IRRIGATION AND ITS RESPONSE TO CLIMATE:
IMPROVING REPRESENTATION OF HUMAN IMPACTS IN HYDROLOGICAL MODELING AND DATA ASSIMILATION SYSTEMS

by

Wanshu Nie

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

January, 2020

Abstract

Irrigation is vital for agricultural production as it resolves the temporal and spatial disconnects between water supply and water demand, especially for semi-arid and arid regions. Globally, irrigation accounts for 70% of global freshwater withdrawals, 40% of which is supplied by groundwater. In the United States, irrigation accounts for ~40% of total freshwater withdrawal and more than 80% of total freshwater consumption. Approximately 60% of irrigated areas are supplied by groundwater resources. Irrigation has led to serious aquifer depletion due to groundwater pumping for many regions around the world and has also modified the land-atmosphere interactions via the surface energy balance. Its importance in altering the water cycle and climate within the Earth system is evident, but the process, along with its impact on the water cycle and climate in the Earth System Models is not well represented.

This dissertation is motivated by the need to improve representation of irrigation in hydrological modeling and data assimilation systems that are used to study, monitor, and predict water resource dynamics. This is done through three specific objectives: (1) establishing the representation of irrigation process in a Land Surface Model (LSM) that accounts for source water partitioning; (2) applying this improved model to a land data assimilation system; (3) using the improved modeling system to study the climate sensitivity of existing irrigation developments. Three separate studies are performed using the Noah-Multiparameterization (Noah-MP) LSM within the

framework of NASA's Land Information System (LIS). The first study implements a groundwater irrigation scheme into Noah-MP and explores key factors that improve model representation of drought and groundwater depletion in the United States High Plains Aquifer (HPA) region. In the second study, terrestrial water storage (TWS) observations from the Gravity Recovery and Climate Experiment (GRACE) mission are assimilated into the model. The individual and combined effects of simulating irrigation and including GRACE data assimilation (GRACE-DA) are assessed for the HPA. The third study extends the application to the Contiguous United States (CONUS), partitions between surface water and groundwater for irrigation, and quantifies climate sensitivity of simulated water use across major irrigation zones of the CONUS.

Acknowledgements

Working towards a PhD at JHU was among the most challenging and rewarding experiences of my life. It is a process through which I realized that common sense is not all that common; that one must practice to embrace both ups and downs; and that the creation of a rich, full and meaningful life is all powerful.

First and foremost, I am sincerely grateful to my extraordinary advisor Dr. Benjamin Zaitchik, who has provided me with a constant source of encouragement, support, and sound advice. Thank you Ben, for raising me to become a scientist and for helping me to build this dream. I would not be where I am if not for you.

Endless thanks I extend to my committee members, Dr. Anand Gnanadesikan, who is also my second reader, and Dr. Darryn Waugh for useful suggestions on my research and continuous guidance along the winding path of my academic career. I would like to thank EPS faculty, in particular, Thomas Haine, who has shaped my fundamental understanding of a modeling world and helped me to transform that knowledge into my research.

I could not do anything without the help of an incredible, incomparable team at NASA, including Drs. Matthew Rodell, Sujay Kumar, Kristi Arsenault, Bailing Li and Augusto Getirana, who have helped to form ideas, solve issues in the code, read versions of the manuscripts, and provided critical feedback. I'm also grateful to Drs. Martha Anderson and Christopher Hain as well

as many others who have provided datasets, information and suggestions that were essential to this research. I appreciate all of their efforts.

In addition, I've been fortunate to have Drs. Guangheng Ni, Ting Sun, and Long Zhao, who are incredible lifelong mentors and friends. They encouraged me when I needed it, guided me when I struggled, and always inspired me to explore and identify myself. Special thanks goes out to Denise Link-Farajali, my wonderful English counselor, for the warmhearted support and encouragement.

I'm grateful to my friends and family in Baltimore, who have brought me so much joy and made my world more beautiful and interesting.

Finally an enduring thank you to my parents, for their unconditional love, as well as to my parents-in-law, for the unceasing support on our marriage. Last but not least, I offer to my husband, Yifan, sincere thanks for his abiding support and company.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The water in the snowpack, rivers, lakes, reservoirs and underground aquifers are vital to the world's ecological systems. These water bodies supply human water needs across multiple sectors, including withdrawals for agriculture, industry, residential use, public-supply, livestock, and aquaculture, among others. Irrigated agriculture, which is the sector with by far the largest consumptive water use and water withdrawals, is responsible for 70% of the global water demand, sustaining 40% of the global food production (Wada et al., 2013).

Above the ground, climate conditions play a major role in determining the distribution and volume of irrigation water applications, as irrigation resolves the temporal and spatial disconnects between water supply and water demand. For instance, for places where the water needs for crops cannot be covered by precipitation, irrigation water has to supplement precipitation in order to meet water demand. For others, such as in a hot and arid environments, temperature may constrain cropping seasons and affect the order and variability of irrigation demand. Irrigation water is sourced from either surface water or groundwater, the availability of which is significantly impacted by geological and topographic conditions. For semi-arid or arid plains with limited rivers and lakes, groundwater can be the primary source of irrigation, as is the case in the U.S. High Plains, whereas in mountainous regions, seasonal snowpack may serve as a major source of irrigated agriculture in

that snow accumulated over higher altitudes in winter can melt the following spring and recharge rivers and lakes. The partitioning between surface water and groundwater withdrawals depends on their availabilities as well as climate conditions.

Climate variabilities and distribution of water bodies influence irrigation water supply, and in turn, irrigation withdrawals modify local hydrology and regional climate. Aquifer depletion has become a major issue related to groundwater sustainability in many regions around the world as the increasing demand in groundwater irrigation cannot be balanced by recharge, a problem further intensified under droughts (Asoka et al., 2017; Famiglietti et al., 2011; Rodell et al., 2018; 2009). Over-utilized groundwater abstraction also induces secondary effects, such as land subsidence due to compaction in the subsurface aquifer system, which can cause permanent loss of aquifer storage capacity, as is the case in the California Central Valley and North China Plain (Changming et al., 2001; Liu et al., 2019). Over-pumping of groundwater also leads to an increase in aquifer contamination, such as seawater intrusion into freshwater aquifers in coastal plains and elevated arsenic concentrations in Bangladesh, which is described as "the largest poisoning of a population in history" (Fendorf et al., 2010).

In addition to influences on groundwater systems, irrigation can have a profound impact on the regional climate via land-atmosphere interaction, as it alters the surface water and energy fluxes. On one hand, irrigation moistens the surface, increasing local latent heat flux and moist static energy in the planetary boundary layer locally and in downwind areas. This results in increased Convective Available Potential Energy and decreased Lifting Condensation Level, thereby increasing the chance for convective cloud development and rainfall. On the other hand, release of latent heat results in surface cooling, therefore restraining the development of convection and potentially inhibiting convective rainfall (Lawston et al., 2015; Mahalov et al., 2016). While both positive and negative

feedbacks exist, the net impact of irrigation on climate is strongly dependent on time, space, regional climate, land cover in surrounding areas, and the amount of irrigation (Pryor et al., 2016).

Despite its importance, the irrigation process, along with its impacts, is not adequately represented in the Earth System Models. Global Hydrological Models (GHMs) are large-scale gridded models that are valuable for resource analysis and future water use projections. They have been designed to consider the water resource impacts of large scale irrigation, but they are not structurally appropriate for coupling with atmospheric models to study climate feedbacks. Land Surface Models, initially built to support land-atmosphere coupling studies, have focused more on irrigation's impact on the atmosphere and often ignore the source of water that is applied in irrigation. Despite the efforts that have been made in both GHMs and LSMs in representing human water consumption (Döll & Siebert, 2002; Lawston et al., 2015; Leng et al., 2015; Pokhrel et al., 2015; Wada, 2015), both types of models are still imperfect and incomplete as current simulations cannot match recent hydrological observations (Nazemi & Wheater, 2015a). Chapter 2 of this dissertation is motivated by the paucity of LSMs that are able to account for groundwater withdrawals for irrigation. The chapter reports on the implementation of a groundwater pumping irrigation scheme into Noah-MP LSM and evaluation of its impact on surface water and energy fluxes in offline simulations. The study particularly focuses on simulation of water withdrawals under drought conditions, when climate forcing and adaptive management both influence water withdrawals and associated water resource impacts. Model development in this chapter, as in subsequent chapters, was conducted using the NASA Land Information System (LIS), which is a software framework for high performance terrestrial hydrology modeling and data assimilation (Kumar et al., 2006). The choice of Noah-MP LSM allows for flexible combinations of parameterization schemes and provides opportunities for future coupling with atmospheric models such as the Weather Research and Forecasting Model (WRF).

Of the numerous intensively irrigated regions among the world, the High Plains Aquifer (HPA) has been chosen as our testbed for detailed investigation. The HPA region underlies an area of ~ 450,000 km^2 in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas and Wyoming. The HPA is dominated by summer convective storms, the rainy season coincides with crop growing season, and the distribution of precipitation is uniform from north to south (475-501 mm/yr). In the HPA, surface water resources are limited to a few rivers and internally drained ephemeral lakes or playas. Groundwater is the primary source of irrigation water. Although the HPA region aquifer system is mixed, with both unconfined and confined units, the irrigation withdrawals are mostly along the southern and eastern margins of the confined Great Plains aquifer, where water levels are near the land surface and are approximately equal to the water levels of the unconfined High Plains aquifer (Miller & Appel, 1997). Ranked as the first aquifer in the United States for total groundwater withdrawals, water-level declines began in parts of the HPA soon after the beginning of substantial irrigation in the 1950s (McGuire, 2014). By 2015, declines of more than 150 ft were found in central and southern parts of the HPA, mainly in Kansas and Texas (McGuire, 2017). However, there is almost no depletion in the northern HPA, because of high natural recharge rates from the Sand Hills and additional recharge from the surface water-fed irrigation from the Platte River (Scanlon & Faunt, 2012). The fact that irrigation withdrawals in the HPA are dominated by groundwater and the fact that impacts on groundwater change can be represented by unconfined groundwater dynamics make the HPA an ideal testbed. These characteristics allow us to simplify the modeling problem by considering only groundwater irrigation, and they justify the application of the Noah-MP LSM to study groundwater levels, as this model only includes a simple unconfined shallow groundwater reservoir. Meanwhile, the intra-regional contrasts in groundwater behavior between the northern and southern HPA make it a challenging region that tests the model's capabilities to capture different recharge behaviors.

Moreover, the HPA experienced a significant drought in 2011 and 2012, offering an opportunity to evaluate model behavior during an extreme climate event.

Human-water interactions include a wide spectrum of anthropogenic interventions (Nazemi & Wheater, 2015a). From a modeling perspective, representing these features for regional to global impact assessments can be achieved in multiple ways. Among them, directly implementing the key processes into the model is the most common way to address that issue, as is done in Chapter 2 of this dissertation. However, due to limitations in data, generalized and simplified algorithms, and empirical parameterizations, there is considerable uncertainty in such models. Including human-water interactions may improve one aspect of a model simulation while at the same time introducing bias to others. Several studies have shown that models equipped with water management routines tend to overestimate groundwater decline, and this tendency has been attributed to the scarcity of observations for model calibration and to underlying uncertainties in model parameters (Scanlon et al., 2018; Tangdamrongsub et al., 2018). We found similar challenges in parts of our study region. Most notably, Noah-MP without the irrigation scheme was found to overestimate groundwater decline in Nebraska during post-drought groundwater recovery. Including groundwater withdrawal for irrigation in the model made this bias worse. These errors could be due to oversimplification of subsurface hydrology parameterizations, problems with soil parameters, and/or errors in meteorological forcing datasets. Evidently, improvements in representing the anthropogenic process itself are not good enough and may introduce more sources of uncertainty to the system.

Another way to improve model performance is through model calibration. Appropriate model calibration can address deficiencies in the model's representations of physical processes and reduce model biases. However, this comes with its own risks of underdetermined parameter sets, incomplete parameter-specific observations, and other well-known calibration challenges (Arsenault

et al., 2018). The growing development in remote sensing technologies has brought a third opportunity for improvements, which is land data assimilation. Land surface conditions such as surface soil moisture, snow water equivalent, snow cover, and near surface temperature can be estimated through hydrological remote sensing. These abundant observations provide valuable information when used in land data assimilation systems, using model-data merging ("assimilation") algorithms adopted from atmospheric or oceanic data assimilation systems (Reichle, 2008). The basic idea of land data assimilation is to introduce satellite observations of predicted model states to the Earth System Models and to produce an optimal estimate of the model states or fluxes based on the combination of the two. Direct insertion, 4DVAR, and the Kalman filter are examples of data assimilation techniques.

Efforts to assimilate observations of TWS anomalies derived from the Gravity Recovery and Climate Experiment (GRACE) satellite system are particularly relevant to the problem of capturing the variability and trends of TWS, especially for groundwater, which is a large and under-observed component of terrestrial water storage variability (Zaitchik et al., 2008). GRACE (and, since 2018, the GRACE-Follow On mission (GRACE-FO)) is a twin-satellite system that flies in a polar orbit 500 kilometers above the Earth. It maps the Earth's gravity field by measuring the distance between the two satellites using GPS and a microwave ranging system. Once changes due to ocean currents, ice sheets, glaciers and the atmosphere are removed, the remaining variation of the gravity field over land is mainly contributed by movement of water. That is how GRACE measurements can be converted to TWS anomalies. GRACE has been applied in exploring many hydrological topics, such as detecting groundwater depletion, evaluating TWS changes related to droughts, and estimating ice sheet loss, among others. However, limited by its instrument errors and the magnitude of TWS variations, meaningful estimates of TWS anomalies from GRACE can only be achieved at relatively coarse temporal (~monthly) and spatial (~ 100,000 km^2) scales. To further extend the value of

GRACE in hydrological applications, a GRACE data assimilation (GRACE-DA) technique has been developed so that the signal detected by GRACE can be disaggregated and downscaled vertically into snow, groundwater, soil moisture, and surface water bodies and horizontally into finer spatial scales using model physics, and in turn to correct for model biases.

GRACE data assimilation has been applied in GHMs and LSMs at a broad range of scales. Improvements in groundwater variability estimates and drought monitoring for various regions over the world are reported by many studies regardless of differences in model physics and DA approaches, choices of GRACE product, and spatial resolution at which the model is operated. However, challenges persist, particularly for regions that have substantial TWS depletion that is not dominated by natural variabilities. For regions that exploit groundwater to meet water resource needs, but do so in excess of natural recharge rates, there will be a negative trend in total water storage over time. Notably, this depletion includes a vertical redistribution of storage due to human water use. For example, irrigation redistributes water from groundwater (or surface water) to soil water before it goes to the atmosphere via evapotranspiration (ET). For groundwater-fed irrigation, the extracted water amount is first applied to the crops on the surface. A small part of the water goes back to the groundwater as return flow through infiltration and drainage, the percentage of which depends on the irrigation efficiency. The major part goes into the soil and taken up by the crops and finally evapotranspires into the atmosphere. In that case, surface and root zone soil moisture will increase while groundwater decreases, resulting in a net decrease in a total water storage. Without representing this process, for places that GRACE observed a decline trend in TWS, a model that includes GRACE-DA will tend to spread the declining storage signal across all vertical water storage components—from surface soils through groundwater—leading to drier surface soil moisture and thereby reduced ET, both of which should in reality be enhanced by

irrigation. This systematic error was emphasized by Girotto et al. (2017) for simulations that assimilated GRACE into Catchment LSM for northwestern India.

The fact that groundwater recovery after drought in Nebraska is underestimated by Noah-MP, and that this underestimation is exacerbated when we add irrigation withdrawal to the model, led us to use GRACE observations to constrain modeled TWS variations and trends via GRACE-DA. The challenges that GRACE-DA faces in correcting for errors in simulating human managed regions highlight the need to improve the model representation of human water consumption and irrigation. In Chapter 3, the individual and combined impact of groundwater pumping irrigation and GRACE-DA are investigated. GRACE data assimilation is implemented for Noah-MP LSM (including irrigation and water withdrawal routines) using an Ensemble Kalman Smoother (EnKS). HPA is again selected as the testbed because of the profound groundwater depletion trend induced by irrigation. A set of four simulations with and without GRACE-DA and with and without groundwater pumping irrigation is performed to isolate and compare the impacts. Groundwater, surface soil moisture, and ET are the primary focuses of model evaluation. Key factors for successful data assimilation and model realism are also discussed in Chapter 3.

Including key processes like water withdrawal for irrigation (Chapter 2) and merging satellite observations into the model (Chapter 3) both serve to improve the modeling of the coupled natural-human water resources system. In particular, the studies presented here answer the questions of whether it is important to represent human-water interactions and what might be an optimal way to utilize data and model resources for this topic. However, these first two studies intentionally used a relatively simple study area: the groundwater-dominated irrigation systems of the HPA. Generalizing the work for application to other areas presented some technical challenges. First, the irrigation scheme introduced in Chapter 2 assumes that irrigation water comes from groundwater only, which

limits the model's potential in simulating surface water dominated or groundwater and surface water mixed regions. Second, areas equipped for irrigation are fed to the model as a static map, which does not reflect changes in irrigated areas due to land use change, climate variability, and/or crop management. To overcome these limitations, in Chapter 4, we further develop datasets with information to partition source water for irrigation and to reflect the changes in potentially irrigated areas. These datasets are generated at five-year intervals based on available observations and are used as inputs to the model to expand the study from the HPA to CONUS.

As a scarcity of fresh water resources has already emerged in recent decades and is likely to increase in the future under climate change and global warming, understanding the effects of climate variability on irrigation water use is important in assessments of sustainability of fresh water resources and in the development of adaptation and management strategies. Therefore, in Chapter 4, we turn our attention from model performance to the model's potential as a tool to understand climate sensitivities of irrigation water use. The goal there is to understand how irrigation water use responds to climate variability, in particular to precipitation, temperature, and their interaction. Theses sensitivities are studied across the different growing seasons, basins, and source water characteristics of the CONUS.

Terrestrial water fluxes and surface energy budget are affected by both climate and human interventions, whereas human water consumption is affected by both climate variability and social, political, and economic factors. This dissertation is designed to improve model representation and scientific understanding of these relationships and feedbacks. The enhanced modeling framework presented through the dissertation can be used in follow-on studies of land-atmosphere interactions and projected water demand under climate change.

Chapter 2

Groundwater withdrawals under drought: reconciling GRACE and Land Surface Models in the United States High Plains Aquifer[1]

Abstract

Advanced Land Surface Models (LSM) offer a powerful tool for studying hydrological variability. Highly managed systems, however, present a challenge for these models, which typically have simplified or incomplete representations of human water use. Here we examine recent groundwater declines in the US High Plains Aquifer (HPA), a region that is heavily utilized for irrigation and that is also affected by episodic drought. To understand observed decline in groundwater and terrestrial water storage during a recent multi-year drought, we modify the Noah-MP LSM to include a groundwater irrigation scheme. To account for seasonal and interannual variability in active irrigated area, we apply a monthly time-varying greenness vegetation fraction

[1] This chapter has been published as: Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Anderson, M. C., & Hain, C. (2018). Groundwater Withdrawals Under Drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer. *Water Resources Research*, *48*(3), 317. https://doi.org/10.1029/2017WR022178

(GVF) dataset within the model. A set of five experiments were performed to study the impact of groundwater irrigation on the simulated hydrological cycle of the HPA and to assess the importance of time-varying GVF when simulating drought conditions. The results show that including the groundwater irrigation scheme improves model agreement with ALEXI ET data, mascon-based GRACE TWS data and depth-to-groundwater measurements in the southern HPA, including Texas and Kansas, and that accounting for time-varying GVF is important for model realism under drought. Results for the HPA in Nebraska are mixed, likely due to the model's weaknesses in representing subsurface hydrology in this region. This study highlights the value of GRACE datasets for model evaluation and development and the potential to advance the dynamic representations of the interactions between human water use and the hydrological cycle.

2.1 Introduction

Globally, irrigation accounts for 70% of global freshwater withdrawals (Frenken & Gillet, 2012; Siebert et al., 2010) and the volume of water extraction has increased significantly since the 1950s (Nazemi & Wheater, 2015a; Steffen et al., 2011). Groundwater (GW) supplies approximately 40% of irrigation water globally and 60% within the United States (Scanlon & Faunt, 2012; Taylor et al., 2013). It is the most reliable and important source of irrigation water for many semi-arid regions, especially during drought when crops most need supplementary water and surface water sources become limited or unavailable. Growing groundwater demand in recent decades has led to concerns of aquifer depletion in many regions around the world (Asoka et al., 2017; Famiglietti et al., 2011; Pokhrel et al., 2015; Richey, Thomas et al., 2015; Rodell et al., 2009; Scanlon & Faunt, 2012; Tiwari et al., 2009; Zeng et al., 2016; Zou et al., 2013) and has also modified the interactions between the groundwater systems and the climate (Case et al., 2013; DeAngelis et al., 2010; Gutman & Ignatov, 1998; Kueppers & Snyder, 2012; Kueppers et al., 2007; Kustu et al., 2010; 2011; Lawston et al.,

2015; Mahalov et al., 2016; Ozdogan et al., 2010; Pokhrel et al., 2012; Pryor et al., 2016; Zeng et al., 2016).

These two critical water cycle implications of irrigation – direct impacts on water resources and influence on weather and climate via the surface energy balance – have motivated two parallel streams in Earth System Model development. The first is concerned with monitoring and projecting water resources, and tends to make use of water balance models that include extensive representation of water management but are less concerned with impacts on the surface energy balance. These Global Hydrological Models (GHMs) are valuable for resource analysis, but they are not structurally appropriate for coupling with atmospheric models to study the climate impacts of irrigation. WBMplus (Wisser et al., 2010), WaterGAP (Alcamo et al., 2003; Döll et al., 2012; 2003; Eicker et al., 2014), PCR-GLOBWB (van Beek et al., 2011; Wada et al., 2013; 2010; 2014) and H08 (Hanasaki et al., 2008b) are examples of this type of model.

The second stream of model development is primarily concerned with surface fluxes through which water management—and particularly irrigation—can alter atmospheric processes. This modeling effort has focused on advanced Land Surface Models (LSM) that simulate the water and energy balances at the land surface and that can be coupled to atmospheric models at regional and global scale. Advanced LSMs include models like HiGW-MAT(Pokhrel et al., 2015), the Variable Infiltration Capacity Model (VIC; (Haddeland et al., 2006; Tang et al., 2009; Wood et al., 1992)), the Community Land Model (CLM; (Lawrence et al., 2011; Leng et al., 2015; 2014; Sacks et al., 2008)), Catchment Land Surface Model (CLSM; Koster et al., 2000) and the Noah and Noah-Multiparameterization Land Surface Model (Noah-MP LSM; Niu et al., 2011). These LSMs typically operate at sub-hourly time steps, and irrigation modules incorporated to the LSMs are designed to simulate discrete irrigation events(Ozdogan et al., 2010; Zaitchik et al., 2005). Since the primary

application area for these models is land-atmosphere interactions, however, it is not uncommon to see LSM irrigation schemes ignore the source of water that is applied in irrigation. Often the water simply appears to meet demand, and there is no effort to account for where the withdrawals occur or what impact they have on groundwater or surface water processes. This approach may be adequate when estimating on-field consumptive water use (Döll & Siebert, 2002; Ozdogan et al., 2010; Yilmaz et al., 2014) or studying land-atmosphere interactions (Evans & Zaitchik, 2008; Lawston et al., 2015; Li et al., 2016; Mahalov et al., 2016; Zaitchik et al., 2005), but it prevents the application of the models to integrated water resource analysis or to evaluate trends in water storage. Indeed, the fact that LSMs typically ignore anthropogenic influences on terrestrial water storage has been applied usefully in studies that diagnose anthropogenically-induced groundwater depletion as the residual between LSM simulations and water storage anomaly observations from the Gravity Recovery And Climate Experiment (GRACE) satellite system (Rodell et al., 2009; Tiwari et al., 2009).

The failure to account for irrigation water source, however, becomes a limitation when LSMs are merged with GRACE water storage anomaly estimates via data assimilation (Girotto et al., 2017; 2016; Houborg et al., 2012; Kumar et al., 2016; B. Li et al., 2012; Zaitchik et al., 2008). As long as natural variability dominates the GRACE water storage signal it is possible to assimilate GRACE observations into an LSM that does not account for anthropogenic impacts on groundwater or surface water reservoirs, but as the anthropogenic signal in GRACE emerges in more and more irrigation-heavy regions around the world, this assumption could break down. Lack of source water accounting is also a broader limitation that prevents useful application of LSM to study or monitor the role of managed surface and groundwater in hydrologic and climatic variability (Asoka et al., 2017; Felfelani et al., 2017; Pokhrel et al., 2015; 2017; Zeng et al., 2016; Zou et al., 2013).

To address this limitation, Leng et al. (2014; 2015) introduced a source water accounting scheme to CLM that withdraws water for irrigation in response to simulated irrigation applications. Like other efforts (Leng et al., 2017; Li et al., 2016; Mahalov et al., 2016; Ozdogan et al., 2010), irrigation applications in this formulation are demand-driven, and source water is withdrawn to meet the application required by the irrigation module. Because the model development was performed in CLM rather than a simplified GHM, Leng et al. (2014; 2015) were able to investigate local hydrological feedbacks influencing irrigation efficiency.

Building on this previous work, we implement a demand-driven sprinkler type irrigation module based on Ozdogan et al. (2010) to Noah-MP in combination with a groundwater withdrawal model of a form similar to Niu et al. (2011). We focus on the High Plain Aquifer (HPA) region of the United States, where irrigation water is drawn almost exclusively from a primarily unconfined aquifer. This simplifies the problem, as we do not account for surface water sources of irrigation water or for confined aquifers, but the tools developed in this study could be extended to include those cases. We then apply the system to study groundwater withdrawals during drought events in 2011 and 2012 that affected a large portion of the HPA. This period of drought was severe enough that irrigated area declined over parts of the HPA, as water limitations, crop growing conditions, or economic stress led farmers to fallow some of their fields (Rippey, 2015; Wallander et al., 2013).

This phenomenon of fallowing formerly irrigated fields exposes another challenge for simulating irrigation in Earth System Models: knowing when an area that is equipped for irrigation is actually being cultivated. Some irrigation modules address this challenge by including a vegetation threshold within the irrigation trigger. If a field is not sufficiently green then it is assumed to be out of production and is not irrigated, even if irrigation infrastructure is known to be present. The default greenness datasets used in many previous studies, however, are climatological averages that

do not capture year-to-year variability that might reflect fields going in and out of production (Lawston et al., 2015; Yilmaz et al., 2014). Here we further advance irrigation modeling in Noah-MP by implementing a satellite-derived time-varying greenness vegetation fraction (GVF) in place of climatological GVF estimates. This allows us to simulate irrigation water use and groundwater withdrawals over a period that includes significant variability in actively irrigated area. Through a series of offline Noah-MP simulations we explore the impacts that including groundwater withdrawal and time-varying GVF has on the model's ability to simulate groundwater levels during an extended drought in a semi-arid and intensively irrigated region.

The objective of this study is to enhance the representation of human water regulation in an advanced LSM, with specific focus on improving the simulation of water and energy fluxes in drought years. This work builds on previous studies that have: (1) studied irrigation impacts on LSM simulation of surface fluxes and meteorology, but without accounting for the impact that irrigation withdrawals have on groundwater (Lawston et al., 2015; Ozdogan et al., 2010; Pei et al., 2016; Pokhrel et al., 2012); (2) studied the impact of groundwater on TWS in global hydrological models that calculate a water balance but do not simulate energy fluxes or support coupling with the atmosphere (Döll et al., 2012; Wada et al., 2014); (3) implemented groundwater withdrawals to an advanced LSM for studies of irrigation dynamics calibrated in a single year (Leng et al., 2014; 2013). In combining groundwater withdrawal accounting with an advanced LSM and time-varying parameters to capture changes during prolonged drought, this study introduces a powerful modeling platform for monitoring and predicting freshwater system are changing under the influence of both climate variability and human water exploitation. This has been identified as a gap in current model capabilities (McDermid et al., 2017; Nazemi & Wheater, 2015b; Pokhrel et al., 2016; Wada, 2015).

2.2 Data and methodology

2.2.1 Model description

All simulations are performed using the Noah-Multiparameterization Land Surface Model (Noah-MP LSM; Niu et al., 2011), version 3.6, within the framework of NASA's Land Information System (LIS; Kumar et al., 2006), version 7.1. LIS is a terrestrial hydrology modeling and data assimilation system that allows for single or ensemble LSM simulations and supports multiple data assimilation techniques and integration of satellite-derived parameter datasets. Noah-MP v3.6 offers multi-physics options including surface/groundwater transfer and storage, dynamic vegetation and frozen soil physics. Noah-MP has delivered improved performance relative to earlier versions of Noah LSM in the simulation of runoff, soil moisture, snow and skin temperature in many river basins across the globe (Cai et al., 2014; 2015; Niu et al., 2011; Yang et al., 2011). Improvements are particularly notable for regions that have active groundwater and snow dynamics. In addition, the model's simple unconfined shallow groundwater scheme (SIMGW; Niu et al., 2007) provides the opportunity to develop a groundwater-based irrigation scheme that represents agricultural water withdrawals.

In this study, we introduce such an irrigation scheme to Noah-MP, running at 0.125° spatial resolution. The approach is based on the sprinkler irrigation scheme developed for the Noah LSM presented by Ozdogan et al. (2010). In this method, irrigation water is applied in the form of supplementary rainfall in order to maintain processes such as canopy interception that occur in sprinkler irrigation systems. The three key rules to trigger the irrigation in this modeling framework are the irrigation location (where to irrigate), timing (when to irrigate) and the amount (how much to irrigate). The Moderate Resolution Imaging Spectroradiometer – International Geosphere Biosphere Program (MODIS-IGBP) land cover dataset (1 km) is used to provide the information for cropland

or other potentially irrigated land class (e.g., grass) and the 500 m high-resolution Global Rain-fed, Irrigated, and Paddy Croplands dataset (GRIPC; Salmon et al., 2015) is used to supply the percent irrigated area within a model grid cell. The scheme determines the timing of irrigation by checking whether it is the growing season. The growing season begins and ends when a certain threshold (GVF threshold hereinafter) within the long-term range of GVF at the grid cell is exceeded. Then the scheme checks if the current root zone soil moisture availability (RZSM) falls below a certain threshold (RZSM threshold hereinafter) and estimates the irrigation water demand based on the RZSM deficit, defined as the water required to raise the current RZSM to field capacity. The RZSM is checked and the water demand is calculated at every time step between 0600 to 1000 LT local time and the irrigation is applied within this time period until the RZSM reaches field capacity. Depth of the root zone is drawn from a static crop rooting depth table.

In the original implementation of the sprinkler irrigation scheme, the source of irrigation water is not specified—it simply appears in order to meet the demand. Here we modify the sprinkler irrigation scheme to account for irrigation water sourced from a local, shallow aquifer. We do this by "pumping" groundwater from the SIMGW aquifer unit: the volume of irrigation water applied in an irrigation event is subtracted from the groundwater aquifer, and the water table depth and groundwater storage are updated accordingly. In this way, the effects of irrigation pumping on groundwater depletion can be explicitly examined. The method is similar to that developed for the Community Land Model by Leng et al. (2014). Details of the groundwater dynamics such as soil infiltration and water table depth equations can be found in Niu et al. (2007). It should be noted that the following assumptions are made in this relatively simple groundwater accounting scheme: 1) the irrigation water is fully contributed by groundwater, so the scheme is not appropriate in regions that use surface water for irrigation; 2) the irrigation water is coming from the local groundwater grid cell; horizontal groundwater flow and redistribution are not considered, as they are not represented

in Noah-MP; 3) the GVF threshold and RZSM threshold are user-specified and should be regarded as empirical parameters to calibrate the simulated irrigation amounts and groundwater dynamics.

2.2.2 Generating time-varying GVF

Previous studies that applied irrigation in the LIS framework have relied on climatological GVF fields (Lawston et al., 2015; Ozdogan et al., 2010), which capture the average seasonal cycle of vegetation but do not include interannual variability. This makes the model incapable of accounting for the response of vegetation to climate variability, especially extreme events such as hard freeze and drought (Case et al., 2012), and it means that the interannual variability in GVF in irrigated areas—which might indicate fallowing or other management changes—is neglected. This limitation has inspired some efforts to generate real-time GVF at weekly or daily scales aiming at improving the simulation of water and energy fluxes at finer temporal resolution (Lawston et al., 2017) and benefiting short-term weather forecasting (Case et al., 2013; James et al., 2009; Jiang et al., 2010). As our application is more concerned with seasonal and interannual variability, we develop a monthly time-varying GVF dataset to allow us to capture systematic changes in actively irrigated area, especially in dry years.

Following Case et al. (2013), the time-varying GVF is calculated in three steps:

1) The maximum NDVI for each land class ($NDVI_V$) is calculated using the collection of monthly MODIS NDVI composites with a spatial resolution of 0.05° from Jan 2002 to Dec 2015. The 2010 MODIS-IGBP land cover dataset (also 0.05° resolution) is used to generate a single distribution of the maximum NDVI of all grid points with the same land use class and identify the 95th percentile of maximum NDVI for each land use class.

2) Using the same approach, the 5th percentile of the maximum NDVI for the barren land use class (NDVIS) is calculated representing zero vegetation coverage.

3) The GVF at each grid point *i* is calculated as:

$$GVF_i = \frac{NDVI_i - NDVI_S}{NDVI_{V,i} - NDVI_S} \tag{2.1}$$

where $NDVI_i$ is the NDVI composite value at grid point *i*. It should be noted that land use change is not considered in generating GVF in our study but may be substantial in places that have experienced rapid urbanization in recent years.

2.2.3 Experimental design

In this study, Noah-MP with the incorporated pumping irrigation scheme is run offline within the LIS framework over the HPA in the western United States. A 3x21 year offline spin-up with irrigation turned off is performed (three times through the period 1995-2015), as the presence of a groundwater model in Noah-MP requires long spin-up to reach an equilibrium water table depth (Cai et al., 2014).

Following spin-up, a set of five simulation experiments are performed for the GRACE period of record, 2002-2015. The experiments are designed to study the impact of irrigation with groundwater withdrawal on the simulated hydrological cycle of the HPA and to assess the importance of time-varying GVF when simulating irrigation during drought conditions (Table 2.1). The first two runs, noIRR_C and IRR_C use a climatologically averaged GVF dataset which is derived by computing the monthly averaged time-varying GVF as described in section 2.2. The remaining three runs noIRR_T, IRR_T and IRR_TR use the time-varying GVF dataset. The noIRR_C and noIRR_T simulations serve as control runs in which irrigation is not represented,

while the IRR_C, IRR_T and IRR_TR simulations account for irrigation. The distinguishing feature of IRR_TR is that the GVF threshold used to define the active irrigation season varies spatially, as a function of the average annual GVF range of the grid cell. This differs from IRR_C and IRR_T, and from previous irrigation studies in LIS, which use a fixed GVF threshold for all irrigation grid cells. All the simulations set the RZSM threshold to 0.45 which means that irrigation is triggered when the RZSM falls below 45% of the soil moisture range from wilting point to field capacity during the growing season.

Table 2. 1 Description of simulations conducted in this study.

Name	Irrigation	GVF dataset	GVF threshold
noIRR_C	No	Climatology	-
IRR_C	Yes		constant
noIRR_T	No	Time-varying	-
IRR_T	Yes		constant
IRR_TR	Yes		function of GVF range

2.2.4 Data

The atmospheric forcing data for all sets of runs are from the National Land Data Assimilation System – Phase 2 (NLDAS-2; Xia et al., 2011; 2012) at 0.125° spatial resolution. NLDAS-2 meteorological fields are downscaled from the North American Regional Reanalysis (NARR) data, supplemented with several in situ observational datasets. The U.S. Geological Survey (USGS)-reported groundwater irrigation use at county level for 2005 and 2010 are used to evaluate simulated groundwater withdrawals for irrigation (https://water.usgs.gov/watuse/data/). The GRACE-derived Terrestrial Water Storage (TWS) anomaly data are obtained from the NASA JPL archive (https://grace.jpl.nasa.gov). TWS estimates derived using both global spherical harmonics

(SH) (Bettadpur, 2007; Wahr et al., 1998) and regional mass concentration (MS) (Rowlands et al., 2005; Watkins et al., 2015) processing approaches are used for TWS anomaly evaluation. SH products are more established and have been used in many previous studies of TWS anomalies and trends. However, recent work has indicated that SH tends to attenuate terrestrial signals and is not optimal for quantifying human water abstractions, especially in dry regions (Döll et al., 2014). The MS approach allows for localized solutions that explicitly define land and ocean. This limits leakage error and has the potential to reduce uncertainties compared to the SH method (Long et al., 2015; Scanlon et al., 2015; 2016). In our study, the GRACE_SH is the ensemble average of TWS anomalies from the Center for Space Research (CSR), NASA Jet Propulsion Laboratory (JPL) and GeoforschungsZentrum Potsdam (GFZ) based on the SH processing approach and the GRACE_MS is calculated from the mascon (MS) TWS anomalies from JPL. For the hot spot areas with intensive groundwater depletion in HPA, we examine the water level data in winter time from available wells over those areas collected by USGS (McGuire, 2014). Finally, to quantify the impact of irrigation on evapotranspiration simulation over the irrigated areas, we compare model outputs against diagnostic evapotranspiration (ET) estimates from the Atmosphere-Land Exchange Inverse (ALEXI; Anderson et al., 2007) surface energy balance model, which are calculated using time differential land surface temperature recorded by the GOES satellites and other satellite-derived inputs.

2.3 Results and discussion

2.3.1 Irrigation water amount

The HPA system, comprising the Ogallala, Brule, and Arikaree Aquifers, is one of the world's largest aquifer system. It covers eight states and underlies an area that accounts for 27% of the irrigated land and 30% of the groundwater used for irrigation in the United States. A satellite-

derived map of irrigation fraction percentage (Salmon et al., 2015) aggregated to 0.125°(Figure 2.1 (a)) shows that major irrigation zones of HPA are located in Nebraska (NE), Kansas (KS) and Texas (TX). Figure 2.1 (b) gives an overview of the simulated 14 year averaged annual irrigation rate over the HPA. Irrigation rates are highest in Kansas and Texas and somewhat lower in Nebraska. A map of groundwater irrigation fraction derived from the USGS water use report for 2000 from Leng et al. (2014) shows that groundwater is the dominant source for irrigation across the HPA (Figure 2.1 (c)). However, there is some large-scale surface water irrigation in Wyoming and Colorado along major rivers. These areas are withheld from all groundwater analyses in this study through the application of a mask to remove all grid cells with less than 50% groundwater contribution to irrigation.

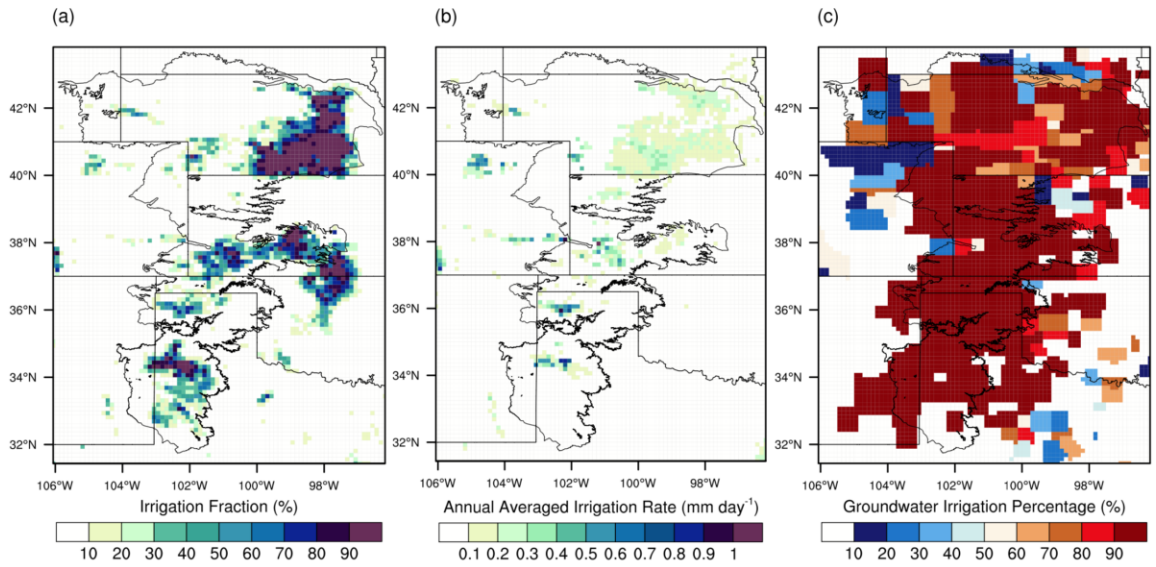

Figure 2. 1 (a) The global rain-fed, irrigated, and paddy croplands (GRIPC) dataset irrigation fraction over the HPA, (b) simulated long term averaged (2002-2015 baseline) annual irrigation rate and (c) percent of irrigation from groundwater derived from USGS water use report for 2000 over the HPA.

USGS estimates of groundwater extraction for irrigation from the HPA are available at county level every five years. Taking an average of the 2005 and 2010 USGS figures and comparing them to our simulated annual irrigated water use indicates that the simulations that include irrigation match the relative distribution of extracted water across the three major irrigating HPA states quite closely, though the simulated quantities are about half of the USGS estimates (Table 2.2). In all simulations and in the USGS data the largest total water use is in Nebraska. Though the averaged rate is much lower in NE than in KS or TX, the irrigated area is very large. The IRR_C and IRR_T simulations produce similar estimates of irrigation amount while IRR_TR is slightly lower. The underestimation relative to USGS is likely due to several factors. First, the simulated irrigation water amount is exactly the amount of water that the root zone soil column needs to reach field capacity. In reality, however, irrigation systems are not perfectly efficient and soils are heterogeneous. This means that the extracted amount of water as reported by USGS isn't fully utilized by the crops. Instead, water loss may occur during delivery to the field, through rapid infiltration processes in macropores that lead to return flow without contributing to plant available water, and in the form of surface runoff that flows out of the region or evaporates. Irrigation inefficiency issue has been addressed in several studies (Döll et al., 2012; Döll & Siebert, 2002; Pokhrel et al., 2015; Wada et al., 2014). Doll et al. (2012) reported that groundwater irrigation efficiency is around 40% globally and Wada et al. (2014) reported that for US, the irrigation water use efficiency (including both surface and groundwater resources) is around 60%. Second, the model assumes uniform soil layer thickness and the irrigation water demand is estimated only to meet the soil moisture deficit in the model's root zone layers. In reality, soil thickness is heterogeneous, and drainage from surface layers may be influenced by hydraulic gradients across these deeper layers. Lastly, irrigation methods vary. The USGS reports that HPA has a small proportion of flood irrigation that mostly occurs along rivers. This system has lower water use efficiency and is not considered in our simulations. In future work

it may be possible to apply recently piloted satellite-based methods to estimate variability in irrigation withdrawals in the absence of continuous *in situ* withdrawal records (Anderson et al., 2015).

Table 2. 2 Comparison of the observed and simulated groundwater irrigation amount for NE, TX, KS and the whole HPA.

Groundwater irrigation withdrawals (km^3)				
AREA	USGS Report	Simulations		
		IRR_C	IRR_T	IRR_TR
NE	6.1	3.2	3.1	2.6
TX	2.6	1.6	1.6	1.4
KS	2.7	1.4	1.4	1.2
HPA	13.3	6.7	6.6	5.6

2.3.2 TWS anomaly and GW depletion

Groundwater, soil moisture, snow and ice, lakes and rivers, and water contained in biomass are the principal components of TWS (Rodell & Famiglietti, 2001). The dominant contributors to TWS anomalies vary among climate regions. Figure 2.2 shows the comparison of simulated TWS anomalies, as contributed only by soil moisture and groundwater, along with GRACE observations for NE, KS, TX and HPA. The figure shows GRACE TWS anomaly solutions derived using both spherical harmonics (SH) functions and mass concentration (mascon; MS) functions. It is evident from Figure 2.2 that both GRACE observation and all the simulations show a decline in TWS from 2011 to 2013 and slow recovery in the following years.

The climate condition in HPA ranges from cooler, more humid conditions in the North to warmer and semi-arid conditions in the South. This gives rise to different TWS response to climate variability in the states of interest. For NE, the simulated TWS decline is relatively similar in simulations with and without irrigation, despite the fact that irrigation extent and intensity in NE is

quite significant. This similarity indicates that TWS depletion due to natural climate variability is larger than that contributed by irrigation withdrawal in NE. Indeed, the contribution of irrigation to TWS variance in NE, estimated as the ratio $(Var\ (TWS)_{IRR\_TR} - Var\ (TWS)_{noIRR\_T})/$ $Var\ (TWS)_{noIRR\_T}$ for cells that are irrigated every year, is 31% for NE, while it is 78% and 66% for TX and KS respectively. However, all the simulations underestimate TWS variability inferred from the GRACE data, including in the period of rising TWS since 2011. This is mainly due to the underestimation of the TWS recovery in winter 2011 and 2012. The reason for the model's failure to capture the TWS recovery in these two winter period has not been identified. It could inherit from meteorological forcing errors, or it could be a result of limitations in model parameters, structure, or process representations that lead to inaccuracies under conditions of extended TWS recharge in the NE portion of the HPA. This apparent error with respect to GRACE TWS highlights the importance of continued model development to investigate possible TWS-relevant biases in soil retention, percolation, groundwater recharge and/or snowpack accumulation and melting.

For TX and KS, the magnitude of the difference between irrigated and non-irrigated simulations is as large or larger than the magnitude of the depletion signal observed by GRACE since 2011. This indicates that groundwater depletion due to irrigation dominates the simulated TWS signal, rather than natural variability. It is interesting to note that IRR_T and IRR_TR—the two simulations that account for both irrigation and vegetation variability—provide the best match with GRACE_MS, which shows a substantial TWS decline in KS and TX portions of the HPA since 2011. Simulations without irrigation (noIRR_C and noIRR_T), meanwhile, match well with the GRACE_SH record, in which TWS depletion in 2011 and 2012 is small and TWS soon recovers in the following years. The anomaly correlation R fields are calculated based on comparison between simulated TWS and both GRACE solutions (Table 2.3). Anomaly correlations with GRACE_SH are

lower for IRR_T and IRR_TR than they are for noIRR_T in NE and KS. The correlations are higher for IRR_T and IRR_TR in TX, but the improvements over noIRR_T are not statistically significant at 95% confidence level (Fisher's Z transform test). However, the anomaly R values calculated relative to GRACE_MS are significantly higher in TX and KS for IRR_T and IRR_TR compared to noIRR_T while still insignificantly lower in NE due to the apparent underestimation of TWS recovery in 2011 and 2012 noted above.

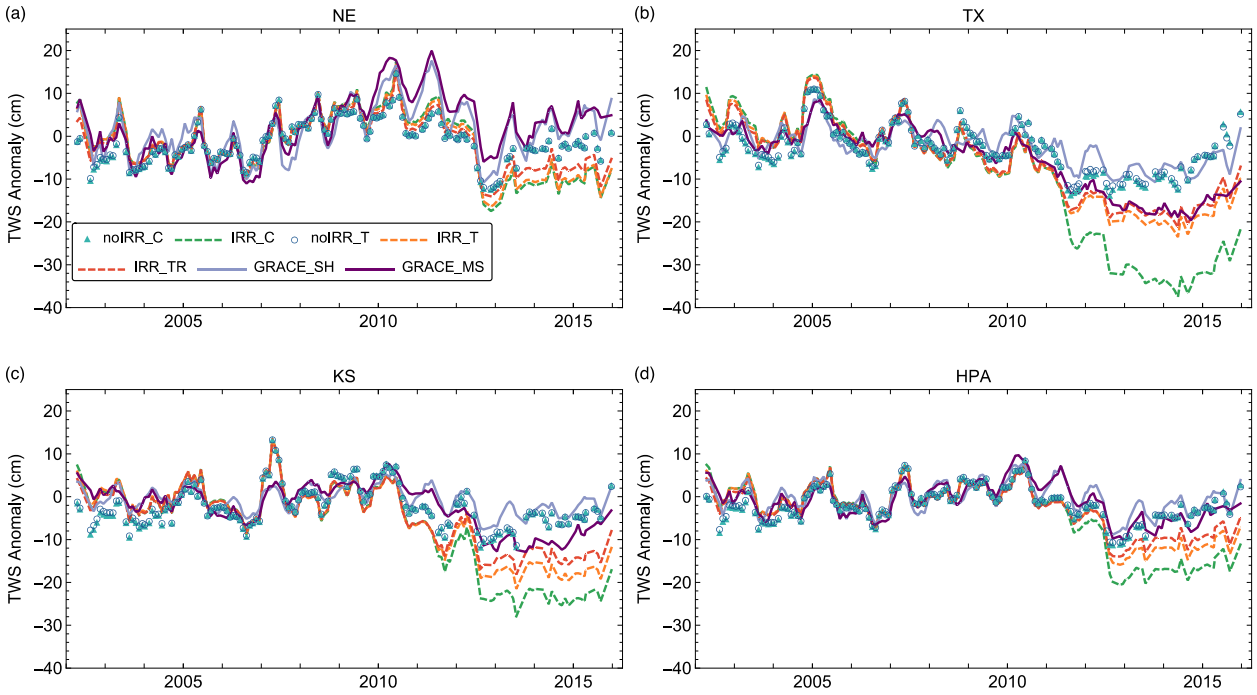

Figure 2. 2 Time series comparison of TWS estimates from the simulations and GRACE observations derived from spherical harmonics functions (GRACE_SH) and mass concentration (GRACE_MS) for (a) NE, (b) TX, (c) KS and (d) the entire HPA.

The differences between the two GRACE observational products introduce some ambiguity to the evaluation of simulated TWS. However, we have higher confidence in GRACE_MS for a region like the High Plains, both because GRACE_SH is known to dilute the gravitational anomaly signal over land and because the correction factors applied to address such shortcomings are based

on LSMs that do not include water withdrawal. For this reason, we focus on GRACE_MS to assess the realism of the irrigation simulations, and we conclude that the groundwater irrigation scheme successfully captures the impact that groundwater-sourced irrigation has on water storage in the southern portion of the HPA, despite the underestimation of the irrigation water amount relative to USGS. The failure of the groundwater withdrawal scheme to improve simulation of TWS in NE reflects the dominance of natural variability in that portion of the basin and, perhaps, model shortcomings in representing subsurface hydrology that are not directly related to irrigation. The large differences between the two GRACE products also highlight the importance of considering the applicability of the observation data for model evaluation in any given region.

Table 2. 3 Correlations of TWS anomalies between GRACE observations and five sets of simulations. Bold front indicates a significant higher correlation compared to the simulations with the same GVF option at the 5% significance level using paired sample t-test.

Anomaly Correlation Coefficient						
	GRACE_SH			GRACE_MS		
	NE	TX	KS	NE	TX	KS
noIRR_C	**0.77**	0.79	0.75	0.74	0.77	0.70
IRR_C	0.65	0.81	0.65	0.62	**0.97**	**0.88**
noIRR_T	0.77	0.80	0.75	0.74	0.78	0.70
IRR_T	0.66	0.83	0.68	0.62	**0.96**	**0.87**
IRR_TR	0.70	0.83	0.70	0.66	**0.95**	**0.85**

TWS results for the simulations with irrigation also suggest that it is important to account for GVF variability when simulating irrigation. IRR_T and IRR_TR outperform IRR_C, as IRR_C generates unrealistically large groundwater depletion in TX and KS during the drought, when irrigation was curtailed due to water shortage. More detailed analysis of this will be provided in section 3.3. IRR_TR, which accounts for different vegetation types and cropping practices when

setting the GVF threshold, tends to reduce TWS depletion relative to IRR_T, which brings simulations into even closer agreement with GRACE_MS.

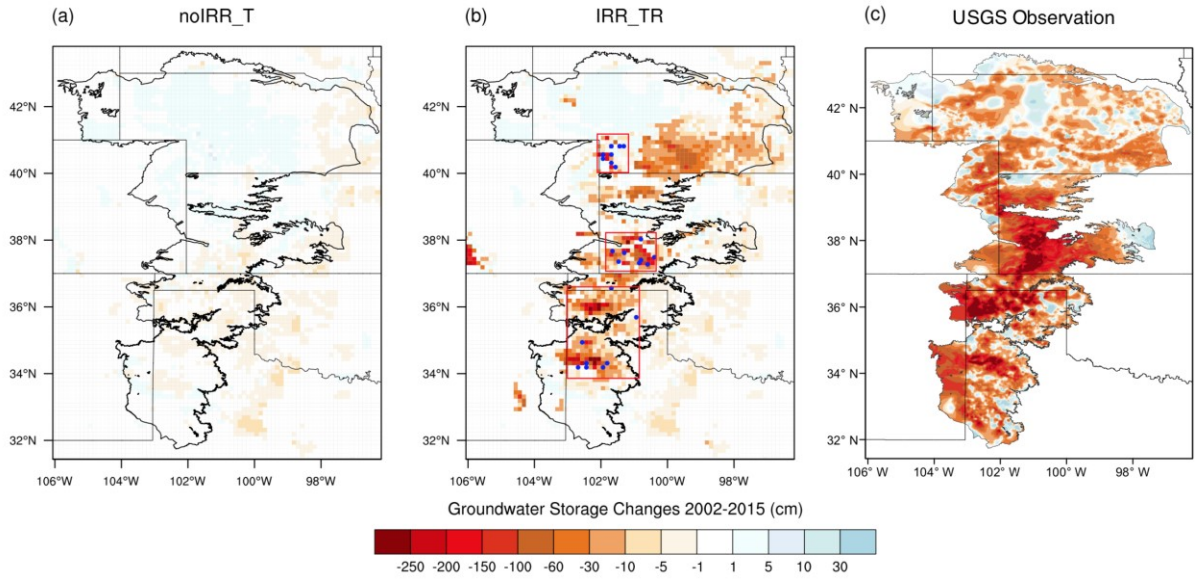

Figure 2. 3 Comparison of the simulated groundwater storage declines for (a) noIRR_T and (b) IRR_TR and (c) USGS observation from 2002 to 2015 over the HPA. Red boxes in (b) represent three decline hot spot area in NE, KS and TX respectively and the blue points are the selected observation sites for water level change comparison (see Figure 4). Groundwater change fields in (c) are generated by interpolating the water-level measurements following McGuire (2017) and then converted by using HPA specific yield dataset (Gutentag et al., 1984), while (a) and (b) are gridded LIS fields.

The magnitude and location of simulated groundwater decline due to irrigation can be examined by comparing the 2002-2015 change in groundwater simulated by the best-performing model without irrigation (noIRR_T; Figure 2.3a) and the best-performing model that includes irrigation (IRR_TR; Figure 2.3b). For noIRR_T, there is a small groundwater decline in TX and KS and a mixed signal in NE. However, the groundwater depletion contributed by irrigation is remarkable. In IRR_TR, the average water table of the TX portion of the HPA dropped by 2.14 m between 2002 to 2015. For KS this figure was 2.07 m and for NE it was 1.54 m.

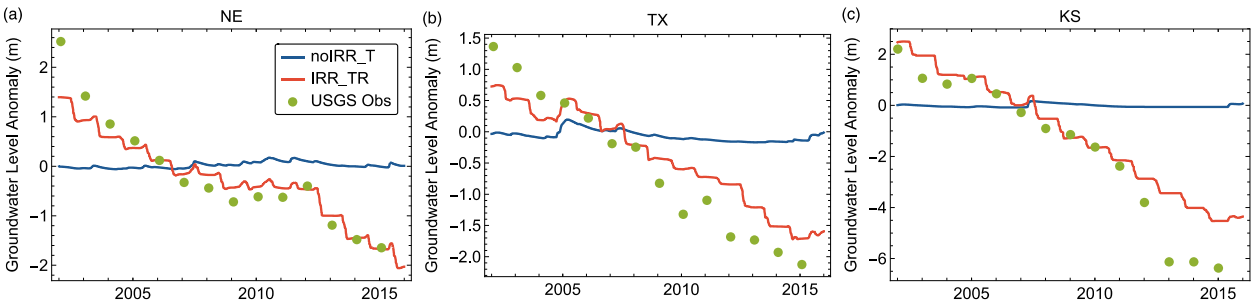

Figure 2. 4 Comparison of simulated daily water-level change (noIRR_T and IRR_TR are shown) with the observed winter time water-level change for the three depletion hot spots in (a) NE, (b) TX and (c) KS.

Incorporating irrigation brings the distribution and the magnitude of groundwater depletion hot spots area over HPA into closer agreement with the USGS observations (Figure 2.3c). The water-level change field for Figure 2.3c is firstly generated following the USGS interpolation method (Khandelwal et al., 2017) and then converted to groundwater change field by introducing the specific yield dataset for HPA (Gutentag et al., 1984). Figure 2.4 shows the area averaged winter time water-level anomaly for three "hot spot" areas, as shown in the red boxes in Figure 2.3b for NE, TX and KS respectively. Eight observational sites (shown in Figure 2.3b) for each area are selected for comparison. For each, we apply an averaging filter by extracting simulated groundwater level for the average of nine grid cells that surround the well location, and then averaging across all well locations within each hot spot. In IRR_TR, the groundwater levels for these three regions continuously drop from 2002 to 2015 at a rate of 20 cm/yr, 18cm/yr and 53cm/yr in NE, TX and KS, respectively. These results are generally consistent with the observations. In NE, this agreement is surprisingly good considering the model's limitations in representing recharge processes. This agreement could reflect the fact that the monitored area is relatively dry and does not include the potentially missing process that causes the bias of averaged TWS anomaly in NE for winter 2011 and 2012. For TX and KS, the model slightly underestimates depletion relative to well observations. This underestimate is

not seen in the simulation of TWS at large scale (Figure 2.2). The discrepancy could be due to well location, as the GRACE TWS estimate is smoothed over space while the wells are located in active irrigation zones, it could also be a function of simplified model hydrology. The wells show water level in the utilized aquifer. In KS, this aquifer lies beneath a shallow, low permeability alluvial aquifer that can inhibit groundwater recharge (Gurdak & Roe, 2010; Katz et al., 2016; McMahon et al., 2006). Noah-MP does not distinguish between these aquifers. The results from noIRR_T also indicate that climate variability alone has almost no impact on the observed groundwater trends in these three intensively irrigated areas.

2.3.3 Irrigation water use in response to the drought

As described above, both observations and our simulations indicate that groundwater withdrawals had a significant impact on TWS in the HPA over the 2002-2015 period. To explore shorter term dynamics of groundwater exploitation under climate variability we examine model and observations in 2011 and 2012, when portions of the HPA were impacted by drought. The signal of these droughts is evident in larger depletions in both TWS (Figure 2.2) and groundwater (Figure 2.4) for TX in 2011 and NE in 2012.

The drought is also evident in a sample comparison between climatological GVF and time-varying GVF in August (Figure 2.5). In 2011, GVF is reduced over TX and part of KS, while in 2012, GVF reduction is widespread and is greatest in NE. The lower values of GVF in these two years are primarily a result of low rainfall, as shown in Figure 2.6 (a). In TX the drought began in 2011, with the annual rainfall reduced from 488 mm in climatology (2002-2015 baseline) to 195 mm in 2011. The drought expanded to the northern HPA in 2012, with a sharp decrease of annual rainfall in NE and KS. The strong correlation of precipitation and GVF anomalies indicates that drought can lead to a vegetation response that is quite different from the climatological response,

which, in turn, emphasizes the importance of realistic representation of GVF for better estimation of land surface fluxes.

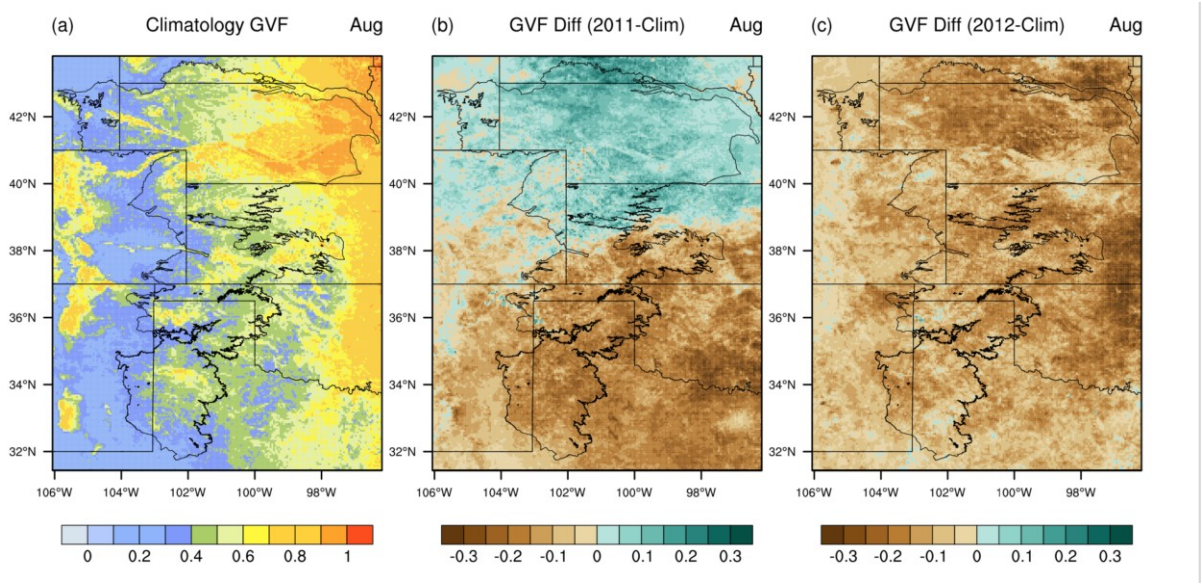

Figure 2. 5 The greenness vegetation fraction (GVF) over the HPA for (a) climatology (2002-2015 baseline) and the difference between climatology and (b) 2011 (2011 - climatology), (c) 2012 (2012 - climatology).

What is less obvious, however, is the impact that drought might have on irrigation water use. On one hand, reduced precipitation and arid conditions increase the demand for irrigation water. On the other, water shortages or climate conditions that are unfavorable for crop growth might lead farmers to take irrigated areas out of production. Further, drought may prompt voluntary or mandatory water usage reductions in some jurisdictions. We examine this with model simulations. First, simulations with climatological GVF (IRR_C; Figure 2.6(b)) indicate a large spike in irrigation intensity in TX in 2011 and in the rest of the HPA in 2012. This occurs because a simulation with climatological GVF responds to the increased demand for irrigation but has no information on changes in vegetation status within potentially irrigated fields. Thus there is no mitigation in water withdrawals due to farmers taking fields out of production.

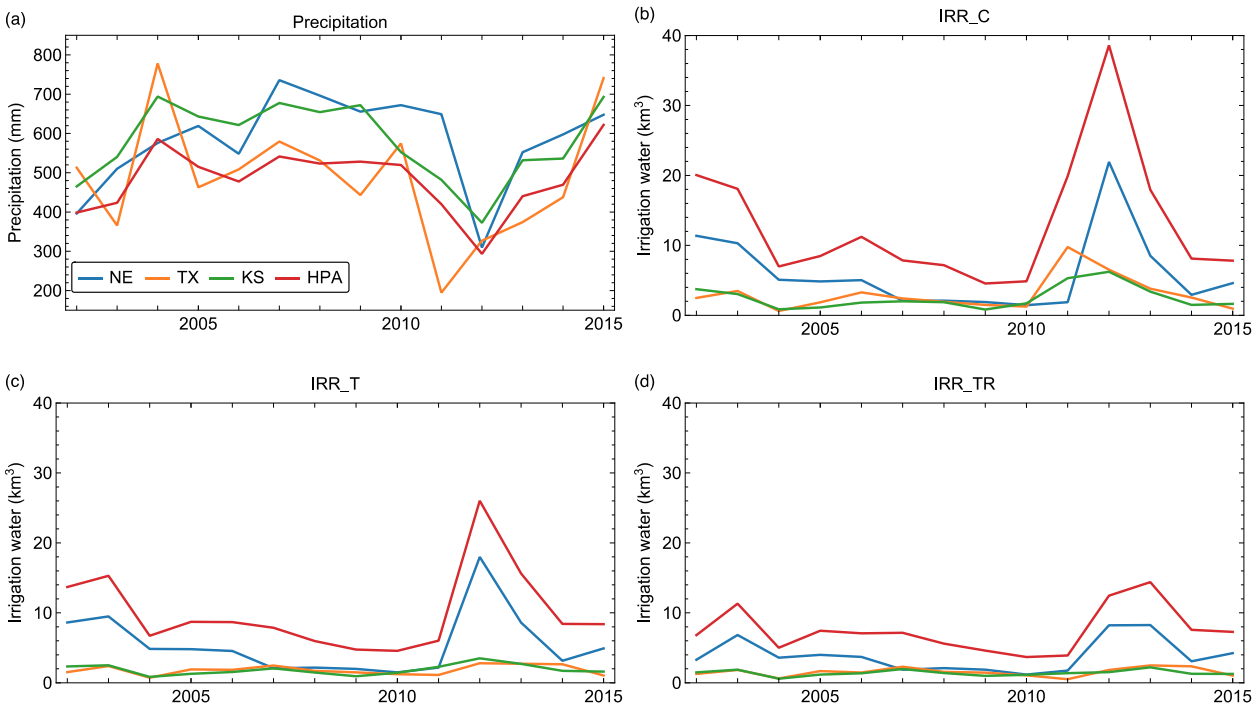

Figure 2. 6 (a) Annual averaged precipitation from the NLDAS-2 forcing data and annual averaged irrigation water amount simulated by (b) IRR_C, (c) IRR_T and (d) IRR_TR for NE, TX, KS and the entire HPA.

This limitation explains why IRR_C greatly overestimates the decline in TWS during the drought (Figure 2.2). Introducing the time-varying GVF reduced this spike in water use (IRR_T; Figure 2.6(c)), and applying a GVF threshold as a function of its range reduced simulated irrigation applications by 70% for TX in 2011, 62.5% for NE and 75.5% for KS in 2012 (IRR_TR; Figure 2.6(d)). USGS irrigation withdrawal data are not available for these years, so we rely primarily on TWS and well comparisons to assess model performance. By these metrics, IRR_TR, which overall produces the best match for TWS anomaly (Figure 2.2) and groundwater, outperforms other simulations and is deemed to have the most reasonable estimation of drought impacts on irrigation water withdrawals, recognizing that off-field losses are not included in the model. The irrigation response to drought can also be quite different from region to region. As can be seen in Figure 2.6,

IRR_T and IRR_TR indicate that irrigation water use actually declined in TX in 2011 relative to 2010, but there was a substantial increase in KS and NE during the 2012 drought. This discrepancy may reflect regional differences in farmers' willingness to grow an irrigated crop under drought conditions.

As an additional evaluation of model representation of irrigation under drought, we compare ET simulated by Noah-MP over irrigated areas to satellite-derived ET estimates from ALEXI (Figure 2.7d-i). To first order, the August climatology and the spatial pattern of the ET differences between climatology and drought years in IRR_TR simulation are in good agreement with the ALEXI. ET variability over the HPA mirrors rainfall variability: below average August rainfall in the southern HPA in 2011 (Figure 2.7b) and most of the HPA in 2012 (Figure 2.7c) has parallels in the ALEXI ET fields for these two years compared to climatology (Figure 2.7e-f). But the result shows that climate conditions drive large-scale ET variability that extends well beyond irrigated areas, and that this rain-fed variability appears to dominate over irrigation-related ET variability when viewed across the entire HPA.

There are, however, systematic differences in ET associated with irrigation in drought years relative to non-drought years. In both 2011 and 2012 only a subset of GRIPC-identified irrigated areas are identified as irrigated according to our GVF threshold (compare Figure 2.7k and 2.7l to Figure 2.7j). When we consider ET from actively irrigated grid cells in each year, as viewed by ALEXI and Noah-MP simulations, we see that there is interannual variability in ET in cells that are >50% irrigated (Figure 2.8), but that irrigation serves to buffer this variability—ET variability is lower in irrigated cells than in non-irrigated cells, and Noah-MP simulations that include irrigation show less variability than those that do not. In capturing this buffering effect, Noah-MP simulations with irrigation tend to draw ET estimates into closer agreement with ALEXI. During drought years,

the impact of irrigation on ET is enhanced. RMSE in ET from irrigated grid cells in August (a period when irrigation is large and rainfall is relatively low) for model versus ALEXI is reduced in irrigated relative to non-irrigated simulations for NE and TX in both 2011 and 2012 and for KS in 2012 (paired sample t-test) (Table 2.4). For example, with irrigation (IRR_T), ET averaged for these active grid cells in August 2012 is increased by 44.2 mm, 55.5 mm and 31.6 mm in NE, KS and TX respectively, compared to noIRR_T . In all cases this leads to a significant reduction in RMSE evaluated against ALEXI. Only in KS in 2011 does the addition of irrigation to Noah-MP drive simulations away from the ALEXI observed ET. Reduced variances of the errors for all three states in the simulations with irrigation for 2012 also indicate that implementing irrigation scheme enhanced the stability of the model during the widespread drought (Levene test). The Taylor diagram (Figure 2.9) also demonstrates that monthly ET simulated by IRR_TR is in better agreement with ALEXI compared to that simulated by noIRR_T for active grid cells in each state with a higher correlation coefficient and reduced bias.

This apparent improvement—defined relative to ALEXI—is only evident in drought years. In wet years there is a small tendency for simulations to overestimate ET in some areas relative to ALEXI, a result that could be attributed to shortcomings in the irrigation routine. On one hand, this overestimation of summertime ET may due to the wet bias of Noah-MP that overestimates the evapotranspiration as compared with ALEXI. On the other, ALEXI is selected as a high performing satellite-derived ET estimate, but its strength is its ability to portray spatial and temporal variability of ET, while the accuracy of its climatology mean ET is less well verified. Besides, ALEXI and Noah-MP capture both irrigated and unirrigated fields in single grid cell, while the cell size of ALEXI and Noah-MP are different. This may also contribute to ET difference in summertime. This model shortcoming should be investigated, but it is not related to our groundwater withdrawal and irrigation algorithm and so it is not considered in this study.

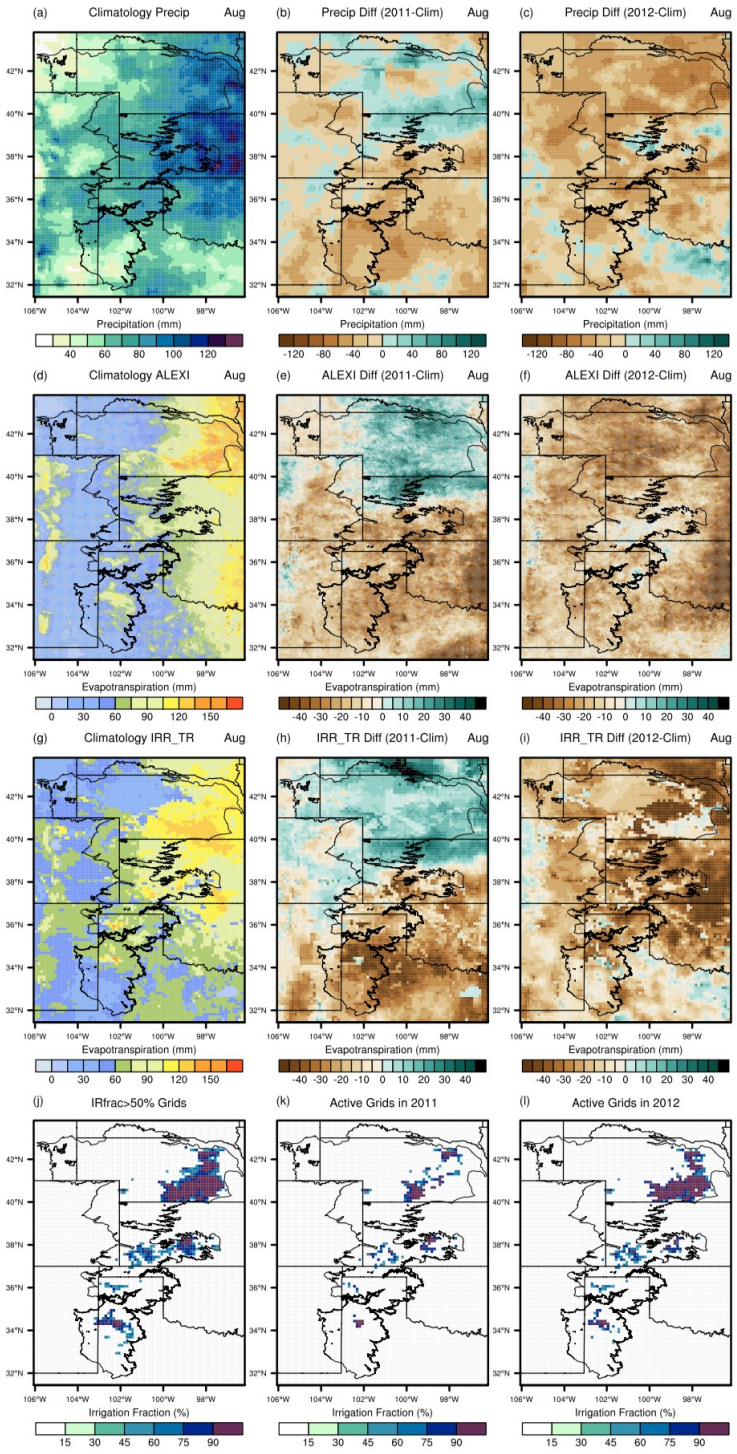

Figure 2. 7 Comparison of (a, d) climatology (2002-2015 baseline) and (b, c, e, f) the difference of precipitation and evapotranspiration (ET) pattern in August during drought period. (g) shows the grids with irrigation fraction greater than 50% from GRIPC and irrigation active grids among them for (h) 2011 and (i) 2012.

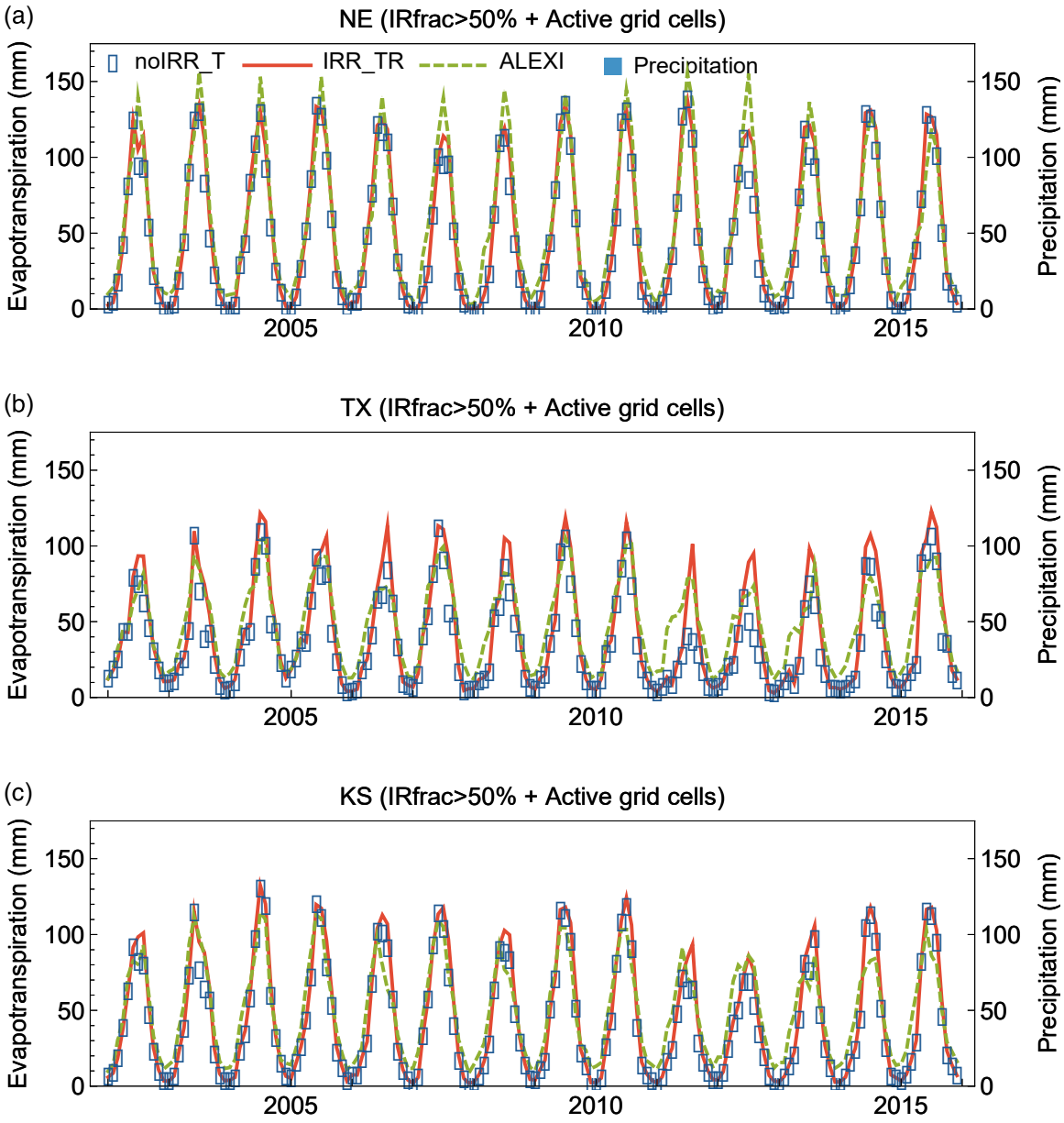

Figure 2. 8 Monthly ET (left axis) of the simulations with time-varying GVF averaged over the actively irrigated grid cells with irrigation fraction larger than 50% for each year in (a) NE, (b) TX and (c) KS. Also shown is the corresponding grid cells' averaged monthly precipitation (blue bar plots) from the NLDAS-2 forcing data (right axis).

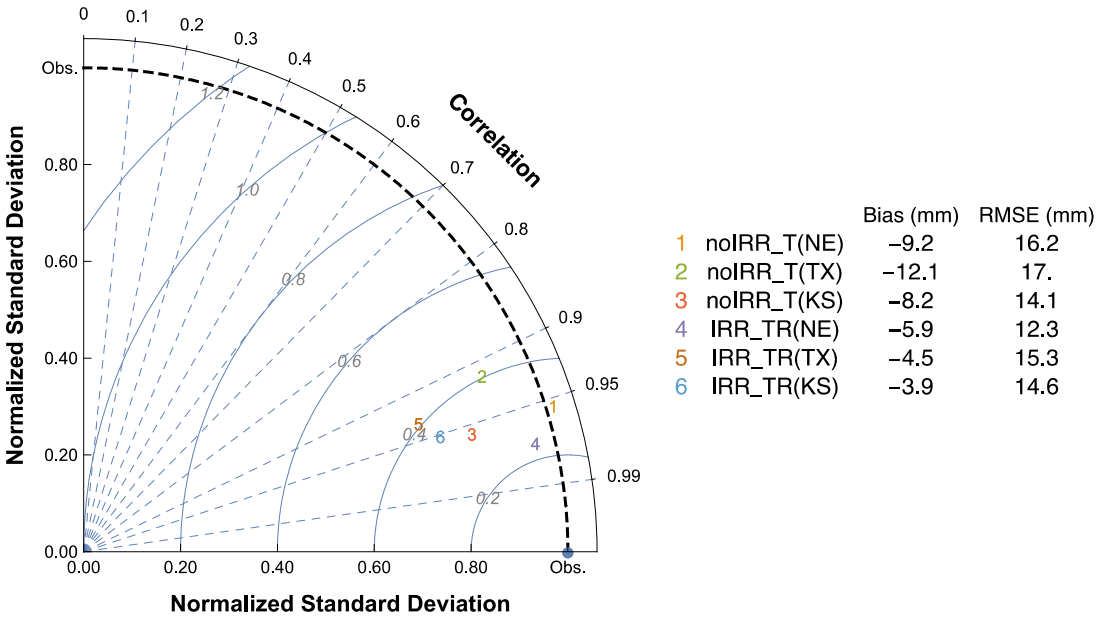

Figure 2. 9 Taylor diagram of monthly ET for NE, TX, and KS, evaluated against ALEXI, along with the bias and RMSE.

Table 2. 4 Evaluation of the evapotranspiration (ET) estimates for irrigation active grids from the simulations against ALEXI. Mean value of ET are calculated with respect to daily averaged ET in August for 2011 and 2012 using data from active irrigation grid cells in the corresponding year. Values denoted with * indicate a significant reduction in error in IRR_T/IRR_TR relative to noIRR_T or IRR_C relative to noIRR_C at the 5% significance level using paired sample t-test. Bold font indicates a reduced variance of the errors at the 5% significance level using Levene test. Italics indicates degradation of simulations with irrigation.

Evapotranspiration for active grids with IRR_Frac >50% in Aug (mm)						
AREA	YEAR	Climatology GVF		Time-varying GVF		
		noIRR_C	IRR_C	noIRR_T	IRR_T	IRR_TR
NE	2011	111.9	116.2*	112.4	117	117.1
	2012	69.2	115.4*	68.8	**113***	107.4*
TX	2011	36.6	114.5	36.6	99*	101.3*
	2012	38.9	105.4	39.1	**94.6***	**95***
KS	2011	64.4	*100.5*	63.8	*93.6*	*93.6*
	2012	53.3	**94.7**	52.5	**84.1***	**76.1***

2.4 Conclusions

This study investigates the impacts of groundwater-fed irrigation on Noah-MP's ability to simulate irrigation-induced terrestrial water storage change during a drought in the HPA region. We modify Noah-MP to include a groundwater pumping irrigation scheme and apply a monthly time-varying GVF dataset to the model, enabling the representation of variability in irrigation water use and groundwater withdrawals. The results show that including irrigation in Noah-MP improves model agreement with GRACE mascon solutions for TWS and well observations of groundwater depletion in the southern HPA, including Texas and Kansas, and that accounting for time-varying GVF is important for model realism under drought. Results for the HPA in Nebraska are mixed, likely due to misrepresentation of subsurface hydrology.

This study points to several areas for future work. First, we find that including time-varying GVF improves the simulation of irrigation variability under drought. By default, many LSMs and regional climate models use climatological values for GVF and related vegetation fields. This simplified approach needs to be examined. Second, we fail to replicate the magnitude of water withdrawals in the HPA reported by USGS. This emphasizes the need to better simulate the inefficiency of current, widely-employed irrigation practices. Our demand-driven approach does not include information on water losses that drive total withdrawals and that may impact redistribution of groundwater. Third, comparisons with observation are fraught with uncertainty due to limited groundwater monitoring efforts and to uncertainty in GRACE estimates of TWS variability. Continued work on in situ and space-borne observing systems are critically important, as are robust, process-based evaluations of existing products to understand their applicability and limitations. Fourth, this is a single-model study that is subject to uncertainties related to Noah-MP physics and parameterizations. As more open source modeling systems implement water management routines it

will be valuable to perform multi-model comparison studies to constrain uncertainties and identify gaps in data and understanding. Fifth, the irrigation schemes introduced here are extremely simplified from a management perspective. Water applications are entirely demand-driven, and do not account for other factors that influence irrigation timing or amount, such as crop type. Finally, the groundwater extraction schemes presented here are generally applicable to unconfined aquifers. Extension to include confined aquifers and surface water extraction is the subject of ongoing model development.

Despite these limitations, the current study demonstrates that accounting for groundwater withdrawals and variability in an irrigated area can improve simulation of groundwater withdrawals under drought. Such methods are required to understand coupled natural-human systems affecting the HPA under climate variability and change. Incorporating groundwater withdrawals in an advanced LSM, as demonstrated here, makes it possible to link coupled natural-human water resource analysis with the study of distributed hydrologic fluxes and land-atmosphere interactions that might feed back onto climate variability.

Acknowledgements, Samples, and Data

The study was supported in part by NASA GRACE Science Team award NNX16AF12G. Different datasets used for model evaluation were obtained from various sources described in section 2.2.4. Computing was supported by the Maryland Advanced Research Computing Center (MARCC). We would like to thank Hamada Badr and NASA LIS support team for help on model development, Guoyong Leng, Joseph Santanello, Patricia Lawston for many useful conversations related to this work.

Chapter 3

Assimilating GRACE into a Land Surface Model in the presence of an irrigation-induced groundwater trend[2]

Abstract

Assimilating terrestrial water storage (TWS) observations from the Gravity Recovery and Climate Experiment (GRACE) mission into Land Surface Models (LSMs) provides an opportunity to disaggregate and downscale GRACE information to finer scales and improve water component estimates in LSMs. However, the performance of GRACE data assimilation (GRACE-DA) is limited by the lack of representation of human activities in most LSMs. To simultaneously improve GRACE-DA and reduce the uncertainties in the modelled anthropogenic processes, we assimilate mascon-based GRACE TWS into the Noah-Multiparameterization LSM that includes groundwater extraction for irrigation. Simulations with and without GRACE-DA and with and without groundwater pumping for irrigation are performed to study the isolated and combined effects of

[2] This chapter has been published as: Nie, W., Zaitchik, B.F., Rodell, M., Kumar, S.V., Arsenault, K.R., Li, B., & Getirana, A.(2019). Assimilating GRACE into a Land Surface Model in the Presence of an Irrigation-induced Groundwater Trend. *Water Resources. Research*. http://doi.org/10.1029/2019WR025363

groundwater irrigation and GRACE-DA on water and energy fluxes over the High Plains Aquifer (HPA). The results reveal that the DA-only simulation may erroneously distribute increments across water storage components and affect the related fluxes through biased feedbacks, while the irrigation-only simulation may overestimate groundwater decline due to shortcomings in the irrigation and groundwater parameterizations. Assimilating GRACE when irrigation is simulated produces the best overall performance for water storage trends over the northern HPA. For the southern HPA, GRACE assimilation with irrigation performs similarly to irrigation-only simulation for water storage components and ET. GRACE assimilation also improves results in non-irrigated regions and can potentially alleviate the overestimation of groundwater trends in regions with greater irrigation uncertainties. This study highlights the potential to advance hydrological data assimilation in the context of anthropogenic water consumption and land-atmosphere interactions.

3.1 Introduction

Groundwater depletion threatens the sustainability of water resources and increases the vulnerability of society and ecosystems in regions around the world (Rodell et al., 2018; Scanlon & Faunt, 2012; Wada et al., 2010; 2012). For most depletion hot spots, groundwater-fed irrigation, often accompanied by episodic droughts, is the leading contributor to the groundwater depletion (Famiglietti, 2014; Feng et al., 2018; Joodaki et al., 2014; Richey, Thomas et al., 2015; Rodell et al., 2009; Scanlon & Faunt, 2012; Shamsudduha et al., 2012; Tangdamrongsub et al., 2018; Zhong et al., 2018). In recent decades, irrigation has expanded significantly (Nazemi & Wheater, 2015a; Steffen et al., 2011), accounting for more than 70% of freshwater withdrawals and more than 90% of total consumptive water use globally (Frenken & Gillet, 2012; Siebert et al., 2010). As agricultural production becomes increasingly dependent on groundwater resources (Siebert et al., 2010), accurate

accounting for groundwater use and availability is essential for predicting regional food and water supplies and enhancing drought preparedness.

Earth system models (ESMs) are powerful tools for studying the distribution of the water resources and their impact on hydroclimate over a broad range of scales. Within the past few years, several leading ESMs have been modified to directly or implicitly represent the influences of anthropogenic processes on the water cycle. Many of these efforts have focused on global hydrological models (GHMs), several of which now include explicit representation of anthropogenic water management (i.e. irrigation, reservoir operations). These balances are modeled as a function of grid cell specific parameters or by downscaling coarser scale inventory datasets. The WaterGAP Global Hydrology Model (WGHM; Alcamo et al., 2003; Döll et al., 2003; 2012; Eicker et al., 2014), the PCRaster Global Water Balance model (PCR-GLOBWB; van Beek et al., 2011; Wada et al., 2010; 2013; 2014), WBMplus (Wisser et al., 2010), H08 (Hanasaki et al., 2008a), and the World-Wide Water Resource Assessment model (W3RA; van Dijk et al., 2011) are examples of this type of model. Such GHMs are well-suited for global water resource analysis, in that they capture a wide range of processes and can be calibrated reasonably efficiently using available data. At the same time, the fact that GHMs use an empirical water budget approach, and the limitation that they are structurally inappropriate for coupling with atmospheric models, restrict their suitability for studying physical processes and climate feedbacks (Nazemi & Wheater, 2015a; Pokhrel et al., 2016).

In contrast to GHMs, land surface models (LSMs) are more physically based and are designed primarily for land-atmosphere coupling. There are many LSMs, such as the Community Land Model (CLM; Lawrence et al., 2011), the Variable Infiltration Capacity model (VIC; Liang, 1994), the Noah (Chen et al., 1996) and Noah-Multiparameterization models (Noah-MP; Niu et al., 2011), that have been modified to account for irrigation (Evans & Zaitchik, 2008; Lawston et al.,

2015; Leng et al., 2013; Ozdogan et al., 2010; Tang et al., 2009; Yilmaz et al., 2014). Few LSMs account for the source of irrigation water and close the water balance in a manner that includes accounting for water withdrawals. Efforts have been made in LSMs such as CLM (Leng et al., 2014), Noah-MP (Nie et al., 2018), HiGW-MAT (Pokhrel et al., 2015) and ORCHIDEE (De Rosnay et al., 2003) to include this accounting.

The GHMs and select LSMs that do represent water withdrawals are able to simulate the groundwater depletion trends in the presence of water use and source accounting. Several recent studies, however, have shown that they tend to overestimate the decline in certain key regions of groundwater depletion. This tendency has been attributed to the scarcity of observations for model calibration and to underlying uncertainties in model parameters (Nie et al., 2018; Scanlon et al., 2018; Tangdamrongsub et al., 2018). For example, WGHM and PCR-GLOBWB overestimated the total water storage decline in China's Hai basin, even when WGHM was specifically calibrated for this basin (Scanlon et al., 2018). Similarly, in a study of the United States High Plains Aquifer (Nie et al., 2018), Noah-MP was found to overestimate groundwater declines in Nebraska, particularly during post-drought groundwater recovery. These errors could be due oversimplification of subsurface hydrology in Noah-MP, problems with soil parameters, and/or errors in meteorological forcing datasets.

In parallel to recent progress on representing human water management, many researchers have worked to address model shortcomings in simulating hydrological states and fluxes through land data assimilation (Reichle, 2008). Land data assimilation systems introduce satellite observations of predicted model states to correct for model biases and protect against drift. Efforts to assimilate observations of terrestrial water storage (TWS) anomalies derived from the Gravity Recovery and Climate Experiment (GRACE) satellite system are particularly relevant to the problem of capturing

groundwater variability and trends (Zaitchik et al., 2008). GRACE (and, since 2018, the GRACE-Follow On mission (GRACE-FO)) was a twin satellite system that monitored variations in Earth's gravity field by measuring the distance between the two co-orbiting satellites via a microwave signal (Tapley et al., 2004). After removing the gravitational effects of atmospheric and oceanic circulations, the gravity anomalies over land are used to infer TWS anomalies (departures from the long term mean), which include changes from snow, soil moisture, groundwater, large lakes, rivers floodplains and wetlands. GRACE-derived products, with their ability to track both natural and anthropogenic influences on TWS, have been applied widely to investigate water budgets, groundwater change, and floods and droughts at a broad range of scales (Andrew et al., 2017; Felfelani et al., 2017; Getirana et al., 2017; Houborg et al., 2012; B. Li et al., 2012; Reager & Famiglietti, 2009; Shamsudduha et al., 2017; Solander et al., 2017; Tang et al., 2017; Thomas et al., 2017).

As powerful as GRACE observations have been for constraining estimates of TWS variability and trends at basin to continental scales, the application of GRACE to water management is limited by its coarse temporal and spatial scales and the fact that GRACE TWS is vertically integrated (Swenson & Wahr, 2006). This has motivated research efforts to disaggregate and downscale the GRACE TWS observation, including through data assimilation. Data assimilation disaggregates GRACE observed TWS vertically into snow, groundwater, soil moisture and surface water bodies and can downscale the observation horizontally into finer spatial scales, using model physics and forcing data to enhance the applications of the GRACE observation (Zaitchik et al., 2008). A number of studies have shown that GRACE data assimilation (GRACE-DA) with GHMs or LSMs can improve simulations of TWS and groundwater variability in major river basins in the U.S. (Eicker et al., 2014; Girotto et al., 2016; Schumacher et al., 2016; Zaitchik et al., 2008), India (Girotto et al., 2017), Australia (Schumacher et al., 2018; Tian et al., 2017), European countries (Li et

al., 2012; Tangdamrongsub et al., 2015) and globally (Li et al., 2019). GRACE-DA is also found to be advantageous for drought monitoring (Bernknopf et al., 2018; Houborg et al., 2012; Kumar et al., 2016; Li et al., 2012; Li & Rodell, 2015), as it provides information on groundwater conditions that are otherwise difficult to ascertain. However, the impacts of GRACE-DA on snow, river discharge and evapotranspiration (ET) are mixed (Forman et al., 2012; Girotto et al., 2017; Kumar et al., 2016).

In recent years, a number of studies have reported on efforts to optimize GRACE-DA performance by investigating 1) the spatial scales of assimilation (Eicker et al., 2014; Schumacher et al., 2016), 2) the representation of horizontal error covariance of models and GRACE observations (Forman & Reichle, 2013; Khaki et al., 2018; 2017), 3) the choice of GRACE products and scaling factors (Kumar et al., 2016; Schumacher et al., 2018) and 4) the choice of data assimilation techniques (Girotto et al., 2016; Khaki et al., 2017; Khaki et al., 2017; Schumacher et al., 2016). Despite these efforts, challenges persist, particularly for regions that have substantial TWS depletion. For instance, Girotto et al. (2017) found that assimilating a GRACE-based TWS declining trend into the Catchment Land Surface Model (CLSM; Koster et al., 2000) in northwestern India improved the model's simulation of TWS and groundwater trends, but it erroneously introduced a negative trend for ET. In reality, ET is enhanced by irrigation in northwestern India, but irrigation was not represented in the model. In order to capture the simultaneous decline in water storage and increase in surface soil moisture leading to enhanced ET, one either needs to apply multi-sensor joint data assimilation to constrain water storage components (Khaki & Awange, 2019; Tian et al., 2017; Zhao & Yang, 2018) and ET flux (Pipunic et al., 2008; Schuurmans et al., 2003), or modify the model to account for the groundwater pumping and irrigation. The latter approach has been applied in some GHM studies, but it has not been done for an advanced LSM.

In this paper, we report on the assimilation of GRACE TWS into an advanced LSM, the Noah-MP model, which has been modified to include simulation of anthropogenic water withdrawals and irrigation. Our objectives are two-fold: first, we investigate whether the performance of GRACE-DA can be improved by the addition of anthropogenic processes to this LSM; second, we examine to what extent GRACE-DA, in turn, can correct biases in the LSM's simulation of irrigation-induced groundwater decline. To isolate these impacts, we perform simulations with and without GRACE-DA and with and without simulation of groundwater pumping and irrigation. In situ groundwater observation and satellite-derived estimates of soil moisture and ET are used to evaluate the simulations. Key factors for successful data assimilation and implications for the application of the integrated Noah-MP modeling system are discussed.

3.2 Data and methods

3.2.1 GRACE TWS observations

This study makes use of the full record of level 3, version *RL05*, monthly gridded mascon GRACE data at 0.5° x 0.5° resolution derived by NASA Jet Propulsion Laboratory (JPL) (Wiese et al., 2016). We select this product because it outperformed other GRACE products in capturing the anthropogenic influences on groundwater over the High Plains in our previous study (Nie et al., 2018). Prior to data assimilation, the corresponding scaling coefficients are applied to GRACE TWS anomalies, which downscale the signal from its native 3° x 3° mascon blocks to the 0.5° x 0.5° grid resolution while conserving the total mass for each block (Wiese et al., 2016). The data assimilation using gridded GRACE data can be performed either in observation space, the resolution at which GRACE gridded data is provided, or in model space, the resolution at which the LSM runs. Here, building on a recent study (Kumar et al., 2016) using NASA's Land Information System (LIS;

Kumar et al., 2006), the simulations are conducted in model space, with the downscaled GRACE TWS anomalies being disaggregated to the model resolution (0.125 °) and then converted to total TWS by adding the mean of TWS from 2004 to 2009 provided by model's baseline open loop (OL) simulations with and without irrigation. This step is performed with the Land Data Toolkit (LDT; Arsenault et al., 2018). These pre-processed observations are then employed in the GRACE-DA simulation. We emphasize that the results presented here are specific to assimilating the JPL mascon product, but the data processing and assimilation method is applicable across other GRACE products.

3.2.2 The Noah-MP land surface model and irrigation scheme

We assimilate GRACE-derived TWS in simulations of the Noah-MP LSM, version 3.6, implemented within the framework of LIS, version 7.2 (software available at https://github.com/NASA-LIS/LISF/). Building on the Noah LSM (Ek et al., 2003), Noah-MP advances its structure by better accounting for vegetation, groundwater, and snow dynamics (Yang et al., 2011), and allowing for a combination of multi-physics options such as dynamic leaf, frozen soil permeability, groundwater, and runoff schemes.

In a previous study (Nie et al., 2018), a demand-driven, sprinkler irrigation scheme with groundwater pumping was introduced to Noah-MP in order to study the observed groundwater and TWS decline over the High Plains Aquifer. The Moderate Resolution Imaging Spectroradiometer – International Geosphere Biosphere Program (MODIS-IGBP; Friedl et al., 2010) land cover dataset (1 km) is used to provide the information for cropland or other potentially irrigated land class (e.g., grass) and the 500 m high-resolution Global Rain-fed, Irrigated, and Paddy Croplands dataset (GRIPC; Salmon et al., 2015) is used to supply the percent irrigated area within a model grid cell. The scheme then determines the timing of irrigation based on the greenness vegetation fraction

(GVF), derived from monthly MODIS NDVI composites and the MODIS-IGBP land cover data set at a 0.05° spatial resolution (Nie et al., 2018) with predetermined thresholds used to define the beginning and end of the growing season. We use our MODIS-derived monthly time-varying GVF dataset, rather than a climatological GVF dataset, as is often applied in land surface models, in order to capture the seasonal and interannual variability of vegetation. This is critical when simulating irrigation water use, especially under drought conditions. We note that the use of static GRIPC and MODIS-IGBP data assumes that land use change and irrigation expansion are negligible over the study period, but the use of time-varying GVF in the irrigation trigger still allows us to capture variability and trends in total actively irrigated areas, within the GRIPC and MODIS-IGBP constraints.

Once irrigation is triggered, the irrigation water amount is calculated as the volume required to bring root zone soil moisture deficit in the irrigated area up to field capacity and is applied at every time step between 0600 to 1000 LT local time until the soil moisture reaches field capacity. Maximum depth of the root zone is drawn from a static crop rooting depth table (Ozdogan et al., 2010), and single crop type (maize) is used for simplicity and because it is the major irrigated crop type in most of High Plains. The irrigation water source is accounted for by subtracting the applied irrigation water from the model's groundwater aquifer storage term. Both water table depth and groundwater storage are updated accordingly. Our demand-driven irrigation approach does not account for non-consumptive irrigation water use (i.e. return flow) and more complex irrigation practices, however, as described in Nie et al. (2018), the above-mentioned modifications were able to capture the effects of irrigation on groundwater and TWS decline over the southern High Plains compared to the simulations without irrigation. These simulations, however, failed to capture some key hydrological responses, such as a post-drought groundwater recovery in Nebraska in 2012-2015,

likely due to limitations in Noah-MP subsurface hydrology, uncertainties associated with soil parameters, and/or errors in the meteorological forcing dataset.

GRACE-DA has the potential to address these limitations, as GRACE-based TWS estimates serve as a constraint that could correct model limitations in simulating the water components while still allowing for the representation of human water practices. More importantly, the existence of irrigation processes helps to better partition the increments estimated from data assimilation into soil water and groundwater due to better representation of model error structure as compared to that without irrigation.

3.2.3 Data assimilation

To assimilate GRACE observation into LIS, we use an ensemble smoother, as described in Zaitchik et al. (2008) and Kumar et al. (2016). In this method, GRACE observations are assimilated into the model at the monthly scale, whenever the observation is available. The JPL mascon products are downscaled to 0.5° resolution by applying the corresponding scaling coefficients and further disaggregated into model space (0.125°) using nearest neighbor interpolation approach, at which the data assimilation is performed. The ensemble smoother creates "model predicted TWS observations" by averaging daily simulated terrestrial water storage in the model at three times (on the 5th, 15th, and 25th days) per month, roughly approximating the frequency of influential GRACE overpasses at any given location. These predicted observations are then used in the ensemble update, in which increments are determined on the basis of relative uncertainty between GRACE observations and the model ensemble:

$$X_{T+}^{i} = X_{T-}^{i} + K_T(Z_T^i - H_T X_{T-}^{i}) \tag{3.1}$$

where X_{T-}^{i} and X_{T+}^{i} denote the i^{th} ensemble member of the state vectors before and after the update for each grid cell. Here the state vector for each ensemble member includes soil moisture components for Noah-MP's four soil layers and one groundwater component. In our study region, snow and surface water have marginal contributions to the magnitude and variation of TWS, as shown in Getirana et al. (2017), so for simplicity they are excluded from the data assimilation structure. The state vector, as mentioned before, is collected as a three time point average sampled within the assimilation time window, T, which follows the calendar month. Z_T^i denotes the i^{th} ensemble member of the observed total TWS and H_T is an operation matrix to map the simulated state vector of water components X_{T-}^{i} onto total TWS in order to make it comparable to the GRACE observations.

The time-dependent Kalman gain matrix, K_T, determines the relative weights of uncertainties in the model and GRACE observations and is estimated by minimizing the error covariance of the updated state vector:

$$K_T = \sum_{XM}(\sum_{MM} + \sum_{ZZ})^{-1} \tag{3.2}$$

where Σ_{XM} denotes the error covariance between the state vector (X_{T-}^{i}) and model prediction of TWS ($H_T X_{T-}^{i}$), which provides the information to vertically disaggregate the increments into each water component; Σ_{MM} denotes the error covariance of the model predicted TWS while Σ_{ZZ} represents the error covariance matrix of the observation.

The way LIS assimilates GRACE observation into the system is through two iterations, which is described in detail in Kumar et al. (2016). During the first iteration, an ensemble of model states is propagated forward for a month without data assimilation and the model predicted monthly

TWS ($H_T X_{T-}^i$) is estimated based on TWS simulated on the 5th, 15th and 25th day of the month. Next, the assimilation increments $K_T(Z_T^i - H_T X_{T-}^i)$ are calculated for the given month. Then, during the second iteration, the ensemble is reinitialized at the beginning of the month and the assimilation increments are divided by the number of days in that month and are evenly applied on a daily basis. The updated model states at the end of the second iteration then serve as the initial conditions for the first iteration for next month, and the above process repeats. When the irrigation option is active with data assimilation, it is applied in both iterations. Applying irrigation and groundwater withdrawal in the first iteration ensures that the impact of irrigation on TWS is included when calculating the predicted observation state vector for GRACE data assimilation. Applying on the second iteration allows for irrigation processes to be informed by GRACE updates to water storage terms. This second iteration is what is included in the final model output. GRACE-derived TWS anomalies are assimilated without application of CDF-matching, in order to preserve information on both amplitude and temporal variability of TWS.

Similar to previous GRACE-DA studies (Girotto et al., 2016; Kumar et al., 2016; Li et al., 2012; Zaitchik et al., 2008), we represent model uncertainties through an ensemble generated by applying perturbations to meteorological forcing parameters and model prognostic state variables, as listed in Table 3.1. Cross correlations of the perturbations are designed within the forcing fields and state vectors based on known associations between meteorological variables and physical links between subsurface water storage components (Reichle et al., 2007). Horizontal correlation lengths of the perturbations were set to 2° in order to represent the error scale of precipitation dynamics (Reichle & Koster, 2003). The temporal correlation of the perturbations was chosen to be 3 days for the forcing fields and 1 day for the model states via a first-order autoregressive model. The perturbation parameters for the four soil moisture states are scaled by the layer thickness so that the

standard deviation of total water volume applied to each layer is constant and identical to that of the groundwater perturbation (Ryu et al., 2009). For the observation error, we simply use a uniform observation error standard deviation of 20 mm, following Wahr et al. (2006), considering both limited uncertainty variability in our relatively small study region and the challenge of providing an accurate estimation of GRACE uncertainty in model space. This assumption of spatially uniform GRACE errors has been proved to produce a reasonable assimilation result in previous studies (Houborg et al., 2012; Kumar et al., 2016; Zaitchik et al., 2008). An ensemble size of 20 members is used for all model integrations as suggested by previous studies (Kumar et al., 2016). Quality-control checks are imposed to constrain the range and to prevent unphysical perturbed values, and perturbation bias corrections following Ryu et al. (2009) were applied to all the perturbed forcing and state variables to avoid biases introduced by the nonlinear processes in the model.

Table 3. 1 Ensemble perturbation parameters in the assimilation simulations. Multiplicative (M) or additive (A) perturbations are applied to the meteorological ("Met") fields: incident shortwave radiation (sw), incident longwave radiation (lw), precipitation (precip), and the model states: four layers of soil moisture content (smc (1) - (4)) and groundwater storage (gws).

Variable	Type	Std dev	Perturbation cross-correlations				
Met- forcings			sw	lw	precip		
sw	M	0.3 (dimensionless)	1	-0.5	-0.8		
lw	A	50 W m^{-2}	-0.5	1	0.5		
precip	M	0.3 (dimensionless)	-0.8	0.5	1		
LSM States			smc (1)	smc (2)	smc (3)	smc (4)	gws
smc (1)	A	1.0 x 10^{-4} m^3 m^{-3}	1	0.6	0.4	0.2	0
smc (2)	A	3.3 x 10^{-5} m^3 m^{-3}	0.6	1	0.6	0.4	0
smc (3)	A	1.7 x 10^{-5} m^3 m^{-3}	0.4	0.6	1	0.6	0
smc (4)	A	1.0 x 10^{-5} m^3 m^{-3}	0.2	0.4	0.6	1	0
gws	A	0.01 mm	0	0	0	0	1

3.2.4 Experimental design

In order to study the isolated and combined impacts of irrigation and GRACE-DA in simulating the modeled water and energy fluxes, four simulations were conducted for the main

GRACE data period (April 2002 to June 2017) at 0.125° spatial resolution with a 15-min time step. All four simulations drew initial conditions for April 2002 from a 63-year Noah-MP spin-up simulation (three times over the period of 1995-2015). Next, two open loop (OL) simulations were conducted: one without irrigation (referred to simply here as OL) and one with irrigation (OL_{irr}). GRACE TWS anomalies were then converted to the absolute TWS estimates (in model space) by adding the 2004-2009 mean TWS from the open loop simulation to the gridded GRACE TWS anomalies in each month, as introduced in Section 2.1. Finally, two GRACE-DA simulations were performed without and with irrigation (DA & DA_{irr}) by using the converted GRACE total TWS based on OL and OL_{irr}, respectively. Comparing OL_{irr} and DA experiments with the baseline OL simulation enables us to see the isolated impact of irrigation or data assimilation while comparisons between DA_{irr} and OL can provide information on how the simulation of human impacts is influenced by GRACE-DA.

All simulations used atmospheric forcing data from the North America Land Data Assimilation System-Phase 2 (NLDAS-2; Xia et al., 2011; 2012) at 0.125° spatial resolution. The model is configured using the MODIS-IGBP land cover at 1 km, State Soil Geographic (STATSGO; Schwarz & Alexander, 1995) soil texture at 1km and the Shuttle Radar Topography Mission elevation at 30 m (SRTM30; Farr et al., 2007). All the datasets are pre-processed to model space by LDT prior to model execution in LIS. For simulations including the irrigation algorithm, GRIPC was used to provide the percentage of irrigated area within each model grid cell (Figure 3.1 (a)). Within irrigated areas, irrigation is triggered by a soil moisture threshold in combination with a GVF requirement: irrigation is only active when GVF exceeds a certain threshold, which is estimated as a function of its range (Nie et al., 2018). This makes the simulation of irrigation highly

sensitive to GVF, so we use monthly time-varying GVF estimates rather than climatological GVF, as is typically used in Noah-MP.

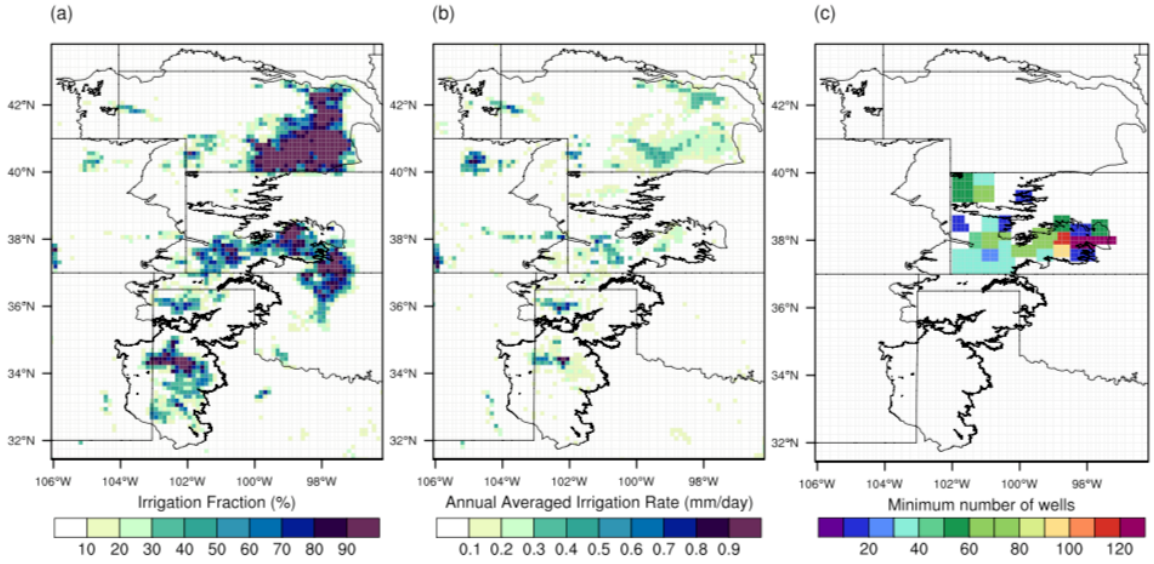

Figure 3. 1 (a) The global rain-fed, irrigated and paddy croplands (GRIPC) data set irrigation fraction over the HPA, (b) the simulated long-term averaged annual irrigation rate in OLirr, and (c) Kansas counties within HPA with the corresponding annual minimum number of the wells that are used for water level measurements, along with the black points indicating the locations of all the observation sites.

3.2.5 Evaluation data and metrics

Groundwater: County averaged annual groundwater levels for Kansas, covering the period of 2003-2016 were obtained from The Water Well Completion Records database (WWC5) of Kansas Geological Survey (KGS, http://www.kgs.ku.edu), and were used to evaluate the groundwater level estimates. Thirty counties, partially or fully located within the High Plains with relatively abundant observations, were selected to represent the area average. As shown in Figure 3.1 (c), most of the counties have measurements from more than 30 wells every year. The simulated groundwater storage is converted to groundwater levels by using the USGS provided specific yield dataset (Gutentag et al., 1984). We then aggregated the simulated water levels for each county and converted

both simulated and observed water levels to anomalies by removing the long-term mean respectively in order to make them comparable. We use the KGS dataset for its extensive spatial coverage and clear documentation.

Soil Moisture: The daily near surface soil moisture measurements from two satellite microwave instruments at 0.25° spatial resolution, covering the period of 2010-2017, are used for surface soil moisture evaluation in this study: 1) the Advanced Scatterometer (ASCAT; Bartalis et al., 2007) sensor onboard of the EUMETSAT MetOp satellite and 2) the Soil Moisture Ocean Salinity (SMOS) mission (Kerr et al., 2001). The ASCAT retrievals are obtained through the Soil Moisture Operational Products System (SMOPS; Liu et al., 2016) of NOAA/NESDIS. The ASCAT product has been found to be more promising than many other microwave-derived soil moisture products at detecting irrigation features (Kumar et al., 2015). The SMOS Level 3 daily soil moisture product is obtained from Centre Aval de Traitment des Données (CATDS; Jacquette et al., 2010), and has been found to have comparable skill to ASCAT for the High Plains (Kumar et al., 2018). The simulated surface layer (10 cm) soil moisture is aggregated to 0.25° spatial resolution and the three datasets are then regridded to a common grid. Analysis is performed only on the dates that all three data are available.

Evapotranspiration (ET): We compare model simulated ET to diagnostic ET estimates from the Atmosphere-Land Exchange Inverse (ALEXI; Anderson et al., 2007) surface energy balance model, implemented using MODIS land surface temperature estimates (Hain & Anderson, 2017), and aggregated from its native 0.04° to our model based 0.125° spatial resolution.

The evaluation is performed for groundwater, surface soil moisture at the respective observation's spatial resolution and ET at the model's spatial resolution. Statistical skill metrics

include the correlation coefficients for the original (R), deseasonalized (R_d) and detrended (R_{dd}) time series with 95% significance tested using Fisher's z transform test (Fisher, 1921). The root-mean-square difference, for the original (RMSD) or unbiased (ubRMSD) time series with 95% significance tested by using paired sample t-test with temporal correlation accounted (Entekhabi et al., 2010). R_d is calculated by removing the seasonal cycle from the time series, whereas the seasonal cycle is calculated as the multiyear average of each calendar month (2002-2017). R_{dd} is then calculated by additionally removing the trend of the time series, calculated by linear least squares fit. We use R to examine the overall mismatch between observations and simulations in terms of trend, seasonality and interannual variability, while R_d and R_{dd} are used to examine the mismatch in trend or interannual variability. The ubRMSD is computed as the RMSD after removing the long-term mean difference. The ubRMSD metric is used for evaluation of water storage terms—TWS and groundwater—because we are concerned with variability in these terms and not with the absolute match to arbitrary model water storage reservoirs. We use RMSD to evaluate the simulation of energy fluxes, such as ET, due to the importance of capturing the magnitude of these fluxes for water balance monitoring and analysis of land-atmosphere coupling.

The Theil-Sen trend metric is used to estimate trends in TWS and ground water and the Mann-Kendall test is performed to identify the statistical significance of the trends. We also use a bias metric to examine the over/underestimation of ET and the triple collocation approach (Dorigo et al., 2010; Stoffelen, 1998) to estimate the root mean square error (RMSE) of simulated surface soil moisture.

3.3 Results and discussion

3.3.1 Terrestrial water storages

Figure 3.2 shows the monthly TWS time series averaged for the High Plains Aquifer (HPA) region and for HPA subareas in Texas, Kansas and Nebraska. These three states dominate total groundwater irrigation in the HPA, according to USGS water use records (https://water.usgs.gov/watuse/data/). In general, the model captures the TWS seasonality and the magnitude of its variation, which is primarily a function of natural fluxes of precipitation and evapotranspiration. However, the OL simulation, which does not include any irrigation or extractions for the groundwater reservoir and no assimilation, failed to simulate the TWS decline for Texas especially after the onset of a severe drought in 2011 (Nielsen-Gammon, 2011). The OL simulated trend for the entire GRACE period for Texas and Kansas was -0.25 cm/yr and -0.07 cm/yr, respectively, compared to GRACE observed trends of -1.54 cm/yr and -0.68 cm/yr. Introducing a groundwater irrigation scheme into the model, OL_{irr}, brings the model into closer agreement with the GRACE observations for Texas (-1.67 cm/yr), and it also results in a negative trend for Kansas (-1.25 cm/yr) that resembles but exceeds the GRACE observation. However, the representation of irrigation also leads to distinct overestimation of decline in Nebraska (-0.46 cm/yr), where GRACE sees a positive trend of 0.98 cm/yr. The positive GRACE trend in this region has been explained as a response to depressed water from the deep drought at the start of the GRACE period (Rodell et al., 2018). Averaged over the HPA, OL actually agrees quite well with GRACE observation. However, this is a result of offsetting errors: OL underestimates TWS decline for Texas and Kansas (Figure 3.2c, e) due to the absence of water withdrawals in the simulation, and underestimates post-drought TWS recovery in Nebraska—likely due to model shortcomings in representing subsurface hydrology such as recharge process (Figure 3.2g). In OL_{irr}, irrigation improves the simulation for Texas and Kansas but exacerbates OL's problems with simulating post-drought TWS recovery in Nebraska. This causes OL_{irr} to have larger TWS errors than OL when averaged over HPA, when compared to GRACE. However, this makes Nebraska a useful example

of how data assimilation can correct for model bias in TWS, while including representation of irrigation in the model can be key to achieving reasonable partitioning of changes in storage vertically through the soil column and groundwater.

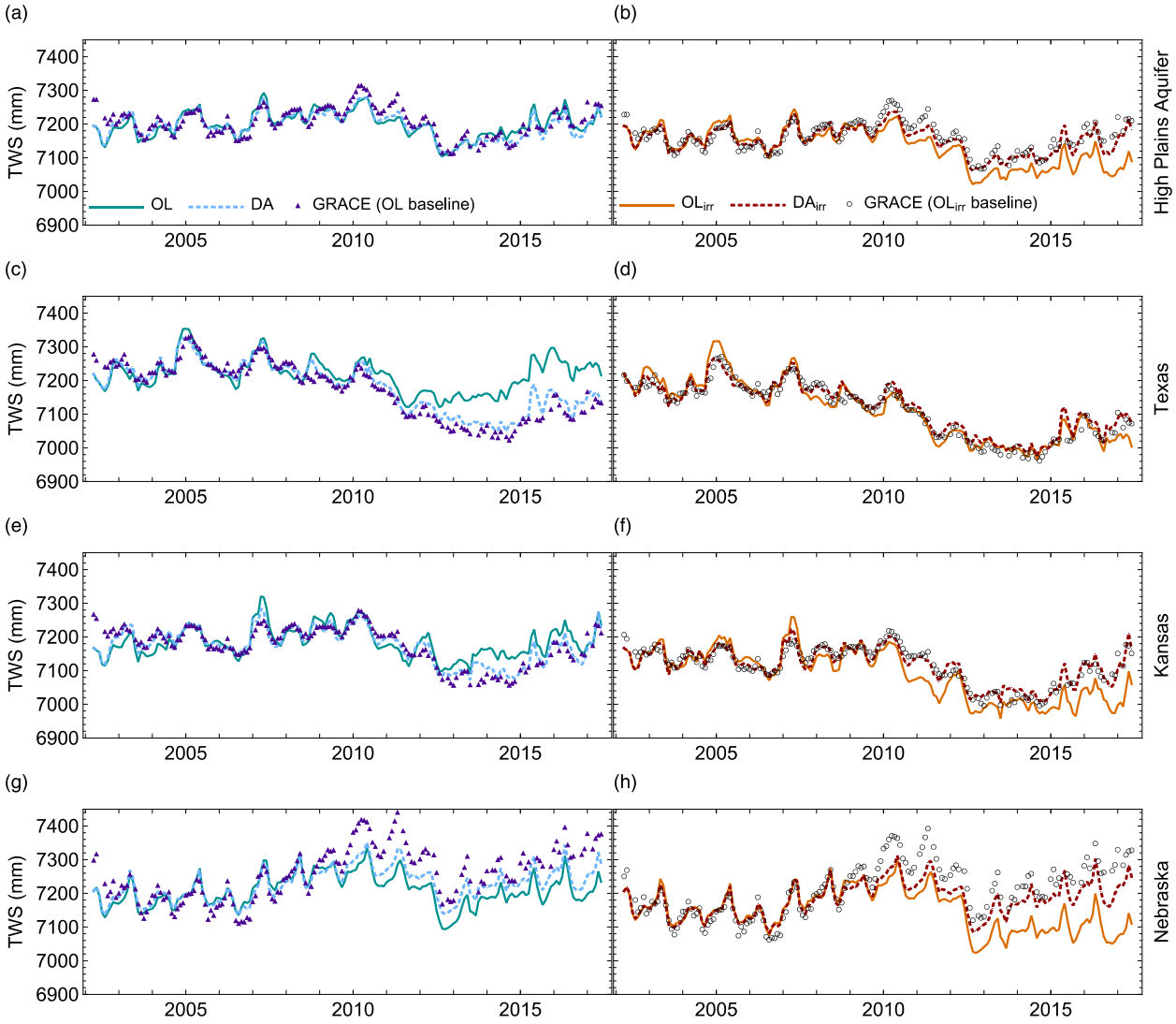

Figure 3. 2 Time series comparison of TWS estimates from the simulations and GRACE observations for (a, b) the entire High Plains Aquifer region, and for the HPA portion of (c, d) Texas, (e, f) Kansas, and (g, h) Nebraska. The JPL GRACE mascon anomaly data with the scaling coefficients are used for evaluation, which are converted into the TWS estimates by adding the TWS baseline simulated by OL (purple triangles, left panel) and OLirr (open circles, right panel), respectively.

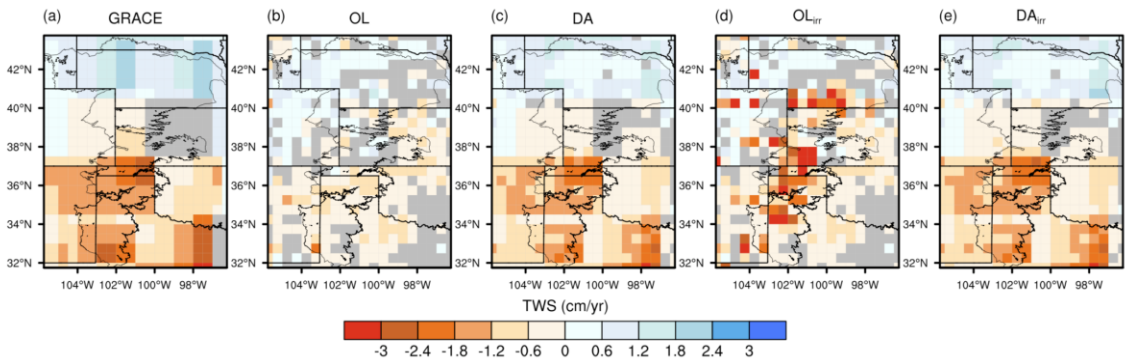

Figure 3. 3 TWS trends in (a) GRACE observations, and (b-e) the model simulations. Gray colors indicate locations where the trends are not statistically significant at 5% significance level, using the Mann-Kendall trend test.

The spatial pattern of the trends can be seen in Figure 3.3. All the simulations are upscaled to 0.5° to match the GRACE data resolution before calculating the trends. An overestimation of the decline by OL_{irr} in Nebraska mainly occurred in the southern portion, where the simulated irrigation rate was relatively high (Figure 3.1b). This localization is consistent with the spatial pattern of OL_{irr} simulated TWS decline in Kansas and Texas: in all cases, OL_{irr} shows trends that are steeper than GRACE in intensively irrigated areas. This is in large part because we are mapping at the gridded GRACE resolution of 0.5°, when the true resolution of the product is closer to 3°. This result illustrates the fact that irrigation represents fine scale variability as compared to GRACE, such that assimilation of GRACE to a model that includes irrigation withdrawals has the potential to downscale GRACE TWS estimates. At the same time, GRACE-DA provides an observational constraint that corrects the errors of both OL and OL_{irr}. As shown in Figure 3.2, the DA simulated TWS time series falls between that of the OL and the GRACE observations generated by adding the OL baseline. In this application, the assimilation results generally fall closer to GRACE than to the OL, reflecting the fact that error estimates for GRACE observations are smaller than that of the model, which are represented by the ensemble spread. The same holds for simulations in the

presence of irrigation (i.e. OL_{irr} and DA_{irr}). Both DA and DA_{irr} improve the overall TWS trend agreement with GRACE observations (Figure 3.3).

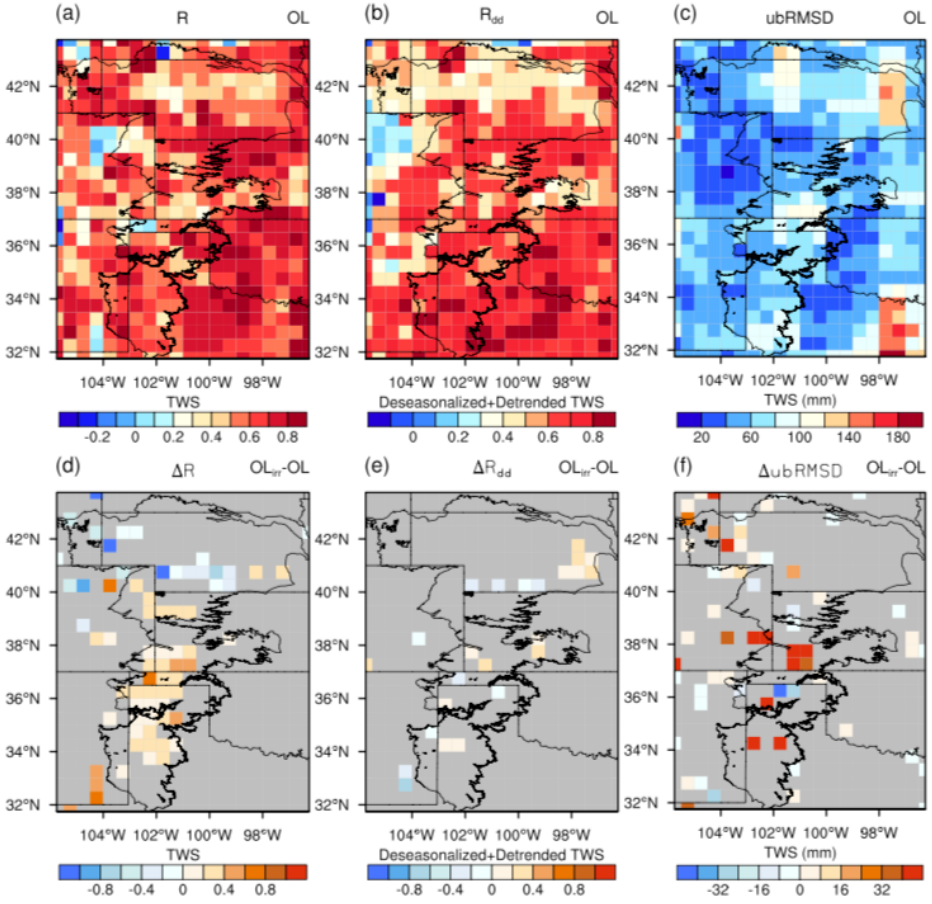

Figure 3. 4 (a) Correlation (R), (b) deseasonalized and detrended correlation (R_{dd}), and (c) unbiased RMSD (ubRMSD) for OL simulated TWS against GRACE observations. Differences in correlation (d), deseasonalized and detrended correlation (e), and ubRMSD (f) between OL_{irr} and OL for TWS. Gray colors in correlation (d-e plots; Fisher's z transform test) and ubRMSD (f plot; paired sample t-test) difference plots indicate locations of insignificant differences at 95% significance level accounting for temporal correlation.

Figure 3.4 shows overall high correlation values in TWS (R_d) and the deseasonalized and detrended TWS (R_{dd}) for the OL simulation against GRACE observations, indicating that the model can capture both seasonality and interannual variability relatively well. This speaks to the quality of the NLDAS-2 meteorological forcing dataset and to the general fidelity of Noah-MP in representing

the large-scale natural water balance. The relatively lower R and R_{dd} values in Nebraska reflect potential weakness in model parameter or physics in this region. OL_{irr} substantially improves the correlation with GRACE for the HPA by capturing the declining trend. It has only a minor impact on improving interannual variability, with the exception of eastern Nebraska around the aquifer boundaries. Despite the improvements in R and R_{dd}, OL_{irr} leads to greater ubRMSD because the magnitude of the trend is overestimated. This is due to uncertainties in irrigation parameters and the sensitivities to the ensemble perturbation of soil moisture, as further discussed in section 3.2.

Despite the fact that there are certain improvements in OL_{irr} and DA simulations against OL for different regions, there are major limitations in both irrigation-only and DA-only systems. For instance, GRACE observes a significant TWS decline in Texas during the GRACE period, and this decline is attributable to the combination of sporadic drought, especially the one during 2011-2012 (Figure 3.2 (c, d)) and the withdrawal of groundwater to support the irrigated agriculture (Rodell et al., 2018; Scanlon & Faunt, 2012). OL is not able to replicate this trend due to the lack of representation of groundwater withdrawal. In this case, via data assimilation, the DA simulation corrects against this error by applying positive increments in the early simulation period and negative increments for the last few years. The sign of these increments is sensitive to the period used to define the TWS baseline. Although these increments correct the TWS trend, they are correcting without proper process representation and can, therefore, introduce erroneous trends to the hydrological fluxes, such as groundwater recharge or ET. For example, Girotto et al. (2017) found that assimilation of GRACE in heavily irrigated areas in India introduces a negative trend in ET, whereas, in reality, ET is observed to be enhanced by irrigation for that region. In these cases, restoring the main trend signal through introducing irrigation processes can prevent the system from introducing biases and improves the representation of surface fluxes by reasonably redistributing water through the model's storage compartments.

On the other hand, in places like Nebraska, e.g., in this study, it is possible that an OL simulation has errors that lead to overestimated TWS decline or failure to capture an increase in TWS (Figure 3.2 (g, h)). In that context, introducing groundwater withdrawal exacerbates model TWS errors, even though the process representation is an improvement, in principle. For cases like this, assimilating GRACE into a simulation with irrigation can constrain TWS while preserving simulation of key processes such as irrigation. This does come at the cost of sacrificing the local water balance, since assimilation adds or removes water from the system.

Figure 3. 5 Monthly time series of differences between DA/OL_{irr}/DA_{irr} and OL simulations for averaged High Plains, Texas, Kansas, and Nebraska groundwater storage, (a-d) GWS, and (e-t) the four model layers of soil water storage, SWS.

3.3.2 Groundwater

Figure 3.5 (a-d) compares the OL simulated groundwater and soil moisture components against the other three simulations. For this region surface water and snow are very limited and have minor impact on the total water storage. Groundwater is the largest storage term in Noah-MP, and the magnitude of difference in ground water storage between OL_{irr}/DA/DA_{irr} and OL is larger than that in the soil moisture layers. The differences between simulations, however, is striking. As expected, OL_{irr} shows the steepest groundwater decline relative to OL. In contrast, the DA simulation knows nothing of irrigation, and the groundwater trends in that simulation are less dramatic and more variable. For Nebraska, the DA simulation has a positive difference in groundwater trend from OL. DA_{irr}, meanwhile, falls in between OL_{irr} and DA; this simulation withdraws water from groundwater, like OL_{irr}, but GRACE serves as a constraint on water withdrawal, and the Ensemble Kalman Smoother update spreads the increments across groundwater and soil moisture storages.

The way that DA and DA_{irr} partition the groundwater increments differently also affects groundwater seasonality. Figure 3.6 shows the averaged seasonal cycle of groundwater storage for the three states within HPA. The least square fit linear trend is removed prior to the calculation of the seasonal cycle. Stronger seasonality is observed in Nebraska and Kansas than in Texas in the OL simulation, likely due to differences in meteorology and the recharge process associated with deep soil conditions. Simulation of seasonal groundwater dynamics is strongly impacted by the irrigation process, as shown in OL_{irr}, leading to a much sharper decline from June to September due to seasonal irrigation, along with a much larger amplitude. DA_{irr} is similar to OL_{irr} in phase and amplitude, but with slightly muted variability due to the assimilation of GRACE-TWS, particularly for Texas and Kansas. However, DA without the irrigation process failed to match this seasonality,

and even introduced a significant phase shift in the timing of groundwater maximum in Texas and Kansas likely, due to unreasonable partitioning of increments between groundwater and other storage terms. The interannual variability of groundwater storage varies from 30 mm to 120 mm depending on the region, time and simulation settings. This interannual variability is, in general, substantially wider than seasonal storage variations and the differences between the seasonal cycles among simulations.

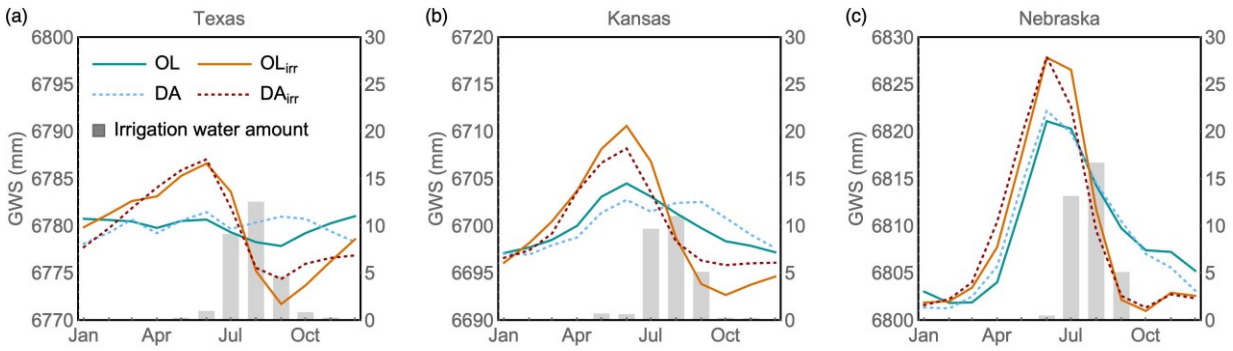

Figure 3. 6 Average seasonal cycle of the groundwater storage from the simulations for (a) Texas, (b) Kansas, and (c) Nebraska within HPA. Also shown is the simulated average monthly irrigation water amount (grey bar plots, right axis).

To evaluate the groundwater interannual variability at a spatial scale that is large enough to reflect the GRACE signal but small enough to examine the skill of downscaling GRACE via the LSM, we selected KGS groundwater level measurements and calculated the winter annual groundwater anomalies for 30 counties within the HPA domain by averaging the available in situ measurements within each county. We note that there are uncertainties in the well measurements, which may come from the variation of the number of wells included in each year, and in the locations and usage of the wells. Therefore, we only include the counties that have relatively more measurements and limited interannual variation in well numbers.

The groundwater recharge to HPA in Kansas is significantly impacted by the existence of thick alluvial aquifers that overlies the HPA. This complicates the relationship between total groundwater storage and groundwater decline in the HPA (Brookfield et al., 2018). Therefore, the extent to which the county averaged water level can represent the total groundwater storage change is sensitive to the depth and location of the wells. Comparisons between simulated and observed groundwater levels are also complicated by model limitations, including uncertainties in irrigation parameters, the irrigation map, and relevant soil and vegetation parameters, and/or the use of highly simplified irrigation rules that do not account for change in irrigation practices in response to water management. To minimize the impact of the extreme cases brought by these uncertainties, we compare the four simulations to county groundwater observations using median metrics for water level correlation, ubRMSD, and trend (Table 3.2). The ratio of the number of counties that yield improvement to all 30 counties is shown as percent improvement in Table 3.2, along with the range of improvements.

The median of the correlation between OL and observations is only 0.3. Correlation is improved for 73% of the counties, with a median of 0.68 when both irrigation and data assimilation are included (DA_{irr}), which is the highest among the simulations. Accounting for groundwater withdrawal also reduces the ubRMSD, as does data assimilation, but the combination of withdrawal and data assimilation does not offer additional improvements; the lowest ubRMSD and the largest percent improvement was found for OL_{irr}, followed by DA_{irr}. Interestingly, all simulations significantly underestimate groundwater decline relative to observations. OL_{irr} is the closest to observations, simply because it has the steepest trend (-2.8 cm/yr), but DA_{irr} yields the largest percent improvement (77%), even though it pulls the trend further from observations by moderating the simulated decline (-2.2 cm/yr in DA_{irr}). The fact that data assimilation degrades simulation of trend relative to observations likely reflects the low resolution of the GRACE

observations and the effects of the processing approach, both of which tends to attenuate localized mass signals as it smooths them over a large area (Landerer & Swenson, 2012). At the same time, the severity of decline at the KGS sites relative to GRACE observed trends might reflect the fact that wells are clustered in active agricultural areas. While the wells appear to be widely distributed across reporting counties (Figure 3.1 (c)), site specific conditions might lead to overrepresentation of active pumping zones. Despite the slight degradation in DA_{irr} compared to OL_{irr} with respect to the long-term trend, DA_{irr} performs better than either OL or DA for Kansas according to other metrics, and it can potentially compensate for OL_{irr} trend uncertainties in cases where groundwater decline is overestimated by irrigation. We also note that the impact of data assimilation on groundwater trend is highly dependent on the scale of irrigation relative to the scale at which GRACE observations are assimilated, and that assimilation algorithms and GRACE-FO retrieval methods currently under development may allow for more effective localization of the trend signal.

Table 3. 2 Median metrics, along with the percent of improvement among all the counties showing Correlation, unbiased RMSD (ubRMSD), and annual trend values for groundwater levels for Kansas counties, comparing each simulation to KSGS observations. The range of improvements are shown in the parenthesis.

	Obs	OL	DA		OL_{irr}		DA_{irr}	
			Median	Percent Improved	Median	Percent Improved	Median	Percent Improved
Correlation	—	0.3	0.63	73% (0.04 - 1.2)	0.57	83% (0 - 1.3)	0.68	73% (0.05 - 1.3)
ubRMSD (m)	—	1.37	1.3	77% (0 - 0.49)	1.22	93% (0 - 1.41)	1.28	77% (0 - 0.54)
Trend (cm/yr)	-17.5	-0.6	-1.8	73% (0.1-12.8)	-2.8	73% (0 - 42.8)	-2.2	77% (0.1 - 13.0)

3.3.3 Soil moisture

As shown in Figure 3.5, the simulation of irrigation in OL_{irr} and DA_{irr} has the effect of seasonally wetting the root zone soil moisture layers—approximately the top one meter of the soil column (Figure 3.5 e-p). This is expected, since the irrigation scheme draws groundwater and applies it at the surface, and is designed to bring root zone soil moisture to saturation in any irrigation event. In the bottom soil layer (Figure 3.5 q-t), irrigation alone (OL_{irr}) has very little impact: drainage from

added irrigation water is small on account of plant transpiration, while changes due to pressure gradients with the underlying aquifer have some impact during dry periods but do not impart any significant trend. In contrast, data assimilation has a larger total volume impact on soil moisture in deeper soil layers, due to the fact that the deeper layers are thicker in Noah-MP. It is also important to note that the way the Kalman gain matrix partitions the increments into groundwater and soil moisture is strongly impacted by the structure of the model. For example, for Texas, when GRACE observes a decreasing trend while the model does not (as in the DA simulation), the system tends to spread this trend to soil moisture at all layers, resulting in synchronous wetter conditions in first 2~3 years and long-term drier conditions during the rest of the period (Figure 3.5 (f,j,n,r)). The distribution of assimilation increments to deep soil moisture in the DA simulation is one of the pitfalls of assimilating GRACE in the absence of appropriate process representation. TWS reductions caused by groundwater withdrawal are erroneously applied as negative increments to soil moisture layers, which would negatively impact the contribution of GRACE-DA to applications like drought monitoring. Accounting for that irrigation process, DA_{irr} simulates wetter conditions in the root zone, with a strong growing season signal due to irrigation. Negative trends are confined to the deepest soil layer, which is 1 meter thick, and groundwater. Accounting for irrigation when calculating the model predicted TWS baseline—i.e., the effect of representing irrigation and groundwater withdrawal in the first iteration of data assimilation before increments are applied—also reduces the potential for data assimilation to introduce erroneous negative increments due to trend mismatch between model predictions and GRACE observation. This can be seen when comparing DA_{irr} to DA.

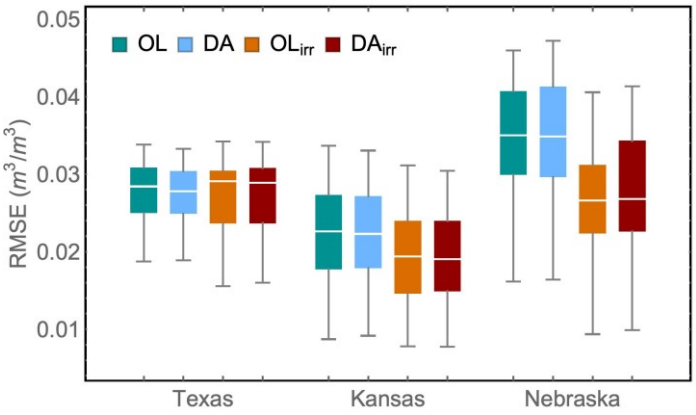

Figure 3. 7 TC-based RMSE estimates of simulated surface soil moisture anomalies for intensively irrigated grid cells (IRfrac > 50%) within Texas, Kansas and Nebraska.

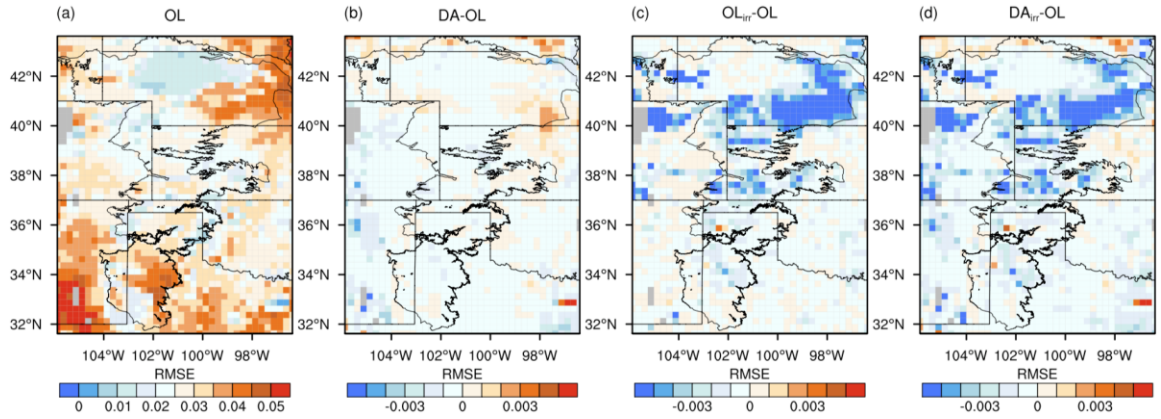

Figure 3. 8 (a) TC-based RMSE estimates of simulated surface soil moisture anomalies for OL and (b, c, d) the difference of RMSE between OL and the other three simulations. Grey colors in (a) indicate locations either masked out for TC analysis due to limited sample size or non-physical RMSE.

Evaluating soil moisture is challenging, as in situ observations are generally limited and unrepresentative of the 0.125° grid box of the simulations, while gridded soil moisture products often represent different soil depth, spatial scales, varying dynamic ranges and climatologies. Given the difficulty of finding a high quality reference dataset, we evaluate the modeled surface soil moisture (10 cm thick surface layer) by applying a triple collocation (TC) approach (Dorigo et al., 2010; Stoffelen, 1998) to estimate the root mean square error (RMSE) of simulated surface soil

moisture. Four sets of triplets of model-based, ASCAT, and SMOS soil moisture anomalies are created via the removal of a single long-term mean. Intensively irrigated grid cells with irrigation fraction higher than 50% are selected for the three states within HPA, and the distribution of the TC-based RMSE for these grid cells is shown in Figure 3.7. In general, TC-based RMSE is much larger for Texas and Nebraska than for Kansas in the OL simulation. With the presence of irrigation, both OL_{irr} and DA_{irr} show significant RMSE reduction in Nebraska, followed by Kansas and Texas. Data assimilation does little to alter the distribution, as the RMSE in OL and OL_{irr} is identical to DA and DA_{irr}, respectively. The improvement of surface soil moisture condition attributed to irrigation is also observed over the HPA domain as a whole. Figure 3.8 shows the spatial distribution of TC-based RMSE for OL simulation and the RMSE difference between OL and the other three simulations. Due to their representation of irrigation, both OL_{irr} and DA_{irr} show a substantial decrease in RMSD over HPA, in general, with the greatest improvement in intensively irrigated regions of Nebraska and Kansas.

3.3.4 Evapotranspiration

To quantify the impact of irrigation and data assimilation on evapotranspiration, we compare the simulated ET to satellite-derived ET estimates from ALEXI at monthly scale (Figure 3.9). Noah-MP without irrigation captures the variation of ET quite well, with averaged correlation coefficient of 0.88 for the HPA. After removing the seasonal cycle, the correlation coefficient degrades to 0.62, indicating less agreement in representing ET interannual variability between ALEXI and Noah-MP, which may stem from uncertainties from both sides; ALEXI ET anomaly estimates are subject to considerable uncertainty in their own right. Although water storage components are significantly modified by irrigation and data assimilation, the impact of either on the area-averaged ET is small.

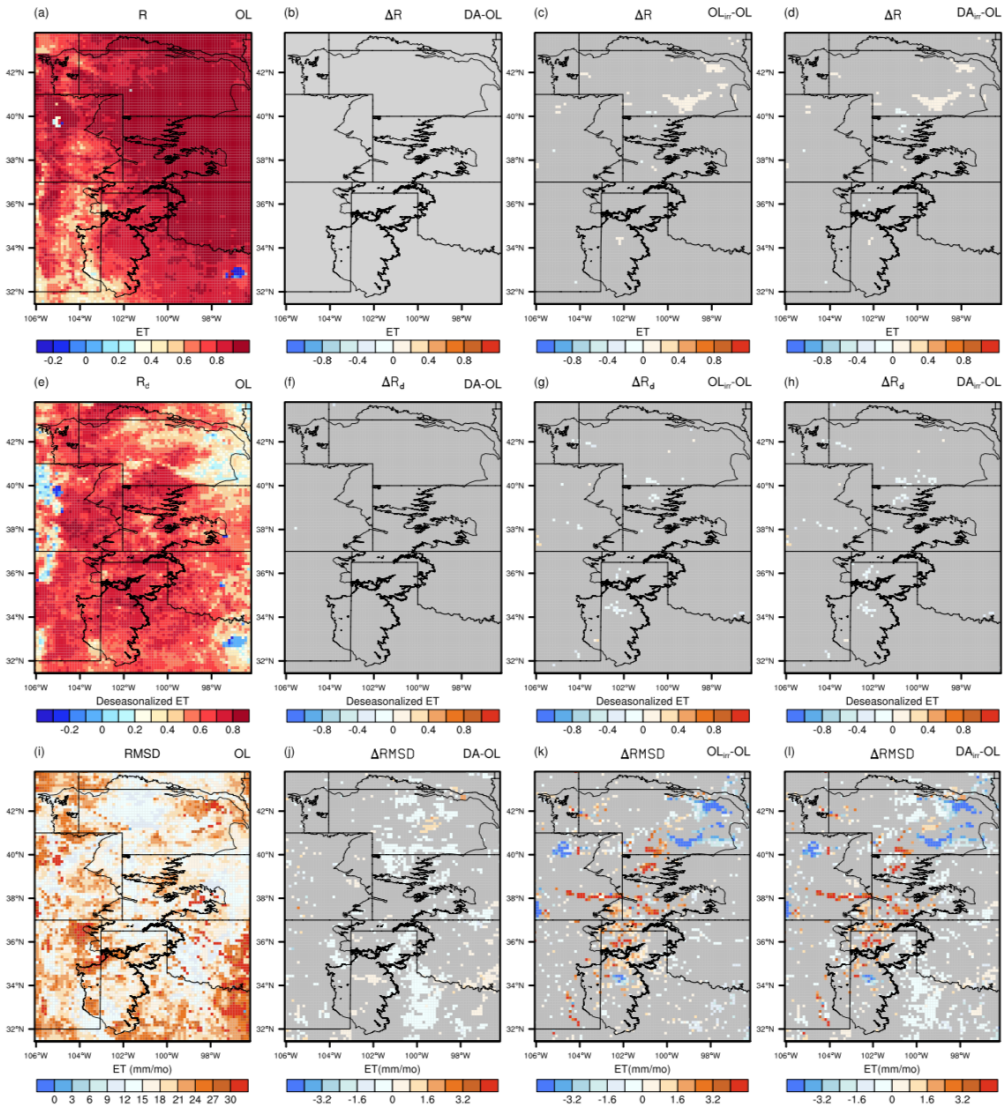

Figure 3. 9 (a) Correlation (R), (e) deseasonalized correlation (R_d), and (i) RMSD for OL-simulated evapotranspiration (ET) against ALEXI estimates. Differences are shown in correlation (b-d), deseasonalized correlation (f-h), and ubRMSD (j-l) between DA/OLirr/DAirr and OL for ET. Gray colors in the correlation difference plots (figures b-d and f-h) indicate locations where the correlation differences are not statistically significant at 5% significance level, using Fisher's z transform test with temporal correlation accounted.

Statistically significant improvements in ET correlation are only seen in southern Nebraska from OL_{irr} and DA_{irr}, while the correlation for these two simulations decreases in intensively irrigated parts of Kansas and Texas after removing the seasonal cycle, which implies that simulated irrigation shifts the interannual variability of ET and draws it further away from the ALEXI ET estimates. Mixed

results are obtained for RMSD. DA slightly reduces RMSD as compared to OL for many non-irrigated grid cells, whereas in OL_{irr} RMSD is reduced for eastern Nebraska and southern Texas but increased in the central HPA due to overestimation of ET. This overestimation of ET in the central HPA in OL_{irr} can be attributed primarily to overestimates in grid cells with moderate irrigation density and could be related to uncertainty in irrigation maps. For intensively irrigated grid cells—those with irrigation fraction higher than 50%—both OL_{irr} and DA_{irr} show similar large bias reductions relative to OL (Figure 3.10). The impact is greatest in Nebraska, with a 60% reduction in DA_{irr}. These simulations also exhibit a reduction in RMSD, the order of which is slightly smaller than that for the bias. Overall, the combination of irrigation and data assimilation improves simulation of ET for the northern HPA and alleviates the overestimation from irrigation in the southern HPA.

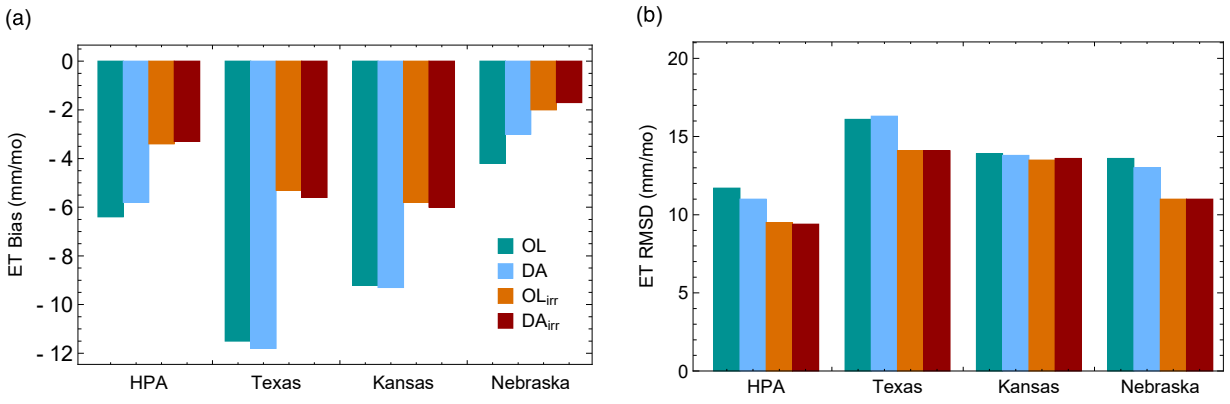

Figure 3. 10 (a) Bias and (b) RMSD of simulated monthly ET against ALEXI for active irrigation grids only.

3.4 Conclusions

To date, GRACE data assimilation systems built on advanced land surface models have not accounted for anthropogenic processes. This is inconsistent with the understanding that GRACE

trends reflect anthropogenic activities in many regions (Rodell et al., 2018; Scanlon et al., 2018). Assimilating GRACE into these models can successfully align the simulation of TWS with that observed by GRACE, but in doing so it can also introduce systematic biases in water storage and water and energy fluxes due to the failure to represent key human related processes. The specific objective of this study was to investigate the ability to enhance the skill of GRACE DA by performing assimilation using a modified version of Noah-MP that includes a groundwater pumping irrigation scheme (Nie et al., 2018).

A set of four simulations are performed to investigate the influence of irrigation-only (OL_{irr}), DA-only (DA) and the integrated (DA_{irr}) configurations on simulation of hydrological processes over the High Plains Aquifer. The results reveal certain limitations in DA-only and irrigation-only simulations. The DA simulation leads to systematic negative soil moisture increments in the southern HPA due to the trend in GRACE TWS observations. This can result in a long-term dry bias during the later GRACE period. The DA simulation also fails to improve simulation of groundwater variability and degrades ET simulation in Texas on account of the erroneous reduction of soil moisture in shallow layers. The OL_{irr} simulation introduces negative trends in TWS and groundwater while wetting the surface soil layers through infiltration of irrigation water. This is qualitatively correct, and it leads to significant improvements in water and energy fluxes for intensively irrigated regions. However, the simulation also tends to overestimate groundwater decline and ET enhancement for areas with moderate irrigation coverage, and the benefit of the simulation's improvements over OL only apply to locations with active irrigation. Assimilating GRACE in the presence of irrigation (simulation DA_{irr}) produces the best overall performance in terms of the water storage trend over the northern HPA while results in performance that is similar to OL_{irr} for water storage components and ET over the southern HPA. In this study, we found limited potential for DA_{irr} to improve simulations beyond the improvements realized by OL_{irr}. This

result, however, is highly context dependent. In regions with greater uncertainty in irrigation maps, larger bias in Noah-MP simulation of irrigation, or a greater ratio of TWS variability in non-irrigated versus irrigated lands, the addition of GRACE-DA to a simulation with irrigation would be expected to provide greater benefit.

We note that the choice of TWS baseline and the reference period used to convert GRACE TWS anomaly values into total TWS can have a major impact on assimilation results. In this study, we added different mean TWS to GRACE anomalies for DA and DA_{irr}. This is important because if we use the mean of TWS generated by OL to convert GRACE anomalies and apply that in DA_{irr}, most of the increments would be positive since the GRACE TWS is derived from a baseline that does not consider irrigation-induced depletion, while the DA_{irr} model prediction does. Selecting baselines derived using consistent modeling options, in this case, with or without irrigation, can potentially avoid such bias and its negative impact on data assimilation. In addition, the choice of the baseline period is also important, as baselining to a period that includes significant trends can introduce systematical biased increments to the data assimilation system. In this study, we selected TWS from 2004 to 2009 as the baseline period, based on prior knowledge that the model uncertainty is likely to be enlarged after the onset of the 2011-2012 drought due to weakness in handling extreme events. The choice of baseline can have a strong influence on data assimilation performance in regions with a significant trend. This risk can be minimized by choosing a baseline based on knowledge of the modeling system and the hydro-climatological conditions of the study area. We also note that the use of time-varying GVF is important when simulating irrigation impacts on water storage during drought. As shown in Nie et al. (2018), failing to account for fields going in and out of production—as can be monitored by satellite-derived GVF—can lead to spurious groundwater trends in the model.

This study is sensitive to the many sources of uncertainty known to affect data assimilation in general. These include the challenge of estimating relative errors between observation and models, the many sensitive parameters included in any advanced land surface model, perturbation settings, and the potential for nonlinear model response to perturbation and assimilation. We examined some of these issues through sensitivity tests when developing the assimilation system, and we have used best estimates of each. Sensitivity tests performed using a range of GRACE error estimates (20 – 100 mm) yielded similar results, as did sensitivity tests that applied 0.1 to 10 times perturbation parameters to states. We conclude that our results are stable to the choice of model error estimate and model perturbation settings for the current DA system. Besides the uncertainties in the data assimilation system, there are several limitations to be addressed in future work. First, surface water irrigation is not accounted for, which is not a major issue in this study because the irrigation water source for the HPA region is mainly from groundwater. But further development and testing on water source accounting as well as the inclusion of the snow and surface water states in the GRACE-DA framework are necessary to extend the area of interest (e.g., regions like California). Second, the groundwater extraction schemes presented here are generally applicable to shallow unconfined aquifers. This assumption is reasonable in this study because most of the withdrawals in HPA are from the unconfined High Plains aquifer system (Ryder, 1996). There are also withdrawals along the southern and eastern margins of the confined Great Plain aquifer, however, the water levels are close to the land surface and are about equal to the water levels of the unconfined High Plains aquifer (Miller & Appel, 1997). Third, the lack of observation datasets for irrigation water use and soil moisture for deep layers limits the ability to thoroughly evaluate and calibrate the system and to make effective conclusions on the disaggregation of TWS components. Fourth, we have only performed assimilation updates in model space. Previous implementations of GRACE-DA in advanced LSMs have sometimes performed the assimilation update in observation space (Forman

et al., 2012; Houborg et al., 2012; B. Li et al., 2012; Zaitchik et al., 2008) and sometimes in model space (Kumar et al., 2016). It will be valuable to compare these approaches in the context of irrigation in the future.

Despite these limitations, the current study demonstrates that GRACE data assimilation can be used to diagnose model weaknesses in process representation, and that it can be effectively improved by adding anthropogenic processes into the model. It also shows the synergy between work to improve model physics and to advance data assimilation techniques. The combination of process-based simulation of groundwater withdrawals and assimilation of GRACE observations led to more meaningful simulation of water storage anomalies, which is critical for monitoring hydrological conditions and for agricultural applications. These improvements are relevant for widely used systems such as NLDAS, the National Climate Assessment – Land Data Assimilation System (NCA-LDAS) and United States Drought Monitor (USDM). Implications for surface water and energy fluxes are also potentially important for simulations of land-atmosphere interactions in coupled weather and climate models.

Acknowledgements

The study was supported in part by the NASA GRACE Science Team award NNX16AF12G. Different data sets used for model evaluation were obtained from various sources described in section 3.2.5. Computational resources were provided by the NASA Center for Climate Simulation (NCCS) at NASA's Goddard Space Flight Center and the Maryland Advanced Research Computing Center (MARCC). We would like to thank the NASA LIS support team for help on model development. The model output and the time-varying GVF dataset are available through Johns Hopkins University Data Archive (https://doi.org/10.7281/T1/4ZCOC4).

Chapter 4

Irrigation water use sensitivity to climate across the Contiguous United States

Abstract

Climate variability and extremes are important drivers of changes in irrigation water use. Efforts to anticipate where and how climate change will affect future water availability can benefit from understanding how irrigation water demand has responded to these drivers to date. Here we modify the Noah-MP land surface model to include a demand-driven irrigation scheme and source water accounting. To account for seasonal to interannual variability in an active irrigated area, we apply a time-varying irrigation fraction map, groundwater irrigation ratio, greenness vegetation fraction derived from multiple remote sensing datasets, and USGS water use reports. We evaluate model performance in terms of irrigation water amount from simulations for the contiguous United States for the period 2002-2017. We focus on drought and explore the relationships between demand-driven irrigation water applications, precipitation and temperature using correlation analysis and multiple linear regression. Key results include: (1) widespread sensitivity of irrigation water use to precipitation, in both model and available observations (which is not surprising); (2) regionally and seasonally variable sensitivity to temperature, suggesting differing potential impacts of warming

on irrigation water demand; (3) interactions between temperature and precipitation sensitivity that differ by location and might reflect impacts of management decisions under climate variability, and these interaction terms might be underestimated by the model; (4) spatial disconnects exist between aquifers that are stressed by pumping and the basins that exhibit strong climate sensitivities, and these disconnections need to be considered in assessing future climate impacts and irrigation-based adaptation. Accounting for these sensitivities could aid assessments of emerging freshwater scarcity and inform efforts to manage the interacting impacts that climate change and irrigation have on U.S. water resources.

4.1 Introduction

Irrigation is vital for agricultural production as it resolves the temporal and spatial disconnects between water supply and water demand, especially for semi-arid and arid regions. It redistributes the water fluxes and surface-energy budgets, thus affecting the interactions between the land and atmosphere, as a result of combined effects from both climate variability and human management. In the United States, irrigation accounts for ~40% of total freshwater withdrawal and more than 80% of total freshwater consumption (Jury & Vaux, 2005). Approximately 60% of irrigated areas are supplied by groundwater resources (Scanlon & Faunt, 2012; Siebert et al., 2010). In recent decades, there has been dramatic groundwater depletion in the southern High Plains and California Central Valley, as observed by both ground-based measurements and Gravity Recovery and Climate Experiment (GRACE) satellite data. In these regions, intensive abstraction at the rates currently practiced cannot be balanced by groundwater recharge.

Irrigation applications, and associated groundwater extractions, are inextricably linked to climate variability. Irrigation demand is affected by precipitation, temperature, and other meteorological variables, such that total pumping for irrigation can vary as a function of climate

conditions in many regions—a phenomenon known as climate-induced pumping (Russo & Lall, 2017). While a scarcity of fresh water resources has already emerged in many parts of the world under current climate conditions, climate change may pose an additional threat to water security should warming and/or a shift in precipitation patterns lead to changes in water supply (Schewe et al., 2014). Climate-induced pumping might exacerbate these water shortages. Additionally, the tendency of an irrigated region to increase water use in response to precipitation shortage or elevated temperature is an indicator of the sensitivity of that region's irrigated agriculture to climate variability or change. Therefore, understanding the impact that climate variability has on irrigation water use is important when assessing the vulnerability of irrigation developments from a perspective of water resource depletion and of changes in water demand. However, spatially and temporally distributed observations of irrigation water extraction and consumption are limited, constraining the assessment of irrigation responses.

One way to overcome this data gap is to simulate irrigation water use through Global Hydrological Models (GHMs) or Land Surface Models (LSMs) implemented with reasonable irrigation parameterizations and assumptions. Modeling efforts have been undertaken to improve the representation of human water use in these models (Nazemi & Wheater, 2015a; Pokhrel et al., 2016; Scanlon et al., 2018; Wada et al., 2017). Nonetheless, limitations exist in current studies, including, in different cases: 1) not accounting for where the withdrawals occur or what impact they have on groundwater or surface water processes; 2) not accounting for interannual changes in irrigation land use or source water partitioning; 3) irrigating based on a climatologically fixed growing season that does not account for management responses to extreme events such as drought; 4) calibrating models using non-physical parameters that may not provide skill in future water use projections.

We have modified an advanced land surface model to include a demand-driven irrigation scheme and source water accounting (Nie et al., 2018). Here, we perform simulations with this modeling system that make use of multiple time-varying data sources (MIrAD-US irrigation fraction, USGS-based groundwater irrigation percentage, and MODIS NDVI-based greenness vegetation fraction) to capture the timing, location, and seasonal to interannual variability of conditions relevant to irrigation water use. Using retrospective model simulations, we quantify the impacts that precipitation and temperature variability have on simulated irrigation quantity. Results are analyzed by region, growing season, and irrigation water source to understand patterns of climate sensitivity, as simulated by the model. Finally, we compare the climate sensitivity of our demand-driven irrigation simulation to climate sensitivity inferred from available data on irrigation water use.

4.2 Data and methods

In previous work, we introduced a demand-driven sprinkler irrigation scheme to the Noah-Multiparameterization Land Surface Model (Noah-MP LSM; Niu et al., 2011), version 3.6, within the framework of NASA's Land Information System (LIS; Kumar et al., 2006). The three key rules to trigger the irrigation in this modeling framework include the irrigation location (where to irrigate), the timing (when to irrigate) and the amount (how much to irrigate). The percent irrigated area is obtained from the MODIS Irrigated Agriculture Dataset (MIrAD-US; Brown & Pervez, 2014) available for 2002, 2007, and 2012 at 250 m spatial resolution, providing the location and fraction of the irrigated area within a model grid cell. Since MIrAD-US is only updated once every five years, we use a single dataset for five years, until the next dataset becomes available; i.e., dataset of 2002 is used for 2002-2006, 2007 is used for 2007-2011, and finally, 2012 for 2012-2017. The time-varying greenness vegetation fraction (GVF), derived from monthly MODIS NDVI composites and MODIS-IGBP land cover data set at 0.05° spatial resolution (Nie et al., 2018), is used to define the

timing of irrigation, which begins and ends when a certain threshold within the long-term range GVF at the grid cell is exceeded. Then the scheme checks if the current root zone soil moisture availability falls below a certain threshold. Once irrigation is triggered, the irrigation water amount is calculated as the volume required to bring the root zone soil moisture deficit up to field capacity, and is applied at every time step between 0600 to 1000 local time. This approach to simulating sprinkler irrigation was first described by Ozdogan et al. (2010).

For our source water partitioning scheme, we assume that irrigation water is sourced from some combination of groundwater and surface water. The ratio of groundwater to total irrigation water is derived from county-level data included in the USGS water use report (https://water.usgs.gov/watuse/data/), which are available at 5-year intervals for 2000-2015. The scheme then subtracts the groundwater irrigation amount from the model's groundwater aquifer storage term. We do not attempt to represent the full complexity of irrigation practices across CONUS in our model. Instead, in order to obtain results that can be cleanly interpreted in terms of climate sensitivity of irrigation demand, we make the following simplifying assumptions: 1) all irrigation is simulated as sprinkler irrigation of row crops; 2) water use efficiency is not considered, as we want to avoid introducing additional uncertainties to the system; 3) crop management and deficit irrigation are not directly accounted for due to limited data sources and coverage. Notwithstanding these simplifications, incorporating time-varying irrigation fractions, groundwater ratios and GVF into our simulation at their greatest available frequency allows the model to capture some management effects, including interannual variability and trends in irrigation water use in response to climate. The simulation is performed over CONUS at 0.125° spatial resolution for the period 2002-2017. Noah-MP is applied in offline (not coupled to the atmosphere) mode, and is

driven by National Land Data Assimilation System-Phase 2 (NLDAS-2; Xia et al., 2012) meteorological fields.

To evaluate the realism of simulated irrigation water use, we compared the model's annual irrigation totals with the USGS reported irrigation amount for the top 10 irrigated states. These estimates are available for 2005, 2010, and 2015. Note that the USGS reported water use for 2005 and 2010 does not distinguish between consumptive and non-consumptive water uses while the report for 2015 includes both total and consumptive water use estimation, which offers a useful point of comparison for our simulated irrigation water use. While our model includes some non-consumptive use in addition to consumptive use—i.e., water returned via recharge to groundwater or as lateral runoff—it does not include conveyance losses or other off-field losses and on-field application inefficiencies that can account for a significant amount of non-consumptive water use in real irrigation settings. The USGS consumptive use estimate in 2015, therefore, provides a useful lower bracket on water use, and we might expect our simulations to fall near that lower bracket.

To explore the relationship between irrigation water use and climate, we examine temporal correlations between in-season meteorological variables—seasonal precipitation (P) and near surface air temperature (T)—and simulated irrigation water use for all actively irrigated grid cells. Rather than using a fixed growing season period or referring to a crop calendar, we define an "effective growing season" for each grid cell. This is done by first locating the peaks in the climatologically-averaged annual cycle of monthly GVF for each grid cell, and then defining the peak month, together with the month before and after it, as an effective growing season. The peaks are detected if they are greater than a defined threshold and can survive Gaussian blurring up to scale 1. Nine major growing season types (GST) are categorized for analysis in which four are defined as the

single season type (spring, summer, autumn, winter) while the remaining five are defined as the dual season type (e.g., "spring and autumn").

We then average simulated irrigation water use, P, and T in actively irrigated grids to sub-basin scale and apply a simple multivariate linear regression model: growing season irrigation water use is the response variable, and P, T and their interaction (P*T) are the three predictor variables. We choose this simple statistical approach because we are attempting to understand general relationships rather than to optimize prediction skill. Six major irrigated river basins are included in the analysis. A GST is included for a river basin if the GST represents more than 10% of the basin area. This criterion led to identification of 14 basin-by-GST combinations. They include three single season GSTs (spring, summer and autumn) and one dual season GST (spring-autumn). The annual time series averaged over each sub-basin are normalized by subtracting the mean and dividing by standard deviation so that the regression coefficients are comparable to each other. Note that for the interaction term, we standardize P and T before multiplying them in order to reduce the structural multicollinearity caused by strong correlation between P and its interaction with T.

Annual calendar year irrigation water use reports are obtained for Kansas and Texas while annual water year irrigation water use is obtained for California in order to evaluate the modelled climate sensitivities. These three states are chosen because they are the only databases that we found that are publicly available. The states also happen to be among the most intensively irrigated and the most agriculturally productive states over CONUS. Statewide averages from 2002 to 2017 for Kansas and Texas are provided by the Kansas Department of Agriculture, Division of Water Resources (per request) and the Texas Water Development Board (http://www.twdb.texas.gov/) based on the data submitted by the irrigation water right owners. Unlike the cases in Kansas and Texas, private groundwater pumping for irrigation is rarely metered in California (Faunt 2009),

therefore, the annual water year estimates from 2002-2015 produced by the California Simulation of Evapotranspiration of Applied Water (Cal-SIMETAW) model are used as a proxy for observations. The model estimates irrigation water demand based on an evapotranspiration demand approach, the aim of which is to improve the accuracy of water use estimates for the California Water Plan (Orang et al., 2013). The correlation and regression analyses are performed for both reported estimates and the corresponding simulated irrigation water use for these three states based on the annual averages without considering the effective growing seasons.

4.3 Results and discussion

4.3.1 Irrigation water use

The largest irrigated areas in CONUS are located in the Lower Mississippi River Basin (LMRB), the High Plains Aquifer crossing the Missouri River Basin (MRB) and the Arkansas-White-Red Basin (AWRB), the California River Basin (CRB), and the Pacific Northwest Basin (PNB) along the Snake River (Figure 4.1 (a)). More than 50% of the irrigation water use is from groundwater for all basins except the CRB and the PNB, which are dominated by surface water irrigation in the early 2000s (Figure 4.1 (b)). Within the short period of 2000-2015, the coverage and locations of irrigation areas and the partitioning of irrigation water sources have changed substantially. As shown in Figure 4.2, irrigation fraction increased about 4% for the LMRB in 2012 compared to 2005, while the percentage contribution of groundwater decreased by 4.5%. In the CRB, in contrast, the irrigation fraction dropped by about 2%, while groundwater contribution to irrigation increased over time by 26%. The dramatic shift towards groundwater for the CRB is likely due to the long-lasting drought that affected the region during this period; the use of groundwater to prevent crop water stress during this drought led to serious groundwater depletion issues and threatened groundwater sustainability (Scanlon & Faunt, 2012). Previous modeling studies (Lawston et al., 2015; Leng et al.,

2014; Pokhrel et al., 2012) often used a static map to represent irrigated area or ignored the source of water that is applied in irrigation, which limited their ability to study long-term irrigation water use variability and its impact on both surface water and groundwater. While the irrigation scheme used in this study is demand-driven, by incorporating time-varying irrigation fraction and groundwater irrigation percentage maps into our model, we are able to indirectly simulate irrigation water use associated with land use change, irrigation expansion, and water management.

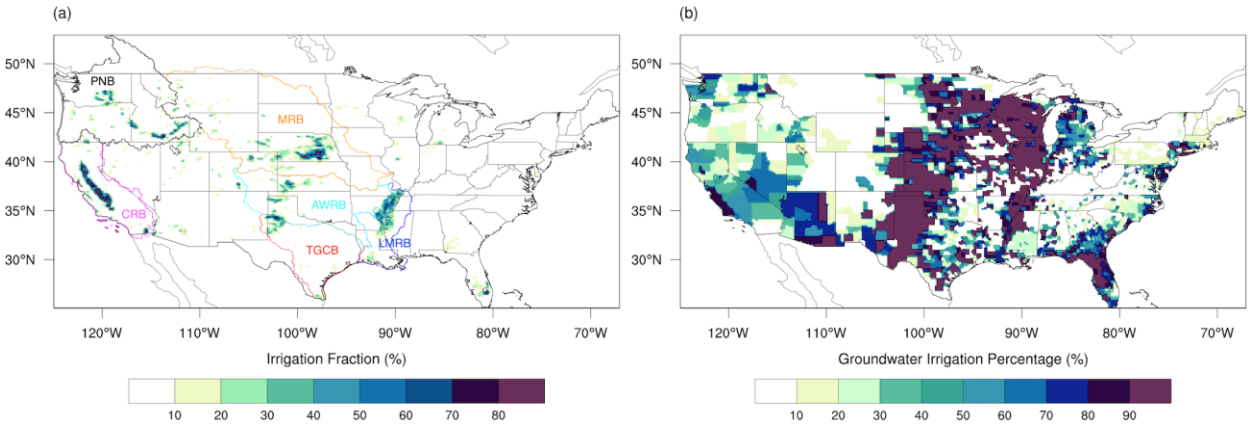

Figure 4. 1 The (a) irrigation fraction from the MODIS Irrigated Agriculture Dataset for the United States (MIrAD-US) for 2002 and (b) percent of irrigation from groundwater derived from the USGS water use report for 2000 over CONUS.

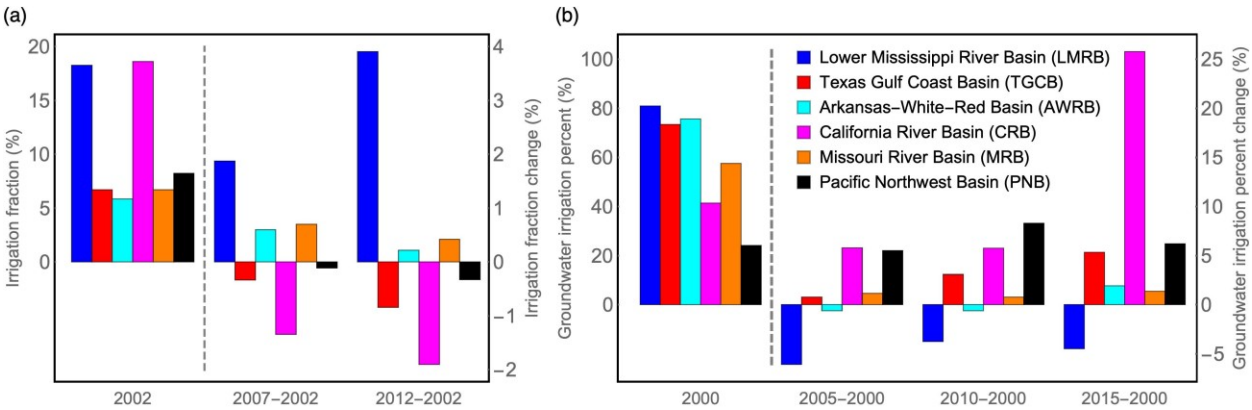

Figure 4. 2 The (a) irrigation fraction and (b) percent of irrigation from groundwater for the reference year (of 2002 (a) and 2000 (b)) and the difference between the reference years and the other available years at a 5-year intervals respectively.

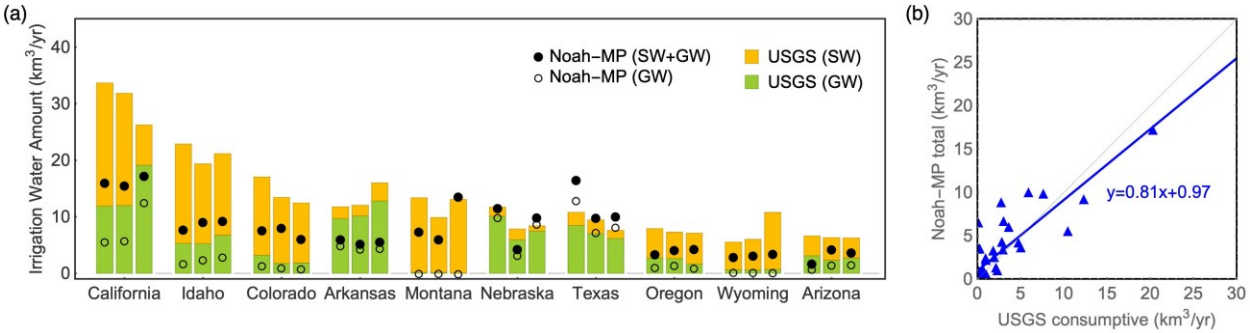

Figure 4. 3 (a) Comparison of the observed and simulated groundwater and surface water irrigation amount for the top 10 irrigated states, selected and ranked based on averaged total irrigation amount for 2005, 2010 and 2015; (b) scatter plot of simulated total against observed consumptive irrigation amount, 2015.

Figure 4.3 compares simulated and USGS reported irrigation water use from groundwater and surface water for the top 10 irrigated states for the years 2005, 2010, and 2015. In general, our simulated annual irrigation water use matches the relative distribution in the USGS report in terms of states and years, though the simulated quantities are mostly lower than the USGS estimates, except for Texas. The underestimation of simulated irrigation water use compared to USGS reports is likely due to inefficient irrigation practices and to off-field losses, which may be comparable to the consumptive water use in real irrigation settings, but are not represented in the model. However, starting with 2015, the USGS report provides consumptive irrigation water use estimates. This offers a better reference for comparison, as the simulated irrigation amounts are mostly consumptive, with only a small proportion returning to water bodies as either surface runoff or groundwater recharge (Figure 4.3 (b)). For example, our simulated total water use for California in 2015 is 17.3 km^3, which is only 66% of the USGS reported total water use, but that matches the USGS reported consumptive water use of 20.3 km^3 quite closely. Overall, the simulated water use regressed on the reported consumptive use has a slope of 0.81, indicating that on average the model underestimates by about 20% relative to consumptive use reports. We consider this to be a reasonable level of agreement considering the uncertainties in the irrigation parameters, method from the model,

measurement errors and limited resolution and precision of USGS estimates. Furthermore, our previous study (Nie et al., 2018) showed that for the groundwater-fed irrigation dominated High Plains Aquifer, including the irrigation scheme improves model agreement with ALEXI ET data, GRACE-derived terrestrial water storage patterns, and depth-to-groundwater measurements in Texas and Kansas, especially under drought conditions. These results suggest that the model, together with its multiple satellite derived observation inputs, is in general able to capture the irrigation response to climate variability, extreme conditions, and their impact on water and energy fluxes.

4.3.2 Influence of irrigation water source

In general, simulated irrigation demand shows significant sensitivity to same-season precipitation in all growing seasons (Figure 4.4). At least 71% of the irrigated grid cells have significant negative correlation with precipitation for single season crops, with the highest (92%) for summer crops, while the correlation with temperature is much less significant, with the significant percent ranging from 3% for winter crops to 50% for summer crops. Most of the grid cells that have significant correlation with temperature are also significantly correlated to precipitation. However, the percent that significant relationships with both P and T is much less for dual season crops than for single season crops, and, more interestingly, the dual season grid cells that are significantly correlated to T are not necessarily significantly correlated to P. The reduced significance effects are mainly attributable to weaker correlations in the later growing season, while we note that the correlation in the earlier season for dual season crops is comparably strong as for single season crops. Overall, simulated climate influences are fairly insensitive to the choice of growing seasons, so the particular choice of the timing and duration of the effective growing season does not greatly affect results (Figure A.1, A.2).

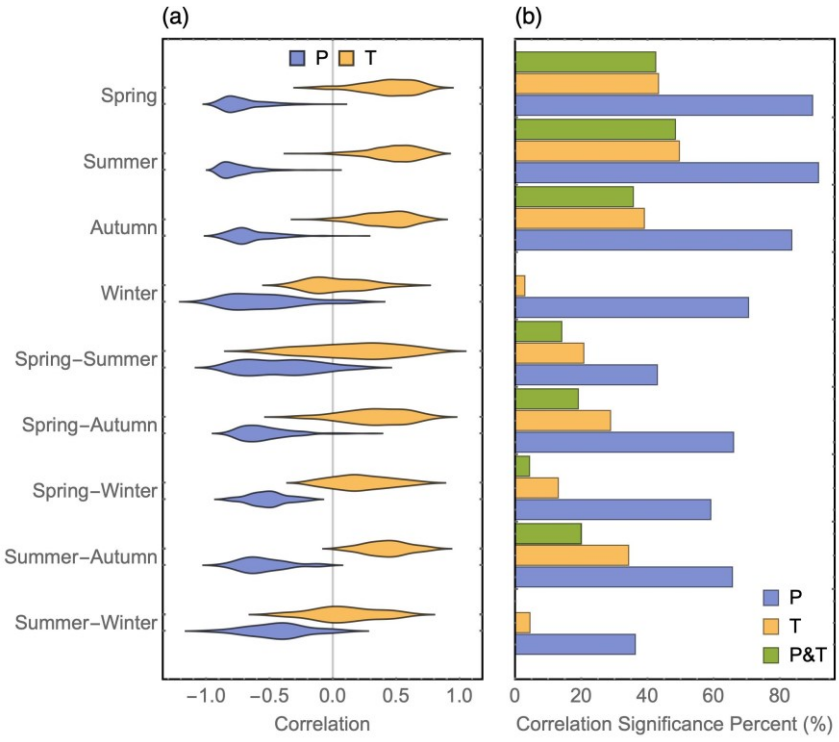

Figure 4. 4 The (a) distribution of correlation coefficients for irrigation water use against precipitation (P) and temperature (T) for active irrigated grid cells and (b) percent of the active irrigated area that has statistically significant correlation values at the 5% level for nine major growing season types. The correlation is calculated based on annual averaged values during the growing seasons.

To investigate connections between irrigation water source and climate-irrigation water use relationships, we further categorize the active irrigated grid cells into three classes for each GST: (i) the surface water dominated area, where the groundwater irrigation percentage is lower than 20%; (ii) the groundwater dominated area, where the groundwater irrigation percentage is higher than 80% and (iii) the mixed area, where the groundwater percentage is in between 20% and 80%. As shown in Figure A.3, there is no clear distinction in the connections between climate and irrigation water source. However, uncertainties exist as the groundwater irrigation ratio gets updated every five years while in reality it may exhibit a strong connection with climate condition, which is not able to be represented in the model because of a lack of data at finer temporal resolution. Nevertheless, previous studies have found that climate variability can have large influence on groundwater storage

change through climate-induced pumping, and this effect may be much stronger and faster than climate's direct effect on groundwater recharge rates (Asoka et al., 2017; Gurdak, 2017; Russo & Lall, 2017). Recognizing the fast response of groundwater to climate through irrigation withdrawal, assessments of irrigation management must consider both irrigation demand sensitivity to climate and the associated aquifer condition as well as groundwater availability.

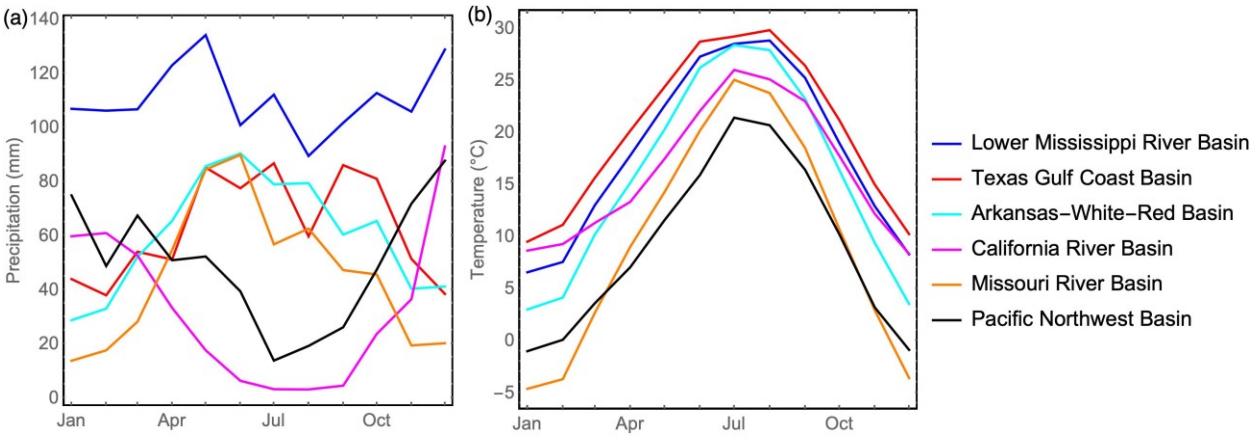

Figure 4. 5 Averaged seasonal cycle of (a) precipitation and (b) temperature in active irrigated area for six major irrigated river basins, 2002-2017.

Table 4. 1 Summary statistics of regression between modeled irrigation water use and climate variables (P, T, and the interactive term) as well as groundwater storage trends for major growing season types for each river basin.

	LMRB	TGCB				AWRB			CRB		MRB	PNB		
	Su	Sp	Au	Su	Sp–Au	Su	Sp	Au	Sp	Su	Su	Su	Sp	Sp–Au
β_T	0.21	0.17	0.19	−0.03	0.26	0.20	0.17	−0.23	0.21	−0.39	0.21**	0.07	0.29**	0.55*
β_P	−0.61*	−0.65**	−0.67*	−0.87*	−0.59*	−0.56*	−0.44	−1.29*	−0.41	−0.50**	−0.75*	−0.83*	−0.71*	−0.71*
$\beta_{P\&T}$	0.18	−0.13	−0.38**	−0.20*	−0.19	−0.13	−0.25	−0.76*	0.21	0.24	−0.06	−0.08	−0.02	0.07
R^2	0.58	0.54	0.61	0.9	0.7	0.77	0.37	0.77	0.4	0.32	0.92	0.71	0.79	0.68
GW_{TR}(cm/yr)	−0.1	−0.5*	−1.3*	−2.2*	−2.1*	−1.9*	−0.4*	−3.9*	−1.1*	−3.8*	−0.6*	−0.7*	−1.0*	−0.3*

4.3.3 Regression analysis

At the basin scale, a linear regression that includes P, T, and a P by T interaction term explains between 32% and 92% of variability in simulated year-to-year irrigation water use (Table

4.1). For some growing seasons and basins, such as summer crops in the TGCB, the AWRB and the MRB, more than 75% of the variance can be captured by the negative relationship between irrigation water use and precipitation. In these situations, the effective growing season coincides with their rainy season (Figure 4.5), so it is intuitive that interannual P variability would have a dominant influence on irrigation applications. For other basins, such as summer crops in the CRB and spring-autumn crops in PNB, however, temperature provides comparable or even greater explanatory power relative to precipitation. The interaction term is only significant for summer crops in the TGCB and autumn crops in the AWRB, both indicating that irrigation water use is more sensitive to the precipitation deficit under higher temperature. The opposite tendency (not statistically significant) is found in the LMRB, the CRB and spring-autumn crops in the PNB, where warmer conditions weaken the link between precipitation and irrigation water use. The CRB stands out among the six river basins, as the regression has particularly poor explanatory power there for both spring crops (40%) and summer crops (32%). The fact that roughly two-thirds of simulated irrigation water use variance is unexplained by P, T and their interaction for the CRB indicates a mixed response of irrigation water use to climate, which is largely attributable to north-to-south spatial variability of climate over the CRB. For summer crops in the Central Valley, where the climate is extremely dry and there is low in-season precipitation, temperature plays a major role in determining irrigation amount. This simulation result is driven primarily by an observed negative relationship between temperature and GVF—higher temperature (drawn from NLDAS2 estimates) is associated with lower GVF (derived from MODIS), and that leads to lower irrigation depths in growing season, resulting in smaller irrigation amount. In the northern CRB, irrigation amount is much smaller, as this region receives greater in-season precipitation. Irrigation water variability in this area, therefore, is dominated by precipitation sensitivity.

Negative simulated groundwater storage trends are found for all basins except for the LMRB, estimated using a Mann-Kendall test to detect monotonic significance at the 5% level and the Theil Sen Slope test for trend magnitude (Table 4.1). Widespread groundwater declines over irrigated areas are mainly attributed to associated groundwater pumping and drought. Extreme groundwater declines are found in the ARWB for autumn crops, the CRB, and the TGCB for summer crops, reaching 3.9, 3.8 and 2.2 cm/yr respectively. The fact that irrigation water use for these basins and seasons is primarily sensitive to precipitation indicate that future warming is less likely to have a direct impact on groundwater sustainability compared to the variability of precipitation. However, localized impact exists such as in the California Central Valley, where there is high temperature sensitivity and alarming groundwater depletion. For basins where the interactions between P and T have relatively larger effects, future warming may influence irrigation water use indirectly, by altering the sensitivity of irrigation demand to precipitation. We note that this analysis does not account for carbon dioxide fertilization, which may also influence the water balance in irrigated fields. It is worth noting that the basins exhibiting greatest climate sensitivity tend not to be in regions that are currently experiencing the greatest water stress. For instance, spring-autumn crops in PNB are highly sensitive to both precipitation and temperature, but under current conditions groundwater depletion in this region is small relative to basins in other parts of the country, in large part because surface water supplies buffer against overexploitation of groundwater. This spatial disconnect between regions of climate sensitive irrigation and regions of current water stress is relevant when assessing sustainability of groundwater resources under climate change, though the lack of a large-scale groundwater depletion signal today does not mean that there is no risk of local, seasonal, or future groundwater pressures.

We note that our simulations are performed at relatively coarse resolution, such that each "irrigated" grid cell contains a mosaic of irrigated and non-irrigated land. Regression results are

sensitive to the choice of which grid cells qualify as "irrigated." As shown in Table A.1 and A.2, the inferred effect of climate variables on irrigation demand shift as the percent irrigated threshold rises. For instance, for summer crops in LMRB, as the threshold is raised from 30% to 50%, variance in irrigation water that can be explained by climate terms increases from 58% to 72%, and the contribution of T and the interaction term increase and become significant while the effect of P becomes marginal. In contrast, for crops in the TGCB and the PNB, sensitivity to precipitation increases for higher irrigation fractions. Greater groundwater declines are found in more intensively irrigated grid cells regardless of the locations and growing season types, implying that simulated groundwater changes are dominated by irrigation.

Overall, simulated irrigation water use is more sensitive to precipitation variability than it is to temperature variability or the interaction between P and T across most of CONUS. Our regressions leave more unexplained variance in basins that include significant internal climate variability, suggesting that major basin delineations do not sufficiently resolve climate sensitivities of irrigation in these regions.

4.3.4 Comparison with observations

Evaluation of the realism of model responses is challenging, on account of both the model's limited ability to account for management decisions and the limited availability of data at a temporal scale fine enough to represent interannual variability in growing season water use. Drawing on publicly available databases, we use statewide calendar year total water use reports for Texas and Kansas, and simulated water year total demand for California as the best available datasets to evaluate our model performance. As shown in Table 4.2, these State reports do not offer a perfect comparison for our model results, as the reports draw on multiple data sources associated with their own uncertainties. They are also quantitatively different from our simulation in that our model is

based on a soil moisture deficit approach that does not explicitly consider irrigation efficiency or human management and adaptation strategies, which may be particularly important under drought conditions. Despite the differences in the magnitude of irrigation water amount, correlations between the simulated and observed irrigation amount are relatively high, indicating that the model can capture observed interannual variability in irrigation amounts. With these caveats in mind, we examine the simulated and reported sensitivity of irrigation applications to P, T and the interaction between P and T in these three states.

Table 4. 2 Summary statistics of statewide annual averaged irrigation water amount, the correlation between the simulated and observed amount and their regression on P, T, and the interaction term for Kansas, Texas, and California.

		Irrigation Water Amount		Multivariate Linear Regression				
		Annual Mean (km^3)	Correlation	β_T	β_P	β_{PT}	R^2	R^2_{adj}
Kansas	Obs.	4.3	0.54*	-0.53*	-0.99*	-0.13	0.73	0.67
	Sim.	6.3		0.13	-0.62*	-0.1	0.56	0.45
Texas	Obs.	10.7	0.52	-0.29	-0.50*	-0.67*	0.55	0.43
	Sim.	15.3		0.16	-0.76*	-0.33*	0.87	0.84
California	Obs.	37.9	0.66*	-0.04	-0.77*	0.53*	0.87	0.83
	Sim.	20.2		-0.07	-0.71*	0.25	0.54	0.39

For Kansas, both P and T have significant impacts on reported irrigation water use, with lower T and P associated with higher irrigation water amount. The interaction term has less effect. However, in the model, the explanatory power of T is much lower and the coefficient even has the opposite sign compared to the regression for the reported estimates. The variability that is explained by the three terms is also substantially lower in the model (56%) than for reported data (73%). However, for reported water use in Texas and California, the effect of T is marginal while the interaction term is significant. Interestingly, the sign of the interaction term differs between Texas and California. In Texas, warmer conditions strengthen the effect of P shortfall on irrigation water use (negative interaction term), while in California warmer conditions reduce the impact of precipitation deficit (positive interaction). The Texas result is more intuitive, as warmer conditions

increase potential evapotranspiration and might be expected to increase the need to irrigate under precipitation deficit. The California result is possibly lead by active management or crop failure in response to dry and hot conditions. As GVF in the intensely irrigated Central Valley is observed to drop under warmer conditions (section 3.3), it is possible that the interaction term captures the fact that crops are less extensive in hot years, resulting in lower sensitivity of total irrigation water to precipitation. This interpretation, however, requires further research. We also note that the interaction term is smaller in the model than in observations for both Texas and California, further suggesting that management decisions not included in our modeling system have some influence on the interaction between P and T.

The net result of these data comparisons is not entirely conclusive. We see that model results and reported estimates are similar in terms of the relative contribution from P and the interaction term. We also see discrepancies between model and reported estimates for the impact of T. These discrepancies could be due to limitations in the model, uncertainties in the reported water use estimates, and noise from off-season data. For example, unlike the analyses performed on the sub-basin scale, that only focused on the effective growing seasons, our comparisons with reported data at state level are performed at annual time scale due to data limitations. The variability of annual P or T may not represent the variability of within-season P or T, especially if the rainy seasons do not overlap the growing seasons, such as in much of California. Moreover, it happens that the three states for which we are able to obtain data are not necessarily "typical" in irrigation norms or climate sensitivity. As shown in Table 4.1, negative simulated irrigation water responses to T are only found for sub-basins within these three states. In all other sub-basins, an increase in T is associated with more irrigation water use. Notably, the decision of whether to perform regression on calendar year or water year can be critical for regions that rely on snow water sources. To demonstrate this, we evaluate sensitivity of the regression results to the choice of calendar year or water year for the top

10 irrigated states (Table A.3). The explained variance for California and Arizona increase by 15% and 46%, respectively, when water year averages are used, because snowpack and snowmelt from the Sierra Nevada is the primary source of water supply in California (California Water Plan Update 2018) and snowmelt-driven flows in the Colorado River serve as the major surface water supply for Arizona (Lahmers & Eden, 2018).

From the perspective of understanding climate sensitivity, then, these comparisons reinforce our understanding that the simulation results must be interpreted as a partially idealized representation of climate influence on irrigation water use. Our model uses a simple demand-driven irrigation algorithm, which is an obvious idealization. The fact that the model is applied using satellite-derived input (vegetation fraction, land cover, irrigated areas) does account for some influence of management decisions on irrigation water use, in that the area actively irrigated and vegetation status are informed by data, but any decisions beyond those basics are not included at all. The comparisons with reported data provide some sense of how realistically the model performs with respect to climate sensitivity—allowing for the caveats associated with the reported irrigation estimates—but simulated irrigation applications and actual irrigation applications are inherently different and must be understood as such. While the reported irrigation water use is a result of mixed impacts of both natural and human factors, the results demonstrate that its connections with temperature either as an independent term or interacting with precipitation are somewhat significant but may be underrepresented in the model. Although the temperature impact might be indirect, making it difficult to represent explicitly in the model, it is worth further exploration because this may become a large source of uncertainty in estimating the water demand under future warming scenarios. Further studies are needed to investigate the extent to which the human decisions and the relationships between precipitation and irrigation water use are sensitive to temperature.

4.4 Conclusions

Conceptually, the link between precipitation and irrigation is straight-forward: when rain supplies more of the water that a crop requires, irrigation applications can be reduced. The link between temperature and irrigation is more nuanced. High growing season temperatures can increase evaporative demand, drying the soil and increasing the need for irrigation applications. But in some regions temperature can impose a constraint on the growing season, potentially altering cropping and management patterns in a manner that can be captured by our satellite-informed, demand-driven irrigation model. Temperature and precipitation also interact, as the sensitivity of irrigation to precipitation might increase at higher temperatures that bring greater evaporative demand, but if high temperatures lead to reduced crop area then the sign of the interaction can flip. Under real-world conditions, these climate sensitivities are further complicated by management decisions and off-field factors influencing irrigation water availability. Even in a relatively simple demand-driven model, however, the influence that climate variability has on irrigation is not always obvious, and the results of modeling exercises can yield insight on regional and seasonal differences in expected climate impacts on irrigation, and hence on water resources.

This study was designed to examine how climate sensitivity of irrigation demand is simulated by a state-of-the-art advanced LSM when equipped with demand-driven irrigation routines and source water accounting. As the model is a simplified and idealized representation of irrigation decisions, the results are primarily intended to demonstrate model response to climate forcing, rather than to provide precise estimates of distributed water use in the CONUS. At the same time, the use of a modeling system that ingests satellite-derived observations of irrigated area and vegetation status means that the results do include some information on cropping decisions, which are in turn under potential influence of climate. Key results of this study include: (1) widespread

sensitivity of irrigation water use to precipitation, in both model and available observations (which is not surprising); (2) regionally and seasonally variable sensitivity to temperature, suggesting differing potential impacts of warming on irrigation water demand; (3) interactions between temperature and precipitation sensitivity that differ by location and might reflect impacts of management decisions and/or crop failure under climate variability, and these interaction terms might be underestimated by the model; (4) spatial disconnects exist between aquifers that are stressed by pumping and the basins that exhibit strong climate sensitivities, and these disconnections need to be considered in assessing future climate impacts and irrigation-based adaptation. For each of these results, management can play a critical role in climate response and resilience strategies.

Future work, for this reason, will focus on addressing the limitations of our model's representation of management. This includes the need to obtain higher frequency maps of irrigation fraction and groundwater irrigation percentage, insomuch as these variables respond to interannual climate variability, the need to consider time-varying source water allocations rather than static estimates that do not consider seasonal or interannual changes in availability, the need to increase spatial resolution to avoid confusion between irrigated and non-irrigated land use types, the need to refine the representation of crops in the model, and the need to consider climate regionalization in categorizing basins. While we believe that simplified numerical experiments like those presented here are important for model development and interpretation, and that they can reveal broad climate sensitivities relevant to water resources, model-informed climate adaptation strategies require refinement of the modeling tools used to simulate coupled natural-human water resource systems.

Acknowledgements

This research is supported in part by the NASA GRACE Science Team award NNX16AF12G. Computational resources were provided by the NASA Center for Climate Simulation (NCCS) at

NASA's Goddard Space Flight Center and the Maryland Advanced Research Computing Center (MARCC). Different data sets used for comparison were obtained from various sources described in section 4.2. The model output will be available through Johns Hopkins University Data Archive (https://archive.data.jhu.edu) once published. We thank Long Zhao for helpful guidance on the statistical techniques.

Chapter 5

Concluding remarks

This thesis has focused on the use of advanced modeling and remote sensing techniques to understand and monitor dynamics of the human-water interface. Specifically, the work has focused on irrigation water use in the Contiguous United States, and done so through the development of a modified version of the Noah-MP Land Surface Model, use of multisource ground-based as well as satellite observations for vegetation, water and energy states and fluxes to improve the model representation of irrigation processes, and application of the GRACE data assimilation technique to reduce system uncertainties and to improve simulation of distribution of water components in the presence of irrigation. Through these efforts, we investigate the links between irrigation water use and climate variability from both modeling and observation perspectives in order to characterize climate sensitivity of water demand across different basins. This approach can provide new insights on the climate vulnerability of irrigation developments and on priorities for improving irrigation water management under climate change.

5.1 Modeling development contributions

In Chapter 2, a demand-driven sprinkler irrigation scheme is implemented into Noah-MP. The presence of a simple groundwater module in Noah-MP allows us to extract the estimated irrigation water amount from a shallow groundwater aquifer, which is then applied onto grid cells that include irrigation as an identified land use. The model is able to simulate groundwater depletion and increased root zone soil moisture via pumping irrigation. However, it tends to overestimate irrigation amount under drought conditions due to unrealistic representation of greenness vegetation fraction, which is a key parameter in simulating the irrigation growing season and irrigation amount. For this reason, we develop monthly time-varying GVF maps derived from MODIS NDVI and MODIS-IGBP land cover datasets. This dataset is input to Noah-MP, such that drought impact can be implicitly represented by interannual variability of GVF and irrigation water amount can be estimated in a realistic manner. We identify prescribed time-varying GVF to be one of the most important parameters for model realism under drought.

In Chapter 3, the GRACE data assimilation framework is implemented into our modified Noah-MP. Extensive sensitivity studies are performed to provide the best set of perturbation parameters for meteorological forcing and land surface water states, to decide the order of triggering irrigation process in a DA system, and to determine reasonable choices of baseline period, which is used to convert GRACE TWS anomalies to total TWS. To our knowledge, this is the first time that GRACE data assimilation is applied to an LSM in the presence of irrigation. Establishing such a modeling and data assimilation system within the framework of LIS will allow researchers take advantage of various GRACE products and to apply the systems to regions that experience intensive irrigation with vulnerable fresh water resources while are in general not able to well simulated by LSMs in their original forms.

One limitation of the above-mentioned developments is that irrigation water was assumed to be supplied exclusively from groundwater. This limits the application of modeling studies to regions where surface water irrigation is not a significant factor. To overcome this limitation and to extend the model's ability to broader scales, in Chapter 4, we develop a gridded dataset that estimates groundwater irrigation percentage at a five year interval, with 0.125° spatial resolution, based on county level USGS water use reports. This allows the model to partition the estimated irrigation amount into groundwater and surface water and to reflect changes in water source division with time. We also update the irrigation fraction dataset every five years by incorporating the MIrAD-US datasets to reflect changes in irrigation equipped areas. The utilization of time-varying irrigation fractions, groundwater irrigation ratios and previously developed time-varying GVF enables the model to capture both natural and human management effects, including seasonal to interannual variabilities and trends in irrigation water use in response to climate.

5.2 Main findings

The first study investigates the groundwater and terrestrial water storage declines in the US High Plains Aquifer, which is intensively irrigated in recent decades and dominated by groundwater-fed irrigation. We modify the Noah-MP LSM to include a groundwater irrigation scheme. In-situ based measurements and satellite retrievals are used to evaluate the model performance. Simulations indicate that including the groundwater irrigation scheme improves modeled evapotranspiration, TWS and groundwater level change in the southern HPA, while the results for the northern HPA are mixed in that an overestimation of TWS decline is found in the model, likely due to its weaknesses in representing subsurface hydrology, problems with soil parameterizations and/or errors in meteorological forcing datasets. We also find that accounting for time-varying GVF instead of climatological GVF is important for model realism under drought.

In the second study, TWS observations from the GRACE mission are assimilated into Noah-MP LSM, and the individual and combined effects of groundwater pumping irrigation and GRACE data assimilation are assessed for the HPA. Results indicate that although GRACE-DA only simulation helps to correct the trends and variations in TWS, it may erroneously distribute the water storage components such as surface soil moisture and affect the related fluxes such as ET through biased feedbacks. However, irrigation-only simulation may over- or under-estimate groundwater decline due to uncertainties in the model parameterizations. Assimilating GRACE when irrigation is simulated produces the best overall performance in water and energy fluxes in that irrigation acts to distribute the water fluxes through proper processes, while GRACE-DA can compensate the model weakness at the cost of sacrificing the water balance.

The third study extends the application to CONUS by further implementing the surface water and groundwater partitioning ratio into the model. The modeled relationship between irrigation water use and climate – precipitation, temperature, and their interaction – is explored and the modeled sensitivities of irrigation water use to climate are compared to the observed ones for three states with publicly available datasets. Results show that, generally, the link between precipitation and irrigation is straight-forward: when rain supplies more of the water that a crop requires, irrigation applications can be reduced. The link between temperature and irrigation is more complex, as is the case for sensitivities to the interaction between temperature and precipitation. There are also differences in climate sensitivities between the model and observations. Comparisons with observed sensitivities for the three states imply that the role of temperature either as an independent term or as a part of the interaction term may be underrepresented in the model.

5.3 Future work

A primary motivation for building the human-water interface into Noah-MP is its capability to be coupled with regional climate models. The land surface is an active, dynamic component of the climate system rather than simply a passive respondent (Koster et al., 2017). Multiple feedbacks exist, such as soil moisture – temperature connections, snow-precipitation-streamflow connections, and dynamic coupling between vegetation phenology, precipitation anomalies and evapotranspiration. Improved initialization of soil moisture anomalies, snow states, and vegetation states through either improved physical structures in the model or data assimilation have proved to have significant impact on hydrological forecast skill (Koster & Walker, 2015; Koster et al., 2010; Lin et al., 2016).

Irrigation increases root zone soil moisture, which further leads to enhanced evapotranspiration, and that can further influence clouds, precipitation and temperature through land-planetary boundary layer coupling processes. Groundwater extraction for irrigation enlarges the gradient between groundwater and deep soil moisture levels, and affects the recharge and discharge rates. The studies performed for this dissertation have provided insights on how irrigation may impact hydrological and energy fluxes, how it may respond to climate variability and extremes and what the key parameters are in controlling model performance. The importance of these improvements in human-water interactions under the context of land-atmosphere interaction can be further studied and evaluated.

Despite the inclusion of time-varying datasets for GVF, irrigation fractions, and groundwater irrigation ratios to reflect the impact of both natural and human factors on irrigation water use, further developments are needed in the irrigation parameterization in order to better simulate the hydroclimatic conditions at a broader range of scales:

1) In our study, all irrigation is simulated as sprinkler irrigation for row crops. This is a reasonable assumption for the HPA, but may introduce uncertainties once applied to CONUS. Different irrigation methods and crops can significantly impact the irrigation water demand due to different water needs, irrigation water efficiency, and irrigation practices. Flood and drip irrigation options are available in LIS but cannot be used simultaneously in a single simulation. Maps of crop types and irrigation methods, or even the percentage of different crops and irrigation methods within each grid cell need to be incorporated to reflect their impact on irrigation water use.
2) A shallow unconfined groundwater reservoir may not be adequate for places with complex groundwater systems such as those including alluvial and confined aquifers. The low permeability in alluvial aquifers may prevent return flow from reaching the main aquifer system, while the extraction from confined aquifers may not affect the flux exchange between deep soil and the confined aquifer. Moreover, lateral flow is not considered in our model, which may lead to biased estimation in groundwater variation if the model is applied in regions that experience significant amount of lateral recharge or is applied at a finer spatial scale. Better groundwater parameterization is needed, such as coupling to MODFLOW, as is more commonly considered in global hydrological models.

Finally, the motive for the irrigation sensitivity study (Chapter 4) is to better understand how climate variability impacts irrigation water demand, how these impacts vary among basins and growing seasons and how the sensitivities are distributed in models compared to observations. This is essential in the context of projecting irrigation water use under global warming. While many global hydrological models and global crop models have simplified objective functions to assess irrigation water use under climate change, these models typically have limited factors that control the

estimation of water use, such as precipitation deficit and warming0associated CO2 fertilization effects. We find that the relationship between irrigation and temperature variability is more complicated, even for our simple system of sprinkler-irrigated row crops. High temperature may result in increased irrigation demand due to increased evaporative demand but also may decrease irrigation demand by altering cropping and management patterns. Its interaction with precipitation can also vary in magnitude and sign, depending on the region and season of irrigation. The different modeled sensitivities among basins and the discrepancies between the modeled and observation-based relationships suggest that large uncertainties may exist in irrigation water use projection, particularly for simplified and spatially uniform parameterizations. There is a need for deeper study of the links between irrigation and climate under both natural variability and variability in management, in order to identify key processes that are not well represented in current Earth system models. Incorporating knowledge from such studies into advanced models will enhance the value of models for climate impact studies.

Appendix A

Supporting information for Chapter 4

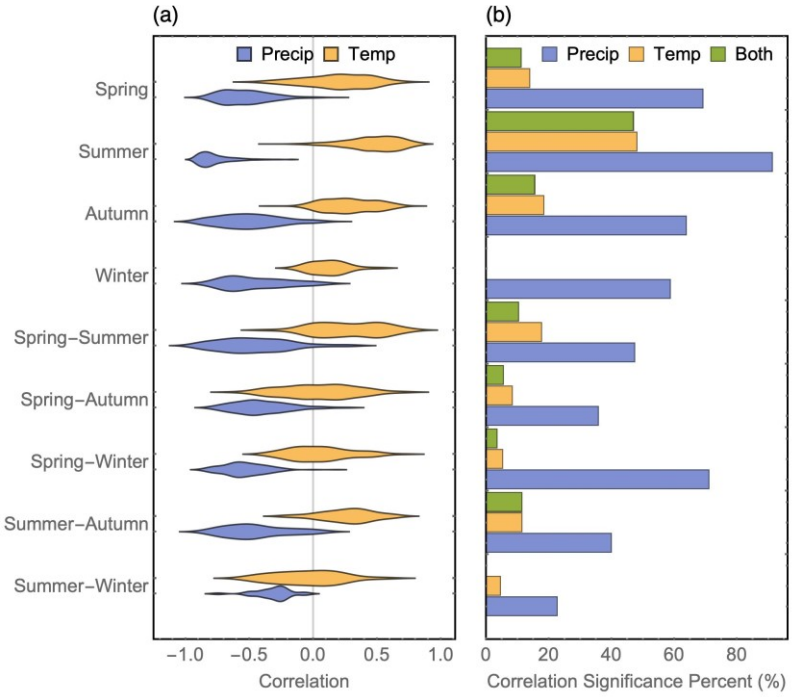

Figure A. 1 The distribution of correlation coefficients for irrigation water use against precipitation (P) and temperature (T) for active irrigated grid cells with "effective growing season" defined as the calendar seasons that include the peaks of climatology GVF.

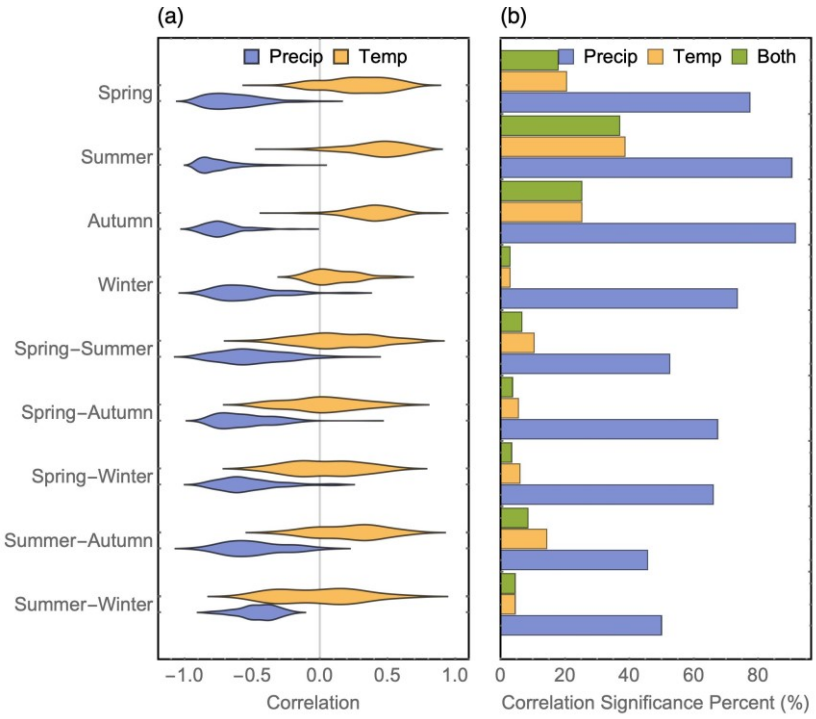

Figure A. 2 The distribution of correlation coefficients for irrigation water use against precipitation (P) and temperature (T) for active irrigated grid cells with "effective growing season" defined as contiguous five months including the month containing the peaks of climatology GVF as well as the two before and after.

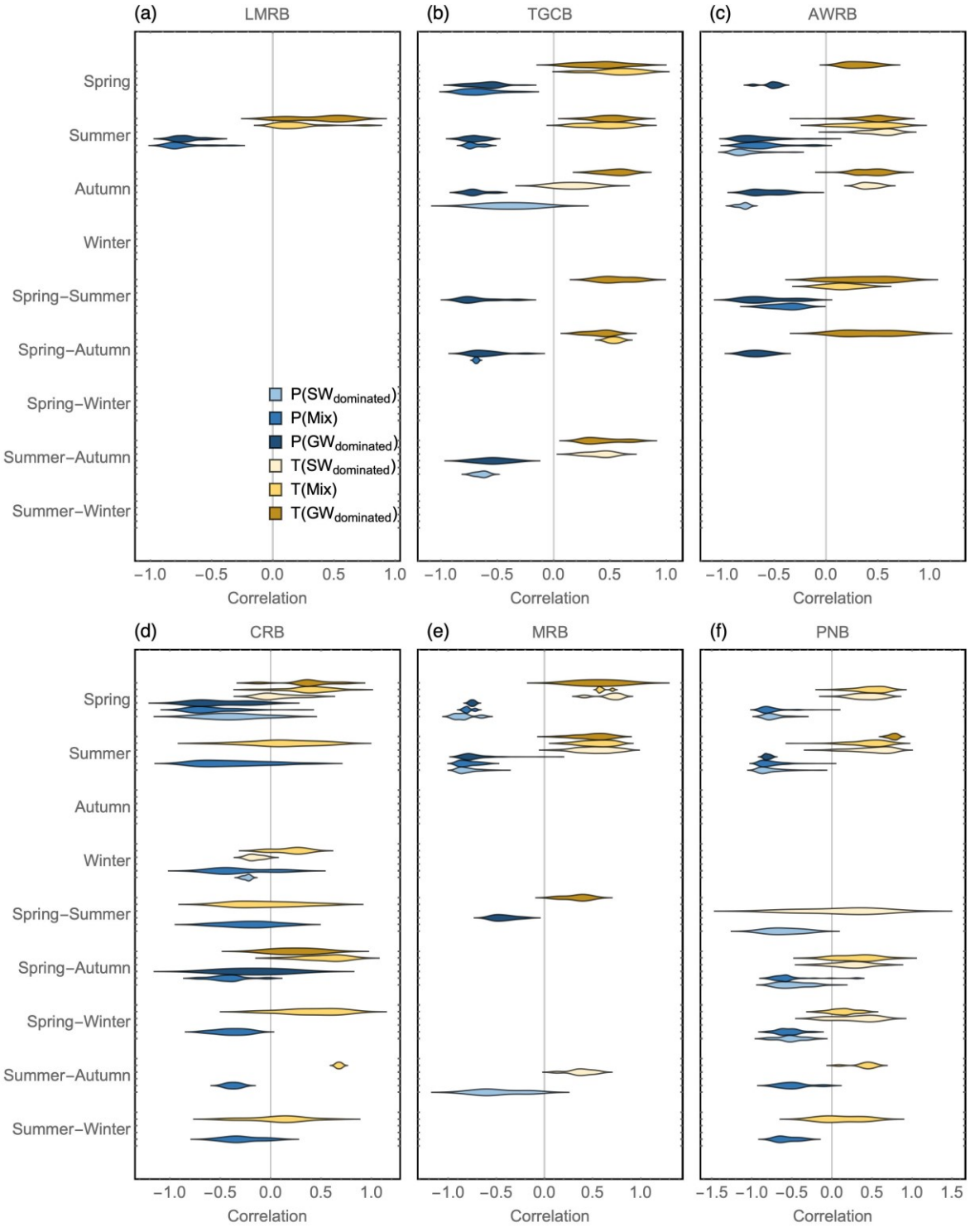

Figure A. 3 The distribution of correlation coefficients for irrigation water use against precipitation (P) and temperature (T) for each river basin categorized by the surface water dominated ($SW_{dominated}$), the groundwater dominated ($GW_{dominated}$), and the mixed (Mix) irrigation.

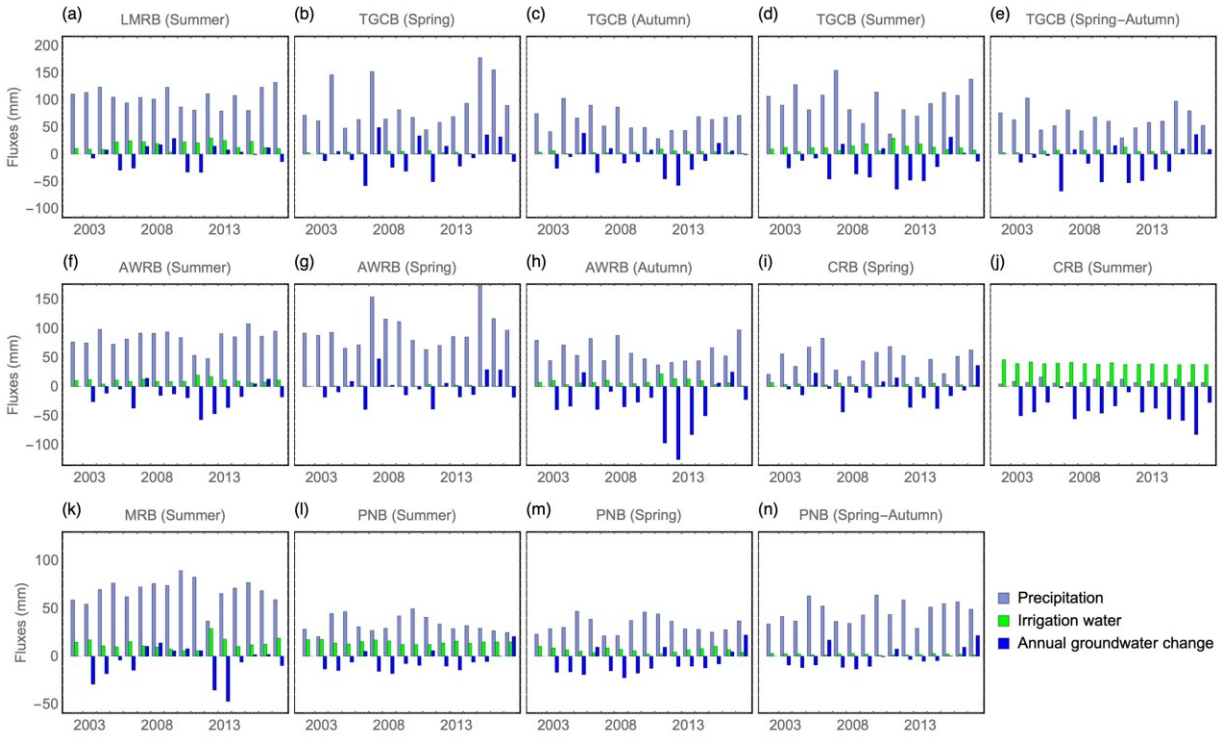

Figure A. 4 Interannual variability of precipitation, irrigation water use during the growing seasons and the year-by-year annual groundwater change for major growing season types for each river basin.

Table A. 1 Summary statistics of regression between irrigation water use and climate variables (P, T, and the interactive term) for major growing season types for each river basin with temporally averaged irrigation fraction over 30% for each grid.

	LMRB	TGCB				AWRB		CRB		MRB	PNB		
	Su	Sp	Au	Su	Sp–Au	Su	Au	Sp	Su	Su	Su	Sp	Sp–Au
β_T	0.51*	0.05	0.2	−0.00	0.12	−0.04	0.21	0.24	−0.53	0.06	0.07	0.24	0.29**
β_P	−0.62*	−0.82*	−0.78*	−0.77*	−0.65*	−0.90*	−0.58*	−0.36	−0.44	−0.48*	−0.88*	−0.67*	−0.59*
$\beta_{P\&T}$	0.57**	−0.18	−0.35**	−0.24	−0.31	−0.05	0.00	0.21	0.10	−0.37*	−0.16	0.07	0.66*
R^2	0.68	0.53	0.67	0.7	0.69	0.78	0.5	0.34	0.28	0.76	0.79	0.68	0.75
GW_{TR}(cm/yr)	−0.4	0.01	−2.5*	−11.*	−10.*	−10.*	−0.4	−5.3*	−10.*	−1.7*	−6.0*	−4.2*	−2.3*

Table A. 2 Summary statistics of regression between irrigation water use and climate variables (P, T, and the interactive term) for major growing season types for each river basin with temporally averaged irrigation fraction over 50% for each grid.

	LMRB	TGCB	AWRB	CRB		MRB	PNB	
	Su	Su	Su	Sp	Su	Su	Su	Sp
β_T	0.84*	−0.02	0.28	0.23	−0.58**	0.04	−0.02	0.23
β_P	−0.24	−0.81*	−0.44*	−0.37	−0.43	−0.56*	−1.04*	−0.81*
$\beta_{P\&T}$	0.69*	−0.27	−0.48**	0.22	0.07	−0.40*	−0.36	−0.04
R^2	0.72	0.7	0.76	0.39	0.27	0.75	0.84	0.78
GW_{TR}(cm/yr)	−0.3	−16.*	−12.*	−8.1*	−12.*	−2.3*	−7.7*	0.14*

References

Alcamo, J., Döll, P., Henrichs, T., & Kaspar, F. (2003). Development and testing of the WaterGAP 2 global model of water use and availability. *Hydrological Sciences Journal*, *48*(3), 317–337. https://doi.org/10.1623/hysj.48.3.317.45290

Anderson, Lo, M. H., Swenson, S., Famiglietti, J. S., Tang, Q., Skaggs, T. H., et al. (2015). Using satellite-based estimates of evapotranspiration and groundwater changes to determine anthropogenic water fluxes in land surface models. *Geoscientific Model Development*, *8*(10), 3021–3031. https://doi.org/10.5194/gmd-8-3021-2015

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A., & Kustas, W. P. (2007). A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology. *Journal of Geophysical Research: Atmospheres*, *112*(D11), 1100. https://doi.org/10.1029/2006JD007507

Andrew, R. L., Guan, H., & Batelaan, O. (2017). Large-scale vegetation responses to terrestrial moisture storage changes. *Hydrology and Earth System Sciences*, *21*(9), 4469–4478. https://doi.org/10.5194/hess-21-4469-2017

Arsenault, K. R., Kumar, S. V., Geiger, J. V., Wang, S., Kemp, E., Mocko, D. M., et al. (2018). The Land surface Data Toolkit (LDT v7.2) – a data fusion environment for land data assimilation systems. *Geoscientific Model Development*, *11*(9), 3605–3621. https://doi.org/10.5194/gmd-11-3605-2018

Arsenault, K. R., Nearing, G. S., Wang, S., Yatheendradas, S., & Peters-Lidard, C. D. (2018). Parameter Sensitivity of the Noah-MP Land Surface Model with Dynamic Vegetation. *Journal of Hydrometeorology*, *19*(5), 815–830. https://doi.org/10.1175/jhm-d-17-0205.1

Asoka, A., Gleeson, T., Wada, Y., & Mishra, V. (2017). Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India. *Nature Geoscience*, *10*(2), 109–117. https://doi.org/10.1038/ngeo2869

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., et al. (2007). Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophysical Research Letters*, *34*(20). https://doi.org/10.1029/2007gl031088

Bernknopf, R., Brookshire, D., Kuwayama, Y., Macauley, M., Rodell, M., Thompson, A., et al. (2018). The Value of Remotely Sensed Information: The Case of a GRACE-Enhanced Drought Severity Index. *Weather, Climate, and Society*, *10*(1), 187–203. https://doi.org/10.1175/wcas-d-16-0044.1

Bettadpur, S. (2007). CSR Level-2 Processing Standards Document for Product Release 04 GRACE 327-742. Center for Space Research.

Brookfield, A. E., Hill, M. C., Rodell, M., Loomis, B. D., Stotler, R. L., Porter, M. E., & Bohling, G. C. (2018). In-Situ and GRACE-based Groundwater Observations: Similarities, Discrepancies, and Evaluation in the High Plains Aquifer in Kansas. *Water Resources Research*, 1–33. https://doi.org/10.1029/2018WR023836

Brown, J. F., & Pervez, M. S. (2014). Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. *Agricultural Systems*, *127*, 28–40.

Cai, X., Yang, Z.-L., David, C. H., Niu, G.-Y., & Rodell, M. (2014). Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin. *Journal of Geophysical Research: Atmospheres*, *119*(1), 23–38. https://doi.org/10.1002/2013JD020792

Cai, X., Yang, Z.-L., Xia, Y., Huang, M., Wei, H., Leung, L. R., & Ek, M. B. (2015). Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed, 1–20. https://doi.org/10.1002/(ISSN)2169-8996

Case, J. L., LaFontaine, F. J., Bell, J. R., Jedlovec, G. J., Kumar, S. V., & Peters-Lidard, C. D. (2013). A Real-Time MODIS Vegetation Product for Land Surface and Numerical Weather Prediction Models. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(3), 1772–1786. https://doi.org/10.1109/TGRS.2013.2255059

Case, J. L., LaFontaine, F. J., Kumar, S. V., & Peters-Lidard, C. D. (2012). P69 Using the NASA-Unified WRF to Assess the Impacts of Real-Time Vegetation on Simulations of Severe Weather.

Changming, L., Jingjie, Y., & Kendy, E. (2001). Groundwater Exploitation and Its Impact on the Environment in the North China Plain. *Water International*, *26*(2), 265–272. https://doi.org/10.1080/02508060108686913

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H. L., Koren, V., et al. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research*, *101*(D3), 7251–7268.

De Rosnay, P., Polcher, J., Laval, K., & Sabre, M. (2003). Integrated parameterization of irrigation in the land surface model ORCHIDEE. Validation over Indian Peninsula. *Geophysical Research Letters*, *30*(19).

DeAngelis, A., Dominguez, F., Fan, Y., Robock, A., Kustu, M. D., & Robinson, D. (2010). Evidence of enhanced precipitation due to irrigation over the Great Plains of the United States. *Journal of Geophysical Research*, *115*(D15), D15115–14. https://doi.org/10.1029/2010JD013892

Dorigo, W., Scipal, K., Parinussa, R. M., Liu, Y. Y., & Wagner, W. (2010). Error characterisation of global active and passive microwave soil moisture data sets.

Döll, P., & Siebert, S. (2002). Global modeling of irrigation water requirements. *Water Resources Research*, *38*(4). https://doi.org/10.1029/2001WR000355

Döll, P., Fritsche, M., Eicker, A., & Müller Schmied, H. (2014). Seasonal Water Storage Variations as Impacted by Water Abstractions: Comparing the Output of a Global Hydrological Model with GRACE and GPS Observations. *Surveys in Geophysics*, *35*(6), 1311–1331. https://doi.org/10.1007/s10712-014-9282-2

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., et al. (2012). Impact of water withdrawals from groundwater and surface water on continental water storage variations. *Journal of Geodynamics*, *59-60*, 143–156. https://doi.org/10.1016/j.jog.2011.05.001

Döll, P., Kaspar, F., & Lehner, B. (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology*, *270*(1-2), 105–134. https://doi.org/10.1016/s0022-1694(02)00283-4

Eicker, A., Schumacher, M., Kusche, J., Döll, P., & Schmied, H. M. (2014). 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. https://doi.org/10.1007/s10712-014-9309-8
Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., & Grunmann, P. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*.
Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918
Evans, J. P., & Zaitchik, B. F. (2008). Modeling the large-scale water balance impact of different irrigation systems. *Water Resources Research*, *44*(8), n/a–n/a. https://doi.org/10.1029/2007WR006671
Famiglietti, J. S. (2014). The global groundwater crisis. *Nature Publishing Group*, *4*(11), 945.
Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., et al. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, *38*(3), n/a–n/a. https://doi.org/10.1029/2010GL046442
Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, *45*(2), 1485. https://doi.org/10.1029/2005RG000183
Felfelani, F., Wada, Y., Longuevergne, L., & Pokhrel, Y. N. (2017). Natural and human-induced terrestrial water storage change: A global analysis using hydrological models and GRACE. *Journal of Hydrology*, 1–43. https://doi.org/10.1016/j.jhydrol.2017.07.048
Fendorf, S., Michael, H. A., & van Geen, A. (2010). Spatial and Temporal Variations of Groundwater Arsenic in South and Southeast Asia. *Science*, *328*(5982), 1123–1127. https://doi.org/10.1126/science.1172974
Feng, W., Shum, C., Zhong, M., & Pan, Y. (2018). Groundwater Storage Changes in China from Satellite Gravity: An Overview. *Remote Sensing*, *10*(5), 674–25. https://doi.org/10.3390/rs10050674
Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, *1*, 3–32.
Forman, B. A., & Reichle, R. H. (2013). The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system. *Water Resources Research*, *49*(11), 7457–7468. https://doi.org/10.1002/2012WR012885
Forman, B. A., Reichle, R. H., & Rodell, M. (2012). Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resources Research*, *48*(1), 872–14. https://doi.org/10.1029/2011WR011239
Frenken, K., & Gillet, V. (2012). Irrigation water requirement and water withdrawal by country. FAO.
Friedl, M. A., Sulla-Menashe, D., Bin Tan, Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168–182. https://doi.org/10.1016/j.rse.2009.08.016
Getirana, A., Kumar, S., Girotto, M., & Rodell, M. (2017). Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability. *Geophysical Research Letters*. https://doi.org/10.1002/2017gl074684
Girotto, M., De Lannoy, G. J. M., Reichle, R. H., & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, *52*(5), 4164–4183. https://doi.org/10.1002/2015wr018417

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., Rodell, M., Draper, C., Bhanja, S. N., & Mukherjee, A. (2017). Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in India. *Geophysical Research Letters*, *44*(9), 4107–4115. https://doi.org/10.1002/2017gl072994

Gurdak, J. J. (2017). Climate-induced pumping. *Nature Geoscience*, *10*(2), 71–71. https://doi.org/10.1038/ngeo2885

Gurdak, J. J., & Roe, C. D. (2010). Review: Recharge rates and chemistry beneath playas of the High Plains aquifer, USA. *Hydrogeology Journal*, *18*(8), 1747–1772. https://doi.org/10.1007/s10040-010-0672-3

Gutentag, E. D., Heimes, F. J., Krothe, N. C., & Luckey, R. R. (1984). Geohydrology of the High Plains Aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas and Wyoming.

Gutentag, E. D., Heimes, F. J., Krothe, N. C., Luckey, R. R., & Weeks, J. B. (1984). Geohydrology of the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming.

Gutman, G., & Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *International Journal of Remote Sensing*, *19*(8), 1533–1543. https://doi.org/10.1080/014311698215333

Haddeland, I., Lettenmaier, D. P., & Skaugen, T. (2006). Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins. *Journal of Hydrology*, *324*(1-4), 210–223. https://doi.org/10.1016/j.jhydrol.2005.09.028

Hain, C. R., & Anderson, M. C. (2017). Estimating morning change in land surface temperature from MODIS day/night observations: Applications for surface energy balance modeling. *Geophysical Research Letters*, *44*(19), 9723–9733. https://doi.org/10.1002/2017gl074952

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., & Tanaka, K. (2008a). An integrated model for the assessment of global water resources – Part 1: Model description and input meteorological forcing. *Hydrology and Earth System Sciences*, *12*(4), 1007–1025. https://doi.org/10.5194/hess-12-1007-2008

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., & Tanaka, K. (2008b). An integrated model for the assessment of global water resources–Part 1: Model description and input meteorological forcing. *Hydrology and Earth System Sciences*, *12*(4), 1007–1025.

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*(7), n/a–n/a. https://doi.org/10.1029/2011WR011291

Jacquette, E., Bitar, Al, A., Mialon, A., Kerr, Y., Quesney, A., Cabot, F., & Richaume, P. (2010). SMOS CATDS level 3 global products over land, *7824*, 78240K.

James, K. A., Stensrud, D. J., Yussouf, N., James, K. A., Stensrud, D. J., & Yussouf, N. (2009). Value of Real-Time Vegetation Fraction to Forecasts of Severe Convection in High-Resolution Models. *Dx.Doi.org*, *24*(1), 187–210. https://doi.org/10.1175/2008WAF2007097.1

Jiang, L., Kogan, F. N., Guo, W., Tarpley, J. D., Mitchell, K. E., Ek, M. B., et al. (2010). Real-time weekly global green vegetation fraction derived from advanced very high resolution radiometer-based NOAA operational global vegetation index (GVI) system. *Journal of Geophysical Research: Atmospheres*, *115*(D11), 650. https://doi.org/10.1029/2009JD013204

Joodaki, G., Wahr, J., & Swenson, S. (2014). Estimating the human contribution to groundwater depletion in the Middle East, from GRACE data, land surface models, and well observations. *Water Resources Research*, *50*(3), 2679–2692. https://doi.org/10.1002/2013wr014633

Jury, W. A., & Vaux, H. (2005). The role of science in solving the world's emerging water problems. *Proceedings of the National Academy of Sciences*, *102*(44), 15715–15720. https://doi.org/10.1073/pnas.0506467102

Katz, B. S., Stotler, R. L., Hirmas, D., Ludvigson, G., Smith, J. J., & Whittemore, D. O. (2016). Geochemical Recharge Estimation and the Effects of a Declining Water Table. *Vadose Zone Journal*, *15*(10), 0. https://doi.org/10.2136/vzj2016.04.0031

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1729–1735.

Khaki, M., & Awange, J. (2019). The application of multi-mission satellite data assimilation for studying water storage changes over South America. *Science of the Total Environment*, *647*(C), 1557–1572. https://doi.org/10.1016/j.scitotenv.2018.08.079

Khaki, M., Ait-El-Fquih, B., Hoteit, I., Forootan, E., Awange, J., & Kuhn, M. (2017). A two-update ensemble Kalman filter for land hydrological data assimilation with an uncertain constraint. *Journal of Hydrology*, *555*, 447–462. https://doi.org/10.1016/j.jhydrol.2017.10.032

Khaki, M., Forootan, E., Kuhn, M., Awange, J., van Dijk, A. I. J. M., Schumacher, M., & Sharifi, M. A. (2018). Determining water storage depletion within Iran by assimilating GRACE data into the W3RA hydrological model. *Advances in Water Resources*, *114*, 1–18. https://doi.org/10.1016/j.advwatres.2018.02.008

Khaki, M., Hoteit, I., Kuhn, M., Awange, J., Forootan, E., van Dijk, A. I. J. M., et al. (2017). Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. *Advances in Water Resources*, *107*, 1–16. https://doi.org/10.1016/j.advwatres.2017.07.001

Khaki, M., Schumacher, M., Forootan, E., Kuhn, M., Awange, J. L., & van Dijk, A. I. J. M. (2017). Accounting for spatial correlation errors in the assimilation of GRACE into hydrological models through localization. *Advances in Water Resources*, *108*, 99–112. https://doi.org/10.1016/j.advwatres.2017.07.024

Khandelwal, A., Karpatne, A., Marlier, M. E., Kim, J., Lettenmaier, D. P., & Kumar, V. (2017). An approach for global monitoring of surface water extent variations in reservoirs using MODIS data. *Remote Sensing of Environment*, *202*, 113–128. https://doi.org/10.1016/j.rse.2017.05.039

Koster, R. D., & Walker, G. K. (2015). Interactive vegetation phenology, soil moisture, and monthly temperature forecasts. *Journal of Hydrometeorology*, *16*(4), 1456–1465.

Koster, R. D., Betts, A. K., Dirmeyer, P. A., Bierkens, M., Bennett, K. E., Déry, S. J., et al. (2017). Hydroclimatic variability and predictability: a survey of recent research. *Hydrology and Earth System Sciences*, *21*(7), 3777–3798. https://doi.org/10.5194/hess-21-3777-2017

Koster, R. D., Mahanama, S. P. P., Yamada, T. J., Balsamo, G., Berg, A. A., Boisserie, M., et al. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters*, *37*(2), n/a–n/a. https://doi.org/10.1029/2009GL041677

Koster, R. D., Suarez, M. J., & Ducharne, A. (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of* ….

Kueppers, L. M., & Snyder, M. A. (2012). Influence of irrigated agriculture on diurnal surface energy and water fluxes, surface climate, and atmospheric circulation in California. *Climate Dynamics*, *38*(5-6), 1017–1029. https://doi.org/10.1007/s00382-011-1123-0

Kueppers, L. M., Snyder, M. A., & Sloan, L. C. (2007). Irrigation cooling effect: Regional climate forcing by land-use change. *Geophysical Research Letters*, *34*(3), L03703. https://doi.org/10.1029/2006GL028679

Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., & Bolten, J. (2018). Information theoretic evaluation of satellite soil moisture retrievals. *Remote Sensing of Environment*, *204*, 392–400. https://doi.org/10.1016/j.rse.2017.10.016

Kumar, S. V., Peters-Lidard, C. D., Santanello, J. A., Reichle, R. H., Draper, C. S., Koster, R. D., et al. (2015). Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. *Hydrology and Earth System Sciences*, *19*(11), 4463–4478. https://doi.org/10.5194/hess-19-4463-2015

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., & Houser, P. R. (2006). Land information system: An interoperable framework for high resolution land surface modeling. *Environmental Modelling & Software*, *21*(10), 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004

Kumar, S. V., Zaitchik, B. F., Peters-Lidard, C. D., Rodell, M., Reichle, R., Li, B., et al. (2016). Assimilation of Gridded GRACE Terrestrial Water Storage Estimates in the North American Land Data Assimilation System. *Journal of Hydrometeorology*, *17*(7), 1951–1972. https://doi.org/10.1175/jhm-d-15-0157.1

Kustu, M. D., Fan, Y., & Robock, A. (2010). Large-scale water cycle perturbation due to irrigation pumping in the US High Plains: A synthesis of observed streamflow changes. *Journal of Hydrology*, *390*(3-4), 222–244. https://doi.org/10.1016/j.jhydrol.2010.06.045

Kustu, M. D., Fan, Y., & Rodell, M. (2011). Possible link between irrigation in the U.S. High Plains and increased summer streamflow in the Midwest. *Water Resources Research*, *47*(3), n/a–n/a. https://doi.org/10.1029/2010WR010046

Lahmers, T., & Eden, S. (2018). Water and Irrigated Agriculture in Arizona.

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*(4). https://doi.org/10.1029/2011wr011453

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., et al. (2011). Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *Journal of Advances in …*, *3*(1), n/a–n/a. https://doi.org/10.1029/2011MS00045

Lawston, P. M., Santanello, J. A., Jr., Franz, T. E., & Rodell, M. (2017). Assessment of irrigation physics in a land surface modeling framework using non-traditional and human-practice datasets. *Hydrology and Earth System Sciences*, *21*(6), 2953–2966. https://doi.org/10.5194/hess-21-2953-2017

Lawston, P. M., Santanello, J. A., Jr., Zaitchik, B. F., & Rodell, M. (2015). Impact of Irrigation Methods on Land Surface Model Spinup and Initialization of WRF Forecasts. *Journal of Hydrometeorology*, *16*(3), 1135–1154. https://doi.org/10.1175/JHM-D-14-0203.1

Leng, G., Huang, M., & Tang, Q. (2015). A modeling study of irrigation effects on global surface water and groundwater resources under a changing climate. *Journal of Advances in …*, *7*(3), 1285–1304. https://doi.org/10.1002/2015ms000437

Leng, G., Huang, M., Tang, Q., Gao, H., & Leung, L. R. (2014). Modeling the Effects of Groundwater-Fed Irrigation on Terrestrial Hydrology over the Conterminous United States. *Journal of Hydrometeorology*, *15*(3), 957–972. https://doi.org/10.1175/jhm-d-13-049.1

Leng, G., Huang, M., Tang, Q., Sacks, W. J., Lei, H., & Leung, L. R. (2013). Modeling the effects of irrigation on land surface fluxes and states over the conterminous United States: Sensitivity to input data and model parameters. *Journal of Geophysical Research: Atmospheres*, *118*(17), 9789–9803. https://doi.org/10.1002/jgrd.50792

Leng, G., Leung, L. R., & Huang, M. (2017). Significant impacts of irrigation water sources and methods on modeling irrigation effects in the ACMELand Model. *Journal of Advances in …*, *9*(3), 1665–1683. https://doi.org/10.1002/2016ms000885

Li, B., & Rodell, M. (2015). Evaluation of a model-based groundwater drought indicator in the conterminous U.S. *Journal of Hydrology*, *526*, 78–88. https://doi.org/10.1016/j.jhydrol.2014.09.027

Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., et al. (2019). Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges. *Water Resources Research*, *55*(9), 7564–7586. https://doi.org/10.1029/2018WR024618

Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D., & van Dam, T. M. (2012). Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, *446-447*, 103–115. https://doi.org/10.1016/j.jhydrol.2012.04.035

Li, J., Mahalov, A., & Hyde, P. (2016). Impacts of agricultural irrigation on ozone concentrations in the Central Valley of California and in the contiguous United States based on WRF-Chem simulations. *Agricultural and Forest Meteorology*, *221*, 34–49. https://doi.org/10.1016/j.agrformet.2016.02.004

Liang, X. (1994). A two-layer variable infiltration capacity land surface representation for general circulation models.

Lin, P., Wei, J., Yang, Z.-L., Zhang, Y., & Zhang, K. (2016). Snow data assimilation-constrained land initialization improves seasonal temperature prediction. *Geophysical Research Letters*, *43*(21), 11–423–11–432.

Liu, J., Zhan, X., Hain, C., Yin, J., Fang, L., Li, Z., & Zhao, L. (2016). NOAA Soil Moisture Operational Product System (SMOPS) and its validations, 3477–3480.

Liu, Z., Liu, P.-W., Massoud, E., Farr, T. G., Lundgren, P., & Famiglietti, J. S. (2019). Monitoring Groundwater Change in California's Central Valley Using Sentinel-1 and GRACE Observations. *Geosciences*, *9*(10), 436–21. https://doi.org/10.3390/geosciences9100436

Long, D., Longuevergne, L., & Scanlon, B. R. (2015). Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resources Research*, *51*(4), 2574–2594. https://doi.org/10.1002/2014WR016853

Mahalov, A., Li, J., & Hyde, P. (2016). Regional impacts of irrigation in Mexico and southwestern U.S. on hydrometeorological fields in the North American Monsoon region. *Journal of Hydrometeorology*, JHM–D–15–0223.1–41. https://doi.org/10.1175/JHM-D-15-0223.1

McDermid, S. S., Mearns, L. O., & Ruane, A. C. (2017). Representing agriculture in Earth System Models: Approaches and priorities for development. *Journal of Advances in …*, *9*(5), 2230–2265. https://doi.org/10.1002/2016MS000749

McGuire, V. L. (2014). Water-Level Changes and Change in Water in Storage in the High Plains Aquifer, Predevelopment to 2013 and 2011–13, 1–24.

McGuire, V. L. (2017). Water-Level Changes and Change in Recoverable Water in Storage, High Plains Aquifer, Predevelopment to 2015 and 2013–15, 1–24.

McMahon, P. B., Dennehy, K. F., Bruce, B. W., Böhlke, J. K., Michel, R. L., Gurdak, J. J., & Hurlbut, D. B. (2006). Storage and transit time of chemicals in thick unsaturated zones under rangeland and irrigated cropland, High Plains, United States. *Water Resources Research*, *42*(3), 157–18. https://doi.org/10.1029/2005WR004417

Miller, J. A., & Appel, C. L. (1997). Ground Water Atlas of the United States: Segment 3, Kansas, Missouri, Nebraska.

Nazemi, A., & Wheater, H. S. (2015a). On inclusion of water resource management in Earth system models – Part 1: Problem definition and representation of water demand. *Hydrology and Earth System Sciences*, *19*(1), 33–61. https://doi.org/10.5194/hess-19-33-2015

Nazemi, A., & Wheater, H. S. (2015b). On inclusion of water resource management in Earth system models – Part 2: Representation of water supply and allocation and opportunities for improved modeling. *Hydrology and Earth System Sciences*, *19*(1), 63–90. https://doi.org/10.5194/hess-19-63-2015

Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Anderson, M. C., & Hain, C. (2018). Groundwater Withdrawals Under Drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer. *Water Resources Research*, *48*(3), 317. https://doi.org/10.1029/2017WR022178

Nielsen-Gammon, J. (2011). The 2011 Texas drought: a briefing packet for the Texas Legislature.

Niu, G.-Y., Yang, Z.-L., Dickinson, R. E., Gulden, L. E., & Su, H. (2007). 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), 15. https://doi.org/10.1029/2006JD007522

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*(D12), D12109–19. https://doi.org/10.1029/2010JD015139

Orang, M. N., Snyder, R. L., Shu, G., Hart, Q. J., Sarreshteh, S., Falk, M., et al. (2013). California Simulation of Evapotranspiration of Applied Water and Agricultural Energy Use in California. *Journal of Integrative Agriculture*, *12*(8), 1371–1388. https://doi.org/10.1016/S2095-3119(13)60742-X

Ozdogan, M., Rodell, M., Beaudoing, H. K., & Toll, D. L. (2010). Simulating the Effects of Irrigation over the United States in a Land Surface Model Based on Satellite-Derived Agricultural Data. *Journal of Hydrometeorology*, *11*(1), 171–184. https://doi.org/10.1175/2009JHM1116.1

Faunt, C. C. (2009). Groundwater availability of the Central Valley aquifer, California. US Geological Survey professional paper.

Pei, L., Moore, N., Zhong, S., Kendall, A. D., Gao, Z., & Hyndman, D. W. (2016). Effects of Irrigation on Summer Precipitation over the United States. *Journal of Climate*, *29*(10), 3541–3558. https://doi.org/10.1175/JCLI-D-15-0337.1

Pipunic, R. C., Walker, J. P., & WESTERN, A. (2008). Assimilation of remotely sensed data for improved latent and sensible heat flux prediction: A comparative synthetic study. *Remote Sensing of Environment*, *112*(4), 1295–1305.

Pokhrel, Y. N., Felfelani, F., Shin, S., Yamada, T. J., & Satoh, Y. (2017). Modeling large-scale human alteration of land surface hydrology and climate. *Geoscience Letters*, 1–13. https://doi.org/10.1186/s40562-017-0076-5

Pokhrel, Y. N., Hanasaki, N., Wada, Y., & Kim, H. (2016). Recent progresses in incorporating human land-water management into global land surface models toward their integration into Earth system models. *Wiley Interdisciplinary Reviews: Water*, *3*(4), 548–574. https://doi.org/10.1002/wat2.1150

Pokhrel, Y. N., Koirala, S., Yeh, P. J. F., Hanasaki, N., Longuevergne, L., Kanae, S., & Oki, T. (2015). Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. *Water Resources Research*, *51*(1), 78–96. https://doi.org/10.1002/2014WR015602

Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P. J. F., Kim, H., et al. (2012). Incorporating Anthropogenic Water Regulation Modules into a Land Surface Model. *Journal of Hydrometeorology*, *13*(1), 255–269. https://doi.org/10.1175/JHM-D-11-013.1

Pryor, S. C., SULLIVAN, R. C., & Wright, T. (2016). Quantifying the Roles of Changing Albedo, Emissivity, and Energy Partitioning in the Impact of Irrigation on Atmospheric Heat Content. *Journal of Applied Meteorology and Climatology*, *55*(8), 1699–1706. https://doi.org/10.1175/jamc-d-15-0291.1

Reager, J. T., & Famiglietti, J. S. (2009). Global terrestrial water storage capacity and flood potential using GRACE. *Geophysical Research Letters*, *36*(23), 1147–6. https://doi.org/10.1029/2009GL040826

Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in Water Resources*, *31*(11), 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001

Reichle, R. H., & Koster, R. D. (2003). Assessing the Impact of Horizontal Error Correlations in Background Fields on Soil Moisture Estimation. *Journal of Hydrometeorology*, *4*(6), 1229–1242. https://doi.org/10.1175/1525-7541(2003)004<1229:atiohe>2.0.co;2

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., & Owe, M. (2007). Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research*, *112*(D9), 1697–14. https://doi.org/10.1029/2006JD008033

Richey, A. S., Thomas, B. F., Lo, M. H., Famiglietti, J. S., Swenson, S., & Rodell, M. (2015). Uncertainty in global groundwater storage estimates in a total groundwater stress framework. *Water Resources Research*, *51*(7), 5198–5216.

Richey, A. S., Thomas, B. F., Lo, M. H., Reager, J. T., Famiglietti, J. S., Voss, K., et al. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resources Research*, *51*(7), 5217–5238. https://doi.org/10.1002/2015wr017349

Rippey, B. R. (2015). The US drought of 2012. *Weather and Climate Extremes*, *10*, 57–64. https://doi.org/10.1016/j.wace.2015.10.004

Rodell, M., & Famiglietti, J. S. (2001). 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. https://doi.org/10.1029/2000WR900306

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M. H. (2018). Emerging trends in global freshwater availability. *Nature*, *37*, 1327. https://doi.org/10.1038/s41586-018-0123-1

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*(7258), 999–1002. https://doi.org/10.1038/nature08238

Rowlands, D. D., Luthcke, S. B., Klosko, S. M., Lemoine, F. G. R., Chinn, D. S., McCarthy, J. J., et al. (2005). Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. *Geophysical Research Letters*, *32*(4), n/a–n/a. https://doi.org/10.1029/2004gl021908

Russo, T. A., & Lall, U. (2017). Depletion and response of deep groundwater to climate-induced pumping variability. *Nature Geoscience*, *10*(2), 105.

Ryder, P. D. (1996). Ground water atlas of the united states: Oklahoma, texas. *US Geological Survey Publication HA*.

Ryu, D., Crow, W. T., Zhan, X., & Jackson, T. J. (2009). Correcting Unintended Perturbation Biases in Hydrologic Data Assimilation. *Journal of Hydrometeorology*, *10*(3), 734–750. https://doi.org/10.1175/2008JHM1038.1

Sacks, W. J., Cook, B. I., Buenning, N., Levis, S., & Helkowski, J. H. (2008). Effects of global irrigation on the near-surface climate. *Climate Dynamics*, *33*(2-3), 159–175. https://doi.org/10.1007/s00382-008-0445-z

Salmon, J. M., Friedl, M. A., Frolking, S., & Wisser, D. (2015). Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *International Journal of …*, *38*, 321–334. https://doi.org/10.1016/j.jag.2015.01.014

Scanlon, B. R., & Faunt, C. C. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Presented at the Proceedings of the …. https://doi.org/10.1073/pnas.1200311109/-/DCSupplemental/pnas.201200311SI.pdf

Scanlon, B. R., Zhang, Z., Reedy, R. C., Pool, D. R., Save, H., Long, D., et al. (2015). Hydrologic implications of GRACE satellite data in the Colorado River Basin. *Water Resources Research*, *51*(12), 9891–9903. https://doi.org/10.1002/2015WR018090

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., et al. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences*, *115*(6), E1080–E1089. https://doi.org/10.1073/pnas.1704665115

Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Di Long, et al. (2016). Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resources Research*, *52*(12), 9412–9429. https://doi.org/10.1002/2016wr019494

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., et al. (2014). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences*, *111*(9), 3245–3250. https://doi.org/10.1073/pnas.1222460110

Schumacher, M., Forootan, E., van Dijk, A. I. J. M., Schmied, H. M., Crosbie, R. S., Kusche, J., & Döll, P. (2018). Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sensing of Environment*, *204*, 212–228. https://doi.org/10.1016/j.rse.2017.10.029

Schumacher, M., Kusche, J., & Döll, P. (2016). A systematic impact assessment of GRACE error correlation on data assimilation in hydrological models. *Journal of Geodesy*, *90*(6), 537–559. https://doi.org/10.1007/s00190-016-0892-y

Schuurmans, J. M., Troch, P. A., Veldhuizen, A. A., Bastiaanssen, W., & Bierkens, M. (2003). Assimilation of remotely sensed latent heat flux in a distributed hydrological model. *Advances in Water Resources*, *26*(2), 151–159.

Schwarz, G. E., & Alexander, R. B. (1995). State soil geographic (STATSGO) data base for the conterminous United States.

Shamsudduha, M., Taylor, R. G., & Longuevergne, L. (2012). Monitoring groundwater storage changes in the highly seasonal humid tropics: Validation of GRACE measurements in the Bengal Basin. *Water Resources Research*, *48*(2). https://doi.org/10.1029/2011wr010993

Shamsudduha, M., Taylor, R. G., Jones, D., Longuevergne, L., Owor, M., & Tindimugaya, C. (2017). Recent changes in terrestrial water storage in the Upper Nile Basin: an evaluation of commonly used gridded GRACE products. *Hydrology and Earth System Sciences*, *21*(9), 4533–4549. https://doi.org/10.5194/hess-21-4533-2017

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation – a global inventory. *Hydrology and Earth System Sciences*, *14*(10), 1863–1880. https://doi.org/10.5194/hess-14-1863-2010

Solander, K. C., Reager, J. T., Wada, Y., Famiglietti, J. S., & Middleton, R. S. (2017). GRACE satellite observations reveal the severity of recent water over- consumption in the United States. *Nature Publishing Group*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-07450-y

Steffen, W., Grinevald, J., Crutzen, P., & McNeill, J. (2011). The Anthropocene: conceptual and historical perspectives. *Philosophical Transactions of the Royal Society of London a: Mathematical, Physical and Engineering Sciences*, *369*(1938), 842–867. https://doi.org/10.1098/rsta.2010.0327

Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research*, *103*(C4), 7755–7766.
Swenson, S., & Wahr, J. (2006). Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters*, *33*(8). https://doi.org/10.1029/2005gl025285
Tang, Q., Rosenberg, E. A., & Lettenmaier, D. P. (2009). Use of satellite data to assess the impacts of irrigation withdrawals on Upper Klamath Lake, Oregon. *Hydrology and Earth System Sciences*, *13*(5), 617–627.
Tang, Y., Hooshyar, M., Zhu, T., Ringler, C., Sun, A. Y., Di Long, & Wang, D. (2017). Reconstructing annual groundwater storage changes in a large-scale irrigation region using GRACE data and Budyko model. *Journal of Hydrology*, *551*, 397–406. https://doi.org/10.1016/j.jhydrol.2017.06.021
Tangdamrongsub, N., Han, S.-C., Tian, S., Schmied, H. M., Sutanudjaja, E. H., Ran, J., & Feng, W. (2018). Evaluation of Groundwater Storage Variations Estimated from GRACE Data Assimilation and State-of-the-Art Land Surface Models in Australia and the North China Plain. *Remote Sensing*, *10*(3), 483. https://doi.org/10.3390/rs10030483
Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., & Weerts, A. H. (2015). 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. https://doi.org/10.5194/hess-19-2079-2015
Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE measurements of mass variability in the Earth system. *Science*, *305*(5683), 503–505.
Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., et al. (2013). Ground water and climate change. *Nature Climate Change*, *3*(4), 322–329. https://doi.org/10.1038/nclimate1744
Thomas, B. F., Famiglietti, J. S., Landerer, F. W., Wiese, D. N., Molotch, N. P., & Argus, D. F. (2017). GRACE Groundwater Drought Index: Evaluation of California Central Valley groundwater drought. *Remote Sensing of Environment*, *198*, 384–392. https://doi.org/10.1016/j.rse.2017.06.026
Tian, S., Tregoning, P., Renzullo, L. J., van Dijk, A. I. J. M., Walker, J. P., Pauwels, V. R. N., & Allgeyer, S. (2017). Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals. *Water Resources Research*, *53*(3), 1820–1840. https://doi.org/10.1002/2016WR019641
Tiwari, V. M., Wahr, J., & Swenson, S. (2009). Dwindling groundwater resources in northern India, from satellite gravity observations. *Geophysical Research Letters*, *36*(18), L18401. https://doi.org/10.1029/2009GL039401
Use of satellite data to assess the impacts of irrigation withdrawals on Upper Klamath Lake, Oregon. (2009). Use of satellite data to assess the impacts of irrigation withdrawals on Upper Klamath Lake, Oregon. *Hydrology and Earth System Sciences*, *13*(5), 617–627. https://doi.org/10.5194/hess-13-617-2009
van Beek, L. P. H., Wada, Y., & Bierkens, M. F. P. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, *47*(7), 15. https://doi.org/10.1029/2010WR009791
van Dijk, A. I. J. M., Renzullo, L. J., & Rodell, M. (2011). Use of Gravity Recovery and Climate Experiment terrestrial water storage retrievals to evaluate model estimates by the Australian water resources assessment system. *Water Resources Research*, *47*(11). https://doi.org/10.1029/2011wr010714

Wada, Y. (2015). Modeling Groundwater Depletion at Regional and Global Scales: Present State and Future Prospects. *Surveys in Geophysics*, *37*(2), 419–451. https://doi.org/10.1007/s10712-015-9347-x

Wada, Y., Beek, L. P. H., & Bierkens, M. F. P. (2012). Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resources Research*, *48*(6), W00L06. https://doi.org/10.1029/2011WR010562

Wada, Y., Bierkens, M., & de Roo, A. (2017). Human-water interface in hydrological modeling: Current status and future directions. *Hydrology and Earth* …. Retrieved from http://www.hydrol-earth-syst-sci-discuss.net/hess-2017-248/hess-2017-248.pdf

Wada, Y., van Beek, L. P. H., van Kempen, C. M., Reckman, J. W. T. M., Vasak, S., & Bierkens, M. F. P. (2010). Global depletion of groundwater resources. *Geophysical Research Letters*, *37*(20), n/a–n/a. https://doi.org/10.1029/2010GL044571

Wada, Y., Wisser, D., & Bierkens, M. F. P. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, *5*(1), 15–40. https://doi.org/10.5194/esd-5-15-2014

Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., et al. (2013). Multimodel projections and uncertainties of irrigation water demand under climate change. *Geophysical Research Letters*, *40*(17), 4626–4632. https://doi.org/10.1002/grl.50686

Wahr, J., Molenaar, M., & Bryan, F. (1998). Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Atmospheres*, *103*(B12), 30205–30229. https://doi.org/10.1029/98JB02844

Wahr, J., Swenson, S., & Velicogna, I. (2006). Accuracy of GRACE mass estimates. *Geophysical Research Letters*, *33*(6), 30205–5. https://doi.org/10.1029/2005GL025305

Wallander, S., Aillery, M., Hellerstein, D., & Hand, M. (2013). The Role of Conservation Programs in Drought Risk Adaptation.

Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, *120*(4), 2648–2671. https://doi.org/10.1002/2014JB011547

Wiese, D. N., Landerer, F. W., & Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resources Research*, *52*(9), 7490–7502. https://doi.org/10.1002/2016WR019344

Wisser, D., Fekete, B. M., Vörösmarty, C. J., & Schumann, A. H. (2010). Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H). *Hydrology and Earth System Sciences*, *14*(1), 1–24. https://doi.org/10.5194/hess-14-1-2010

Wood, E. F., Lettenmaier, D. P., & Zartarian, V. G. (1992). A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research: Atmospheres*, *97*(D3), 2717–2728. https://doi.org/10.1029/91JD01786

Xia, Y., Ek, M., Wei, H., & Meng, J. (2011). Comparative analysis of relationships between NLDAS-2 forcings and model outputs. *Hydrological Processes*, *26*(3), 467–474. https://doi.org/10.1002/hyp.8240

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, *117*(D3), n/a–n/a. https://doi.org/10.1029/2011jd016048

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research*, *116*(D12), D12110–16. https://doi.org/10.1029/2010JD015140

Yilmaz, M. T., Anderson, M. C., Zaitchik, B., Hain, C. R., Crow, W. T., Ozdogan, M., et al. (2014). Comparison of prognostic and diagnostic surface flux modeling approaches over the Nile River basin. *Water Resources Research*, *50*(1), 386–408. https://doi.org/10.1002/2013WR014194

Zaitchik, B. F., Evans, J., & Smith, R. B. (2005). MODIS-Derived Boundary Conditions for a Mesoscale Climate Model: Application to Irrigated Agriculture in the Euphrates Basin. *Monthly Weather Review*, *133*(6), 1727–1743. https://doi.org/10.1175/mwr2947.1

Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008). 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. https://doi.org/10.1175/2007JHM951.1

Zeng, Y., Xie, Z., & Zou, J. (2016). Hydrologic and climatic responses to global anthropogenic groundwater extraction. *Journal of Climate*, JCLI–D–16–0209.1–50. https://doi.org/10.1175/JCLI-D-16-0209.1

Zhao, L., & Yang, Z.-L. (2018). Multi-sensor land data assimilation: Toward a robust global soil moisture and snow estimation. *Remote Sensing of Environment*, *216*, 13–27. https://doi.org/10.1016/j.rse.2018.06.033

Zhong, Y., Zhong, M., Feng, W., Zhang, Z., Shen, Y., & Wu, D. (2018). Groundwater Depletion in the West Liaohe River Basin, China and Its Implications Revealed by GRACE and In Situ Measurements. *Remote Sensing*, *10*(4), 493–16. https://doi.org/10.3390/rs10040493

Zou, J., Xie, Z., Yu, Y., Zhan, C., & Sun, Q. (2013). Climatic responses to anthropogenic groundwater exploitation: a case study of the Haihe River Basin, Northern China. *Climate Dynamics*, *42*(7-8), 2125–2145. https://doi.org/10.1007/s00382-013-1995-2

Biography

Wanshu Nie received a Bachelor of Engineering degree in Hydraulic Engineering from Tsinghua University in 2013 and a Master of Engineering degree in the same place in 2015. She then enrolled in a Ph.D. program at Johns Hopkins University, in the Department of Earth and Planetary Science. Starting in Jan 2020, Wanshu will work on soil moisture data assimilation under Dr. Christa Peters-lidard as NASA-GESTAR supported postdoctoral researcher.