Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California

Agarwal, Vibhor ; Akyilmaz, Orhan ; Shum, CK; Feng, Wei; Yang, Ting-Yi ; Forootan, Ehsan; Syed, Tajdarul Hassan ; Haritashya, Umesh K. ; Uz, Metehan

Published in:
Science of the Total Environment

DOI (link to publication from Publisher):
10.1016/j.scitotenv.2022.161138

Publication date:
2023

Document Version
Accepted author manuscript, peer reviewed version

Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California.

Vibhor Agarwal[1,2,3], Orhan Akyilmaz[4], C.K. Shum[3,5], Wei Feng[5,6], Ting-Yi Yang[7], Ehsan Forootan[8], Tajdarul Hassan Syed[9], Umesh K. Haritashya[2], Metehan Uz[3]

[1] Department of Earth Sciences, College of Wooster, USA

[2] Department of Geology and Environmental Geosciences, University of Dayton, USA

[3] Division of Geodetic Science, School of Earth Sciences, The Ohio State University, USA

[4] Department of Geomatic Engineering, Istanbul Technical University, Turkey

[5] Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, China

[6] School of Geospatial Engineering and Science, Sun Yat-sen University, China;

[7] Polaris Geospatial Services, Columbus, Ohio, USA;

[8] Department of Planning, Aalborg University, Denmark;

[9] Department of Earth Sciences, Indian Institute of Technology Kanpur, India

Corresponding Author: Vibhor Agarwal, agarwal.282@osu.edu agarwal.vibhor.phd@gmail.com

Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California.

Vibhor Agarwal[1,2,3], Orhan Akyilmaz[4], C.K. Shum[3,5], Wei Feng[5,6], Ting-Yi Yang[7], Ehsan Foroota[8], Tajdarul Hassan Syed[9], Umesh K. Haritashya[2], Metehan Uz[3]

[1] Department of Earth Sciences, College of Wooster, USA

[2] Department of Geology and Environmental Geosciences, University of Dayton, USA

[3] Division of Geodetic Science, School of Earth Sciences, The Ohio State University, USA

[4] Department of Geomatic Engineering, Istanbul Technical University, Turkey

[5] Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, China

[6] School of Geospatial Engineering and Science, Sun Yat-sen University, China;

[7] Polaris Geospatial Services, Columbus, Ohio, USA;

[8] Department of Planning, Aalborg University, Denmark;

[9] Department of Earth Sciences, Indian Institute of Technology Kanpur, India

Corresponding Author: Vibhor Agarwal, agarwal.282@osu.edu agarwal.vibhor.phd@gmail.com

Abstract

California's Central Valley, one of the most agriculturally productive regions, is also one of the most stressed aquifers in the world due to anthropogenic groundwater over-extraction primarily for irrigation. The groundwater depletion is further exacerbated by climate-stressed droughts. Gravity Recovery and Climate Experiment (GRACE) satellite gravimetry has demonstrated the feasibility of quantifying global groundwater storage changes at uniform monthly sampling, though at a coarse resolution and is thus impractical for effective water resources management). Here, we employ the Random Forest machine learning algorithm to establish empirical relationships between GRACE-derived groundwater storage and in-situ groundwater level variations over the Central Valley during 2002–2016 and achieved downscaling of GRACE-observed groundwater storage changes from 666 km to 5 km. Validations of our modeled groundwater level with *in situ* groundwater level indicate excellent Nash-Sutcliffe Efficiency coefficients ranging from 0.94–0.97. In addition, the modeled groundwater trends have good agreements with two independent measurements of vertical land subsidence measured rates using GPS, and CryoSat-2 radar altimetry. Our estimated groundwater loss is about 30 km^3 during 2002–2016, which agrees well with previous studies. We find the maximum groundwater storage losses of -5.7 ± 1.2 $km^3\ yr^{-1}$ and -9.8 ± 1.7 $km^3\ yr^{-1}$ occurred during the extended drought periods of January 2007-December 2009, and October 2011-September 2015, respectively. We observed that Central Valley experienced groundwater recharges during abrupt winter flood episodes. The 5-km resolution Central Valley-wide groundwater storage trends reveal that groundwater depletion occurs mostly in southern San Joaquin Valley and is collocated with sites showing severe land subsidence due to aquifer compaction from groundwater over withdrawal.

Keywords: Machine Learning, Groundwater, GRACE, Remote Sensing

1. Introduction

Groundwater is an important freshwater resource that meets agricultural, industrial, and domestic needs (Siebert et al., 2010; Wada et al., 2014; Zekster and Everett, 2004). Over the past few decades, several aquifers worldwide such as Central Valley, High Plains, Indus Plain, middle East, and others, have faced unprecedented human-induced stress due to the population growth, expansion of the irrigated areas, and other economic activities causing a drastic increase in groundwater usage (Bierkens and Wada, 2019; Famiglietti, 2014). Groundwater abstraction and outflow exceeding groundwater recharge over a long period of time and in large areas has been reported as the main cause of groundwater depletion (Konikow and Kendy, 2005; Wada et al., 2010). Groundwater depletion can lead to global water security and environmental issues (Famiglietti, 2014; Wada et al., 2010). There is an urgent need for quantifying long-term groundwater storage variations (GWS) at frequent temporal samplings that can help characterize the groundwater depletion in these stressed regions.

Several approaches for quantifying GWS variations have been applied (e.g., Bierkens and Wada, 2019). Groundwater levels from *in-situ* ground wells provide essential information about stresses acting on the aquifers and play a key role in developing groundwater models (Taylor and Alley, 2001). Continuous groundwater level observations may further help quantify GWS and predict future trends in storage (Butler et al., 2013; Sun et al., 2013). However, it is infeasible to use only these data for quantifying regional GWS for several reasons. Firstly, monitoring wells required to accurately estimate groundwater levels are expensive to install and maintain. Therefore, several aquifers have poor coverage of such wells. Secondly, spatio-temporal gaps in the coverage of ground wells might necessitate the interpolation of groundwater level data, leading to interpolation errors (Ahamed et al., 2022; Thomas et al., 2017). Thirdly, uncertainties in the value of storage coefficient at well sites might translate into errors in GWS (Scanlon et al., 2012; Alam et al., 2021). Since 2002, the Gravity Recovery and Climate Experiment (GRACE) twin-satellite mission gravimetry data have enabled a continuous global Terrestrial Water Storage (TWS) record for over a decade and a half, at a spatial resolution larger than 333 km (half-wavelength) and monthly sampling, e.g., Frappart et al., (2018). Innovative processing of GRACE data has enabled the uniform global quantification of GWS change for the first time by removing surface water storage changes using hydrologic data and model outputs (Famiglietti et al., 2011; Rodell et al., 2009), as well as data assimilation (50 km resolution in Mehrnegar et al., (2021); 12.5 km resolution in Schumacher et al., (2018)). However, due to the limited spatial resolution and the associated errors in disaggregating GRACE-derived TWS (Scanlon et al., 2012), the application of GRACE data directly for groundwater assessment is not feasible at the local scale (Alley and Konikow, 2015), including in Central Valley in California which is the subject of the present study. Moreover, most GRACE-based groundwater studies estimate GWS variations by removing soil moisture estimates simulated by Land Surface Models (LSMs) from GRACE observations of terrestrial water storage (Scanlon et al., 2012). However, LSMs do not simulate irrigation water use; hence soil moisture values will be particularly erroneous in the Central Valley, where groundwater irrigation is predominant (Famiglietti et al., 2011).

Other methods of GWS computation include the water balance method, where several hydro-meteorologic quantities, such as the difference between stream inflow and outflow, precipitation, and evapotranspiration, along with several of the storage changes (soil moisture, snow water equivalent, reservoir storage) are computed for a given aquifer (Ahamed et al., 2022; Xiao et al., 2017). Several remote sensing, in-situ, and modeled datasets can be used to estimate the above quantities. However, these quantities might be subject to sources of uncertainties that creep into the water balance equation (Ahamed et al., 2022; Bierkens and Wada, 2019). Further, vertical deformation from GPS and Interferometric Synthetic Aperture Radar (InSAR) can provide regional estimates of groundwater storage changes (Ojha et al., 2018) or trends in the case of InSAR. With the availability of Sentinel-1 data since 2014, this method holds promise for revealing aquifer dynamics and storage variations at high spatial resolutions (Castellazzi et al., 2016; Vasco et al., 2021). However, often in lieu of missing a continuous and uniform time series, InSAR land velocity or trend/subsidence estimates could potentially be biased due primarily to interannual or longer signals in the land deformation data.

Machine Learning (ML) has been used for solving several non-linear complex problems in geoscience, e.g., Dramsch et al. (2020) and Sun and Scanlon (2019), as it does not require the knowledge of exact physical relationships between input and target variables. Machine Learning can also be used to estimate GWS variations at a higher resolution if GRACE-derived TWS variations can be downscaled to model in-situ groundwater level variations. A suitable combination of hydro-meteorological variables should be identified as input variables to build a robust machine-learning algorithm to model groundwater variations (Adamowski and Chan, 2011). Several studies in the past have incorporated machine learning models like Artificial Neural Network (ANN) model, Random Forest, Boosted Regression Tree, and Deep Learning to downscale GRACE satellite data to produce GWS variations at high resolution (Chen et al., 2019; Chen et al., 2020; Miro and Famiglietti, 2018; Rahaman et al., 2020).

Quantifying GWS variations is especially important for Central Valley. Here, ever-increasing irrigation demands, limited availability of surface water, and climate extremes such as prolonged and intensified droughts resulting from climate change have forced farmers to depend more on groundwater. As a result of the continuing groundwater depletion, several adverse impacts such as falling groundwater levels, decreasing groundwater yields, increase in pumping costs, degrading water quality, and damage to the aquatic ecosystems and wetlands have been observed (Faunt, 2009; Faunt and Sneed, 2015; Konikow, 2015). San Joaquin Valley, a major agricultural region in Central Valley, has witnessed the largest share of such adverse impacts, which have become more severe during prolonged and recurrent droughts in California.

Several of the methods mentioned above have been applied to quantify the GWS in Central Valley. Famiglietti et al., (2011) used GRACE-derived TWS variations and other hydrological variables to quantify GWS variations during 2002-2011. Scanlon et al., (2012) used updated GRACE processing and in-situ groundwater level variations to compute groundwater depletion from 2002-2011. Ojha et al., (2018) used vertical deformation derived from InSAR to derive the storage changes. Alam et al., (2021) used a combination of GRACE, wells, water balance, and hydrological modeling to quantify GWS variations from 2003-2019. Ahamed et al., (2022) used remote sensing data and an ensemble of water balance methods to quantify groundwater storage variations in Central Valley during 2002-2020. Miro and Famiglietti, (2018) implemented ANN using GRACE-derived TWS variations along with hydro-meteorologic variables to model annual amplitudes of GWS variations at 4 km spatial resolution over a small portion of the San Joaquin Valley in Central Valley during 2002-2010. All the above studies have confirmed the continued loss of groundwater losses along with dramatic rates of subsidence due to groundwater overdrafts during the last two decades.

All the above methods, except those utilizing in-situ wells and machine learning techniques, have limited capability to model GWS variations at high spatial resolutions at frequent temporal intervals. Groundwater levels in Central Valley can reflect complex variations due to withdrawal for irrigation, recharge due to partial infiltration of irrigation water, surface water impoundment, or precipitation. Further, climate extremes such as drought have put unprecedented stress on groundwater reserves which might be reflected in the groundwater fluctuations (Faunt, 2009). Compared to interpolation and kriging, more robust approaches are needed to fill the spatio-temporal gaps in in-situ groundwater levels and possibly obtain regional storage variations (Alam et al., 2021; Thomas et al., 2017). Previous machine learning-based approaches (Miro and Famiglietti, 2018) can be further expanded to the whole of Central Valley to cover broader spatio-temporal scales and improve the accuracy of modeled results using a suitable choice of input variables, models and better approaches for training the machine learning model.

The primary objective of this study is to downscale GRACE-derived GWS variations in Central Valley, California, using the Random Forest machine learning algorithm. We chose the period from October 2002-September 2016, which covers most of the operational phase of GRACE satellite data. We use GRACE along with hydro-meteorologic/geologic data as input and in-situ groundwater level data as the target data for the model. Further, the Central Valley has a record of geodetic measurements from GPS, extensometers, and remote sensing observations, which have been used to quantify the subsidence due to groundwater overdraft (Ojha et al., 2018; Sneed and Brandt, 2015). These data can provide us with ancillary information against which we can further validate our modeled results as a part of our second objective. Primarily, we compared the modeled groundwater level with the vertical deformation obtained from GPS and altimeter and obtained an inelastic storage coefficient for a portion of Central Valley. This approach of combining multiple hydrological and geodetic data can further enhance our understanding of aquifer dynamics, which is important for regional studies such as the one presented here. A machine learning approach might help provide relevant local-scale data on groundwater depletion to Groundwater Sustainability Agencies (GSAs) to make informed management decisions required to support the goals of the Sustainable Groundwater Management Act.

2. Study area

The Central Valley aquifer system in California covers an area of 52,000 km^2 (Figure 1) and produces one-fourth of the food in the US (Faunt, 2009). Central Valley is primarily semi-arid and most precipitation occurs during the winter and early spring months and not in summer when it is needed for irrigation and drinking (Jasechko et al., 2020). San Joaquin Valley is the major agricultural region and surface water quantity here depends on seasonal snowmelt from the Sierra Nevada in the East and Sacramento Valley in the North, which varies from year to year. Consequently, supplies for irrigation must be met through diverted surface water sources, and through groundwater from confined and unconfined aquifers. Groundwater is, therefore, an essential/persistent freshwater source accounting for up to 40% or more of the required water supply in Central Valley.

Central Valley lost approximately 113 km^3 of groundwater in the 20^{th} century and 20 percent of this depletion is estimated to be contributing to land subsidence (Faunt, 2009). Consequently, groundwater levels have been declining since the 1930s when the first in-situ measurement was made (Bertoldi, 1989; Williamson et al., 1989). Groundwater storage losses from GRACE satellite observations and Central Valley Hydrological Model for the first decade of the 21^{st} century is 25-30 km^3 (Konikow, 2013).

As groundwater depletion continues in Central Valley and other nearby regions, Sustainable Groundwater Management Act was passed in 2014 in California to promote better groundwater management and governance. Through this act, more emphasis is laid on the sustenance of groundwater resources for all regions by optimizing the water consumption by agricultural and other sectors. This issue is extremely critical for Central Valley as impacts of depletion here are visible from the 1920s on the local scale.

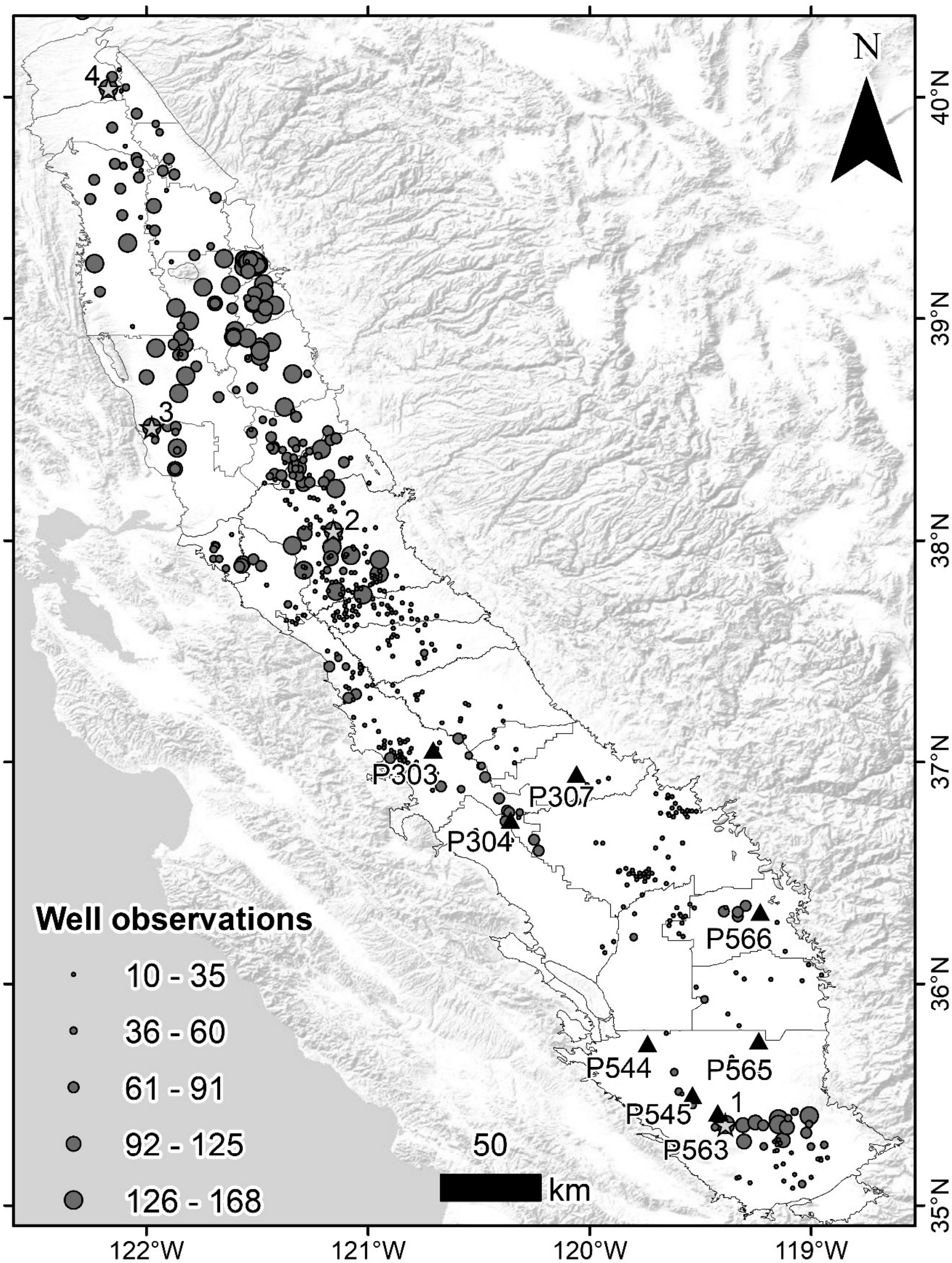
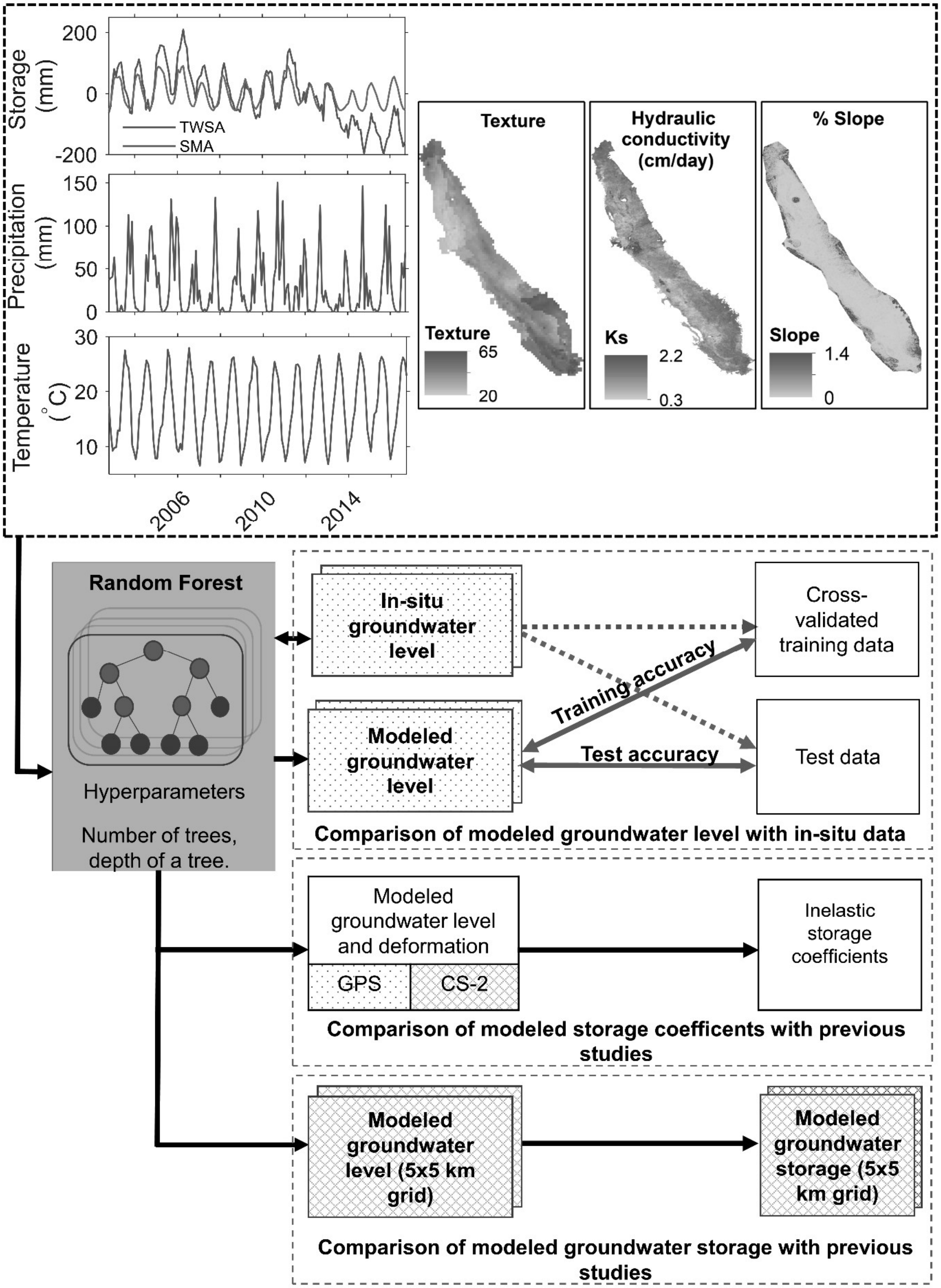


Figure 1. Location of Central Valley and two major basins, Sacramento (black) and San Joaquin (blue) Valley in north and south, respectively. The location of wells used in this study

and the number of measurements over the study period is also shown with filled red circles. 7 GPS sites used in this study are shown by black triangles. The green stars represent the wells used for plotting in Figure

3. Data and Methods

We implemented the Random Forest (RF) machine learning (ML) model and followed the data processing workflow as in Figure 2.


Storage (mm)
200
0
-200
TWSA
SMA
Precipitation (mm)
150
100
50
0
Temperature (°C)
30
20
10
2006
2010
2014
Texture
Texture
65
20
Hydraulic conductivity (cm/day)
Ks
2.2
0.3
% Slope
Slope
1.4
0
Random Forest
Hyperparameters
Number of trees, depth of a tree.
In-situ groundwater level
Modeled groundwater level
Cross-validated training data
Test data
Training accuracy
Test accuracy
Comparison of modeled groundwater level with in-situ data
Modeled groundwater level and deformation
GPS
CS-2
Inelastic storage coefficients
Comparison of modeled storage coefficents with previous studies
Modeled groundwater level (5x5 km grid)
Modeled groundwater storage (5x5 km grid)
Comparison of modeled groundwater storage with previous studies

197

Figure 2. The workflow used for modeling groundwater level

3.1. Input and target data

3.1.1. Precipitation and temperature

Here monthly precipitation and temperature data is obtained from Parameter-Elevation Regression on Independent Slopes Model (PRISM) dataset at 4 km spatial resolution (Daly et al., 2008). PRISM simulates the spatial variations of the weather and climate using in-situ data. It uses a “weighted regression scheme” to account for different physiographic features and climate regimes when providing final estimates of precipitation and temperature.

Since precipitation can take a few months to recharge groundwater, we use 0, 1, 2, 3, and 4-month lags for precipitation labeled as PPT0, PPT1, PPT2, PPT3, and PPT4, respectively, in this study.

3.1.2. Terrestrial Water Storage and Soil Moisture Variations

For the computation of TWS, we used the latest GRACE data product, the Release (RL) 06 Level 2 (L2) monthly gravity field solutions provided by the University of Texas at Austin Center for Space Research (CSR). This solution consists of monthly spherical harmonic coefficients (SHC) complete to degree and order 60. This truncation represents low pass filtering in the spatial domain, causing a limited spatial resolution of GRACE data due to signal dampening. Consequently, the above processing step causes GRACE signal to represent 666 km (half-wavelength) resolution on the ground. The post-processing involves standard steps such as replacing the degree d the zonal degree 2 coefficients from satellite laser ranging solutions, correcting for Glacial Isostatic Adjustment (GIA) process using a forward model, destriping using the Swenson method (Swenson and Wahr, 2006), and smoothing using a Gaussian filter of 300 km half-radius. Further signal leakage correction is performed by the iterative forward modeling approach (Chen et al., 2014). More detailed descriptions for GRACE post-processing are available in supplementary section 1. We finally obtained monthly TWS anomaly grids oversampled at 0.25° resolution.

We obtained the monthly soil moisture from the GLDAS Noah Land Surface Model L4 monthly 0.25° x 0.25° V2.1 (GLDAS_NOAH025_M) [accessed October 2020]. We compute soil moisture anomaly (SMA) by removing the mean soil moisture over the study period. We further computed TWSA-SMA, which provides useful information on spatio-temporal groundwater storage variations continuously over the study period covering the whole Central Valley. However, it is with the coarsest resolution of 0.25° amongst the predictor variables.

3.1.3. Saturated hydraulic conductivity (K)

Saturated hydraulic conductivity data is available at 1 km spatial resolution (Zhang et al., 2019). To our knowledge, this is the only publicly available global dataset at such fine resolution.

3.1.4. Texture

Faunt et al., (2009) compiled texture data from the lithological drill holes, which range in depth from 12 to 1200 feet below the ground level. Faunt et al., (2009) used this textural data to simulate the geological model for Central Valley Hydrologic Model.

3.1.5. Percent Slope

Percent slope is derived from the National Elevation Dataset (NED) at 1/3 arc second (~10 m) resolution.

3.1.6. Groundwater level

The target variable against which we train for our machine learning model is the in-situ groundwater level obtained from the California Department of Water Resources (DWR) California Statewide Groundwater Elevation Monitoring (CASGEM) database (DWR CASGEM, 2021 a, b) and the United States Geological Survey (http://water.usgs.gov/ogw/data.html). Though Central Valley consists of ~10,000 wells, we chose 586 wells for the entire Central Valley with good spatio-temporal coverage over our study period. We only chose a well if it has at least biannual measurement or continuous measurement over a shorter time scale within our study period (Figure 1).

3.1.7. Deformation data

The vertical deformation data from GPS and CryoSat-2 (CS2) radar altimeter was not used for ML model development but rather as independent data to validate our modeled groundwater level results. GPS data is available from https://sideshow.jpl.nasa.gov/pub/JPL_GPS_Timeseries/repro2018a/post/point/, NASA Jet Propulsion Laboratory (JPL), California Institute of Technology. We also use the CS2 low-resolution mode (LRM) radar altimetry data sensing solid Earth deformation time series in an innovative method applied to Central Valley (Yang, 2020). CS2 data was waveform retracked and spatially interpolated to obtain the 2-D vertical deformation maps for the southern San Joaquin Valley (Figure S1).

Finally, to overcome the problem of mismatch in the spatial resolution of various input and target variables, all the inputs except the TWSA-SMA are interpolated at the ground well locations using the 'scatteredInterpolant' function in MATLAB and all input variables were aggregated to monthly sampling. The variable TWSA-SMA is used without further interpolations or resampling at the 0.25° spacing interval.

3.2. Machine Learning Modeling

3.2.1. Random Forest

Random Forest is a robust model which has shown the capability to produce highly accurate results for several geological applications, e.g., Hengl et al., (2018) and Tyralis et al., (2019).

Random Forest is an ensemble of decision trees (DTs) consisting of decision-making units known as nodes arranged in the form of a tree. Each DT is trained by passing data down from the node at the top (root node) to the leaf node (the node at which splitting stops), each splitting of a node result in two child nodes. Out of all the input variables at a node, the one chosen for splitting should be such that the child nodes are "purer", i.e., homogeneous in terms of the target variable than the parents. The metric that is commonly used in regression problems is the sum of squared error (SSE) between all the observations at a particular node, and the mean of all the observations. Thus, SSE should be lower for the child nodes compared to the parent nodes for a valid split.

In the algorithm described in Breiman, (2001), the observations are randomly sampled with replacement at each DT, a process known as bagging. Approximately two-thirds of observations are used for model building in each DT and are known as "in-bag" samples. The remaining one-third of the samples are called "out of bag" (OOB) samples used for internal validation by the RF model. Each DT has a different combination of in-bag and OOB data, and by combining predictions on OOB data from each DT, we can get a secondary validation of whether our RF model is over-fitted. Randomness in an RF is further increased by only selecting a few input variables for each DT, reducing the correlation between individual DTs and preventing overfitting.

3.2.2. Development of model

During the development of any machine learning model, a small portion of the dataset is isolated, known as the test dataset. The remaining dataset is then split into a training dataset, using which a model is built, and a validation dataset, against which the accuracy of the model is evaluated. This process of developing and fine-tuning the model on the training and validation dataset is iterative and repeated until the desired number of steps or accuracy is achieved. The predictive accuracy of the model is evaluated against the independent test dataset. Here we randomly select a test dataset spread throughout the study period and it constitutes 20% of the overall dataset. Previous studies have used 10-44% of the overall dataset as test data (Rajaee et al., 2019).

3.2.3. Cross-validation of model

Several studies (Hawkins et al., 2003; Molinaro et al., 2005) have pointed out that for smaller sample sizes, a single validation dataset does not provide an unbiased estimate of model performance. We, therefore, use the k-fold cross-validation technique wherein a training dataset is further split into multiple folds, each containing a unique combination of training and validation dataset. A separate model is then built and evaluated for each fold of the data. This way, model parameters are optimized for the entire dataset and overfitting is minimized. Finally, an ensemble of models is formed, and their predictive accuracy is quantified with the test dataset. Since k-fold cross validation makes machine learning model development slower, we use k=5 or 5-fold for model development.

3.2.4. Hyperparameter optimization

Random Forest model has several hyperparameter values which need to be initialized by the user (Biau and Scornet, 2016; Probst and Boulesteix, 2017). They include the number of decision trees, the number of samples in the leaf node, and the number of variables to consider for splitting in each decision tree. While previous studies have attempted to improve machine learning predictions by increasing the complexity of model architecture (Nourani et al., 2013; Seyoum et al., 2019; Yin et al., 2022) or by optimizing the number of input variables (Rajaee et al., 2019; Tyralis and Papacharalampous, 2019), fewer studies have implemented strategies for optimizing hyperparameters. While random search and grid search algorithms for hyperparameter optimization are time-consuming and might not lead to the best hyperparameters (Feurer and Hutter, 2019; Yin et al., 2021), we fine-tune the machine learning models by implementing a Bayesian Hyperparameter Optimization (Snoek et al., 2012). This optimization algorithm first builds a probability model of the objective function (such as RMSE) during model training using different hyperparameters. It then uses the Bayesian distribution to find the most promising hyperparameter to evaluate the true (actual) objective function.

3.2.5. Assessment of model accuracy and feature importance

The modeled results are validated against selected *in situ* groundwater level observations located in the Central Valley using statistical estimates, correlation coefficient, root mean squared error (RMSE), Nash-Sutcliffe efficiency (NSE) coefficient, and scaled RMSE (R*).

Supplementary section contains detailed information on these quantities. Correlation quantifies the interdependence between two datasets. It ranges in value from −1 to +1, which represents a perfect negative and positive relationship, respectively, while a value of 0 represents no relationship. RMSE quantifies the standard deviation of residuals of the best fit line between observed and modeled values. NSE has been used to quantify the predictive power of hydrological models (Nash and Sutcliffe, 1970) and ranges from −∞ +1. Values below 0 suggest unacceptable predictions, while above 0 are good predictions, with 1 being the perfect prediction.

We also compute the feature importance by permuting out of bag (OOB) observations (Breiman, 2001). The underlying concept of this approach is that permuting the values of the most influential predictor should lead to the most increase in modeling error.

To further understand the dependence of modeling accuracy of the model on the input variables, we use the drop-column method (Jyolsna et al., 2021; Parr et al., 2020). We consider the model developed above after Bayesian Hyperparameter Optimization using all the input variables as the base model. Models are retrained without the dropped input variables and the increase in RMSE on test data compared to the base model is noted.

3.3. Computation of inelastic storage coefficient and groundwater storage

We use the modeled monthly groundwater level variations obtained above and the vertical deformation data from GPS and CS-2 altimeter to obtain the inelastic storage coefficient S_{kv}. The formula for computing S_{kv} is mentioned in Supplementary section 3. Since GPS measures daily vertical deformation, we averaged them to monthly values when correlating them with monthly groundwater level.

For unconfined aquifers, the storage coefficient is the specific yield (S_y)—the volume of water released due to drainage from an unconfined aquifer per unit decline in groundwater level. In the Central Valley, typical values range from 0.06 to 0.3 (Faunt, 2009). We obtained the groundwater storage in terms of equivalent water height (EWH) for the whole of Central Valley by multiplying groundwater level changes with the specific yield of 0.1 for the unconfined wells (<60 m deep) (Faunt, 2009). The groundwater storage, when multiplied by the area of Central Valley (~52,000 km^2), gives the volumetric GWS estimate of Central Valley.

4. RESULTS

4.1 Overall results

Modeled results show a high accuracy for both San Joaquin and the Sacramento Valley (Figure 3a). For San Joaquin Valley, correlation, root mean square error, NSE, and R* for training (test) data are 0.99(0.97), 1.35 (2.72), 0.99 (0.95), and 0.12(0.21), respectively. For Sacramento valley, correlation, root mean square error, Nash Sutcliffe efficiency, and normalized RMSE for training (test) data is 0.99(0.95), 1.21(2.12), 0.98(0.94), and 0.14(0.26), respectively. Additional validations of modeled results with respect to the out-of-bag data are provided in Supplementary file (Figure S2, Table S1).

For the computation of feature importance in Figure 3b, we summed the contributions from different lags of precipitation, i.e., PPT0, PPT1, PPT2, PPT3 and PPT4, in terms of one variable, PPT, to make it easier for analysis (Figure S3, we show feature importance considering all lag components for precipitation). For San Joaquin Valley, texture, hydraulic conductivity, slope, GWS (TWSA-SMA), temperature and precipitation are the most important features in decreasing order. For Sacramento Valley, precipitation is the most important, followed by hydraulic conductivity and temperature. Texture, slope and groundwater storage (TWSA-SMA) show almost similar importance.

Using the drop-column method, we find that groundwater storage causes the most increase in RMSE compared to the base model for both Sacramento and San Joaquin valley (Table S2). Removal of geological factors, texture and hydraulic conductivity, along with topographic slope, also significantly increases the RMSE of the models.

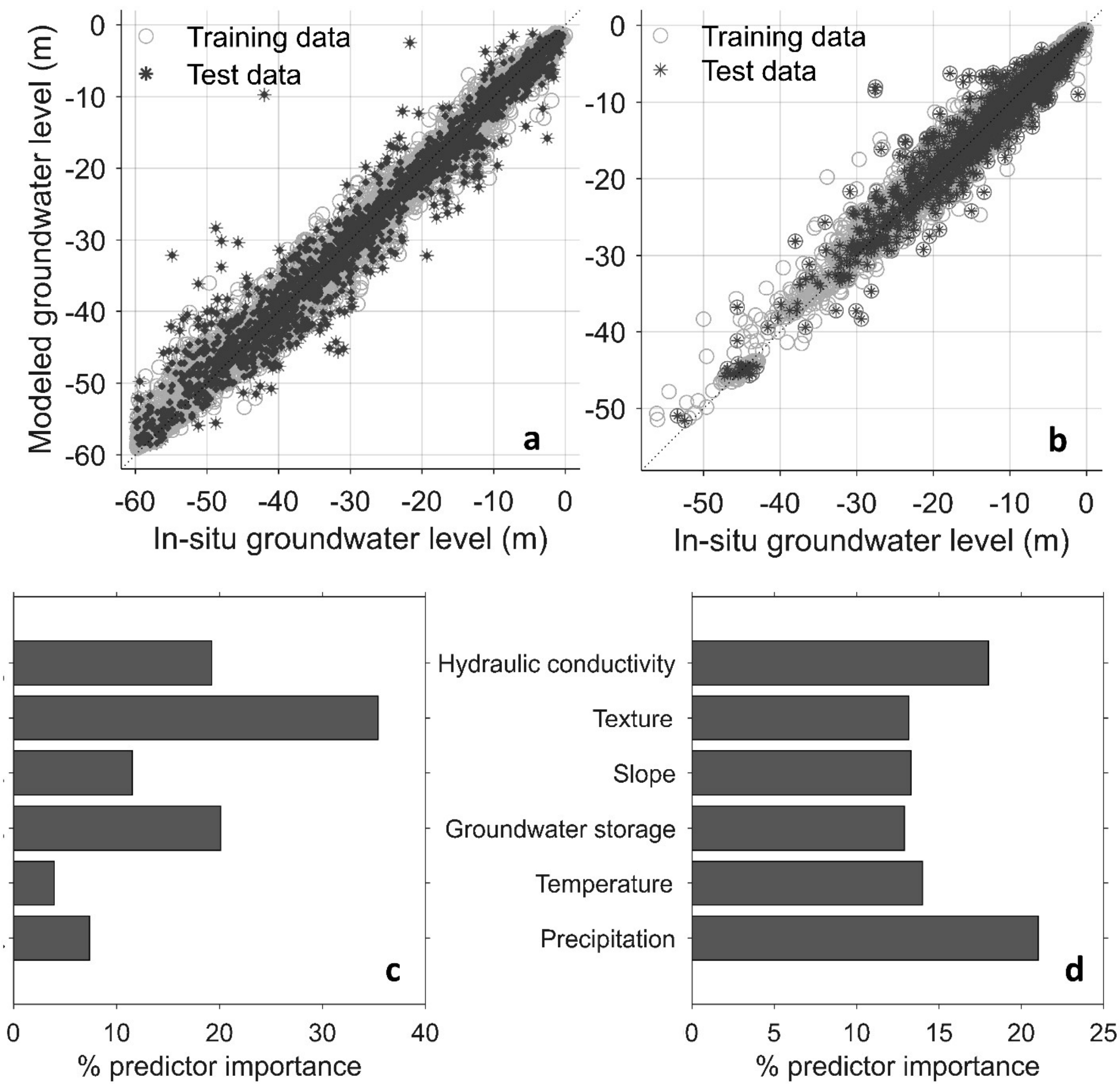


Figure 3. The accuracy assessment for machine learning modeling. Correlation plots

between the modeled results and in-situ groundwater level variations for training and test data for (a) San Joaquin and (b) Sacramento Valley respectively. Feature importance plots for (c) San Joaquin and (d) Sacramento Valley.

The modeled results also compare well with the *in-situ* groundwater level from several ground wells (Figure 1, Figure S3) as they show similar seasonality and trends, and the largest groundwater level declines can be seen during the drought periods (Figure 4, Figure S4). Some mismatches can be seen, and they indicate that the modeled results are not perfect. These reflect remaining unmitigated errors or noise in in-situ data which cannot be modeled. Wells in San Joaquin valley generally show higher declines than those in Sacramento valley. We can also effectively fill the data gaps in in-situ groundwater levels through machine learning modeling.

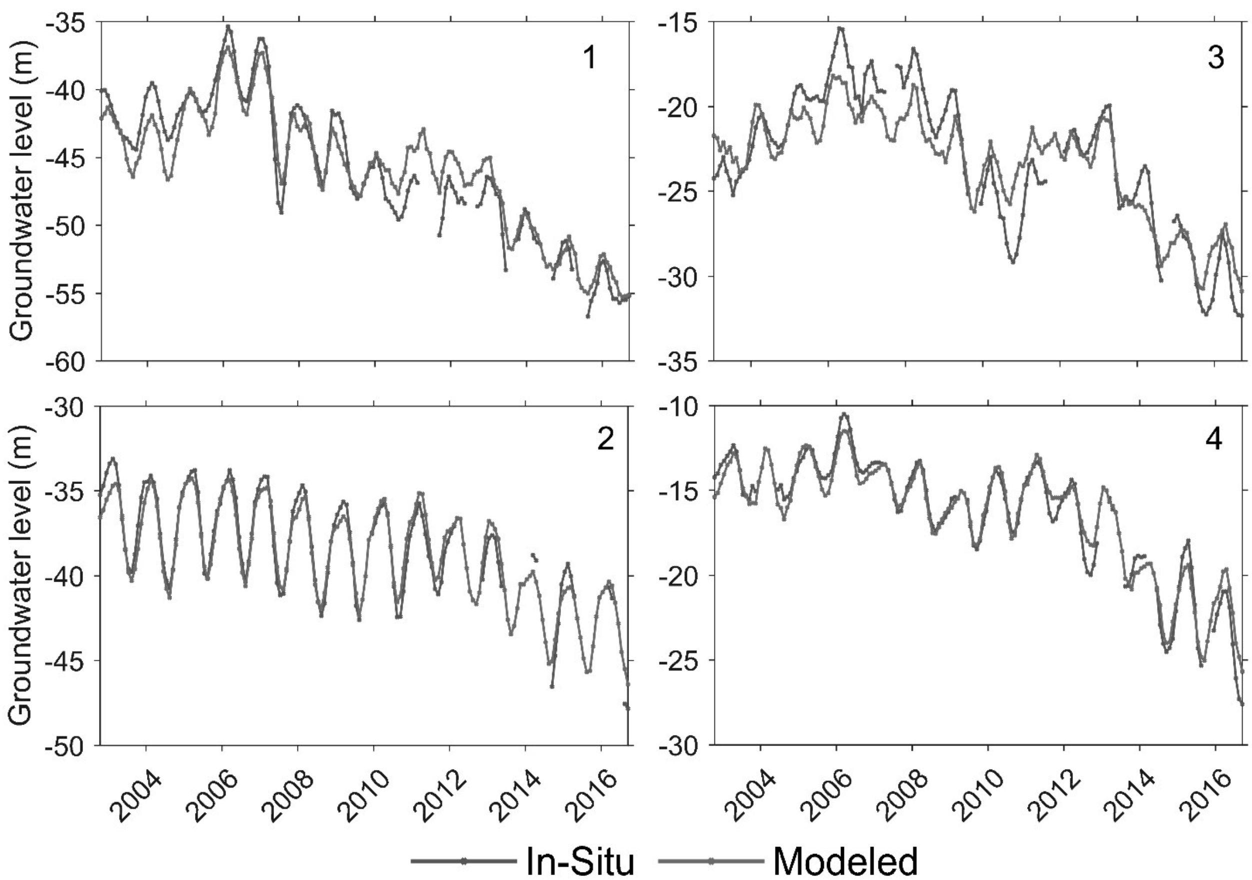


Figure 4. Modeled and in-situ groundwater level time series for wells in San Joaquin (left) and Sacramento valley (right). The location of the wells can be seen in Figure 1. Table S3 shows the statistic).

4.2 Comparison of modeled results with vertical deformation data

Inelastic storage coefficients from vertical deformation measured at GPS sites and modeled groundwater level varies from 0.15-4.02×10^{-2} for GPS sites P544 and P303, respectively (Figure 5a-b; Table 1). In addition, we find a good correlation between the long-term subsidence and the modeled groundwater level at the selected GPS locations (Table 1). Skv computed from groundwater level and CS2 varies among the subbasins. The mean S_{kv} over the subbasins is 5.89×10^{-2}.

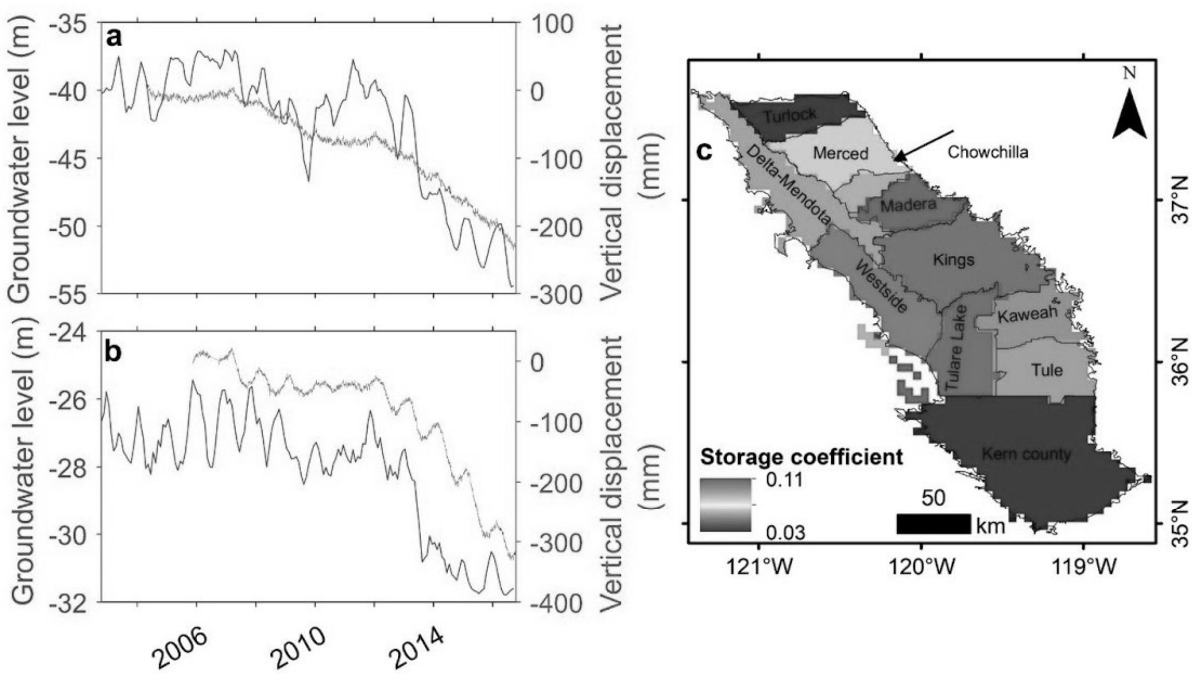


Figure 5. Computation of inelastic storage coefficient. (a) and (b) shows modeled groundwater level and vertical deformation from GPS at P304 and P545 (shown in Figure 1). (c) shows the

inelastic storage coefficients for subbasins computed from modeled groundwater level and deformation data from CS-2 altimeter.

GPS	S_{kv} (This study)	Correlation between groundwater level and deformation from GPS	S_{kv} (Ojha et al., 2019)
P303	3.46	0.90	1.87
P304	0.9	0.96	1.38
P307	1.94	0.89	1.14
P544	0.15	0.85	0.19
P565	4.02	0.91	-
P566	0.86	0.86	0.76
P545	0.42	0.94	0.33
P563	0.38	0.96	-

Table 1. Computation of Skv from modeled groundwater level and vertical deformation. Skv from Ojha et al., (2019) is shown for reference

4.3 Spatio-temporal variations of groundwater variations

Central Valley lost approximately 30 km^3 of groundwater from October 2002 - September 2016 (Figure 6a). The most rapid decline in groundwater occurs during the two drought periods, January 2007- December 2009, and October 2011 - September 2015 (Table). These periods of decline usually follow or happen during phases of low/negative annual precipitation anomalies (Figure 6b). Periods of positive annual precipitation anomalies usually are followed by periods of increase in groundwater storage.

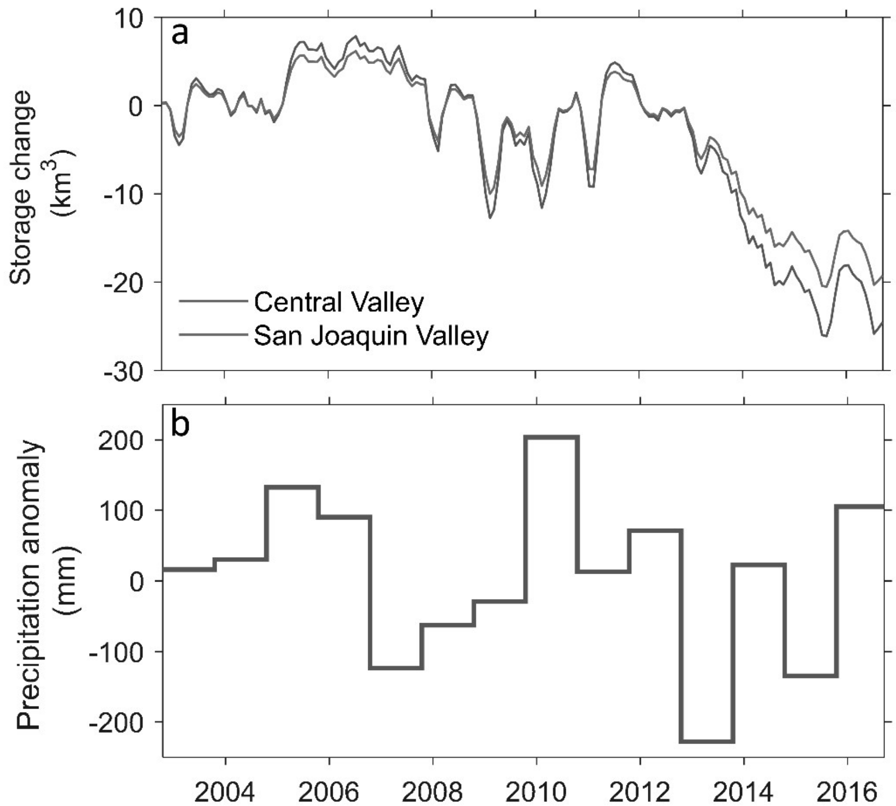


Figure 6. (a) Temporal variations of groundwater storage in Central Valley and San Joaquin Valley (shown for comparison), and (b) annual precipitation anomalies in the Central Valley.

Time period	Annual groundwater volume loss (km^3 yr^{-1})		
	This Study	**Previous Results**	**Study**
April 2006 - September 2009	-5.1 ± 1.2	-7.8 ± 0.8 -4.2 ± 0.3	(Scanlon et al., 2012) (Xiao et al., 2017)
April 2006 – March 2010	-4.2 ± 1.0	-6.0 ± 1.5 -4.1 ± 0.2	(Famiglietti et al., 2011)
January 2007 – December 2009	-5.7 ± 1.2	-7.1 ± 2.4 -5.5 ± 0.3 -6 -(3-10)	(Ojha et al., 2018) (Xiao et al., 2017) (Alam et al., 2021) (Ahamed et al., 2022)
October 2011 – September 2015	-7.6 ± 1.5 (San Joaquin Valley only) -9.8 ± 1.7	-6.1 * (San Joaquin Valley only) -7 -(6-17)	(Ojha et al., 2019) (Alam et al., 2021) (Ahamed et al., 2022)
October 2012 – September 2016	-7.7 ± 1.8	-10.0 ± 0.2	(Xiao et al., 2017)

Table 2. Comparison of GWS loss obtained from this study with previously published estimates.

GWS declines over San Joaquin valley are more prominent than declines over Sacramento valley (Figure 7). The decline during the latter drought period can be seen in wider areas and have a higher magnitude compared to the declines during the former period. Groundwater depletion can be seen mainly in Tulare Lake, Tule, and Kern subbasins, although lower groundwater depletion can be observed in Kings, Westside, and Kaweah subbasins.

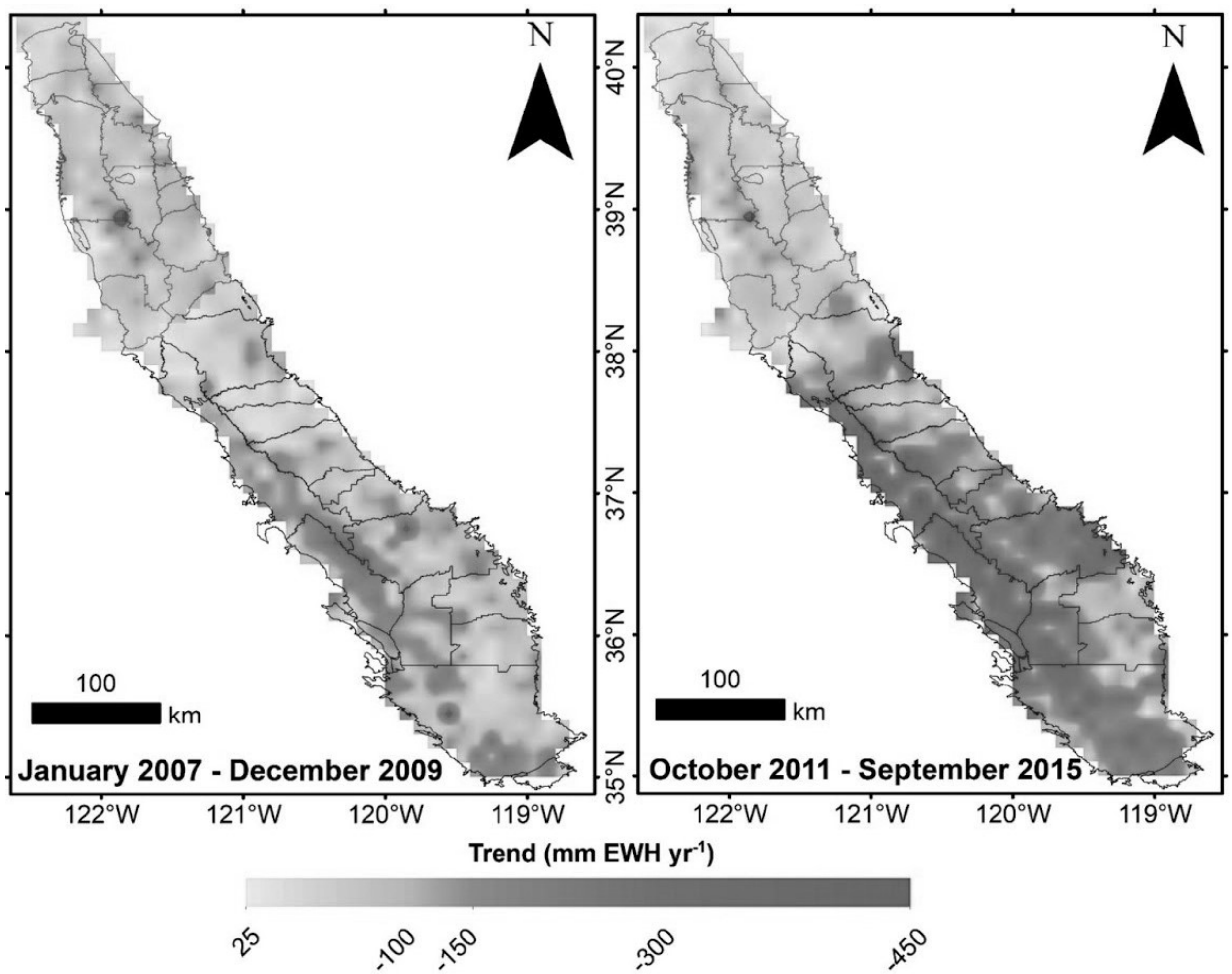


Figure 7. Spatial variations in modeled groundwater storage trends at 5-km resolution for (a) January 2007- December 2015 (b) October 2011–September 2015

Discussion

1. Choice of predictor variables in ML modeling

Several hydrological and geological datasets have been used in past studies for modeling groundwater storage variations in previous studies. These include temperature, precipitation, soil type, soil moisture, land cover, evapotranspiration, canopy water, and surface runoff (Jyolsna et al., 2021; Milewski et al., 2019; Seyoum et al., 2019; Sun et al., 2013; Yin et al., 2022). The choice of these predictors depends on the study area and type of aquifer and obviously on the availability of reliable data. For example, evapotranspiration is difficult to model reliably for heavily irrigated regions like Central Valley (Allen et al., 2011) and was not used. Surface water changes for Central Valley might be available at sufficient resolution through the Surface Water Ocean Topography (SWOT) mission to be launched after 2022. Several of the predictors chosen in this study, such as precipitation, temperature, topographic slope, and texture, have also been used in the numerical groundwater models as they hold importance for hydrological balances (Faunt, 2009; Faunt et al., 2016).

RF moel also estimated that input data used in modeling has different importance for Sacramento and San Joaquin Valley (Figure 3b), suggesting that different processes are ongoing in the two regions. Texture or percentage of coarse-grained material is an important indicator of the lithological variations in Central Valley. While Sacramento valley shows fine-grained texture as it majorly consists of sediments derived from fine-grained volcanic rocks, San Joaquin Valley shows spatial variation in texture from east to west. The eastern region near the Sierra Nevada has coarser-grained sediments, making this region a good aquifer. The western part near the Coast Ranges has a fine-grained texture, being richer in shale. San Joaquin Valley and Tulare Basin consist of alternating layers of coarse and fine material, creating a mix of confined, unconfined, and semi-confined units.

Saturated hydraulic conductivity (K) describes the ease with which water moves through the pore spaces in the soil and is considered an important quantity in groundwater modeling

(Sanchez-Villa et al., 2006). Its order of importance is second and third for Sacramento and San Joaquin valley based on permutation of OOB data. Together, hydraulic conductivity and texture provide important geological information about groundwater flow patterns in the whole Central Valley at high spatial resolutions and the removal of both predictors causes a significant increase in RMSE of models. Variations in topographic slopes lead to differences in groundwater recharge (Satapathy and Syed, 2015). Its importance is, however, lesser than other geological variables.

TWSA-SMA, though at coarse resolution, provides valuable information about the continuous spatio-temporal groundwater storage variations over the last decade and a half. However, removal of this predictor alone causes the most increase in RMSE for models built for Sacramento and San Joaquin valley. Based on OOB permutation, it is of mid-importance. The above two findings seem contradictory. However, they can be explained by the fact that this predictor has crucial information for modeling groundwater variations, though at the lowest resolution of all predictors. Therefore, the permutation of this predictor might not significantly affect the predictions, while its removal affects the modeling results.

Precipitation is the dominant source of groundwater recharge and is, therefore, most significant for determining groundwater level patterns in Sacramento Valley based on OOB permutation. In contrast, San Joaquin Valley majorly depends on other surface water sources for recharge and anthropogenic sources might influence the groundwater withdrawal significantly. Its significance is, therefore, least for San Joaquin Valley. However, based on the drop-column method, precipitation has the least impact if removed from modeling in both Sacramento and San Joaquin valley.

Temperature consists of seasonal signals, which might also help capture seasonal groundwater signals. It is the least important for San Joaquin valley based on OOB permutation. However, its removal significantly affects the accuracy of modeling in Sacramento valley.

2. Accuracy of machine learning results

Our study achieved high accuracy for both cross-validated and test data in Sacramento and San Joaquin valleys (Figure 3a). Therefore, we minimized the overfitting, which reduces the confidence of machine learning results (Roelofs, 2018). Overall results indicate excellent NSE coefficients at 0.94–0.97 for validation data, revealing superior model predictive capability. These results are similar to Agarwal (2020), that used only 180 wells for modeling in Central Valley using the Random Forest model. As our accuracy estimates are similar to Agarwal, (2020), we can conclude that Random Forest can accommodate additional data without sacrificing accuracy.

An earlier study in southern San Joaquin Valley by Miro and Famiglietti, (2018) also used ANN and therefore we compared similarities and differences between this study, Agarwal, (2020) and Miro and Famiglietti, (2018). Miro and Famiglietti (2018) obtained validation NSE ranging from 0.039 to 0.751 when modeling annual groundwater storage variations in southern San Joaquin valley using ANN. We obtained a better validation NSE of 0.95 for San Joaquin Valley when modeling monthly groundwater variations using random forest. Even Agarwal (2020) obtained a validation NSE of 0.86 using ANN. This is despite the fact that our study used similar predictors such as precipitation, temperature, and topographic slope from the same source as Miro and Famiglietti, (2018). We have processed GRACE L2 data along with leakage correction, while Miro and Famiglietti, (2018) used GRACE L3 monthly mass grids. A possible reason for the lower accuracy in their study might be because they model groundwater storage for each year, leaving less spatio-temporal data for modeling groundwater storage. Miro and Famiglietti (2018) use kriging to interpolate groundwater level changes for each year, a process that might lead to further errors (Deutsch, 2003; Sun et al., 2009). Since these kriged groundwater levels were used for training the model, kriging errors can further propagate in the modeled groundwater storage variations. The choice of geological variables like texture and/or hydraulic conductivity might have further improved the accuracy. Further, since Random Forest is less prone to overfitting compared to ANN (Agarwal, 2020), Miro and Famiglietti, (2018) could have considered random forest for comparison with their results.

We have directly used groundwater level as the output variable in ML modeling and then used the modeled results to compute groundwater storage. Several studies in the past have focused on groundwater level modeling and forecasting using fuzzy logic, ANN, Support Vector Machine, and other computational algorithms (see a comprehensive review by Rajaee et al., (2019) and references therein). Nonetheless, several studies have been conducted lately on modeling groundwater storage using RF and other algorithms.

Jyolsna et al. (2021) obtained a correlation of 0.50-0.83 when modeling TWSA in different Indian aquifers using RF, while Seyoum and Milewski, (2017) found a correlation of 0.86 when modeling TWS in Northern High Plains using ANN. Koch et al., (2019) obtained a correlation of 0.78 when modeling depth to shallow water table for aquifers in Denmark using RF. Seyoum et al., (2019) obtained an NSE value of 0.45 when downscaling groundwater level anomalies for glacial aquifers in Illinois using a two-stage boosted regression tree. Rahaman et al., (2020) downscaled GRACE-derived groundwater storage variations in Northern High Plains to 0.25° resolution and produced NSE in the range of 0.5-0.8 using RF. Although our study and past studies have used different output variables related to groundwater in machine learning modeling, better validation accuracy achieved in this study might also be attributed to improvements in model choice and model development as well as the choice of input predictors.

The modeled groundwater level fits closely to the in-situ data from individual ground wells. They can capture the long-term decline in groundwater, accelerated depletion during the two drought periods, recovery during the wet years, and seasonal variations, which are essential for groundwater modeling in Central Valley (Ahamed et al., 2022).

3. Spatio-temporal variations in groundwater storage

The groundwater storage losses or drought 1 range from 27 km^3 (Scanlon et al., 2012) to 29 km^3 (Ahamed et al., 2022), while losses for drought 2 are 71 km^3 (Ahamed et al., 2022). Groundwater storage losses from the ensemble water balance method (Ahamed et al., 2022) range from 8 to 31 km^3 for drought 1 and vary from 22 to 67 km^3 for drought 2. Variations in the estimate are due to different combinations of remote sensing, in-situ, and model data used in the water balance approach. Xiao et al., (2017) estimated groundwater storage loss of 16.5 km^3 and 40.0 km^3 during drought 1 and 2, respectively, using the water balance approach, which also matched with the estimates from GRACE in their study. Ojha et al., (2018) estimated groundwater loss of 21.32 ± 7.2 km^3 during drought 1. Ojha et al., (2019) estimated that San Joaquin valley lost 24.2 ± 9.3 km^3 lost groundwater from October 2011 to September 2015 based on GRACE data. Based on the GPS vertical deformation data, groundwater loss was 29.25 ± 8.7 km^3 for the same region and period (Ojha et al., 2019). Groundwater storage losses for droughts 1 and 2 are 17.1 ± 3.6 and 39.2 ± 5.1 km^3, respectively, from this study which lie within the range of previous estimates.

There is significant variability of storage losses for similar time periods using similar approaches. The causes of the variations include different methods and datasets along with their errors. Scanlon et al., (2012) used a distributed specific yield ranging from 0.05-0.3 in their study to estimate groundwater storage variations from in-situ groundwater levels (Faunt, 2009). Since regions with high groundwater level declines in southern San Joaquin valley have higher specific yields, it might be one of the reasons for higher groundwater storage estimated by their study. Water balance approach also has errors related to input variables, such as evapotranspiration which was identified as the most uncertain variable (Xiao et al., 2017; Ahamed et al., 2022). Estimates of regional groundwater storage from in-situ groundwater level data will require significant spatio-temporal interpolation due to issues with coverage in many regions (Figure 1). GRACE-derived TWS is also affected by several errors during data processing, which might also impact our machine learning model.

4. Comparison with vertical deformation data

Several past studies have combined groundwater levels from in-situ wells with geodetic observations from GPS and InSAR to obtain inelastic storage coefficient. Calculated inelastic storage coefficients for individual subbasins in southern San Joaquin valley from this study is comparable to past studies (Ojha et al., 2018). Ojha et al., (2018) computed S_{kv} of 4.08 x 10^{-2} for the whole of Central Valley, with San Joaquin having a higher S_{kv}. Ojha et al., (2019) computed a mean value of as 2.3 × 10^{-2}, while Smith et al., (2017) reported a computed mean with the range of 2.3 × 10^{-2} - 11.0 × 10^{-2} using estimates of aquifer compaction modeling for the San Joaquin Valley. These estimates compare to 5.8 x 10^{-2} from our study.

At GPS sites, P304 and P545, vertical deformation can be seen mostly in times of drought with the groundwater level dropping. Between drought periods, the groundwater level was rising due to the availability of surface water; hence, little deformation occurred. Further, at the well site near P304, the lowest water level was recorded in 1992 at 45 m below the land surface (Faunt egt al., 2016). At the end of drought1, and most of the drought 2, modeled groundwater level at the site of P304 was below the previous lowest level (pre-consolidation stress level). The correlation between subsidence and long-term groundwater levels suggests that groundwater overdraft was the cause of the subsidence (Liu et al., 2019). Further analysis could be done with long-term modeled groundwater level data and vertical deformation data for other sites to understand the aquifer compaction. Higher groundwater depletion can be combined with geological models to study sites that might be further vulnerable to subsidence.

Significant groundwater depletion can be seen for subbasins in the Tulare basin and western San Joaquin valley for both the droughts. These regions have also been subjected to subsidence (Faunt et al., 2016, Sneed et al., 2013; Farr et al., 2015). It is an expected consequence because this region requires water for *intensive* irrigation and drinking water needs. Due to climate extremes such as droughts, surface water has dwindled over the years. Consequently, groundwater from the deeper confined aquifers is usually extracted and the overlying aquitard belonging to the Corcoran clay layer undergoes compaction. Due to the continued groundwater losses in this region exacerbated during droughts, irreversible compaction of the clay layers results in subsidence signals and might reflect the permanent loss in groundwater (Smith et al., 2017; Vasco et al., 2019).

It is important to note that the groundwater storage variations reflect the balance between groundwater recharge and abstractions in an area or region and directly reflect groundwater depletion. Magnitude and rate of subsidence, on the other hand, might also depend on the hydraulic and mechanical properties of the aquifer along with the past stress regime in the region. Our results are, therefore, an important contribution to the study of localized groundwater variations in the Central Valley for the study period longer than one and a half decades.

5. Future perspectives

The approach presented here has some limitations. The Random forests, like most of the other machine learning models, cannot predict reliably outside the training range (Hengl et al., 2019). We might not be able to use the model developed here in a new region with different geology and groundwater conditions. We will have to build a new model which can be applied to a specific region. However, this limitation is also applicable to numerical groundwater models or the water balance approach. We propose building deep learning networks incorporating larger datasets and wider regions to model more complex variations in the future.

Further, unlike the water balance method of Ahamed et al., (2022), which predicted groundwater storage variations for 2002-2020 in Central Valley, our method is currently limited by the temporal coverage of the GRACE. The GRACE mission operated from 2002-2017, followed by a gap of 1 year, after which GRACE-FO was launched. Several studies have filled the data gap using deep learning (e.g., Uz et al., 2021), and modeled GRACE data from such studies can be used to extend the study for a longer time.

Conclusions

This study advances the application of remote sensing data in the field of hydrological sciences by demonstrating an effective and improved downscaling of GRACE-estimated groundwater storage variations in Central Valley to a spatial resolution of 5 km using Random Forest ML approach and other hydrologic, meteorologic, and geologic datasets. We applied it in the Central Valley region, which has developed an ever-increasing groundwater demand for irrigation given the lack of surface water supplies within most parts and has also been impacted by two droughts during our study period. Making the information about local-scale groundwater variations across Central Valley will be crucial to help twitch the groundwater management as per the plans of SGMA.

We obtained good modeling accuracy for San Joaquin and Sacramento Valley, proving that Random Forest is a robust machine learning model for such applications. We obtained similar or better prediction accuracy than other studies implementing machine learning to quantify groundwater storage variations, possibly because of the choice of predictors, choice and development of machine learning models. Development of better models, including deep learning, can further improve modeling. However, the Random Forest model developed here is suited for studies wherein predictor importance is required.

We also suggest new approaches for validating machine learning modeled results by comparing long-term modeled groundwater level changes with vertical deformation from GPS and CS-2 altimeter. The produced inelastic storage coefficient is an important aquifer mechanical reflecting deformation caused due to groundwater withdrawal. Since 2014, Sentinel-1 can provide information about continuous vertical deformation using Interferometric Synthetic Aperture Radar (InSAR) technique. Using a similar approach as in this study, new information about the aquifer dynamics using Sentinel-1, GRACE-FO, and in-situ groundwater level data can be generated.

Central Valley exhibits groundwater loss of ~ 30 km^3 during October 2002 - September 2016; however, there are periods of depletion and recharge during or followed by precipitation. Maximum amount of groundwater depletion occurs during the drought of January 2007-December 2009 and October 2011-September 2015, with rates of -5.7 ± 1.2 and -9.8 ± 1.7 $km^3 yr^{-1}$, respectively. We produced groundwater depletion maps at 5 km resolution for these drought periods that can identify groundwater overdraft areas. These areas have also exhibited land subsidence.

We conclude that the resulting modeled time series of groundwater storage variations at 5 km resolution over a decade and a half time period is effective for practical groundwater resources management.

‘References

Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. Journal of Hydrology 407, 28–40. https://doi.org/10.1016/J.JHYDROL.2011.06.013

Agarwal, V. (2021). Machine Learning Applications for Downscaling Groundwater Storage Changes Integrating Satellite Gravimetry and Other Observations [Doctoral dissertation, Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1609914912574523

A, G., Wahr, J., Zhong, S., 2013. Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: an application to Glacial Isostatic Adjustment in Antarctica and Canada. Geophysical Journal International 192, 557–572. https://doi.org/10.1093/gji/ggs030

Ahamed, A., Knight, R., Alam, S., Pauloo, R., Melton, F., 2022. Assessing the utility of remote sensing data to accurately estimate changes in groundwater storage. Science of The Total Environment 807, 150635. https://doi.org/10.1016/J.SCITOTENV.2021.150635

Alam, S., Gebremichael, M., Ban, Z., Scanlon, B.R., Senay, G., Lettenmaier, D.P., 2021b. Post-Drought Groundwater Storage Recovery in California’s Central Valley. Water Resources Research 57, e2021WR030352. https://doi.org/10.1029/2021WR030352

Alam, S., Gebremichael, M., Li, R., Dozier, J., Lettenmaier, D.P., 2020. Can Managed Aquifer Recharge Mitigate the Groundwater Overdraft in California's Central Valley? Water Resources Research 56, e2020WR027244. https://doi.org/10.1029/2020WR027244

Alley, W.M., Konikow, L.F., 2015. Bringing GRACE Down to Earth. Groundwater 53, 826–829. https://doi.org/10.1111/gwat.12379

Bertoldi, G.L., 1989. Ground-water resources of the Central Valley of California.

Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25, 197–227. https://doi.org/10.1007/S11749-016-0481-7/FIGURES/4

Bierkens, M.F.P., Wada, Y., 2019. Non-renewable groundwater use and groundwater depletion: a review. Environmental Research Letters 14, 063002. https://doi.org/10.1088/1748-9326/AB1A5F

Breiman, L., 2001. Random forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Butler, J.J., Stotler, R.L., Whittemore, D.O., Reboulet, E.C., 2013. Interpretation of Water Level Changes in the High Plains Aquifer in Western Kansas. Groundwater 51, 180–190. https://doi.org/10.1111/J.1745-6584.2012.00988.X

Castellazzi, P., Martel, R., Galloway, D.L., Longuevergne, L., Rivera, A., 2016. Assessing groundwater depletion and dynamics using GRACE and InSAR: Potential and limitations. Groundwater 54, 768–780. https://doi.org/10.1111/gwat.12453

Chen, J., Li, J., Zhang, Z., Ni, S., 2014. Long-term groundwater variations in Northwest India from satellite gravity measurements. Global and Planetary Change 116, 130–138. https://doi.org/10.1016/j.gloplacha.2014.02.007

Chen, C., He, W., Zhou, H., Xue, Y., Zhu, M., 2020. A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. Scientific Reports 10. https://doi.org/10.1038/s41598-020-60698-9

Chen, L., He, Q., Liu, K., Li, J., Jing, C., 2019. Downscaling of GRACE-Derived Groundwater Storage Based on the Random Forest Model. Remote Sensing 11, 2979. https://doi.org/10.3390/RS11242979

Cheng, M.K., Ries, J.R., 2018. Monthly estimates of C20 from 5 SLR satellites based on GRACE RL06 models.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28, 2031–2064. https://doi.org/10.1002/joc.1688

Department of Water Resources California Statewide Groundwater Elevation Monitoring (DWR CASGEM), 2021b. Periodic groundwater level measurements. [Data file]. https://data.cnra.ca.gov/dataset/periodic-groundwater-level-measurements.

Department of Water Resources California Statewide Groundwater Elevation Monitoring (DWR CASGEM), 2021b. Continuous groundwater level measurements. [Data file]. https://data.cnra.ca.gov/dataset/continuous-groundwater-level-measurements.

Dramsch, J.S., 2020. 70 years of machine learning in geoscience in review. Advances in Geophysics 61, 1–55. https://doi.org/10.1016/BS.AGPH.2020.08.002

Erban, L.E., Gorelick, S.M., Zebker, H.A., 2014. LETTER • OPEN ACCESS Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam. https://doi.org/10.1088/1748-9326/9/8/084010

Famiglietti, J.S., 2014. The global groundwater crisis. Nature Climate Change 2014 4:11 4, 945–948. https://doi.org/10.1038/nclimate2425

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

Farr, T.G., Jones, C., Liu, Z., 2015. Progress Report: Subsidence in the Central Valley, California.

Farr, T.G., Liu, Z., 2014. Monitoring Subsidence Associated with Groundwater Dynamics in the Central Valley of California Using Interferometric Radar, in: Remote Sensing of the Terrestrial Water Cycle. Wiley Blackwell, pp. 397–406. https://doi.org/10.1002/9781118872086.ch24

Faunt, C.C., 2009. Groundwater Availability of the Central Valley Aquifer, California: U.S. Geological Survey Professional Paper, 1766.

Faunt, C.C., Belitz, K., Hanson, R.T., 2009. Development of a three-dimensional model of sedimentary texture in valley-fill deposits of Central Valley, California, USA. Hydrogeology Journal 2009 18:3 18, 625–649. https://doi.org/10.1007/S10040-009-0539-7

Faunt, C.C., Sneed, M., 2015. Water availability and subsidence in California's Central Valley. San Francisco Estuary and Watershed Science 13. https://doi.org/10.15447/SFEWS.2015V13ISS3ART4

Faunt, C.C., Sneed, M., Traum, J., Brandt, J.T., 2016. Water availability and land subsidence in the Central Valley, California, USA. Hydrogeology Journal 24, 675–684. https://doi.org/10.1007/s10040-015-1339-x

Feng, W., 2019. GRAMAT: a comprehensive Matlab toolbox for estimating global mass variations from GRACE satellite data. Earth Science Informatics 12, 389–404. https://doi.org/10.1007/s12145-018-0368-0

Feurer, M., Hutter, F., 2019. Hyperparameter Optimization, in: Automated Machine Learning. Springer, Cham, pp. 3–33. https://doi.org/10.1007/978-3-030-05318-5_1

Fowler, H.J., Wilby, R.L., 2007. Beyond the downscaling comparison study. International Journal of Climatology 27, 1543–1545. https://doi.org/10.1002/joc.1616

Frappart, F., Ramillien, G., 2018. Monitoring Groundwater Storage Changes Using the Gravity Recovery and Climate Experiment (GRACE) Satellite Mission: A Review. Remote Sensing 10, 829. https://doi.org/10.3390/rs10060829

Hawkins, D.M., Basak, S.C., Mills, D., 2003. Assessing model fit by cross-validation. Journal of Chemical Information and Computer Sciences 43, 579–586. https://doi.org/10.1021/CI025626I/ASSET/IMAGES/LARGE/CI025626IF00006.JPEG

Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B.M., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ 2018. https://doi.org/10.7717/PEERJ.5518

Hoffmann, J., Galloway, D.L., Zebker, H.A., 2003. Inverse modeling of interbed storage parameters using land subsidence observations, Antelope Valley, California. Water Resources Research 39. https://doi.org/10.1029/2001WR001252

Jasechko, S., Perrone, D., 2020. California's Central Valley Groundwater Wells Run Dry During Recent Drought. Earth's Future 8, e2019EF001339. https://doi.org/10.1029/2019EF001339

Jekeli, C., 1981. Modifying Stokes' function to reduce the error of geoid undulation computations. Journal of Geophysical Research 86, 6985. https://doi.org/10.1029/JB086iB08p06985

Jyolsna, P.J., Kambhammettu, B.V.N.P., Gorugantula, S., 2021. Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes. Hydrological Sciences Journal 66, 874–887. https://doi.org/10.1080/02626667.2021.1896719

Koch, J., Berger, H., Henriksen, H.J., Sonnenborg, T.O., 2019. Modelling of the shallow water table at high spatial resolution using random forests. Hydrology and Earth System Sciences 23, 4603–4619. https://doi.org/10.5194/HESS-23-4603-2019

Konikow, L.F., 2015. Long-Term Groundwater Depletion in the United States. Groundwater 53, 2–9. https://doi.org/10.1111/gwat.12306

Konikow, L.F., Kendy, E., 2005. Groundwater depletion: A global problem. Hydrogeology Journal 13, 317–320. https://doi.org/10.1007/s10040-004-0411-8

Kuhn, M., Johnson, K., 2013. Over-Fitting and Model Tuning, in: Applied Predictive Modeling. Springer New York, pp. 61–92. https://doi.org/10.1007/978-1-4614-6849-3_4

Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. Water Resources Research 48, 4531. https://doi.org/10.1029/2011WR011453

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 (Basel) 9, 436. https://doi.org/10.3390/geosciences9100436

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 (Basel) 9, 436. https://doi.org/10.3390/geosciences9100436

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

Mehrnegar, N., Jones, O., Singer, M.B., Schumacher, M., Jagdhuber, T., Scanlon, B.R., Rateb, A., Forootan, E., 2021. Exploring groundwater and soil water storage changes across the CONUS at 12.5 km resolution by a Bayesian integration of GRACE data into W3RA. Science of The Total Environment 758, 143579. https://doi.org/10.1016/J.SCITOTENV.2020.143579

Miro, M., Famiglietti, J., 2018. Downscaling GRACE Remote Sensing Datasets to High-Resolution Groundwater Storage Change Maps of California's Central Valley. Remote Sensing 10, 143. https://doi.org/10.3390/rs10010143

Molinaro, A.M., Simon, R., Pfeiffer, R.M., 2005. Prediction error estimation: a comparison of resampling methods. Bioinformatics 21, 3301–3307. https://doi.org/10.1093/BIOINFORMATICS/BTI499

Nash, J.E., Sutcliffe, J. V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nourani, V., Hosseini Baghanam, A., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. Journal of Hydrology 514, 358–377. https://doi.org/10.1016/J.JHYDROL.2014.03.057

Ojha, C., Shirzaei, M., Werth, S., Argus, D.F., Farr, T.G., 2018. Sustained Groundwater Loss in California's Central Valley Exacerbated by Intense Drought Periods. Water Resources Research 54, 4449–4460. https://doi.org/10.1029/2017WR022250

Ojha, C., Werth, S., Shirzaei, M., 2019. Groundwater Loss and Aquifer System Compaction in San Joaquin Valley During 2012–2015 Drought. Journal of Geophysical Research: Solid Earth 124, 3127–3143. https://doi.org/10.1029/2018JB016083

Parr, T., Wilson, J.D., Hamrick, J., 2020. Nonparametric Feature Impact and Importance. arXiv 1–17.

Parr, T., Wilson, J.D., Hamrick, J., n.d. Nonparametric Feature Impact and Importance.

Probst, P., Boulesteix, A.L., 2017. To tune or not to tune the number of trees in random forest? Journal of Machine Learning Research 18, 1–8. https://doi.org/10.48550/arxiv.1705.05654

Rahaman, M., Thakur, B., Kalra, A., Li, R., Maheshwari, P., 2019. Estimating High-Resolution Groundwater Storage from GRACE: A Random Forest Approach. Environments 6, 63. https://doi.org/10.3390/environments6060063

Rajaee, T., Ebrahimi, H., Nourani, V., 2019. A review of the artificial intelligence methods in groundwater level modeling. Journal of Hydrology 572, 336–351. https://doi.org/10.1016/J.JHYDROL.2018.12.037

Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature 460, 999–1002. https://doi.org/10.1038/NATURE08238

Roelofs, R., 2019. Measuring Generalization and Overfitting in Machine Learning. University of California, Berkeley.

Sanchez-Vila, X., Guadagnini, A., Carrera, J., 2006. Representative hydraulic conductivities in saturated groundwater flow. Reviews of Geophysics 44, 3002. https://doi.org/10.1029/2005RG000169

Satapathy, I., Syed, T.H., 2015. Characterization of groundwater potential and artificial recharge sites in Bokaro District, Jharkhand (India), using remote sensing and GIS-based techniques. Environmental Earth Sciences 74, 4215–4232. https://doi.org/10.1007/S12665-015-4474-8/TABLES/10

Scanlon, B.R., Longuevergne, L., Long, D., 2012. Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. Water Resources Research 48. https://doi.org/10.1029/2011WR011312

Schumacher, M., Forootan, E., van Dijk, A.I.J.M., Müller Schmied, H., 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

Seyoum, W.M., Milewski, A.M., 2017. Improved methods for estimating local terrestrial water dynamics from GRACE in the Northern High Plains. Advances in Water Resources 110, 279–290. https://doi.org/10.1016/j.advwatres.2017.10.021

Siebert, S., Henrich, V., Frenken, K., Burke, J., 2010. Update of the Global Map of Irrigation Areas to version 5.

Siebert, S., Henrich, V., Frenken, K., Burke, J., n.d. Update of the Global Map of Irrigation Areas to version 5.

Smith, R., Knight, R., Fendorf, S., 2018. Overpumping leads to California groundwater arsenic threat. Nature Communications 9, 1–6. https://doi.org/10.1038/s41467-018-04475-3

Smith, R.G., Knight, R., Chen, J., Reeves, J.A., Zebker, H.A., Farr, T., Liu, Z., 2017. Estimating the permanent loss of groundwater storage in the southern San Joaquin Valley, California. Water Resources Research 53, 2133–2148. https://doi.org/10.1002/2016WR019861

Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian optimization of machine learning algorithms, in: Advances in Neural Information Processing Systems. pp. 2951–2959.

Sun, A.Y., 2013. Predicting groundwater level changes using GRACE data. Water Resources Research 49, 5900–5912. https://doi.org/10.1002/WRCR.20421

Swenson, S., Chambers, D., Wahr, J., 2008. Estimating geocenter variations from a combination of GRACE and ocean model output. Journal of Geophysical Research: Solid Earth 113, 8410. https://doi.org/10.1029/2007JB005338

Swenson, S., Wahr, J., 2006. Post-processing removal of correlated errors in GRACE data. Geophysical Research Letters 33, L08402. https://doi.org/10.1029/2005GL025285

Taylor, C.J., Alley, W.M., 2001. Ground-water-level monitoring and the importance of long-term water-level data, Circular. https://doi.org/10.3133/CIR1217

Taylor, R.G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., Konikow, L., Green, T.R., Chen, J., Taniguchi, M., Bierkens, M.F.P., Macdonald, A., Fan, Y., Maxwell, R.M., Yechieli, Y., Gurdak, J.J., Allen, D.M., Shamsudduha, M., Hiscock, K., Yeh, P.J.F., Holman, I., Treidel, H., 2012. Ground water and climate change. Nature Climate Change 2012 3:4 3, 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

Tyralis, H., Papacharalampous, G., 2017. Variable Selection in Time Series Forecasting Using Random Forests. Algorithms 2017, Vol. 10, Page 114 10, 114. https://doi.org/10.3390/A10040114

Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. Water 2019, Vol. 11, Page 910 11, 910. https://doi.org/10.3390/W11050910

USGS Professional Paper 1766: Groundwater Availability of the Central Valley Aquifer, California [WWW Document], n.d. URL https://pubs.usgs.gov/pp/1766/ (accessed 4.1.21).

Vasco, D.W., Farr, T.G., Jeanne, P., Doughty, C., Nico, P., 2019. Satellite-based monitoring of groundwater depletion in California's Central Valley. Scientific Reports 9, 1–14. https://doi.org/10.1038/s41598-019-52371-7

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. 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, 15–40. https://doi.org/10.5194/ESD-5-15-2014

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: Solid Earth 103, 30205–30229. https://doi.org/10.1029/98JB02844

Wahr, J., Swenson, S., Velicogna, I., 2006. Accuracy of GRACE mass estimates. Geophysical Research Letters 33, L06401. https://doi.org/10.1029/2005GL025305

Williamson, A.K., Prudic, D.E., Swain, L.A., 1989. Ground-water flow in the Central Valley, California: Professional Paper 1401-D.

Xiao, M., Koppa, A., Mekonnen, Z., Pagán, B.R., Zhan, S., Cao, Q., Aierken, A., Lee, H., Lettenmaier, D.P., 2017. How much groundwater did California's Central Valley lose during the 2012-2016 drought? Geophysical Research Letters 44, 4872–4879. https://doi.org/10.1002/2017GL073333

Yang, T.-Y., 2020. Satellite Altimetry Applications on Lake Ice Thickness and Land Subsidence. The Ohio State University.

Yin, W., Fan, Z., Tangdamrongsub, N., Hu, L., Zhang, M., 2021a. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE: a case study over the state of Victoria, Australia. J Hydrol 602, 126735. https://doi.org/10.1016/j.jhydrol.2021.126735

Yin, W., Fan, Z., Tangdamrongsub, N., Hu, L., Zhang, M., 2021b. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE – A case study over the state of Victoria, Australia. Journal of Hydrology 602, 126735. https://doi.org/10.1016/J.JHYDROL.2021.126735

Yin, W., Zhang, G., Liu, F., Zhang, D., Zhang, X., Chen, S., 2022. Improving the spatial resolution of GRACE-based groundwater storage estimates using a machine learning algorithm and hydrological model. Hydrogeology Journal 30, 947–963. https://doi.org/10.1007/S10040-021-02447-4/FIGURES/11

Zektser, Igor S., Everett, L.G., 2004. Groundwater resources of the world and their use, UNESCO IHP-VI series on groundwater no. 6. Paris, France: UNESCO.

Zhang, Y., Schaap, M.G., Zha, Y., 2018. A High-Resolution Global Map of Soil Hydraulic Properties Produced by a Hierarchical Parameterization of a Physically Based Water Retention Model. Water Resources Research 54, 9774–9790. https://doi.org/10.1029/2018WR023539