating a value-added product that can be used to break into new sectors. A survey of over 3,000 executives and decision-makers at 30 prominent international companies showed that businesses are actively seeking ways to integrate data-driven decision-making into their current operational frameworks [*LaValle et al.*, 2011]. However, many of these professionals do not have experience with analytics and cite a lack of understanding of data-driven decision-making as a main impediment to adoption [*LaValle et al.*, 2011]. One critical step remote sensing firms might take towards moving companies forward in this task is to reduce the time to value [*LaValle et al.*, 2011], that is, help customers understand upfront the specific ways in which remote data can be used to inform their decision processes. To do so, remote sensing companies might "start with the question, not the data," and outline the insights that need to be generated first [*LaValle et al.*, 2011] in order for a potential customer to achieve value. While there is customer pull for data to drive decision-making, the burden of proof is on commercial remote sensing providers to demonstrate value for particular applications. Demonstrating tangible business value for potential new customers will require some level of investment in business know-how, creativity, and product development; however, this knowledge will provide the basis for developing new products and services likely to expand the current customer base.

Potential customers interested in remote sensing data have a wide array of data sources to choose from – internet of things data, consumer sales information, traffic patterns, weather predictions, etc. Remote sensing data are unique in that they provide both complete spatial coverage (as opposed to discrete data points) as well as temporal patterns (as opposed to static-like map data). We believe that a focus on curated medium spatial and temporal resolution data (data that are collected daily with a spatial resolution between 1 and 5 meters) offers commercial remote sensing firms a unique space within the industry and with their potential consumer base. This "space" is shown in grey in Figure 4.2. As seen in the figure, emerging remote sensing companies are competing with two main sources – free U.S. government-provided Landsat data and DigitalGlobe's low-cost imagery. Among the current data offerings of these competitors, there is a gap or niche with respect to the spatiotemporal nature of the data. Remote sensing companies can harness this spatiotemporal uniqueness of their product and provide information solutions that change in both space and time. Building intelligence offerings that capitalize on *spatiotemporal resolution*, instead of just spatial resolution, could enable new commercial remote sensing entrants to carve out a competitive position. In addition to filling the gap in current data offerings shown in grey in Figure 4.2, medium spatiotemporal resolution remote sensing data have few to no replacement products and measure different phenomena than internet of things sources.

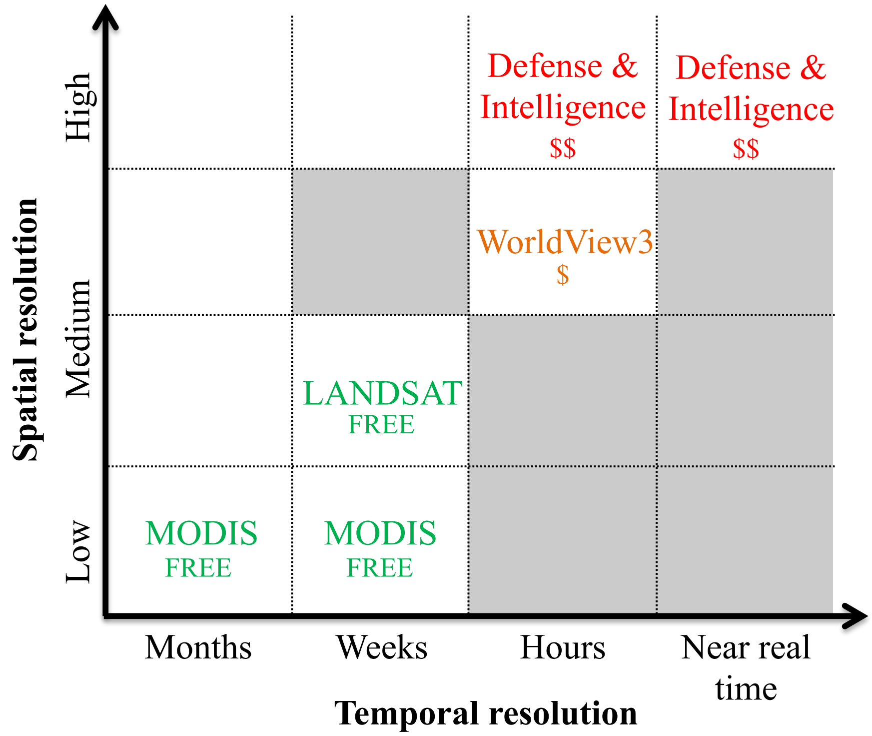

Figure 4.2 The spatial and temporal resolution of existing remote sensing data sources and opportunities for new private sector entrants (shown in grey).

Data analytics built on commercial remote sensing data may directly address client needs and allow companies to be more sustainable to market fluctuations. So called "big data analytics" are already proving to be crucial sources of advantage that help companies better measure and understand their business environment [*McAfee and Brynjolfsson*, 2012]. Recent studies have indicated that data-driven decision-making has made companies more productive, more profitable and increased their output by about 3-6% [*McAfee and Brynjolfsson*, 2012]. A 2007 study that examined cost-revenue tradeoffs at the Indian Space Research Organization (ISRO) found that ISRO's revenue would have increased from $40 million to $223.5 million if it had provided and directly charged for thematic information and mapping rather than just raw data access [*Geospatial World*, 2013; *Sankar*, 2007]. The study also notes that an additional $204 million in cost savings was achieved by data purchasers who opted to use and develop remote sensing-derived maps instead of traditional surveying [*Geospatial World*, 2013; *Sankar,* 2007].

Within the commercial remote sensing industry, delivering analytics-based products could generate additional revenue and facilitate diversification of offerings. Analytics will draw upon expertise interpreting raw remote sensing data, a skill that very few would-be customers possess. Moreover, regardless of whether another remote sensing satellite is ever launched, analytics are already required to process and utilize the large volume of data that currently exists [*Private Conversation,* 2016]. And in the unlikely event of a market failure in the satellite side of the industry, remote sensing companies could still apply analytics to free data sources (i.e., Landsat).

4.4 A case study of Walmart's operations and supply chain management

To understand how these strategies could be implemented and to quantify the potential value-added of remote sensing data in new sectors and assess future demand, we have performed a case study of Wal-Mart Stores, Inc. ("Walmart"). Walmart's geographically extensive and well-oiled supply chain, its pioneering use of information technology, and its financial strength are all qualities that make Walmart an ideal company to study the applications and value-added of remote sensing data.

In terms of its earnings, Walmart grossed over $14 billion USD in profit and generated more than $482 billion USD in the 2015 fiscal year, making it the world's largest company by revenue [*Fortune,* 2016]. Studies on Walmart have shown they have achieved their competitive advantage through two main mechanisms – adopting new technologies and securing low prices from suppliers [*Brea-Solis et al.*, 2012]. In pursuing these strategies, Walmart has come to be considered as an industry leader in supply chain operations and technology adoption [*Traub*, 2012; *Wang et al.*, 2016]. It was one of the first companies to promote electronic data interchange (EDI), to require suppliers to participate in an innovation-driven data sharing platform, and to utilize cutting-edge operational methods to maintain a "superbly efficient" supply chain [*University Alliance*, 2016]. Moreover, Walmart has already enhanced many areas of their operations through the use of novel data analyses, shipping and inventory management techniques, for example, tracking temperature changes, weather events and consumer shopping preferences [*Ozment,* 2014; *Helman,* 2014]. Energy use and energy costs have also been relatively optimized through in-store instrumentation, use of alternative energy sources, efficient truck routing, and experimentation with demand reduction technologies [*Helman*, 2014].

Our study focuses on Walmart's supply chain, its principal supply chain management strategies, and finally where additional efficiencies could be gained through the use of remote sensing data. Table 4.2 details the main components of Walmart's supply chain – production and product delivery, distribution center storage, transit to Walmart stores and in-store stocking and sales – and the strategies Walmart employs to decrease costs and move goods more efficiently. From the upstream end of Walmart's supply chain, products move from manufacturing facilities at various supplier locations around the world and are then transported to Walmart-owned distribution centers [*Wulfraat*, 2016]. The supplier manages this first step of the supply chain. Walmart does not directly control the movement of goods into its distribution centers but manages the location and operation of the distribution centers themselves [*Greenspan*, 2015]. From distribution centers, Walmart trucks take goods to the store where they are sold [*Wulfraat,* 2016]. The flow of information within the supply chain moves in the opposite direction to the flow of goods. Information on each good from Walmart's point-of-sale (POS) systems is shared with suppliers via "Retail Link," an electronic data interchange (EDI) platform that is controlled in Walmart's headquarters in Bentonville, Arkansas [*Wailgum,* 2007]. Retail Link has been innovative since its inception, and has propelled Walmart to its place as a market leader in B2B communications and data processing [*Wailgum,* 2007]. Retail Link implementation also caused secondary benefits, such as encouraging Walmart's suppliers to apply analytics to their product sales to justify product placement. Suppliers have "relatively small sets of products and significant vested interest in seeing those products' performances optimized" [*Waller and Boccasam,* 2013]. The main supply-chain strategies that Walmart employs are shown in Table . Italicized entries will be discussed in Section 4.5 below. While each of these has led to operational improvements, operational challenges still persist. Principal among these are risk across all levels of the supply chain, costs associated with movement and storage of product, and uncertainty in demand predictions.

Table 4.2 Walmart's supply chain key strategies and challenges

Supply chain stages	Key strategies and benefits	Existing challenges
Manufacturing, production and product delivery	*Bulk purchasing power* • Constant competition for Walmart's business and their drive for low prices requires continuous innovation and cost-cutting by the vendor	• Lack of knowledge of supplier's suppliers (*upstream suppliers*) and the risks they face • Natural disasters and other supply chain disruptors (strikes, etc.)
Distribution centers	*Vendor-managed inventory (VMI)* • Vendor is responsible for moving products from its production facilities to Walmart Distribution Centers • Pushes vendors to streamline their inventory management and delivery services	• Transportation costs • Inventory size and turnover • Sales forecasting errors
Transit to stores	*Cross-docking* • Decreases inventory in distribution centers • Quickens time to consumer (and time to sale) of goods	• Transportation costs • Store site selection
In-store stocking	*RetailLink* • Seamless data exchange between vendors and Walmart • Saves time any money by automating the order of new inventory, based on in-store sales trends • Vendors compete for shelf and store placement via data analyses on sales and store data	• Stock shortages • Demand (sales forecasting) uncertainty • Back room stockpiles and slowdowns

Sources: *Wailgum*, 2007; *Soni,* 2015; *Waller and Boccasam*, 2013; *Wulfraat,* 2016.

4.5 Remote sensing applications for Walmart

Of the potential areas for use of remote sensing data within Walmart's supply chain and operational structure, four main options were identified and are shown in Table . The first row of Table explains potential applications of various forms of geospatial information to optimize the supply chain downstream from the distribution centers. Geospatial data from the communities around Walmart stores – such as information on outdoor parties, sporting events, car washing, gardening, public park utilization, when and where people are shopping, etc. – could enhance sales forecasts by developing a better understanding of the spatio-temporal patterns of the individuals and their shopping habits within a community. Walmart could use this information to predict purchase of specific goods or buying trends in time. A second application of remote sensing data that would aid in demand forecasting is relevant to the days following a natural or an environmental event or disaster, such as a large storm or tornado, localized flooding, or sewer backups. Knowledge of the type and quantity of different types of damage – to roofs, driveways, porches, patios, etc. – would aid Walmart in expediting and streamlining the movement of the necessary goods (such as tents, fans, home appliances, etc.) to the correct locations. In general, a more accurate understanding and prediction of demand has the ability to minimize stock shortages in stores and improve e-commerce through targeted online campaigns. Moving goods only when they are truly needed cuts down on transportation costs and last minute order filling from stock shortages. Similar types of location- and temporally-specific information near a potential new store would make the site-selection process more efficient through a more complete picture of local demand. The second row of Table focuses on everything upstream of the distribution centers in Walmart's supply chain. Appropriately processed remote sensing data could help suppliers anticipate potential disruptions, transit blockages, port delays and other relevant geospatial information (such as information on raw materials). A few examples of possible actions related to remote sensing information are: (a) rerouting shipments following information on strikes, the spatial extent and type of road closures/blockages or knowledge of outbound shipment delays due to port traffic; (b) quicker exchanging of upstream suppliers if a hurricane or natural disaster hits or if remote sensing information shows a production slowdown. This could help optimize sub-tier suppliers' delivery of products on time and could cut costs overall. Lower production and transportation costs translate to lower costs for Walmart.

Importantly, Table 4.3 provides examples of remote sensing data that would be relevant for Walmart at multiple levels of its supply chain and also identifies data needs that could uniquely be met by the spatial and temporal resolution of remote sensing data.

Table 4.3 Applications of remote sensing data to Walmart's supply chain

	Potential uses within supply chain	**Examples of relevant remote sensing data**
Downstream	*Sales or demand forecasting* • Location-specific and temporally-relevant information incorporated into existing sales forecasting models • Predictive disaster relief *Store site selection* • Business intelligence on and consumer activity at potential competitors • Secondary source checks on location-specific information	*Consumer activity in region* • Local events (soccer games, swim meets, park utilization, outdoor parties) • Prevalence and timing of consumer shopping at competitors' stores *Post disaster and storm relief information* • Outdoor damage classification and location information (damage to roofs, gardens, patios, driveways) • Flood depths and extents • Location of sewer backups
Upstream	*Supply chain streamlining, risk mitigation* • Embedded as additional data available to vendors in RetailLink EDI platform *Upstream suppliers* • Monitoring of upstream suppliers by vendors and/or Walmart mitigates risk of shocks to supply chain originating before suppliers	*Potential supply chain disruptions* • Monitoring of activity at production facilities of upstream suppliers • Strikes/road blockages (duration and spatial extent) • Port traffic • Natural disaster impacts

Sources: *Smithson*, 2015; *University Alliance*, 2016; *Greenspan*, 2015.

4.6 Valuing geospatial data within Walmart's sales forecasting capacity

To better understand the impact that remote sensing data may have on Walmart's supply chain, this section estimates the value of applying remote sensing data to Walmart's sales forecasting capacity. As discussed above, sales forecasting improvements would have a number of cost-saving and revenue-generating impacts on Walmart's supply chain. Instead of measuring the internal cost-saving dynamics directly, we will value the reduction in uncertainty in sales forecasting that results from the addition of proxy geospatial information.

Classic valuation methods for geospatial data determine socioeconomic impacts or direct economic benefits of data use by either (a) assessing one specific application of geospatial data at one static point in time or by (b) relying on surveys and interviews of stakeholders to quantify value [*NASA Applied Sciences Program*, 2012; *Hertzfeld et al.*, 2003; *Loomis et al.*, 2015; *Forney et al.*, 2012; *Bernknopf and Shapiro*, 2015]. Value of information (VOI) methods, for example, estimate by how much geospatial information changes an individual's beliefs about uncertainties and then quantify how this impacts decision-making [*NASA Applied Sciences Program*, 2012]. VOI approaches are typically applied to a single decision and can be subjective in nature. Other methods, such as revealed preferences valuation, employ more robust numerical approaches (i.e. hedonic pricing) to monetize the effect geospatial data has on user behavior [*NASA Applied Sciences Program,* 2012]. These methodologies typically require assumptions about the direct and indirect impacts of new information on a decision [*NASA Applied Sciences Program,* 2012]. Cost-benefit and cost-effectiveness analyses quantify the tradeoffs between the costs associated with use of geospatial data and their known or projected impacts [*NASA Applied Sciences Program,* 2012]. Of these, cost-effectiveness analysis is helpful in cases when the benefits of a particular dataset or geospatial-based intervention cannot be fully monetized [*NASA Applied Sciences Program,* 2012]. Multivariate regression has also been employed to value geospatial data [*NASA Applied Sciences Program,* 2012]. Overall, most studies seek to estimate the benefits or impacts with and without geospatial information and then compare the two in order to quantify the value-added of the use of the information [*NASA Applied Sciences Program*, 2012; *Bernknopf and Shapiro*, 2015].

This study builds on these methods by employing an alternate approach based on a statistical analysis of four years of Walmart store-level data. We apply multivariate regression, a common approach to sales forecasting, and examine its predictive capacity with and without the addition of proxy geospatial data [*Stergiou et al.,* 1997]. Our empirical approach quantifies the impact of proxy geospatial data on uncertainty in sales forecasts using real store-level data. Generally, the idea is to compare two models – one that contains proxy geospatial information and a second baseline model without this data [*NASA Applied Sciences Program*, 2012]. In this study, we follow this approach and use four years of weekly store-level sales data from 45 stores randomly distributed across the United States that Walmart made publicly available on Kaggle.com [*Kaggle*, 2014]. This data does not contain geolocation information but includes predictor variables of interest for each store for each week of sales. These are shown in Table 4.4.

Table 4.4 Time series of independent variables, weekly (2/2010-7/2013)

Independent variables	Symbol
In-store indicators	
Holidays	HOL
Type of store (superstore, etc.)	STOR
Size of store (ft^2)	SQFT
Mark downs offered ($)	MKDN
Season (winter, summer, etc.)	SEAS
Economic variables	
Unemployment rate (%)	UNP
Fuel price ($/gallon)	FPR
Consumer Price Index, CPI	CPI
Environmental variables	
Outside temperature (°F)	TEMP

A linear multivariate regression model of these variables takes the following form:

$$Y_{ti} = a + b_1 x_{ti} + b_2 x_{2ti} + b_3 x_{3ti} \dots + b_8 x_{8ti} + e_{ti} \qquad (4.1)$$

$$for\ i = 1{:}\,45\ stores\ and\ t = 1{:}\,n\ weeks$$

Where Y, is the value of the dependent variable – store sales – at time t and for store i, x_1 to x_8 are the independent variables and e is an error term that is assumed to be sampled independently from a normal distribution [*Stergiou et al.*, 1997].

Our modeling approach is based on the idea that incorporating key environmental and geospatial variables into sales forecasting models would improve Walmart's ability to predict its sales. Table 4.3 outlined a host of possible examples of this. In this study, we utilize available data on temperature provided by Walmart in the Kaggle dataset as a proxy for remote sensing data. This is appropriate for two main reasons – it has the same spatial and temporal qualities of remote sensing data for Walmart and, while temperature data is widely available from many information sources, it is something that can be (and often is) remotely sensed and could be provided by remote sensing companies. The spatial and temporal nature of the temperature data also matches with our second strategic recommendation in the first part of this study – that remote sensing companies should focus on the provision of medium spatial and temporal resolution data. Temperature variations day to day and within the region surrounding a store may and do impact purchasing behavior of Walmart customers.

To quantify the specific impact of temperature on sales forecasting, we created three nested linear, multivariate regression models that are conceptually defined by Equations 4.2-4.4 and follow the form of Equation 4.1. The first of these models (Model 1) includes all independent variables in Table 4.4, with the exception of temperature. Model 1 serves as our base model (Equation 4.2). We then developed two additional models to study different approaches in quantifying temperature in the context of sales. Model 2 (Equation 4.3) treats temperature as a continuous variable that is defined in units of degrees (°F). This is intended to quantify the impact that including temperature change has on sales forecasting accuracy. Model

3 (Equation 4.4), incorporates temperature in two binary variables that indicate if the temperature that day was below freezing (less than 32°F) or particularly warm (greater than 90°F). The assumption in this model is that extreme temperature may have a larger effect on sales volumes that intermediate temperatures. It also allows us to examine whether highly tailored (curated) information is better for business decision making.

$$SALES_{ti} = f(Independent\ Variables_{ti}) \tag{4.2}$$

$$SALES_{ti} = f(Independent\ Variables_{ti}, Temperature_{ti}) \tag{4.3}$$

$$SALES_{ti} = f(Independent\ Variables_{ti}, Temperature_{Hot_{ti}}, Temperature_{Cold_{ti}}) \tag{4.4}$$

$$for\ i = 1{:}\,45\ stores\ and\ t = 1{:}\,n\ weeks$$

All three multivariate regression models were fit in R using ordinary least squares and are presented below in Table [*R Core Team*, 2016]. We employed a Likelihood Ratio Test to compare the relative goodness of fit between Model 1 and Model 2 and between Model 1 and Model 3. More information on the concept of likelihood and likelihood ratio tests can be found in *Cox and Hinkley* [1974]. The likelihood ratio test results, via the Chi-squared test statistic, for each model comparison are presented below in column five of Table 4.5. In both cases, the models with temperature were significantly more predictive than the base model. We also tested the relative contribution of temperature to the explanatory power of the sales forecasting models using the *relaimpo* package in R [*Grömping,* 2006]. This package employs the *Lindeman, Merenda and Gold* [1980)] approach to assessing the relative importance of each independent variable [*Grömping,* 2006]. Column four of Table 4.5 shows that temperature explains 0.828% of weekly store sales, while a temperature less than 32°F would explain 0.751% of weekly store sales (the impact of a temperature greater than 90°F was not significant). Both of these variables are significant within each respective regression model according to a t-test (Table 4.5 column 3).

Table 4.5 Statistical test results for multivariate regression of Models 1-3

Model	adj. R^2	b_{TEMP} [Pr(>|t|)]	Percent variance explained, TEMP	Likelihood ratio test Pr(>Chi) vs. base	Mean absolute error
1. Base	0.5736	n/a	n/a	n/a	$276,240.70
2. Base + TEMP	0.5879	$6,876 [7.17e-12***]	0.828%	4.708e-12***	$269,842.80
3. Base + TEMP_C + TEMP_H	0.5793	-$160,338[1] [3.1e-05***]	0.751%	8.34e-05***	$274,553.00

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; [1]Only the variable TEMP_C was significant

Using the empirical results in Table 4.5, we can draw inferences about the impact that geospatial information could have on sales forecasting. Incorporating our proxy geospatial datum, temperature, provided a solid prediction of about $6,400 more in sales per week per store. Walmart operates over 4,600 stores across the United States [*Walmart*, 2016]. If temperature information had a similar impact on each store, Walmart could predict upwards of $500,000,000 more in sales each year. This reduction in uncertainty of sales estimates could translate into cost savings and revenue generation, as Walmart would more efficiently store and move its goods to keep them in stock to meet demand. In addition, the results from Model 3 show us that value may be gained by further processing geospatial data based upon knowledge of consumer behavior. Retailers know that cold weather affects sales volume. Bringing in knowledge of a client's business along with advanced analytical methods could better curate information for the business intelligence needs of a potential client.

4.7 Conclusions

The market for satellite-based remote sensing data and services has undergone considerable change over the past decade. Although the U.S. government has traditionally driven this sector, recent changes within the industry and growth in the availability of COTS technology have made cheaper and smaller satellites possible, and led to rapid growth in the commercial remote sensing industry. However, despite investor speculation and growth in revenue, the use of today's less-expensive remote sensing imagery has not yet spread into new sectors. To understand these issues better, we examined the health of the commercial remote sensing market, finding that a key reason for the lag in the growth and diversification of the consumer base for remote sensing data is the disconnect between the business intelligence needs of potential clients and the data services that remote sensing companies are providing.

Given our belief that a healthy market should have diverse revenue streams and be robust to internal and external market fluctuations, we recommend three strategies for remote sensing firms to consider in order to create a healthier B2B market for their products. First, it is critical for remote sensing suppliers to develop a deeper understanding of the business needs of new customers outside traditional insurance, agriculture, resources, and government communities; this understanding can help identify key entry points within potential clients' decision processes where remote sensing data could add value. Second, our research suggests that products based on medium spatial and temporal resolution data and data products offer opportunities to address currently unmet business needs, and provide differentiation for commercial suppliers relative to free remote sensing data and existing market competitors. Third, we recommend a focus on data analytics to support business intelligence; companies have an urgent need for better data refinement capabilities even at today's level of commercial remote sensing capacity.

We also showed that remote sensing companies could generate a value-added proposition for their products in new sectors. We performed a case study of Walmart's supply chain, which included a statistical analysis of sales forecasting. Because Walmart has a significant breadth of product offerings and is considered a global corporate leader, the lessons learned from case study of Walmart are applicable to many other companies and sectors. We found that sales forecast uncertainty could be significantly reduced from using a single proxy remote sensing data product, and we speculated that it could be reduced even more through the provision of relevant business intelligence derived from remote sensing data.

Overall, a large potential for remote sensing industry growth exists in new market verticals. To date, penetration into new verticals has been limited by both lack of capability within potential clients to utilize remote sensing data, and the inability of remote sensing companies to supply business intelligence. Our analysis shows a path for new remote sensing entrants to create value for these clients and monetize their offerings.

4.8 References

BBC Research (2016), Remote Sensing Industry Expanding Commercial Opportunities, Reports BBC Research, BBC Research LLC. Market Wired. http://www.marketwired.com/press-release/remote-sensing-industry-expanding-commercial-opportunities-reports-bcc-research-2113909.htm

Bernknopf, R. and C. Shapiro (2015), Economic Assessment of the Use Value of Geospatial Information, ISPRS International Journal of Geo-Information. 2015, 4.

Brea-Solís, H., Casadesus-Masanell, R. and E. Grifell-Tatjé (2012), Business Model Evaluation: Quantifying Walmart's Sources of Advantage, Harvard Business School. Working Paper 13-039. 6 November 2012.

Burgel, O. and G. Murray (2000), The International Market Entry Choices of Start-Up Companies in High-Technology Industries, Journal of International Marketing: Summer 2000, Vol. 8, No. 2, pp. 33-62.

Cox, D. and D. Hinkley (1974). Theoretical statistics. Chapman and Hall. ISBN 0-412-12420-3

Dalby, R. (2015), Space Market Disruption: The Smallsat Revolution, Avascent. White Paper. October 2015. www.avascent.com

Darrow, B. (2015), How sensors are helping mining companies save money, Fortune. 27 July 2015. http://fortune.com/2015/07/27/rockwell-automation-mining/

de Selding, P. (2016), The state of the satellite industry in 5 charts, SpaceNews Magazine. 20 June 2016. http://www.spacenewsmag.com/feature/the-state-of-the-satellite-industry-in-5-charts/

DigitalGlobe (2014), Data Sheet - WorldView-3, DS-WV3 09/14. www.digitalglobe.com

DigitalGlobe (2016), Data Sheet - WorldView-2, DS-WV2 02/16. www.digitalglobe.com

Dillow, C (2015), Why small satellites are so big right now, 4 August 2015. Fortune. http://fortune.com/2015/08/04/small-satellites-newspace/

eoPortal (2016), Satellite Missions Database, Earth Observation Portal. https://directory.eoportal.org

Forney, W.M., Raunikar, R.P., Bernknopf, R.L., and Mishra, S.K. (2012), An economic value of remote-sensing information — Application to agricultural production and maintaining groundwater quality: U.S. Geological Survey Professional Paper 1796, United States Geological Survey. 60 p.

Fortune (2016), The Fortune 2016 Global 500, Fortune. Time Inc. http://beta.fortune.com/global500/

Foust, J. (2013), Smallsat constellations: the killer app?, The Space Review. Space News. 1 July 2013. http://www.thespacereview.com/article/2322/1

GeoConnections (2016), Value Findings Study Report, Canadian Geospatial Data Infrastructure. Information Product 48e doi:10.4095/297711

Geospatial World (2013), Economic Value of Geospatial Data: The great enabler, Geospatialworld.net. 22 May 2013.

Gissot, A. *et al.* (2015), Is the Space Industry Experiencing a Speculative Bubble? Venture Capital in the Space Sector, Toulouse Business School.

Greenspan, R. (2015), Walmart: Inventory Management, The Panmore Institute. http://panmore.com/walmart-inventory-management

Grömping, U. (2006), Relative Importance for Linear Regression in R: The Package relaimpo, Journal of Statistical Software. Vol 17:1.

Hellman, C. (2014), How Walmart Became A Green Energy Giant, Using Other People's Money, Forbes. http://www.forbes.com/sites/christopherhelman/2015/11/04/walmarts-everyday-renewable-energy/#36115074894e

Hertzfeld, H. *et al.* (2003), Weather Satellites and the Economic Value of Forecasts: Evidence from the Electric Power Industry, 54th International Astronautical Conference. Bremen, Germany. September 2003.

Hope, B. (2016), Tiny Satellites: The Latest Innovation Hedge Funds Are Using to Get a Leg Up, The Wall Street Journal. 14 Aug. 2016. http://www.wsj.com/articles/satellites-hedge-funds-eye-in-the-sky-1471207062

Joint Chiefs of Staff (2013), Joint Intelligence, Joint Publication 2-0. 22 October 2013. http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf

Kaggle (2014), Walmart recruiting – Store sales forecasting, Kaggle Inc. https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting

Kearns, J. (2015), Satellite Images Show Economies Growing and Shrinking in Real Time, Bloomberg BusinessWeek. 8 July 2015. http://www.bloomberg.com/news/features/2015-07-08/satellite-images-show-economies-growing-and-shrinking-in-real-time

Keith, A. (2015), Emerging Programs, Markets Drive Earth Observation Growth, Earth Imaging Journal. 21 February 2015. http://eijournal.com/print/articles/emerging-programs-markets-drive-earth-observation-growth

Kramer, H. (2016), WorldView-4, eoPortal Directory. https://directory.eoportal.org

Krebs, G. (2016), Spacecraft: Earth Observation – USA, Gunter's Space Page. http://space.skyrocket.de/

Lal, B. *et al.* (2015), Global Trends in Space Volume 2 Trends by Subsector and Factors that Could Disrupt Them, IDA Paper P-5242, Vol. 2. IDA Science and Technology Policy Institute. www.ida.org

LaValle, S. *et al*. (2011), Big Data, Analytics and the Path From Insights to Value, MIT Sloan Management Review. Magazine Winter 2011. Massachusetts Institute of Technology.

de Leeuw, J. *et al.* (2014), The Potential Uptake of Remote Sensing in Insurance: A Review, Remote Sensing. doi:10.3390/rs61110888

Lindeman, R., Merenda, P. and R. Gold (1980), Introduction to Bivariate and Multivariate Analysis, Scott Foresman. Glenview, IL.

Loomis, J. *et al.* (2015), Valuing Geospatial Information: Using the Contingent Valuation Method to Estimate the Economic Benefits of Landsat Satellite Imagery, Photogrammetric Engineering & Remote Sensing. Vol. 81. No 8. August 2015.

McAfee, A. and E. Brynjolfsson (2012), Big Data: The Management Revolution, Harvard Business Review. October 2012.

NASA Applied Sciences Program. (2012), Measuring Socioeconomic Impacts of Earth Observations – A Primer, National Aeronautics and Space Administration. Earth Sciences Division. www.nasa.gov

Navulur, K., Pacifici, F. and B. Baugh (2013), Trends in Optical Commercial Remote Sensing Industry [Industrial Profiles], IEEE Geoscience and Remote Sensing Magazine, vol. 1, no. 4, Dec. 2013. doi: 10.1109/MGRS.2013.2290098

NOAA (2015), What is remote sensing? National Oceanic and Atmospheric Association. http://oceanservice.noaa.gov/facts/remotesensing.html

Office of Technology Assessment (1994), Remotely Sensed Data: Technology, Management and Markets, OTA-ISS-604. Congress of the United States.

Ozment, D. (2014), Walmart's Approach to Renewable Energy, Walmart Today. http://blog.walmart.com/sustainability/20140508/walmarts-approach-to-renewable-energy

Planet Labs (2015), Planet Labs Specifications: Spacecraft Operations & Ground Systems, Version 1.0. Planet Labs. www.planet.com

The Tauri Group (2016), 2016: State of the Satellite Industry Report, Satellite Industry Association. June 2016. www.sia.org

RS Metrics (2016), Products, https://www.rsmetrics.com/

Orbital Insight (2016), Solutions, https://orbitalinsight.com/

Pohl, C. and J. L. Van Genderen (1998), Review article Multisensor image fusion in remote sensing: Concepts, methods and applications, International Journal of Remote Sensing, 19:5, 823-854, DOI: 10.1080/014311698215748

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Sankar, U. (2007), The Economics of India's Space Programme: An Exploratory Analysis, Oxford University Press.

Shepherd, M. (2016), When It Comes To U.S. Weather Forecasting: Private, Public Or Both? Forbes. 7 June 2016. http://www.forbes.com/sites/marshallshepherd/2016/06/07/when-it-comes-to-u-s-weather-forecasting-private-public-or-both/

Smithson, N. (2015), Walmart: Operations Management 10 Decisions, Productivity, Panmore Institute. http://panmore.com/walmart-operations-management-10-decisions-areas-productivity-case-study-analysis

Soni, P. (2015), Analyzing Walmart – The World's Largest Retailer, Market Realist. http://marketrealist.com/2015/02/analyzing-walmart-worlds-largest-retailer/

Spaceknow (2016), Spaceknow trading indexes, https://spaceknow.com/

Stergiou, K.I., Christou, E.D., and G. Petrakis (1997), Modelling and forecasting monthly fisheries catches: comparison of regression, univariate and multivariate time series methods, Fisheries Research. 29.

Thenkabail, P., Gamage, M. and V. Smakhtin (2004), The use of remote sensing data for drought assessment and monitoring in southwest Asia, Colombo, Sri Lanka: International Water Management Institute (IWMI) v, 25p. (IWMI Research Report 85). doi: 10.3910/2009.086

Traub, T. (2012), Wal-Mart Used Technology to Become Supply Chain Leader, arkansasbusiness.com Arkansas Business Limited Partnership, July 2, 2012.

University Alliance (2016), Walmart: Keys to Successful Supply Chain Management, University of San Francisco. http://www.usanfranonline.com/resources/supply-chain-management/walmart-keys-to-successful-supply-chain-management/#

USGS (2015), Landsat—Earth observation satellites: U.S. Geological Survey Fact Sheet 2015–3081, United States Geological Survey. United State Department of the Interior. November 2015. http://dx.doi.org/10.3133/fs20153081.

Ursa Space Systems (2016), Applications, http://ursaspace.com/

Wailgum, T. (2007). "How Wal-Mart lost its technology edge". CIO. IDG. http://www.cio.com/article/2437953/strategy/how-wal-mart-lost-its-technology-edge.html

Waller, M. and PV Boccasam (2013), How Data Sharing Drives Supply Chain Innovation, IndustryWeek. http://www.industryweek.com/supplier-relationships/how-sharing-data-drives-supply-chain-innovation

Walmart (2016), Our Locations, http://corporate.walmart.com/our-story/our-locations

Wang, X. Wu, Y., Liang, L. et al. (2016), Service outsourcing and disaster response methods in a relief supply chain, Annals of Operations Research. 240: 471. doi:10.1007/s10479-014-1646-y

Wulfraat, M. (2016), The Walmart Distribution Center Network in the United States, MWPVL International Inc. http://www.mwpvl.com/html/walmart.html

5. Conclusions

This dissertation is motivated by the principle that data availability and scientific analysis are fundamental for effective natural resources management. The research in this thesis presents approaches that can enhance water management institutions' ability to measure groundwater resources and to manage groundwater extractions. This is achieved through the development of a downscaling model and a hydrologically-based framework, both of which are applied to California's Central Valley. This research draws from a diverse scientific body of work in numerical modeling, remote sensing science, hydrology and public policy. Further, groundwater management in California can be seen as a case study of a region that requires management under resource scarcity. The lessons and takeaways presented can be applied worldwide to similar natural resource management issues subject to scarcity and changing climate dynamics.

In Chapter 2, research was presented that developed a robust, artificial neural network model that downscales GRACE gridded land datasets (~1 degree) to higher-resolution (~4km) groundwater storage change estimates. The model utilized GRACE estimates of variations in total water storage and a series of widely available hydrologic variables (PRISM precipitation and temperature data, DEM-derived slope and NRCS soil type) to derive spatial patterns in groundwater behavior. The model was able to effectively simulate groundwater storage change over the central and southern portions of the Central Valley with Nash-Sutcliffe model efficiency coefficient (NSE) values ranging from 0.0391 to 0.7511. This study also showed that the model required richer estimations of groundwater data (kriged datasets) for improved calibration and validation performance. The output of the downscaling model – high-resolution maps of groundwater storage change – illustrates the high heterogeneity in groundwater behavior and the tendency for more dramatic declines in the groundwater table to occur in the southern and western portions of the San Joaquin Valley and Tulare Groundwater Basins.

Overall, the extension of GRACE data by means of numerical downscaling represents a unique contribution to the scientific remote sensing community and advances the state of current remote sensing-based hydrologic science. This approach departs from data assimilation methods in that it is model-independent and thus offers more flexibility in data scarce environments and with changing input data products (i.e. new data releases or alternate remote sensing products). This has implications for worldwide applicability in developing regions, where models and dense monitoring networks may not be available. This neural network model also constitutes an alternative, numerical approach to improving the resolution of remote sensing products and offers a hybrid solution between low-resolution GRACE data and sparse groundwater monitoring networks.

In Chapter 3, a three-step framework is introduced that quantifies sustainable yield. First, we present a modified Hill's method that is more appropriate for California's highly variable hydrology. Basin managers can use this method to derive sustainable yield values – where sustainability can take the classic definition of zero annual average groundwater level change or represent any other groundwater level change (depletion) that basin managers consider tolerable. Second, we present an approach to integrating groundwater quality, land subsidence and surface water undesirable results into the quantification of sustainable yield. This method is flexible to varying types of input data and can incorporate complex relationships between the undesirable results and groundwater levels. Finally, the long-term implications of three different definitions for sustainable yield are assessed through an empirical groundwater balance that is projected to 2040. The results of these three scenarios show that there are tradeoffs to be had between groundwater availability, future climate uncertainty and socio-economic preferences that must be carefully weighed. This work contributes a framework that groundwater agencies can use to assess the long-term impacts of sustainable yield-based management scenarios that take into account both future groundwater availability and potential undesirable effects.

Chapter 4 examined the health of the commercial remote sensing market, finding that a key reason for the lag in the growth and diversification of the consumer base for remote sensing data is the disconnect between the business intelligence needs of potential clients and the data services that remote sensing companies are providing. The future of commercial remote sensing has important implications for natural resource management, among many other sectors, and its long-term success as a private industry will depend on its ability to diversify. Overall, a large potential for remote sensing industry growth exists in new market verticals. To date, penetration into new verticals has been limited by both lack of capability of potential clients to utilize remote sensing data, and the inability of remote sensing companies to supply business intelligence. Our analysis shows a path for new remote sensing entrants to create value for these clients and monetize their offerings.

5.1 Contributions

The following summarizes the main contributions from each chapter:

Chapter 2:

- Advances the state of current remote sensing-based hydrologic science by numerically downscaling GRACE data products.
- Develops a robust, artificial neural network model that downscales GRACE gridded land datasets (~1 degree) to higher-resolution (~4km) groundwater storage change estimates.

- Produces a series of annual high-resolution groundwater maps that offer a hybrid solution between low-resolution GRACE data and sparse groundwater monitoring over the GRACE time period (2002-present).

Chapter 3:

- Enhances the scientific integrity of groundwater management in California.
- Develops a methodology to calculate sustainable yield under SGMA that is flexible to varying input data, threshold values for undesirable results and tolerable levels of groundwater depletion.
- Assesses the sustainability of sustainable yield approaches to groundwater management.
- Offers a timely and easily implementable methodology for GSAs to calculate sustainable yield, manage extraction volumes and consider the tradeoffs of various management strategies.

Chapter 4:

- Develops a novel approach to quantify the value of geospatial data for decision makers.
- Provides an unbiased assessment of a rapidly developing branch of remote sensing – privately-owned and funded small satellites.
- Recommends three principle strategies for private satellite firms to grow their business, these strategies are also applicable to and highly useful for larger organizations like NASA and firms in related industries that provide geospatial data offerings.

5.2 Future Work

The following areas for future work are suggested extensions of the current research in this dissertation.

The downscaling model presented in Chapter 2 could be adjusted to include InSAR and GPS data as inputs, under the hypothesis that these geodetic measurements will improve the horizontal resolution and vertical accuracy of the model. InSAR and GPS data are already being used to enhance the horizontal resolution and vertical accuracy of GRACE data [*Sasgen et al*., 2013; *Matthews,* 2014]. Researchers have also shown that GPS observations of land surface deformation can reveal patterns of snow load variations over time [*Ouellette et al.,* 2013]. *Argus et al.* [2014] utilized inverted GPS observations of vertical motion to estimate surface water thickness at a quarter of the resolution of GRACE. In a similar way, inverted GPS data could be applied as an input to the neural network, offering finer resolution estimates of water storage variations. InSAR data has been used to look at patterns of subsidence and land deformation that have resulted from intensive groundwater withdrawals in California's San Joaquin Valley [*Sneed et al.*, 2013]. Specifically, this type of data would provide information on groundwater response to pumping and help to better refine the downscaling process.

A second extension of the downscaling approach could be to combine it with a physically-based model. The results of Chapter 2 showed that an empirically-based downscaling approach was improved through the addition of kriging, which provided a spatially continuous estimate of groundwater behavior. Instead of kriging, the downscaling model could be calibrated to the output of a high-resolution physical model, such as the Central Valley Hydrologic Model (CVHM) [*Faunt,* 2009]. This would also allow investigation of the model's ability to forecast in time and predict future groundwater storage changes over the model's spatial domain [*French et al.*, 1992; *Luk et al.*, 2000; *Ramírez et al.*, 2005; *Sun,* 2013]. It would also extend the maps produced to the entire Central Valley, rather than just the San Joaquin.

A third extension of the downscaling model would link the research in Chapters 2 and 3. Understanding how and where groundwater levels are changing in time is just one piece of the puzzle when it comes to better managing groundwater resources in a region. A second key component is insight into the causes behind these changes. Both climate change (via precipitation and potential evapotranspiration) and water management (via extraction, irrigation and land use) alter terrestrial water storage. In light of the newly mandated Groundwater Sustainability Plans (GSPs) for California, knowledge of the spatial distribution of these causal factors constitutes an important mechanism by which management plans can be adapted to the specific drivers within a groundwater basin. To achieve this, the high-resolution maps generated in Chapter 2 could be used to parse out the spatial distribution of the climatic and human impacts on groundwater change (conceptually following the methodology from *Ferguson and Maxwell*, 2012). This would produce a map showing the distribution in space and the magnitude of the impact of these factors. Using the framework presented in Chapter 3 and knowledge of the key drivers of groundwater depletion for a specific GSA, more appropriate management interventions (such as artificial recharge, pumping restrictions, etc.) could be determined.

The framework presented in Chapter 3 should be assessed in comparison to other sustainable yield approaches. Chapter 3 recommends a modification of Hill's method, but a variety of other approaches exist [*Loáiciga*, 2016]. These could each be incorporated into the framework to calculate a baseline sustainable yield value and used to assess long-term groundwater trends. In addition to incorporating these supplementary sustainable yield calculations, as more data and more models become available for groundwater management in the region, the robustness of the framework over the long-term should also be investigated.

Chapter 3 is an important contribution to the literature on sustainable yield for California and also serves as a template for newly formed GSAs to quantify groundwater extraction volumes (i.e. sustainable yield) year to year. Currently, neither California's State Water Resources Control Board nor Department of Water Resources (the two oversight bodies for SGMA) have released recommendations for sustainable yield as specific or comprehensive as those in Chapter 3. This utility of this chapter for widespread use requires clear communication and outreach to both of these agencies as well as to GSAs. With this outreach, the framework could be implemented or modified to meet the needs of individuals GSAs or be used a basis for a DWR Best Management Practice for sustainable yield calculation.

The research shown in Chapter 4 could be extended in two main ways. The first would be to size the current market for publicly provided data (NASA, NOAA) and assess the primary users of existing data products. A similar three-part strategy to what was proposed in Chapter 4 may be applicable to these government organizations, but a deeper dive into the main customers of publicly provided remote sensing data would offer more insight. The second main extension of Chapter 4 is related to the geospatial valuation component of the paper. Here there is a growing, but small, body of literature [*Forney et al.*, 2012; *NASA Applied Sciences Program*, 2012; *Loomis et al.*, 2015]. While Chapter 4 was not able to directly use satellite-provided data for the valuation model, an extension of this work could. This can be achieved through a case study of a company or industry with more granularity of sales data.

Overall, there is a clear and growing trend in multi-disciplinary approaches, based on publicly available data (either remotely sensed or in situ) and scientific models, that enhance and support decision making. Limitations in the utility of this research for end users and management agencies still exist. These limitations lay in the ability of decision makers to properly distill the breadth and complexity of scientific information to something that is convincing and actionable. Instead of solely providing model output, data products or academic research, scientists and researchers have the responsibility to continue engaging in work that is able to provide the form of information that is suitable for decision makers without reducing the scientific integrity of the solution or data offered. This requires outreach to decision makers, integrated, multi-disciplinary approaches, timely research and strong communication. Many researchers have been successful at this. The body of work presented in this dissertation could be further enhanced and more widely applied by following these lessons as well as the examples set by scientists who continue to make lasting impacts in their field.

5.3 References

Argus, D. *et al*. (2014), Seasonal variation in total water storage in California inferred from GPS observations of vertical land motion, Geophysical Research Letters. 41, 6.

Faunt, C., *et al.* (2009), Groundwater Availability of the Central Valley Aquifer, California, U.S. Geological Survey. Professional Paper 1766.

Ferguson, I. and R. Maxwell (2012), Human impacts on terrestrial hydrology: climate change versus pumping and irrigation. Environmental Research Letter. 7.

Forney, W.M., Raunikar, R.P., Bernknopf, R.L., and Mishra, S.K., (2012). "An economic value of remote- sensing information — Application to agricultural production and maintaining groundwater quality: U.S. Geological Survey Professional Paper 1796" United States Geological Survey. 60 p.

French, M., W. Krajewski, and R. Cuykendall (1992), Rainfall forecasting in space and time using a neural network, Journal of Hydrology 137.

Loáiciga, H.A. (2016), The Safe Yield and Climate Variability: Implications for Groundwater Management, Groundwater. doi: 10.111/gwat.12481

Loomis, J. *et al*. (2015). "Valuing Geospatial Information: Using the Contingent Valuation Method to Estimate the Economic Benefits of Landsat Satellite Imagery". Photogrammetric Engineering & Remote Sensing. Vol. 81. No 8. August 2015.

Luk, K.C. *et al.* (2000), A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting, Journal of Hydrology. 227:1-4.

Matthews, S. (2014), Improved groundwater monitoring through GRACE, Water Wheel Magazine. South African Water Research Commission.

NASA Applied Sciences Program. (2012). "Measuring Socioeconomic Impacts of Earth Observations – A Primer". National Aeronautics and Space Administration. Earth Sciences Division. www.nasa.gov

Ouellette, K. *et al.* (2013), Estimating snow water equivalent from GPS vertical site-position observations in the western United States, Water Resources Research. 49, 5.

Ramírez, M. *et al.* (2005), Artificial neural network technique for rainfall forecasting applied to the São Paulo region, Journal of Hydrology. 301:1-4.

Sasgen, I. *et al.* (2013), Antarctic ice-mass balance 2003 to 2012: regional reanalysis of GRACE satellite gravimetry measurements with improved estimate of glacial-isostatic adjustment based on GPS uplift rates, The Cryosphere, 7, 1499-1512.

Sneed, M. *et al.* (2013), Land subsidence along the Delta-Mendota Canal in the northern part of the San Joaquin Valley, California, 2003–10, U.S. Geological Survey Scientific Investigations.

Sun, A. (2013), Predicting groundwater level changes using GRACE data, Water Resources Research 49: 5900-5912.

6. Appendix A – Basics of Neural Networks

Techniques for spatial data analysis are complex and highly vulnerable to poor data quality and analysis. In addition, most of these techniques were developed in the 1960s and 1970s when computational power was significantly limited [*Turban et al.*, 2008]. Artificial neural networks (ANN), in contrast, offer the ability to efficiently and comprehensively handle large, diverse and noisy spatial datasets by drawing on the rapidly growing fields of computer science and systems analysis. ANNs were originally modeled after the neurons in the human brain and are now applied to a multitude of fields, showing strength in pattern classification, clustering and categorization, function approximation, forecasting and optimization. Their architecture offers four main advantages - machine learning, computational speed, high level of flexibility (not based on rigid assumptions), and robustness (in the face of noisy input data) [*Turban et al.*, 2008]. In essence, ANNs function by deriving non-linear, empirical relationships between a set of input and output variables. ANNs can then utilize these relationships to generate new estimates of the output variable from an alternate set of input data. More specifically, ANNs are a distributed information structure that consists of a network of weighted connections between input and output data. These connections are governed by numerical equations (termed processing elements) that are fit to best simulate the output data using the inputs provided. The number of input and output parameters is not restricted. The design of a neural network is based up the specific properties of these processing elements, the pattern of connections between them, and the algorithm used to determine the weights between connected processing elements. A variety of designs, or structures, of ANNs have been developed for different numerical purposes.

The study presented in Chapter 2 utilized a two-layer feedforward neural network structure, which has proven to be successful at solving multi-dimensional mapping problems

[*Turban et al.*, 2008]. Feedforward neural networks consist of a series processing elements (also called neurons) and the connections between them. These are arranged into a series of layers – an input layer, an output layer and a number of hidden layers in between. The number of processing elements in the input and output layer is dependent upon the number of variables in the input and output datasets, while the number of processing elements in the hidden layers can be adjusted depending on the complexity of the problem. This, along with the number of hidden layers, is problem-specific and is typically selected on a trial and error basis [*Zhu*, 2000]. Figure A1 shows the general structure of a two-layer feedforward network, as well as the various components of the processing elements and connections.

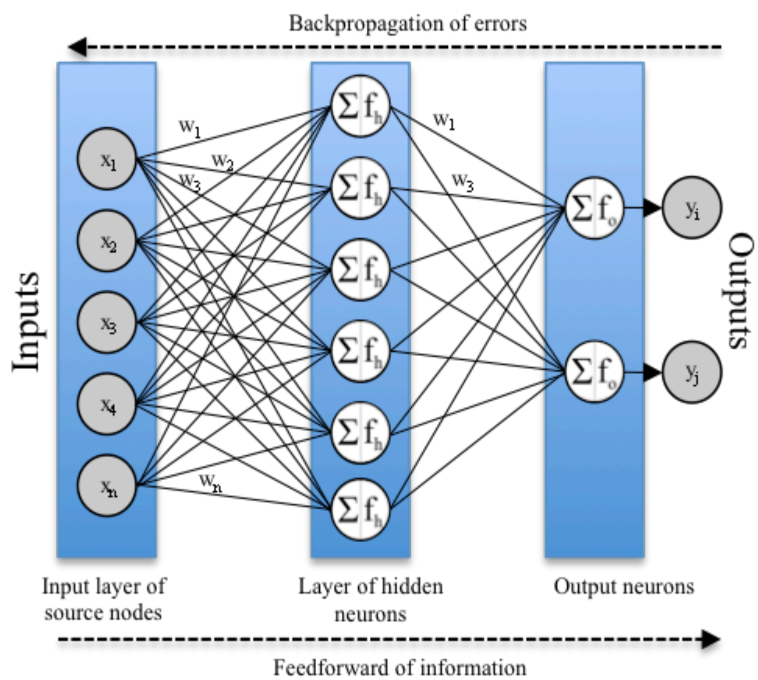

Figure A1 General structure of a two-layer feedforward neural network, with *n* input nodes, one layer of six hidden neurons and one layer of two output neurons.

Inputs to a neural network generally correspond to a set of attributes, while outputs indicate the solutions of the particular problem being addressed. In the case of an environmental problem, the inputs could come in the form of environmental variables such as precipitation, slope or canopy coverage, with the output or variable of interest being soil moisture, for example. Inputs and outputs are connected by a series of processing elements. The overall behavior of the neural network model is characterized by the weights ($w_{i,j}$), or connection parameters, that represent the strength of the connection between two processing elements. Each processing element is composed of a summation function and a transfer function. The summation function is represented mathematically as:

$$y_i = \sum_{i=1}^{n} x_i w_{ij} \qquad \text{(A1)}$$

Where y_j represents the output from neuron j; x_i is the input to neuron j and w_{ij} is the weight between them. The transfer function allows the neural network to activate or deactivate neurons based on the value of the summation function. Transfer functions are typically sigmoidal functions in which the outputs range from 0 to 1. In this way, the neural network numerically represents the activation level of each neuron without output values growing increasingly large [*Turban et al.* 2008]. A typical sigmoidal transfer function is:

$$y_T = 1/(1 + e^{-y_i}) \qquad \text{(A2)}$$

Where y_T is the transformed output and y_i is the output from neuron i.

Training, or network learning, is the process by which the neuron's interconnection weights are determined. The learning process involves the following steps: 1) Selection of initial weight vector; 2) Computation of outputs; 3) Comparison of generated outputs to training targets (error); 4) Adjustment of weights and repeat [*Turban et al.*, 2008]. The mechanism by which the network learns and systematically updates new weights is dependent upon its learning algorithm. This study employed a backpropagation algorithm for feedforward networks. Backpropagation, in general, is a method of non-linear optimization that seeks to minimize the sum of the error, i.e. the difference between the training data and the output of the network, by finding an optimal set of weights. Weights are adjusted using gradient descent, so that as the network learns, the sum of the error gradually decreases to a minimum value.

6.1 References

Turban, E. *et al.* (2008), Business Intelligence: A Managerial Approach, Pearson Prentice Hall. Upper Saddle River, N.J.

Zhu, A.-Xing (2000), Mapping soil landscape as spatial continua: The Neural Network Approach, *Water Resources Research*: 663.