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

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

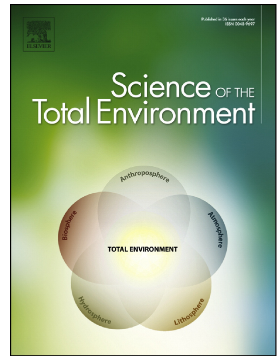
*Published in:*
Science of the Total Environment

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

*Publication date:*
2023

*Document Version*
Version created as part of publication process; publisher's layout; not normally made publicly available

Journal Pre-proof




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

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

PII: S0048-9697(22)08241-9

DOI: https://doi.org/10.1016/j.scitotenv.2022.161138

Reference: STOTEN 161138

To appear in: *Science of the Total Environment*

Received date: 6 August 2022

Revised date: 19 December 2022

Accepted date: 19 December 2022

Please cite this article as: V. Agarwal, O. Akyilmaz, C.K. Shum, et al., Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California, *Science of the Total Environment* (2022), https://doi.org/10.1016/j.scitotenv.2022.161138

# 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[4]

[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. Groundwater depletion is further exacerbated by climate-driven 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 spatial downscaling of GRACE-observed groundwater storage changes from a few hundred 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 secular components of modeled groundwater show good agreements with those of vertical displacements observed by GPS, and CryoSat-2 radar altimetry measurements and is perfectly consistent with findings from previous studies. Our estimated groundwater loss is about 30 $km^3$ from 2002 to 2016, which also agrees well with previous studies in Central Valley. We find the maximum groundwater storage loss rates of $-5.7 \pm 1.2$ $km^3$ $yr^{-1}$ and $-9.8 \pm 1.7$ $km^3$ $yr^{-1}$ occurred during the extended drought periods of January 2007-December 2009, and October 2011-September 2015, respectively while Central Valley also experienced groundwater recharges during prolonged flood episodes. The 5-km resolution Central Valley-wide groundwater storage trends reveal that groundwater depletion occurs mostly in southern San Joaquin Valley collocated with severe land subsidence due to aquifer compaction from excessive 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 consumption (Bierkens and Wada, 2019; Famiglietti, 2014). Climate change might affect the natural recharge cycle of groundwater reservoirs by altering the precipitation and evapotranspiration patterns significantly. Climate extremes such as floods and droughts might drastically increase or decrease the recharge (Taylor et al., 2012). Groundwater abstraction and outflow exceeding groundwater recharge over a long period of time and in large areas have been reported as the main causes of groundwater depletion (Konikow and Kendy, 2005; Wada et al., 2010). Groundwater depletion can lead to global water security and environmental issues, food security issues (Famiglietti, 2014; Wada et al., 2010) which could trigger mass emigration. There is an urgent need for quantifying long-term groundwater storage (GWS) changes at frequent temporal samplings that can help in better management of groundwater resources and characterize the groundwater depletion in these stressed regions.

Quantifying GWS changes 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 approaches for quantifying GWS changes have been applied in the past (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 (Faunt, 2009; Taylor and Alley, 2001). However, it is infeasible to use only these data for quantifying regional GWS changes as several aquifers have poor coverage of such wells owing to high cost of their installation and maintenance. Moreover, spatio-temporal gaps in the groundwater level data might necessitate their interpolation, which might lead to additional errors (Ahamed et al., 2022; Thomas et al., 2017). Further, uncertainties in the value of storage coefficients at well sites might translate into errors when computing GWS changes (Alam et al., 2021; Scanlon et al., 2012). Another approach to quantify GWS changes is using data from Gravity Recovery and Climate Experiment (GRACE) twin-satellite gravimetry mission. GRACE has enabled a continuous and uniform global Terrestrial Water Storage (TWS) record for the time span starting from April2002 to Oct 2017, at the "true" spatial resolution longer than 666 km (full-wavelength) and monthly sampling (Frappart and Ramillien, 2018). Innovative processing of GRACE data has enabled the uniform global quantification of GWS change 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 (e.g., 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). In Central Valley, Famiglietti et al. (2011) is the first study which used GRACE-derived TWS changes and other hydrological variables to quantify GWS changes during 2002-2011. Scanlon et al. (2012) used updated GRACE processing and *in situ* groundwater level variations to compute groundwater depletion from 2002 to 2011. The above studies estimated GWS changes by removing soil moisture estimates simulated by Land Surface Models (LSMs) from GRACE-derived TWS (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).

Vertical deformation observed during droughts from Interferometric Synthetic Aperture Radar (InSAR) has also been inverted to derive GWS changes. Recent studies have used a combination of *in situ*, satellite, and modeling data to quantify GWS changes. Alam et al. (2021) used a combination of GRACE, *in situ* wells, water balance and hydrological modeling to quantify GWS variations during 2003-2019. Ahamed et al. (2022) used remote sensing data and an ensemble of water balance methods to quantify GWS changes in Central Valley during 2002-2020. While all these studies have confirmed the continued loss of GWS along with dramatic rates of subsidence during the last two decades, all the techniques except those incorporating *in situ* groundwater levels have limited capability to model GWS changes 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. Climate extremes such as drought which have put unprecedented stress on groundwater reserves are also reflected in the groundwater fluctuations (Faunt, 2009). This necessitates the incorporation of the *in situ* groundwater level data in the groundwater models. Therefore, we propose to use Machine Learning (ML), an effective data-driven approach, to estimate GWS changes at a higher spatial resolution by downscaling GRACE-derived GWS changes to model *in situ* groundwater level variations. ML has been used for solving several non-linear complex problems in geoscience, (e.g., Berner et al. (2020); Chen et al. (2021); Dramsch (2020); Sun and Scanlon (2019)), as it does not require the knowledge of exact physical relationships between input and response variables. Further, ML methods can jointly use different types of data with different units, scales and accuracy, and is thus suitable for empirically modeling complex hydrological processes, such as basin-wide groundwater variations. Several studies in the past have incorporated ML algorithms to downscale GRACE data and produce GWS changes at high resolution for various aquifers (Chen et al., 2019; Chen et al., 2020; Miro and Famiglietti, 2018; Rahaman et al., 2020).

The primary objective of this study is to downscale GRACE-derived GWS changes in Central Valley, California, using the Random Forest ML algorithm to model and simulate monthly groundwater level and GWS changes at spatial resolution as fine as 5km. This study contrasts with Miro and Famiglietti (2018) which used Artificial Neural Networks (ANN) to model annual GWS changes in the time period 2003-2010 for a portion of San Joaquin Valley. We chose the period from October 2002 to September 2016, which covers most of the operational phase of GRACE satellite data. GRACE data beyond November 2016 was excluded to avoid errors due to the accelerometer data transplant; the accelerometer instrument onboard one of the twin satellites (GRACE-B) had thermal issues and was no longer operational until the end of mission (Bandikova et al., 2019). We use GRACE data along with hydro-meteorologic/geologic data as input and *in situ* groundwater level data as the response variable for developing the RF model. Further, the Central Valley has a record of geodetic measurements from *in situ* GPS, synthetic aperture radar interferometry, extensometers, and others, which have been used to quantify the subsidence due to groundwater overdraft (Ojha et al., 2018; Sneed and Brandt, 2015). While groundwater level change and land subsidence are two different physical processes, the subsidence measurements data can be used to qualitatively compare or validate our ML-modeled groundwater levels. We then validate the ML-modeled groundwater level using GPS vertical deformation data and basin-wide subsidence rate measured by a radar altimeter over Central Valley, CA (Yang, 2020). Here we compute inelastic storage coefficients using geodetic satellite subsidence measurements for severely subsiding regions in Central Valley for validation. This approach of combining multiple hydrological and geodetic data can further enhance our understanding of aquifer dynamics. The ultimate goal of this study is to verify the feasibility of using ML-downscaled GWS change over the whole Central Valley. We compare our results and with estimates from prior studies which can further validate the overall results. A ML approach, such as the one presented here, is hypothesized to be able to produce local-scale groundwater level storage/level information for Groundwater Sustainability Agencies to make informed management decisions under Sustainable Groundwater Management Act (SGMA).

The rest of the paper is organized as follows. The study area is introduced, data and methodology along with the details of model building and validation are described in section 2. The numerical results and comparisons with previous studies are presented in section 3. The findings of the study as well as the main limitations and the future perspectives are discussed in section 4. Finally, conclusions are drawn in section 5.



## 2. Materials and Methods

### 2.1 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 most 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. Sacramento Valley in the north also receive more precipitation than San Joaquin Valley. Consequently, supplies for irrigation in San Joaquin Valley must be met through diverted surface water sources from Sacramento Valley, 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 20th century and 20% 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 losses from GRACE satellite observations and Central Valley Hydrological Model during the first decade of the 21st century is estimated at 25-30 $km^3$ (Konikow, 2013).

As groundwater depletion continues in Central Valley and other nearby regions, California's Sustainable Groundwater Management Act (SGMA) was enacted in 2014 to promote better groundwater management, governance and thus sustainability. 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 have been visible since 1920s at the local scale.

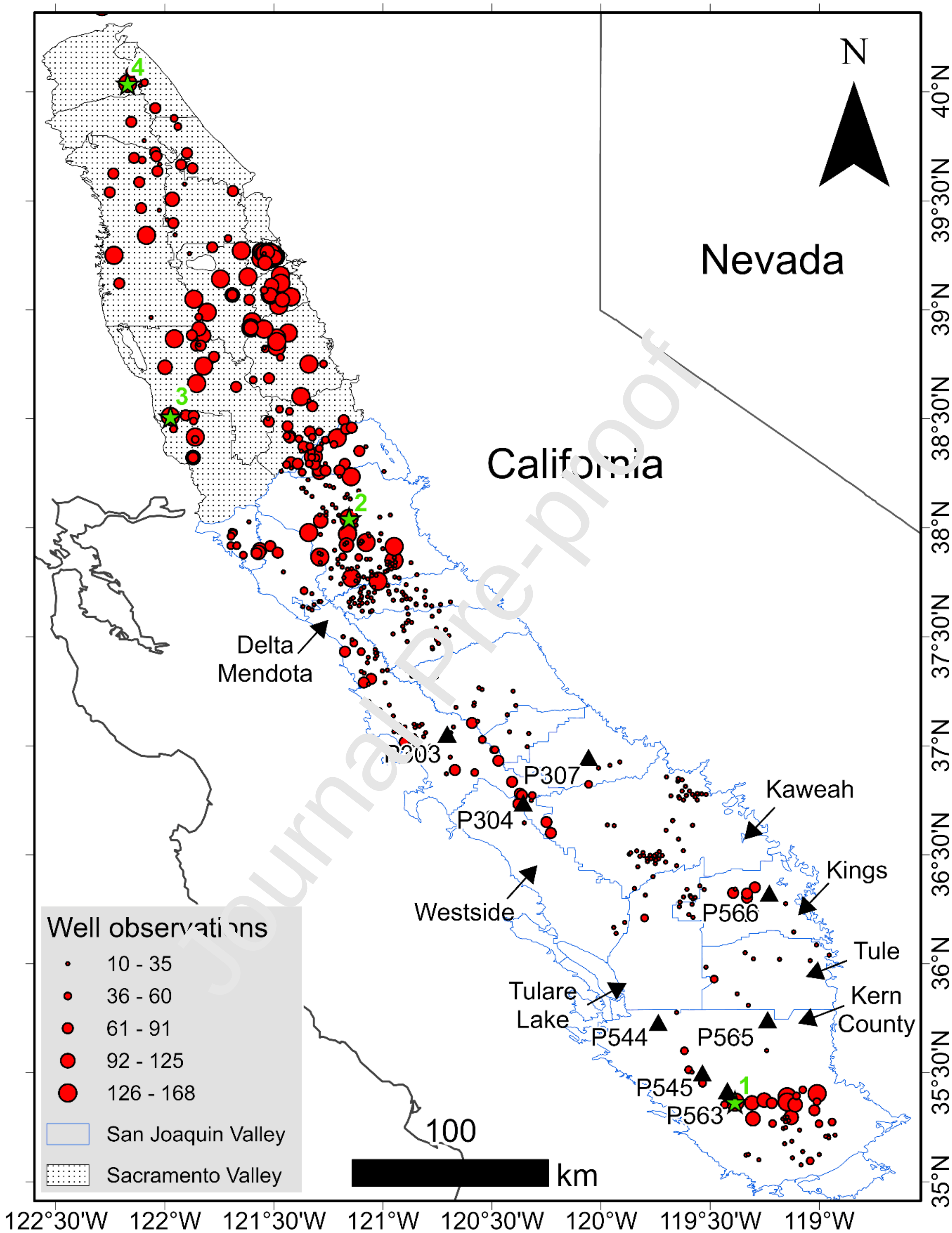



**Figure 1**. Location of Central Valley and borderlines of the two major basins, Sacramento (shaded) and San Joaquin (blue) Valley in north and south, respectively. Some of the subbasins

in southern San Joaquin Valley are marked in the figure. The location of wells used in this study and the number of measurements over the study period is also shown with solid red circles. 7 GPS sites used in this study are shown by solid black triangles. The green stars represent the wells used for validation studies and for plotting in Figure 6.

**2.2 Study design**

We adopt an empirical approach based on Random Forest Machine Learning algorithm to downscale the GRACE data. We develop two RF models, one for San Joaquin Valley, and other for Sacramento Valley. Our downscaling approach includes the following three steps (Figure 2):

(1) Modeling the groundwater level variations at the *in situ* well sites using the RF algorithm. We use GRACE-derived TWS and other hydrometeorological variables as input and monthly groundwater level data from *in situ* wells as response variables (Section 2.3) for model development and validation (Section 2.4). Once the model is trained and validated, *in situ* groundwater level data is no longer needed for steps (2) and (3)

(2) Comparing vertical deformation data with the modeled groundwater level variations from RF model created in (1) (Section 2.5).

(3) Downscale GRACE data and obtain GWS at regular 5 x 5 km grids covering Central Valley using the two RF models in (1) (Section 2.6).

Using the GRACE downscaling methodology explained above, we also propose an approach suitable for spatial downscaling of coarse resolution GRACE-derived GWSA to higher resolution GWSA for the entire Central Valley.

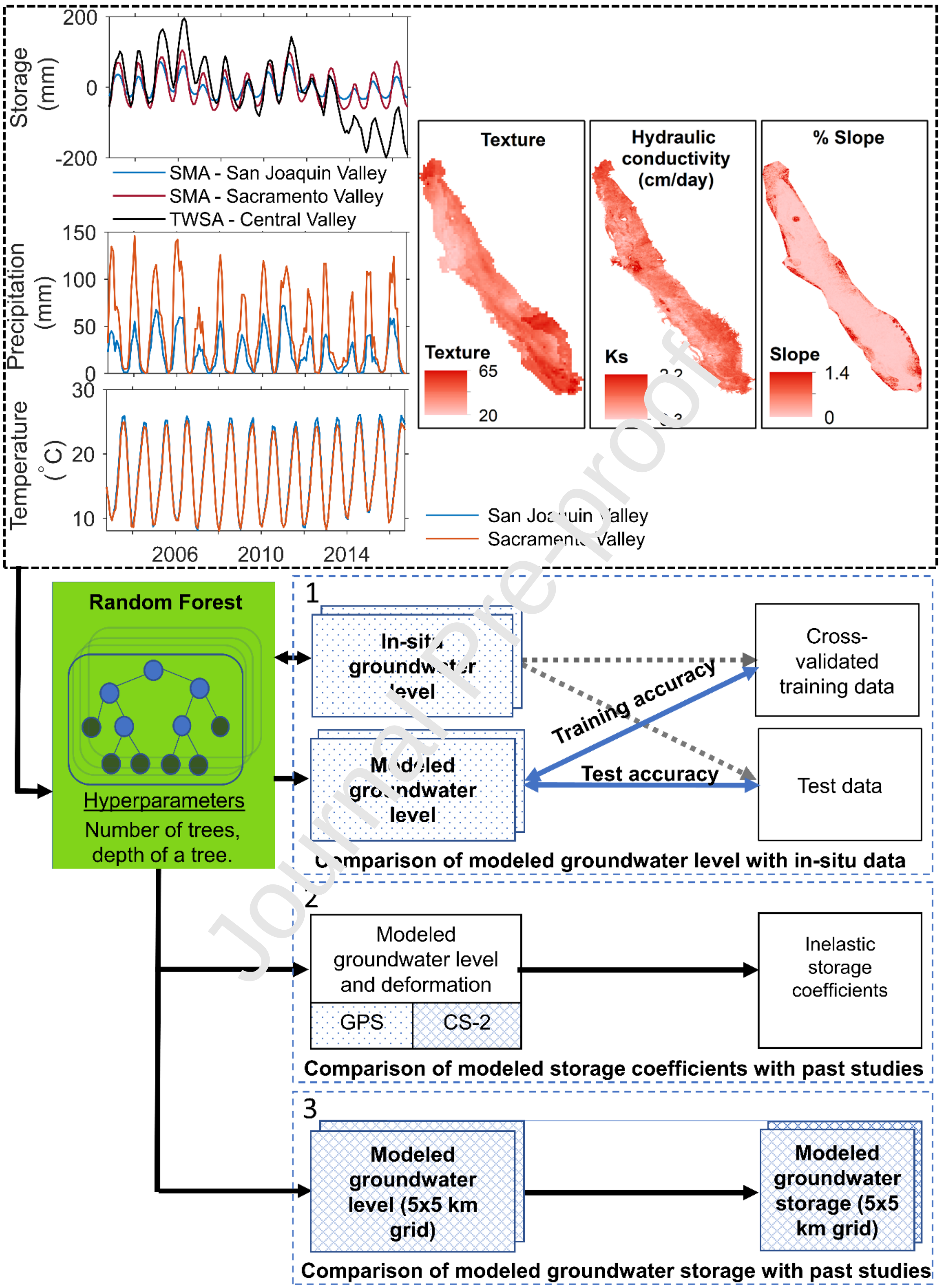

Storage (mm)
200
0
-200
SMA - San Joaquin Valley
SMA - Sacramento Valley
TWSA - Central Valley
Precipitation (mm)
150
100
50
0
Temperature (°C)
30
20
10
2006
2010
2014
Texture
Texture
65
20
Hydraulic conductivity (cm/day)
Ks
% Slope
Slope
1.4
0
San Joaquin Valley
Sacramento Valley
Random Forest
Hyperparameters
Number of trees, depth of a tree.
1
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
2
Modeled groundwater level and deformation
GPS
CS-2
Inelastic storage coefficients
Comparison of modeled storage coefficients with past studies
3
Modeled groundwater level (5x5 km grid)
Modeled groundwater storage (5x5 km grid)
Comparison of modeled groundwater storage with past studies

**Figure 2**. The workflow used for downscaling GRACE data and modeling groundwater changes using Random Forest

### 2.3 Data and Pre-Processing

| Variable name | Source | Data type | Resolution | |
|---|---|---|---|---|
| | | | **Spatial** | **Temporal** |
| **Precipitation** (PPT) | PRISM (Daly et al., 2008) | Modeled | 4 km | Monthly |
| **Temperature** (TEMP) | | | | |
| Terrestrial Water Storage Anomaly | UTCSR GRACE L2 | Remote sensing | 0.25° | Monthly |
| Soil Moisture Anomaly | GLDAS Noah LSM | Modeled | | |
| **Groundwater Storage Anomaly (GWSA)** | TWSA – SMA | | | |
| **Saturated hydraulic conductivity (K)** | (Zhang et al., 2019) | Modeled | 1 km | Static |
| **Texture (TEX)** | (Faunt et al., 2010) | *in situ* | 1 km | |
| **Percent Slope (SLP)** | National Elevation Dataset (NED) | Remote sensing | 10 m | |
| **Groundwater level (GWL)** | CASGEM/ USGS | *in situ* point data | - | Monthly |

**Table 1:** Input and response variables of the ML model

We use six hydrometeorological and geological input variables for developing the RF model (Table 1). Variables such as temperature, precipitation, soil type, soil moisture, land cover, evapotranspiration, canopy water, transmissivity and surface runoff among others are frequently used as input variables by previous downscaling studies involving GRACE data (Jyolsna et al., 2021; Milewski et al., 2019; Seyoum et al., 2019; Sun et al., 2013; Yin et al., 2022). The choice of these input variables depends on the study area and type of aquifer and obviously on the availability of reliable data.

Temperature and precipitation are important meteorological variables and affect the hydrologic cycle. PRISM (Parameter Elevation Regression on Independent Slopes Model), the source of precipitation and temperature data, simulates the spatial variations of the weather and climate using *in situ* observations. It uses a "weighted regression scheme" to account for different physiographic features and climate regimes when providing final estimates of precipitation and temperature. Precipitation is an important water source, especially for Sacramento Valley. Since precipitation can take a few months to recharge groundwater (Milewski et al., 2019), we have used lagged values of precipitation as input variables. We used temperature as a proxy for evapotranspiration due to the difficulty of modeling evapotranspiration for an irrigated region like Central Valley (Ahamed et al., 2022; Xiao et al., 2017).

For the computation of TWSA, 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 (UTCSR). This solution consists of monthly spherical harmonic coefficients (SHC) complete to degree and order 60. This truncation already represents low pass filtering in the spatial domain, resulting in the "true" GRACE spatial resolution at 666 km (full-wavelength). Monthly SHC are then post-processed to retrieve TWS changes with respect to a reference field, e.g., to the long-term mean of the monthly solutions in the study period. The post-processing involves standard steps such as replacing the zonal degree 2 coefficients from satellite laser ranging solutions (Cheng and Ries, 2018), correcting for Glacial Isostatic Adjustment (GIA) process using a forward model (A et al., 2013), destriping using the Swenson and Wahr (2006) method, and smoothing using a Gaussian filter with a half-width radius of 300 km (Jekeli, 1981). 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 (TWSA) grids oversampled at 0.25° (~25 km) resolution with respect to the mean over the study period. We also compare the TWSA obtained from this study with the TWSA from CSR Mascon (Figure S1) which also is oversampled to a resolution of 0.25°.

We obtained the monthly soil moisture storage 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 from the monthly soil moisture values. We further computed TWSA-SMA, which provides useful information on spatio-temporal GWS changes continuously over the study period covering the whole Central Valley. This input variable called GWSA has the coarsest resolution of 0.25° amongst the input variables.

Saturated hydraulic conductivity (K) describes the ease with which water moves through the pore spaces in the soil and is considered as an important quantity in groundwater modeling (Mace et al., 2000; Sanchez-Villa et al., 2006). We use the data from Zhang et al. (2019) which is the only publicly available global dataset at such fine resolution. Texture data represent the percentage of coarse-grained material. This information is computed every 15 m from lithological drill holes ranging in depth from 12 to 1200 feet below the ground level (Faunt et al., 2010). At a given site, we have used average of texture values over the range of depth for which texture information was available. Texture is an important indicator for lithological variations within 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. Besides, texture is useful to determine groundwater-flow rate as well as the magnitude and distribution of aquifer-system compaction. Topographic slope is also an important topographical parameter as its variations can also lead to differences in runoff characteristics and thus groundwater recharge (Satapathy and Syed, 2015).

The response variable against which we train our ML model is the *in situ* groundwater level obtained from the California Department of Water Resources (DWR) (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).

## 2.4 Machine Learning modeling

### 2.4.1 Downscaling method

We use Random Forest for downscaling GRACE data as it is a robust model which has shown the capability to produce highly accurate results for several geological and hydrological applications, (e.g., Hengl et al. (2018) and Tyralis et al. (2019)). Several statistical and machine learning methods, such as Multiple Linear Regression (Mukherjee et al., 2018; Sun et al., 2020), Artificial Neural Networks (Agarwal, 2021; Miro and Famiglietti, 2018), Boosted Regression Tree (Milewski et al., 2019; Seyoum et al., 2019), and Random Forest (Jyolsna et al., 2021; Rahaman et al., 2019) have been used by previous downscaling studies in different regions. In this study, we chose RF for following reasons: First, RF is a simple, straightforward model consisting of an ensemble of decision trees (DTs). It does not involve input data scaling, can handle categorical variables and missing values in the input variables in contrast to ANN.

Second, RF uses approximately two-thirds of observations for model building in each DT ("in-bag" samples), while remaining one -third ("out of bag" (OOB) samples) are 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 modeling accuracy of RF. 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. Third, user needs to optimize one fewer hyperparameter e.g., than Boosted Regression Tree. Hyperparameters are values describing the model architecture and these values are to be set by the user before running the model. For RF model, these hyperparameters 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 (Biau and Scornet, 2016; Probst and Boulesteix, 2017). For Boosted Regression Tree, shrinkage factor is the additional hyperparameter which needs to be optimized.

**2.4.2 Implementing Random Forest model**

At each of the *in situ* well site location (numbered 1,2, 3, …n in Figure 3), we consider the months for which groundwater level observation (GWL) exist (green box under GWL column in Figure 3). For GWL at month *t*, we extract value of input variable for month *t* using 'scatteredInterpolant' function with bilinear interpolation in MATLAB (yellow rows in Figure 3). For static variables, such as TEX, K, and SLP, we take the same values for all months at a given *in situ* well location. For precipitation, we extracted values for months *t*, *t*-1, *t*-2, *t*-3, and *t-4* for the output GWL at month *t*, and these input variables are labeled as PPT0, PPT1, PPT2, PPT3, and PPT4, respectively. For coarse resolution input variables such as GWSA, we did not use any interpolation while extracting the values at the GWL sites, rather the value at the 0.25° grid closest to the GWL site is adopted. Thereafter, we generated the input-output data pattern by spatio-temporal matching of the existing groundwater level observations and corresponding input variables for all *in situ* wells for all months. We developed independent RF models for San Joaquin and Sacramento valley using data for all *in situ* wells located within each region's boundary. The *in situ* GWL data also has missing values denoted by red boxes under the GWL column (Figure 3). These values were not filled in, rather after we obtain the validated model (described in section 2.4.4), we feed the input variables corresponding to all such sites and times (greyed rows in Figure 3) to the validated model to obtain modeled GWL.

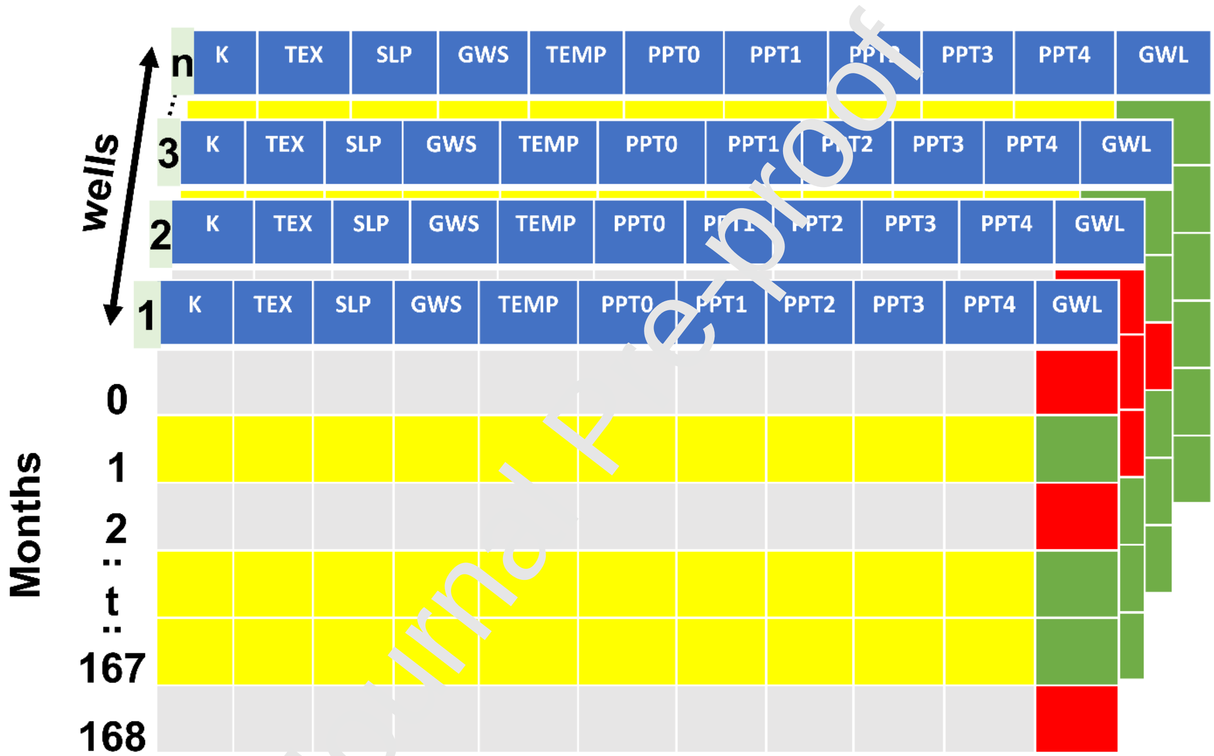



**Figure 3**: Prepared input-output data patterns for Random Forest model. We compile data for the wells (numbered *1,2, 3, …, n*) separately for San Joaquin and Sacramento Valley for all 168 months covering our study period. Here *n* represents the number of wells in the respective region. Column GWL represents the response variable, while rest of the columns represent the input variables.

During model development and training, we isolated 20% of the data as test dataset which is not used for the estimation of the model parameters. For creating this test dataset, we draw equally from all the groundwater wells at some randomly selected specific epochs distributed throughout their coverage. Remaining 80% of the data is used for the training of the model and we implemented a *k*-fold cross-validation technique on this data. In this technique, any data is further split into '*k*' folds each of which contains a unique combination of training and validation dataset. In each fold, model parameters are estimated using its own training data and the accuracy of the model is evaluated by its own validation dataset. To obtain best accuracy for validation data (in terms of lowest RMSE), we need to optimize the model hyperparameters through training process. Hyperparameter optimization is usually performed by iterative approaches and is a computationally expensive process. Commonly used methods such as random search and grid search algorithms, are time-consuming and might not lead to the best set of hyperparameters (Feurer and Hutter, 2019; Yin et al., 2021). Therefore, we fine-tune RF model by implementing a more advanced method, the (BHO) Bayesian Hyperparameter Optimization (Snoek et al., 2012). This optimization algorithm first builds a "surrogate" probability model of the RMSE and then use Bayesian methods on surrogate model to find the most promising hyperparameter on actual RMSE (Feurer and Hutter, 2019; Shahriari et al., 2016).

Applying the BHO algorithm, the model hyperparameters are optimized for the dataset and overfitting is avoided which can severely affect the accuracy on test dataset. Cross-validation is important especially for smaller sample sizes (such as the one in this study); a single validation dataset does usually not provide an unbiased estimate of model performance (Hawkins et al., 2003; Molinaro et al., 2005). The groundwater level responses from optimized model has lowest possible RMSE for test dataset.

### 2.4.3 Model validation and feature importance

***Model validation***

The responses from optimized model are validated against *in situ* groundwater level observations (both training and test data) located in the Central Valley based on commonly used statistical performance metrics such as correlation coefficient, root mean squared error (RMSE), Nash-Sutcliffe efficiency (NSE) coefficient, and scaled RMSE (R*). Supplementary section contains detailed information (including formulae) on these quantities. These metrics, when computed and compared on training and test data, ensure that ML model is not overfitted and generalizes well to the test data and other "unseen" data. Obviously, the RF model resulting with the highest correlation coefficient and NSE values and the lowest RMSE both for training and test data is chosen as the final (optimized) model.

***Feature Importance***

We 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 on the input variables, we use the drop-column method (Jyolsna et al., 2021; Milewski et al., 2019; Parr et al., 2020). We consider the model developed above after training with Bayesian Hyperparameter Optimization using all the input variables as the base model. Models are retrained after removing one input variable at each time and the increase in RMSE on test data compared to the base model is noted for the corresponding dropped/removed variable. Then the obtained increase of RMSE for each variable is normalized by dividing it to the sum of RMSE increases obtained for all input variables.

**2.5 Comparing modeled groundwater level variations with vertical deformation**

Vertical deformation data from GPS and CryoSat-2 (CS-2) radar altimeter was not used for ML model development but rather as independent data to compare against the modeled monthly groundwater level results. We obtained GPS data from https://sideshow.jpl.nasa.gov/pub/JPL_GPS_Timeseries/repro2018a/post/point/, NASA Jet Propulsion Laboratory (JPL), California Institute of Technology. Since GPS measures daily vertical deformation, we averaged them to monthly values for correlating them with monthly modeled groundwater level. We also use the solid Earth vertical deformation time series in Central Valley from CS-2 low-resolution mode (LRM) radar altimetry data generated through an innovative altimeter data processing method (Yang, 2020). CS-2 data was waveform retracked and spatially interpolated to obtain the 2-D vertical deformation maps for the southern San Joaquin Valley (Figure S2).

To obtain modeled groundwater level responses at GPS locations from the RF model, we extracted the input variables at the GPS locations for each month within the GPS data coverage. This extraction process for variables is similar to the process we described in Section 3.2.2. Similarly, we extracted the values of input variables at the grid locations of CS-2 altimeters. Hence, running our RF model with these input data we obtain the monthly groundwater level variations at the locations of GPS sites and CS-2 grids. We then combined these modeled monthly groundwater level variations with vertical deformation data from GPS and CS-2 altimeter to obtain the inelastic storage coefficient $S_{kv}$. The formula used for computing $S_{kv}$ is given in Supplementary section 3.

### 2.6 Downscaling GRACE GWSA to 5 km grid resolution

We start with generating 0.05° (~5 km) grids covering the San Joaquin and Sacramento Valley followed by extracting the value of input variables at each of these grid locations. This extraction process for variables is similar to the one mentioned in Section 2.4.2. Thereafter, we simulated groundwater level responses from the RF models at each of the grid points for San Joaquin and

Sacramento Valley, respectively. These modeled groundwater level responses cover all the 168 months of our study period at each of the grid locations. Finally, we obtained groundwater level anomalies (GWLA) at each grid location and in each month *t* (GWLA (*t*)) using the following relationship

$$\mathrm{GWLA}(t) = \mathrm{GWL}(t) - \overline{\mathrm{GWL}} \qquad (1)$$

where, GWL(*t*) is the modeled groundwater level at the grid location for the *t*-th (t = 1, 2, …, 168) month, and $\overline{\mathrm{GWL}}$ is the mean GWL over the whole study period.

We further obtained the GWS changes in terms of equivalent water height (EWH) for the whole of Central Valley by multiplying GWLA with the specific yield ($S_y$) value of 0.1 for the unconfined wells (<60 m deep) as suggested by Faunt (2009) and Miro and Famiglietti (2018). Specific yield represents the volume of water released due to drainage from an unconfined aquifer per unit decline in groundwater level. It ranges from 0.06 to 0.3 in Central Valley. The GWS in terms of EWH, when multiplied by the area of Central Valley (~52,000 $km^2$), gives the volumetric estimate of GWS changes in Central Valley.

It is further important to note here that GWSA obtained from the GRACE-derived TWSA and soil moisture anomaly from GLDAS data has been oversampled at a resolution of 0.25°. After we integrated the coarse resolution GRACE-derived GWSA in the ML model with other hydrological variables, and followed the methodology described in this section, we obtain GWS changes at 0.05 resolution. This methodology results in the spatial downscaling of regional GWS changes from GRACE to GWSA at the local scale.

## 3. Results

### 3.1 Overall results

***Validation of modeling results***

The results from RF models show high accuracy for both San Joaquin and the Sacramento Valley (Figure 4). For San Joaquin Valley, correlation coefficient, RMSE, 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. The same metrics computed over Sacramento valley for training (test) data are 0.99 (0.95), 1.21 (2.12), 0.98 (0.94), and 0.14 (0.26). Additional validations of model results with respect to the out-of-bag data are provided in Supplementary file (see Figure S3, Table S1). Boosted Regression Tree, also based on decision tree architecture, is slightly more prone to overfitting in our study, as seen by worse test accuracy for both Sacramento and San Joaquin Valley (Table S2).

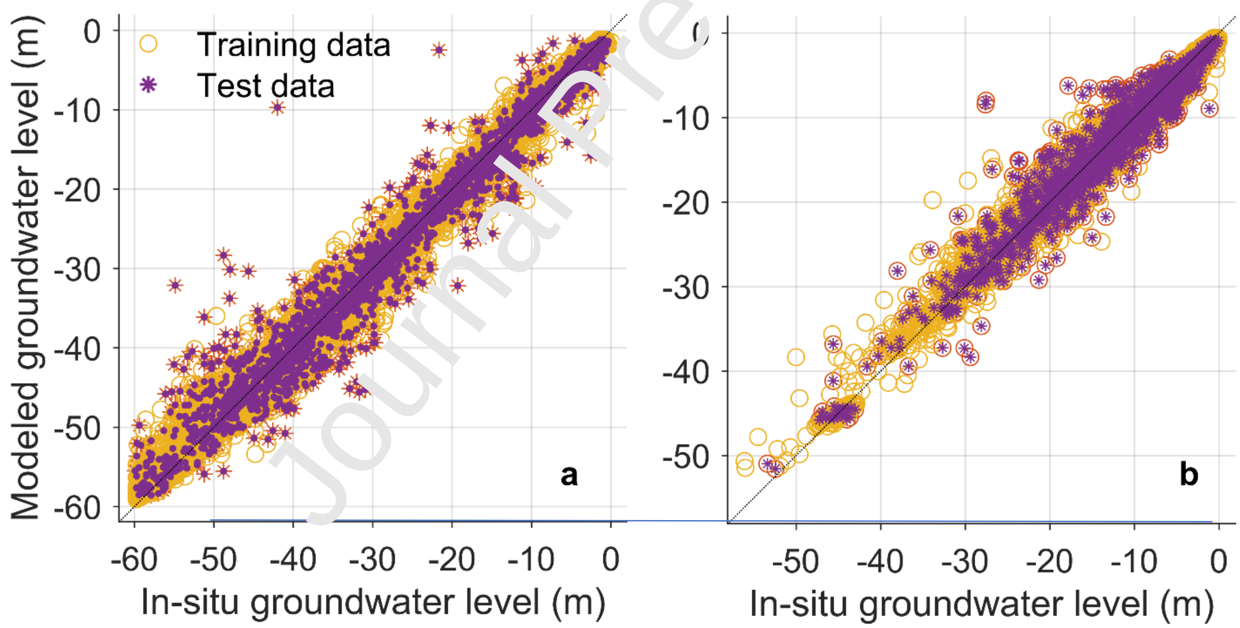



**Figure 4**. The accuracy assessment for RF machine learning modeling. Correlation plots between the modeled results and *in situ* groundwater level variations for training and test data for (a) San Joaquin and (b) Sacramento Valley, respectively.

***Evaluation of Feature Importance***

We test the usefulness of RF model in assessment of the feature importance over both Sacramento and San Joaquin valleys. Based on the permutation of OOB data, RF model gives relative importance for different input variables (Figure 5). In San Joaquin Valley, texture, hydraulic conductivity, slope, coarse resolution GRACE-derived GWSA, temperature and precipitation are the most important input features in decreasing order. In Sacramento Valley, precipitation is the most important variable, followed by hydraulic conductivity and temperature. Texture, slope, and GRACE-derived GWSA inputs show almost similar importance.

Precipitation is the primary source of groundwater recharge in Central Valley. Temperature affects evapotranspiration and groundwater extraction and will therefore affect the seasonal groundwater variations. Sacramento Valley receives high precipitation in winter months which causes groundwater recharge, while water is abstracted during summer months. This explain the higher importance of precipitation and temperature in Sacramento Valley. For San Joaquin Valley, both meteorological variables, temperature and precipitation, have relatively low importance based on OOB permutation. The valley does not receive enough precipitation and must depend on diverted surface water for its irrigation needs. Groundwater is abstracted heavily for irrigation purposes throughout the year for yearlong cropping patterns. Irrigation water is also responsible for some groundwater recharge which is hard to quantify. All these facts suggest that groundwater variations show complex seasonal variations, hence yielded low importance for temperature based on OOB permutation. Sacramento Valley, in contrast, has balanced GWS regime which suggests consistent seasonal signals and therefore much more importance of temperature is observed.

Using the drop-column method, we find that GWSA causes the most increase in RMSE compared to the base model for both Sacramento and San Joaquin valley (Table S3). Based on OOB permutation, it is of mid-importance for both San Joaquin and Sacramento Valley. The above two findings seem contradictory. However, they can be explained by the fact that this input variable has crucial information for modeling groundwater variations, though at the coarsest resolution among all input variables. Therefore, the permutation of this variable might not significantly affect the accuracy, while its removal affects the modeling results. Therefore, coarse resolution GRACE-derived GWSA is an important input for modeling and used for downscaling process.

Removal of geological factors, texture, and hydraulic conductivity, along with topographic slope, also significantly increases the RMSE of the models. Hydraulic conductivity and texture provide important 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 after their removal. These variables also show high importance when evaluated with OOB permutations. Highest importance (based on OOB permutation) of texture and hydraulic conductivity for San Joaquin Valley reflect that these variables capture the complex geology of the San Joaquin Valley.

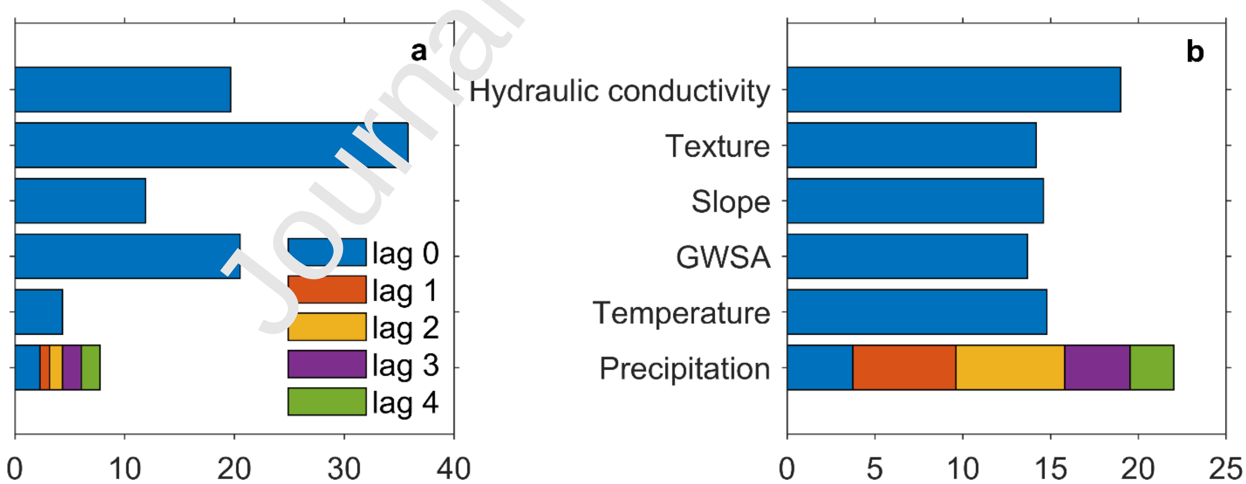



**Figure 5**. Feature importance plots based on OOB permutations for (a) San Joaquin and (b) Sacramento Valley.

***Modeling groundwater level time series***

The modeled and *in situ* groundwater level time series match tightly (Figure 1, Figure S4) as they show similar seasonality and trends, and the largest groundwater level declines can be seen during the drought periods (Figure 6, Figure S5). Some mismatches can be seen, and they indicate that the modeled results are not perfect or are not generalized too closely while avoiding overfit. Wells in San Joaquin valley generally show higher declines than those in Sacramento valley. Further, we also effectively fill the data gaps in *in situ* groundwater level time series through ML modeling.

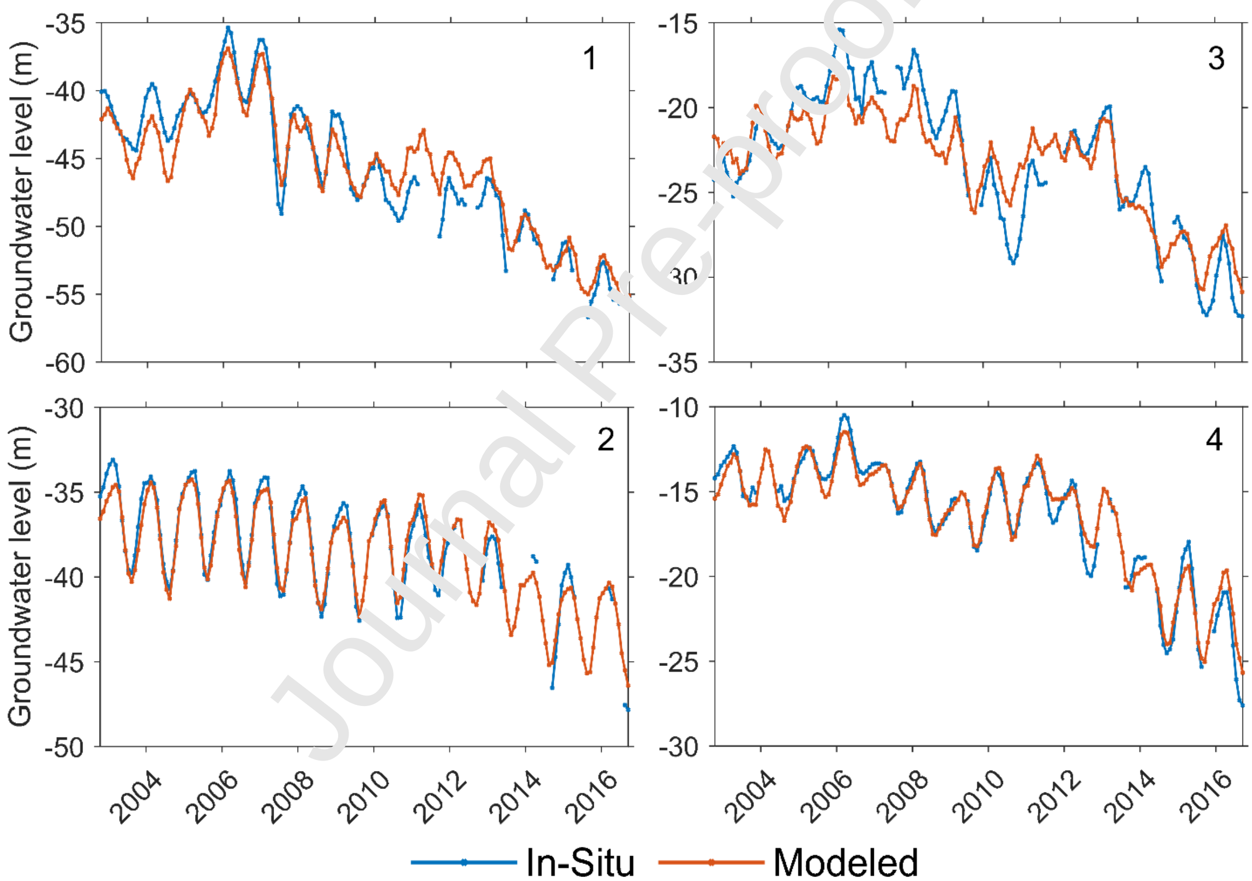



**Figure 6**. Modeled and *in situ* groundwater level time series for wells in San Joaquin (left) and Sacramento valley (right). The location of the wells can be seen in Figure 1. Table S4 shows the statistics.

**3.2 Comparison of RF modeling results with vertical deformation data**

Inelastic storage coefficients from vertical deformation measured at GPS sites and modeled groundwater level using RF varies from 0.15-4.02×$10^{-2}$ for GPS sites P544 and P303, respectively (Figure 7a-7b; Table 2). In addition, we find a good correlation between the long-term subsidence and the modeled groundwater level at the selected GPS locations (Table 2) consistent with the findings of Liu et al. (2019). $S_{kv}$ computed from groundwater level and CS-2 varies among the subbasins with a mean value of 5.89×$10^{-2}$ for the whole Central Valley (Figure 7).

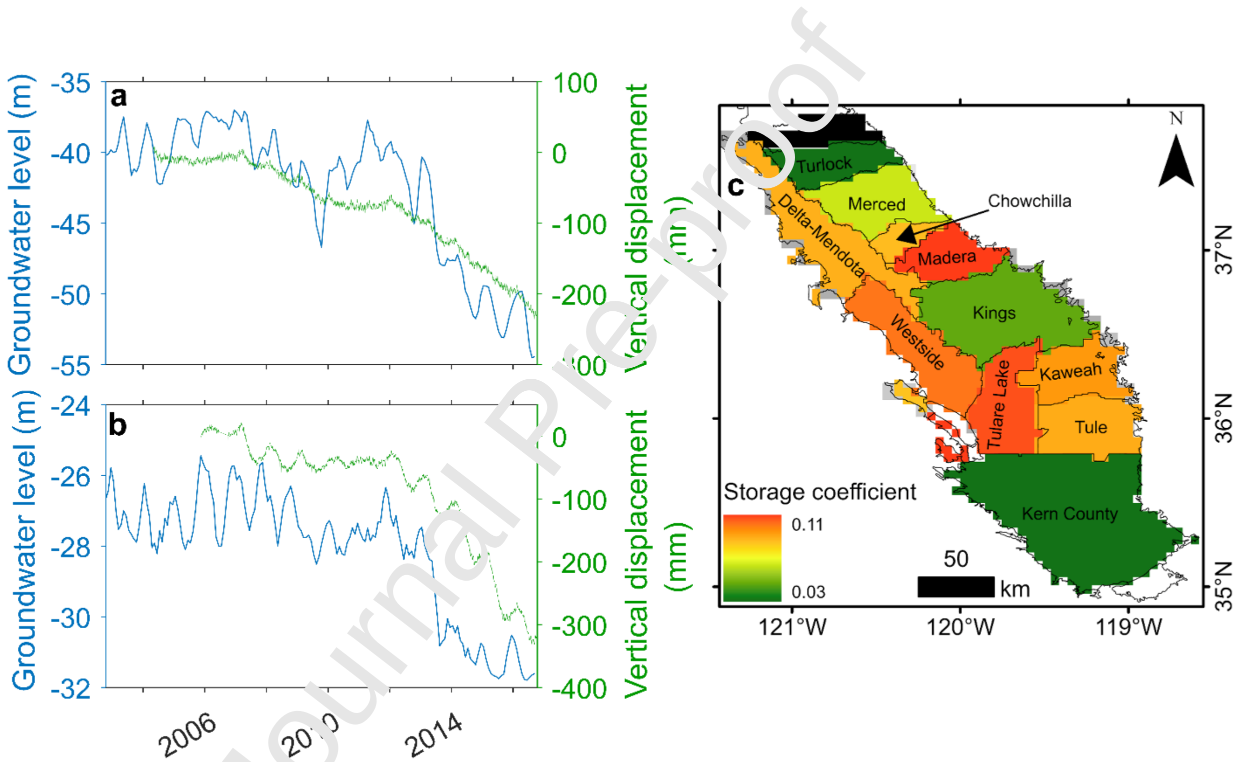



**Figure 7.** Computation of inelastic storage coefficient. (a) and (b) shows modeled groundwater level and vertical displacement from GPS at sites P304 and P545 (shown in Figure 1), respectively. (c) the inelastic storage coefficients for subbasins computed from modeled groundwater level and vertical displacement data from CS-2 altimeter.

| GPS | $S_{kv}$ | Correlation between groundwater | $S_{kv}$ |
|---|---|---|---|

| | (This study) | level and deformation from GPS | (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 2**. $S_{kv}$ computed from modeled groundwater level and vertical deformation. $S_{kv}$ from Ojha et al., (2019) is shown for reference.

**3.3 Spatio-temporal variations of groundwater storage during the two drought periods**

Central Valley lost approximately 30 km$^3$ of groundwater from October 2002 to September 2016 (Figure 8a). The most rapid decline in groundwater occurs during the two drought periods, January 2007- December 2009, and October 2011 - September 2015 (Table 3). These periods of decline usually follow or happen during phases of low/negative annual precipitation anomalies (Figure 8b). Periods of positive annual precipitation anomalies (2010-2012 and during 2016) usually are followed by periods of increase or recovery in GWS.

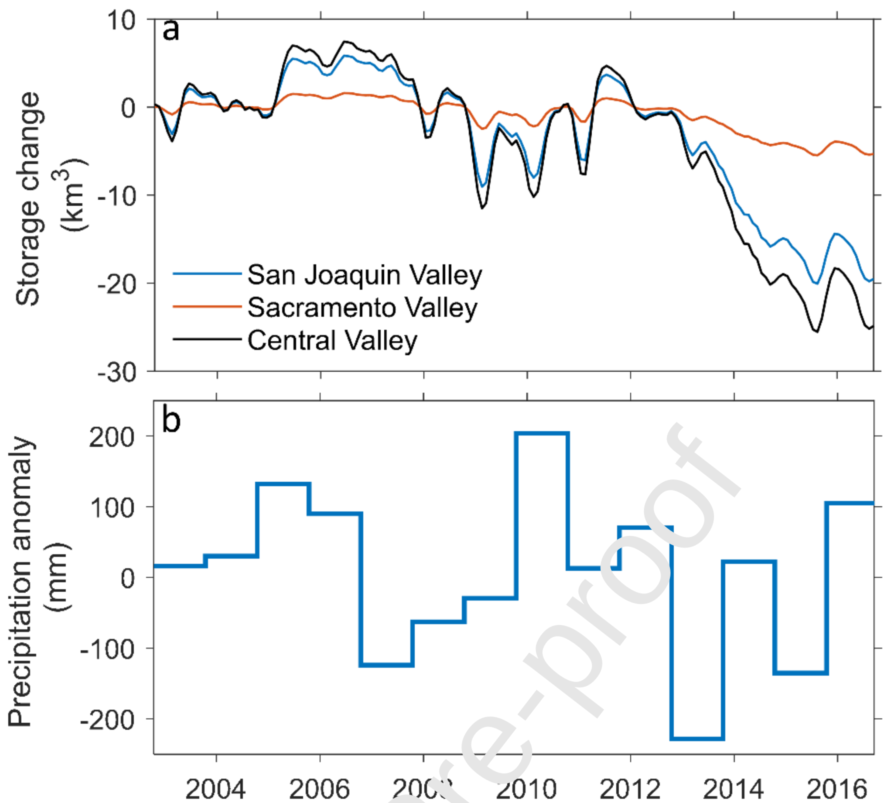



**Figure 8.** (a) Temporal variations of groundwater storage in Central, San Joaquin and Sacramento Valley, and (b) annual precipitation anomalies in the Central Valley.

| Time period | Annual groundwater volume loss (km$^3$ yr$^{-1}$) | | |
|---|---|---|---|
| | This Study | Results from Previous Studies | Reference Study |
| April 2006 - September 2009 | -5.1 ± 1.2 | -7.8 ± 0.8<br>-4.2 ± 0.3 | (Scanlon et al., 2012)<br>(Xiao et al., 2017) |
| April 2006 – March 2010 | -4.2 ± 1.0 | -6.0 ± 1.5 | (Famiglietti et al., 2011) |

| | | | |
|---|---|---|---|
| **January 2007 – December 2009** | -5.7 ± 1.2 | -7.1 ± 2.4<br>-5.5 ± 0.3<br>-6<br>-(3-10) | (Ojha et al., 2018)<br>(Xiao et al., 2017)<br>(Alam et al., 2021)<br>(Ahamed et al., 2022) |
| **October 2011 – September 2015** | -7.6 ± 1.5 (San Joaquin Valley only)<br>-9.8 ± 1.7 | -6.1 * (San Joaquin Valley only)<br>-7<br>-(6-17) | (Ojha et al., 2019)<br>(Alam et al., 2021)<br>(Ahamed et al., 2022) |
| **October 2012 – September 2016** | -7.7 ± 1.8 | -10.0 ± 0.2 | (Xiao et al., 2017) |

**Table 3**. Comparison of GWS loss obtained from this study with previously published estimates.

Groundwater declines over San Joaquin valley are more prominent than those over Sacramento valley, especially during the second drought period (Figure 8). In San Joaquin Valley, 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, Delta Mendota, and Westside subbasins, although lower groundwater depletion can be observed in Tule, Kern County, and Kaweah subbasins (Figure 1, Figure 9).

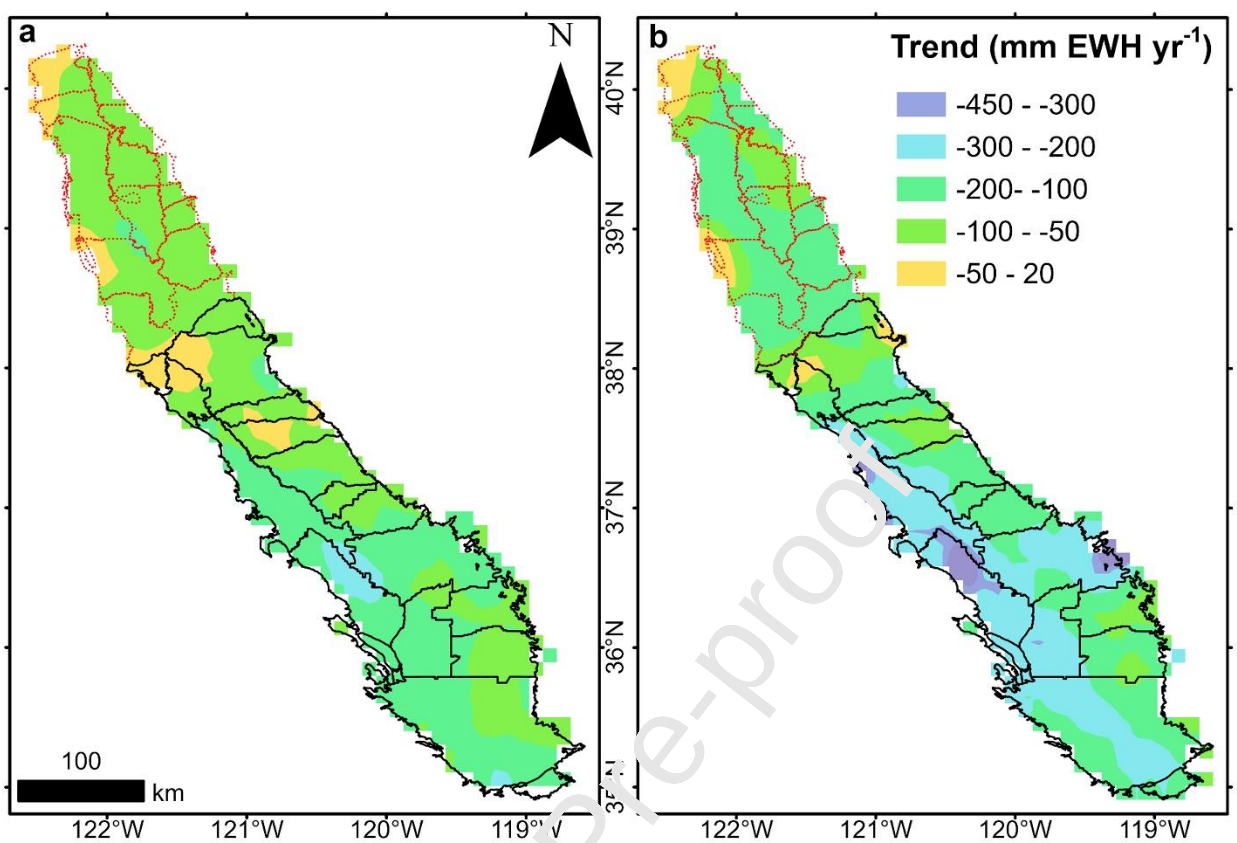



**Figure 9**. Spatial variations in modeled groundwater storage trends at 5-km resolution between (a) January 2007- December 2009 (b) October 2011–September 2015

## 4. Discussions

### 4.1 Machine Learning modeling

Our study achieved high accuracy for both training and test data in Sacramento and San Joaquin valleys (Figure 4). This suggest that downscaling of GRACE data to model groundwater level variations at sites of *in situ* wells was successful. The model development and training process adopting cross-validation scheme, avoided overfitting. Overfitting can reduce the confidence of ML results, which be a challenge for downscaling studies as we seek to model groundwater variations at higher resolutions (Roelofs, 2018). Previous studies using GRACE data for downscaling have obtained good accuracies for training data, but the accuracy on test data was significantly degraded (Jyolsna et al., 2021; Koch et al., 2019; Miro and Famiglietti, 2018; Rahaman et al., 2019; Seyoum and Milewski, 2017; Seyoum et al., 2019). Better accuracy achieved in this study might also be attributed to the choice of input variables such as texture and hydraulic conductivity which are important in groundwater modeling studies along with GRACE, removal of which causes the highest increase in RMSE of the models. We also found that input data used in modeling has different importance for Sacramento and San Joaquin Valley (Figure 5), suggesting that different processes are ongoing in the two regions.

In the following, we compare results from this study with those from other downscaling studies in Central Valley. Agarwal (2021), used only 180 wells for modeling in Central Valley through the Random Forest approach. As our accuracy estimates are similar to Agarwal (2021), we can conclude that Random Forest can accommodate additional data without sacrificing accuracy. Miro and Famiglietti (2018) used ANN to model annual GWS changes in southern San Joaquin Valley. We therefore compare similarities and differences between this study, Agarwal (2021) and Miro and Famiglietti (2018). Miro and Famiglietti (2018) obtained test NSE ranging from 0.039 to 0.751 when modeling GWS changes in southern San Joaquin valley using ANN. We obtained a better NSE (0.95) for test data in San Joaquin Valley when modeling monthly groundwater variations using RF. Even Agarwal (2021) obtained a NSE value of 0.86 for test data with ANN. It is worth noting that even though our study used similar input variables such as precipitation, temperature, and topographic slope from the same source as Miro and Famiglietti (2018), we obtained more accurate results. 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 Miro and Famiglietti (2018) modeled annual GWS changes, leaving less spatio-temporal data for modeling GWS changes. Moreover, they 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 krigged groundwater levels were used for training their ANN model, errors due to kriging interpolation can further propagate in the modeled GWS variations. We therefore propose better approaches for study design, model training and validation schemes with a potential of further improvement in accuracy for future studies.

### 4.2 Spatiotemporal variations in groundwater storage during the two drought periods

Several studies in the past have quantified GWS changes in Central Valley within different time periods (Table 3). Here, we focus on the two drought events; namely drought 1 (January 2007-December 2009), and drought 2 (October 2011 - September 2015) and compare our results with those from previous studies. The previously estimated mean GWS losses for drought 1 ranges from 19 $km^3$ (Alam et al., 2021) to 29 $km^3$ (Ahamed et al., 2022), while those for drought 2 vary from 28 $km^3$ (Alam et al., 2021) to 71 $km^3$ (Ahamed et al., 2022). Differences in the above estimates are due to different combinations of remote sensing, *in situ*, and model data used in the water balance approach. Xiao et al. (2017) estimated GWS 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 their estimates from GRACE. Ojha et al. (2018) estimated GWS 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 $km^3$ of groundwater from October 2011 to September 2015 based on GRACE data. Based on the GPS vertical deformation data, groundwater loss of 29.25 ± 8.7 $km^3$ was estimated for the same region and period (Ojha et al., 2019). GWS 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 differences between the estimated GWS losses for similar time periods. The causes of these differences include using different methods and datasets along with their associated errors. Scanlon et al. (2012) used a distributed specific yield ranging from 0.05-0.3 (Faunt, 2009) to estimate groundwater storage variations from *in situ* groundwater levels. Since regions with high groundwater level declines in southern San Joaquin valley have higher specific yields, it might have led to overestimation of groundwater storage changes in that region. 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 GWS changes from *in situ* groundwater level data will require significant spatio-temporal interpolation due to issues with coverage in many regions (Figure 1). However, GRACE-derived TWSA used as input variable in modeling is also affected by several errors during data processing, which might also have negative impacts on our ML model.

### 4.3 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 mean value within 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 well 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 concurrent with the dropping groundwater levels. 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 et al., 2016). At the end of drought 1, and during most of the drought 2 period, modeled groundwater level at the site of P304 was below the previous lowest level (pre-consolidation stress level). The high correlation (~0.71) 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. Regions showing higher groundwater depletion can be combined with information from geological models to identify potential sites that might be further vulnerable to subsidence.

Significant groundwater depletion can be seen for subbasins in Tulare basin and western San Joaquin valley for both drought events. 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 has been 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 GWS changes 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.

**4.4 Limitations and Future Studies**

The downscaling approach presented here shows promise for general applicability, however there are some caveats. For example, no ML method including the one we used (i.e., RF) can be used for extrapolation, or in other words, accurate predictions ahead/ forecasting cannot be made without any prior knowledge or assumptions such as continuation of the long-term trend seen during the training data period also out of the training regime which may not be realistic under today's climate change conditions (Milly et al., 2008). As a future work, we propose building deep neural networks incorporating larger datasets and wider regions combined with mechanistic approaches (Razavi, 2021) to model more complex variations and to provide at least short-term forecasts of groundwater variations with commensurate accuracy.

Further, unlike the water balance method of Ahamed et al. (2022), which has modeled GWS variations for the longest period, 2002-2020, so far 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., 2022), and availability of modeled GRACE data from such studies can be used to extend the study for a longer time. This is beyond the scope of this study and left for a future work.

Even with the mentioned limitations, the results obtained from this study are useful for geodesy, hydrogeology, and for further downscaling studies. We present a simple, yet robust approach for downscaling GRACE data utilizing diverse hydrometeorological and geological data and addressed the complex groundwater modeling problem. GRACE temporal gravity data is a powerful tool to quantify regional GWS changes (Famiglietti, 2014), while *in situ* wells are very useful for precise measurements of groundwater level. This study further uses this data to produce a downscaled gridded GWSA product useful to resource managers. For example, how GWS varies with changes in precipitation regimes and human abstractions for each subbasin in Central Valley can be quantified reliably by the approach presented in this study. This can be a useful first indicator of future availability of these resources. We model the impact of climate extremes such as droughts on GWSA and correlated the variations with vertical deformation to obtain storage coefficients which will be useful for modeling aquifers.  Several machine learning and hydrological models require the continuous availability of the groundwater level data for calibration and this study can fulfil that need.

## 5 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 severe droughts during our study period. Producing reliable 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 achieved high modeling accuracy for San Joaquin and Sacramento Valley, proving that Random Forest is a robust machine learning model for such downscaling applications. We obtained comparable or better prediction accuracy than previous 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 feature 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 with higher spatial resolution using Sentinel-1, GRACE-FO, and *in* situ groundwater level data can be generated.

Central Valley exhibits groundwater storage 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 because of ground water decline.

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. Though this study addresses the spatial downscaling of GWS changes, the temporal downscaling is also likely to gain more importance in the near future considering the ever-increasing impacts of climate change.

## References


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., 2021. 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

Bandikova, T., McCullough, C., Kruizinga, G.L., Save, H., Christophe, B., 2019. GRACE accelerometer data transplant. Advances in Space Research 64, 623–644. https://doi.org/10.1016/J.ASR.2019.05.021

Berner, L.T., et al, 2020. Summer warming explains widespread but not uniform greening in the Arctic tundra biome. Nat. Commun. 11, 4621. https://doi.org/10.1038/s41467-020-18479-5

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

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

Chen, Y., et al. 2021. Thermokarst acceleration in Arctic tundra driven by climate change and fire disturbance. One Earth 4, 1718–1729. https://doi.org/10.1016/j.oneear.2021.11.011

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

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.

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., 2010. Development of a three-dimensional model of sedimentary texture in valley-fill deposits of Central Valley, California, USA. Hydrogeology Journal18:3 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

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

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

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., 2013, Groundwater depletion in the United States (1900−2008): U.S. Geological Survey Scientific Investigations Report 2013-5079, 63 p., http://pubs.usgs.gov/sir/2013/5079

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

Mace, R., Chowdhury, A., Anaya, R., Way, S., 2000. Groundwater Availability of the Trinity Aquifer, Hill Country Area, Texas: Numerical Simulations through 2050: Texas Water Development Board, 172 p

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

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: whither water management? Science 319, 573–574.

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

Mukherjee, A., Ramachandran, P., 2018. Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India : Analysis of comparative performances of SVR, ANN and LRM. Journal of Hydrology, 558, 647–658. https://doi.org/10.1016/J.JHYDROL.2018.02.005

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

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.

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

Razavi, S., 2021. Deep learning, explained: Fundamentals, explainability, and bridgeability to process-based modelling. Env. Mod. And Software, 144, 105159. https://doi.org/10.1016/j.envsoft.2021.105159

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

Shahriari, B., Swersky, K., Wang, Z., Adams, R., de Freitas, N., 2010. Taking the human out of the loop: A review of Bayesian optimization. Proceedings of the IEEE 104(1), 148–175

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

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

Sneed, M., Brandt, J., Solt, M., 2013. Land Subsidence along the Delta-Mendota Canal in the Northern Part of the San Joaquin Valley, California, 2003–10. USGS Scientific Investigations Report 2013-5142.

Sneed, M., Brandt, J.T., 2015. Land subsidence in the San Joaquin Valley, California, USA, 2007–2014. Proceedings of IAHS 372, 23–27. https://doi.org/10.5194/PIAHS-372-23-2015

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

Sun, A.Y., Scanlon, B.R., 2019. How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. Environmental Research Letters 14, 073001. https://doi.org/10.1088/1748-9326/AB1B7D

Sun, Z., Long, D., Yang, W., Li, X., Pan, Y., 2020. Reconstruction of GRACE Data on Changes in Total Water Storage Over the Global Land Surface and 60 Basins. Water Resources Research, 56(4), e2019WR026250. https://doi.org/10.1029/2019WR026250

Sun, Y., Kang, S., Li, F., & Zhang, L., 2009. Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. Environmental Modelling & Software, 24(10), 1163–1170. https://doi.org/10.1016/J.ENVSOFT.2009.03.009

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

Uz, M., Atman, K.G., Akyilmaz, O., Shum, C.K., Keleş, M., Ay, T., Tandoğdu, B., Zhang, Y., Mercan, H., 2022. Bridging the gap between GRACE and GRACE-FO missions with deep learning aided water storage simulations. Science of The Total Environment 830, 154701. https://doi.org/10.1016/J.SCITOTENV.2022.154701

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

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. [Doctoral dissertation, Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1593533552523635

Yin, W., Fan, Z., Tangdamrongsub, N., Hu, L., Zhang, M., 2021. 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

## Author Contributions

**Vibhor Agarwal**: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing- Original draft preparation, Writing- Reviewing and Editing.

**Orhan Akyilmaz**: Conceptualization, Methodology, Resources, Writing- Reviewing and Editing, Supervision.

**CK Shum**: Conceptualization, Writing- Reviewing and Editing, Supervision, Project administration, Funding acquisition.

**Wei Feng**: Software, Writing- Reviewing and Editing, Supervision.

**Ting-Yi Yang**: Resources, Writing- Reviewing and Editing, Visualization. Ehsan Forootan: Writing- Reviewing and Editing.

**Tajdarul Hassan Syed**: Writing- Reviewing and Editing.

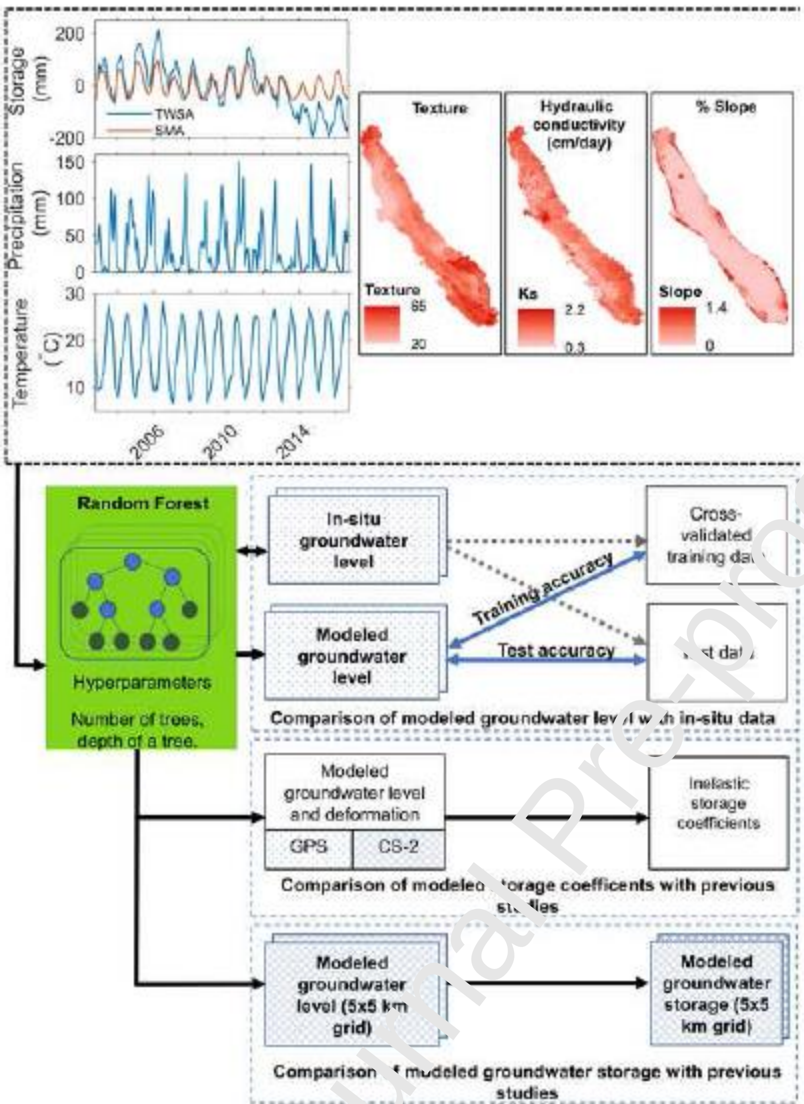
**Umesh Haritashya:** Writing- Reviewing and Editing.

**Metehan Uz:** Resources.

**Declaration of interests**

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:




**Graphical abstract**

Highlights

1. Machine Learning approach to integrate in-situ groundwater and remote sensing data
2. Groundwater storage variations for one and a half-decade at 5 km resolution
3. Model the impact of two droughts on groundwater storage variations
4. Novel application for broader applicability of GRACE gravimetry data