ABSTRACT

Title of dissertation: ESTIMATION OF TERRESTRIAL WATER STORAGE IN THE WESTERN UNITED STATES USING SPACE-BASED GRAVIMETRY, GROUND-BASED SENSORS, AND MODEL-BASED HYDROLOGIC LOADING

Gaohong Yin, Doctor of Philosophy, 2020

Dissertation directed by: Associate Professor Barton A. Forman
Department of Civil and Environmental Engineering

Accurate estimation of terrestrial water storage (TWS) is critically important for the global hydrologic cycle and the Earth′s climate system. The space-based Gravity Recovery and Climate Experiment (GRACE) mission and land surface models (LSMs) have provided valuable information in monitoring TWS changes. In recent years, geodetic measurements from the ground-based Global Positioning System (GPS) network have been increasingly used in hydrologic studies based on the elastic response of the Earth′s surface to mass redistribution. All of these techniques have their own strengths and weaknesses in detecting TWS changes due to their unique uncertainties, error characteristics, and spatio-temporal resolutions. This dissertation investigated the potential of improving our knowledge in TWS changes via merging the information provided by ground-based GPS, GRACE, and LSMs. First, the vertical displacements derived from ground-based GPS, GRACE, and NASA Catchment Land Surface Model (Catchment) were compared to analyze

the behavior and error characteristics of each data set. Afterwards, the ground-based GPS observations were merged into Catchment using a data assimilation (DA) framework in order to improve the accuracy of TWS estimates and mitigate hydrologic state uncertainty. To the best of our knowledge, this study is the first attempt to assimilate ground-based GPS observations into an advanced land surface model for the purpose of improving TWS estimates. TWS estimates provided by GPS DA were evaluated against GRACE TWS retrievals. GPS DA performance in estimating TWS constituent components (i.e., snow water equivalent and soil moisture) and hydrologic fluxes (i.e., runoff) were also examined using ground-based in situ measurements. GPS DA yielded encouraging results in terms of improving TWS estimates, especially during drought periods. Additionally, the findings suggest a multi-variate assimilation approach to merge both GRACE and ground-based GPS into the LSMs to further improve modeled TWS and its constituent components should be pursued as a new and novel research project.

ESTIMATION OF TERRESTRIAL WATER STORAGE IN THE WESTERN UNITED STATES USING SPACE-BASED GRAVIMETRY, GROUND-BASED SENSORS, AND MODEL-BASED HYDROLOGIC LOADING

by

Gaohong Yin

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:
Associate Professor Barton A. Forman, Chair/Advisor
Associate Professor Hubert J. Montas, Dean's Representative
Associate Professor Kaye L. Brubaker
Assistant Professor Allison C. Reilly
Dr. Bryant D. Loomis

Foreword

Materials presented in Chapter 3 and Chapter 4 of this dissertation are based on two journal paper to be published as peer-reviewed articles in the next few months. The dissertation presented here is carried out in its entirety by Gaohong Yin. Please see the list of publications below.

Yin, G., Forman, B. A., Loomis, B. D., Luthcke, S. B. (2020) Comparison of Vertical Surface Deformation Estimates Derived from Space-based Gravimetry, Ground-based GPS, and Model-based Hydrologic Loading over Snow-dominated Watersheds in the United States. Journal of Geophysical Research: Solid Earth. (Minor revisions).

Yin, G., Forman, B. A., Wang, J., Loomis, B. D., Luthcke, S. B. Assimilation of Ground-based GPS Observations of Vertical Displacement into a Land Surface Model. Water Resources Research. (In preparation).

This research work is partially funded by NASA (grant NNX16AT03G) as well as through funding from the NASA GRACE-FO Science Team (NNX16AF17G).

Acknowledgments

I owe my gratitude to all the people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever. First and foremost I'd like to thank my advisor, Dr. Forman for giving me an invaluable opportunity to study in the United States and work on this interesting and challenging project over the past four years. He encourages me to be an independent researcher, but he is also always providing great help and insightful suggestions on my study and research. Dr. Forman not only taught me knowledge but also many valuable skills to support me as I continue pursuing my career goals. It has been a pleasure to work with and learn from Dr. Forman.

I would like to express my gratitude to Dr. Loomis from NASA's Goddard Space Flight Center. He provided great help in my research with patience, and he always can highlight the key point that I need to confront during my research. I have learned a lot from his enthusiasm and professional attitudes toward work. I would also like to thank Dr. Scott B. Luthcke, who is my co-author, and provides great support to my research activities.

I am deeply grateful to Dr. McCuen, who is always there willing to offer help and advice for both classes and professional career. Whenever I tried to seek help from Dr. McCuen, he never hesitated to help and shared many useful documents with me, inspiring my thoughts of an academic career. He is keen, rigorous, and has high expectations for all of his students, which motivates my progress. I would also like to thank Dr. Brubaker, Dr. Reilly, and Dr. Montas for agreeing to serve

on my dissertation committee and for sparing their invaluable time to review the manuscript.

Further, I want to express my thank to my colleagues: Jongmin Park, Jawairia A. Ahmad, Jing Wang, and Lizhao Wang. We have spent a lot of time together, and they have enriched my graduate life in many ways. They provided a lot of help in my research, courses, job searching, and all aspects of life. Without their support, I could not get through many difficult times and have had so wonderful of a graduate experience. I would also like to thank Dr. Yuan Xue for helping me adapt to my new life in the United States, and she keeps providing help from many aspects. I would like to thank Dr. Jing Tao and Dr. Yonghwan Kwon for providing me with many research suggestions. I would like to thank all of my friends at UMD for many meaningful discussions and enjoyable time spent together.

I owe my deepest thanks to my family – my mother and father, my sister and nephew – who have always stood by me, supported my decisions, and guided me through difficulties. They provided me a harbor that I could take a rest, get energy, and prepare with great courage to overcome future challenges.

Table of Contents

List of Tables vii

List of Figures viii

List of Abbreviations xii

1 Introduction 1
- 1.1 Terrestrial Water Storage . . . 1
- 1.2 Research Objectives . . . 4
- 1.3 Dissertation Structure . . . 5

2 Literature Review 7
- 2.1 GRACE in Terrestrial Hydrology . . . 7
- 2.2 Land Surface Models in Terrestrial Hydrology . . . 11
- 2.3 Ground-based GPS in Terrestrial Hydrology . . . 14
  - 2.3.1 GPS-based Displacement Study . . . 15
    - 2.3.1.1 Early GPS-based Surface Displacement Applications 16
    - 2.3.1.2 GPS Applications in the Himalayas . . . 18
    - 2.3.1.3 GPS Applications in the Western United States . . . 19
  - 2.3.2 GPS-Based Terrestrial Hydrologic Loading . . . 21
    - 2.3.2.1 GPS-based TWS using Green's Function . . . 22
    - 2.3.2.2 Other Applications of GPS-based Terrestrial Hydrologic Loadings . . . 25
- 2.4 Data Assimilation . . . 26

3 Comparison of Vertical Surface Deformation Estimates Derived from GRACE, Ground-based GPS, and the NASA Catchment Land Surface Model over Snow-dominated Watersheds in the Western United States 31
- 3.1 Motivation and Objectives . . . 31
- 3.2 Study Area and Methodology . . . 33
  - 3.2.1 Study Area . . . 33
  - 3.2.2 GPS Data . . . 33
  - 3.2.3 GRACE Data . . . 36

3.2.4 Land Surface Model . . . 39
3.2.5 SNOTEL Data . . . 42
3.2.6 Evaluation Metrics . . . 42
3.3 Results and Discussion . . . 43
3.3.1 Monthly Vertical Displacement Analysis . . . 43
3.3.2 Seasonally-Adjusted Variation Analysis . . . 52
3.3.3 Surface Deformation and Hydrologic Loadings . . . 56
3.3.4 Inter-annual Change Since Late 2010 . . . 60
3.3.5 Analysis of Discrepancies and Error Sources . . . 64
3.4 Conclusions . . . 67

4 Assimilation of Ground-based GPS Observations of Vertical Displacement into the NASA Catchment Land Surface Model 70
4.1 Motivation and Objectives . . . 70
4.2 Data and Methods . . . 72
4.2.1 Prognostic Land Surface Model . . . 72
4.2.2 GPS Observations . . . 74
4.2.3 Data Assimilation Framework . . . 77
4.2.3.1 Ensemble Kalman Filter . . . 77
4.2.3.2 Observation Operator . . . 79
4.2.4 Evaluation Approach . . . 79
4.2.4.1 GRACE TWS Retrievals . . . 80
4.2.4.2 SNOTEL SWE Measurements . . . 81
4.2.4.3 SCAN Soil Moisture Measurements . . . 81
4.2.4.4 GRDC Runoff Measurements . . . 82
4.2.4.5 Evaluation Metrics . . . 83
4.3 Results and Discussion . . . 85
4.3.1 Terrestrial Water Storage . . . 85
4.3.2 Snow Water Equivalent . . . 88
4.3.3 Soil Moisture . . . 94
4.3.4 Runoff . . . 97
4.3.5 Analysis of Assimilation Increments . . . 100
4.3.6 Normalized Innovation Sequence . . . 103
4.4 Conclusions . . . 105

5 Overall Conclusions and Future Work 109
5.1 Overall Conclusions . . . 109
5.2 Future Work . . . 112
5.2.1 Using GPS as an Independent Measurement to Estimate TWS Change . . . 112
5.2.2 Robustness Experiment of a GPS DA Framework . . . 112
5.2.3 Multivariate Data Assimilation for Estimating TWS . . . 113

Bibliography 114

List of Tables

3.1 Correlation coefficient, R, of vertical displacement time series derived from (1) GPS versus GRACE, (2) GPS versus Catchment, and (3) GRACE versus Catchment, and RMS reduction after removing GRACE and Catchment from the GPS signal, respectively. The selected stations correspond to the minimum (MPUT and RG09), median (LMUT and MC01), and maximum (FORE and P728) correlation coefficient between GPS- and GRACE-derived vertical displacement in the Great Basin and Upper Colorado watersheds. 46

4.1 Ensemble perturbation parameters for meteorological boundary conditions and model prognostic variables [a]. 73
4.2 Experimental design and evaluation strategy. 84
4.3 Watershed-averaged correlation coefficient (R) and unbiased root-mean-square difference (ubRMSD) computed between modeled TWS anomalies (i.e., the OL and GPS DA) and GRACE TWS retrievals. Bold font indicates statistically significant differences between the OL and DA at a 5% significance level based on the t-test. 87
4.4 Correlation coefficient, R, for three-month averaged runoff estimates derived from the OL, GPS DA, and GRACE DA simulations relative to the *in situ* GRDC measurements. Bold font indicates statistically significant differences between the OL and DA at a 5% significance level. 99

List of Figures

1.1 Definition of terrestrial water storage, which includes surface water, soil moisture, snow, ice, groundwater, and vegetation biomass over land. 2

2.1 Diagram of how GRACE and GRACE-FO measure gravity based on changes in satellite separation distance. The figure is accessible from https://gracefo.jpl.nasa.gov/resources/50/how-grace-fo-measures-gravity/. 8
2.2 Conceptual schematic of the components of TWS in the NASA Catchment Land Surface Model (Catchment) adapted from Forman et al. [1] where 1 represents the Catchment's profile soil moisture deficit, 2 is root-zone soil moisture excess, 3 is surface soil moisture excess, 4-6 are individual snow layers, and 7 is canopy interception. 13
2.3 Conceptual model of hydrologic loading impact on crustal deformation that can be directly measured with ground-based GPS sensors . 15

3.1 Study area including the Great Basin and Upper Colorado watersheds along with the GPS permanent stations shown as light blue dots. Three representative stations in each of the two watersheds used for discussion in Section 3.3.1 are shown as red upper triangles along with their four letter codes. 34
3.2 Flowchart of GPS, GRACE, and NASA′s Catchment model data preprocessing. 37
3.3 Vertical displacements caused by adding a uniform disc load with radius of 14 km and equivalent water height (EWH) of 1m. The vertical dashed line represents the edge of the disc. 41
3.4 Comparison of monthly-averaged vertical displacements derived from ground-based GPS (blue), GRACE TWS retrievals (dark green), and Catchment estimates (red) for GPS station locations in column (a) the Great Basin watershed and column (b) the Upper Colorado watershed. 45

3.5 Correlation coefficient maps of vertical displacements derived from (a) GPS versus GRACE; (b) GPS versus Catchment; and (c) GRACE versus Catchment for all examined GPS stations in the Great Basin and Upper Colorado watersheds. m = watershed-averaged R; GB = Great Basin; and UC = Upper Colorado. 49
3.6 Maps of RMS reduction of vertical displacements derived from (a) removing GRACE time series from GPS time series; (b) removing Catchment time series from GPS time series. A large RMS reduction value represents a better consistency between two data sets. 51
3.7 Network-wide mean vertical displacement time series derived from ground-based GPS (blue), GRACE (dark green), and Catchment (red) in the (a) Great Basin watershed and (b) Upper Colorado watershed. The mean±1 std (standard deviation) are shaded with corresponding transparent colors. 52
3.8 Example of removing mean seasonal cycle from processed monthly vertical displacement derived from ground-based GPS (blue), GRACE (dark green), and Catchment (red), of which (a) the processed monthly vertical displacement time series; (b) the mean seasonal cycle to be removed from (a); and (c) the seasonally-adjusted vertical displacement time series. The right column shows scatter plots of the seasonally-adjusted vertical displacements derived from (d) GPS versus GRACE; (e) GPS versus Catchment; and (f) GRACE versus Catchment at the example station. The correlation coefficient R values are listed in the lower right-hand corner, and the black lines in (d)-(f) represent the 1:1 line. 53
3.9 Network-wide mean seasonally-adjusted vertical displacement time series derived from ground-based GPS (blue), GRACE (dark green), and Catchment (red) in the (a) Great Basin watershed and (b) Upper Colorado watershed. The mean±1 std (standard deviation) are shaded with corresponding transparent colors. 54
3.10 Correlation coefficients of seasonally-adjusted vertical displacement after removing the mean seasonal cycle for (a) GPS versus GRACE; (b) GPS versus Catchment; and (c) GRACE versus Catchment for all examined GPS station locations in the Great Basin and Upper Colorado watersheds. The gray dots in (b) represent stations that provides a seasonally-adjusted correlation coefficient not statistically different from zero. m = watershed-averaged seasonally-adjusted R; GB = Great Basin; and UC = Upper Colorado. 55

3.11 SWE time series derived from ground-based SNOTEL stations in the (a) Great Basin and (b) Upper Colorado watershed, with individual station time series in gray and station ensemble mean time series in red. (c) shows the annually-averaged, maximum SWE derived from Catchment (gray scale), SNOTEL (red circles), and the annually-average amplitude of ground-based GPS vertical displacement (blue squares). The annual amplitude of GPS vertical displacement is calculated by dividing the difference between the maximum and minimum of each year by two. 58
3.12 Slope of linear regression fitted to GRACE-derived TWS change in units of cm/yr (a) before (2003-2009), (b) during (2010-2011) and (c) after (2012-2016) the 2010-2011 La Niña event following the method described in Han [2]. 63

4.1 Study area including the Great Basin and Upper Colorado watersheds along with ground-based stations of GPS (circle), SNOTEL (triangle), and GRDC (diamond) used in the study marked. The red squares indicate the location of the example pixels discussed in Section 4.3.2 for the SWE analysis. 74
4.2 Average changes (computed as difference between September 1 and March 1 estimates) across the years 2003 through 2015 in (a) interpolated GPS-based vertical displacement and (b) modeled TWS using Catchment. Note that the average vertical motions in (a) are all positive given the general loss of mass from the near-peak period of the snow season in March through the end of the summer in September. . 76
4.3 Watershed-averaged TWS anomaly time series from the OL (red), GPS DA (blue), and GRACE TWS retrievals (black dots) for the (a) Great Basin and (b) Upper Colorado watersheds. Each line represents the respective ensemble mean, and the ensemble range is shaded with corresponding transparent color. The error bars represent the time-invariant standard deviation of the GRACE observation error. 86
4.4 Time series of SWE derived from the OL (gray solid line), GPS DA (blue solid line), GRACE DA (red dash line), and SNOTEL measurements (black dots) at (a) pixel 1 located in the northeast of Great Basin; (b) pixel 2 located in the east of Upper Colorado basin. Pixel locations are highlighted in Figure 4.1. The monthly-averaged ensemble means from the OL and DA simulations are shown relative to the monthly-averaged SNOTEL SWE measurements. 89

4.5 Spatial distributions of correlation coefficient, R, for SWE between (a) the OL and SNOTEL; (b) GPS DA and SNOTEL; (c) GRACE DA and SNOTEL, and the difference in R calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SNOTEL measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SNOTEL measurements. 91
4.6 Spatial distributions of unbiased root-mean-square difference, ubRMSD, for SWE between (a) the OL and SNOTEL; (b) GPS DA and SNOTEL; (c) GRACE DA and SNOTEL, and the difference in ubRMSD calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SNOTEL measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SNOTEL measurements. 93
4.7 Spatial distributions of correlation coefficient, R, for surface soil moisture between (a) the OL and SCAN; (b) GPS DA and SCAN; (c) GRACE DA and SCAN, and the difference in R calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SCAN measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SCAN measurements. 95
4.8 Spatial distributions of unbiased root-mean-difference, ubRMSD, for surface soil moisture between (a) the OL and SCAN; (b) GPS DA and SCAN; (c) GRACE DA and SCAN, and the difference in R calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SCAN measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SCAN measurements. 96
4.9 Watershed-averaged, monthly increments for state variables *catdef* (gray) and SWE (black) in the (a) Great Basin watershed via GPS DA; (b) Upper Colorado watershed via GPS DA; (c) Great Basin watershed via GRACE DA; and (d) Upper Colorado watershed via GRACE DA. 101
4.10 GPS DA innovation statistics for the Great Basin (open marker) and Upper Colorado (filled marker) watersheds. Different marker shapes represent different observation error standard deviations varying from 2 mm to 6 mm. 105

List of Abbreviations

AOD	Atmosphere and Ocean De-aliasing
Catchment	NASA Catchment Land Surface Model
CF	Center of Figure
CM	Center of Mass
CWU	Central Washington University
DA	Data Assimilation
DLR	Germany Aerospace Center
EASE	Equal-Area Scalable Earth
EnKF	Ensemble Kalman Filter
EWH	Equivalent Water Height
GAGE	Geodesy Advancing Geosciences and EarthScope
GEOS-5	Goddard Earth Observing System Version 5
GIA	Glacial Isostatic Adjustment
GLDAS	Global Land Data Assimilation System
GPS	Global Positioning System
GRACE	Gravity Recovery and Climate Experiment
GRACE-FO	Gravity Recovery and Climate Experiment Follow-on
GRDC	Global Runoff and Data Center
GSFC	Goddard Space Flight Center
KBRR	K-band Range Rate
LSM	Land Surface Model
MERRA2	Modern-Era Restrospective Analysis for Research and Applications
NCEP	National Centers for Environmental Prediction
NMT	New Mexico Tech
NOAA	National Oceanic and Atmospheric Administration
NRCS	Natural Resources Conservation Service
OL	Open Loop
PF	Particle Filters
PREM	Preliminary Reference Earth Model
RMS	Root Mean Square
SCAN	Soil Climate Analysis Network
SH	Spherical Harmonics
SNODAS	Snow Data Assimilation System
SNOTEL	SNOw TELemetry
SWE	Snow Water Equivalent
TWS	Terrestrial Water Storage
ubRMSD	unbiased root mean square error

Chapter 1: Introduction

1.1 Terrestrial Water Storage

Terrestrial water storage (TWS; Figure 1.1) is defined as the sum of all forms of water on land, which includes surface water, soil moisture, snow, ice, groundwater, and vegetation biomass [3, 4]. The spatial and temporal distributions of TWS and its constituent components are critically important for the global hydrologic cycle and the Earth′s climate system [4]. Systematic monitoring of TWS changes across a variety of spatial scales (i.e., local, regional, continental, and global) is valuable for studies in climate change, water resources management, hazard mitigation, and food security [5].

It is challenging to monitor TWS variation using the traditional method of in situ measurements, especially at a large spatial scale due to the constraints in techniques, economy, and the limited data sharing among nations [5]. Ground-based networks do not provide a direct measurement of TWS, instead, they measure the constituent components of TWS such as snow, soil moisture, and groundwater. Moreover, the collocation issue among different networks (i.e., stations location mismatch among different networks), as well as the small footprint that can be represented by each measurement station (<10 m^2), limits the feasibility of using

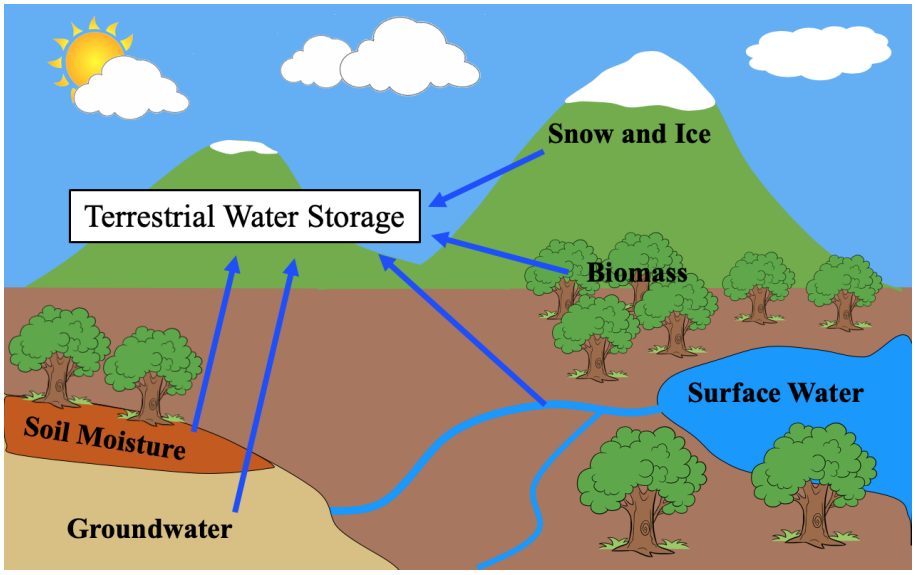

Figure 1.1: Definition of terrestrial water storage, which includes surface water, soil moisture, snow, ice, groundwater, and vegetation biomass over land.

ground-based networks for TWS monitoring [6].

Land surface modeling has achieved significant progress in the past two decades in representing terrestrial hydrologic variations [7] by providing continuous estimates of TWS and its components in both space and time. However, it is still far short of being able to provide accurate estimates of TWS at the global scale [7]. The common problem for land surface models (LSMs) is the lack of complete, all encompassing physical representations of the hydrological system such as groundwater pumping, irrigation practices, or the management of reservoirs [8]. Many LSMs do not include a deep ($\geq$10 m) groundwater component, which generally results in incomplete information about TWS changes [9]. In addition, the representation of interactions between surface and subsurface components, land and atmosphere components, the quality and availability of observations used in the models, and the computational capability also impede the efficacy of using LSMs to study TWS changes [7, 10].

The launch of the Gravity Recovery and Climate Experiment (GRACE) mission in March 2002 has provided an unprecedented opportunity to study the spatial and temporal variation of TWS at global and regional scales [11]. GRACE is a twin-satellite mission that detects the Earth′s gravity field changes. GRACE provides highly accurate (∼10–100 mm error) [12, 13], column-integrated TWS estimates at a spatial resolution of approximately 400 km and an approximate temporal resolution of monthly [11]. GRACE-based TWS retrievals have been successfully used in studies of drought monitoring [14, 15], flood potential [16, 17], groundwater depletion [18, 19], and ice mass loss [20, 21]. With the great success achieved by GRACE, the GRACE Follow-On (GRACE-FO) mission was launched in May 2018 as a successor of GRACE to continue tracking the Earth′s water movement. However, the coarse spatio-temporal resolution provided by GRACE and GRACE-FO retrievals limits their application to to fine-scale ($<10^5$ km^2; sub-monthly) TWS analyses.

In recent years, geodetic observations of surface displacement measured by the ground-based Global Positioning System (GPS) network have been increasingly used in hydrologic studies based on the elastic response of the Earth′s surface to mass redistribution [22]. Ground-based GPS observations have been compared against GRACE TWS retrievals in terms of vertical displacement and TWS anomaly (i.e., after removing long-term average) across a range of spatial scales, and a good consistency was found in most regions [22–26]. The potential of using ground-based GPS observations to bridge the gaps between GRACE and GRACE-FO missions was also investigated [27, 28]. Unfortunately, GPS-based surface displacements are not only influenced by the changes of TWS but also non-hydrologic loading effects

such as atmospheric and non-tidal ocean loading and tectonic (e.g., earthquake) effects. Additionally, the ground-based GPS network is sparse in most regions across the globe, which limits a broader application of ground-based GPS in estimating TWS.

Therefore, it is found that different data sets bring different pieces of information about TWS change based on their unique spatial and temporal resolution along with measurement characteristics. A combination of information derived from multiple data sources may provide an optimal estimate of TWS change.

1.2 Research Objectives

The overarching goal of this dissertation is to better understand the spatial and temporal features of TWS, and its components, at a watershed scale ($\sim 10^5$ km^2). Considering the aforementioned advantages and disadvantages of different techniques in estimating TWS, this study proposed to merge ground-based GPS observations into an advanced land surface model to improve the accuracy of estimated TWS as well as reduce the prediction uncertainty. The point-scale measurement nature of ground-based GPS observations provide the opportunity of characterizing TWS changes at a finer resolution in space and time relative to GRACE. By harnessing a Bayesian merger process, ground-based GPS can serve as a bridge to fill the gaps in spatial resolution between point-scale measurements of TWS and GRACE TWS retrievals, as well as the temporal gap between GRACE and GRACE-FO. More specifically, the dissertation is aimed to address the following scientific

questions.

1. How much variability in ground-based GPS observations of vertical displacement are associated with changes in TWS? Further, can ground-based GPS observations of vertical displacement be used as a proxy for representing changes in TWS at the watershed scale after removing the effects of non-hydrologic loading?

2. Will the merger of ground-based GPS observations of vertical displacement into an advanced land surface model improve the accuracy of modeled TWS estimates? Further, how will GPS assimilation perform in regions where snow is a significant component of the hydrologic cycle?

1.3 Dissertation Structure

The dissertation is structured as follows. Chapter 2 provides a literature review of the implementation of ground-based GPS, GRACE, and hydrologic models in terrestrial hydrology.

Chapter 3 compares the vertical displacements derived from ground-based GPS, GRACE, and an advanced land surface model to demonstrate the capability of using ground-based GPS to represent TWS variation. The discrepancies between different data sets and the error sources in each data set were analyzed. This chapter formed a manuscript entitled "Comparison of Vertical Surface Deformation Estimates Derived from Space-based Gravimetry, Ground-based GPS, and Model-based Hydrologic Loading over Snow-dominated Watersheds in the United States"

which is under review with the Journal of Geophysical Research: Solid Earth.

Chapter 4 proposes an integration of ground-based GPS observations of vertical displacement into an advanced land surface model using a data assimilation (DA) framework to improve the accuracy of estimated TWS and its constituent components, as well as mitigate the estimation uncertainty. This chapter formed a manuscript entitled "Assimilation of Ground-based GPS Observations of Vertical Displacement into a Land Surface Model" for submission to Water Resources Research.

Chapter 5 summarizes the overall findings in this dissertation, highlighted new and novel contributions made to the hydrologic research community, and discussed future work directions.

Chapter 2: Literature Review

2.1 GRACE in Terrestrial Hydrology

GRACE is a joint mission between the National Aeronautics and Space Administration (NASA) and the German Aerospace Center (DLR) that was launched in 2002 with the goal of mapping the Earth′s time-variable gravity field [29]. The launch of GRACE has provided a remarkable opportunity to monitor TWS changes across regional to global scales. GRACE consists of two identical satellites in identical, near-polar orbits at a altitude of ∼450 km and a separation distance between the two satellites of ∼220 km [29]. Due to the separation distance, the two satellites regularly experienced small, but different, orbital perturbations caused by the local mass variations on the Earth directly below the satellites (Figure 2.1). The range between the two satellites was measured using a K-band microwave ranging (KBR) system. The K-band range rate (KBRR) measurements, which is the first temporal derivative of KBR, was then used to solve for the spatial and temporal mass variations associated with the gravity field. The non-gravitational effects in the measurements (e.g., atmospheric drag, solar pressure) were removed using an accelerometer mounted at the center of mass of each satellite [30].

The most common method used to process GRACE range rate measurements

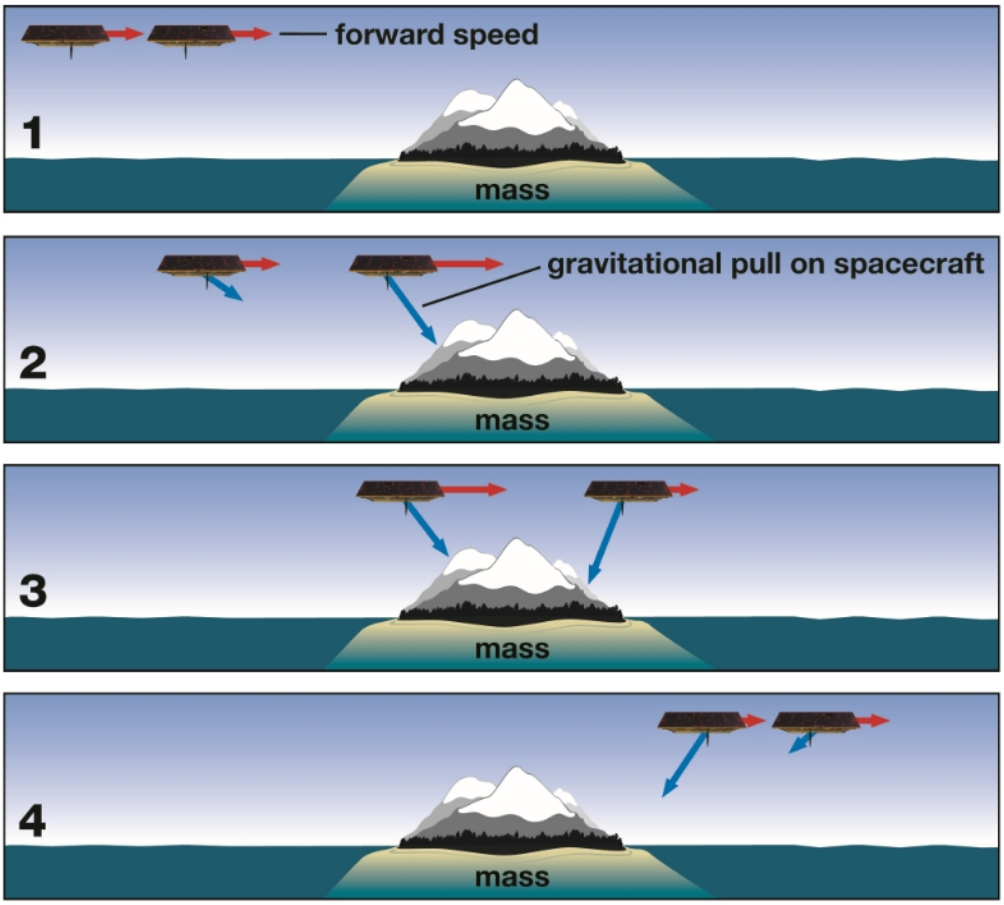

Figure 2.1: Diagram of how GRACE and GRACE-FO measure gravity based on changes in satellite separation distance. The figure is accessible from https://gracefo.jpl.nasa.gov/resources/50/how-grace-fo-measures-gravity/.

is via parameterization of the Earth′s gravity field using global spherical harmonics (SH) functions [31]. In recent years, an alternative approach, the mass concentrations (a.k.a., mascons) solution, has received increasing interest [32]. Scanlon et al. [33] summarized the basic difference between the SH and mascons solutions; SH solutions cannot distinguish between land and ocean areas, which results in a non-negligible signal leakage effect along coastal areas. On the contrary, the mascons solution can explicitly define land versus ocean regions, which can reduce the leakage errors and increase the spatial localization of TWS anomalies.

Another quandary in the processing of the range rate observations is that as the degree of the SH increases, the spatial resolution in the recovered TWS signal is improved, but the uncertainty also increases. Due to the poor observability of the east-west component of the gravity gradient associated with the near-polar orbit, GRACE data shows clear north-south stripes in the gravity field in the SH solutions [33]. Therefore, various methods were used in post-processing steps to reduce the noise in SH coefficients such as truncating high degree and order SH coefficients and applying a destriping filter [33]. The method used to reduce the noise in SH coefficients often causes signal loss [33]. Therefore, a scaling factor is used to restore the lost signal associated with post-processing [34]. Mascons, on the other hand, provide a more optimal solution to balance the trade-off between noise reduction and signal loss. Constraint equations are applied during the least square inversion in order to provide estimates of mascon parameters while suppressing the correlated errors and minimizing signal attenuation. Therefore, there is no requirement of post-processing for the mascons solution [35].

Numerous studies have been performed to investigate TWS changes at the spatial and temporal scales that can be detected by GRACE, and the efficacy of vertically disaggregating TWS into its components such as groundwater, soil moisture, or snow was also examined [10, 36, 37]. Since the launch of GRACE, both spherical harmonic solutions and mascons solutions have been used for hydrologic studies. A number of research studies have been conducted to study the long-term trends in TWS due to climate variability and anthropogenic activities [38]. Significant TWS depletion has been found in northwestern India [39–41], north-

ern China [42], southwestern Australia [43], and California′s Central Valley [18, 44] mainly associated with groundwater depletion as related to agricultural irrigation and domestic consumption. Further, GRACE-based TWS can be used as a remote sensing-based drought indicator to help monitor the climate change induced TWS depletion such as the 2011 drought in Texas [15] and the multi-year drought in southeastern Australia [43]. TWS depletion along the Greenland ice sheet [20,21,45], Antarctica [20,46,47], and the Tibetan Plateau [48] caused by increased glacier melt (and snow melt) were also revealed using GRACE TWS retrievals. There are also regions exhibiting increasing trends in TWS. Ahmed et al. [49] investigated the variation in water availability across Africa and found an increasing trend of TWS over western and central Africa caused by the warming of the tropical Atlantic Ocean, which intensified Atlantic monsoons and thus brought more precipitation on shore. Rodell et al. [50] mapped the trend of TWS globally using GRACE TWS retrievals, and categorized the driver for these trends into four categories as natural interannual variability (e.g., El Niño and La Niña), unsustainable groundwater consumption, climate change, and a combination thereof. Results showed an increase of terrestrial water storage in far-northern North America and Eurasia, and significant non-frozen-freshwater losses in the mid-latitudes. Human impact is also evident in the world′s irrigated agricultural regions, where a rapid decrease of freshwater storage was detected [50].

In addition to studying total freshwater changes, GRACE-based TWS has also been used together with *in situ* measurements and/or hydrologic models to monitor important elements of the hydrologic cycle such as soil moisture [51], evapotran-

spiration [52], and precipitation [53]. However, due to the coarse spatial resolution provided by GRACE, the application of GRACE TWS retrievals focused on regional, continental, and global scales with difficulties at implementing the TWS retrievals into smaller scale studies.

2.2 Land Surface Models in Terrestrial Hydrology

Prior to the launch of GRACE, there were few observations of TWS available, especially for large spatial scale applications. Land surface model, which models physically-based energy and water fluxes, is an important tool for global TWS studies. After the launch of GRACE, a retrieval of global TWS change with high accuracy became feasible for the first time. However, multiple GRACE TWS products using different data processing approaches based on different assumptions became available, and the efficacy of the different GRACE products have not been systematically investigated. Utilizing modeled TWS in a Bayesian merger can provide *a priori* information to combine a suite of different TWS products so as to improve the estimates of TWS change [38]. Further, the land surface models add value to the GRACE TWS retrievals by providing estimates of the energy and water fluxes along with estimates of the constituent components of TWS that GRACE does not provide.

The earliest LSMs were rather simple and prescribed the land surface moisture conditions, and did not explicitly include land-atmosphere interactions [54]. With developments over the past three decades, LSMs have become more compre-

hensive in representing more interactions and feedback between physical, biological, and chemical processes [55]. The Global Land Data Assimilation System (GLDAS) product is a joint effort by scientists at the NASA Goddard Space Flight Center (GSFC) and the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) [56]. GLDAS estimates land surface states (e.g., shallow groundwater, soil moisture, and snow water equivalent) by merging ground-based and space-based observations into a land surface model using a data assimilation framework [56]. GLDAS-1 utilizes four different LSMs: (1) Mosaic [57], (2) Noah [58], (3) the Community Land Model (CLM) [59], and (4) the Variable Infiltration Capacity model (VIC) [60].

GLDAS simulations based on the four different LSMs have been successfully used for estimating global and regional hydrologic cycles, spatiotemporal variations of TWS, and monitoring climate extremes such as drought and flood [4, 15, 61, 62]. However, changes in TWS from the LSMs have many large uncertainties due to errors in forcing (a.k.a, boundary conditions), parameterization of physical processes, and initial conditions [15]. Long et al. [38] evaluated the uncertainty in TWS estimates derived from the four LSMs (i.e., Mosaic, Noah, CLM, and VIC) using the three cornered hat approach. In the three cornered hat approach, TWS derived from GRACE, LSMs, and Global Hydrological Model (GHM) were used as three different sets of data, and the individual variance was separated using the cross correlation among the three data sets. Results show that Noah provides the lowest uncertainty (23.7 mm) followed by VIC (27.3 mm), and Mosaic has the highest uncertainty of 32 mm. To analyze the cause of the uncertainty in modeled TWS

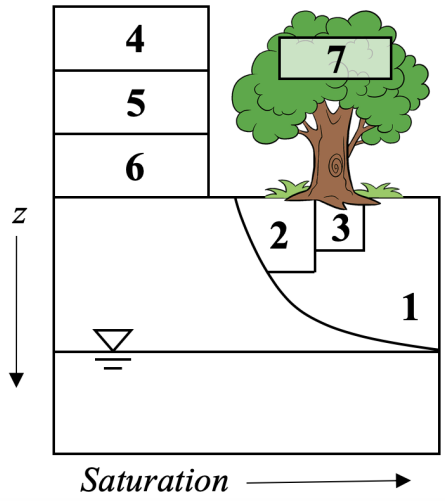

Figure 2.2: Conceptual schematic of the components of TWS in the NASA Catchment Land Surface Model (Catchment) adapted from Forman et al. [1] where 1 represents the Catchment's profile soil moisture deficit, 2 is root-zone soil moisture excess, 3 is surface soil moisture excess, 4-6 are individual snow layers, and 7 is canopy interception.

relevant to the boundary conditions, precipitation is considered to be the dominant source of mass, and its impact on land surface states typically varies with season and climate [63, 64]. All four LSMs include only a shallow (<10 m) groundwater component and no dynamic surface water routing or representation of glaciers [65], which hampered the performance of each model in estimating TWS. Furthermore, the simple one or two-layer snow models in each of the LSMs can lead to unrealistic runoff volume and timing. Dirmeyer et al. [66] assessed the representation of land surface heat balance and water storage, and their mutual impact on the performance of the four LSMs. It was shown that no individual model can accurately represent the interaction between latent heat and soil moisture, and that the multi-model mean generally provides a better estimate than most or all of individual models.

In an effort to improve the shortcomings of these LSMs, GLDAS upgraded

the Mosaic model to the Catchment Land Surface Model (Catchment) [54, 67] (Figure 2.2). Catchment is the land model component of the Goddard Earth Observing System, version 5 (GEOS-5). It was developed in response to a weakness in conventional LSMs that assume a soil-layer-based vertical discretization [68, 69]. The conventional, layer-based LSMs often assume a uniform horizontal structure of soil spanning over tens of kilometers, which neglects much of the horizontal variability of soil type, soil moisture, and its subsequent effects on evaportranspiration and runoff [67]. Catchment improved the conventional LSMs by employing a topographically-based hydrologic catchment as the fundamental land surface unit with an average area of approximately 4000 km^2. Sub-grid soil moisture heterogeneity was then treated explicitly by dividing the catchment into dynamic fractions of saturated, unsaturated, and wilting areas [68]. In addition, a three-layer snow model was used in Catchment, which accounts for snow melting and refreezing, dynamic changes in snow density, and other snow-related processes [54].

2.3 Ground-based GPS in Terrestrial Hydrology

The Earth deforms in an instantaneous, reversible, elastic fashion to surface loading variations caused by either hydrological components such as snow and ice, or non-hydrologic loading such as changes in atmospheric pressure. In terms of TWS changes, ground-based GPS observes downward motion (i.e., negative vertical displacement) of the Earth′s surface when water loading increases, and uplift displacement (i.e., positive vertical displacement) when water loading reduces (Figure 2.3).

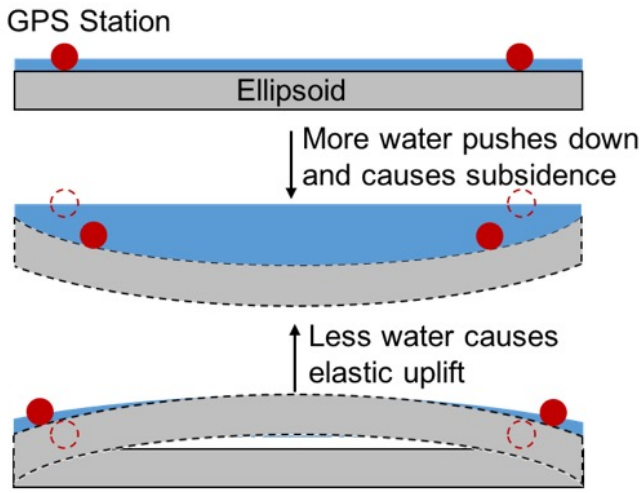

Figure 2.3: Conceptual model of hydrologic loading impact on crustal deformation that can be directly measured with ground-based GPS sensors

Farrell [70] proposed the loading theory in the form of Green′s function to quantified the relationship between mass variation and three-dimensional (3D) surface displacement. Therefore, terrestrial hydrologic mass redistribution can be resolved using ground-based GPS observations of surface displacement after accounting for the effects of non-hydrologic loading (e.g., atmosphere, non-tidal ocean, and plate motion). More details of the Green′s function were discussed in Section 3.2.4. Due to the larger response in vertical displacement (relative to horizontal displacement) for the same concentrated mass load, vertical displacement is much more frequently applied in hydrologic studies [71].

2.3.1 GPS-based Displacement Study

A number of studies have been conducted to compare surface displacements derived from ground-based GPS with that from GRACE in order to explore the us-

age of geodetic deformation in representing continental-scale water storage changes. GRACE-based vertical displacement was computed based on the elastic response of the Earth′s surface to TWS change by using GRACE TWS retrievals. Non-hydrological loadings embedded in the GPS observations were explicitly separated as a part of post-processing routine. However, systematic discrepancies between surface displacements derived from GRACE and GPS still exist. The surface displacements computed from GRACE retrievals are spatially smoothed, which only reflect the water loading changes that induce effects over large spatial areas. On the other hand, the GPS-based displacements are point-scale measurements that are significantly influenced by local effects, technical artifacts, and inadequate environmental models used for removing non-hydrologic loading effects [72].

2.3.1.1 Early GPS-based Surface Displacement Applications

Early studies showed that vertical displacements computed from GRACE TWS retrievals are in good agreement with GPS-based observations of surface displacements only in regions with high mass changes. Davis et al. [73] demonstrated the high-consistency between the vertical displacements derived from GRACE and ground-based GPS over the Amazon River Basin where GRACE suggests a $\sim$13 mm annual change in vertical displacement. King et al. [74] compared GRACE-based and GPS-based vertical displacements across multiple locations where GRACE-derived signals are large or small. Results showed good agreement between GRACE and GPS for sites that are surrounded mainly by land, but not for sites with moder-

ate to small GRACE signals such as islands. van Dam et al. [23] compared vertical displacements over Europe, and a poor agreement was found between GRACE-based and GPS-based vertical displacements in terms of both amplitude and phase. The largest discrepancy was detected at stations near the coast, suggesting the presence of systematic errors in the GPS processing procedure (e.g., the mismodeling of ocean tidal loading signals).

With improved GPS measurement and processing techniques, a more accurate estimate of GPS positioning in three-dimensional space and more precise modeling of atmospheric pressure and tidal loading became available. Using GPS data with reduced spurious signals, Tregoning et al. [75] compared three-dimensional displacements derived from GRACE and GPS. A far greater consistency was achieved than the results from van Dam et al. [23] and King et al. [74]. However, a number of stations still showed poor consistency between the vertical displacements derived from GPS and GRACE. A potential explanation was that GPS observations are dominated by local water mass change, which does not cause large-scale influence that can be easily detected by GRACE. Therefore, Tesmer et al. [76] compared the vertical deformation derived from GRACE and GPS at locations where the surface displacements are not dominated by local effects or technical artifacts. Due to the pre-selection of GPS stations, as well as the improved accuracy in GPS processing techniques, a considerable enhancement was found compared to previous studies.

2.3.1.2 GPS Applications in the Himalayas

Bettinelli et al. [77] compared the three-dimensional displacements derived from GPS with GRACE in the Himalayas by assuming an elastic, half-space Earth model. Results showed that it is difficult to simultaneously find agreement in both amplitude and phase of displacements derived from GPS and GRACE for both vertical and horizontal components. Chanard et al. [78] further explored the inconsistency between GRACE-based and GPS-based surface displacements, and found that using the PREM (Preliminary Reference Earth Model) layered spherical Earth model (instead of an elastic half-space model) can enhance the agreement of displacements derived from GRACE and GPS. Other studies [26, 79] compared surface displacements between GRACE and GPS in both vertical and horizontal directions and revealed a good consistency in terms of both amplitude and phase in the Himalayas region. Additionally, long-term mass loss in the Himalaya region detected by GRACE was reported and explained by increased melting of snow and ice [26,79]. Saji et al. [80] also found subsidence in the sub-Himalaya region with the surrounding areas showing uplift, which is related to the unsustainable consumption of groundwater associated with anthropogenic activities.

With different Earth models, the load love numbers used to represent the response of the Earth′s structure to surface mass loading variations will be different. Wang et al. [81] compared the load love numbers computed for the elastic Earth models between PREM [82], iasp91 [83], and ak135 [84]. Results showed that the load love numbers for iasp91 and ak135 models are very close, but they both showed

large differences with the load love numbers as computed from the PREM model for degrees around 200 and higher. The large differences can be explained by the difference in elastic structure between PREM and iasp91 (or ak135) in the near-surface portion of the Earth′s crust, especially for a discontinuity at a depth of 220 km shown in the PREM model but not represented in the iasp91 or ak135 models [81]. Gu et al. [85] computed GRACE-based vertical displacements based on these Earth models (i.e., PREM, ispa91, and ak135) and found that the influence of these different Earth models is small in terms of mean annual amplitudes of the computed vertical displacements.

2.3.1.3 GPS Applications in the Western United States

Due to the severe drought in the western United States starting around 2012, as well as the existence of a dense GPS network [86], the region has been well-studied using GPS observations. Wahr et al. [71] monitored the level of Lake Shasta in northern California using GPS data, and demonstrated that incorporating horizontal motion (along with vertical displacement) can improve and extend the analysis related to water mass variation. The combination of horizontal and vertical displacements helps to infer whether the nearby hydrologic loading (i.e., water mass) is concentrated in a small region such as a single lake. Amos et al. [87] explored the groundwater loss in central California (including a large portion of California′s Central Valley) using GPS-based displacements. The vertical displacement uplift rate derived from GPS is slightly larger than the value derived from GRACE, but

they are in overall good agreement. Tan et al. [88] compared the vertical displacements derived from GRACE, GPS, and hydrologic models in California′s Central Valley. In general, a good agreement was found between the three data sets at most station locations. However, irregular peak periods were detected in the GPS-based vertical displacement time series at some stations, which was explained as the result of local groundwater pumping. Therefore, the application of GPS observations in regions with large groundwater changes needs to be carefully considered. In regions where surface displacement changes are caused by groundwater losses, the Earth′s surface exhibits a poroelastic response to groundwater changes in the near-field (i.e., region close to groundwater pumping location) above the aquifer, but is dominated by the elastic response in the far-field (i.e., region far from groundwater pumping location). Therefore, Green′s functions based on elastic response assumption can only be correctly used to estimate groundwater mass variation when the GPS station is located far from localized groundwater pumping activities [89]. More details of the poroelastic and elastic response were discussed in Section 3.3.5. Knappe et al. [90] separated GPS-based observations of vertical displacement into regional and local components, and then compared them with both GRACE-based vertical displacements and *in situ* SWE measurements from the SNOw TELemetry (SNOTEL) network in the northern Rockies. Results demonstrated the potential of using GPS-based observations of surface displacement to bridge the spatial scales between point-scale measurements from SNOTEL and relatively coarse-scale retrievals from GRACE.

GPS-based surface displacements were also used to study hydrologic loading

variations in other parts of the globe, such as western African (which has a distinct monsoon season) [24], the North China Plain (which has smaller seasonal variations) [91–93], and Australia (which experienced a significant water cycle budget imbalance over 2010-2015) [2]. Overall, a good consistency was detected between GPS-based, GRACE-based, and model-based surface displacements, which provides a solid empirical foundation for the proposed study of using GPS-based observation of surface displacement to estimate hydrologic loading variations with a particular focus on snow.

2.3.2 GPS-Based Terrestrial Hydrologic Loading

Research has shown repeatedly that terrestrial water mass change signals embedded in ground-based GPS observations of surface displacement can be inverted for use in hydrologic loading studies. The estimation of TWS change using GPS observations of surface displacement independently is referred to more simply as "inversion process" hereafter in this dissertation. Due to the limited accuracy and resolution of monthly GRACE gravity solutions at low degrees, many studies have attempted to use the global spherical harmonics parameterization technique to jointly use GPS measurements along with GRACE observations to provide estimates of mass redistribution across the globe. Kusche et al. [94] used a physically-motivated, regularized least squares method to invert GPS measurements into hydrologic loading space up to degree and order 7. It has been shown that after removing the effects of atmospheric loadings (associated with changes in surface pressure), the

remaining GPS signal provides similar patterns as the GRACE TWS retrievals as well as estimates from a global hydrologic model. Rietbroek et al. [95] estimated weekly hydrologic loading variations up to spherical harmonics degree and order 30 by combining data from GRACE gravimetry, GPS-based deformation, and modeled ocean bottom pressure. The strategy used by Rietbroke et al. [95] simultaneously solves for low degree surface loading, geocenter motion, and mitigates mass conservation effects related to the ocean model. Rietbroek et al. [28] further explored the plausibility of using GPS-derived surface loading to fill in the temporal gaps between the GRACE and GRACE-FO missions. Results from Rietbroek et al. [28] showed that the accuracy of the mass redistribution derived from ground-based GPS is lower than that derived from GRACE, but there is no doubt that GPS-based observations of surface displacement contain valuable information related to hydrological loading phenomena that can be used in hydrologic studies [28].

2.3.2.1 GPS-based TWS using Green′s Function

In recent years, an increasing number of studies have attempted to use ground-based GPS observations of vertical displacement to estimate TWS changes based on Green's functions proposed by Farrell [70]. Most studies are conducted across relatively small regions where a relatively dense GPS network exists, most notably in the western United States. Argus et al. [89] used GPS-based vertical displacements to resolve TWS changes across California with a spatial resolution of 50 km. The GPS-based TWS anomaly revealed that the seasonal variations in California were

dominated by the loading of snow and surface water. Additionally, changes in TWS derived from GPS are, on average, 50% larger than from the NLDAS-Noah hydrology model, which is explained by the larger amounts of snow and surface water storage that are not properly captured by the hydrologic model. Argus et al. [22] continued to use GPS-based observations of vertical displacement to study the severe drought in California between 2012 and 2015. The combination of GPS-based TWS variations and SWE estimates from the Snow Data Assimilation System (SNODAS) indicated a large loss of subsurface water (i.e., soil moisture and groundwater) in the Sierra Nevadas. However, the large loss of subsurface water was absent in the hydrology models. Borsa et al. [96] studied drought in the western United States during the year 2013 by estimating TWS using GPS-based surface displacement with a spatial resolution of approximately 200-300 km. Results showed that GPS can be used to estimate hydrologic loading changes caused by both dry and wet climate patterns. The GPS network in the western United States provides high precision measurements of surface deformation along with high sampling density and sensor stability, and thus, can be used to monitor long-term regional climate change [96].

By applying the Green′s functions, GPS has also been used to study TWS changes across Washington and Oregon [27], southwestern United States [97], and the lower Three Rivers headwater region in China [98]. It was shown that GPS is capable of estimating TWS variations independently at a finer spatial resolution than GRACE-based TWS anomalies. Milliner et al. [99] inverted GPS-based three-dimensional displacements for TWS to track daily changes in TWS caused by Hurricane Harvey's path across the Gulf Coast region. Results demonstrated that

it is feasible to use ground-based GPS for near-real-time flood alerts and remote monitoring of water storage during hurricanes and other severe storms. Ferreira et al. [62] was the first large-scale attempt to infer TWS from GPS-based vertical displacements using Green′s functions. The study was conducted over South America where only a limited number of GPS stations are available for use. A synthetic experiment was first performed; results showed the potential of using inverted TWS to fill the temporal gaps between the GRACE and GRACE-FO missions while providing TWS estimates at a spatial resolution of about 300 km with a daily temporal resolution. However, in the natural (i.e., non-synthetic) experiment using the actual ground-based GPS observations of vertical displacement, missing observations of daily-averaged vertical displacement in the GPS record often occurred. Therefore, the actual GPS-based TWS experiment performance was inferior to the synthetic experiment. That is, the experiment yielded a large root-mean-square error (RMSE) when compared against TWS estimates from GLDAS-Noah. Adusumilli et al. [100] studied the TWS changes across the contiguous United States over a decade (2007 to 2017) using ground-based GPS vertical displacements. The study quantified the influence of the El Niño/Southern Oscillation on TWS changes across the United States as well as the impact of atmospheric rivers on water storage in the western United States.

2.3.2.2 Other Applications of GPS-based Terrestrial Hydrologic Loadings

The capability of using Green′s functions to estimate TWS using ground-based GPS observations is limited when applied at a large spatial scale (>1,000,000 km^2) where ground-based GPS coverage is not as dense as that available in California [101]. Han et al. [101] developed an algorithm to estimate mass loading variation from GPS-based vertical displacement in Australia using regional spherical (Slepian) basis functions and a spectral domain approach. The inverted TWS change reflects mass variation at not only seasonal and inter-annual time scales, but also sub-monthly scales, thereby providing a good agreement when compared against GRACE-based TWS anomalies and GLDAS-Noah estimates of soil moisture. This consistency demonstrated the capability of the proposed method in resolving hydrologic loadings via surface displacement when the GPS network is relatively sparse. Instead of acquiring a direct estimate of TWS changes, ground-based GPS observations can also be converted into a drought index in regions with a limited number of GPS stations [6, 102]. Furthermore, a few studies explored the usage of GPS-based data to infer TWS component information such as snow [103, 104] or groundwater [105, 106]. The partitioning of TWS into its components using *in situ* measurements [103] or modeled data [104] offers opportunities to further disaggregate the GPS-based TWS change into its constituent components.

2.4 Data Assimilation

GRACE provides retrievals of column-integrated TWS change across the entire globe, but at a coarse spatial (∼400 km) and temporal resolution (∼monthly). Land surface model estimates of land surface states and fluxes are provided consistently at a fine spatial (∼1 km) and temporal (∼hourly) scale. The accuracy, however, is limited by inadequate model physics as well as uncertainty in model parameters and boundary conditions. In an effort to harness all of the available observations while mitigating the errors and uncertainties in each data set, an optimal estimation of terrestrial water storage change should integrate both the model estimates and GRACE retrievals using a data assimilation technique [1,3,14,68]. Through an assimilation framework, GRACE-based TWS can be effectively downscaled to finer spatial and temporal scales, as well as disaggregated into its individual components such as snow and soil moisture by leveraging the information content from a hydrologic (land surface) model. Meanwhile, the accuracy of modeled land surface states can be improved, and model uncertainty can be simultaneously reduced by merging GRACE TWS retrievals into the model.

Zaitchik et al. [68] was the first published study to assimilate GRACE TWS retrievals into the NASA Catchment Land Surface Model using an ensemble-based smoother. The study was conducted over the Mississippi River basin. The land surface model integration was conducted twice over the course of the same month (i.e., requires rewinding the land surface model): the first time to compute the differences between the observed and modeled TWS over the course of an entire

month, and the second time to update TWS for that month. DA-based estimates of shallow groundwater suggested improved agreement with *in situ* groundwater measurements in comparison to the case of using the model without assimilation (a.k.a., Open Loop). That is, DA provided reduced errors and increased temporal correlations relative to the Open Loop (OL). As the TWS change in the Mississippi River basin is little influenced by snow, snowpack variables were simply updated without direct calculation of SWE increments from the ensemble statistics. Forman et al. [1] used a similar DA approach within a snow-dominated region in northwestern Canada in order to explore the impact of GRACE TWS assimilation on regional snowpack characterization. Model estimates from OL and DA simulations were compared against independent SWE measurements, and a modest improvement in SWE statistics was found for the GRACE DA simulations. The same method was also applied across North America [14] and western and central Europe [65] for drought monitoring. In these studies, GRACE TWS retrievals were averaged over entire watersheds and a spatially- and temporally-uniform observation error was assumed. A later study by Forman et al. [107] evaluated the impact of the spatial scale of the GRACE TWS retrievals on DA performance, and provided some guidance on the treatment of spatial error correlations in those TWS retrievals. Results demonstrated superior performance for TWS anomaly assimilation at sub-basin scales compared to basin scales. It was also recommended that one assimilates TWS retrievals at the smallest spatial scale that the observations errors can be considered as uncorrelated, which for the spherical harmonics TWS retrievals is generally $\geq$150,000 km^2.

There are also a number of studies investigating the merger of GRACE TWS retrievals into land surface models using an ensemble Kalman Filter (EnKF), which applies increments at the end of the assimilation window without the need for rewinding the land surface model as in the smoother method [108–110]. Su et al. [108] studied the impact of GRACE DA on SWE and snow depth estimates over North America. Results revealed that in the basins where snowfall is a large contributor to the water cycle, the inclusion of GRACE-based TWS anomalies into DA can significantly enhance the estimates of SWE. However, in some high latitude regions, DA did not improve, or mitigate, the accuracy of the estimated SWE. Eicker et al. [109] successfully demonstrated the potential of including GRACE TWS as a gridded product ($5^\circ \times 5^\circ$) instead of using basin or sub-basin average into the data assimilation framework while explicitly computing the full error covariance of GRACE TWS used over the Mississippi River basin. Moreover, the proposed method allows one to simultaneously calibrate model parameters and perform state updates. Tangdamrongsub et al. [110] added to the previous studies by investigating the performance of GRACE TWS assimilation when a land surface model has significant uncertainty due to unreliable forcing data or poor model parameters. It was shown that GRACE assimilation is able to correct model errors associated with forcing data and parameter calibration issues by drawing the estimated TWS toward the observed TWS. Based on the method proposed by Eicker et al. [109], Schumacher et al. [111] systematically evaluated the impact of GRACE error correlation structure on GRACE assimilation. Results showed that the spatial error correlation used for GRACE TWS retrievals has a substantial influence on the estimates of model states

and model parameters. However, implementing a full GRACE error covariance matrix does not always improve assimilation estimates, and thus, the specification of the GRACE error correlation structure remains critically important.

Recently, GRACE data assimilation using an EnKF was further explored using GRACE TWS retrievals on a finer $1^\circ \times 1^\circ$ grid [3,112]. Besides the gridded GRACE TWS product and the spatially-distributed GRACE error used, Kumar et al. [112] also examined the impact of GRACE scaling factors used to recover (enhance) signal amplitude that was damped during GRACE data pre-processing (prior to assimilation). Girotto et al. [3] investigated the potential to improve soil moisture and groundwater estimates using a GRACE assimilation framework by introducing a new scheme to compute and apply the analysis increments (i.e., the amount of update added to model prognostic variables) that is less sensitive to the specific conditions on a single day within a single month. Results clearly showed improvement in the estimates of shallow groundwater while the influence on root-zone and surface layer soil moisture is relatively small. Girotto et al. [113] further investigated the benefits and drawbacks of GRACE DA by conducting an assimilation experiment over India. The capability of GRACE DA in overcoming model errors caused by a lack of representation of groundwater extraction and irrigation was specifically examined. Results showed that GRACE DA provided an improved correlation coefficient between estimated and *in situ* measured groundwater compared to the OL simulation. However, due to the lack of deep aquifer hydrological and irrigation processes in model, the assimilation unrealistically reduced evapotranspiration, which suggested the necessity of adding anthropogenic related processes (e.g., groundwater pumping

applied in the context of agricultural irrigation) into the model. The merging of GRACE TWS retrievals into different hydrologic models for different application purposes has been increasingly explored across different parts of the globe such as groundwater depletion in Iran [114] and surface soil moisture estimation in western Africa [115]. Similarly, Khaki et al. [116] assessed the performance of different data assimilation techniques on merging GRACE data into a hydrologic model over Australia. It was concluded that there was no single technique that can fit every application and every model in an optimal manner. Therefore, it is crucial to find a filter that is robust to the system error covariance as well as one that is able to generate representative ensembles that serve as a reasonable, low-rank approximation of the true probability distributions [116].

Chapter 3: Comparison of Vertical Surface Deformation Estimates Derived from GRACE, Ground-based GPS, and the NASA Catchment Land Surface Model over Snow-dominated Watersheds in the Western United States

3.1 Motivation and Objectives

Ground-based GPS, GRACE, and land surface models all have their advantages and pitfalls in representing TWS variations in accordance with their uncertainties and spatio-temporal characteristics. In order to provide a more accurate estimate of TWS anomaly (i.e., after removing the long-term mean; term is used interchangeably with "TWS changes" hereafter) by combining these data sets through a Bayesian merger, it is important to first carefully analyze the behaviors and error characteristics of each data set. For ground-based GPS, it is critically important to remove the non-hydrologic loading effects (e.g., tectonic motions, atmospheric pressure, and non-tidal ocean loading) embedded in the observations. Therefore, a comparison of displacements derived from GPS, GRACE, and the NASA Catchment Land Surface Model (Catchment) was conducted to analyze the TWS variation characteristics reflected by the different data sets, as well as demonstrating that the

non-hydrologic loadings have been properly removed from the ground-based GPS observations. Considering the relatively large amplitude of the vertical displacement observations relative to the horizontal displacements [24, 79], vertical displacement was made the focus of comparison and discussion in this study. Previous studies compared vertical displacements derived from GPS and GRACE, and found that ground-based GPS vertical displacements can be used to represent TWS changes over the western United States. However, most studies focused on California where a very dense ground-based GPS network is available [71,88,89]. Relatively few studies have been conducted over regions of complex terrain (i.e., mountains) where snow is a significant source of water supply [117, 118] and only moderate ground-based GPS networks are available. Therefore, two snow-dominated regions in the western United States – the Great Basin and Upper Colorado watersheds – are studied in this work.

Data sources as well as the processing procedure of each data set are introduced in Section 3.2. The comparison of seasonal and seasonally-adjusted changes in vertical displacements derived from different data sources are provided in Sections 3.3.1 and 3.3.2. The connection between hydrologic processes and vertical displacement is discussed in Section 3.3.3. The influence of atmospheric circulation, particularly the La Niña event in 2010-2011, on vertical displacement of the land surface is discussed in Section 3.3.4. Section 3.3.5 summarizes the possible explanations for discrepancies between vertical displacements derived from GPS, GRACE, and the LSM as well as the error sources in each data set.

3.2 Study Area and Methodology

3.2.1 Study Area

The study was conducted in the western United States and includes the Great Basin and Upper Colorado watersheds as shown in Figure 3.1. The watershed shapefiles used in Figure 3.1 were acquired from the Watershed Boundary Dataset provided by United States Geological Survey (USGS; https://water.usgs.gov/GIS/huc.html). The Great Basin watershed covers a total area of about 367,000 km^2 and includes almost all of Nevada, a large portion of Utah, and a small region of California, Oregon, Idaho, and Wyoming. The Great Basin watershed is bounded by the Sierra Nevadas in the west and the Wasatch Mountain range in the east, and the less distinct ridges of the Snake River Plain in the north [119]. The Upper Colorado watershed is to the east of the Great Basin watershed and covers a large area of Utah and Colorado and portions of Wyoming, New Mexico, and Arizona with a basin area of approximately 293,000 km^2. The Upper Colorado watershed is bounded by the Rocky Mountains in the east, Wasatch Mountain Range in the west, and drains to Lees Ferry in Arizona [118]. The two watersheds generally have warm summers and cold winters with snow being a critical reservoir of fresh water [120].

3.2.2 GPS Data

The ground-based GPS data used in this study is the Level 2 daily position data from the Plate Boundary Observatory (PBO; https://www.unavco.org/data/data.html)

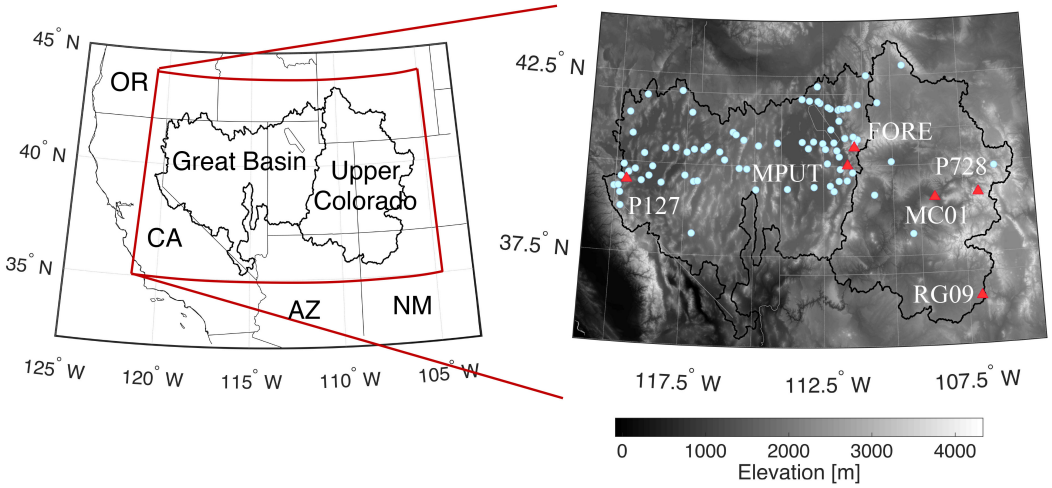

Figure 3.1: Study area including the Great Basin and Upper Colorado watersheds along with the GPS permanent stations shown as light blue dots. Three representative stations in each of the two watersheds used for discussion in Section 3.3.1 are shown as red upper triangles along with their four letter codes.

project aimed at quantifying Earth′s three-dimensional deformation across the boundary between the Pacific and North American plates [86]. Data was archived by the Geodesy Advancing Geosciences and EarthScope (GAGE) Facility at UNAVCO. Data processed by the New Mexico Tech analysis center was used instead of the combined data solution from the two data analysis centers because a network-wide vertical shift in December 2011 was observed in the GPS data processed by the Central Washington University analysis center [96], but not in data processed by the New Mexico Tech analysis. In order to ensure that the vertical displacements observed in the GPS record reflects only variations in hydrologic loading, and thus, can be compatible with GRACE-derived vertical displacements, all non-hydrologic loading effects must first be removed.

In this study, GPS stations providing records during the study period from January 2003 to March 2016 in the study area were first found. Afterwards, only

stations with continuous displacement records (i.e., without an offset event) that cover at least 50% of the study period were selected for use. The offset in displacement is a sudden step-like shift of displacement associated with an earthquake, changing of antenna, or other miscellaneous reasons. The occurrence of offsets was provided by UNAVCO [86]. Earth vertical displacement caused by non-hydrologic variations such as glacial isostatic adjustment (GIA) and tectonic effects can, at times, appear in the GPS record of vertical displacement. The ground-based GPS data provided by the New Mexico Tech analysis center included the removal of some of the non-hydrologic loading (e.g., solid earth tide), this study continued removing the remaining non-hydrological loadings effects that are not accounted for in the Level 2 GPS observations of vertical displacement as follows. For example, based on previous studies [2,6,96], a long-term linear trend can be used to remove the secular trend effect (e.g., plate motion) without an evident impact on the analysis of seasonal and inter-annual change. In addition, the GRACE retrieval used in this study accounted for the effects of atmosphere and non-tidal ocean using the atmosphere and ocean de-aliasing (AOD) model. Therefore, the AOD model was similarly used to remove atmosphere and non-tidal ocean effects from the ground-based GPS observations using the monthly-averaged AOD Level-1B Release 05 (GAC solution with degree and order 100) from the German Research Centre for Geosciences (GFZ; https://www.gfz-potsdam.de/en/aod1b/) [121]. Due to the strong sensitivity of GPS observations to crustal movement, GPS observations larger than three times the standard deviation of the time series (relative to the mean) were considered as outliers and removed from the record. Afterwards, the processed daily GPS vertical

displacements were converted into monthly-averages in accordance with the dates used in the corresponding GRACE solutions. The processed, monthly GPS vertical displacements were then used to compare against GRACE and the hydrologic model (more details in Section 3.3.1). Figure 3.2 illustrates the general processing stream of the ground-based GPS observations. One factor that has considerable influence on the observed vertical displacement is groundwater pumping with more details discussed in section 3.3.5. However, according to Konikow et al. [122], most of the Great Basin and Upper Colorado watersheds do not have significant groundwater pumping issues, and therefore, the groundwater pumping effect is not explicitly considered in this study.

In the Great Basin, there are 109 GPS stations with vertical displacement records between January 2003 and March 2016. After removing stations with relatively short time spans (i.e., records less than half of the period from January 2003 to March 2016), a total of 77 GPS stations remained that were then used in the study. For the Upper Colorado region, 14 GPS stations are available, and after completing quality control, a total of 10 GPS stations remained. The locations of the GPS stations used in the study are shown in Figure 3.1.

3.2.3 GRACE Data

GRACE Level 2 monthly TWS mass concentration (mascon) solutions from NASA′s Goddard Space Flight Center (GSFC; https://neptune.gsfc.nasa.gov/grace/) [123] were used in this study. The rationale for utilizing the GRACE mascons in this

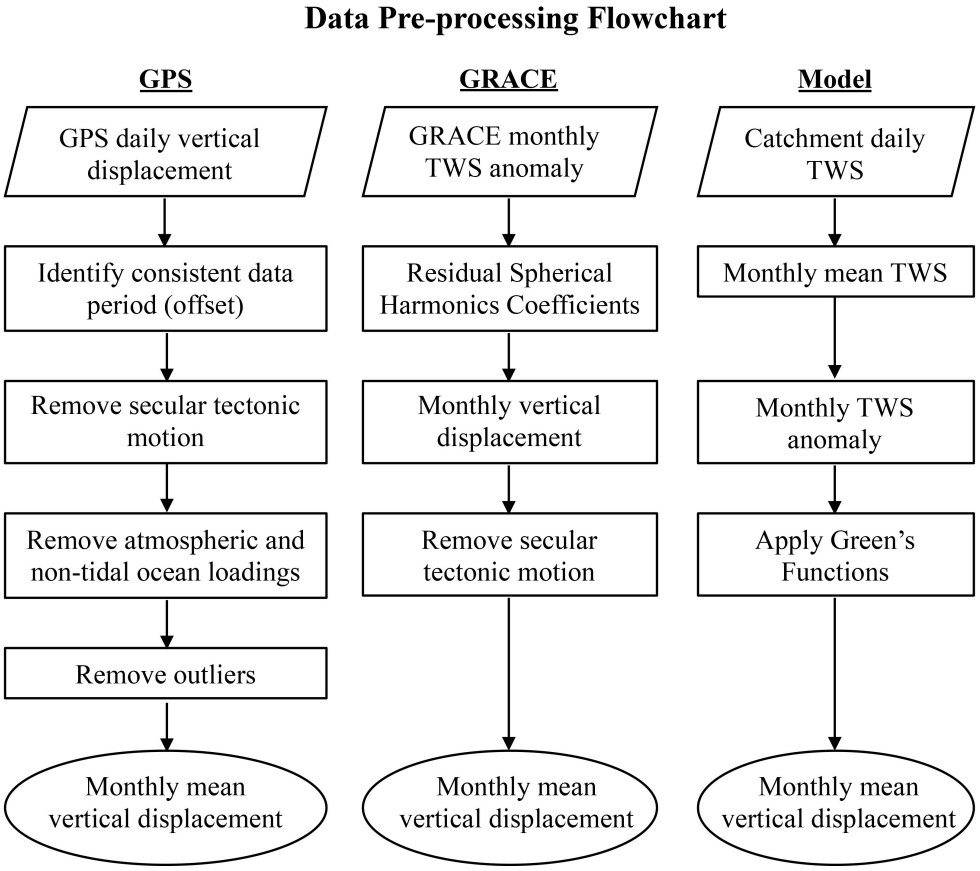

Figure 3.2: Flowchart of GPS, GRACE, and NASA′s Catchment model data pre-processing.

study is the ability to optimally reduce errors by using the data noise covariance [124] when inverting for the gravity parameters, as opposed to estimating unconstrained spherical harmonic estimates that then need to be filtered, where the filter tends to indiscriminately reduce both noise and signal. Therefore, the mascon technique provides a more optimal solution to balance the trade-offs between noise reduction and signal attenuation compared to the spherical harmonics (SH) solutions. The GSFC mascon product provides monthly estimates of time-variable gravity in the form of equivalent water height (EWH) for a global set of $1^\circ \times 1^\circ$ mascons. The monthly mascon solution for the period from January 2003 to March 2016 (without GIA removed) was used, and it was converted into normalized spherical hamonics up to degree 60 to prepare for the computation of vertical displacement. The geocenter correction (i.e., replacing spherical harmonic coefficients of degree 1) has been applied in the GSFC mascon product [12], which allows for the comparison of vertical displacement derived from GPS and GRACE in the center of figure (CF) frame. GRACE C20 terms (spherical harmonic coefficients of degree 2 and order 0) have also been replaced using results from the Satellite Laser Ranging measurements in the mascon product in order to obtain a higher accuracy in the C20 terms [125]. Assuming a homogeneous reaction of the Earth′s crust to mass loading changes, the induced vertical displacement was represented [23, 94] using spherical harmonic coefficients for the gravity field and load Love numbers via:

$$dr(\theta,\phi) = R\sum_{l=1}^{\infty}\sum_{m=0}^{l}\widetilde{P}_{lm}(cos\theta)(\Delta\widetilde{C}_{lm}\cos m\phi + \Delta\widetilde{S}_{lm}\cos m\phi)\frac{h_l^{'}}{1+k_l^{'}} \tag{3.1}$$

where $dr(\theta, \phi)$ is the vertical displacement in the radial direction; θ and ϕ are co-latitude (i.e., the difference between 90°) and the latitude) and east longitude, respectively; R is the average Earth radius; $\widetilde{P}_{lm}$ is the fully normalized Legendre functions of degree l and order m; $\Delta\widetilde{C}_{lm}$ and $\Delta\widetilde{S}_{lm}$ are the spherical harmonic coefficients of the time-variable Earth′s gravity field relative to a long-term average derived from GRACE; and $h_l^{'}$ and $k_l^{'}$ are the elastic load Love numbers (i.e., dimensionless parameters that characterize the elastic response of the Earth to surface mass loading). After converting the GRACE TWS retrieval into vertical displacements at each GPS station location, a linear trend was also removed from each displacement time series in the same way as for GPS data in order to remove any remaining secular trend artifacts.

3.2.4 Land Surface Model

The NASA Catchment Land Surface Model [54, 67], a process-simulating numerical model, was used to calculate the redistribution of hydrological mass loadings that was then used to compute vertical displacement. Traditional LSMs discretize soils into vertical layers, but the horizontal structure of soil is assumed to be uniform spanning over tens of kilometers [67]. This assumption neglects the impact of horizontal variability of soil moisture and its effects on evapotranspiration and runoff. Catchment improves upon traditional LSMs by importing an explicit treatment of subgrid soil moisture variability, and thus, explicit modeling of evaportranspiration and runoff mechanisms [54, 67]. Hydrological catchments were used as the

fundamental land surface unit instead of a traditional grid in order to account for subgrid heterogeneity. Modeled daily TWS was projected to the Equal-Area Scalable Earth (EASE) grid with a spatial resolution of 25 km $\times$ 25 km. Catchment was run on the University of Maryland supercomputing cluster (Deepthought2) using meteorological fields provided by the Modern Era Retrospective Analysis for Research Application version 2 product (MERRA-2) [126] as boundary conditions. Catchment-derived TWS was converted into vertical displacement using Green$'$s function [70, 71] based on the Preliminary Reference Earth Model (PREM) [82]. Assuming a disc loading with angular radius α (i.e., half of the angular diameter, a angular measurement of how large the disc appears from the center of the Earth), the vertical displacement with respect to the distance to the center of the disc can be calculated as:

$$\begin{aligned} dr &= \sum_{l=0}^{\infty} h_l' \Gamma_l \frac{4\pi G R}{g(2l+1)} P_l(\cos\lambda) \\ \Gamma_l &= \frac{1}{2}[P_{l-1}(\cos\alpha) - P_{l+1}(\cos\alpha)] \qquad l > 0 \\ \Gamma_0 &= \frac{1}{2}(1-\cos\alpha) \end{aligned} \tag{3.2}$$

where P_l is the Legendre polynomials for degree $l \in [0, \infty]$, G is Newton$'$s gravitational constant, g is the gravitational acceleration at the Earth$'$s surface, and λ is the angular distance of the observation location to the center of the disc loading. In this study, $h_1^{'}$ = -0.269 was used corresponding to the CF frame [94].

The spatial resolution of 25 km $\times$ 25 km at which Catchment was run is roughly the same area as a disc with a radius of 14 km. The vertical displacement in elastic

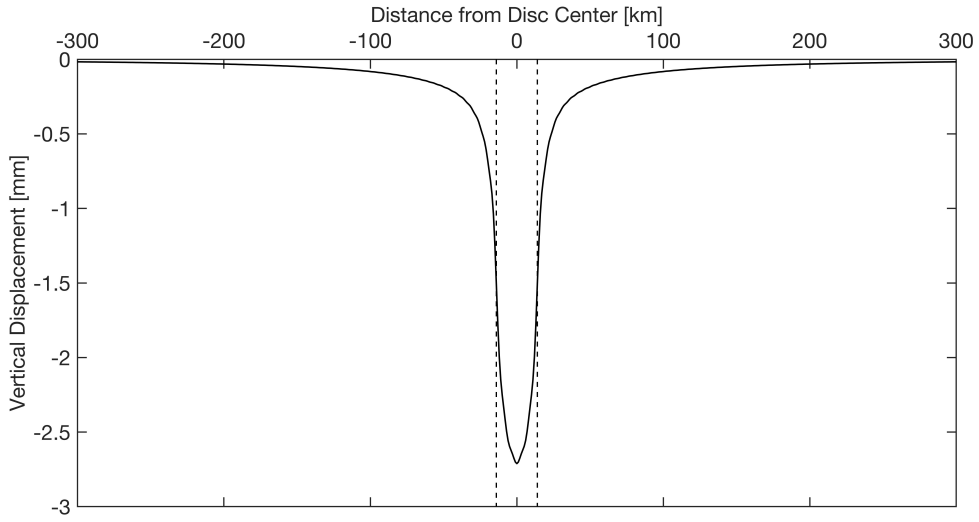

Figure 3.3: Vertical displacements caused by adding a uniform disc load with radius of 14 km and equivalent water height (EWH) of 1m. The vertical dashed line represents the edge of the disc.

response to adding a uniform disc load with radius 14 km and equivalent water height (EWH) of 1 m is shown in Figure 3.3. The negative vertical displacement value represents a downward motion of the surface. Crustal deformation caused by the disc load is most significant at the center with vertical displacement around 2.7 mm and then decreases rapidly away from the center. At the edge of the disc, the resulting vertical displacement change is about 1.5 mm, and when the distance from the disc center becomes 400 km, the influence becomes less than 0.01 mm. Based on the vertical displacement change versus distance relationship, the Catchment-based vertical displacement at each GPS station within the study domain was computed by summing up the impact of TWS change of each 25 km $\times$ 25 km pixel on land onto this GPS stations. The Catchment-derived daily vertical displacements were then converted to monthly vertical displacements using the same time span as used in the GRACE solutions.

3.2.5 SNOTEL Data

Ground-based measurements of SWE provided by the SNOwpack TELemetry (SNOTEL; https://www.wcc.nrcs.usda.gov/snow/) network were used to analyze the relationship between the hydrologic loading and surface deformation in Section 3.3.3. SNOTEL is operated by the Natural Resources Conservation Service (NRCS) National Water and Climate Center, with most stations installed in high mountain areas where access is difficult or restricted. Stations located inside the study area providing SWE records less than half of the study period were excluded from the analysis, and 105 and 106 stations were remained for the Great Basin and Upper Colorado watersheds, respectively, as shown in Figure 3.11. The daily measurements of SWE were averaged to monthly SWE according to the GRACE retrieval dates.

3.2.6 Evaluation Metrics

In order to analyze the agreements between different vertical displacement time series derived from ground-based GPS, space-based GRACE, and model-based Catchment, the correlation coefficients (R) of vertical displacement of (1) GPS versus GRACE; (2) GPS versus Catchment; and (3) GRACE versus Catchment were computed. The same computation was repeated after removing the mean seasonal cycle from each data set, respectively, in order to evaluate the agreement between the inter-annual variations of the three different data sets. Note that the correlation coefficient can only reflect the agreement of two time series in terms of variations in phase, whereas the differences in amplitudes are not assessed. Therefore, the

weighted root mean square (RMS) reduction as defined in equation 3.3 [23], which takes both amplitude and phase differences into consideration, was also used. RMS reduction computes the variance reduction of subtracting one time series from another and thus can be used to analyze the seasonal agreement between the vertical displacements derived from two data sets (e.g., GPS and GRACE). A large positive RMS reduction value represents a good agreement between two time series in amplitude and phase.

$$RMS_{reduction} = \frac{RMS_A - RMS_{A-B}}{RMS_A} \times 100\% \tag{3.3}$$

where RMS_A is the variance of data set A (i.e., GPS), RMS_{A-B} is the variance of the time series after subtracting data set B from A where data set B is either GRACE- or Catchment-based monthly time series in this study.

3.3 Results and Discussion

3.3.1 Monthly Vertical Displacement Analysis

Monthly vertical displacements derived from GRACE, GPS, and Catchment in the Great Basin and Upper Colorado watersheds were acquired through the procedure outlined in Section 3.2. For each watershed, three different GPS stations corresponding to the maximum, median, and minimum correlation coefficient between GRACE- and GPS-based vertical displacements were selected as representative examples as shown in Figure 3.4. Selected station information and statistical values

are summarized in Table 3.1. The six representative stations shown in Figure 3.4 all display the expected seasonality in vertical displacement time series, with negative values found during the winter due to snow accumulation, and positive values in the summer due to snow ablation. For the Great Basin, the lowest R between GRACE- and GPS-derived vertical displacement is found at station MPUT with a value of 0.59, whereas the highest R is found at station FORE with a value of 0.89. For the Upper Colorado watershed, the corresponding lowest and highest R values are 0.48 at station RG09 and 0.86 at station P728, respectively. The exceptionally small R values related to GPS-based vertical displacement at station RG09 is notable. From the time series of GPS vertical displacement at station RG09 in Figure 3.4, an obvious decrease of annual peak displacement from 2003 to early 2011 followed by an increase of upward deformation until 2016 can be observed. However, this change is not captured by GRACE or Catchment.

Time series amplitudes estimated from GPS are consistently larger than for GRACE or Catchment. For all 77 GPS stations located in the Great Basin watershed, an average annual response amplitude of 4.7±2.6 mm (shown as mean ± standard deviation) was derived from the GPS-based vertical displacement records, which is on average 2.4 times larger than the amplitude from GRACE (1.9±0.8 mm). The amplitude of vertical displacements estimated from Catchment (2.6±0.9 mm) is between the scale of GPS and GRACE. Similar amplitude characteristics are found for the stations in the Upper Colorado watershed of which the corresponding amplitudes are 4.3±2.3 mm, 1.7±0.6 mm, and 2.1±1.0 mm for GPS, GRACE, and Catchment, respectively. The phase differences between the vertical displacement

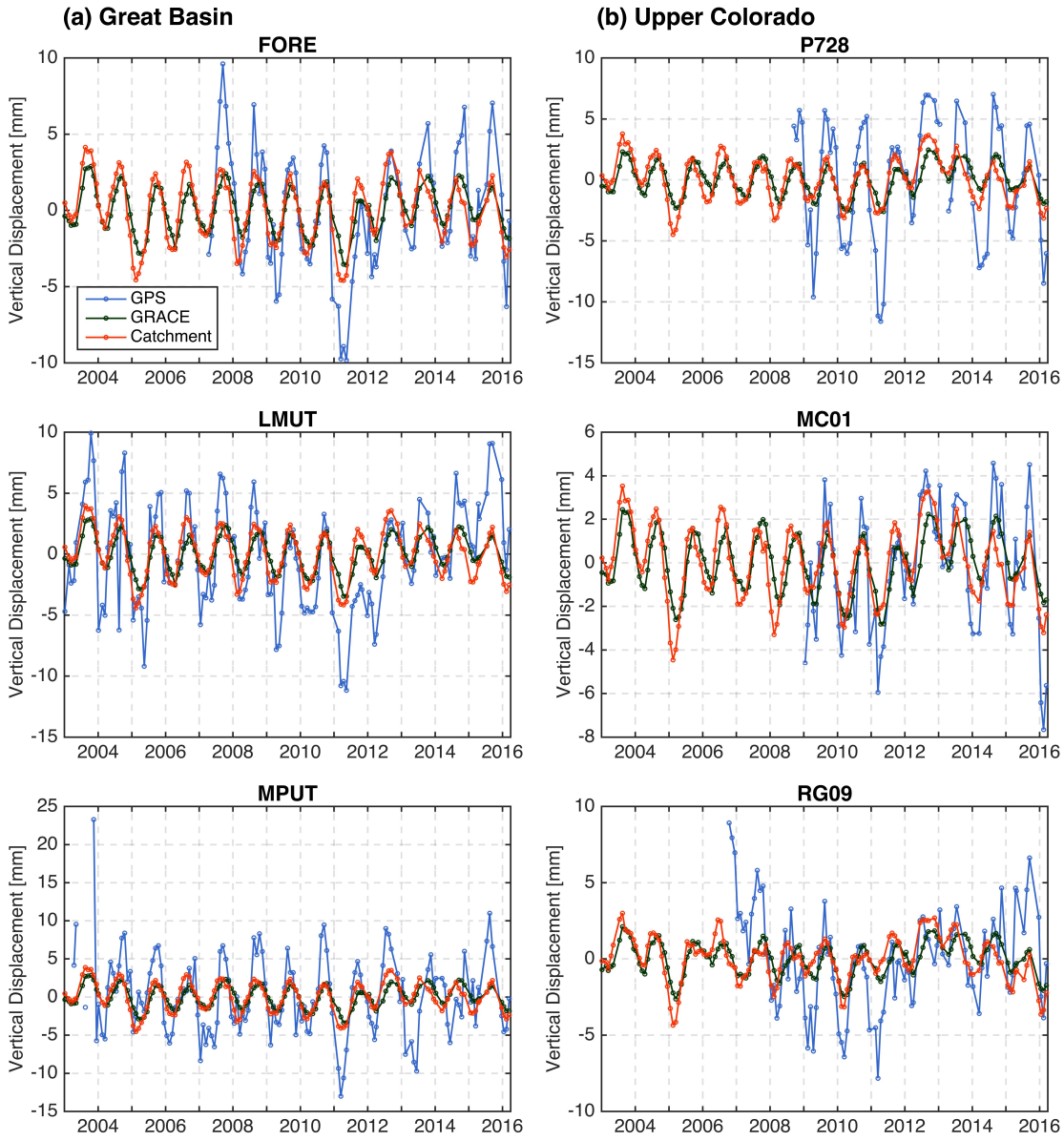

Figure 3.4: Comparison of monthly-averaged vertical displacements derived from ground-based GPS (blue), GRACE TWS retrievals (dark green), and Catchment estimates (red) for GPS station locations in column (a) the Great Basin watershed and column (b) the Upper Colorado watershed.

Table 3.1: Correlation coefficient, R, of vertical displacement time series derived from (1) GPS versus GRACE, (2) GPS versus Catchment, and (3) GRACE versus Catchment, and RMS reduction after removing GRACE and Catchment from the GPS signal, respectively. The selected stations correspond to the minimum (MPUT and RG09), median (LMUT and MC01), and maximum (FORE and P728) correlation coefficient between GPS- and GRACE-derived vertical displacement in the Great Basin and Upper Colorado watersheds.

	Station	Location	R (unitless)			RMS reduction (%)	
			GPS vs. GRACE	GPS vs. Catchment	GRACE vs. Catchment	GPS vs. GRACE	GPS vs. Catchment
Great Basin	FORE	40.51°N 111.38°W	0.89	0.80	0.86	31	33
	LMUT	40.26°N 111.93°W	0.77	0.70	0.86	23	27
	MPUT	40.02°N 111.63°W	0.59	0.61	0.86	13	17
Upper Colorado	P728	39.18°N 106.97°W	0.86	0.84	0.81	22	27
	MC01	39.09°N 108.53°W	0.75	0.83	0.79	29	38
	RG09	36.30°N 107.06°W	0.48	0.35	0.77	11	6

time series derived from three data sets were also investigated. The timing of the maximum vertical displacement (i.e., less water loading) and minimum vertical displacement (i.e., more water loading) varies by year and station location. For each station, if over 50% of the yearly peak or yearly trough values are found in the same month, the station is defined to experience a maximum or minimum, respectively, in that month. For GRACE-based vertical displacement, all examined station locations experienced maximum displacement during September or October, and minimum displacement during March or April. For Catchment, the peaks are found in August or September, and the troughs are in February or March for all station locations, which are both one month earlier than for the GRACE-based time series. The one month earlier peak found in the Catchment-based vertical displacement could be attributed to a lack of a dynamic surface water routing module, differences in snow melt processes (e.g. Catchment snow often melts earlier than snow-covered area retrievals suggest) [127], precipitation errors in the boundary conditions, or model structure errors related to other physical processes in Catchment [128]. A combination of these errors likely results in surface water exiting the basin earlier than in reality. As for GPS, the maxima and minima timings show slightly greater variability than GRACE or Catchment, with peaks ranging from August to October, and troughs ranging from February to April. It is difficult to say with certainty, but the larger spread in the GPS maxima and minima is likely due to the enhanced spatiotemporal resolution of the GPS observations.

The correlation coefficient maps of: (1) GPS versus GRACE, (2) GPS versus Catchment, and (3) GRACE versus Catchment including all examined GPS stations

in the two watersheds are shown in Figure 3.5. The average correlation coefficient between GPS and GRACE pairs is 0.77 for the Great Basin and 0.73 for the Upper Colorado watershed. Over 89% of all the examined GPS stations in the study area (i.e., in both Great Basin and Upper Colorado watersheds) provide an R value greater than 0.70. The correlation coefficient between GPS and Catchment is comparable but slightly lower as compared to GPS versus GRACE, with the averaged values of 0.76 and 0.72 in the Great Basin and Upper Colorado watersheds, respectively. The percentage of stations that provide a R value larger than 0.70 is 80%. The consistency between GRACE and Catchment is generally the highest among the three comparison cases for both the Great Basin (R = 0.87) and Upper Colorado (R = 0.81) watersheds. All stations located in the study area yield R values larger than 0.75. In all three comparison cases, a consistently lower agreement was found in the Upper Colorado watershed, which is due to the limitation of Catchment in modeling dynamic surface water routing of the Colorado river [3]. Most stations with large R values are found in the north of the Great Basin watershed located in the Bear and Great Salt Lake sub-basins that experience longer snow seasons than in areas at lower elevations and closer to the basin outlet.

The RMS reduction as introduced in Section 2.5 is another metric used to indicate the agreement between two data sets through the subtraction of one data set time series from the other. Removing the GRACE and Catchment time series, respectively, from the GPS signal for each station leads to the computation of the RMS reduction (Figure 3.6). A positive RMS reduction suggests a decrease in variance, and the larger the RMS reduction value, the better the agreement. For

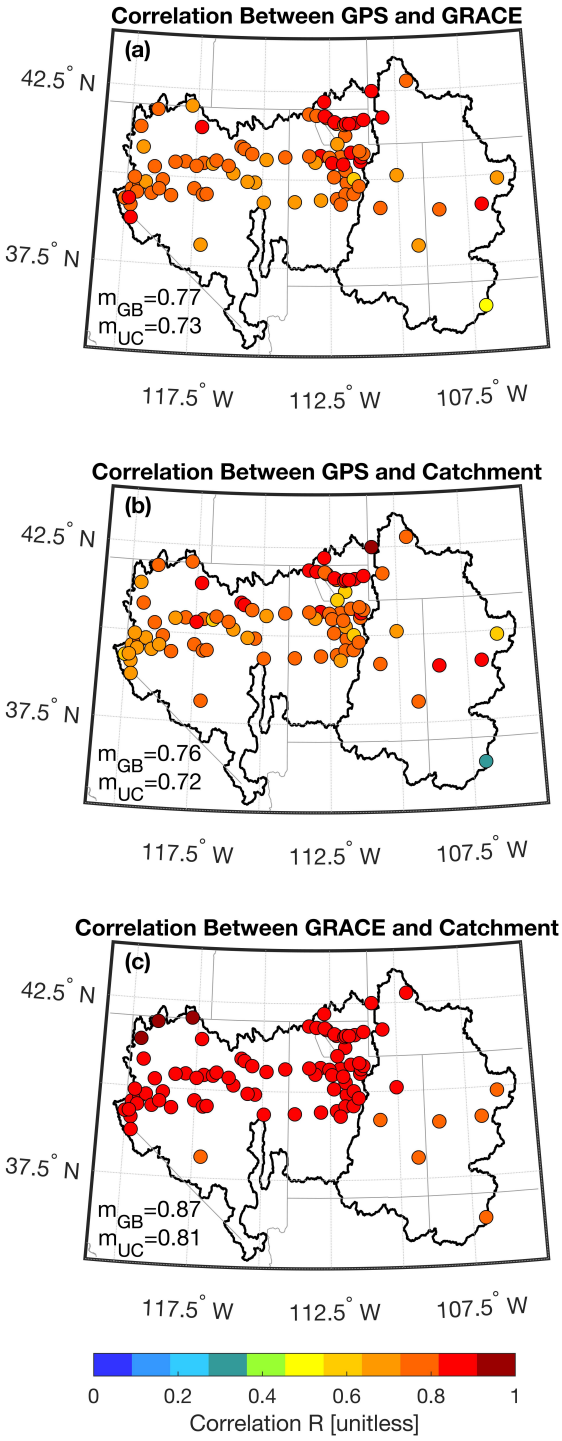

Figure 3.5: Correlation coefficient maps of vertical displacements derived from (a) GPS versus GRACE; (b) GPS versus Catchment; and (c) GRACE versus Catchment for all examined GPS stations in the Great Basin and Upper Colorado watersheds. m = watershed-averaged R; GB = Great Basin; and UC = Upper Colorado.

all of the examined stations in the two watersheds, there are no negative RMS reduction values found for either the GRACE or Catchment cases, which indicates a general agreement between the GPS, GRACE, and Catchment datasets. Removing Catchment from the GPS signal (Figure 3.6b) shows relatively larger RMS reduction values in comparison with the case of removing GRACE (Figure 3.6a) from the GPS signal. This behavior is slightly different from the results of the correlation analysis that suggested better consistency between GPS versus GRACE as compared to GPS versus Catchment. This difference is attributable to the fact that the correlation coefficient R considers only the seasonal phase consistency between two data sets, while RMS reduction takes both seasonal phase and amplitude into consideration. When the amplitude difference between two different data sets is significant, RMS reduction will be small even when the seasonal phase is identical. The relatively larger R values and smaller RMS reduction between GPS and GRACE (as compared to GPS versus Catchment) indicates that GPS and GRACE are in good agreement in terms of seasonal variation, whereas the amplitude differences cannot be neglected.

Following Knappe et al. [90], the network-wide mean vertical displacement time series derived from ground-based GPS, GRACE, and Catchment using all stations were calculated as to represent the regional signal as illustrated in Figure 3.7. In the Great Basin watershed, the network mean time series derived from ground-based GPS shows a relatively consistent variation before 2011 and then an abrupt step that can be observed in early 2011 that is then followed by an increase of upward deformation after 2011. In the Upper Colorado watershed, a similarly abrupt step in the 2010-2011 winter can be detected from the GPS time series, but the following

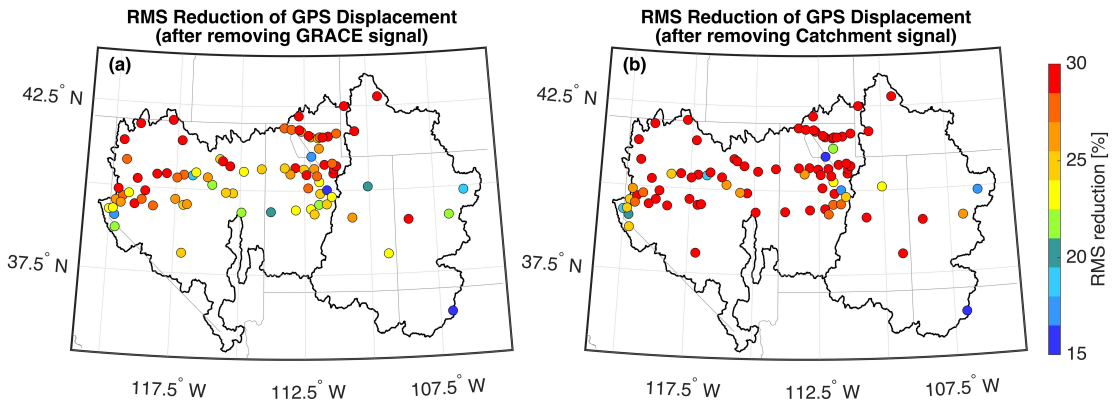

Figure 3.6: Maps of RMS reduction of vertical displacements derived from (a) removing GRACE time series from GPS time series; (b) removing Catchment time series from GPS time series. A large RMS reduction value represents a better consistency between two data sets.

rebound is not as apparent. The correlation coefficients of (1) GPS versus GRACE, (2) GPS versus Catchment, and (3) GRACE versus Catchment were computed using the network mean time series. The computed R values are 0.83, 0.83, and 0.88 in the Great Basin, and 0.81, 0.81 and 0.82 in the Upper Colorado Basin, respectively. A slightly higher consistency is found in the Great Basin than in the Upper Colorado watershed, which is consistent with the previous individual station analysis. As discussed earlier, an approximately one month phase lag was found between the GRACE and Catchment time series. Adjusting for the one month phase lag and re-computing the correlation coefficients between mean vertical displacement time series derived from GRACE and Catchment, the R values increased from 0.88 to 0.91 in the Great Basin, and from 0.82 to 0.89 in the Upper Colorado.

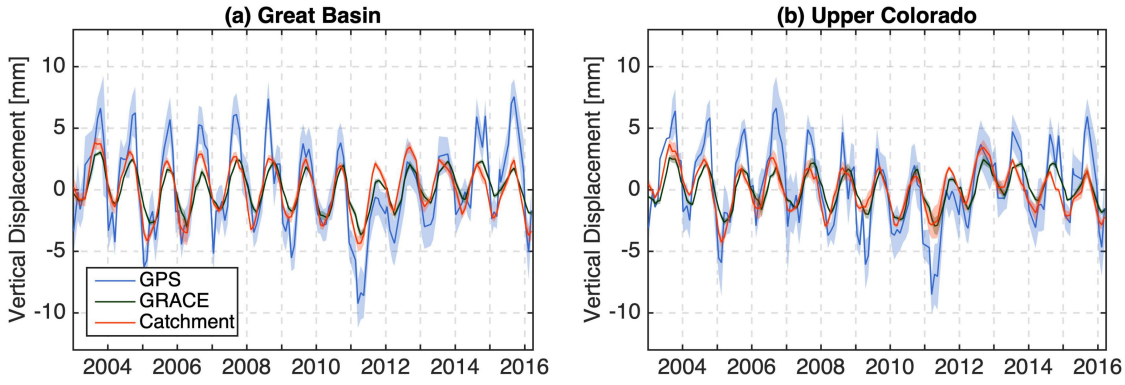

Figure 3.7: Network-wide mean vertical displacement time series derived from ground-based GPS (blue), GRACE (dark green), and Catchment (red) in the (a) Great Basin watershed and (b) Upper Colorado watershed. The mean±1 std (standard deviation) are shaded with corresponding transparent colors.

3.3.2 Seasonally-Adjusted Variation Analysis

The potential of using ground-based GPS vertical displacements to represent TWS inter-annual variations was investigated via the removal of the mean seasonal cycle (i.e., the multi-year average of each month) of vertical displacement embedded within the time series derived from GPS, GRACE, and Catchment. Figure 3.8(a)-(c) used GPS station LMUT located in the Great Basin watershed as an example to explain the process of removing the mean seasonal cycle. Again, the network-wide mean of seasonally-adjusted vertical displacement time series following Knappe et al. [90] were calculated for GPS, GRACE, and Catchment as shown in Figure 3.9. A positive, residual vertical displacement (i.e., after removing the mean seasonal cycle) represents a larger than multi-year averaged displacement for that particular month, which indicates less hydrologic loading than during expected climatology. On the contrary, a negative residual vertical displacement corresponds to more hydrologic loading for that particular month. For both the Great Basin and Upper

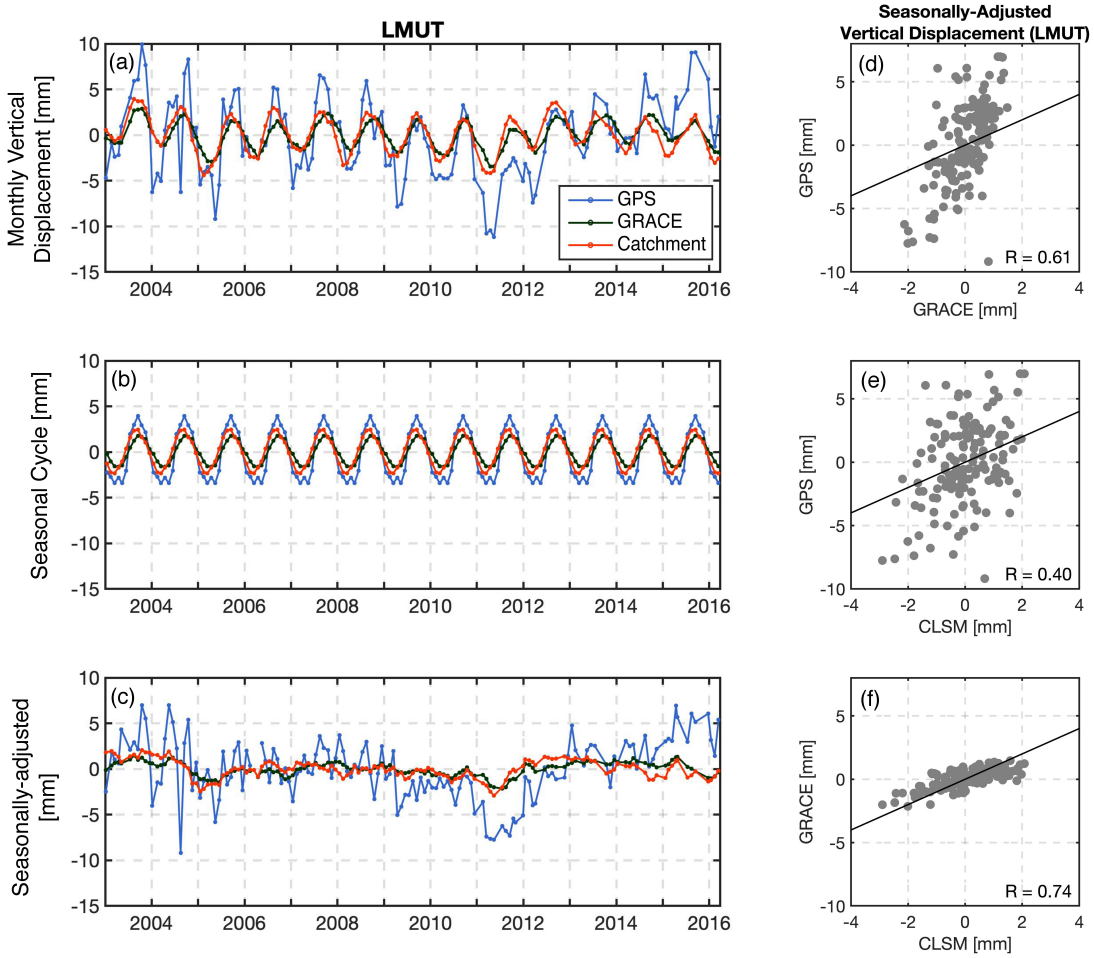

Figure 3.8: Example of removing mean seasonal cycle from processed monthly vertical displacement derived from ground-based GPS (blue), GRACE (dark green), and Catchment (red), of which (a) the processed monthly vertical displacement time series; (b) the mean seasonal cycle to be removed from (a); and (c) the seasonally-adjusted vertical displacement time series. The right column shows scatter plots of the seasonally-adjusted vertical displacements derived from (d) GPS versus GRACE; (e) GPS versus Catchment; and (f) GRACE versus Catchment at the example station. The correlation coefficient R values are listed in the lower right-hand corner, and the black lines in (d)-(f) represent the 1:1 line.

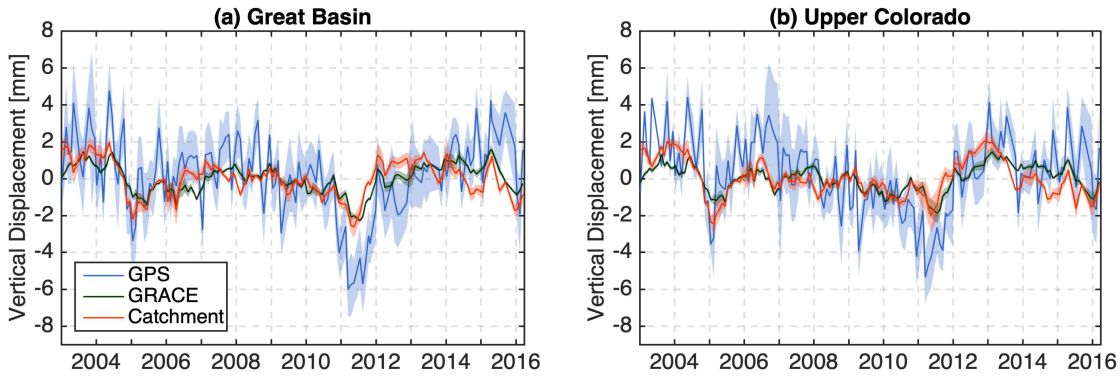

Figure 3.9: Network-wide mean seasonally-adjusted vertical displacement time series derived from ground-based GPS (blue), GRACE (dark green), and Catchment (red) in the (a) Great Basin watershed and (b) Upper Colorado watershed. The mean±1 std (standard deviation) are shaded with corresponding transparent colors.

Colorado watersheds, two distinguishable negative troughs were captured during the 2004-2005 winter and 2010-2011 winter from estimates derived from GPS, GRACE, and Catchment. The residual displacement before and after the two evident troughs (i.e., minimum, negative vertical displacements) are continuous, positive values indicating potential drought periods, which are coincident with the percent area of drought time series provided by the U.S. Drought Monitor [129] (data accessable from https://www.drought. gov/drought/states).

Considering the consistency of residual vertical displacements derived from GPS, GRACE, and Catchment, the seasonally-adjusted correlation coefficient was calculated. Figure 3.8(d)-(f) shows the scatter plots and R values between the seasonally-adjusted vertical displacements derived from GPS, GRACE, and Catchment at station LMUT. The same analysis was conducted for all examined stations and the R values were subsequently mapped as shown in Figure 3.10. In all three comparison cases, a relatively good consistency was found in the northeastern areas

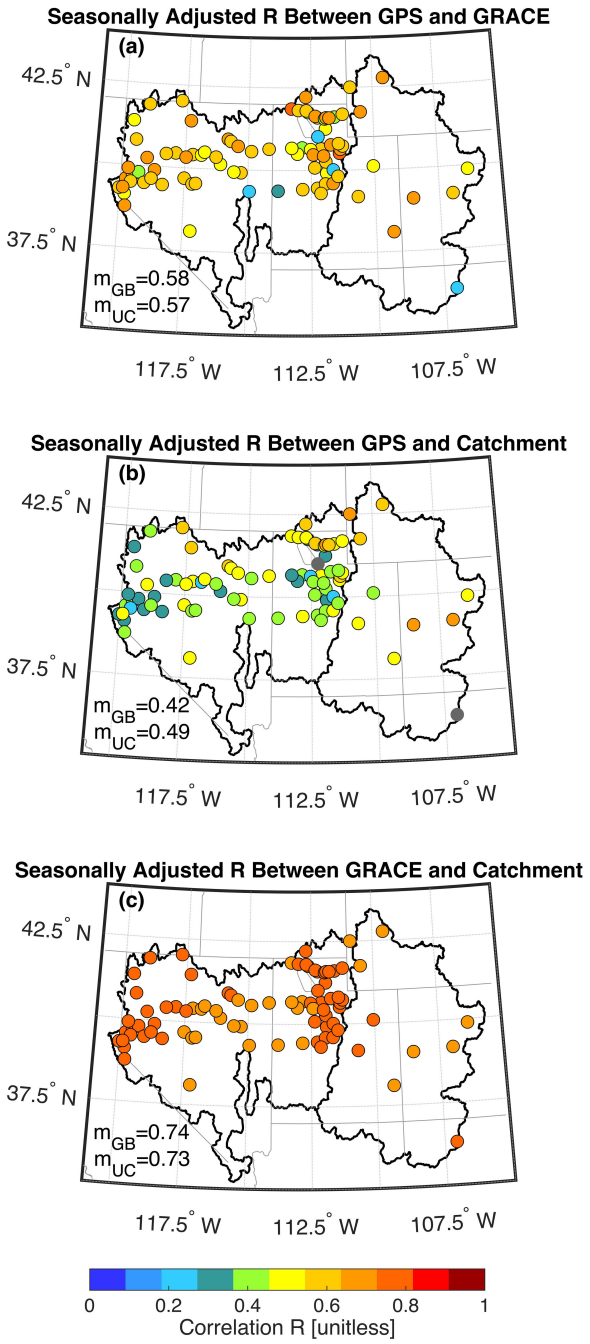

Figure 3.10: Correlation coefficients of seasonally-adjusted vertical displacement after removing the mean seasonal cycle for (a) GPS versus GRACE; (b) GPS versus Catchment; and (c) GRACE versus Catchment for all examined GPS station locations in the Great Basin and Upper Colorado watersheds. The gray dots in (b) represent stations that provides a seasonally-adjusted correlation coefficient not statistically different from zero. m = watershed-averaged seasonally-adjusted R; GB = Great Basin; and UC = Upper Colorado.

of the Great Basin watershed. The seasonally-adjusted R between GPS and Catchment generally provides the lowest agreement in representing the inter-annual variation of vertical displacement, and shows a clear southward decreased consistency. For the Great Basin watershed, the average seasonally-adjusted R between GPS and GRACE is 0.58, and the value for GPS versus Catchment is 0.42. The agreement between GRACE and Catchment is the highest (R = 0.74). The comparison results are similar in the Upper Colorado watershed with average seasonally-adjusted R values of 0.57, 0.49, and 0.73, respectively for the cases of (1) GPS versus GRACE, (2) GPS versus Catchment, and (3) GRACE versus Catchment.

3.3.3 Surface Deformation and Hydrologic Loadings

Processed GPS observations of vertical displacement (i.e., after removing non-hydrological loadings) can help reflect TWS changes across local and regional scales. Knappe et al. [90] found that seasonal precipitation patterns dominate the displacement observed by individual GPS stations. Therefore, in order to better to investigate the hydrologic cause of seasonal changes in the GPS time series, the most important hydrologic boundary condition – precipitation – is further explored here. Note that during the winter season in this region most of the mass (precipitation) falls as snow. Hence, precipitation runoff is minimal during the winter season. Further, most vegetation is dormant during this period, hence, evaportranspiration is minimal during the winter season. Therefore, careful consideration of precipitation (snowfall) is most important in the context of hydrologic loading (mass accumula-

tion) because there are relatively few outflows of water during the winter season.

SNOTEL-based SWE were used to examine the impacts of snowfall on the vertical displacement time series at both watershed- and sub-watershed scales. The annually-averaged maximum magnitude of SWE derived from SNOTEL at each station is shown in Figure 3.11(c) in order to illustrate which areas receive the most snow relative to other measured areas. Additionally, the annual amplitude of GPS vertical displacement and the annually-averaged maximum magnitude derived from Catchment SWE estimates are also plotted in Figure 3.11(c) as auxiliary information sources. Generally speaking, deep snow can be detected in the western and northeastern regions of the Great Basin, and the northern and eastern part of the Upper Colorado watershed, which are mostly mountainous. Meanwhile, a larger GPS amplitude was observed in these areas, especially in the western Great Basin watershed. For each watershed, the SNOTEL-based SWE time series for each station were stacked from which the computed mean time series was as the network-wide mean SWE as shown in Figure 3.11 (note that the SNOTEL-based network-wide mean SWE do not represent watershed-averaged SWE due to the bias associated with most SNOTEL stations being installed in regions with deep snow). Both watersheds exhibit clear seasonality in snow accumulation during the winter and snow ablation in the spring. The network-wide mean SWE and network-wide mean vertical displacement time series derived from GPS, GRACE, and Catchment yield a large, negative correlation indicating that snow is an important hydrologic loading in the study area. More specifically, R values against SNOTEL-based SWE in the Great Basin watershed are -0.74, -0.80, and -0.88 for GPS, GRACE, and Catchment,

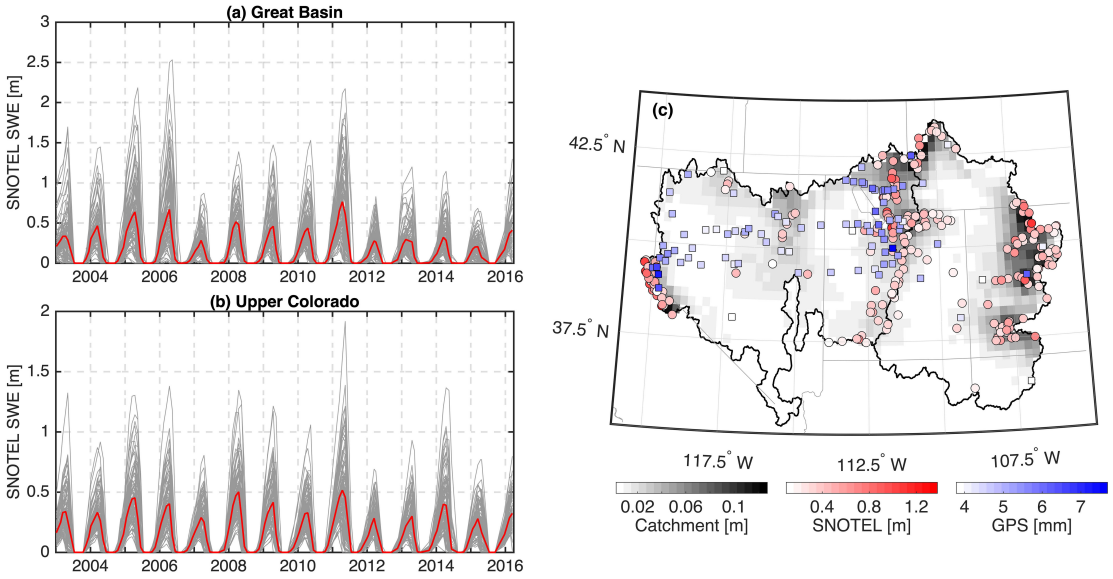

Figure 3.11: SWE time series derived from ground-based SNOTEL stations in the (a) Great Basin and (b) Upper Colorado watershed, with individual station time series in gray and station ensemble mean time series in red. (c) shows the annually-averaged, maximum SWE derived from Catchment (gray scale), SNOTEL (red circles), and the annually-average amplitude of ground-based GPS vertical displacement (blue squares). The annual amplitude of GPS vertical displacement is calculated by dividing the difference between the maximum and minimum of each year by two.

respectively. R values in the Upper Colorado watershed are -0.73, -0.77, and -0.81 for GPS, GRACE, and Catchment, respectively.

Considering the point-scale nature of ground-based GPS measurements, vertical displacement derived from GPS is expected to better represent local (∼10 km) hydrologic loading variations than GRACE and Catchment, which have coarser spatial resolution. The Great Basin watershed is used as a representative example to investigate the capability of GPS-based vertical displacement in inferring local hydrologic loading changes due to snow. Following the methods described in Knappe et al. [90], the local variability in vertical displacement was extracted by subtracting the network-wide mean, vertical displacement time series from individ-

ual GPS-based monthly vertical displacement time series. The correlation coefficient between the residual vertical displacement (i.e., residual after removing the network-wide mean vertical displacement) and SWE time series was then calculated for all GPS-SNOTEL pairs whose separation distance is less than 10 km in the watershed. Note that watershed-averaged SWE was not removed from individual SNOTEL SWE time series because at each SNOTEL station, the measurements of SWE represent the truly local snow mass changes rather than the spatially-integrated effects as measured by ground-based GPS. The same procedure was also conducted for vertical displacement derived from GRACE and Catchment. The correlation coefficients that are statistically significant to a level of significance α=0.05 are used for analysis. It is found that the residual vertical displacement is generally negatively correlated with SNOTEL SWE, with network-wide averaged R values of -0.46, 0.04, -0.28 using the residual vertical displacement derived from GPS, GRACE, and Catchment, respectively. The anti-correlated behavior is not detected when using GRACE-derived residual vertical displacement due to the relatively coarse spatial resolution of GRACE TWS retrievals ($\sim$400 km) that cannot resolve local hydrological loading variations at $\sim$10 km spatial scale. On the contrary, the more negative correlation between GPS-derived residual vertical displacement and SNOTEL SWE demonstrates the capability of ground-based GPS observations in providing information related to local-scale changes in TWS.

However, the negative correlations between residual vertical displacement and SWE are not especially strong for a variety of reasons including: (1) vertical displacement is influenced by changes in TWS, of which SWE is only a small component,

hence, changes in TWS not related to SWE could weaken the correlation between vertical displacement and SWE; (2) the method of extracting local variability in displacement by removing the network-wide mean time series from all of the stations is not always ideal; (3) SNOTEL SWE measures mass loading resting on a snow pillow, and any mass moving off the sensor (e.g., via snow melt or wind redistribution) generally remains in the surrounding region, and by construct, is measured in the integrated TWS signal but not measured in the local SNOTEL SWE signal [90]; and (4) the separation distance between GPS and SNOTEL stations is not the only factor that influences the correlation, e.g., other factors such as elevation difference between the stations should also be considered. Errors in the collection and processing of GPS and SNOTEL measurements can also introduce additional discrepancies.

3.3.4 Inter-annual Change Since Late 2010

As discussed in Section 3.1, a significant, negative vertical displacement followed by an increase in upward deformation can be clearly observed in the GPS vertical displacement time series starting in late 2010. A similar negative vertical displacement (and a corresponding rebound) can also be detected in most of the GRACE signals even though they are not as pronounced as for GPS. In Catchment-derived vertical displacement, the abnormally large, negative displacements seen in late 2010 as well as the prolonged vertical rebound is not captured at most locations. According to the National Climate Report [130] for December 2010, Nevada

and Utah had the wettest December in the 116-year record due to winter storms resulting from deep low pressure systems that developed in a strong west-to-east flow in the jet stream circulation. The movement of cold fronts and low pressure systems were influenced by two large-scale atmospheric circulation patterns, which are the La Niña and the Arctic Oscillation.

In order to quantitatively compare the increase of upward deformation since late 2010 as shown in the vertical displacement time series derived from GPS, GRACE, and Catchment, an uplift rate was fitted to the vertical displacement time series from October 2010 to March 2016 using the seasonally-adjusted vertical displacements derived from GPS, GRACE, and Catchment. A statistical t test was then conducted on the fitted uplift rate to examine if the rate is statistically different from zero or not. Based on the t test (level of significance $\alpha = 0.05$), 90% of GPS station locations yielded a positive uplift rate that is statistically different from zero using GPS-derived time series. The percentage when using GRACE-derived time series was 97%. For Catchment-derived time series, none of the station locations provided an uplift rate that is significantly different from zero. Hammond et al. [131] studied vertical motions in Nevada using GPS observations and a similar increased uplift velocity was observed; the reason for the step before and after 2011 was explained by heavy precipitation in the winter of 2010-2011 with a five-year drought that followed. According to Adusumilli et al. [100] there was a strong La Niña event during 2010 to 2011. A negative correlation coefficient between the Oceanic Niño Index (ONI) and TWS anomaly was reported in most regions of our study area, which suggests an increased TWS anomaly during the La Niña event.

In order to have a more straightforward view of the effect of La Niña on TWS changes, the TWS change rates before (2003-2009) and after (2012-2016) the La Niña event were estimated individually following a method from Han [2] using TWS retrievals from GRACE. The estimated characteristics of TWS changes before, during, and after La Niña are shown in Figure 3.12. Before La Niña (2003-2009), there is no significant water mass loss or gain in either of the two watersheds. Most regions experience a negative water balance (i.e., less precipitation than evapotranspiration) $\sim$-0.5 cm/yr except for the eastern portion of the Upper Colorado watershed. During the La Niña event (2010-2011), a significant increase in TWS was observed in all areas of the Great Basin and over half of the Upper Colorado relative to the water storage during 2008-2009. The magnitude of increase in water storage decays when moving from west to east, which is consistent with previous studies showing a weaker anti-correlation between ONI and TWS anomaly towards the east [100], and is relatively uncorrelated in the Colorado basin [132]. The TWS changes in the western portion of the Great Basin was around 8 cm/yr whereas the southeastern corner of the Upper Colorado had a negative water balance of around -3 cm/yr. After La Niña (2012-2016), the region effectively begins to dry out with a decrease of TWS of approximately -2.5 cm/yr in the western Great Basin. The areas less influenced by La Niña undergo a relatively smaller decrease in TWS. In areas showing a negative TWS change in the Upper Colorado during La Niña, an increase in TWS around 1.5 cm/yr is observed after the La Niña event (2012-2016).

The TWS changes before, during, and after the La Niña event that occurred in 2010-2011 using GRACE-derived TWS retrievals show good consistency with

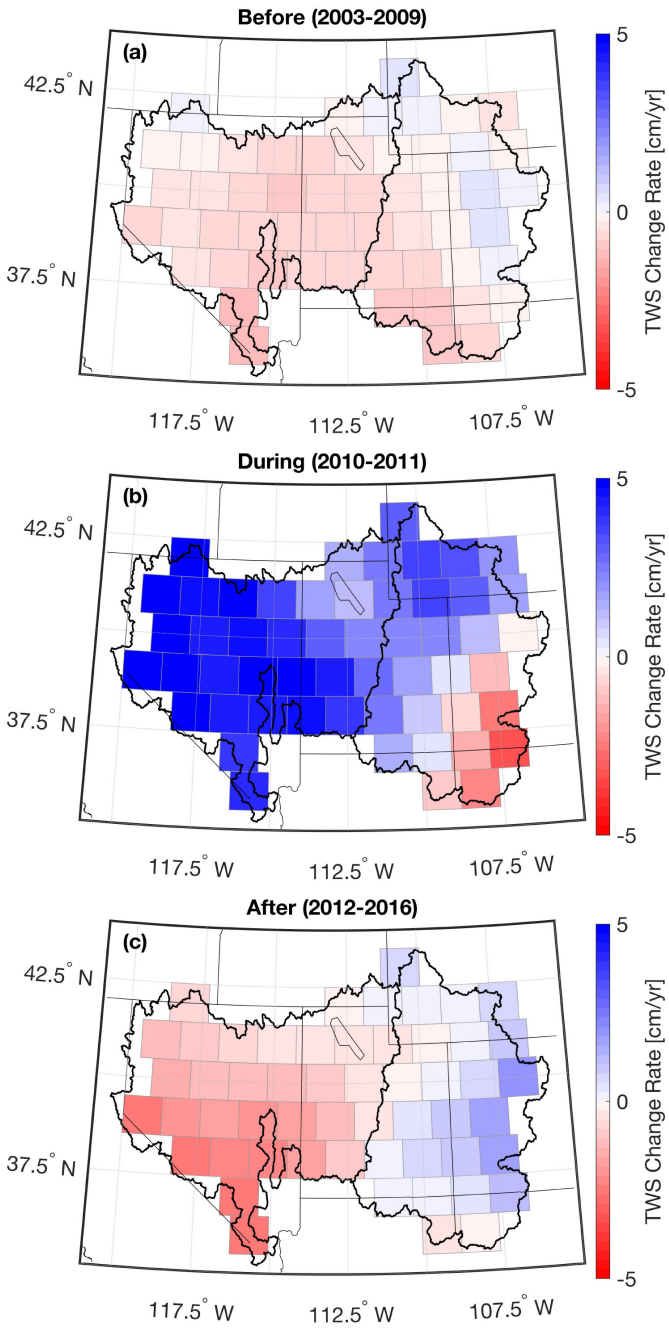

Figure 3.12: Slope of linear regression fitted to GRACE-derived TWS change in units of cm/yr (a) before (2003-2009), (b) during (2010-2011) and (c) after (2012-2016) the 2010-2011 La Niña event following the method described in Han [2].

the changes in vertical displacement derived from GPS. During the La Niña event (2010-2011), a significant water increase resulted in an extreme, negative vertical displacement detected in most of GPS observations. The extended water loss after the La Niña event (2012-2016) resulted in the increase of upward deformation in GPS observations of vertical displacement. The good agreements between TWS change and GPS-derived vertical displacement demonstrates that GPS observations can be a good indicator of TWS change at relatively fine spatial scales. Catchment-derived vertical displacements show better consistency with GRACE-derived vertical displacements with larger R values; however, the prolonged drought after the La Niña event cannot be accurately reflected by the Catchment-based estimates even though some stations show an extremely large negative water balance in early 2011. Therefore, ground-based GPS deformation observations are a valuable data source for investigating TWS variations, especially at a relatively fine spatial resolution.

3.3.5 Analysis of Discrepancies and Error Sources

The comparison of vertical displacements derived from GPS, GRACE, and Catchment shows that a good agreement often exists with GPS time series showing a larger amplitude, which is approximately 2.4 times larger than the GRACE-derived vertical displacement time series amplitude. The amplitude difference between vertical displacement derived from ground-based GPS and space-based GRACE is also found in other studies. Khan et al. [133] estimated a scale factor of $\sim$2.5 between the vertical displacement derived from GPS and GRACE in Greenland. Fu et al. [79]

provides a scale factor of ∼1.1 in the Amazon and ∼1.2 in Southeast Asia between the GPS- and GRACE-based vertical displacement. The relatively larger amplitude in the GPS signal (relative to GRACE and Catchment) is anticipated as GPS is more sensitive to short wavelength (i.e., at the spatial scale of less than ∼100 km) local mass load variations as compared to long wavelength variations [76,79]. On the contrary, GRACE can only sense long wavelength loading variations at monthly time scales, which results in a broader spatial scale (but smaller amplitude) of vertical displacement. Tesmer et al. [76] compared vertical deformations from GRACE and GPS across the globe using 131 GPS stations whose vertical displacement time series are not dominated by local effects, and an improved consistency between GPS and GRACE was witnessed. As for Catchment, the model only includes shallow groundwater storage, and hence, when converted, does not consider vertical deformation associated with changes in deep groundwater storage. Therefore, a smaller amplitude of vertical displacements via Catchment (relative to GPS) is anticipated.

The errors in GPS observations, GRACE retrievals, and Catchment estimates as well as the errors introduced during the processing procedure were investigated. GPS provides a direct measurement of surface deformation, which includes not only TWS variation but also non-hydrologic changes that may not be removed precisely or completely [86, 96]. A number of studies explored the possible causes of errors in the ground-based GPS observations such as orbit mismodeling [134], errors in ocean loading model [23, 76], and height changes caused by a negative relationship between the tropospheric zenith delay and the station heights in the original GPS system of equations [76]. In many studies comparing GRACE and GPS vertical

displacements, the atmospheric and non-tidal ocean loadings were added back to GRACE to maintain consistency between the comparisons [2, 76, 135]. However, in this study, the focus is on the hydrologic loadings in the GPS signals, and thus, the AOD Level 1B product (GAC) used to remove atmospheric and non-tidal ocean loading effects from GRACE was removed from the GPS signal instead. Considering the scale difference in the AOD model relative to the GPS observations, the removed atmospheric and non-tidal ocean loading may not properly capture the local effects, and thus, may result in a discrepancy between the GPS and GRACE estimates.

Another factor that may cause errors in the estimates of vertical displacement, but one that is not explicitly accounted for in this study, is groundwater pumping. GPS stations located on top of aquifers can, at times, respond non-elastically to large groundwater changes due to land subsidence and subsurface consolidation, which is opposite to Earth′s elastic response to snow [22]. When land consolidation occurs, the soil porosity decreases irreversibly in an non-elastic manner, and the Earth′s crust will not rebound correspondingly. Therefore, a careful selection of GPS stations is required in order to remove stations located in areas where land surface subsidence occurs, especially in regions with significant groundwater pumping activities such as the Central Valley in California. Although the study area used here is not significantly influenced by groundwater pumping activities, there will still be some scattered regions with groundwater pumping activities that merit further investigation. Therefore, a closer investigation into each single GPS station will help remove those stations influenced by local groundwater storage changes.

During the conversion of GRACE-derived TWS to vertical displacement, a

homogeneous reaction of the Earth′s crust to mass loading is assumed, which is not always appropriate and may result in some systematic regional effects [76]. Further, as discussed before, GRACE-based TWS only detects mass load variations that cause long wavelength effects. Therefore, local mass variations may not be detected by GRACE, and thus, may not be represented in the GRACE-based vertical displacement time series.

As for the Catchment-derived vertical displacement, two main sources of error that should be highlighted are: (1) error in Catchment-modeled TWS and (2) error in vertical displacement estimates via Green′s function. Catchment only simulates shallow groundwater, which may result in significant errors especially during severe drought years [14, 65]. As for the second error source, the conversion of Catchment TWS to vertical displacement is conducted using Green′s function assuming a PREM Earth model. The reaction of crustal deformation to changes in mass loading is assumed homogeneous, which may be unrealistic in some areas.

3.4 Conclusions

This study compared vertical displacements derived from ground-based GPS, space-based GRACE TWS retrievals, and model-based Catchment TWS estimates for the Great Basin and Upper Colorado watersheds in the United States. A good agreement between vertical displacements derived from the three distinct data sources was found, i.e., the $R_{GPS-GRACE}$, $R_{GPS-Catchment}$, and $R_{GRACE-Catchment}$ values are greater than 0.8 using the network-wide mean vertical displacements derived

from GPS, GRACE, and Catchment. When considering just GPS and GRACE, over 89% of all stations show a correlation coefficient of R>0.70. The RMS reduction after removing GRACE from the GPS signals is consistently positive, indicating a good agreement between GRACE- and GPS-based vertical displacement time series. The highest consistency is found between GRACE and Catchment with all examined stations providing R>0.70. As the information embedded in the GRACE TWS retrieval is related to TWS changes associated with hydrologic loading only, the positive correlation between GPS and GRACE indicates potential in merging the two distinct, yet complementary, information sources to better characterize TWS.

A noticeable change in vertical displacements in early 2011 was observed in most GPS time series, followed by an extended period of upward deformation. A similar pattern was observed in a number of GRACE-derived vertical displacement time series, but not in the Catchment-derived time series. Through an analysis of water balance inter-annual variation using TWS retrievals from GRACE, it was found that heavy precipitation induced by a strong west-to-east flow in the jet stream circulation from 2010 to 2011 influenced a large portion of the study area, especially the Great Basin, causing large, negative vertical displacements from late 2010 to early 2011 associated with a La Niña event. A prolonged drought after the La Niña event resulted in a rebound of vertical displacements. When accounting for the capability of GPS vertical displacement observations in reflecting TWS variability, ground-based GPS is considered a valuable source of information for water storage change at a finer spatial and temporal resolution relative to GRACE or global-scale land surface models. The ground-based GPS sensors can also help bridge the spatial

scales between point-scale (e.g., SNOTEL) and satellite-scale (e.g., GRACE-FO) estimates of terrestrial mass change.

Additional research on the improvement of GPS processing procedures is needed to reduce the potential errors discussed in section 3.5 (e.g., atmospheric loading modeling and groundwater changes) in order to provide a more accurate time series of vertical displacement in response to changes in hydrologic loading. An investigation of incorporating horizontal displacement measured by ground-based GPS into the hydrologic loading analysis should also be explored in order to better resolve hydrologic loading changes in smaller regions such as lakes or other surface water impoundments [71]. Despite their limitations, given the related information content found in the GPS observations and the GRACE TWS retrievals, this work suggests that there is great potential to combine the two distinct information sources as part of a Bayesian merging procedure to be completed in a follow-on study with the ultimate goal of improving TWS estimation at finer spatial and temporal scales.

Chapter 4: Assimilation of Ground-based GPS Observations of Vertical Displacement into the NASA Catchment Land Surface Model

4.1 Motivation and Objectives

The comparison of vertical displacements conducted in Chapter 3 demonstrated the capability of ground-based GPS in representing TWS changes. In this chapter, the utilization of ground-based GPS observations to improve the accuracy of modeled TWS estimates, and its constituent components, is explored. Previous studies have successfully merged GRACE TWS retrievals into land surface models using a data assimilation (DA) framework. Through the DA technique, GRACE TWS retrievals were effectively downscaled in horizontal (i.e., finer spatial resolution) and vertical (i.e., into constituent TWS components) directions, as well as in time (i.e., finer temporal resolution) [1,3,68]. Ground-based GPS observations provide point-scale measurements that are difficult to use in the context of spatially-continuous TWS variations. Additionally, as with GRACE TWS retrievals, GPS observations of vertical displacement reflect column-integrated effects of hydrologic loading (i.e., TWS changes). Therefore, DA is anticipated to be an effective tech-

nique in bridging the spatial gap between GPS point-scale observations and gridded TWS retrievals, as well as make full use of TWS change information as represented by GPS in order to provide individual TWS component estimates such as snow and soil moisture.

The purpose of this study is to further demonstrate the capability of using ground-based GPS observations of vertical displacement to generate a more accurate TWS estimates as well as reduce TWS uncertainty using a DA approach. The potential of GPS DA to improve TWS and terrestrial hydrological components at sub-regional and sub-monthly scales is explored. To our knowledge, this is the first attempt to merge ground-based GPS into a DA framework to improve the prediction accuracy of TWS. The DA experiments are conducted over the same two snow-dominated basins in the western United States (i.e., Great Basin and Upper Colorado watersheds) studied in Chapter 3. Estimated TWS anomalies and their individual component in the terrestrial water cycle are assessed using satellite data and *in situ* measurements. Data sets and the DA framework are introduced in Section 4.2. The TWS anomaly analysis is conducted in Section 4.3.1. Snow water equivalent (SWE), soil moisture, and runoff evaluations are conducted from Section 4.3.2 to Section 4.3.4. Analysis increments and the normalized innovation (NI) sequence are discussed in Sections 4.3.5 and 4.3.6. Section 4.4 summarizes the work and proposes directions for future work.

4.2 Data and Methods

4.2.1 Prognostic Land Surface Model

The prognostic land surface model used in this study is the NASA Catchment Land Surface Model (Catchment) [54,67], which is the land model component of the Goddard Earth Observing System, version 5 (GEOS-5). Different from conventional soil-layer-based LSMs, Catchment explicitly models the horizontal heterogeneity of soil moisture and its impact on evapotranspiration and runoff within each hydrologic catchment [54]. Catchment hydrology model such as TOPMODEL [136] was used to diagnose root zone soil moisture distributions from the morphology of the catchment and prognostic variables of soil moisture. The derived distributions allow the separation of the catchment into hydrological regimes and therefore the heterogeneity of soil moisture horizontally was accounted for in this Catchment model. Further, Catchment employs a three-layer snow model [137] to estimate snowmelt and refreezing processes, which provides the opportunity to estimate TWS in regions where snow is a significant contributor to the hydrologic cycle. In addition, Catchment can model shallow groundwater variations [3, 138], but during an extended drought period, Catchment may fail to capture the complete dynamic range of groundwater, and hence, TWS [14, 65]. Additionally, Catchment does not account for surface water impoundments, anthropogenic water management, or dynamic surface water routing [1, 3], which weakens its performance in regions where surface water storage change is an important component of TWS change.

Table 4.1: Ensemble perturbation parameters for meteorological boundary conditions and model prognostic variables [a].

					Cross-correlations		
	Unit	Type	Standard Deviation	t_{corr}(day)	pcp	sw	lw
pcp		M	0.5	3	-	-0.8	0.5
sw		M	0.3	3	-0.8	-	-0.5
lw	W m^{-2}	A	20	3	0.5	-0.5	-
catdef	mm	A	0.15	1			
swe	mm	M	0.0012	1			

[a] pcp = precipitation; sw = shortwave radiation; lw = longwave radiation; catdef = catchment deficit; swe = snow water equivalent; M = multiplicative perturbation; A = additive perturbation; t_{corr} = temporal correlation

Three prognostic variables are used to represent the vertical soil moisture profile and include the catchment deficit ($catdef$), root-zone excess ($rzexc$), and surface excess ($srfexc$) variables. $catdef$ is the primary subsurface prognostic variable in Catchment and it is defined as the average depth of water that would be needed to bring the catchment to full saturation [54, 68]. Additionally, $rzexc$ and $srfexc$ represent the amount of water in the root-zone layer (0-100 cm) and surface layer (0-5 cm), respectively, in excess of the water that would be present if the soil moisture profile was at equilibrium.

In this study, Catchment was run with a time increment of 7.5 minutes from 1 January 2003 to 31 December 2015 on the 25 km Equal-Area Scalable Earth Grid 2.0 (EASE-Grid 2.0) [139]. Meteorological fields provided by the Modern Era Retrospective Analysis for Research Application version 2 product (MERRA-2) [126] are used as the boundary conditions to Catchment. The perturbation settings for selected meteorological forcing and model prognostic variables are shown in Table 4.1 following Girotto et al. [3]. Model prognostic variables were perturbed in order to

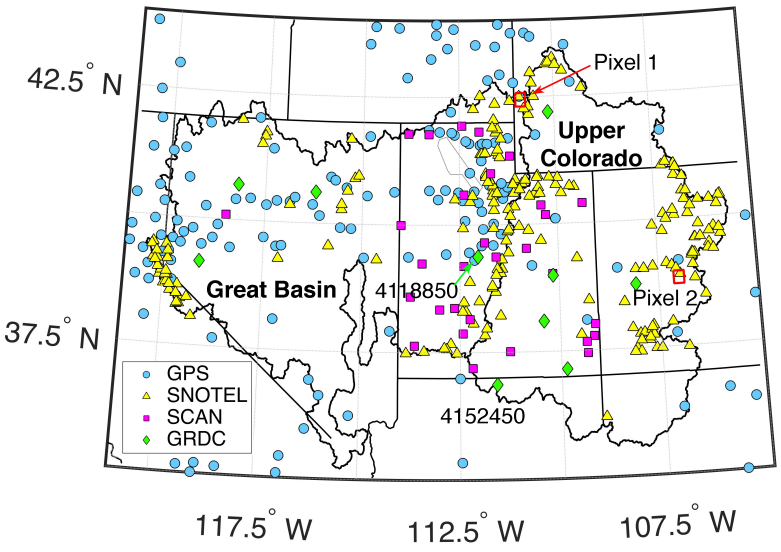

Figure 4.1: Study area including the Great Basin and Upper Colorado watersheds along with ground-based stations of GPS (circle), SNOTEL (triangle), and GRDC (diamond) used in the study marked. The red squares indicate the location of the example pixels discussed in Section 4.3.2 for the SWE analysis.

represent the uncertainties associated with model structure.

4.2.2 GPS Observations

Level 2 daily position data processed by the New Mexico Tech analysis center from the Plate Boundary Observatory (PBO; https://www.unavco.org/data/data.html) project was used to provide ground-based GPS observations of vertical displacement. It is noted that the variations in vertical displacement observations are not only influenced by water mass changes but also non-hydrologic loadings such as atmospheric and non-tidal ocean loading effects. Therefore, a pre-processing procedure as outlined in Chapter 3 was conducted to remove the impact of non-hydrologic loadings.

Due to the point-scale observational nature coupled with limited spatial avail-

ability of the ground-based GPS networks, inverse distance weighting (IDW) interpolation was used to fit a continuous surface of GPS estimates of vertical displacement onto the EASE-Grid 2.0. An interpolation technique was needed here in order to provide a spatially-continuous estimate of vertical displacement (i.e., without gaps associated with GPS data sparsity) as to yield an updated estimate of TWS (see discussion in Section 4.2.3.1) without spatial discontinuities thereby yielding a more realistic, posterior estimate of TWS across the entire study domain. The spatially interpolated fields of vertical displacement were then utilized in the DA framework.

It is understood that the IDW interpolation is a deterministic method, which does not provide estimates of error uncertainty and is sensitive to outliers. More advanced, stochastic interpolation methods such as kriging could provide interpolation results with indicators of errors. However, considering the sparsity of observations in space in conjunction with the uncertainty in the fitted variograms used by the kriging algorithm, IDW was considered as the most applicable and appropriate technique for use here. Future work on the impact of spatial interpolation methods on the DA results could be explored, but was considered secondary in terms of importance in this study and well beyond the scope of work.

To predict the vertical displacement at a single pixel, all available GPS observations within a distance of 400 km are used during interpolation with the weight given to each observation based on the square of the inverse distance (Equation 4.1).

$$dr(p) = \frac{\sum_{j=1}^{n} w_j dr_j}{\sum_{j=1}^{n} w_j} \tag{4.1}$$

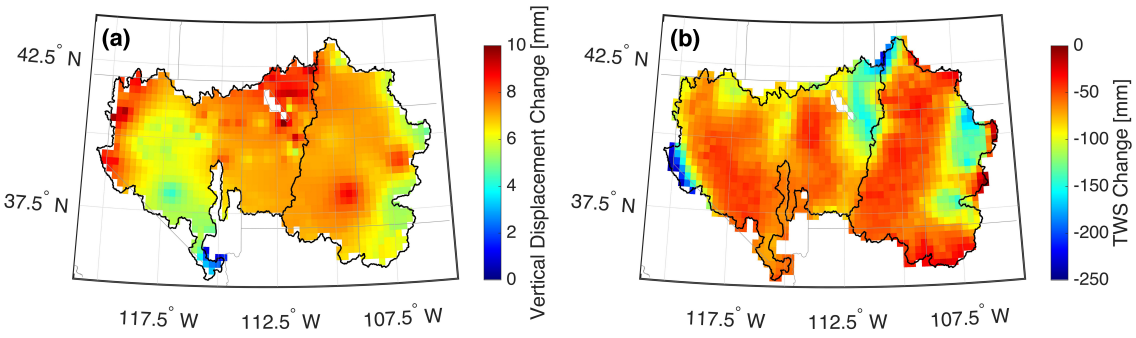

Figure 4.2: Average changes (computed as difference between September 1 and March 1 estimates) across the years 2003 through 2015 in (a) interpolated GPS-based vertical displacement and (b) modeled TWS using Catchment. Note that the average vertical motions in (a) are all positive given the general loss of mass from the near-peak period of the snow season in March through the end of the summer in September.

$$w_j = \frac{1}{d_j^2}$$

where $dr(p)$ represents the interpolated vertical displacement at pixel location p; n is the number of GPS stations located within a distance of 400 km to the pixel p with available observations of vertical displacement; w_j is the weight given to each observation; dr_j represents vertical displacement at each available station j; and d_j is the distance between the GPS station j and the pixel p. A total of 145 GPS stations located in the area range of 35°N to 44°N, and 105°W to 121°W (locations shown in Figure 4.1) were used during spatial interpolation after the non-hydrologic loading contributors were removed. The vertical displacement changes from March 1 to September 1 were calculated for each year using the interpolated results, and the average changes across the years 2003 through 2015 are shown in Figure 4.2. An uplift motion in the vertical displacement from March through August was found across the entire study area, which indicates a decrease of TWS associated with

snow ablation and runoff processes. The spatial characteristics of the uplift motion are generally in good agreement with the TWS changes reflected by the ensemble mean of the Open Loop simulation as shown in Figure 4.2(b). A large amount of uplift (∼8-10 mm) was found in the western and northeastern portions of the Great Basin as well as the eastern Upper Colorado where a significant amount of snow accumulation occurs during the winter season and subsequently ablates between March and September, which results in the corresponding uplift during this time period.

4.2.3 Data Assimilation Framework

4.2.3.1 Ensemble Kalman Filter

In order to assimilate the spatially-interpolated, daily vertical displacements derived from the ground-based GPS observations into the land surface model, a one-dimensional (1-D) ensemble Kalman filter (EnKF) framework was applied in this study. The EnKF is a Monte-Carlo based approximation of the Kalman filter, and it is suitable for moderately nonlinear land surface models [140,141]. The EnKF has been successfully used for soil moisture and snow data assimilation [142,143], but this is the first known attempt at applying it to GPS observations of vertical displacement in the context of TWS. Estimates of model state and error covariance matrices (required for determining the relative weights of the model forecast and observations) based on the mean and spread of ensemble members are used during the Bayesian merger of the land surface model with the ground-based GPS observations.

The EnKF performs propagation and update steps sequentially over the course of the study period. During the propagation phase, the prior estimate of model states at time t for each ensemble replicate i (i.e., x_t^{i-}), is computed with the land surface model by using the posterior (a.k.a. updated) estimates of model states at time $t-1$ (i.e., x_{t-1}^{i+}). When an observation becomes available at time t, an update step is performed such that the updated estimate of model states at time t (i.e., x_t^{i+}) is computed as

$$x_t^{i+} = x_t^{i-} + K_t[y_t + v^i - H(x_t^{i-})] \tag{4.2}$$

where $i \in [1, N]$ represents each ensemble replicate of the ensemble size N; K_t is the Kalman gain matrix; y_t is the observation vector; v^i is a temporally uncorrelated, Gaussian-distributed random perturbation with zero mean added to the observation that represents the observation error; and $H(\cdot)$ is the observation operator that maps the model states to the observations space. K_t controls the partitioning of the update assigned to each model state variable and can be calculated as

$$K_t = Cov[x_t^-, H(x_t^-)][Cov[H(x_t^-), H(x_t^-)] + C_{vv}]^{-1} \tag{4.3}$$

where $Cov[x_t^-, H(x_t^-)]$ is the error covariance between the prior estimates of model states and the predicted observations; $Cov[H(x_t^-), H(x_t^-)]$ is the error (sample) covariance of the predicted observations; and C_{vv} is the observation error covariance.

In this study, N=24 ensemble replicates were used and the increments (i.e., $K_t[y_t + v^i - H(x_t^{i-})]$) were only applied to *catdef* and *swe* explicitly following the

GRACE TWS data assimilation procedure outlined in Girotto et al. [138]. An observation error standard deviation of 5 mm, including errors associated with GPS data pre-processing, was used in this study. More discussion on the selection of the observation error standard deviation is provided in Section 4.3.6.

4.2.3.2 Observation Operator

Previous studies demonstrated that Green′s function can serve as an effective, linear operator to map TWS changes into vertical displacement in order to quantify the elastic response of the Earth′s surface to mass redistribution [22, 96]. Therefore, in this study, a forward model based on Green′s function [70, 71] was used as the observation operator, $H(\cdot)$, to convert modeled TWS anomaly (geophysical model state space) into vertical displacement (observation space) using the Preliminary Reference Earth Model (PREM) [82]. More details on Green′s function used in this study are found in Section 3.2.4. Prior to applying the observation operator, the modeled TWS via Catchment was converted into TWS anomaly by subtracting the long-term average from 1 January 2003 to 31 December 2015 at each grid cell. The obtained TWS anomaly was subsequently used to calculate the model prediction of vertical displacement.

4.2.4 Evaluation Approach

Modeled TWS estimates from the OL and GPS DA are evaluated using GRACE TWS retrievals. Modeled SWE, soil moisture, and runoff were evaluated against *in*

situ measurements. A notable gap in the results analysis is the lack of a groundwater measurement comparison. Based on results from Fan et al. [144], the groundwater table depth in the study area is deep with most areas having a groundwater table depth greater than 80 m. However, the capability of Catchment in modeling groundwater is limited [14] with the main focus of Catchment in modeling groundwater limited to the top few meters. Therefore, an *in situ* groundwater analysis is not conducted in this study, but rather is implicit in the TWS retrievals discussed below.

4.2.4.1 GRACE TWS Retrievals

The Level-3 GRACE monthly $1^\circ \times 1^\circ$ land gridded TWS retrieval based on the Release-05 (RL05) spherical harmonics provided by the Center for Space Research (CSR) at the University of Texas Austin is used in this study (http://grace.jpl.nasa.gov) [145]. The GRACE solution from CSR were truncated to degree 60 and smoothed using a 300 km wide Gaussian filter. The GRACE solution is used to evaluate the skill of modeled TWS anomaly derived from the OL and GPS DA. Daily estimates of TWS anomaly derived from the OL and GPS DA are converted into a monthly-average based on the GRACE TWS retrieval time periods, and then these results are evaluated using GRACE TWS retrievals. Additionally, a GRACE DA experiment following the methods of Girotto et al. [3] was similarly conducted to help evaluate GPS DA performance. Similar as GPS DA, only variables $catdef$ and swe were update explicitly in the GRACE DA simulation [138].

4.2.4.2 SNOTEL SWE Measurements

Ground-based measurements of SWE provided by the SNOwpack TELemetry network (SNOTEL; https://www.wcc.nrcs.usda.gov/snow/) were used to evaluate the skill of modeled SWE from the OL and DA simulations, including both GPS DA and GRACE DA. SNOTEL is operated by the Natural Resources Conservation Service (NRCS) National Water and Climate Center. Stations providing SWE records for less than half of the study period were excluded from the analysis, and as a result, 105 and 106 stations were found in the Great Basin and Upper Colorado basins, respectively, as shown in Figure 4.1. OL and DA estimates of SWE are compared against SNOTEL SWE measurements at a monthly timescale. If there is more than one SNOTEL station located inside a given model pixel, the station-average SWE measurements were used in the evaluation. When using statistical metrics to evaluate the consistency between modeled SWE and SNOTEL-based SWE, coincident zeros were excluded. That is, statistics were only computed when SWE was present in the model estimates or the SNOTEL measurements, which provides a more rigorous statistical snow comparison by excluding coincident snow-free periods.

4.2.4.3 SCAN Soil Moisture Measurements

Comparing modeled soil moisture with *in situ* measurements is complex due to the localized nature of the *in situ* measurements of soil moisture coupled with measurement network sparsity [146]. With that said, surface soil moisture measurements measured at top 5 cm depth provided by the Soil Climate Analysis Network (SCAN;

https://www.wcc.nrcs. usda.gov/ftpref/) were used to evaluate modeled surface soil moisture (0-5 cm surface layer) at a monthly timescale. Following quality control steps in Liu et al. [147], measurements that exceed a physically plausible range (i.e., 0-100%) or measured during frozen soil conditions were excluded. Stations providing records less than three years during the study period were similarly excluded from the evaluation in order to yield a statistically meaningful evaluation. Additionally, stations with inconsistent records that are most likely caused by changes in sensor calibration and/or sensor re-installation, or other miscellaneous factors, were excluded from the analysis. After completing the quality control procedure, 19 and 11 stations were retained in the Great Basin and Upper Colorado watersheds, respectively. Due to the relatively sparse SCAN network, there are no model pixels collocated with more than one SCAN station.

It is worth noting that a root zone soil moisture evaluation was not conducted here. The root zone soil moisture in Catchment is highly dependent on the groundwater table depth [54]. However, Catchment only considers shallow groundwater, and as a result, the interface between root zone soil moisture and groundwater may not be adequately represented in this study domain given the relatively deep groundwater conditions as mentioned earlier in Section 4.2.4.

4.2.4.4 GRDC Runoff Measurements

Model estimates of runoff were also evaluated using ground-based measurements. Monthly-averaged runoff data provided by the Global Runoff Data Center

(GRDC; https: //www.bafg.de/GRDC/EN/Home/homepage_node.html) was used in the study. Ten GRDC stations (Figure 4.1) corresponding to a range of upland areas, including four stations in the Great Basin and six stations in the Upper Colorado watersheds, were used in this study. Station 4152450 located in the southwest of the Upper Colorado basin is noted to be the outlet gauge of the Upper Colorado watershed. The ability of model estimates to represent runoff at different spatial scales was also explored. Due to the lack of a dynamic river routing scheme in Catchment, for each runoff station, a three-month average surface runoff was used during evaluation in order to better mitigate the phase lag between model estimates and the *in situ* measurements [1, 148]. More discussion on runoff comparison is provided in Section 4.3.4.

4.2.4.5 Evaluation Metrics

For the TWS analysis, both OL and GPS DA results are compared against GRACE-based TWS retrievals. As for the SWE, soil moisture, and runoff analyses, estimates of SWE, surface soil moisture, and runoff derived from the OL, GPS DA, and GRACE DA are evaluated against *in situ* measurements. Table 4.2 summarizes the evaluation strategy for each of the modeled variables. To quantify the performance of model estimates versus measurements, the correlation coefficient (R) and unbiased root-mean-square difference (ubRMSD) are used. R is used to reflect the agreement in variations between model estimates and measurements, but the amplitude differences between different data sets are not accounted for. The

Table 4.2: Experimental design and evaluation strategy.

Variable	**Experiment**	**Evaluation Data Source**
TWS Anomaly	OL and GPS DA	GRACE TWS retrieval
SWE	OL, GPS DA, and GRACE DA	SNOTEL
Surface Soil Moisture	OL, GPS DA, and GRACE DA	SCAN
Runoff	OL, GPS DA, and GRACE DA	GRDC

ubRMSD considers both amplitude and phase difference; a large amplitude difference between different data sets results in a large ubRMSD. Because satellite-based and *in situ* measurements both contain errors and cannot represent the "truth", the term "difference" is used here rather than "error". The ubRMSD is computed as the RMSD after removing the long-term mean difference between the model estimates and measurements [149] and can be computed as

$$ubRMSD = \sqrt{\frac{1}{N_t}\sum_{t=1}^{N_t}(x_{est,t} - x_{meas,t})^2 - (\overline{x}_{est} - \overline{x}_{meas})^2} \tag{4.4}$$

where N_t denotes the total number of months; $x_{est,t}$ is the ensemble mean estimate derived from the OL or DA at time t; $x_{meas,t}$ is the measurement obtained at time t; $\overline{x}_{est}$ and $\overline{x}_{meas}$ represent the time-averaged variables obtained from the model estimates and measurements, respectively.

The statistical difference in watershed-averaged skill (i.e., R and ubRMSD) between the OL and DA simulations are tested using the student′s t-test as shown in Section 4.3.1. As for testing the difference in two dependent correlation coefficients derived from the OL and DA (Section 4.3.4), the Hotelling-Williams test [150] is used. For the Hotelling-Williams test, the null hypothesis is that two dependent correlations are equal (i.e., $H_0 : R_{12} = R_{13}$). In this study, R_{12} represents the corre-

lation coefficient between the measurements (*in situ* or satellite retrievals) and the OL results, and R_{13} represents the correlation coefficient between the measurements and the DA results. A t-statistic can be calculated as

$$t_{N_t-3} = (R_{12} - R_{13}) \sqrt{\frac{(N_t - 1)(1 + R_{23})}{2\frac{N_t-1}{N_t-3}|R| + \overline{R}^2(1 - R_{23})^3}} \tag{4.5}$$

where R_{23} is the correlation between the OL and DA results; $\overline{R} = \frac{R_{12}+R_{13}}{2}$; and $|R| = 1 - R_{12}^2 - R_{13}^2 - R_{23}^2 + 2R_{12}R_{13}R_{23}$. If the p-value based on the computed t-statistic for a given N_t is less than the given level of significance, the null hypothesis (H_0) is rejected, indicating that correlation coefficient R values from the OL and DA are statistically different.

4.3 Results and Discussion

4.3.1 Terrestrial Water Storage

The watershed-averaged TWS anomaly derived from the OL and GPS DA simulations, as well as the independent GRACE TWS retrievals are shown in Figure 4.3. The daily estimates from the OL and GPS DA are converted into monthly averages in accordance with the GRACE TWS retrieval periods. The three different data sets show good consistency in capturing the seasonal variation of TWS changes. The amplitude of TWS anomaly provided by GPS DA is generally larger than the OL, with a tendency to add water mass during the winter season and remove mass during the summer season. The ensemble range in the GPS DA case is much more

narrow than in the OL case thereby suggesting reduced uncertainty in the estimated TWS when using the DA approach.

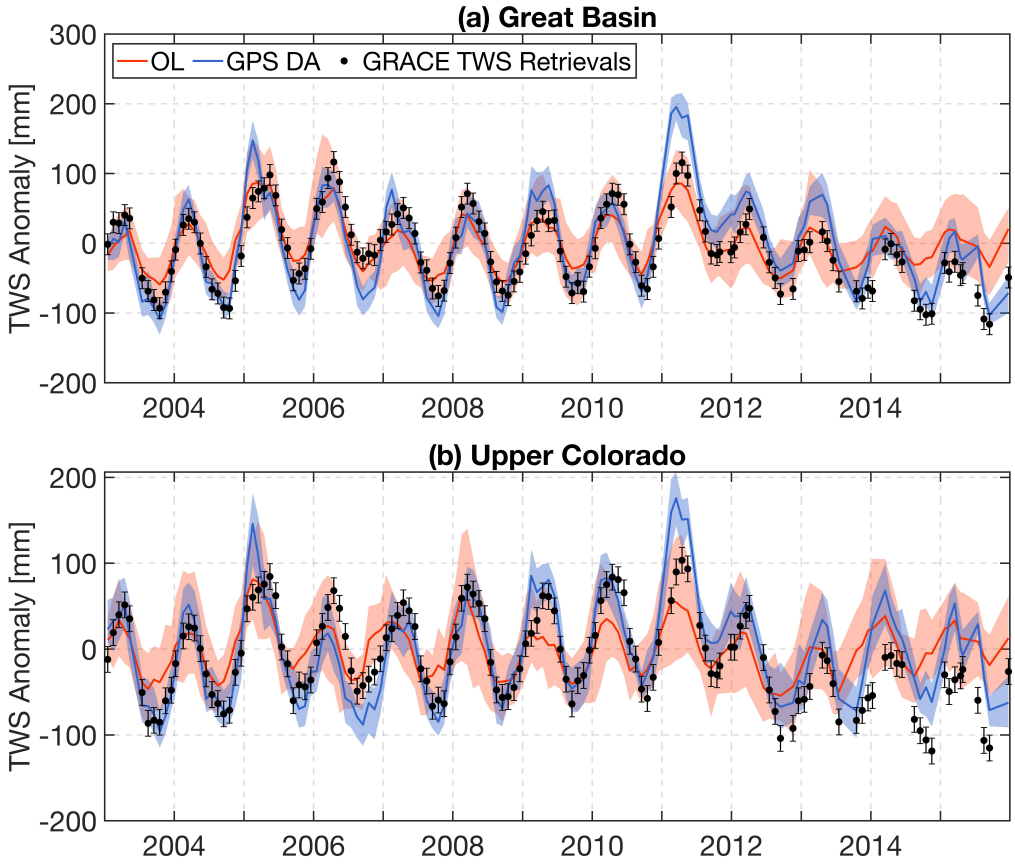

Figure 4.3: Watershed-averaged TWS anomaly time series from the OL (red), GPS DA (blue), and GRACE TWS retrievals (black dots) for the (a) Great Basin and (b) Upper Colorado watersheds. Each line represents the respective ensemble mean, and the ensemble range is shaded with corresponding transparent color. The error bars represent the time-invariant standard deviation of the GRACE observation error.

Both GRACE TWS retrievals and GPS DA estimates reveal a larger than normal TWS anomaly in the 2010-2011 winter, which was followed by a reduction in the amount of total water mass after 2011. The relatively larger TWS anomaly is represented in the OL, but the corresponding decrease in TWS was not witnessed. According to NOAA National Centers for Environmental Information [151], Nevada and Utah experienced one of the wettest October-to-March six-month pe-

Table 4.3: Watershed-averaged correlation coefficient (R) and unbiased root-mean-square difference (ubRMSD) computed between modeled TWS anomalies (i.e., the OL and GPS DA) and GRACE TWS retrievals. Bold font indicates statistically significant differences between the OL and DA at a 5% significance level based on the t-test.

Periods	Basins	OL		GPS DA	
		R	ubRMSD (mm)	R	ubRMSD (mm)
2003/01 - 2015/12	Great Basin	0.76	36	**0.79**	**43**
	Upper Colorado	0.64	42	**0.74**	44
2010/09 - 2015/12	Great Basin	0.46	36	**0.82**	**31**
	Upper Colorado	0.36	41	**0.70**	**36**

riods between October 2010 and March 2011. Hammond et al. [131] also reported particularly heavy precipitation in the Sierra Nevadas in the winter of late 2010 to early 2011 that was followed by a five-year drought until late 2015. Therefore, it is reasonable to suppose that GRACE TWS retrievals and GPS DA estimates successfully detected the drought period from 2011-2015, but that the OL failed.

The correlation coefficient R and ubRMSD between modeled TWS anomalies (i.e., OL and GPS DA) and GRACE TWS retrievals are computed for each GRACE pixel for the entire study period and the drought period (i.e., after mid-2011). Afterwards, the watershed-averaged R and ubRMSD are computed as shown in Table 4.3. For the Great Basin watershed, GPS DA exhibits a slightly larger R than the OL (R = 0.79 for GPS DA and R = 0.76 for the OL), but also a larger ubRMSD (ubRMSD = 43 mm for GPS DA and ubRMSD = 36 mm for the OL) for the entire study period. However, for the drought period, GPS DA greatly improves the consistency between modeled TWS anomaly and GRACE TWS retrievals, yielding an R value of 0.82 as compared to R = 0.46 for the OL estimates and a smaller ubRMSD value

of 31 mm relative to the ubRMSD = 36 mm for the OL simulation.

As for the Upper Colorado watershed, modeled TWS anomaly derived from GPS DA consistently shows better agreements with the GRACE TWS retrievals than did the OL. GPS DA provides a slightly larger ubRMSD relative to the OL (ubRMSD = 44 mm for GPS DA and ubRMSD = 42 mm for the OL) at the Upper Colorado watershed for the entire study period, but the difference is not statistically significant at the 5% level based on the t-test. It is also noticed that both the OL and GPS DA exhibit relatively inferior performance (i.e., smaller R values) for the Upper Colorado watershed than for the Great Basin watershed, which may be explained by the larger impact of surface water storage changes (e.g., rivers and reservoirs) on TWS that is not fully represented in the physics of Catchment [3]. In addition, the less dense GPS measurement network in the Upper Colorado watershed could also account for the smaller improvements relative to the more densely-instrumented Great Basin watershed. The significant improvements in the TWS anomaly estimates provided by GPS DA during the drought period help compensate for the limited dynamic range of the Catchment model in reproducing extreme TWS anomalies due to a lack of deep (i.e., >10 meters) groundwater storage.

4.3.2 Snow Water Equivalent

Monthly averaged SWE estimates derived from the OL and DA at the 625 km^2 EASE-Grid 2.0 pixel scale are compared against point-scale ($\sim$1 m^2) SNOTEL-based measurements. The number of pixels with at least one SNOTEL station collocated

are 63 and 74 in the Great Basin and Upper Colorado watersheds, respectively. In addition to the OL and GPS DA results, GRACE DA estimates of SWE using the methods of Girotto et al. [3] are also presented in the evaluation of GPS DA performance. For each watershed, a typical, example pixel was used to illustrate the temporal variation of SWE derived from the OL, GPS DA, and GRACE DA simulations in conjunction with the SNOTEL measurements (Figure 4.4). The locations of the selected pixels are shown in Figure 4.1.

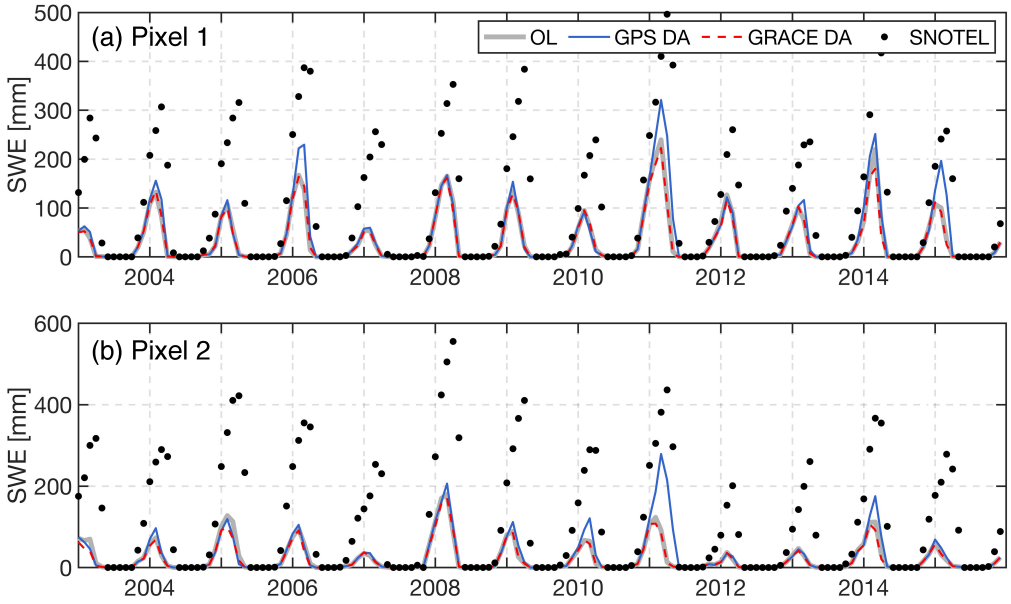

Figure 4.4: Time series of SWE derived from the OL (gray solid line), GPS DA (blue solid line), GRACE DA (red dash line), and SNOTEL measurements (black dots) at (a) pixel 1 located in the northeast of Great Basin; (b) pixel 2 located in the east of Upper Colorado basin. Pixel locations are highlighted in Figure 4.1. The monthly-averaged ensemble means from the OL and DA simulations are shown relative to the monthly-averaged SNOTEL SWE measurements.

Modeled SWE (both OL and DA) and SNOTEL measurements present similar seasonal variations of snow accumulation and ablation, but the amplitude derived from the model estimates is consistently smaller than for the SNOTEL measure-

ments (Figure 4.4). SWE estimates from the OL and GRACE DA do not show much difference from one another except that GRACE DA tends to reduce the peak SWE in some years such as the 2014 winter season. On the contrary, GPS DA consistently increases the peak SWE as compared to the OL, thereby attempting to better match the amplitude of the modeled SWE toward the *in situ* measurements. An example of this amplitude increase is displayed in the 2010-2011 winter at both example pixel locations in Figure 4.4. Additionally, GPS DA consistently extended the peak SWE occurrence to one or two months later relative to the OL and GRACE DA simulations in many years, especially for pixel #2 located in the Upper Colorado watershed, thereby yielding better agreement with the SNOTEL measurements. That is, GPS DA had a tendency to extend the snow season relative to the OL and GRACE DA simulations.

The spatial distribution of the statistical metrics (i.e., R and ubRMSD) for the OL and DA estimates versus SNOTEL SWE, as well as the differences in skill between DA and OL, are shown in Figure 4.5 and Figure 4.6. The consistency between modeled SWE and SNOTEL measurements shows considerable variation across space with R values in the range of 0.05 to 0.86 for the OL and GRACE DA, and 0.05 to 0.87 for GPS DA. Modeled SWE derived from the OL, GPS DA, and GRACE DA all suggest better agreement with the SNOTEL measurements in the Upper Colorado watershed, especially in the northern and eastern regions where deep snow regularly occurs during the winter season. Smaller correlation coefficients between model estimates and *in situ* measurements are mostly found in the central portion of the Great Basin watershed where there is little (ephemeral) or no snow.

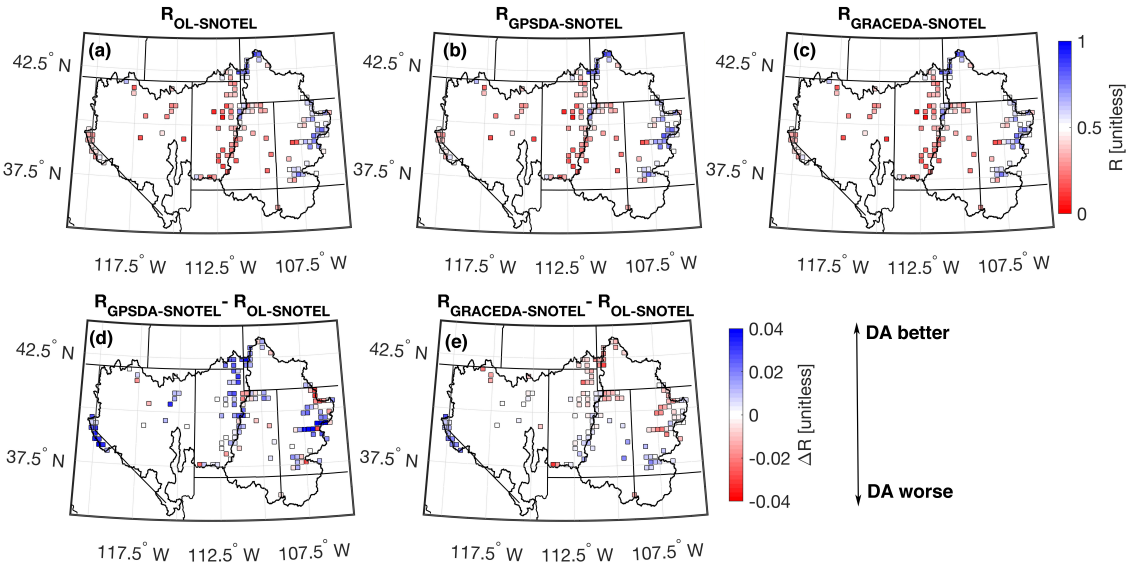

Figure 4.5: Spatial distributions of correlation coefficient, R, for SWE between (a) the OL and SNOTEL; (b) GPS DA and SNOTEL; (c) GRACE DA and SNOTEL, and the difference in R calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SNOTEL measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SNOTEL measurements.

It is important to note that the existence of representativeness error between the *in situ* measurements (spatial support of $\sim$1 m^2) versus the model estimates (spatial support of $\sim 10^8$ m^2) and that it is not expected that the measurements and model estimate would be in perfect agreement. Furthermore, errors in the boundary conditions (i.e., precipitation) and model structure error (e.g., no blowing snow processes included) can result in further discrepancies between the two.

Comparing the performance of the OL and DA simulations by subtracting the R value obtained via the OL case from the DA simulations (i.e., GPS DA and GRACE DA), a positive difference in R suggests that DA enhanced the consistency between modeled SWE and the SNOTEL measurements. For the GPS DA case, improvements are found in 76% and 69% of pixels with collocated SNOTEL stations

in the Great Basin and Upper Colorado basins, respectively. The improvements are mainly located in mountainous areas with deep snow during the winter season. For GRACE DA, the percentages of improved pixels are only 54% and 41% in the Great Basin and Upper Colorado basins, respectively.

A similar analysis was conducted using ubRMSD as the evaluation metric (Figure 4.6). It is found that large ubRMSD values are mostly located in mountainous areas such as the northern and eastern regions of the Great Basin watershed (Figure 4.6(a) through 4.6(c)). Comparing the skill differences between the OL and DA simulations, GPS DA again shows improved consistency with the *in situ* measurements in areas covered by deep snow during the winter season, especially in regions where large ubRMSD values were originally detected in the OL. There is no significant skill difference in the middle of the Great Basin watershed where there is little (ephemeral) or no snow during the year. The percentages of pixels showing improved skill (i.e., reduced ubRMSD) as compared to the OL are 81% and 92% for the Great Basin and Upper Colorado basins, respectively, in terms of the GPS DA case. As for GRACE DA, slight reductions in ubRMSD can be found in the western Great Basin relative to the OL, whereas in the Upper Colorado basin almost all pixels collocated with SNOTEL stations show increased ubRMSD (i.e., DA worse) when assimilating the GRACE TWS retrievals. The corresponding percentages of improvement via GRACE DA are only 46% and 9%, indicating inferior performance in SWE estimation, in general, as compared to OL.

Both R and ubRMSD results demonstrated the capability of GPS DA in improving SWE estimates relative to SNOTEL-based measurements, especially in re-

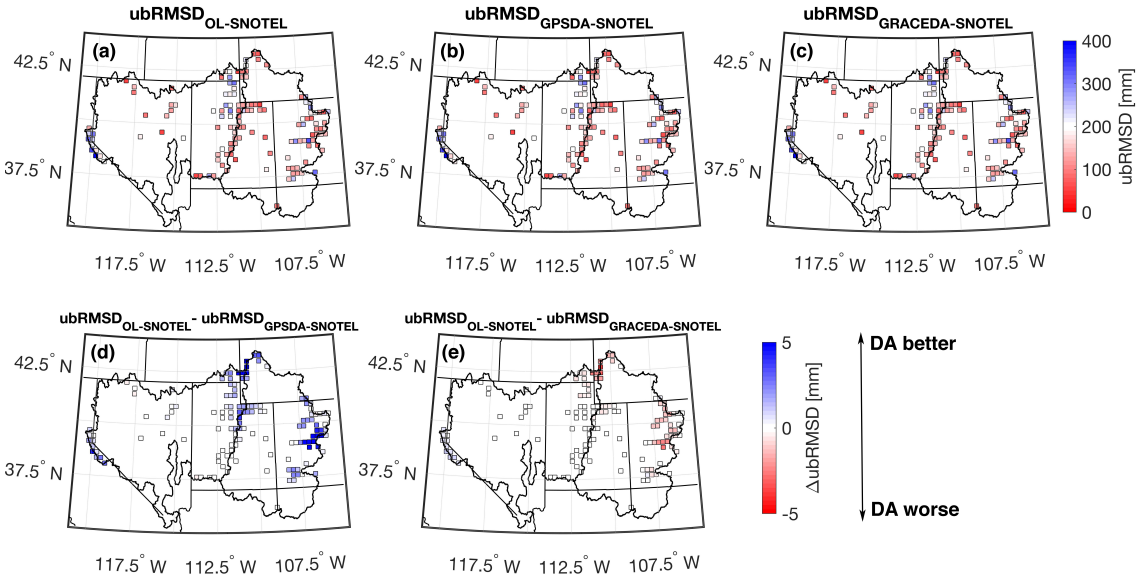

Figure 4.6: Spatial distributions of unbiased root-mean-square difference, ubRMSD, for SWE between (a) the OL and SNOTEL; (b) GPS DA and SNOTEL; (c) GRACE DA and SNOTEL, and the difference in ubRMSD calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SNOTEL measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SNOTEL measurements.

gions with deep snow. The better performance of GPS DA relative to the OL and GRACE DA highlights the advantages of incorporating information content from the GPS observations into the land surface model. Despite the fact that GPS DA improved SWE estimates in many pixels relative to the OL or GRACE DA, there are still a number of pixels exhibiting relatively weak consistency (i.e., $R < 0.5$) between modeled SWE and SNOTEL-based SWE. Analogously, poor consistency between modeled SWE and *in situ* measurements were also found in Xue et al. [143] and were partially explained by the presence of representativeness error between Catchment and SNOTEL. Model structure error in Catchment can also account for some of the inconsistencies. Furthermore, the highly heterogeneous and complex

terrain across the study area can introduce large spatial variations in the relatively coarse-scale meteorological boundary conditions (e.g., precipitation and temperature) used for model simulation, which can result in further exacerbation of the differences between model estimates and *in situ* measurements.

4.3.3 Soil Moisture

Modeled surface (top 0-5 cm) soil moisture were compared against SCAN measurements of soil moisture at 5 cm depth from the surface. The maps of evaluation metrics of correlation coefficient R and ubRMSD for both the OL and DA simlations, as well as the difference in skill between DA and the OL, are shown in Figure 4.7 and Figure 4.8. It is noted that comparing model estimates (at 625 km^2 pixel grid) against sparse, point-scale measurements is difficult due to representativeness error, especially for surface soil moisture that is highly variable in space and time. Additionally, modeled surface soil moisture is a depth-averaged soil moisture content estimated across the top 5 cm whereas SCAN measures the soil moisture at a depth of 5 cm from the surface, which results in further discrepancies between the modeled and measured soil moisture.

Based on the statistical skill, both DA simulations did not improve the surface soil moisture estimates relative to the OL. The R values for the OL are in the range of 0.42 to 0.91 for the study area with a network mean R value of 0.72. GRACE DA provides a smaller network-mean R value of 0.64, but does not differ in a statistically significant sense based on a t-test with a significance level of 5%.

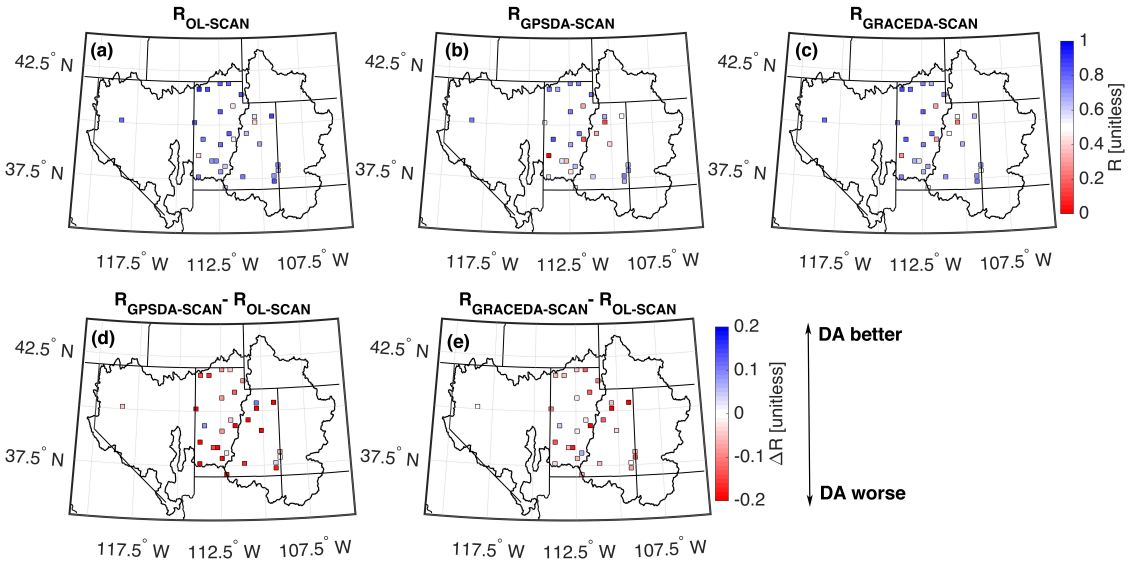

Figure 4.7: Spatial distributions of correlation coefficient, R, for surface soil moisture between (a) the OL and SCAN; (b) GPS DA and SCAN; (c) GRACE DA and SCAN, and the difference in R calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SCAN measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SCAN measurements.

For GPS DA, a statistically degraded network mean R (R = 0.56) was detected relative to the OL, and there is one negative R (R = -0.01) value found at one SCAN station location. Considering the ubRMSD, the network mean ubRMSD derived from the DA simulations are both not statistically different from the OL with ubRMSD values of 4.2%, 4.5%, and 4.9% respectively for the OL, GRACE DA, and GPS DA. Comparing the skill between the OL and DA at each SCAN location, only three out of a total of 30 stations examined show improved R values via GRACE DA while the number for GPS DA is four. When considering ubRMSD, the number of stations with improved ubRMSD (i.e., smaller ubRMSD) relative to the OL are three and eight for GRACE DA and GPS DA, respectively. Although the OL shows a better consistency with SCAN measurements, it is found that the

OL shows a limited variation range of surface soil moisture relative to the *in situ* measurements (figure not shown). However, GPS DA tends to add a larger dynamic range to the surface soil moisture estimates, which provides a better agreement with *in situ* measurements in terms of amplitude.

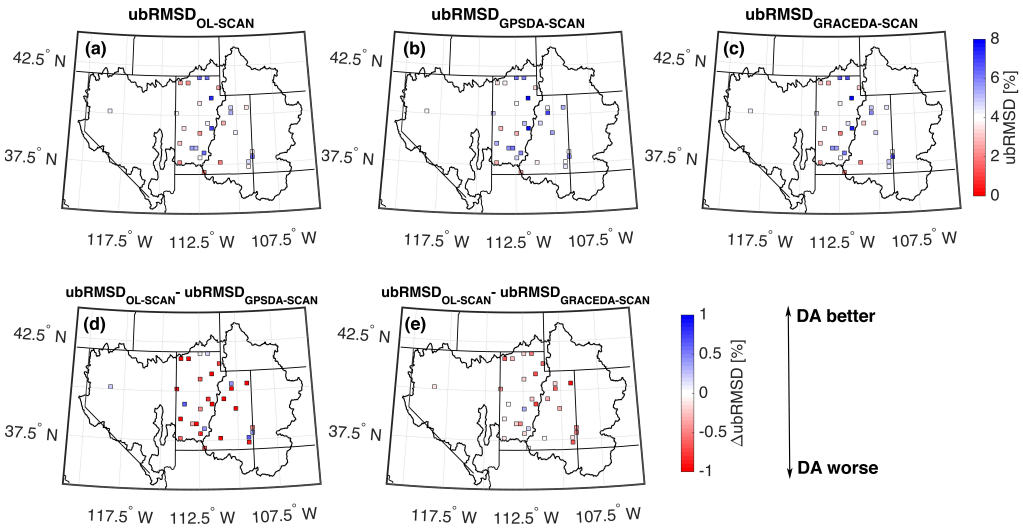

Figure 4.8: Spatial distributions of unbiased root-mean-difference, ubRMSD, for surface soil moisture between (a) the OL and SCAN; (b) GPS DA and SCAN; (c) GRACE DA and SCAN, and the difference in R calculated as (d) GPS DA minus OL; and (e) GRACE DA minus OL. In (d) and (e), a positive difference (blue color) suggests that DA has a better agreement with the SCAN measurements as compared to the OL. On the contrary, a negative difference (red color) suggests the OL agrees better with the SCAN measurements.

The degraded surface soil moisture estimates provided by GRACE DA relative to the OL are similar to the results reported in in previous studies [3, 152]. Surface soil moisture content is strongly influenced by the relatively high-frequency meteorological boundary conditions. However, GRACE TWS retrievals are provided at a monthly timescale, that include components other than surface soil moisture, and in turn, are less impacted by changes in surface soil moisture. As for GPS DA, GPS observations of vertical displacement are influenced by the amount of TWS

change as well as the distance between the locations where TWS change occurs and where the GPS station is installed. In other words, GPS observations are more sensitive to TWS changes close to the GPS station location (i.e., short-wavelength mass changes; see Figure 3.3). Since the water content in the top 5 cm of soil is not a dominant component of TWS coupled with the fact that there are no GPS station located close to the SCAN stations, the influence of the surface soil moisture changes may be rather unsubstantial on the GPS observations of vertical displacement, and thus, may not be easily disentangled from the GPS observations. Therefore, it is hypothesized here that the ground-based GPS observations have limited sensitivity to surface soil moisture changes, which results in the degraded performance, in general, to the surface soil moisture estimates. However, an increase in surface soil moisture dynamics via GPS DA simulation demonstrates the potential for future work to evaluate GPS DA surface soil moisture estimates in a region with a more spatially-dense GPS network.

4.3.4 Runoff

Improving hydrologic flux estimates through data assimilation is often more challenging than improving hydrologic state estimates [68]. For example, an increased estimate of soil moisture via data assimilation may lead to an overestimation of evaporation in order to calibrate the model simulation [153]. Additionally, Catchment does not include dynamic surface water routing or lake storage routines, which makes it difficult to appropriately evaluate the estimated runoff against river

gauge measurements. In order to better compensate for a lack of dynamic surface routing in Catchment, a number of adjustments were made prior to conducting a runoff analysis. Namely, for each measured river gauge location, the upland contributing pixels to that gauge were first located, and then the daily-averaged runoff at that gauge was computed by summing up the surface runoff estimates from the upland contributing pixels. Afterwards, a three-month average was computed for the runoff time series derived from the simulations (both OL and DA) and subsequently compared against the corresponding three-month averaged GRDC *in situ* measurements. The quarterly runoff analysis was used instead of daily or monthly analysis because Catchment routes runoff to the outlet instantaneously without explicitly considering the residence time due to river routing that can be on the order of weeks to months [1]. A quarterly average helps mitigate the phase difference between the *in situ* measurements and model estimates without altering the volumetric flow. The location information of stations used in this study, along with the R values computed between three-month average model estimates and *in situ* measurements at these stations, are listed in Table 4.4.

Overall, runoff estimates provided by all simulations (both the OL and DA) are not consistent with the GRDC *in situ* measurements at most stations in terms of seasonal variation. The mediocre agreement between the simulations and the *in situ* measurements can be largely explained, in part, by the strong regulation of rivers in the study area that are extensively dammed, which is not explicitly modeled in Catchment. In addition, errors in the boundary conditions (i.e., precipitation) in conjunction with model structure error further exacerbate the differences between

Table 4.4: Correlation coefficient, R, for three-month averaged runoff estimates derived from the OL, GPS DA, and GRACE DA simulations relative to the *in situ* GRDC measurements. Bold font indicates statistically significant differences between the OL and DA at a 5% significance level.

Watershed	Station ID	Lat (°N)	Lon (°W)	Upland Area (km^2)	R		
					OL	GPS DA	GRACE DA
Great Basin	4118110	39.15	119.10	6734	0.33	0.32	0.39
	4118410	40.61	116.20	13088	0.59	**0.47**	0.60
	4118850	39.37	112.04	13377	-0.15	**0.05**	-0.09
	4118440	40.69	118.20	40155	0.39	0.31	0.39
Upper Colorado	4152553	42.19	110.16	10127	0.30	0.22	0.31
	4152620	38.09	110.41	10772	0.71	**0.37**	0.71
	4152650	38.75	108.08	14577	0.47	0.47	0.52
	4152600	37.15	109.86	59570	0.38	0.26	0.45
	4152550	38.99	110.15	116160	0.52	0.50	0.55
	4152450	36.86	111.59	289562	-0.06	**0.20**	0.03

modeled and measured runoff.

GRACE DA provides slightly larger R values when compared against the *in situ* measurements relative to the OL, but these R values are not statistically different from the OL simulation using a significance level of 5%. GPS DA, in general, shows the lowest consistency with the *in situ* measurements, yielding smaller R values. But it is also found that at the two stations (i.e., stations 4118850 and 4152450) where the OL provides negative correlation coefficient with *in situ* measurements, GPS DA shows positive R values although the R values are still small. A closer look at the runoff time series at station 4152450, which is located at the outlet of the Upper Colorado and downstream of the Glen Canyon Dam, the seasonal variation in runoff can be detected, but is not clearly evident given the strong river regulation applied. Considering the R values at stations with different upland areas, unfortunately, there is no clear pattern that clarifies the impact of the size of the upland

contributing area on the estimation accuracy of runoff.

4.3.5 Analysis of Assimilation Increments

Investigating the analysis increments (i.e., difference between posterior and prior in Equation 4.2) provides valuable information for understanding when and how much water is added to (or subtracted from) the state variables during the data assimilation update. In this study, only *catdef* and *swe* are explicitly updated. As a result, the watershed-averaged, monthly increments applied to these two variables are analyzed here (Figure 4.9). GPS DA generally shows larger temporal variations in the increments as well as larger amplitudes in comparison with GRACE DA for both *catdef* and *swe*. It is revealed that GPS DA tends to add SWE during the winter season in both the Great Basin and Upper Colorado watersheds whereas GRACE DA, in general, removes mass by subtracting SWE during most winter seasons. As discussed in Section 4.3.2, the SWE estimates derived from GPS DA usually contains a larger amplitude than the OL or GRACE DA. Notably, GPS DA and GRACE DA also exhibit different behaviors in the *catdef* increments that control the amount of mass in the subsurface. GPS DA, in general, applied a negative increment to *catdef* (i.e., added mass) during the winter season, thereby increasing the subsurface water storage during the winter relative to the OL. On the contrary, GRACE DA shows a positive or near-zero *catdef* increment during most winter seasons, which resulted in decreasing the subsurface water storage. Additionally, both GPS DA and GRACE DA applied positive increments to *catdef*

during the no-snow season in most years, but GPS DA shows a two or three month earlier application of these positive increments. As is clearly shown in Figure 4.9, most of the analysis increments are directed toward *catdef*, which is intuitive given that subsurface storage is by far the largest contributor to TWS when viewed across the entire study period and entire study domain.

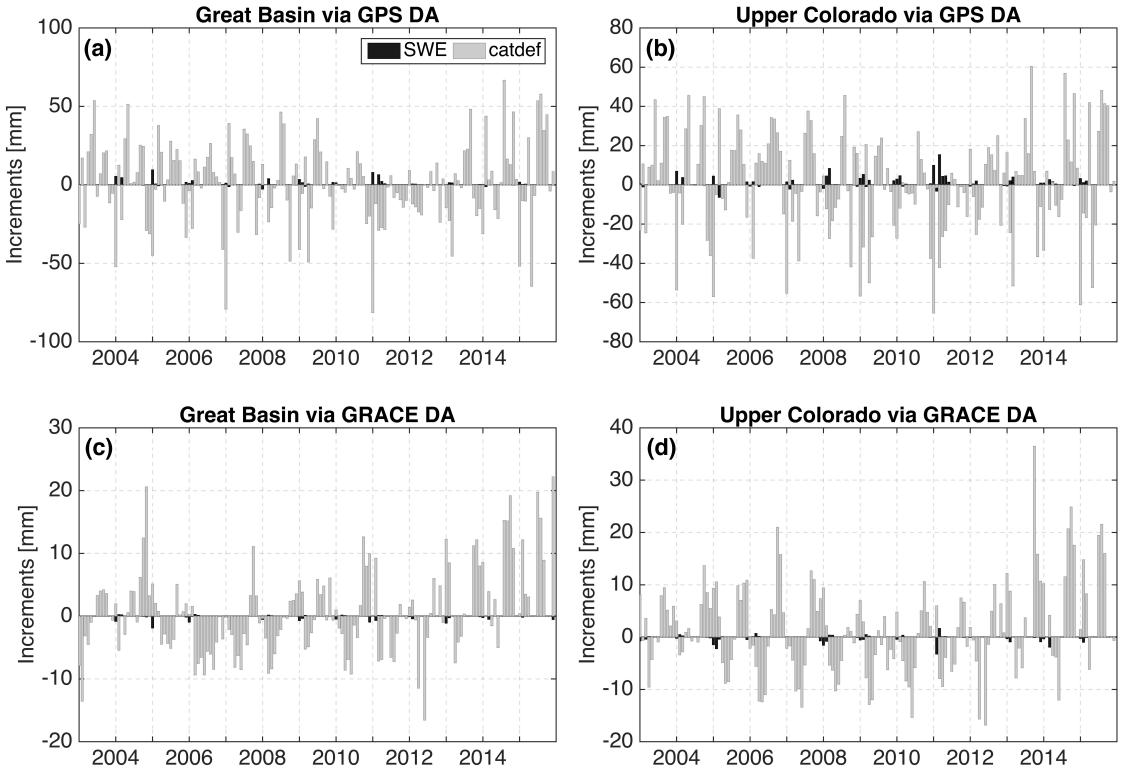

Figure 4.9: Watershed-averaged, monthly increments for state variables *catdef* (gray) and SWE (black) in the (a) Great Basin watershed via GPS DA; (b) Upper Colorado watershed via GPS DA; (c) Great Basin watershed via GRACE DA; and (d) Upper Colorado watershed via GRACE DA.

Recall that in the GPS DA case, ground-based GPS observations of vertical displacement were used during assimilation. Alternatively in the GRACE DA case, the observation used for the model update was a satellite-derived GRACE TWS retrieval. Examining the Kalman gain, K, computed for *catdef* and SWE at each

pixel, GPS DA generally provides a positive K for $catdef$ and a negative K for SWE, which is in concordance with the physical relationship between the errors in water mass changes and the errors in vertical displacement. As for GRACE DA, a physically intuitive positive cross-correlation between the errors in SWE and TWS and a negative cross-correlation between the errors in $catdef$ and TWS are detected. Based on results shown in Chapter 3, GPS-based observations generally provide the largest amplitude while the GRACE-based retrievals exhibited the smallest amplitude. During the winter season, the observed vertical displacement is more negative than the model prediction, which results in a negative innovation, and thus, a positive SWE increments given the negative relationship manifested in the Kalman gain. As for GRACE DA, GRACE-based TWS is usually smaller in magnitude than in the model predictions, and thus, also results in a negative innovation; however, considering the positive Kalman gain as generally computed, a negative increment (i.e., $K_t[y_t + v^i - H(x_t^{i-})]$) was ultimately added to the posterior model estimates. The different signs of $catdef$ increments derived from GPS DA and GRACE DA can be similarly explained. The $catdef$ increment phase difference between GPS DA and GRACE DA is related to the phase difference in the seasonal, minimum TWS as reflected by GPS and GRACE. Chapter 3 compared vertical displacements derived between ground-based GPS and GRACE TWS at GPS stations located in this study area, and it was found that GPS-based vertical displacements exhibited larger variations during the seasonal minimum period from August to October while the seasonal minimum period in the GRACE TWS retrievals generally occurred between September and October.

The time-integrated, watershed-averaged increments were computed for the Great Basin and Upper Colorado watersheds, respectively, as $\sum_{t=1}^{N_t}(\frac{1}{N_p}\sum_{p=1}^{N_p} incr_{p,t})$, where $incr$ is the increment for $catdef$ or SWE and N_p is the number of pixels located inside the watershed. GPS DA exhibited positive SWE and $catdef$ increments in both the Great Basin (57 mm for SWE; 94 mm for $catdef$) and Upper Colorado (94 mm for SWE; 77 mm for $catdef$) watersheds. GRACE DA, on the other hand, collectively removed SWE in both the Great Basin (-10 mm) and Upper Colorado (-16 mm) watersheds and increased $catdef$ (i.e., reduced subsurface water storage) by 89 mm and 207 mm in the Great Basin and Upper Colorado watersheds, respectively. If one only considers the drought period after 2011, both GPS DA and GRACE DA effectively reduced the total water storage by -89 mm and -177 mm, respectively, in the Great Basin watershed as compared to the OL. The corresponding change in total water storage for the Upper Colorado watershed was -100 mm (GPS DA) and -222 mm (GRACE DA).

4.3.6 Normalized Innovation Sequence

The normalized innovation (NI) sequence is used here to investigate the applicability of the assigned error parameters used with the assimilated observations in this study. Assuming that Catchment and the observation operator are both linear (which is not the case for Catchment), and that all errors are mutually independent and Gaussian (which may or may not be true), the NI sequence should possess a standard Gaussian distribution with zero mean and a standard deviation of one if

the assimilation was optimal in a minimized variance sense [154]. In this study, given that these assumptions are not completely fulfilled, the investigation of the NI sequence can still provide useful insight into the efficacy of the GPS DA procedure [1, 143]. Using the ensemble mean estimates as x_t^-, the watershed averaged NI was computed as a function of time via

$$NI_t = \frac{y_t - Hx_t^-}{\sqrt{Cov[H(x_t^-), H(x_t^-)] + C_{vv}}} \tag{4.6}$$

where the numerator is the innovation representing the difference between the observations and the predicted (modeled) observations, and the denominator is the square root of the combination of the observation prediction error covariance (background error) and observation error covariance. Afterwards, the time-averaged mean ($\overline{NI}$) and temporal standard deviation of NI (σ_{NI}) were computed using the watershed-averaged NI sequence. By testing different values of the observation error standard deviation used during the GPS DA simulations (varying from 2 mm to 6 mm), the corresponding $\overline{NI}$ and σ_{NI} were computed for the Great Basin and Upper Colorado watersheds as illustrated in Figure 4.10.

The most notable feature to emerge from Figure 4.10 is the positive $\overline{NI}$ produced by all tests using different observation error standard deviations. The positive $\overline{NI}$ suggests that model predictions contain a negative bias relative to the ground-based GPS observations. Considering the σ_{NI}, when increasing the observation error standard deviation, a decreased spread of NI was found with a smaller σ_{NI} as expected. That is, as more noise is embedded in the GPS measurements during

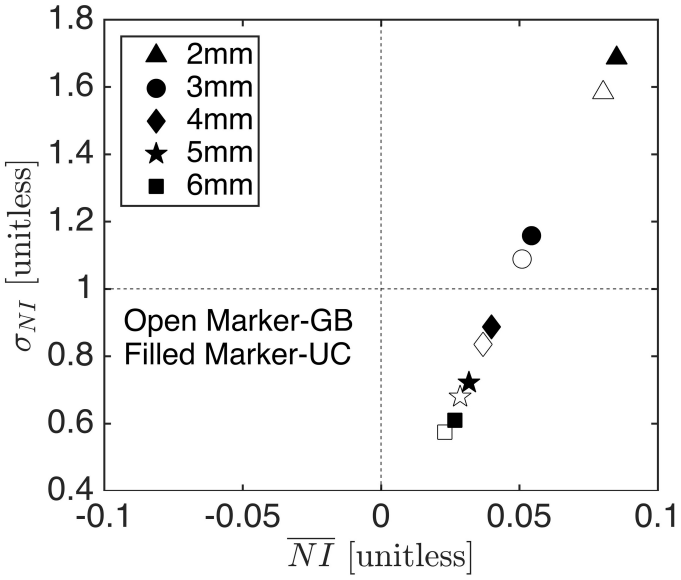

Figure 4.10: GPS DA innovation statistics for the Great Basin (open marker) and Upper Colorado (filled marker) watersheds. Different marker shapes represent different observation error standard deviations varying from 2 mm to 6 mm.

the update, the assimilation procedure relies less heavily on the observation information content, and in turn, yields a less variable posterior estimate. Figure 4.10 suggests that an observation error standard deviation of 6 mm may be too large thereby causing too small of a weight given to the observations during the update step. Conversely, an observation error standard deviation of 2 mm may be too small, which yields too large of an update toward the GPS measurements. An error observation standard deviation between 3 mm and 5 mm is suggested here as the most appropriate selection based on the results of the innovation sequence analysis.

4.4 Conclusions

Ground-based GPS provides a geodetic measurement of vertical displacement that is indirectly related to terrestrial water storage. Based on the elastic response

of the Earth′s surface to mass redistribution, ground-based GPS was used to estimate TWS changes with unique spatial and temporal resolution characteristics. This study was the first known attempt to assimilate spatially-interpolated, vertical displacements derived from ground-based GPS observations into an advanced land surface model in order to provide more accurate estimates of TWS and its constituent components. GPS DA provides an unique opportunity to merge sub-monthly localized information of TWS changes detected by GPS observations into land surface models, which is not currently feasible using GRACE TWS retrievals.

GPS DA results were evaluated by using satellite-based retrievals of TWS and *in situ* measurements of SWE, soil moisture, and runoff. It was found that GPS DA can successfully enhance TWS anomaly estimates (relative to the OL) when compared against GRACE TWS retrievals, especially during the drought period after 2011 when Catchment failed to capture the full dynamic range in TWS. GPS DA also provides improved SWE estimates when compared against SNOTEL-based SWE measurements with approximately 76% and 69% of all pixels collocated with SNOTEL stations showing improved R values in the Great Basin and Upper Colorado watersheds, respectively. The corresponding percentages of improved pixels are 81% and 92% in terms of ubRMSD. The improved agreement between GPS DA SWE estimates and SNOTEL-based SWE is generally found in regions with deep snow where orographic precipitation effects are most pronounced.

Despite the encouraging results shown for TWS anomaly and SWE estimates, GPS DA reveals reduced skill in surface soil moisture and runoff estimates. For surface soil moisture, the reduced performance can be explained by the limited

sensitivity of GPS observations of vertical displacement to changes in the top 5 cm of soil moisture content, especially when the GPS station is located far from the location where the surface soil moisture is measured. In terms of runoff, both the OL and DA simulations exhibited limited skill when compared against *in situ* measurements of river discharge. This is, in part, due to river regulation that is not explicitly modeled in Catchment. A lack of dynamic surface water routing and other model structure errors in Catchment along with errors and uncertainty in the meteorological boundary conditions (i.e., precipitation) also contribute to the relatively poor agreement. Additionally, the density and distribution of the GPS stations also influences the TWS change information implicit in the spatially-interpolated vertical displacements. Both the surface soil moisture and runoff comparisons suggest the need for a dense, ground-based GPS network in order to achieve more accurate results.

Due to the existence of a deep groundwater table [144] in the study domain coupled with the limited capability of Catchment in modeling deep groundwater [14], analyses of modeled groundwater (and modeled root zone soil moisture that is closely linked to the modeled groundwater) were not conducted in this study. More work towards evaluating GPS DA performance of root zone soil moisture and groundwater estimates is needed in the future. Additionally, the impact of the density and distribution of GPS stations on the spatially interpolated vertical displacement observations, and on the GPS DA results themselves, needs further exploration to provide more complete guidance on the use of ground-based GPS observations during assimilation.

Despite these potential pitfalls, the overall performance of GPS DA is encouraging given the improved model estimates of TWS anomalies and SWE. The improvement in modeled SWE when compared against point-scale measurements of SNOTEL SWE demonstrated the advantages of ground-based GPS observations in representing localized water mass variations. The opposite behavior, yet potentially complementary, witnessed between GRACE DA and GPS DA during the comparison of *in situ* measurements of SWE and runoff highlights the potential to simultaneously merge both GPS-based vertical displacements and GRACE-based TWS anomalies into a land surface model. Such a multi-variate assimilation framework could further improve the accuracy of modeled TWS and its constituent components while also reducing TWS uncertainty.

Chapter 5: Overall Conclusions and Future Work

5.1 Overall Conclusions

This dissertation studied terrestrial water storage (TWS) using ground-based GPS geodetic observations, space-based GRACE satellite retrievals, and an advanced land surface model. The scientific questions answered in this work include: (1) Can geodetic observations (i.e., ground-based GPS observations of vertical displacement) be used to reflect TWS changes based on the elastic response of the Earth′s surface to water mass redistribution? (2) Does the merger of ground-based GPS observations into an advanced land surface model improve the accuracy of TWS estimates as well as reduce TWS uncertainty.

In Chapter 3, the potential of using ground-based GPS observations of vertical displacement to represent TWS changes was demonstrated. Non-hydrologic loading effects (e.g., tectonic motion and atmospheric loading) were first removed from the GPS observations in order to ensure that the variations in ground-based GPS observations are only caused by TWS changes. Afterwards, a comparison of vertical displacements derived from GPS, GRACE, and Catchment was conducted. Results showed a good consistency between the vertical displacements derived from the three different data sets in terms of seasonal variation. Noticeably, a large in-

crease in TWS in early 2011 followed by an extended drought period was reflected by vertical displacement time series derived from GPS and GRACE, but not by the vertical displacement time series derived from Catchment. The large increase in TWS was caused by the heavy precipitation from late 2010 to early 2011 associated with a La Niña event. These findings suggested that ground-based GPS can be used to study TWS changes, and can help compensate for the limited capability of land surface models in accurately representing a large dynamic change in TWS.

With the results derived in Chapter 3, Chapter 4 systematically merged the ground-based GPS vertical displacements into an advanced land surface model (i.e., NASA Catchment) using a one-dimensional EnKF. Estimated TWS and its constituent components (i.e., SWE and surface soil moisture) were evaluated against GRACE TWS retrievals and *in situ* measurements. Additionally, GPS DA performance in estimating hydrologic fluxes (i.e., runoff) was also examined. Results showed that GPS DA improved the accuracy of estimated TWS anomalies, especially during an extended drought period after mid-2011. GPS DA also yielded improved SWE estimates relative to OL and GRACE DA with the greatest improvements in agreement between the SWE estimates and the *in situ* measurements generally found in regions with deep snow where orographic precipitation effects are most pronounced. However, GPS DA subsequently degraded the performance in estimating surface soil moisture and runoff. The possible explanation for the reduced skill is the limited sensitivity of GPS observations to water changes in the top 5 cm soil layer as well as a lack of sufficient spatial density in the GPS network in this study area.

In summary, this dissertation demonstrated that ground-based GPS vertical displacement can be used as an effective data source to study TWS changes. Systematically integrating the GPS observations into a land surface model using a data assimilation framework can harness the information content in the GPS observations and, in turn, provide better estimates of TWS. The 2017 Decadal Survey for Earth Science and Applications from Space identified the quantitative improvement of our understanding of the hydrological cycle and water resource characterization as a key scientific objective. The research discussed in this dissertation can help contribute to achieving this objective while enhancing our knowledge of the hydrological cycle at basin and sub-basin spatial scales. Merging multiple sources of information from ground-based GPS, GRACE, and a land surface model can be used for broad-scale studies such as drought and large-scale hurricanes, as well as local-scale studies such as reservoir storage monitoring. Furthermore, the DA approach can effectively downscale TWS information from the ground-based GPS and GRACE data into TWS constituent components such as snow, soil moisture, and shallow groundwater, as well as hydrological fluxes such as runoff, which provides discretized information for water resource planning and management.

5.2 Future Work

5.2.1 Using GPS as an Independent Measurement to Estimate TWS Change

Ground-based GPS observations can be used to represent TWS changes after removing non-hydrologic effects. With a relatively dense GPS network, GPS observations can be mapped into TWS space based on Green′s function that quantitatively explains the elastic response of the surface to mass redistribution. Future work could investigate the potential of using ground-based GPS to estimate TWS at a finer spatial and temporal resolution relative to GRACE. The impact of the density and distribution of GPS networks could be explored using a synthetic experiment. The anticipated results will provide an opportunity to further study TWS changes at sub-regional scales.

5.2.2 Robustness Experiment of a GPS DA Framework

The performance of GPS DA is influenced by a variety of factors, including the GPS observations used during the update step and the establishment of observation errors and model parameters. Therefore, the robustness of the GPS DA system could be examined by (1) applying GPS DA to different regions of the globe using a variety of different GPS network densities and distributions; (2) altering the spatial interpolation method used in this study (e.g., kriging versus IDW); and (3) tuning the parameters related to the model and observation errors. The DA results could

be evaluated using satellite observations and *in situ* measurements as done in this dissertation, or could be explored as part of a synthetic experiment.

5.2.3 Multivariate Data Assimilation for Estimating TWS

Assimilating individual observations into a land surface model can improve the estimates of some hydrologic variables, but at the same time, may degrade the performance of other variables. For example, the assimilation of soil moisture information from L-band Soil Moisture and Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) satellites improved modeled surface soil moisture but degraded groundwater estimates [155] whereas GRACE DA only enhanced groundwater estimates [3]. Results presented in this dissertation showed that GPS DA improved SWE, but degraded runoff estimates. Considering the benefits and pitfalls of univariate data assimilation, future work aimed at improving land surface models could be explored by integrating multi-sensor observations into models. TWS retrievals from GRACE and GRACE-FO missions, vertical displacements from ground-based GPS, and soil moisture retrievals from SMOS and SMAP missions could be used simultaneously in a multivariate DA framework. It is anticipated that multivariate data assimilation can take advantage of each type of data as well as mitigate the potential weaknesses introduced by each platform thereby leading to a more robust DA system.

Bibliography

[1] Barton A Forman, Rolf H Reichle, and Matthew Rodell. Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resources Research*, 48(1), 2012.

[2] Shin-Chan Han. Elastic deformation of the Australian continent induced by seasonal water cycles and the 2010–2011 La Niña determined using GPS and GRACE. *Geophysical Research Letters*, 44(6):2763–2772, 2017.

[3] Manuela Girotto, Gabriëlle JM De Lannoy, Rolf H Reichle, and Matthew Rodell. Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, 52(5):4164–4183, 2016.

[4] Tajdarul H Syed, James S Famiglietti, Matthew Rodell, Jianli Chen, and Clark R Wilson. Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resources Research*, 44(2), 2008.

[5] Wondwosen M Seyoum, Dongjae Kwon, and Adam M Milewski. Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system. *Remote Sensing*, 11(7):824, 2019.

[6] Clara C Chew and Eric E Small. Terrestrial water storage response to the 2012 drought estimated from GPS vertical position anomalies. *Geophysical Research Letters*, 41(17):6145–6151, 2014.

[7] Eric F Wood, Joshua K Roundy, Tara J Troy, LPH Van Beek, Marc FP Bierkens, Eleanor Blyth, Ad de Roo, Petra Döll, Mike Ek, James Famiglietti, et al. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research*, 47(5), 2011.

[8] Jesper Overgaard, Dan Rosbjerg, and MB Butts. Land-surface modelling in hydrological perspective? a review. 2006.

[9] Guo-Yue Niu, Zong-Liang Yang, Robert E Dickinson, Lindsey E Gulden, and Hua Su. Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. *Journal of Geophysical Research: Atmospheres*, 112(D7), 2007.

[10] Matthew Rodell and James S Famiglietti. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, 37(5):1327–1339, 2001.

[11] Byron D Tapley, Srinivas Bettadpur, John C Ries, Paul F Thompson, and Michael M Watkins. GRACE measurements of mass variability in the Earth system. *Science*, 305(5683):503–505, 2004.

[12] Sean Swenson, Don Chambers, and John Wahr. Estimating geocenter variations from a combination of GRACE and ocean model output. *Journal of Geophysical Research: Solid Earth*, 113(B8), 2008.

[13] John Wahr, Sean Swenson, and Isabella Velicogna. Accuracy of GRACE mass estimates. *Geophysical Research Letters*, 33(6), 2006.

[14] Rasmus Houborg, Matthew Rodell, Bailing Li, Rolf H Reichle, and Benjamin F Zaitchik. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, 48(7), 2012.

[15] Di Long, Bridget R Scanlon, Laurent Longuevergne, Alexander Y Sun, D Nelun Fernando, and Himanshu Save. GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. *Geophysical Research Letters*, 40(13):3395–3401, 2013.

[16] Jianli L Chen, Clark R Wilson, and Byron D Tapley. The 2009 exceptional amazon flood and interannual terrestrial water storage change observed by GRACE. *Water Resources Research*, 46(12), 2010.

[17] JT Reager and JS Famiglietti. Global terrestrial water storage capacity and flood potential using GRACE. *Geophysical Research Letters*, 36(23), 2009.

[18] James S Famiglietti, Minhui Lo, Sing L Ho, James Bethune, KJ Anderson, Tajdarul H Syed, Sean C Swenson, Caroline R de Linage, and Matthew Rodell. Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, 38(3), 2011.

[19] Katalyn A Voss, James S Famiglietti, MinHui Lo, Caroline De Linage, Matthew Rodell, and Sean C Swenson. Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water resources research*, 49(2):904–914, 2013.

[20] Isabella Velicogna. Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE. *Geophysical Research Letters*, 36(19), 2009.

[21] Scott B Luthcke, H Jay Zwally, Waleed Abdalati, DD Rowlands, RD Ray, RS Nerem, FG Lemoine, JJ McCarthy, and DS Chinn. Recent Greenland ice mass loss by drainage system from satellite gravity observations. *Science*, 314(5803):1286–1289, 2006.

[22] Donald F Argus, Felix W Landerer, David N Wiese, Hilary R Martens, Yuning Fu, James S Famiglietti, Brian F Thomas, Thomas G Farr, Angelyn W Moore, and Michael M Watkins. Sustained water loss in California's mountain ranges during severe drought from 2012 to 2015 inferred from GPS. *Journal of Geophysical Research: Solid Earth*, 122(12), 2017.

[23] Tonie van Dam, J Wahr, and David Lavallée. A comparison of annual vertical crustal displacements from GPS and Gravity Recovery and Climate Experiment (GRACE) over Europe. *Journal of Geophysical Research: Solid Earth*, 112(B3), 2007.

[24] Samuel Nahmani, Olivier Bock, Marie-Noëlle Bouin, Alvaro Santamaría-Gómez, Jean-Paul Boy, Xavier Collilieux, Laurent Métivier, Isabelle Panet, Pierre Genthon, Caroline Linage, et al. Hydrological deformation induced by the West African Monsoon: Comparison of GPS, GRACE and loading models. *Journal of Geophysical Research: Solid Earth*, 117(B5), 2012.

[25] Yuning Fu, Jeffrey T Freymueller, and Tim Jensen. Seasonal hydrological loading in southern Alaska observed by GPS and GRACE. *Geophysical Research Letters*, 39(15), 2012.

[26] Yuning Fu and Jeffrey T Freymueller. Seasonal and long-term vertical deformation in the Nepal Himalaya constrained by GPS and GRACE measurements. *Journal of Geophysical Research: Solid Earth*, 117(B3), 2012.

[27] Yuning Fu, Donald F Argus, and Felix W Landerer. GPS as an independent measurement to estimate terrestrial water storage variations in Washington and Oregon. *Journal of Geophysical Research: Solid Earth*, 120(1):552–566, 2015.

[28] Roelof Rietbroek, Mathias Fritsche, Christoph Dahle, Sandra-Esther Brunnabend, Madlen Behnisch, Jürgen Kusche, Frank Flechtner, Jens Schröter, and Reinhard Dietrich. Can GPS-derived surface loading bridge a GRACE mission gap? *Surveys in Geophysics*, 35(6):1267–1283, 2014.

[29] G Ramillien, James S Famiglietti, and J Wahr. Detection of continental hydrology and glaciology signals from GRACE: a review. *Surveys in geophysics*, 29(4-5):361–374, 2008.

[30] K-W Seo, CR Wilson, JS Famiglietti, JL Chen, and M Rodell. Terrestrial water mass load changes from Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, 42(5), 2006.

[31] John Wahr, Mery Molenaar, and Frank Bryan. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth*, 103(B12):30205–30229, 1998.

[32] Scott B. Luthcke, T.J. Sabaka, B.D. Loomis, A.A. Arendt, J.J. McCarthy, and J. Camp. Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. *Journal of Glaciology*, 59(216):613–631, 2013.

[33] Bridget R Scanlon, Zizhan Zhang, Himanshu Save, David N Wiese, Felix W Landerer, Di Long, Laurent Longuevergne, and Jianli Chen. Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resources Research*, 52(12):9412–9429, 2016.

[34] Felix W Landerer and SC Swenson. Accuracy of scaled GRACE terrestrial water storage estimates. *Water resources research*, 48(4), 2012.

[35] Stuart B Andrews, Philip Moore, and Matt A King. Mass change from GRACE: a simulated comparison of Level-1B analysis techniques. *Geophysical Journal International*, 200(1):503–518, 2014.

[36] M Rodell and JS Famiglietti. Detectability of variations in continental water storage from satellite observations of the time dependent gravity field. *Water Resources Research*, 35(9):2705–2723, 1999.

[37] M Rodell and JS Famiglietti. The potential for satellite-based monitoring of groundwater storage changes using GRACE: the High Plains aquifer, Central US. *Journal of Hydrology*, 263(1-4):245–256, 2002.

[38] Di Long, Yun Pan, Jian Zhou, Yang Chen, Xueyan Hou, Yang Hong, Bridget R Scanlon, and Laurent Longuevergne. Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. *Remote sensing of environment*, 192:198–216, 2017.

[39] Matthew Rodell, Isabella Velicogna, and James S Famiglietti. Satellite-based estimates of groundwater depletion in India. *Nature*, 460(7258):999, 2009.

[40] Di Long, Xi Chen, Bridget R Scanlon, Yoshihide Wada, Yang Hong, Vijay P Singh, Yaning Chen, Cunguang Wang, Zhongying Han, and Wenting Yang. Have GRACE satellites overestimated groundwater depletion in the Northwest India Aquifer? *Scientific reports*, 6:24398, 2016.

[41] Jianli Chen, Jin Li, Zizhan Zhang, and Shengnan Ni. Long-term groundwater variations in Northwest India from satellite gravity measurements. *Global and Planetary Change*, 116:130–138, 2014.

[42] Wei Feng, Min Zhong, Jean-Michel Lemoine, Richard Biancale, Hou-Tse Hsu, and Jun Xia. 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):2110–2118, 2013.

[43] Marc J Leblanc, Paul Tregoning, Guillaume Ramillien, Sarah O Tweed, and Adam Fakes. Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia. *Water resources research*, 45(4), 2009.

[44] Bridget R Scanlon, Laurent Longuevergne, and Di Long. Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *Water Resources Research*, 48(4), 2012.

[45] JL Chen, CR Wilson, and BD Tapley. Satellite gravity measurements confirm accelerated melting of Greenland ice sheet. *science*, 313(5795):1958–1960, 2006.

[46] Isabella Velicogna and John Wahr. Measurements of time-variable gravity show mass loss in Antarctica. *science*, 311(5768):1754–1756, 2006.

[47] JL Chen, CR Wilson, D Blankenship, and BD Tapley. Accelerated Antarctic ice loss from satellite gravity measurements. *Nature Geoscience*, 2(12):859, 2009.

[48] Xi Chen, Di Long, Yang Hong, Chao Zeng, and Denghua Yan. Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the Upper Brahmaputra River basin? *Water Resources Research*, 53(3):2431–2466, 2017.

[49] Mohamed Ahmed, Mohamed Sultan, John Wahr, and Eugene Yan. The use of GRACE data to monitor natural and anthropogenic induced variations in water availability across Africa. *Earth-Science Reviews*, 136:289–300, 2014.

[50] Matthew Rodell, JS Famiglietti, DN Wiese, JT Reager, HK Beaudoing, Felix W Landerer, and M-H Lo. Emerging trends in global freshwater availability. *Nature*, 557(7707):651, 2018.

[51] Sitotaw Z Yirdaw, Kenneth R Snelgrove, and Clement O Agboma. GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian Prairie. *Journal of Hydrology*, 356(1-2):84–92, 2008.

[52] M Rodell, JS Famiglietti, Jianli Chen, SI Seneviratne, P Viterbo, S Holl, and CR Wilson. Basin scale estimates of evapotranspiration using GRACE and other observations. *Geophysical Research Letters*, 31(20), 2004.

[53] Ali Behrangi, Alex S Gardner, John T Reager, and Joshua B Fisher. Using GRACE to constrain precipitation amount over cold mountainous basins. *Geophysical Research Letters*, 44(1):219–227, 2017.

[54] Randal D Koster, Max J Suarez, Agnès Ducharne, Marc Stieglitz, and Praveen Kumar. A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of Geophysical Research: Atmospheres*, 105(D20):24809–24822, 2000.

[55] Guo-Yue Niu, Zong-Liang Yang, Kenneth E Mitchell, Fei Chen, Michael B Ek, Michael Barlage, Anil Kumar, Kevin Manning, Dev Niyogi, Enrique Rosero, et al. The community noah land surface model with multiparameterization options (noah-mp): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, 116(D12), 2011.

[56] Matthew Rodell, PR Houser, UEA Jambor, J Gottschalck, K Mitchell, C-J Meng, K Arsenault, B Cosgrove, J Radakovich, M Bosilovich, et al. The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85(3):381–394, 2004.

[57] Randal D Koster and Max J Suarez. Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *Journal of Geophysical Research: Atmospheres*, 97(D3):2697–2715, 1992.

[58] MB Ek, KE Mitchell, Y Lin, E Rogers, P Grunmann, V Koren, G Gayno, and JD Tarpley. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research: Atmospheres*, 108(D22), 2003.

[59] Yongjiu Dai, Xubin Zeng, Robert E Dickinson, Ian Baker, Gordon B Bonan, Michael G Bosilovich, A Scott Denning, Paul A Dirmeyer, Paul R Houser, Guoyue Niu, et al. The common land model. *Bulletin of the American Meteorological Society*, 84(8):1013–1024, 2003.

[60] Xu Liang, Dennis P Lettenmaier, Eric F Wood, and Stephen J Burges. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, 99(D7):14415–14428, 1994.

[61] X Mo, JJ Wu, Q Wang, and H Zhou. Variations in water storage in China over recent decades from GRACE observations and GLDAS. *Natural Hazards and Earth System Sciences*, 16(2):469–482, 2016.

[62] Vagner G Ferreira, Christopher E Ndehedehe, Henry C Montecino, Bin Yong, Peng Yuan, Ahmed Abdalla, and Abubakar S Mohammed. Prospects for imaging terrestrial water storage in South America using daily GPS observations. *Remote Sensing*, 11(6):679, 2019.

[63] Balázs M Fekete, Charles J Vörösmarty, John O Roads, and Cort J Willmott. Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate*, 17(2):294–304, 2004.

[64] Jon Gottschalck, Jesse Meng, Matt Rodell, and Paul Houser. Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. *Journal of Hydrometeorology*, 6(5):573–598, 2005.

[65] Bailing Li, Matthew Rodell, Benjamin F Zaitchik, Rolf H Reichle, Randal D Koster, and Tonie M van Dam. Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, 446:103–115, 2012.

[66] Paul A Dirmeyer, Randal D Koster, and Zhichang Guo. Do global models properly represent the feedback between land and atmosphere? *Journal of Hydrometeorology*, 7(6):1177–1198, 2006.

[67] Agnès Ducharne, Randal D Koster, Max J Suarez, Marc Stieglitz, and Praveen Kumar. A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. *Journal of Geophysical Research: Atmospheres*, 105(D20):24823–24838, 2000.

[68] Benjamin F Zaitchik, Matthew Rodell, and Rolf H Reichle. Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin. *Journal of Hydrometeorology*, 9(3):535–548, 2008.

[69] Youlong Xia, David Mocko, Maoyi Huang, Bailing Li, Matthew Rodell, Kenneth E Mitchell, Xitian Cai, and Michael B Ek. Comparison and assessment of three advanced land surface models in simulating terrestrial water storage components over the United States. *Journal of Hydrometeorology*, 18(3):625–649, 2017.

[70] WE Farrell. Deformation of the Earth by surface loads. *Reviews of Geophysics*, 10(3):761–797, 1972.

[71] John Wahr, Shfaqat A Khan, Tonie van Dam, Lin Liu, Jan H van Angelen, Michiel R van den Broeke, and Charles M Meertens. The use of GPS horizontals for loading studies, with applications to northern California and southeast Greenland. *Journal of Geophysical Research: Solid Earth*, 118(4):1795–1806, 2013.

[72] J Ray, Z Altamimi, X Collilieux, and Tonie van Dam. Anomalous harmonics in the spectra of GPS position estimates. *GPS solutions*, 12(1):55–64, 2008.

[73] JL Davis, P Elósegui, JX Mitrovica, and ME Tamisiea. Climate-driven deformation of the solid Earth from GRACE and GPS. *Geophysical Research Letters*, 31(24), 2004.

[74] Matt King, Philip Moore, Peter Clarke, and David Lavallée. Choice of optimal averaging radii for temporal GRACE gravity solutions, a comparison with GPS and satellite altimetry. *Geophysical Journal International*, 166(1):1–11, 07 2006.

[75] Paul Tregoning, C Watson, Guillaume Ramillien, Herbert McQueen, and Jianjun Zhang. Detecting hydrologic deformation using GRACE and GPS. *Geophysical Research Letters*, 36(15), 2009.

[76] Volker Tesmer, Peter Steigenberger, Tonie van Dam, and Torsten Mayer-Gürr. Vertical deformations from homogeneously processed GRACE and global GPS long-term series. *Journal of Geodesy*, 85(5):291–310, 2011.

[77] Pierre Bettinelli, Jean-Philippe Avouac, Mireille Flouzat, Laurent Bollinger, Guillaume Ramillien, Sudhir Rajaure, and Som Sapkota. Seasonal variations of seismicity and geodetic strain in the Himalaya induced by surface hydrology. *Earth and Planetary Science Letters*, 266(3-4):332–344, 2008.

[78] K Chanard, JP Avouac, G Ramillien, and J Genrich. Modeling deformation induced by seasonal variations of continental water in the Himalaya region: Sensitivity to Earth elastic structure. *Journal of Geophysical Research: Solid Earth*, 119(6):5097–5113, 2014.

[79] Yuning Fu, Donald F Argus, Jeffrey T Freymueller, and Michael B Heflin. Horizontal motion in elastic response to seasonal loading of rain water in the Amazon Basin and monsoon water in Southeast Asia observed by GPS and inferred from GRACE. *Geophysical Research Letters*, 40(23):6048–6053, 2013.

[80] Ajish P Saji, PS Sunil, KM Sreejith, Param K Gautam, K Vijay Kumar, M Ponraj, S Amirtharaj, Rose Mary Shaju, SK Begum, CD Reddy, et al. Surface deformation and influence of hydrological mass over Himalaya and North India revealed from a decade of continuous GPS and GRACE observations. *Journal of Geophysical Research: Earth Surface*, 2019.

[81] Hansheng Wang, Longwei Xiang, Lulu Jia, Liming Jiang, Zhiyong Wang, Bo Hu, and Peng Gao. Load love numbers and green's functions for elastic earth models prem, iasp91, ak135, and modified models with refined crustal structure from crust 2.0. *Computers & Geosciences*, 49:190–199, 2012.

[82] Adam M Dziewonski and Don L Anderson. Preliminary reference Earth model. *Physics of the earth and planetary interiors*, 25(4):297–356, 1981.

[83] BLN Kennett and ER Engdahl. Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2):429–465, 1991.

[84] Brian LN Kennett, ER Engdahl, and R Buland. Constraints on seismic velocities in the earth from traveltimes. *Geophysical Journal International*, 122(1):108–124, 1995.

[85] Yanchao Gu, Dongming Fan, and Wei You. Comparison of observed and modeled seasonal crustal vertical displacements derived from multi-institution gps and grace solutions. *Geophysical Research Letters*, 44(14):7219–7227, 2017.

[86] Thomas A Herring, Timothy I Melbourne, Mark H Murray, Michael A Floyd, Walter M Szeliga, Robert W King, David A Phillips, Christine M Puskas, Marcelo Santillan, and Lei Wang. Plate Boundary Observatory and related networks: GPS data analysis methods and geodetic products. *Reviews of Geophysics*, 54(4):759–808, 2016.

[87] Colin B Amos, Pascal Audet, William C Hammond, Roland Bürgmann, Ingrid A Johanson, and Geoffrey Blewitt. Uplift and seismicity driven by groundwater depletion in central California. *Nature*, 509(7501):483, 2014.

[88] Weijie Tan, Danan Dong, Junping Chen, and Bin Wu. Analysis of systematic differences from GPS-measured and GRACE-modeled deformation in Central Valley, California. *Advances in Space Research*, 57(1):19–29, 2016.

[89] Donald F Argus, Yuning Fu, and Felix W Landerer. Seasonal variation in total water storage in california inferred from GPS observations of vertical land motion. *Geophysical Research Letters*, 41(6):1971–1980, 2014.

[90] E Knappe, R Bendick, HR Martens, DF Argus, and WP Gardner. Downscaling vertical GPS observations to derive watershed-scale hydrologic loading in the Northern Rockies. *Water Resources Research*, 55(1):391–401, 2019.

[91] Renli Liu, Jiancheng Li, Hok Fok, CK Shum, and Zhao Li. Earth surface deformation in the north china plain detected by joint analysis of GRACE and GPS data. *Sensors*, 14(10):19861–19876, 2014.

[92] Linsong Wang, Chao Chen, Jinsong Du, and Tongqing Wang. Detecting seasonal and long-term vertical displacement in the North China Plain using GRACE and GPS. *Hydrol. Earth Syst. Sci*, 21:2905–2922, 2017.

[93] Shuya Li, Wenbin Shen, Yuanjin Pan, and Tengxu Zhang. Surface seasonal mass changes and vertical crustal deformation in North China from GPS and GRACE measurements. *Geodesy and Geodynamics*, 11(1):46–55, 2020.

[94] JEJO Kusche and EJO Schrama. Surface mass redistribution inversion from global GPS deformation and Gravity Recovery and Climate Experiment

(GRACE) gravity data. *Journal of geophysical research: solid earth*, 110(B9), 2005.

[95] Roelof Rietbroek, M Fritsche, S-E Brunnabend, I Daras, Jürgen Kusche, Jens Schröter, F Flechtner, and R Dietrich. Global surface mass from a new combination of GRACE, modelled OBP and reprocessed GPS data. *Journal of Geodynamics*, 59:64–71, 2012.

[96] Adrian Antal Borsa, Duncan Carr Agnew, and Daniel R Cayan. Ongoing drought-induced uplift in the western United States. *Science*, 345(6204):1587–1590, 2014.

[97] Shuanggen Jin and Tengyu Zhang. Terrestrial water storage anomalies associated with drought in southwestern USA from GPS observations. *Surveys in Geophysics*, 37(6):1139–1156, 2016.

[98] Bao Zhang, Yibin Yao, Hok Fok, Yufeng Hu, and Qiang Chen. Potential seasonal terrestrial water storage monitoring from GPS vertical displacements: A case study in the lower three-rivers headwater region, China. *Sensors*, 16(9):1526, 2016.

[99] Chris Milliner, Kathryn Materna, Roland Bürgmann, Yuning Fu, Angelyn W Moore, David Bekaert, Surendra Adhikari, and Donald F Argus. Tracking the weight of Hurricane Harvey's stormwater using GPS data. *Science advances*, 4(9):eaau2477, 2018.

[100] S Adusumilli, AA Borsa, MA Fish, HK McMillan, and F Silverii. A decade of water storage changes across the contiguous United States from GPS and satellite gravity. *Geophysical Research Letters*, 2019.

[101] Shin-Chan Han and S Mahdiyeh Razeghi. GPS recovery of daily hydrologic and atmospheric mass variation: A methodology and results from the Australian continent. *Journal of Geophysical Research: Solid Earth*, 122(11):9328–9343, 2017.

[102] VG Ferreira, HC Montecino, CE Ndehedehe, B Heck, Z Gong, SRC de Freitas, and M Westerhaus. Space-based observations of crustal deflections for drought characterization in Brazil. *Science of the total environment*, 644:256–273, 2018.

[103] Karli J Ouellette, Caroline de Linage, and James S Famiglietti. Estimating snow water equivalent from GPS vertical site-position observations in the western United States. *Water resources research*, 49(5):2508–2518, 2013.

[104] Thomas L Enzminger, Eric E Small, and Adrian A Borsa. Accuracy of snow water equivalent estimated from GPS vertical displacements: A synthetic loading case study for western US mountains. *Water Resources Research*, 54(1):581–599, 2018.

[105] Param K Gautam, Saksham Arora, Suresh Kannaujiya, Anjali Singh, Ajanta Goswami, and Prashant K Champati. A comparative appraisal of ground water resources using GRACE-GPS data in highly urbanised regions of Uttar Pradesh, India. *Sustainable Water Resources Management*, 3(4):441–449, 2017.

[106] Chandrakanta Ojha, Susanna Werth, and Manoochehr Shirzaei. Groundwater loss and aquifer system compaction in San Joaquin Valley during 2012–2015 drought. *Journal of Geophysical Research: Solid Earth*, 124(3):3127–3143, 2019.

[107] Barton A Forman and Rolf H Reichle. The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system. *Water Resources Research*, 49(11):7457–7468, 2013.

[108] Hua Su, Zong-Liang Yang, Robert E Dickinson, Clark R Wilson, and Guo-Yue Niu. Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. *Journal of Geophysical Research: Atmospheres*, 115(D10), 2010.

[109] Annette Eicker, Maike Schumacher, Jürgen Kusche, Petra Döll, and Hannes Müller Schmied. Calibration/data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an ensemble Kalman filter: First results. *Surveys in Geophysics*, 35(6):1285–1309, 2014.

[110] N Tangdamrongsub, SC Steele-Dunne, BC Gunter, PG Ditmar, and AH Weerts. Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin. *Hydrology and Earth System Sciences*, 19(4):2079–2100, 2015.

[111] Maike Schumacher, Jürgen Kusche, and Petra Döll. A systematic impact assessment of GRACE error correlation on data assimilation in hydrological models. *Journal of Geodesy*, 90(6):537–559, 2016.

[112] Sujay V Kumar, Benjamin F Zaitchik, Christa D Peters-Lidard, Matthew Rodell, Rolf H Reichle, Bailing Li, Michael Jasinski, David Mocko, Augusto Getirana, Gabrielle De Lannoy, et al. Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System. *Journal of Hydrometeorology*, 17(7):1951–1972, 2016.

[113] Manuela Girotto, Gabriëlle JM De Lannoy, Rolf H Reichle, Matthew Rodell, Clara Draper, Soumendra N Bhanja, and Abhijit Mukherjee. Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in India. *Geophysical research letters*, 44(9):4107–4115, 2017.

[114] M Khaki, E Forootan, Michael Kuhn, Joseph Awange, AIJM van Dijk, M Schumacher, and MA Sharifi. Determining water storage depletion within

Iran by assimilating GRACE data into the W3RA hydrological model. *Advances in water resources*, 114:1–18, 2018.

[115] Hahn Chul Jung, Augusto Getirana, Kristi R Arsenault, Sujay Kumar, and Issoufou Maigary. Improving surface soil moisture estimates in west Africa through GRACE data assimilation. *Journal of Hydrology*, 2019.

[116] M Khaki, Ibrahim Hoteit, Michael Kuhn, Joseph Awange, E Forootan, Albert IJM Van Dijk, M Schumacher, and C Pattiaratchi. Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. *Advances in Water Resources*, 107:301–316, 2017.

[117] Daniel J McEvoy, Justin L Huntington, John T Abatzoglou, and Laura M Edwards. An evaluation of multiscalar drought indices in Nevada and Eastern California. *Earth Interactions*, 16(18):1–18, 2012.

[118] Christine A Rumsey, Matthew P Miller, David D Susong, Fred D Tillman, and David W Anning. Regional scale estimates of baseflow and factors influencing baseflow in the Upper Colorado River Basin. *Journal of Hydrology: Regional Studies*, 4:91–107, 2015.

[119] Bill Fiero. *Geology of the Great Basin*. University of Nevada Press, 2009.

[120] Murray C Peel, Brian L Finlayson, and Thomas A McMahon. Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences discussions*, 4(2):439–473, 2007.

[121] Frank Flechtner, Henryk Dobslaw, and Elisa Fagiolini. AOD1B product description document for product release 05 (rev. 4.4, december 14, 2015). *Technical Note*, 2014.

[122] Leonard F Konikow. *Groundwater depletion in the United States (1900-2008)*. US Department of the Interior, US Geological Survey Reston, Virginia, 2013.

[123] BD Loomis, SB Luthcke, and TJ Sabaka. Regularization and error characterization of GRACE mascons. *Journal of Geodesy*, pages 1–18, 2019.

[124] TJ Sabaka, DD Rowlands, SB Luthcke, and J-P Boy. Improving global mass flux solutions from Gravity Recovery and Climate Experiment (GRACE) through forward modeling and continuous time correlation. *Journal of Geophysical Research: Solid Earth*, 115(B11), 2010.

[125] MK Cheng and JC Ries. Monthly estimates of C20 from 5 SLR satellites based on GRACE RL05 models, GRACE Technical Note 07. *Austin: Center for Space Research*, 2012.

[126] Ronald Gelaro, Will McCarty, Max J Suárez, Ricardo Todling, Andrea Molod, Lawrence Takacs, Cynthia A Randles, Anton Darmenov, Michael G Bosilovich, Rolf H Reichle, et al. The modern-era retrospective analysis for research and

applications, version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454, 2017.

[127] Ally M Toure, Rolf H Reichle, Barton A Forman, Augusto Getirana, and Gabrielle JM De Lannoy. Assimilation of MODIS snow cover fraction observations into the NASA catchment land surface model. *Remote sensing*, 10(2):316, 2018.

[128] Youlong Xia, Kenneth Mitchell, Michael Ek, Brian Cosgrove, Justin Sheffield, Lifeng Luo, Charles Alonge, Helin Wei, Jesse Meng, Ben Livneh, et al. Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *Journal of Geophysical Research: Atmospheres*, 117(D3), 2012.

[129] Mark Svoboda, Doug LeComte, Mike Hayes, Richard Heim, Karin Gleason, Jim Angel, Brad Rippey, Rich Tinker, Mike Palecki, David Stooksbury, et al. The drought monitor. *Bulletin of the American Meteorological Society*, 83(8):1181–1190, 2002.

[130] State of the Climate NOAA National Centers for Environmental Information. National climate report for december 2010, 2011.

[131] William C Hammond, Geoffrey Blewitt, and Corné Kreemer. GPS imaging of vertical land motion in California and Nevada: Implications for Sierra Nevada uplift. *Journal of Geophysical Research: Solid Earth*, 121(10):7681–7703, 2016.

[132] Shengnan Ni, Jianli Chen, Clark R Wilson, Jin Li, Xiaogong Hu, and Rong Fu. Global terrestrial water storage changes and connections to ENSO events. *Surveys in Geophysics*, 39(1):1–22, 2018.

[133] Shfaqat Abbas Khan, John Wahr, Michael Bevis, Isabella Velicogna, and Eric Kendrick. Spread of ice mass loss into northwest Greenland observed by GRACE and GPS. *Geophysical Research Letters*, 37(6), 2010.

[134] Martin Horwath, Axel Rülke, Mathias Fritsche, and Reinhard Dietrich. Mass variation signals in GRACE products and in crustal deformations from GPS: a comparison. In *System earth via geodetic-geophysical space techniques*, pages 399–406. Springer, 2010.

[135] Rong Zou, Qi Wang, Jeffrey Freymueller, Markku Poutanen, Xuelian Cao, Caihong Zhang, Shaomin Yang, and Ping He. Seasonal hydrological loading in southern tibet detected by joint analysis of GPS and GRACE. *Sensors*, 15(12):30525–30538, 2015.

[136] Keith J Beven and Michael J Kirkby. A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, 24(1):43–69, 1979.

[137] Marc Stieglitz, Agnès Ducharne, Randy Koster, and Max Suarez. The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *Journal of Hydrometeorology*, 2(3):228–242, 2001.

[138] Manuela Girotto, Rolf H Reichle, Matthew Rodell, Qing Liu, Sarith Mahanama, and Gabriëlle JM De Lannoy. Multi-sensor assimilation of SMOS brightness temperature and GRACE terrestrial water storage observations for soil moisture and shallow groundwater estimation. *Remote sensing of environment*, 227:12–27, 2019.

[139] Mary J Brodzik, Brendan Billingsley, Terry Haran, Bruce Raup, and Matthew H Savoie. EASE-Grid 2.0: Incremental but significant improvements for Earth-gridded data sets. *ISPRS International Journal of Geo-Information*, 1(1):32–45, 2012.

[140] Rolf H Reichle, Jeffrey P Walker, Randal D Koster, and Paul R Houser. Extended versus ensemble Kalman filtering for land data assimilation. *Journal of hydrometeorology*, 3(6):728–740, 2002.

[141] Gabriëlle JM De Lannoy, Rolf H Reichle, Paul R Houser, Kristi R Arsenault, Niko EC Verhoest, and Valentijn RN Pauwels. Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *Journal of Hydrometeorology*, 11(2):352–369, 2010.

[142] Rolf H Reichle and Randal D Koster. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters*, 32(2), 2005.

[143] Yuan Xue, Barton A Forman, and Rolf H Reichle. Estimating snow mass in North America through assimilation of Advanced Microwave Scanning Radiometer brightness temperature observations using the Catchment land surface model and support vector machines. *Water Resources Research*, 54(9):6488–6509, 2018.

[144] Ying Fan, H Li, and Gonzalo Miguez-Macho. Global patterns of groundwater table depth. *Science*, 339(6122):940–943, 2013.

[145] SC Swenson. GRACE monthly land water mass grids NETCDF RELEASE 5.0. *CA, USA*, 2012.

[146] Randal D Koster, Zhichang Guo, Rongqian Yang, Paul A Dirmeyer, Kenneth Mitchell, and Michael J Puma. On the nature of soil moisture in land surface models. *Journal of Climate*, 22(16):4322–4335, 2009.

[147] Qing Liu, Rolf H Reichle, Rajat Bindlish, Michael H Cosh, Wade T Crow, Richard de Jeu, Gabrielle JM De Lannoy, George J Huffman, and Thomas J Jackson. The contributions of precipitation and soil moisture observations to

the skill of soil moisture estimates in a land data assimilation system. *Journal of Hydrometeorology*, 12(5):750–765, 2011.

[148] Randal D Koster, Sarith PP Mahanama, Ben Livneh, Dennis P Lettenmaier, and Rolf H Reichle. Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nature Geoscience*, 3(9):613–616, 2010.

[149] Dara Entekhabi, Simon Yueh, Peggy E O'Neill, Kent H Kellogg, Angela Allen, Rajat Bindlish, Molly Brown, Steven Chan, Andreas Colliander, Wade T Crow, et al. SMAP handbook–soil moisture active passive: Mapping soil moisture and freeze/thaw from space. 2014.

[150] James H Steiger. Tests for comparing elements of a correlation matrix. *Psychological bulletin*, 87(2):245, 1980.

[151] State of the Climate NOAA National Centers for Environmental Information. State of the climate: National climate report for march 2011, 2011.

[152] Natthachet Tangdamrongsub, Shin-Chan Han, In-Young Yeo, Jianzhi Dong, Susan C Steele-Dunne, Garry Willgoose, and Jeffrey P Walker. Multivariate data assimilation of GRACE, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates. *Advances in Water Resources*, 135:103477, 2020.

[153] Jeffrey P Walker and Paul R Houser. Hydrologic data assimilation. In *Advances in water science methodologies*, pages 45–68. CRC Press, 2005.

[154] Raman Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on automatic control*, 15(2):175–184, 1970.

[155] Gabrielle De Lannoy and R Reichle. Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrology and Earth System Sciences*, 20(12):4895–4911, 2016.