A New Hybrid Water Balance and Machine Learning Approach for Groundwater Withdrawal Prediction using Integrated Multi-Temporal Remote Sensing Datasets

Sayantan Majumdar[1], Ryan Glen Smith[1], James J Butler[2], and Venkataraman Lakshmi[3]

[1]Missouri University of Science and Technology
[2]Kansas Geological Survey, University of Kansas
[3]University of Virginia

November 21, 2022

Abstract

Effective monitoring of groundwater withdrawals is necessary to help mitigate the negative impacts of aquifer depletion. In this study, we develop a holistic approach that combines water balance components with a machine learning model to estimate groundwater withdrawals. We use both multi-temporal satellite and modeled data from sensors that measure different components of the water balance at varying spatial and temporal resolutions. These remote sensing products include evapotranspiration, precipitation, and land cover. Due to the inherent complexity of integrating these data sets and subsequently relating them to groundwater withdrawals using physical models, we apply random forests- a state of the art machine learning algorithm- to overcome such limitations. Here, we predict groundwater withdrawals per unit area over a highly monitored portion of the High Plains aquifer in the central United States at 5 km resolution for the years 2002-2019. Our modeled withdrawals had high accuracy on both training and testing datasets (R[?] 0.99 and R[?] 0.93, respectively) during leave-one-out (year) cross-validation with low Mean Absolute Error (MAE) [?] 4.26 mm and Root Mean Square Error (RMSE) [?] 13.57 mm for the year 2014. Moreover, we found that even for the extreme drought year of 2012, we have a satisfactory test score (R[?] 0.79) with MAE [?] 10.34 mm and RMSE [?] 27.04 mm. Therefore, the proposed hybrid water balance and machine learning approach can be applied to similar regions for proactive water management practices.

A New Hybrid Water Balance and Machine Learning Approach for Groundwater Withdrawal Prediction using Integrated Multi-Temporal Remote Sensing Datasets

S. Majumdar [1], R. Smith [1], J.J. Butler Jr. [2], V. Lakshmi [3]

[1] Missouri University of Science and Technology, Rolla, MO, USA
[2] Kansas Geological Survey, University of Kansas, Lawrence, KS, USA
[3] University of Virginia, Charlottesville, VA, USA

Corresponding author: Sayantan Majumdar (s.majumdar@mst.edu)

Key Points:

- Groundwater withdrawals are not actively monitored in most places of the world at a scale necessary to implement sustainable solutions.
- Various multi-temporal remote sensing data are integrated into a machine learning framework to effectively predict groundwater withdrawals.
- The results over the High Plains Aquifer, Kansas, USA show that this approach is applicable to similar regions having sparse in-situ data.

Abstract

Effective monitoring of groundwater withdrawals is necessary to help mitigate the negative impacts of aquifer depletion. In this study, we develop a holistic approach that combines water balance components with a machine learning model to estimate groundwater withdrawals. We use both multi-temporal satellite and modeled data from sensors that measure different components of the water balance at varying spatial and temporal resolutions. These remote sensing products include evapotranspiration, precipitation, and land cover. Due to the inherent complexity of integrating these data sets and subsequently relating them to groundwater withdrawals using physical models, we apply random forests— a state of the art machine learning algorithm— to overcome such limitations. Here, we predict groundwater withdrawals per unit area over a highly monitored portion of the High Plains aquifer in the central United States at 5 km resolution for the years 2002-2019. Our modeled withdrawals had high accuracy on both training and testing datasets ($R^2 \approx 0.99$ and $R^2 \approx 0.93$, respectively) during leave-one-out (year) cross-validation with low Mean Absolute Error (MAE) ≈ 4.26 mm and Root Mean Square Error (RMSE) ≈ 13.57 mm for the year 2014. Moreover, we found that even for the extreme drought year of 2012, we have a satisfactory test score ($R^2 \approx 0.79$) with MAE ≈ 10.34 mm and RMSE ≈ 27.04 mm. Therefore, the proposed hybrid water balance and machine learning approach can be applied to similar regions for proactive water management practices.

Plain Language Summary

Groundwater is an essential component of global water resources and is the largest source of Earth's liquid freshwater. It is extensively used for drinking water and to support global food production. Consequently, groundwater consumption has significantly increased owing to the pressing demands for water, food, and energy primarily driven by the increasing global population. Despite its critical role in the water-food-energy nexus, very few regions in the United States (US) or elsewhere actively monitor their groundwater withdrawals (also known as extraction or pumping) for implementing sustainable water management solutions. We develop a hybrid approach combining water balance components measured using openly available remote sensing products with a machine learning model to estimate groundwater withdrawals. This framework automatically learns the inter-relationships among these variables and groundwater withdrawals. Our study area is a portion of the High Plains Aquifer in Kansas (central US) where overpumping has caused substantial groundwater storage loss. Also, a large amount of groundwater pumping data is available for validation. Our results indicate good accuracy even for extreme drought years. Thus, this approach should be applicable to similar regions having sparsely or moderately available groundwater pumping data, enabling water managers to proactively implement sustainable solutions addressing water security issues.

1 Introduction

Groundwater constitutes nearly 30% of total global freshwater reserves (Schneider et al., 2011) and is a key component of the water-food-energy nexus (Smajgl et al., 2016). Globally, about half of the drinking water is supplied by groundwater (Gleeson et al., 2019). Moreover, groundwater is extensively used in agricultural activities and the demand for it is increasing due to the rise in the global population, dietary shifts, and climate change (Margat & van der Gun, 2013). Monitoring of groundwater withdrawals (pumping) is needed for quantifying aquifer depletion and to provide essential information for building groundwater models to manage the resource. However, most places in the world, including the United States (US), do not actively monitor their groundwater withdrawals at a high-enough spatial resolution for implementing sustainable water management policies. Without that monitoring, it will be exceedingly difficult to address the negative impacts of groundwater overuse, which include permanent aquifer storage loss, land subsidence and water contamination (Butler et al., 2018; Chaussard & Farr, 2019; Erban et al., 2013; Galloway & Burbey, 2011; MardanDoost et al., 2019; Smith et al., 2017, 2018).

Due to the availability of voluminous amounts of satellite data, it is possible to monitor large regions using remote sensing techniques (Frappart & Bourrel, 2018; Leidner & Buchanan, 2018). In the hydrologic remote sensing domain, there is a multitude of spaceborne missions that provide satellite products for assessing different quantities related to the global water cycle. The most prominent of these products quantify total water storage— GRACE (Gravity Recovery and Climate Experiment) and GRACE-FO (GRACE- Follow On), terrestrial evapotranspiration— MODIS (Moderate Resolution Imaging Spectroradiometer), soil moisture— SMAP (Soil Moisture Active Passive), precipitation— GPM (Global Precipitation Measurement), TRMM (Tropical Rainfall Measuring Mission), and land cover— USDA-NASS (United States Department of Agriculture- National Agricultural Statistics Service) (Boryan et al., 2011; Chan et al., 2016; MardanDoost et al., 2019; Nie et al., 2018; Yi et al., 2018).

Each of these datasets is related in some way to hydrologic fluxes that impact the groundwater system. A great deal of research has been done on using these datasets to estimate groundwater fluxes. A number of studies have used GRACE to estimate changes in total water storage, then subtract the components of soil moisture, surface water, and snow water to estimate fluxes in groundwater storage (Famiglietti et al., 2011; Rodell et al., 2007, 2009). While useful for basin- or continental-scale studies, the resolution of GRACE is too coarse (~400 km) for local estimates of groundwater flux Surface water availability and land use (which is a proxy for water demand) have also been tied to groundwater withdrawals in previous studies (Faunt, 2009). Many land use datasets are now derived from remote sensing, and a growing number of studies are using remote sensing estimates of land use to estimate irrigated area (Deines et al., 2017; Ozdogan & Gutman, 2008), although this has not been tied directly to groundwater extraction. Others have used in-situ data in a water balance approach to estimate groundwater fluxes (Butler et al., 2016, 2018).

Satellite methods have been used to estimate groundwater storage change at high resolution (~100 m) utilizing land subsidence estimates from Interferometric Synthetic Aperture Radar (InSAR, e.g., Hoffmann et al., 2001; Smith et al., 2017), but this approach is typically limited to regions with confined or semi-confined aquifers (where pressure changes are highest) and highly compressible sediments (Smith and Majumdar, 2020). In spite of the mature research in many of these individual fields, very rarely are the various remotely sensed datasets combined to estimate groundwater fluxes, which limits their ability to estimate these fluxes to very specific cases. In this study, we seek a generalizable approach that utilizes satellite data to estimate groundwater withdrawals at the local scale (5 km resolution) without prior knowledge of withdrawals. We accomplish this by integrating diverse satellite datasets that are related to different components of the water balance.

Integrating satellite data for estimating the different water balance components can be immensely challenging due to the varying spatial and temporal resolutions (Tamayo-Mas et al., 2018). Moreover, we lack methods to estimate several key parameters, such as recharge, groundwater inflow/outflow, and surface water withdrawals, to 'close the loop' of the water balance and directly estimate groundwater withdrawals. Furthermore, existing water balance estimates have been shown, in many cases, to exhibit spatial bias, limiting their accuracy to some extent (Hashemi et al., 2017). These roadblocks have limited our ability to estimate groundwater withdrawals except in the special cases noted previously. In addition, traditional approaches to calibrate physical models that incorporate these diverse datasets are challenging because the models quickly become overly complex and computationally intensive, making the parameter estimation procedure required to develop accurate estimates infeasible (Becker et al., 2019; Moeck et al., 2018; Seibert et al., 2019; Tamayo-Mas et al., 2018).

In this research, rather than using traditional physical models, we leverage the correlations between various water balance measurements and groundwater withdrawals in a machine learning framework that learns the relationship between the various datasets and uses them in a predictive fashion. Here, we apply random forests (Belgiu & Drăguț, 2016; Breiman, 2001), a state of the art machine learning algorithm, to develop local scale (5 km) estimates of groundwater withdrawals over Kansas (part of the High Plains Aquifer in the central US). Regarding the satellite and modeled datasets, we include the water balance products from MODIS, PRISM (Parameter-elevation Regressions on Elevation Slopes Model), and USDA-NASS acquired over this region between 2002 and 2019. Here, we use PRISM as it is specific to the US and the precipitation estimates are interpolated from a dense array of rain gauges thereby providing higher accuracies than TRMM and GPM (Cannon et al., 2017; Hashemi et al., 2017). Finally, this novel hybrid water balance and machine learning framework is validated against the annual in-situ groundwater pumping data available over this area and the results are also compared to GRACE and GRACE-FO Total Water Storage (TWS) (Landerer & Swenson, 2012; Swenson, 2012).

The remainder of this paper is divided into five main sections. We first discuss the characteristics of the study area (section 2) followed by the details of our methods (section 3). We then provide an in-depth analysis of the results (section 4) and end with a discussion of the broader implications and applicability of this work to other regions (sections 5 and 6).

2 Study Area

The High Plains Aquifer (HPA) in the central US is one of the world's most extensive and productive aquifers helping to support agricultural demand in the US and across the globe. It spans over eight states and its water is predominantly used for irrigation of crops, including wheat, corn, cotton, and soybean. However, the HPA is heavily stressed (mostly in the central and southern parts) primarily due to groundwater over-pumping driven by the large-scale demand for freshwater and increasing variations in seasonal evapotranspiration, precipitation, and floods resulting from climate change (Smidt et al., 2016). More specifically, the situation in the state of Kansas, particularly in the western area, is critical as 90% of its irrigation water is supplied by the HPA (Butler et al., 2018). In addition, recent trends show that the water table is declining rapidly, and some regions are at risk of completely depleting their aquifer. This would significantly impact the agricultural productivity of the state in the near future thereby also hindering its economic vitality (Butler et al., 2018; Deines et al., 2020).

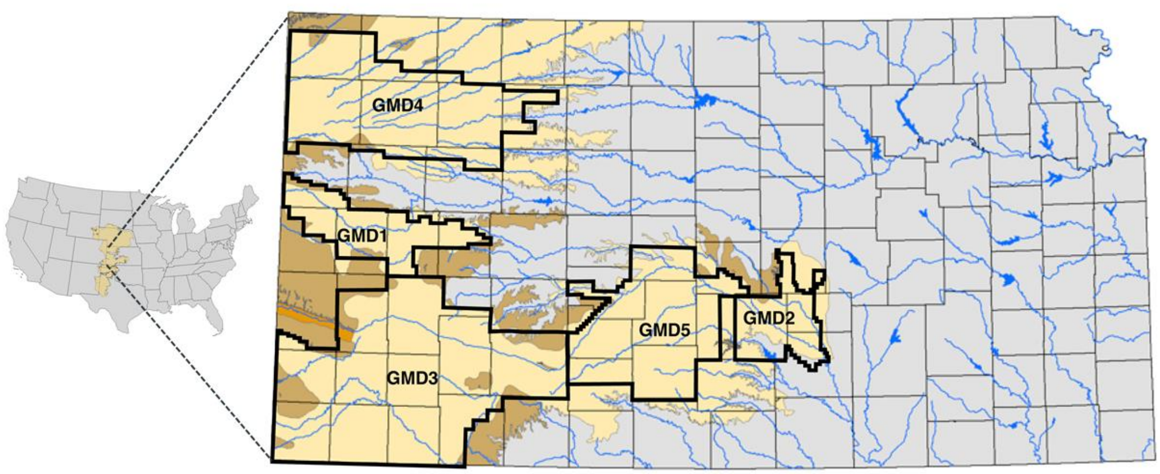


Figure 1. Map of conterminous US highlighting the study area, Kansas, as well as the state's five Groundwater Management Districts (GMDs), all of which overlie the HPA.

Currently, the only feasible way to mitigate these negative impacts of extreme groundwater withdrawals is to reduce groundwater extraction (Butler et al., 2020). This is because surface water is scarce in western Kansas and the water table decline is driven by groundwater pumping. Over 95% of the non-domestic pumping wells in the Kansas High Plains aquifer are metered and pumping volumes must be reported annually (Butler et al., 2018). The pumping data are available through the Water Information Management and Analysis System (WIMAS) database that can be accessed at the Kansas Geological Survey (Wilson, 2019). Since "data" is the crux of any machine learning model (Hastie et al., 2013), we chose the entire state of Kansas as our study area, which is highlighted in Figure 1. Considering the severe aquifer stress in the region and the need for proactive management solutions, we believe that this is the perfect test area for our proposed hybrid water balance and machine learning framework.

3 Methodology

The overall workflow is shown in Figure 2; the main steps are data acquisition (step 1) using the Google Earth Engine (GEE) platform (Gorelick et al., 2017), pre-processing (step 2), and the implementation of the hybrid water balance and machine learning model (step 3) using Python 3. After that, we perform model evaluation and analyze the groundwater withdrawal predictions using the available in-situ groundwater pumping geo-database (GDB). Here, the input time series data consist of evapotranspiration (ET), precipitation (P), and density of the following land-use types: agriculture (AGRI), urban (URBAN), and surface water (SW). We provide a detailed description of our workflow involving the different processing blocks and the time series data specifications in subsections 3.1 and 3.2.

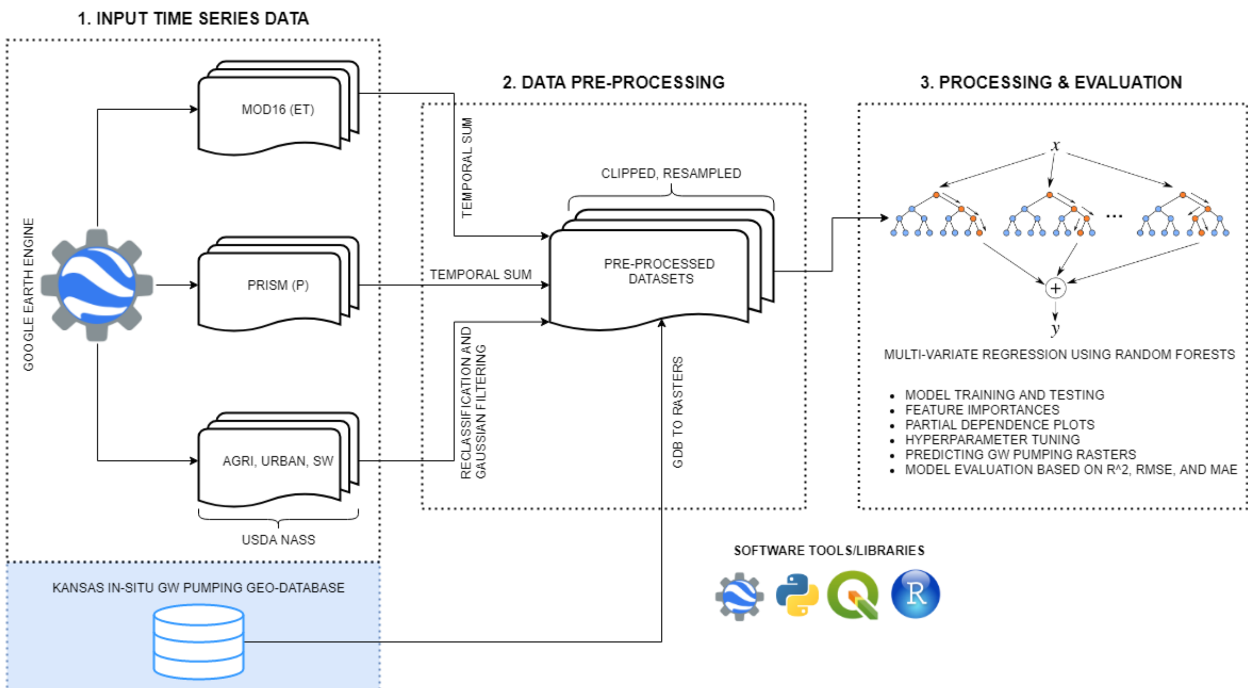


Figure 2. Steps involved in the proposed workflow for predicting groundwater (GW) withdrawals. The machine learning model is implemented in step 3 after data downloading and pre-processing. It is a fully automated and reproducible framework developed using open source or freely available tools, libraries, and data. The icons were downloaded from the respective official websites and IU Digital Science Center (2013) (random forest figure). Also, QGIS (QGIS Development Team, 2019) is used for visualization.

3.1 Data Acquisition and Pre-Processing

GEE, a cloud-based platform for large-scale geospatial data processing and analysis, is accessible to researchers and operational users (Gorelick et al., 2017). We used GEE for seamless and efficient downloading of our datasets which include MODIS (ET), PRISM (P), and land-cover data (used to produce AGRI, URBAN, SW) from USDA-NASS. The specific product names along with the corresponding description are given in Table 1. Moreover, for the MOD16 and PRISM products, we restricted the time span between Apr 1 and Sep 30 (inclusive) of each year (the typical growing season (Butler et al., 2019). Our assumption was that the land cover does not change significantly between years over our study period, so we selected the USDA-NASS Cropland Data Layer (CDL) for the year 2015 only. While many farmers rotate crops, we assume that the rotations did not significantly change the percentage of each 5 km x 5 km pixel that was cultivated over our study period. This dataset does not distinguish between dryland and irrigated crops. To compensate for this, we included precipitation and actual evapotranspiration as input datasets. By combining these datasets, our algorithm is able to identify pixels with low precipitation and high ET, and thus high groundwater demand.

The Kansas in-situ groundwater pumping data were obtained from the WIMAS database (Wilson, 2019). We discarded pixels with average extraction rates of more than 1000 mm per year. These rates only occurred in two pixels, which had high-capacity wells that were withdrawing large volumes of water directly beneath the Missouri River, and are thus more representative of surface water withdrawals than groundwater withdrawals. The in-situ groundwater pumping data have not been released for 2019 and hence, this year is kept solely for forecasting.

We took the temporal sum of both the ET and P bands between April and September of each year thereby obtaining the total ET and P for that period for each year. In addition, the original USDA-NASS CDL dataset (covering the conterminous US) for the year 2015, which classifies land use based on the specific crop type, or water, urban, etc., was reclassified according to Table 2. Thereafter, this reclassified dataset was split into separate binary rasters (AGRI, SW, and URBAN) with pixel values 0 and 1 corresponding to the absence and presence of the respective class (OTHER class is ignored). The Kansas groundwater pumping database (GDB) was automatically converted into yearly shapefiles and subsequently rasters (5 km spatial resolution, UTM 14N projection) using GDAL/OGR Python APIs, GeoPandas and Rasterio (GDAL/OGR contributors, 2019; GeoPandas developers, 2019; Gillies, 2013; Oliphant, 2006). All of these rasters (ET, P, AGRI, URBAN, and SW) were reprojected to UTM 14N, resampled to 5 km spatial resolution, and clipped using the intersection of the groundwater pumping rasters and the Kansas GMD boundary file shown in Figure 1. Each 5 km x 5 km pixel contains the total annual groundwater withdrawal within that pixel for years 2002-2018. The value was computed as total volume of groundwater pumping in that region divided by the area, 25 km^2.

Table 1. Specifications of the time series datasets used in the workflow.

Product Name	Time Period	Pre-processing Applied	Brief Description
MOD16A2.006: Terra Net Evapotranspiration 8-Day Global 500m	2002-2019 (Apr-Sep)	Temporal sum over the 'ET' band	The ET data are available as cumulative 8-day composite images at 500 m spatial resolution (Running & Mu, 2015).
PRISM Monthly Spatial Climate Dataset AN81m	2002-2019 (Apr-Sep)	Temporal sum over the 'ppt' band	The PRISM group provides daily and monthly gridded climate datasets for the conterminous USA. Here we use the monthly datasets available at about 4 km spatial resolution.
USDA-NASS Cropland Data Layers	2015 (Annual)	Reclassification of the 'cropland' band and Gaussian filtering	CDL is an annual crop-specific land cover data layer (30 m spatial resolution) created for the continental USA developed using MODIS and ground-truth data (USDA-NASS, 2015).
Kansas in-situ Groundwater Pumping Data	2002-2018 (Annual)	Groundwater pumping data to rasters	The Kansas groundwater pumping database (GDB) is publicly available (Wilson, 2019). The GDB contains point data across all counties in Kansas.

Table 2. Reclassification table for the pre-processed USDA-NASS CDL data. The original CDL raster data values lie in the interval [1, 254].

Original Class Interval(s)	Reclassified	Label
(0, 59], (66, 77], (203, 255]	1	AGRI
(110, 111]	2	SW
(120, 124]	3	URBAN
(111, 112], (59, 61], (130, 195]	4	OTHER

Furthermore, we applied a Gaussian filter available from the Scipy library (Jones et al., 2001) over each of the AGRI, URBAN, and SW rasters to create smoothed rasters representing the density of each land-use type within a given radius, the size of which is a function of the standard deviation (σ). Here, we set $\sigma = 3$ for AGRI and URBAN, and $= 5$ for SW. The filtered data were then normalized over the interval [0, 1] where the values represent the density of these classes, i.e., pixel values close to 0 have very few of the respective class within the specified window, and vice-versa. Finally, the pre-processed datasets were organized into a Pandas DataFrame object (McKinney, 2010). It is noteworthy that feature ordering affects model performance in any decision-tree-based algorithm (Breiman, 2001; Olson, 2015). In order to have consistent feature ordering, the CSV file was sorted based on the attribute names (Breiman, 2001; Olson, 2015). The feature or attribute names in alphabetical order are AGRI, ET, P, SW, and URBAN where the ET and P are in mm and AGRI, URBAN, and SW are measures of land type density and are unitless. The groundwater pumping values (mm) are marked as the label set (groundwater pumping is the target variable for prediction). It should be noted that our proposed framework provides the flexibility to alter any of these parameters (σ, spatial resolution, reclassification values) as per the requirement of the specific application. Moreover, in our initial setup, we chose these specific values for model workability purposes. However, observing the effects of modifying these pre-processing parameters could provide further insights into the spatial and temporal mechanisms driving groundwater demand and usage.

3.2 Our Approach-Random Forests

Random Forests is a widely used machine learning algorithm that has been extensively applied in the remote sensing domain, primarily for classification purposes with non-linear or multimodal datasets (Belgiu & Drăguţ, 2016). Although in this work, we use it for multi-variate regression, the general idea remains the same where the algorithm employs a set of Classification and Regression Trees (CART) for prediction. Essentially, the random forests algorithm behaves as an ensemble or meta estimator where it uses averaging across all the CARTs to improve the predictive accuracy and control over-fitting (Breiman, 2001). To test the performance of our model, we split the original data into training and testing data because the random forests follow a supervised machine learning approach. Here, two main hyperparameters are involved– the number of estimators or trees (*n_estimators*) and the number of features to consider during the best split operation (*max_features*). If *max_features* = *n_features* (number of actual features), then the sub-sample size is equal to the original input sample size where the samples are drawn with replacement. We have iterated our model using different values for both these hyperparameters. Also, the random_state parameter controls the randomness of the sample bootstrapping and therefore, remained fixed throughout the workflow for model reproducibility (Pedregosa et al., 2011).

In our final model setup, we chose *n_estimators* = 500 which is a commonly used value for remote sensing studies (Belgiu & Drăguţ, 2016; Smith et al., 2018), *random_state* = 0, *max_features* = 5, and the other hyperparameters as scikit-learn defaults (Pedregosa et al., 2011). The parameters *n_estimators* and *max_features* were optimized to minimize the Root Mean Square Error (RMSE) on our testing (validation) datasets (described in section 4). We split the training and testing data considering a leave-one-out (year) cross-validation method where the user is able to choose a particular year as pure test and remaining years as training data respectively.

Regarding model evaluation, we considered the feature importances, partial dependence plots, and different error metrics for the predictions such as the coefficient of determination (R^2), RMSE, and Mean Absolute Error (MAE). The feature or variable importances are values (sum up to 1) signifying the impact of each variable (the higher the value, the more important is the feature). Although these can be calculated in several ways, in our approach, the feature importance, also called Gini importance or mean decrease in importance (MDI) is defined as the total decrease in node impurity (weighted by the probability of reaching a particular node) which is averaged over all the trees (Breiman et al., 1984). Currently, this is the most computationally efficient method and is widely used in the scientific community (Breiman et al., 1984; Pedregosa et al., 2011). Partial dependence plots in conjunction with the feature importances can be useful to interpret the effect of specific variables on the predicted outcome. Partial dependence plots determine the effect on the prediction outcomes when one variable is varied while accounting for all possible values from the rest (Hastie et al., 2013). The equation for partial dependence plots is given below:

$$\tilde{f}(x) = \frac{1}{n}\sum_{i=1}^{n} f(x, x_{ic}) \qquad (1)$$

where, $\tilde{f}(x)$ is the partial dependence function, n is the number of rows in the dataset, $f(x, x_{ic})$ is the random forest model, x is the predictor variable of interest, and x_{ic} denotes the values of all other variables.

This allows us to quantify how each input variable impacts groundwater pumping. For example, we expect regions or years with low precipitation and high evapotranspiration to have high groundwater pumping values. If this is verified in our partial dependence plots, we can have confidence that our model is capturing the dynamics of the hydrologic system. In our proposed workflow, we use the scikit-learn library (Pedregosa et al., 2011) to obtain the different partial dependence plots which are discussed in section 4. Lastly, we also perform residual diagnostics to observe the error distribution and perform normality checks (Hastie et al., 2013).

4 Results and Analysis

In order to better assess our proposed model, we perform two main test cases– one with 2014 (average to slightly dry year) as a pure validation dataset and the other with 2012 (extreme drought). Additionally, we also perform model evaluation based on multiple validation datasets. Here, we describe the results and the corresponding analysis in each of these scenarios.

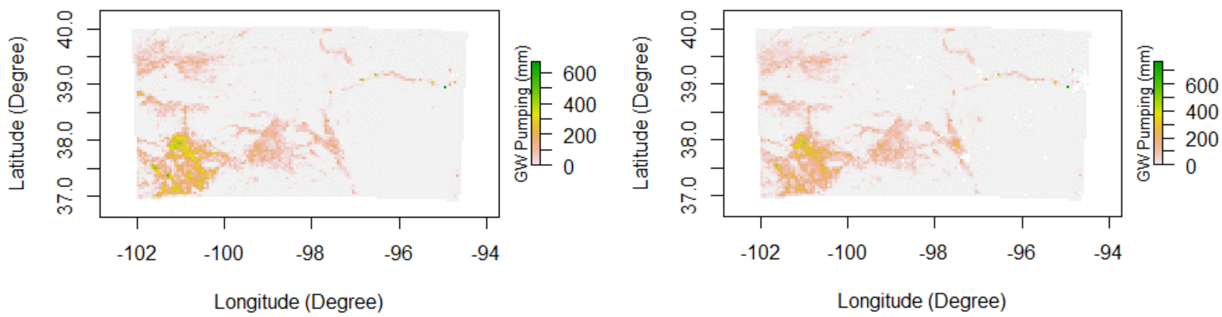
4.1 Test Case I– Year 2014

In this case, we select the pre-processed datasets for the year 2014 as pure validation data for performing leave-one-out cross-validation. The number of rows, n, is 142,600 for the training dataset and 8,912 for the testing or validation dataset. Accordingly, we obtain an overall training score, $R^2 \approx 0.99$. Moreover, the error metrics shown in Table 3 suggest that we are achieving good prediction results for the validation dataset (or testing dataset), $R^2 \approx 0.93$, RMSE ≈ 13.57 mm, MAE ≈ 4.26 mm, and normalized MAE ≈ 0.22 (the MAE is normalized by the average annual pumping).

Table 3. Error metrics (rounded to 2 decimal places) for Test Case I.

YEAR	R^2	RMSE (mm)	MAE (mm)	Normalized MAE
2002	0.98	8.18	2.94	0.12
2003	0.99	6.29	2.19	0.10
2004	0.99	5.44	1.85	0.10
2005	0.99	4.20	1.51	0.08
2006	0.99	5.95	2.00	0.10
2007	0.98	6.60	2.38	0.13
2008	0.99	5.38	1.72	0.09
2009	0.99	4.09	1.42	0.08
2010	0.99	4.52	1.59	0.08
2011	0.98	9.31	3.10	0.13
2012	0.98	8.06	3.02	0.13
2013	0.99	5.77	1.78	0.09
2014	**0.93**	**13.57**	**4.26**	**0.22**
2015	0.99	3.97	1.41	0.08
2016	0.99	4.46	1.58	0.10
2017	0.98	5.53	1.71	0.10
2018	0.99	4.61	1.62	0.10

The actual and predicted groundwater pumping rasters for 2014 are given in Figure 3 (a) and (b), respectively.



(a) (b)

Figure 3. (a) Actual and (b) Predicted groundwater (GW) pumping rasters for 2014. This plot was created using R (R Core Team, 2019). The no data values in the predicted raster are introduced by the ET dataset which had some pixels (falling inside the study area) assigned as no data.

Next, we check the goodness of fit based on different plots (R Core Team, 2019) involving predicted values and residual diagnostics. The 1:1 relationship between the actual and predicted groundwater pumping values is given in Figure 4 which shows that the predictions closely follow the actual values for 2014 ($R^2 \approx 0.93$ and Residual Standard Error (RSE) ≈ 13.34 mm).

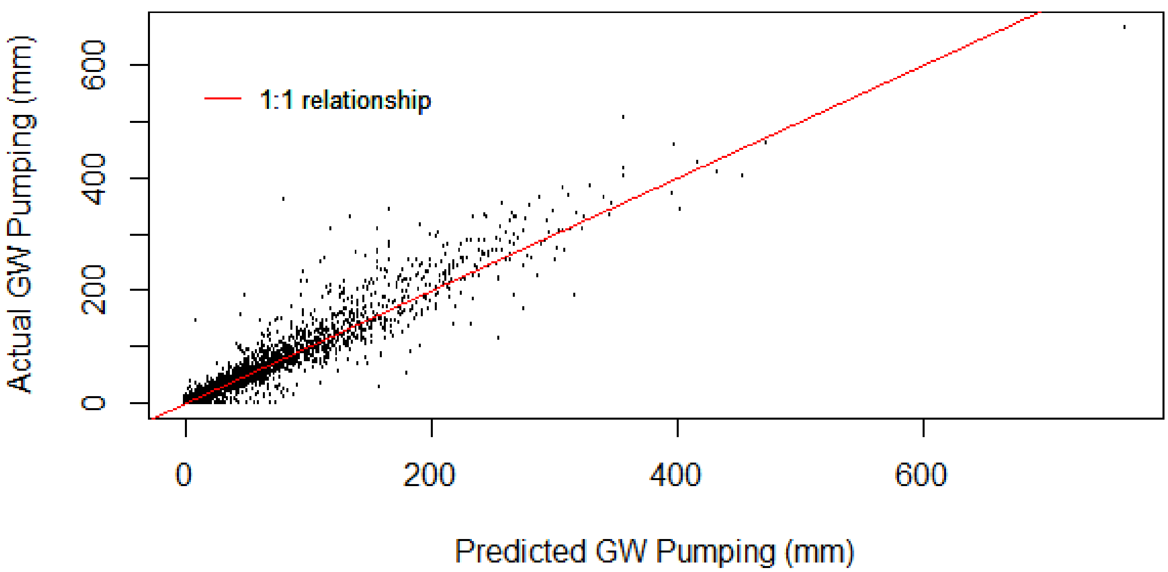


Figure 4. Scatter plot showing the actual and predicted groundwater (GW) pumping values (2014).

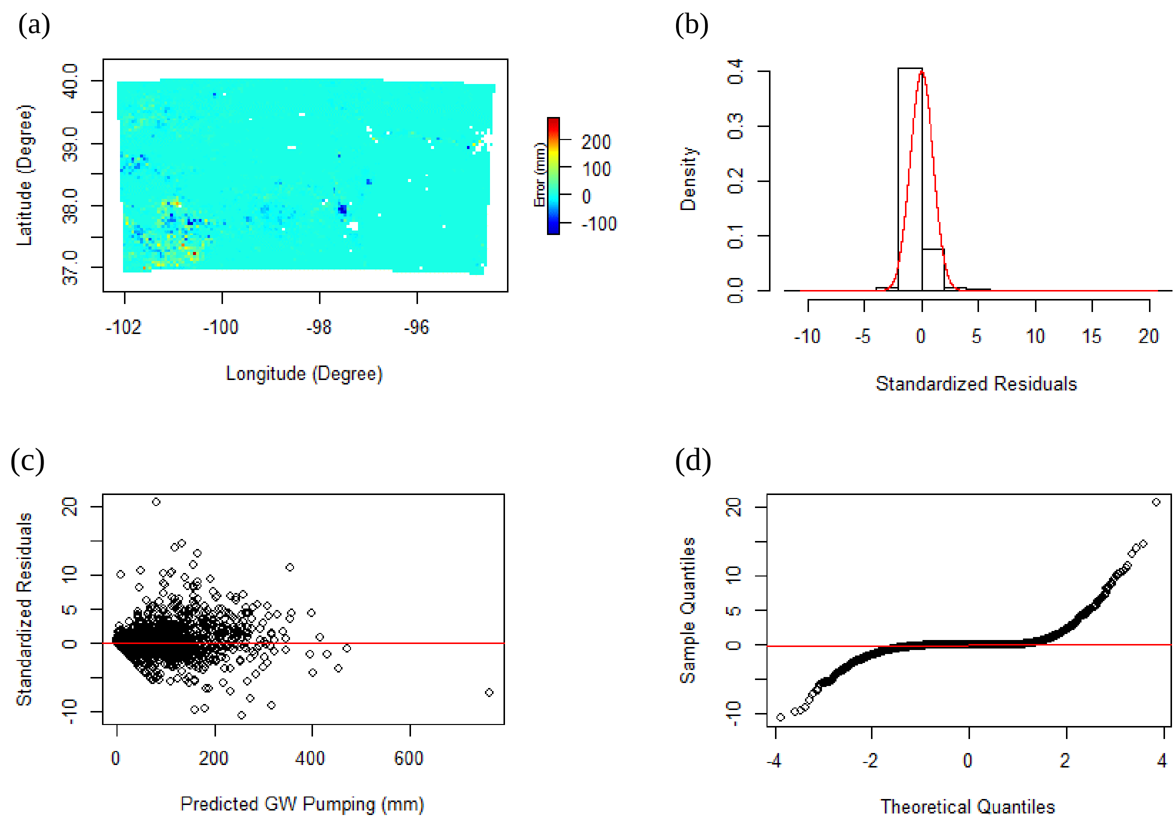


Figure 5. Residual diagnostics (actual - predicted) for 2014 showing (a) Groundwater (GW) pumping error raster, (b) Standardized residual histogram with the red line representing the Gaussian probability density function (PDF), (c) Standardized Residuals vs Predicted, and (d) Q-Q plots.

The model residual diagnostics are shown in Figure 5 (a)-(d). Figure 5 (a) shows some regions in the western portion of the study area that have pockets of high residual error. Figure 5 (b) and (c) show the standardized residuals, obtained by dividing the residuals by their standard deviation. These should lie in the [-2, 2] interval for normally distributed residuals (Hastie et al., 2013). We can see from these plots that there is a slight bias, with the residuals centered around 0 (average residual of 0.40 mm) and no significant trends in the residuals. While about 96% of the data lie within the [-2, 2] interval and Figure 5 (b) seems to follow a nearly normal distribution from the first inspection, the data do not follow a perfect normal distribution. Figure 5 (d) shows a Q-Q plot, which can be used to determine normality. When the curve follows the horizontal line (marked in red), the distribution is normal. As observed, the curve follows a straight line between quantiles -2 and 2, where most of the data lie, but the outliers have some skew.

Some deviation from normality is expected as we are possibly overfitting the random forests model which also agrees with the lower test score ($R^2 \approx 0.93$) than training score ($R^2 \approx 0.99$). However, given that there is no significant bias or trend in the residuals, we consider the model predictions to be robust.

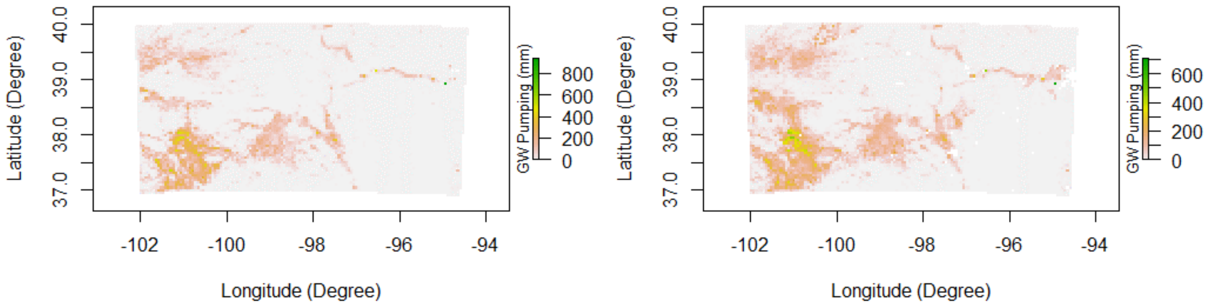
4.2 Test Case II– Year 2012

Similar to the 2014 dataset, we selected 2012 as pure validation dataset and the rest as training. Here, the model training score, $R^2 \approx 0.99$ is also similar. The error metrics for this test case are shown in Table 4.

Table 4. Error metrics (rounded to 2 decimal places) for Test Case II.

YEAR	R^2	RMSE (mm)	MAE (mm)	Normalized MAE
2002	0.98	9.01	3.41	0.14
2003	0.99	6.33	2.27	0.10
2004	0.99	5.61	1.90	0.10
2005	0.99	4.36	1.57	0.09
2006	0.99	6.27	2.17	0.10
2007	0.98	6.97	2.57	0.14
2008	0.99	5.83	1.82	0.09
2009	0.99	4.02	1.42	0.08
2010	0.99	4.64	1.65	0.08
2011	0.98	9.78	3.42	0.14
2012	**0.79**	**27.04**	**10.34**	**0.44**
2013	0.99	5.86	1.81	0.09
2014	0.99	5.12	1.62	0.08
2015	0.99	4.14	1.47	0.08
2016	0.99	4.68	1.66	0.10
2017	0.98	5.59	1.75	0.10
2018	0.99	4.70	1.67	0.10

The actual vs predicted groundwater pumping plots in Figure 6 (a) and (b) for the year 2012 indicate that even for extreme drought conditions, our approach produces satisfactory results. With $R^2 \approx 0.79$, RMSE ≈ 27.04 mm, MAE ≈ 10.34 mm, and normalized MAE ≈ 0.44 this test case has significantly lower accuracy relative to Test Case I. This is caused by a lack of training data during extreme drought years, as 2012 was the driest year in our dataset (Lin et al., 2017). Thus, by holding 2012 out of the training, our random forest model had no dataset from which to learn how groundwater pumping responds to extreme drought. We chose to hold this year out as an extreme test of our model. An R^2 of 0.79 is still quite high given the extreme nature of the hold-out year, so we consider the model performance on this year as a validation of this model's robustness.



(a) (b)

Figure 6. (a) Actual and (b) Predicted groundwater (GW) pumping rasters for 2012.

The actual vs predicted groundwater pumping values for 2012 are shown in Figure 7. As expected, there is significantly higher scatter in Test Case II relative to Test Case I. The lower $R^2 \approx 0.79$ and higher RSE ≈ 26.90 mm for Test Case II also suggest that the predicted groundwater pumping values deviate more from the actual ones.

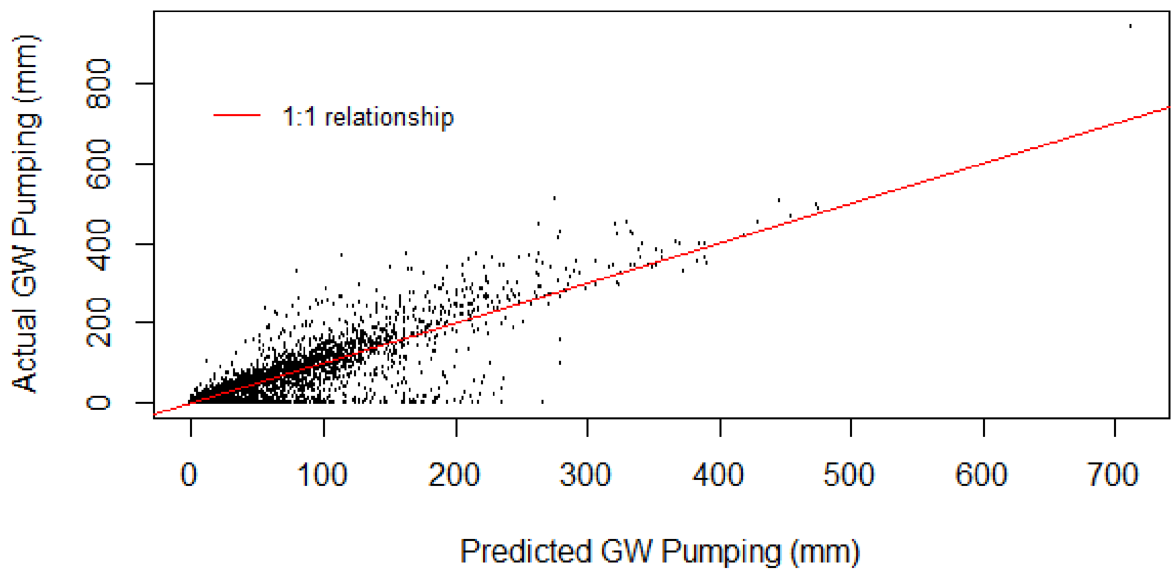


Figure 7. Scatter plot showing the actual and predicted groundwater (GW) pumping values (2012).

The model residual diagnostics are depicted in Figure 8 (a)-(d). As shown in Figure 8 (a), there are several areas in the western portion of the study area with high residuals. Figure 8 (b) and (c) show some bias (with an average residual value of -2.34 mm), and Figure 8 (d) shows that the residuals are not normally distributed though about 95% of the standardized residuals lie in the [-2, 2] interval. These observations are in concordance with the lower testing score of the random forest model, $R^2 \approx 0.79$ when compared to the high training score, $R^2 \approx 0.99$ signifying bias in the final estimates. Still, considering the extreme drought scenario, the year 2012 is an outlier with significantly higher groundwater pumping than observed in the training datasets. Given these limitations, we consider the model performance for this year to be reasonable.

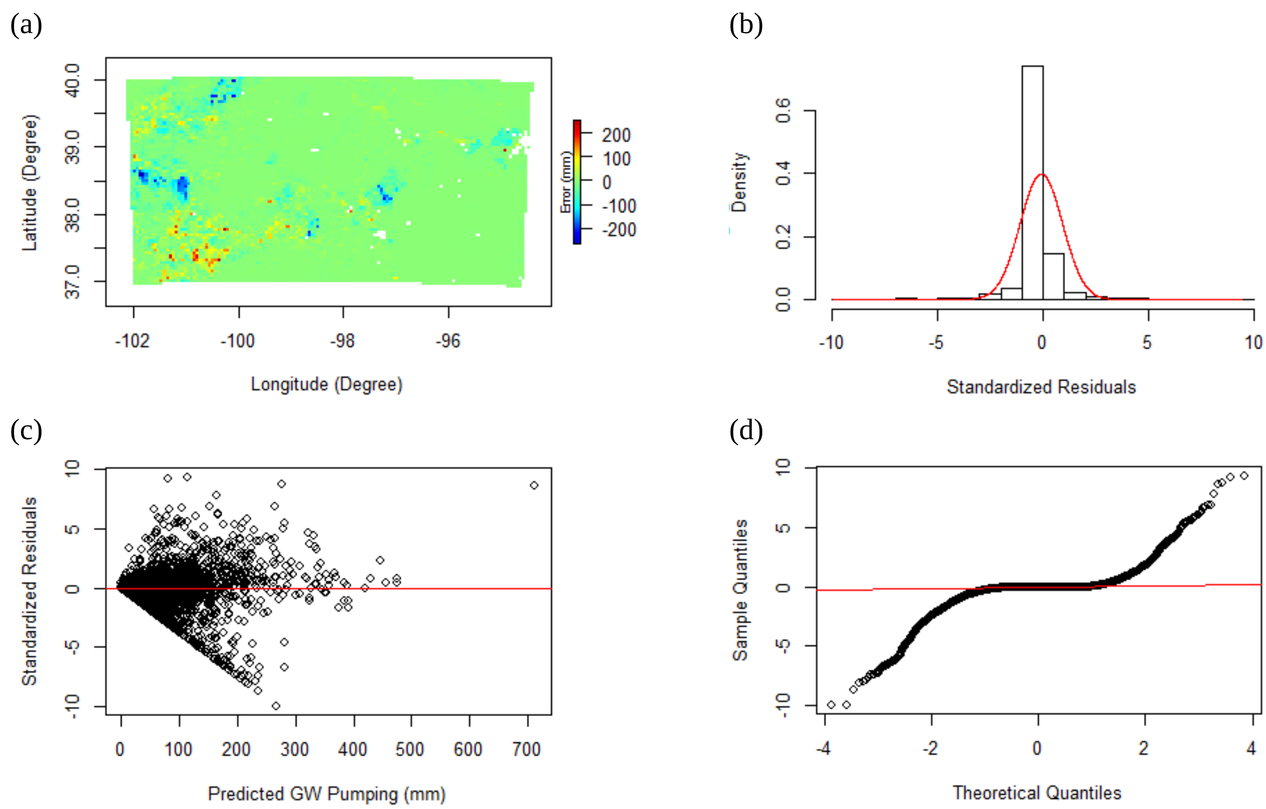


Figure 8. Residual diagnostics for 2012 showing (a) Groundwater (GW) pumping error raster, (b) Standardized residual histogram with the Gaussian PDF highlighted in red, (d) Standardized Residuals vs Predicted, and (f) Q-Q plots.

4.3 Test Case III– Using Multiple Validation Datasets

In order to further assess our model, we held the years 2011 to 2018 out of our model and used the datasets from 2002-2010 for training. Here, we only check the error metrics as given in Table 5. The overall training and testing scores for this scenario are $R^2 \approx 0.98$ and $R^2 \approx 0.78$, respectively, along with RMSE ≈ 23.50 mm, MAE ≈ 8.22 mm, and normalized MAE ≈ 0.42. These error metrics highlight that our approach performs reasonably well even when multiple years are excluded from training.

Table 5. Error metrics (rounded to 2 decimal places) for Test Case III.

YEAR	R^2	RMSE (mm)	MAE (mm)	Normalized MAE
2002	0.98	9.32	3.53	0.14
2003	0.98	7.12	2.55	0.12
2004	0.98	6.81	2.30	0.12
2005	0.99	5.00	1.86	0.10
2006	0.99	6.37	2.32	0.11
2007	0.98	7.29	2.82	0.15
2008	0.99	6.19	2.01	0.10
2009	0.99	5.49	1.94	0.11
2010	0.99	5.94	2.22	0.11
2011	**0.71**	**33.66**	**12.42**	**0.51**
2012	**0.73**	**30.80**	**12.05**	**0.51**
2013	**0.87**	**18.87**	**6.05**	**0.30**
2014	**0.90**	**16.40**	**5.60**	**0.29**
2015	**0.79**	**20.05**	**7.13**	**0.41**
2016	**0.75**	**21.58**	**7.78**	**0.47**
2017	**0.77**	**20.51**	**7.22**	**0.43**
2018	**0.76**	**20.57**	**7.48**	**0.47**

5 Discussion

We chose the random forest model developed during Test Case I as the final one for the groundwater pumping prediction, as it used the most training data and was able to learn from both wet and dry years and is most likely to be effective for future predictions. After the model evaluation phase, we observed that the feature importances in all three test cases were similar, with the most important features being the land-use classes, i.e. SW, AGRI, and URBAN. This finding is in agreement with groundwater withdrawals predominantly occurring in agricultural areas with limited surface water availability (Butler et al., 2018). More specifically, the feature importances (rounded to 2 decimal places) of SW, AGRI, URBAN, ET, and P are 0.31, 0.29, 0.22, 0.09, and 0.09 in descending order, respectively. This metric gives higher importance to spatially variable but temporally static predictors (i.e. land use) than spatio-temporal variable predictors (i.e. P and ET). The main reason for this is that groundwater withdrawals vary more spatially than they do temporally, so spatial predictors explain most of the variance in groundwater pumping values. However, this metric does not provide a complete picture of variable importance, as P and ET are critical variables for quantifying any temporal variations in groundwater, including the effect of drought or wet years, which is one of the key goals of this study. To further explore the importance of each variable, we employ partial dependence plots which indicate a plausible agreement between the model behavior and the expected heuristic relationships (Figure 9 (a)-(e)). Groundwater pumping increases with higher AGRI and ET values and decreases with increased P and SW, as would be expected. Also, the groundwater pumping varies significantly with the URBAN predictor. This could be attributed to the variability in the groundwater demand of urban areas in Kansas.

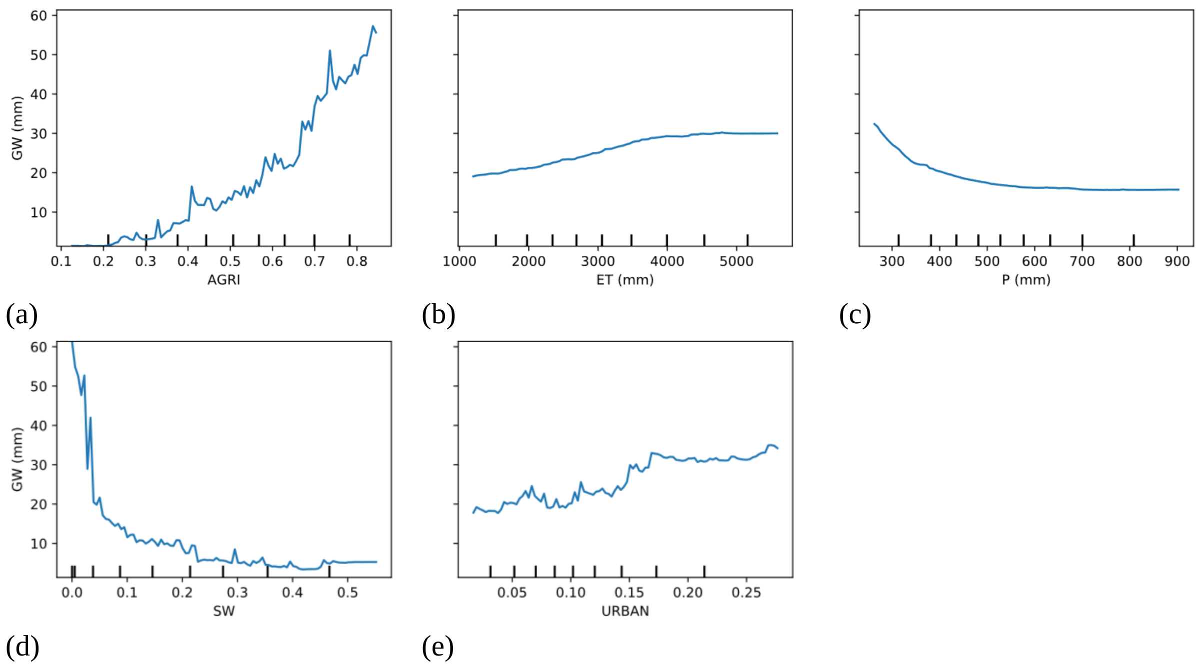


Figure 9. Partial dependence plots relating groundwater (GW) pumping values to (a) AGRI, (b) ET, (c) P, (d) SW, and (e) URBAN. These were generated using the scikit-learn library (Pedregosa et al., 2011). Here AGRI, SW, and URBAN are unitless.

While our model was reasonably accurate, it did display bias in some regions (Figure 5 (a) and 8 (a)), particularly in west-central Kansas, where we are over-predicting pumping, most notably during the major drought year of 2012 (Figure 8 (a)). This could be related to the transmissivity of the aquifer — the aquifer thickness in many of the areas with over-predicted pumping has reached a point that it can no longer support large-scale pumping for irrigated agriculture (Butler et al., 2018). Thus, these areas have high groundwater demand due to the drought, but do not have sufficient resources to meet the demand. We would expect those areas to have lower ET, but the dependence of pumping on ET resulting from the random forests may not be great enough to identify those areas.

The model we developed provides reasonable estimates of groundwater pumping in Kansas. Since the model was trained on climate (i.e., precipitation and evapotranspiration) data and land-use patterns specific to that region, the relationships it learned between those predictor variables and groundwater pumping estimates will likely hold in other regions with similar climates and land-use patterns. However, it is less transferrable to regions with significantly different climates or land use patterns.

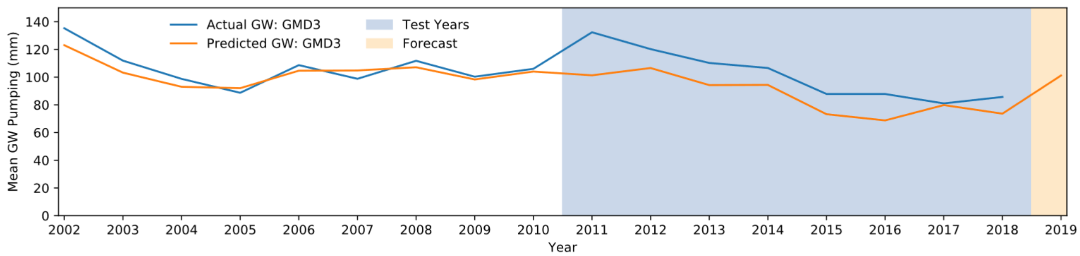
5.1 Temporal trends in groundwater pumping and total water storage

As a final step of our model evaluation, we used Test Case III (2011-2018 as pure test data) to compare (Figure 10) the groundwater pumping values (both actual and predicted) for the entire study area and three specific GMDs (GMD 3, 4, and 5) to the GRACE and GRACE-FO TWS measurements (Landerer & Swenson, 2012; Swenson, 2012). GMDs 3, 4, and 5 cover southwest, northwest, and south-central regions of Kansas (Butler et al., 2016) and are under significant stress due to heavy groundwater pumping. GMDs 3 and 4 are in a semi-arid climate, while GMD5 is in a sub-humid climate. Due to their varying climate and high water demand, these GMDs have conditions similar to a range of other aquifers that are under heavy water stress, so the performance of our model over these regions is indicative of model performance over regions with similar conditions globally.

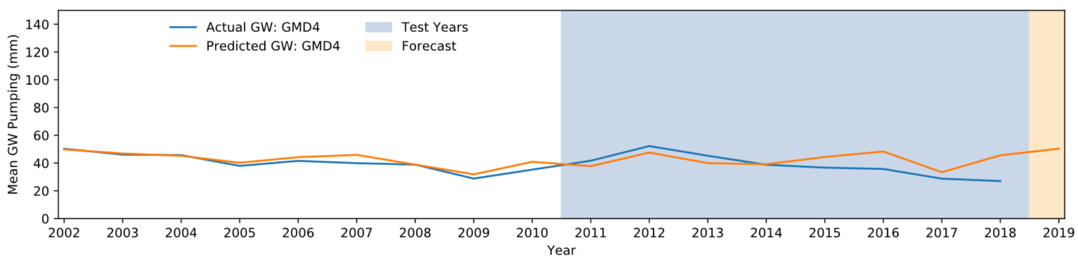
From Figure 10 (a), we observe that the groundwater pumping predictions for GMD 3 are lower than the actual values. However, the predictions for GMD 4 and 5 match quite well with the actual pumping measurements (Figure 10 (b)-(c)). GMD 3 is likely underpredicted because, in general, the actual values of pumping in GMD 3 are the highest of the study area. Random forests tends to produce conservative estimates that lean towards the mean and thus underpredict the highest estimates. However, the predictions are still reasonably close, and the other GMD in a semi-arid climate (GMD 4) matches closely with observed withdrawals. Thus, we consider the model to perform well in both semi-arid and sub-humid climates.

Since the TWS measurements in Figure 10 (e) are too coarse to compare at the GMD scale, we use the groundwater pumping estimates from the entire state of Kansas (Figure 10 (d)) for the comparison. The pattern for the mean annual groundwater pumping values for Kansas is in general agreement with that for the TWS values. Note that groundwater pumping most closely matches the change in TWS during each growing season (Spring to Fall) for any given year. Essentially, this highlights the effectiveness and robustness of our model even when we leave out several years of data from training.

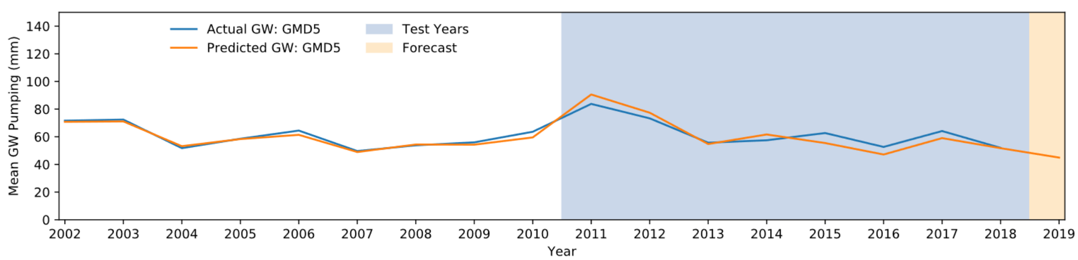
Many studies have found data from the GRACE to be a strong predictor of changes in groundwater storage at regional scales, i.e. 100s of km (Landerer & Swenson, 2012; Rodell, 2004; Rodell et al., 2007, 2009; Tiwari et al., 2009). For this reason, we initially included GRACE TWS as a predictor in our model. However, we found that our model performed better without GRACE data, so it was ultimately removed. We believe that the value of GRACE in groundwater storage changes is diminished in our study both due to its coarse resolution (400 km) (Miro & Famiglietti, 2018) relative to the resolution of our model (5 km), and also due to its correlation with precipitation, which is another strong indicator of drought in Kansas. We consider it likely that GRACE would be a more useful predictor for extending assessments to broader regions, where GRACE is a more consistent indicator of drought than precipitation, which has different normal levels for different regions.



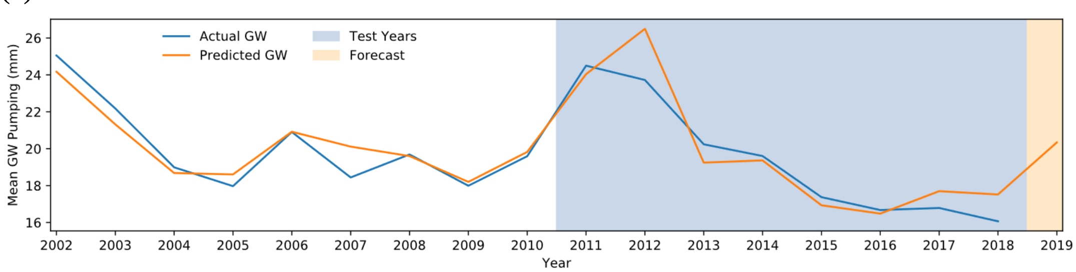
(a)



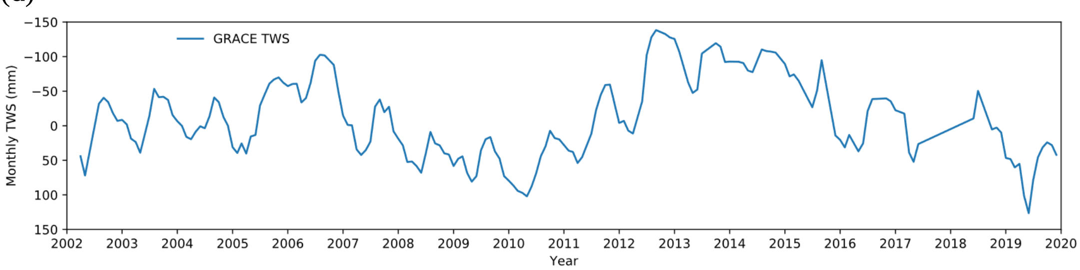
(b)



(c)



(d)



(e)

Figure 10. Actual vs predicted mean annual groundwater (GW) pumping for (a) GMD 3, (b) GMD 4. (c) GMD 5, and (d) entire Kansas. Monthly GRACE/GRACE-FO TWS between 2002 and 2019 (e) are included for comparison with the entire Kansas plot. Here, years 2011-2018 are test data (Test Case III) and forecasts are made for 2019.

6 Conclusions

In this research, we developed a new hybrid water balance and machine learning framework for predicting groundwater withdrawals at a local scale (5 km spatial resolution). This holistic approach integrates various openly available satellite products that are related in different ways to groundwater withdrawals. These include evapotranspiration (MODIS), precipitation (PRISM), and land-cover data (USDA-NASS). We deployed a machine learning framework using random forests to build a self-learnable hybrid water balance method that incorporates these datasets along with in-situ groundwater extraction data (Wilson, 2019). Our workflow is fully automated and has been implemented using open-sourced or freely available software tools and libraries.

A thorough model assessment showed that our proposed approach works satisfactorily with predictions having high R^2 values and low RMSE and MAE. In this work, we considered three test scenarios to validate our method, two of which included the extreme drought period of 2012. Moreover, the random forest feature importances agree with the expected dependencies of groundwater withdrawals.

As climate change, dietary shifts, and population growth increase the global stress on groundwater resources, it is critical to develop sustainable groundwater management practices, yet the volume of groundwater withdrawals, a key quantity for such efforts, is poorly constrained in many regions. Although one of us has repeatedly called for greater monitoring of groundwater withdrawals (Butler et al., 2016, 2018, 2020), that has proven challenging in many areas. This method enables water managers to predict groundwater withdrawals from anthropogenic and climate drivers— such knowledge could lead to proactive implementation of sustainable solutions related to groundwater withdrawal practices.

Acknowledgments

We would like to acknowledge all the open-source software and data communities for making their resources publicly available. Particularly, we would like to thank NASA/JPL (https://grace.jpl.nasa.gov/data-analysis-tool/, https://doi.org/10.5067/MODIS/MOD16A2.006), USDA (https://nassgeodata.gmu.edu/CropScape/), and the PRISM group (http://www.prism.oregonstate.edu/) for providing all the necessary remote sensing datasets. The work of the third author was supported, in part, by the Kansas Water Plan under the Ogallala-High Plains Aquifer Assessment Program (OHPAAP) and the United States Department of Agriculture (USDA) and the United States National Science Foundation (NSF) under USDA-NIFA/NSF INFEWS subaward RC108063UK. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the OHPAAP, USDA, or NSF. We thank Brownie Wilson of the Kansas Geological Survey for helping us access the WIMAS database (http://hercules.kgs.ku.edu/geohydro/wimas/). The entire project code along with the preprocessed datasets are available on https://github.com/montimaj/HydroMST. Finally, we are

grateful to our colleagues and families for their continuous motivation and support. We confirm that there is no conflict of interest among the authors of this manuscript.

References

Becker, R., Koppa, A., Schulz, S., Usman, M., aus der Beek, T., & Schüth, C. (2019). Spatially distributed model calibration of a highly managed hydrological system using remote sensing-derived ET data. *Journal of Hydrology*, *577*, 123944. https://doi.org/10.1016/j.jhydrol.2019.123944

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Boryan, C., Yang, Z., Mueller, R., & Craig, M. (2011). Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto Int.*, *26*(5), 341–358. https://doi.org/10.1080/10106049.2011.562309

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees* (1st ed.). CRC Press.

Butler, J. J., Bohling, G. C., Whittemore, D. O., & Wilson, B. B. (2020). A roadblock on the path to aquifer sustainability: underestimating the impact of pumping reductions. *Environmental Research Letters*, *15*(1), 014003. https://doi.org/10.1088/1748-9326/ab6002

Butler, J. J., Whittemore, D. O., Reboulet, E., Knobbe, S., Wilson, B. B., & Bohling, G. C. (2019). *High Plains Aquifer Index Well Program: 2018 Annual Report*. Kansas Geological Survey. http://www.kgs.ku.edu/Hydro/Publications/2019/OFR19_19/OFR2019-19.pdf

Butler, J. J., Whittemore, D. O., Wilson, B. B., & Bohling, G. C. (2016). A new approach for assessing the future of aquifers supporting irrigated agriculture. *Geophysical Research Letters*, *43*(5), 2004–2010. https://doi.org/10.1002/2016GL067879

Butler, J. J., Whittemore, D. O., Wilson, B. B., & Bohling, G. C. (2018). Sustainability of aquifers supporting irrigated agriculture: a case study of the High Plains aquifer in Kansas. *Water International*, *43*(6), 815–828. https://doi.org/10.1080/02508060.2018.1515566

Cannon, F., Ralph, F. M., Wilson, A. M., & Lettenmaier, D. P. (2017). GPM Satellite Radar Measurements of Precipitation and Freezing Level in Atmospheric Rivers: Comparison With Ground-Based Radars and Reanalyses. *Journal of Geophysical Research: Atmospheres*, *122*(23), 12,747-12,764. https://doi.org/10.1002/2017JD027355

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M. H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., … Kerr, Y. (2016). Assessment of the SMAP Passive Soil Moisture Product. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4994–5007. https://doi.org/10.1109/TGRS.2016.2561938

Chaussard, E., & Farr, T. G. (2019). A New Method for Isolating Elastic From Inelastic Deformation in Aquifer Systems: Application to the San Joaquin Valley, CA. *Geophysical Research Letters*, *46*(19), 10800–10809. https://doi.org/10.1029/2019GL084418

Deines, J. M., Kendall, A. D., & Hyndman, D. W. (2017). Annual Irrigation Dynamics in the U.S. Northern High Plains Derived from Landsat Satellite Data. *Geophysical Research Letters*, *44*(18), 9350–9360. https://doi.org/10.1002/2017GL074071

Deines, J. M., Schipanski, M. E., Golden, B., Zipper, S. C., Nozari, S., Rottler, C., Guerrero, B., & Sharda, V. (2020). Transitions from irrigated to dryland agriculture in the Ogallala Aquifer: Land use suitability and regional economic impacts. *Agricultural Water Management*, *233*, 106061. https://doi.org/10.1016/j.agwat.2020.106061

Erban, L. E., Gorelick, S. M., Zebker, H. A., & Fendorf, S. (2013). Release of arsenic to deep groundwater in the Mekong Delta, Vietnam, linked to pumping-induced land subsidence.

Proceedings of the National Academy of Sciences, *110*(34), 13751–13756. https://doi.org/10.1073/pnas.1300503110

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*(3), 1–4. https://doi.org/10.1029/2010GL046442

Faunt, C. C. (Ed.). (2009). Groundwater availability of the Central Valley Aquifer, California. In *U.S. Geological Survey Professional Paper 1766*. https://doi.org/10.3133/pp1766

Frappart, F., & Bourrel, L. (Eds.). (2018). *The Use of Remote Sensing in Hydrology*. MDPI. https://doi.org/10.3390/books978-3-03842-910-4

Galloway, D. L., & Burbey, T. J. (2011). Review: Regional land subsidence accompanying groundwater extraction. *Hydrogeol. J.*, *19*(8), 1459–1486. https://doi.org/10.1007/s10040-011-0775-5

GDAL/OGR contributors. (2019). *GDAL/OGR Geospatial Data Abstraction software Library*. Open Source Geospatial Foundation. https://gdal.org

GeoPandas developers. (2019). *GeoPandas 0.6.0*. http://geopandas.org/

Gillies, S. (2013). *Rasterio: geospatial raster I/O for Python programmers*. Mapbox. https://github.com/mapbox/rasterio

Gleeson, T., Villholth, K., Taylor, R., Perrone, D., & Hyndman, D. (2019). Groundwater: a call to action. *Nature*, *576*(7786), 213–213. https://doi.org/10.1038/d41586-019-03711-0

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Hashemi, H., Nordin, M., Lakshmi, V., Huffman, G. J., & Knight, R. (2017). Bias Correction of Long-Term Satellite Monthly Precipitation Product (TRMM 3B43) over the Conterminous United States. *Journal of Hydrometeorology*, *18*(9), 2491–2509. https://doi.org/10.1175/JHM-D-17-0025.1

Hastie, T., Tibshirani, R., & Friedman, J. (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer New York. https://books.google.com/books?id=yPfZBwAAQBAJ

Hoffmann, J., Zebker, H. A., Galloway, D. L., & Amelung, F. (2001). Seasonal subsidence and rebound in Las Vegas Valley, Nevada, observed by Synthetic Aperture Radar Interferometry. *Water Resources Research*, *37*(6), 1551–1566. https://doi.org/10.1029/2000WR900404

IU Digital Science Center. (2013). *Harp Random Forests*. https://dsc-spidal.github.io/harp/docs/examples/rf/

Jones, E., Oliphant, E., & Peterson, P. (2001). *SciPy: Open Source Scientific Tools for Python*. http://www.scipy.org/

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*(4). https://doi.org/10.1029/2011WR011453

Leidner, A. K., & Buchanan, G. M. (Eds.). (2018). *Satellite Remote Sensing for Conservation Action*. Cambridge University Press. https://doi.org/10.1017/9781108631129

Lin, X., Harrington, J., Ciampitti, I., Gowda, P., Brown, D., & Kisekka, I. (2017). Kansas Trends and Changes in Temperature, Precipitation, Drought, and Frost-Free Days from the 1890s to 2015. *Journal of Contemporary Water Research & Education*, *162*(1), 18–30. https://doi.org/10.1111/j.1936-704X.2017.03257.x

MardanDoost, B., Brookfield, A. E., Feddema, J., Sturm, B., Kastens, J., Peterson, D., & Bishop, C. (2019). Estimating irrigation demand with geospatial and in-situ data: Application to the high plains aquifer, Kansas, USA. *Agricultural Water Management*, *223*, 105675. https://doi.org/10.1016/j.agwat.2019.06.010

Margat, J., & van der Gun, J. (2013). *Groundwater around the World: A Geographic Synopsis* (1st ed.). CRC Press. https://www.crcpress.com/Groundwater-around-the-World-A-Geographic-Synopsis/Margat-Gun/p/book/9781138000346#googlePreviewContainer

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 51–56).

https://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf
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*(1), 143. https://doi.org/10.3390/rs10010143
Moeck, C., von Freyberg, J., & Schirmer, M. (2018). Groundwater recharge predictions in contrasted climate: The effect of model complexity and calibration period on recharge rates. *Environmental Modelling & Software*, *103*, 74–89. https://doi.org/10.1016/j.envsoft.2018.02.005
Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Anderson, M. C., & Hain, C. (2018). Groundwater Withdrawals Under Drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer. *Water Resources Research*, *54*(8), 5282–5299. https://doi.org/10.1029/2017WR022178
Oliphant, T. E. (2006). *A Guide to NumPy*. Trelgol Publishing. https://books.google.com/books?id=fKulSgAACAAJ
Olson, R. (2015). *Feature order affects the performance of tree-based classifiers*. Github. https://github.com/scikit-learn/scikit-learn/issues/5394
Ozdogan, M., & Gutman, G. (2008). A new methodology to map irrigated areas using multi-temporal MODIS and ancillary data: An application example in the continental US. *Remote Sensing of Environment*, *112*(9), 3520–3537. https://doi.org/10.1016/j.rse.2008.04.010
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn:Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf
QGIS Development Team. (2019). *QGIS Geographic Information System*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org/
R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.r-project.org/
Rodell, M. (2004). Basin scale estimates of evapotranspiration using GRACE and other observations. *Geophysical Research Letters*, *31*(20), L20504. https://doi.org/10.1029/2004GL020873
Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro, J., & Wilson, C. R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, *15*(1), 159–166. https://doi.org/10.1007/s10040-006-0103-7
Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*(7258), 999–1002. https://doi.org/10.1038/nature08238
Running, S., & Mu, Q. (2015). *University of Montana and MODAPS SIPS - NASA. (2015). MOD16A2 MODIS/Terra Evapotranspiration 8-day L4 Global 500m SIN Grid.* NASA LP DAAC. http://doi.org/10.5067/MODIS/MOD16A2.006
Schneider, S. H., Root, T. L., & Mastrandrea, M. D. (Eds.). (2011). Water Resources. In *Encyclopedia of Climate and Weather* (2nd ed.). Oxford University Press. https://doi.org/10.1093/acref/9780199765324.001.0001
Seibert, J., Staudinger, M., & van Meerveld, H. J. (2019). *Validation and Over-Parameterization—Experiences from Hydrological Modeling* (pp. 811–834). https://doi.org/10.1007/978-3-319-70766-2_33
Smajgl, A., Ward, J., & Pluschke, L. (2016). The water–food–energy Nexus – Realising a new paradigm. *Journal of Hydrology*, *533*, 533–540. https://doi.org/10.1016/j.jhydrol.2015.12.033
Smidt, S. J., Haacker, E. M. K., Kendall, A. D., Deines, J. M., Pei, L., Cotterman, K. A., Li, H., Liu, X., Basso, B., & Hyndman, D. W. (2016). Complex water management in modern agriculture: Trends in the water-energy-food nexus over the High Plains Aquifer. *Science of The Total Environment*, *566–567*, 988–1001. https://doi.org/10.1016/j.scitotenv.2016.05.127
Smith, R., 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*(3), 2133–2148. https://doi.org/10.1002/2016WR019861

Smith, R., Knight, R., & Fendorf, S. (2018). Overpumping leads to California groundwater arsenic threat. *Nat. Commun.*, *9*(1), 2089. https://doi.org/10.1038/s41467-018-04475-3

Swenson, S. C. (2012). *GRACE MONTHLY LAND WATER MASS GRIDS NETCDF RELEASE 5.0. Ver. 5.0*. PO.DAAC, CA, USA. https://doi.org/10.5067/TELND-NC005

Tamayo-Mas, E., Bianchi, M., & Mansour, M. (2018). Impact of model complexity and multi-scale data integration on the estimation of hydrogeological parameters in a dual-porosity aquifer. *Hydrogeology Journal*, *26*(6), 1917–1933. https://doi.org/10.1007/s10040-018-1745-y

Tiwari, V. M., Wahr, J., & Swenson, S. (2009). Dwindling groundwater resources in northern India, from satellite gravity observations. *Geophysical Research Letters*, *36*(18), L18401. https://doi.org/10.1029/2009GL039401

USDA-NASS. (2015). *Published crop-specific data layer*. USDA National Agricultural Statistics Service Cropland Data Layer. https://nassgeodata.gmu.edu/CropScape/

Wilson, B. B. (2019). *Water Information Management and Analysis System (WIMAS) for the Web*. Kansas Geological Survey. http://hercules.kgs.ku.edu/geohydro/wimas/query_setup.cfm

Yi, Z., Zhao, H., & Jiang, Y. (2018). Continuous Daily Evapotranspiration Estimation at the Field-Scale over Heterogeneous Agricultural Areas by Fusing ASTER and MODIS Data. *Remote Sens.*, *10*(11), 1694. https://doi.org/10.3390/rs10111694