Large-scale groundwater monitoring in Brazil assisted with satellite-based artificial intelligence techniques

Clyvihk Renna Camacho[1,2], Augusto Getirana[3,4], Otto Corrêa Rotunno Filho[2], Maria Antonieta A. Mourão[1]

1 Geological Survey of Brazil, Belo Horizonte, Brazil

2 Civil Engineering Department, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

3 Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, United States

4 Science Applications International Corporation, Greenbelt, MD, United States

Abstract – Here, we develop and test an artificial intelligence (AI)-based approach to monitor major Brazilian aquifers. The approach combines Gravity Recovery and Climate Experiment (GRACE) data and ground-based hydrogeological measurements from Brazil's Integrated Groundwater Monitoring Network at hundreds of wells distributed in twelve aquifers across the country. We tested model ensembles based on three AI approaches: Extreme Gradient Boost, Light Gradient Boosting Model and CatBoost, followed by a Linear Regression (LR) step. The approach is further boosted with wavelet and seasonal decomposition processes applied to GRACE data. To determine the AI-based model's sensitivity to data availability, we propose four experiments combining hydrogeological measurements from different aquifers. Groundwater storage estimates from the Global Land Data Assimilation System (GLDAS) are used as benchmark. A sensitivity analysis shows that the LR-based model ensemble is the best suited and to reproduce groundwater storage change in all studied Brazilian aquifers. Results show that the proposed approach outperforms GLDAS in all experiments, with an RMSE value of 2.68cm for the experiment that covers all monitored wells in Brazil. GLDAS resulted in RMSE=6.76cm. Using our AI model outputs, we quantified the groundwater storage change of two major aquifers, Urucuia and Bauru-Caiuá, over the past two decades: -31km^3 and -6km^3, respectively. Water loss is driven by a prolonged drought across most of the country and intensification of groundwater pumping for irrigation. This study demonstrates that combining satellite data and AI can be a cost-effective alternative to monitor poorly equipped aquifers at the continental scale, with possible global replicability.

Key Points

- An Artificial Intelligence (AI)-based model was built to monitor groundwater in Brazilian aquifers using satellite gravimetry data
- AI-based groundwater changes outperformed Global Land Data Assimilation System (GLDAS) estimates in all proposed experiments
- Results show that satellite-based AI techniques can be an effective solution for groundwater monitoring in poorly equipped regions

1. Introduction

Proper aquifer monitoring at different scales faces enormous difficulties related to geological factors such as the complexity and diversity of formations and their corresponding structures. Difficulties are accentuated by complexities related to hydraulic properties of aquifers, recharge zones, groundwater exploration, land use and land cover change, as well as meteorological and climate variability. There is definitely a demand for global and operational hydrogeological monitoring, knowing that groundwater is the largest unfrozen freshwater stock on the planet and tightly connected to surface water, reservoir and lakes (Condon et al., 2021). In 2002, over 1.5 billion people were estimated to be directly supplied by groundwater (Alley *et al.*, 2002). This number has risen to 2 billion people in 2020 (UNESCO, 2022). It is estimated that 43% of the total water used in irrigation has underground origin (Siebert et al., 2010). Countries such as the United States and India use approximately 25% and 40% of groundwater resources to supply their respective needs (Getirana et al., 2021), resulting in significant aquifer depletions (e.g., Rodell *et al.* 2018; Nie *et al.* 2019). In Brazil, about 57% of its municipalities have groundwater supply to some extent (IBGE, 2020). Human activities and climate change have been changing the hydrological cycle, which, in turn, may have an impact on aquifers worldwide (Chagas et al., 2022; Getirana et al., 2021, 2022; Richey et al., 2015; Rodell et al., 2018). Therefore, it is essential to understand groundwater spatiotemporal dynamics to ensure its sustainable use, enabling an optimal management that can affect the various sectors of society such as agriculture, power generation and water supply.

Groundwater monitoring networks have been based on observation wells associated with the creation of conceptual and mathematical models (Condon et al., 2021). In recent decades, various regional and global hydrogeological models have been developed and reported in literature (Condon et al., 2021; Gleeson et al., 2021; de Graaf et al., 2015, 2017; Kollet et al., 2018; Maxwell et al., 2015; Reinecke et al., 2019). In addition, large-scale hydrological model outputs, such as those produced by the Global Land Data Assimilation System (GLDAS; Rodell et al., 2003), can be used as a tool to approximate the hydrogeological behavior in regions with a lack of monitoring. Among GLDAS models, the Catchment land surface model (CLSM; Koster et al., 2000) uses atmospheric boundary conditions associated with a partition of the Earth's surface into defined topographic basins, modeling hydrological processes and explicitly represents the groundwater dynamics in a simplified way (Getirana et al., 2020; Li et al., 2019). Such models are robust in their structure and have the ability to provide the behavior of surface water and groundwater at continental and global scales. However, these models still need to be adjusted for optimal regional use (Getirana et al., 2020).

A new frontier has been opened for the study of groundwater by data provided by the Gravity Recovery and Climate Experiment (GRACE) (Tapley et al., 2004) and GRACE Follow On (GRACE-FO) missions, which measure changes on global gravitational forces. Among those changes, there are those promoted by the water cycle. They can be mapped by satellites and later converted into terrestrial water storage (TWS) variability. Several studies have used data from GRACE missions to capture regional groundwater behavior and to assess measurements related to groundwater levels (Andrew et al., 2017; Frappart and Ramillien, 2018; Getirana et al., 2020; Scanlon et al., 2012, 2018). In Brazil, the use of GRACE data to understand the water behavior can be found in recent studies (Getirana, 2016; Hu et al., 2017; Gonçalves et al., 2020; Getirana et al., 2021).Li et al., (2019)assimilated GRACE data into a hydrological model globally and analyzed groundwater variations, comparing model results to *in situ* observations.

There is a clear contribution of GRACE data assimilation (DA) into hydrological models in the representation and prediction of hydrological processes (Getirana, Jung, et al., 2020; Getirana, Rodell, et al., 2020; Girotto et al., 2017; Jung et al., 2019; Kumar et al., 2016; Zaitchik et al., 2008). Nevertheless, new tools based on the so-called artificial intelligence (AI) algorithms have also proved to be very efficient for the pattern recognition of groundwater behavior worldwide (Afzaal et al., 2020; Huang et al., 2019; Iqbal et al., 2021; Lähivaara et al., 2019; Ren et al., 2021; Tao et al., 2022; Zhang et al., 2020). AI algorithms, associated with GRACE-based TWS variations can be of great value in the survey of aquifers. Groundwater studies using AI and GRACE data have been carried out for some years (Gemitzi & Lakshmi, 2017; Sun, 2013; Sun et al., 2019). Wave decomposition methods are also very useful in hydrological studies for flow prediction, seasonal analysis or even hydrogeological studies (Ashraf et al., 2022; Basu et al., 2022; Erkyihun et al., 2016; Qi & Neupauer, 2008). Hybrid use of AI and wavelet decomposition techniques turned out to be an important and active research area, resulting in more accurate models in water resources applications, due to its great ability to discriminate non-stationary and nonlinear trends that occur at different scales in groundwater time series (e.g., Tao et al., 2022). For example, Andrew et al. (2017)

presented the possibility of disaggregating GRACE data using wavelets as a viable path to study groundwater under different observational spatiotemporal scales.

Brazil's Integrated Groundwater Monitoring Network (RIMAS), conceived and built by the Geological Service of Brazil, initiated in 2010 and is currently composed of 409 wells, monitoring 24 aquifers across the country. The distribution of wells across the monitoring network is not homogeneous, leading to constraints in monitoring the spatiotemporal variability of the nation's aquifers. Also, only porous, free or semi-confined aquifers have been monitored by RIMAS. Such a sparse network is substantially less dense than those found in other large countries, such as the U.S. and India, which have more than 16,000 and 22,000 wells, respectively (Getirana et al., 2021). That leads to a limited groundwater monitoring in Brazil, restricting our knowledge on their dynamics and limiting the management and optimized use of the nation's aquifers. The absence of data also limits the development and parameterization of hydrological and hydrogeological models to monitor Brazil's water resources, resulting in inaccurate water flow and storage calculations, affecting various sectors of society such as agriculture, energy generation and domestic water supply (Getirana et al., 2021).

Considering the limitations described above, this work presents a methodology that combines point-based *in situ* groundwater measurements and spatially distributed satellite-based TWS, in addition to wavelet and seasonal decomposition techniques and AI as tools to understand the behavior of large aquifers in Brazil. Complementarily, a trend analysis was applied allowing us to estimate the water storage change in two major Brazilian aquifers: the Urucuia and Bauru-Caiuá. The main advantage of the proposed methodology is the use of a hybrid model (wave decomposition + ensemble model) with the application of four different AI techniques. Science questions addressed in this paper focus on (1) whether large-scale groundwater spatiotemporal variability can be estimated using GRACE data in an AI framework; (2) how accurate these estimates are compared to existing model-based estimates; and (3) based on such estimates, how major Brazilian aquifers have changed in the past two decades.

There have been a few attempts to simulate global-scale groundwater dynamics (Gleeson et al., 2021; de Graaf et al., 2015; Li et al., 2019; Maxwell et al., 2015; Reinecke et al., 2019). These models vary as a function of the numerical representation of physical processes and data assimilated into the modeling system. Here, GLDAS-based groundwater simulations are used as the benchmark to determine the potential of the proposed methodology. GLDAS simulations are those derived from CLSM with the assimilation of GRACE data (Li et al., 2019). We considered such simulations as our benchmark because they have been comprehensively evaluated globally, are widely used and easy operational availability. Also, it is currently the only temporally continuous and spatially distributed groundwater product freely and routinely available over Brazil. We expect that the proposed methodology can be used for the management of large Brazilian aquifers, in addition to enabling the monitoring of groundwater in places where monitoring networks are precarious, inexistent, or with heterogeneous hydrogeological and climatic conditions.

2. Case study and Datasets

2.1. *In situ* data from aquifers

Aquifers monitored by RIMAS total 2.84 million km^2, or 34% of the Brazilian territory. Their sizes vary from 884km^2 (Missão Velha) to 774,385km^2 (Içá), with monitoring coverage varying from 3 (Ronuro) to 71 (Urucuia) wells. The RIMAS monitoring began in 2010 and its spatial distribution is shown in Fig. 1. Groundwater measurements used here spans from August 2010 to June 2020. The monitoring network density varies significantly, with Missão Velha being the aquifer with the highest density of wells (180km^2/well) and Içá with the lowest density (130,000km^2/well). Such a heterogeneous density is mostly explained by the way the network is installed, based on the following criteria: sedimentary aquifers, socioeconomic importance, water use for public supply, natural vulnerability and risk aspects, spatial representativeness of the aquifer and existence of wells for monitoring (Mourão, 2009). Effective porosity (n_e) values for each aquifer were estimated based on available data found in the literature (please refer to Supporting Table S1 for a full list), varying from 0.03 in the Cabeças aquifer to 0.18 in the Alter do Chão aquifer.

RIMAS is designed based on wells equipped with automatic level meters collecting data at the hourly step, which are subsequently subjected to consistency, treatment, and availability processes (http://rimasweb.cprm.gov.br/layout/). The estimated error of the measurements is the minimum resolution of the equipment, which varies between 0.01cm and 1.5cm. More specifically, porous, free, semi-confined and wells in areas of crystalline rocks are addressed in this study, focusing on the responses that different lithologies might produce and how that information could be translated into the building AI model approach we are developing in-here. Geological and hydrogeological description of the aquifers monitored by RIMAS can be found at http://rimasweb.cprm.gov.br/layout/apresentacao.php.

The RIMAS dataset is available at the hourly time step. Monthly groundwater level change (dh_i) at each well was computed by first converting the time series to monthly means, then subtracting the value in the previous month (η_{i-1}) from the subsequent one (η_i), as follows:

$$dh_i = \eta_i - \eta_{i-1} \quad (1)$$

where i stands for months of the time series. Each dh_i value was then subtracted by their respective long-term mean, and multiplied by their corresponding aquifer n_e value, resulting in the time series used as input in our approach, named hereafter as ΔGWS_{OBS} [cm]. n_e was used to convert groundwater level measurements to vertical water storage column. Equation 2 shows the calculation of ΔGWS_{OBS} for each month:

$$\Delta GWS_{OBS}(i) = [dh_i - \mathrm{mean}(dh)].n_e \quad (2)$$

Importantly, the RIMAS data observed here represent the variation in groundwater storage (ΔGWS).

2.2. Terrestrial water storage

The Gravity and Recovery and Climate Experiment (GRACE) satellites mapped Earth's gravity field from April 2002. Temporal variations in gravity can be used to infer changes in total terrestrial water storage (TWS; Li et al., 2019). GRACE RL06 Mascon data, processed by the Center of Space Research (CSR; Save et al., 2016), is retrieved at 0.25-degre spatial resolution and monthly time step from April, 2002 to present with gaps throughout the period. Initially, monthly GRACE-based TWS values had uncertainties estimated at 1cm for areas equal to or greater than 400,000km^2 (Swenson et al., 2003). However, such estimates had a significant improvement, as described in Ditmar (2018), obtaining more refined results for TWS approximations at 0.25-degree spatial resolution. Such Mascon-based products have lower errors compared to spherical harmonics (Rowlands et al., 2010). Even maintaining a resolution limited by the nature of the GRACE data and uncertainties in the TWS of 1cm, these estimates allow a more detailed study of hydrological and hydrogeological basins with dimensions smaller than those indicated by Save et al. (2016) as demonstrated in Melati et al. (2019) and Gonçalves et al. (2020), both carried out in Brazil.

2.3. GLDAS-based groundwater

The CLSM with GRACE-DA was chosen among the different GLDAS products for presenting an explicit and more accurate representation groundwater storage (GWS). CLSM is a state-of-the-art energy and water balance model of the Earth's surface, designed for use in models of global earth systems. The model simulates a dynamic water table with a spatial distribution related to the topography of the basin (Bechtold et al., 2019). CLSM does not model variations in the water table. Instead, GWS is derived from the subtraction of the water stored in the root zone from that stored in the vertical soil profile, whose capacity is determined by the CLSM bedrock depth parameter. The model returns GWS anomalies, among other hydrological variables (Li et al., 2019). CLSM outputs are available daily at 0.25°. CLSM-based GWS (GWS_{CLSM}) was also converted to variations in storage (ΔGWS_{CLSM}), following the approach applied to RIMAS. Details about the model configuration and global evaluation can be found in Li et al. (2019).

3. Methodology

Briefly, the methodology follows four steps. First, wave decomposition (wavelet and seasonal) on the TWS values for Brazil. Second, interpolation of the values obtained by the wavelets for the original time scale. Third, the decomposition results are associated with the RIMAS measurements, according to latitude, longitude and time. Hydrogeological characteristics derived from the Hydrogeological Map of Brazil (HMB) (Diniz et al., 2014) are also associated with RIMAS wells according to latitude and longitude. Finally, the dataset is inserted into an AI model to approximate the groundwater storage (ΔGWS_{OBS}) values obtained by the RIMAS wells (Fig. 2). The input data in the model are GRACE-based TWS, TWS decompositions (wavelet and seasonal) and HMB's hydrogeological description. ΔGWS estimates from the CLSM GRACE-DA (ΔGWS_{CLSM}) are used as the benchmark. Here, we considered the nearest CLSM grid point of each well.

A sensitivity analysis of AI models was conducted, looking for models that best represent the groundwater storage change from GRACE satellite observations. The metrics used in selecting the models are described in Section 3.2. The sensitivity test evaluated 40 different architectures for twelve AI models (Fig. 3) and the architecture with the best calibration metrics was selected for further evaluation.

GRACE-based TWS can be decomposed into five main components: snow water equivalent, canopy interception, soil moisture, surface water and groundwater. Previous studies have attempted to decompose GRACE signals using hydrological models (e.g., Getirana et al., 2017; Scanlon et al., 2018). Here, we assume that GRACE components can be disaggregated and studied through data decomposition techniques.

Wavelet transform (WT) is a technique that has proven to be effective for capturing nonlinear relationships in time series (Tao et al., 2022). It removes noise in the data and allows a better performance in AI models. Several hydrogeological studies have combined these techniques and demonstrated the ability to approximate variations in groundwater levels from different data sources (Barzegar et al., 2017; Ebrahimi & Rajaee, 2017; Khalil et al., 2015; Moosavi et al., 2013, 2014; Yosefvand & Shabanlou, 2020; Zare & Koch, 2018). Here, GRACE-based TWS was decomposed using two techniques: the wavelet transform and the seasonal decomposition.

WT is a mathematical tool to decompose functions hierarchically, and can be considered as a technique for transforming a signal, sampled in the time domain, into a frequency-scaled domain, defining different components of the signal frequency spectrum (Stollnitz et al., 1995). WT consists of approximating a function by a linear combination of basic functions (also called wavelets), obtaining a representation of the original function. The application of wavelets does not necessarily require the stationarity of the time series as a prerequisite, being appropriate for the analysis of irregularly distributed and extreme events (Torrence & Compo, 1998). For non-continuous functions, the use of the discrete wavelet decomposition (DWT) is recommended (Daubechies, 1992), as in the case of this study. To apply the wavelet transform, the highest possible decomposition level of the TWS signal was tested, resulting in five levels. The wavelet family chosen for the decomposition was db3 (Daubechies, 1992). The signal extension model was observed, seeking the best possible application of DWT. The data normalization mode adopted for the WT was the *antireflect*, signal is extended by reflecting anti-simmetrically about the edge sample (PyWavelets, 2022). After decomposing the TWS with DWT, the results are a sequence compressed in one of the dimensions. As the TWS is being treated in three dimensions (i.e., latitude, longitude and time), and decomposed into the time dimension, the transform results reduce the time dimension. The time scale reduction occurs because DWT employs a grid where the mother wavelet is scaled by power two, expressing the results of each decomposition level as half of the previous level. Details on the DWT approach adopted here can be found in Rhif et al. (2019). As the TWS input data has 194 data points per time series, DWT returns the approximation values A5 (194 values) and details, namely D1 (96 values), D2 (48 values), D3 (24 values), D4 (12 values) and D5 (6 values). This method adapts a smooth variation of values for locations without data. This procedure requires the application of a mathematical function that minimizes the curvature of the surface, obtaining a result where the response is smooth and the surface passes exactly through the given entry points (Dierckx & Schumaker, 1994; Marcuzzo et al., 2012).

TWS time series has also passed through the seasonal decomposition method, which returns a moving average around an established window value. The chosen window size was 12 months, seeking to observe the annual variations in the data. The model is additive and suggests that the components are added together as follows (Perktold et al., 2022):

$$TWS_{[t]} = T_{[t]} + S_{[t]} + e_{[t]} \tag{3}$$

The results are represented in three outputs at each time step $[t]$, seasonality (S), trend (T) and residual (e). T is an increasing or decreasing value in the series, S is the short-term repetitive cycle in the series, and e is the random variation of the series.

HMB-based hydrogeological characteristics inserted into the model are: geological group, lithological description of the group, the type of aquifer, the degree of fracturing and the productivity of the aquifer. These non-numerical data were converted into zero and one values (0, 1) by the *one-hot* function (Scikit-learn, 2021) to be better used in the model.

Finally, the assessment of groundwater dynamics in two major aquifers in Brazil (Bauru-Caiuá and Urucuia) involved the application of statistical tests, Mann-Kendall (Kendall, 1948; Mann, 1945; Sneyers, 1991) and Sen (Sen, 1968), at a confidence level of 95%. The tests were applied over the RIMAS measurements between the years 2010 and 2020. The Sen test provided slope values for the observed data in the wells, which were subsequently interpolated using the ordinary kriging method (Ahmed et al., 2008; PyKrige, 2022) onto the GRACE data grid. Kriging was chosen because it is a regression method widely used in geostatistics. It assumes data collected from a given population are correlated in space (Peeters et al., 2010; Ruybal et al., 2019; Verdin et al., 2015). The change in groundwater volume was determined by multiplying the trends based on Sen's test and the simulated values for each aquifer, resulting in groundwater storage from 2002 to 2022.

3.1. Ensemble model

AI models were used in this research due to their flexibility to handle diverse data inputs, such as satellite data (GRACE) and the hydrogeological map of Brazil. AI models possess the capability to model nonlinear relationships and approximate any nonlinear mapping with high accuracy, without any prior assumptions about data properties (Khosravi et al., 2011). Previous studies have shown that AI techniques are effective in groundwater analysis (Malakar et al., 2021; Razavi et al., 2012; Sun, 2013; Tao et al., 2022).

As an attempt to identify the most effective AI model for simulating changes in groundwater storage based on GRACE data, a sensitivity analysis was performed. Various models and combinations of AI models were tested and evaluated (refer to section 4.1 for more details). The models examined include the Multi-Layer Perceptron (MLP), Long Sort-term Memory (LSTM), Bidirectional LSTM, Random Forest (RF), Support-vector Machine (SVM), Extreme Gradient Boosting (XGB), Light Gradient Boosting Model (LGBM), and CatBoost Model (CtB). Additionally, linear models such as Ordinary Least Squares (OLS), Linear Regression (LR), Bayesian Ridge (BR), and Stochastic Gradient Descent (SGDRegressor) were also assessed. Despite unsatisfactory outcomes from individual models, ensemble models were constructed, demonstrating superior metrics in terms of calibration and validation. Consequently, a ensemble model was chosen for the study, based on the AI principle that combining weaker models can yield a stronger model (Géron, 2019).

The ensemble model is composed of three different AI models. They are the Decision Tree (DT technique); the Extreme Gradient Boost (XGB; Chen and Guestrin, 2016), the Light Gradinet Boosting Model (LGBM; Ke et al., 2017) and the CatBoost (CtB; Prokhorenkova et al., 2017), followed by a Linear Regression (LR; Sklearn, 2023a) step. Fig. 2 shows the data processing flow and the architecture of the ensemble model. The input data in the model are the TWS, decompositions (wavelet and seasonal), the hydrogeological characteristics depicted by the hydrogeological map and the position of the well in space (latitude and longitude). The extreme values are removed from the input dataset by quantile threshold evaluation, respecting the criteria of being greater than 99% and less than 1% of the ΔGWS_{OBS}. Finally, the data are normalized. It should be noted that the ΔGWS_{OBS} values are not normalized. After the initial processes, the data are inserted into the DT models for the first approximations of the ΔGWS_{OBS} values from the input data. The results obtained after processing in the DT models are then inserted into the Linear Regression that finalizes the approximations of the ΔGWS_{OBS} observed in the wells. It is important to note that the model architecture presented in this research is not the only feasible design and was established through experimentation.

The architecture of the AI models used was:

XGB: estimators 3000; learning rate 0.001; sampling set 1; maximum depth 7; *XGBRegressor – gbtree*; early stop 150 steps.

LGBM: regression *boosting_type gbdt*; l1 and l2 metrics; learning rate of 0.001; layer fraction of 0.9; *bagging_fraction* 0.7; *bagging_freq* 20; maximum depth 8; number of sheets 128; *max_bin* 512; number of interactions 3000; early stop 150 steps.

CtB: number of iterations 3000, learning rate 0.01, depth 7, RMSE rating metric, *bagging_temperature* 0.01, *od_type*: Iter and *od_wait*: 20.

The data entered in the models are divided into two sets. The first set is composed of 80% of the data for model calibration (i.e., training and testing) and the second one is composed of the remaining 20% of the data for model validation. Note that the second set was selected from the initial data set at random. The calibration data is divided into 80% for training and 20% for testing. After training and testing with the first set, the model is retested with the validation wells. This procedure was performed to observe the real capacity of the model to adapt to different datasets. The hyperparameters of the models had their initial adjustment by the Hyperactive algorithm (Blanke, 2021). However, the final adjustment was done manually.

3.2. Evaluation of the ensemble model

The RIMAS network monitors the largest porous aquifers in Brazil, as illustrated in Fig. 1. To assess the ability of the proposed model to reproduce observed groundwater storage change, the metrics described in this section were used. In seeking a realistic comparison for the study areas with results obtained from a robust and widely tested model, ΔGWS_{CLSM} was used as baseline. Four experiments were carried out, all using the same model described in the previous section. The experiments were designed to quantify the model's ability to adjust to different datasets, training and test batch sizes, and different hydrological and hydrogeological conditions. The experiments are described below.

Experiment E1: RIMAS and GRACE data used in Li et al. (2019), plus data obtained from the HMB. For this experiment, 60 wells with 4504 monthly *in situ* measurements were used.

Experiment E2: In this experiment, all wells contained in the RIMAS database with more than 24 months of measurements were analyzed, as well as GRACE-based TWS and HMB data referring to the wells. In this experiment, 373 wells were used, resulting in 16,487 monthly *in situ* measurements.

Experiment E3: RIMAS and GRACE data used in the work by Li et al. (2019), plus MHB data and selected the wells by monitored aquifer. Eight aquifers were selected, as follows: Alter do Chão, Parecis, Urucuia, Bauru-Caiuá, Guarani, Cabeças, Poti and Serra Grande. The Cabeças, Poti and Serra Grande aquifers were included in the same model as Poti and Serra Grande present similar effective porosity (n_e) and close spatial distribution, renamed Cabeças/Serra Grande. Experiment E3 was not performed in the Içá, Missão Velha, and Mauruti aquifers, as there are only two wells in this aquifer in the dataset used by Li et al. (2019), which made it not possible to execute the model.

Experiment E4: In this experiment, all wells contained in the RIMAS database with more than 24 months of measurements were analyzed, in addition to the GRACE and HMB data referring to the wells. Groundwater data was separated by monitored aquifer. Eleven aquifers were selected, as follows: Alter do Chão, Parecis, Urucuia, Bauru-Caiuá, Guarani, Cabeças, Poti, Serra Grande, Içá, Missão Velha and Mauruti. As in experiment E3, the Cabeças, Poti and Serra Grande aquifers were included in the same model as Poti and Serra Grande present similar effective porosity and close spatial distribution. The same procedure was performed for the Missão Velha and Mauruti aquifers, included in the same model, as they have a very close spatial distribution, wells in the same GRACE pixel, renamed as Araripe.

In experiments E3 and E4, not all monitored aquifers were considered, as the number of wells used in Li et al. (2019) is very small or non-existent in several of them. However, aquifers were selected in all Brazilian regions. It is worth noting that the input data in the model is related to the RIMAS data in spatial and temporal scales, since each input data in the model has as reference the location of a well on a given date. Therefore, the input variables agree in temporal and spatial scales with the target variable (ΔGWS_{obs}). These used the 20% subset of the validation data.

The following statistical metrics were adopted for error evaluation of experiments: the mean absolute error (MAE), the root mean square error (RMSE), the Nash-Sutcliffe Efficiency (NSE) and the Kling-Gupta Efficiency (KGE) were selected. The metrics are defined in the equations below.

$$\mathrm{MAE} = \sum_{i=1}^{n} |x_i - y_i| \tag{4}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{5}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{y}_i)^2} \tag{6}$$

where x_i stands for the *in situ* measurements, y_i the estimated values by the model, $\bar{y}_i$ the average of the model estimates, and *n* is the number of observations. For MAE and RMSE, the ideal values are zero. For the NSE, values closer to one indicate a better adjusted model. The KGE is a reformulation of NSE, according to the expression:

$$\mathrm{KGE} = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{7}$$

where *r* is the Pearson correlation coefficient between the model results and the observed values, β is the ratio of the mean of the calculated values to the mean of the observations, and α is the ratio of the standard deviation of the calculated values to the observed. The optimal values for r, α, β and KGE is one. KGE values above -0.41 indicate that the model presents better results than the long-term average of the evaluated series.

4. Results and Discussion

4.1. Sensitivity analysis

The sensitivity analysis of AI models indicates the feasibility of approximating groundwater storage values using GRACE data, as previously demonstrated by Sun (2013). While these models do not explicitly account for the interaction process between groundwater and the medium, they are capable of tracking storage variations and qualitatively and quantitatively indicating areas with groundwater gains or losses. LR-based ensemble model show the best results for both calibration and validation steps. Such models were not subjected to the stochastic process associated with the neural networks, and had a lower computational cost compared to DT models.

The combined performance of the ensemble models surpassed that of the individual AI models, as depicted in Figure 3. All models were tested using the same calibration and validation datasets obtained from E2, which encompassed the largest available dataset (Supporting Table S2 summarizes the results of the sensitivity analysis).

The sensitivity analysis shows that individual linear models and neural networks performed less effectively compared to the DT models. The relatively inferior performance of the individual linear models can be attributed to the non-linear relationship between the GRACE data and ground-based measurements. Regarding the neural networks, it is plausible that stochastic processes during model execution, the selected model architectures, or even the specific neural network models themselves may have contributed to the relatively lower performance observed. The SVM, RF, SGDRegressor, and OLS models were not efficient in the simulations performed in calibration.

The superior performance of DT models (i.e., LGBM and CtB) may be associated with recursive nature of models, which can be regarded as representations of the decision-making process, where a dataset is depicted by a tree-like structure (Negnevitsky, 2002). Moreover, DT models demonstrate robustness in handling insufficient training data and categorical variables, such as data derived from HMB . As evidenced by RMSE results (see Fig. 3a), there is a significant improvement in overall performance when DT models are integrated with other models. This reinforces the idea that the combination of weaker models can lead to the development of a robust and improved model (Géron, 2019).

KGE values show that even if some models present an underestimation, most of the simulations are above the -0.41 value, indicating that they exceed the long-term average of the series under study (Fig. 3b). The r values indicate that the evaluated models can effectively reproduce the gain/loss between the simulated and observed values in the phase (Fig. 3d). The α values demonstrated a high standard deviation for neural networks (i.e., MLP, LSTM) (Fig. 3e). Finally, the RMSE, NSE, and MAE (Figs. 3a, 3f and 3g) values indicate that as DT models are added to the joint model the metrics are improved.

It worth noting that the LR-based ensemble model (i.e., LR[XGB, LGBM, CtB]) selected for simulating experiments E1, E2, E3, and E4, exhibited overall superior performance metrics. This model combines three models that employ the decision tree technique followed by LR. Although the individual LR model may not produce strong results when directly processing the input data, its effectiveness in refining the accuracy of the ensemble model is demonstrated. This can be attributed to the initial step undertaken by the decision tree models, which excel at capturing complex nonlinear relationships with greater capacity compared to other models. Consequently, the decision tree models provide the linear regression model with input data that exhibit a closer linear relationship, resulting in enhanced model performance.

4.2. Experiments E1 and E2

In experiments E1 and E2, the small errors were concentrated in the central and northeastern region of Brazil (Figs. 4a-4b), areas with a greater number of wells, while the largest errors were concentrated in the northern portion of the country, where there are fewer RIMAS wells inserted into the models. Experiments E1 and E2 presented MAE values below 2cm and RMSE below 3cm, in water column. The validation of these experiments resulted in MAE below 3cm and RMSE below 4.5cm. An interesting aspect that draws special attention in examining such results is the model's ability to simulate ΔGWS_{OBS} in wells inserted in an environment of crystalline rocks in southeastern Brazil, presented in the E2 results (Fig. 4b), where validation wells have RMSE below 3cm. NSE values for E1 and E2 were 0.87 and 0.65, and KGE values of 0.34 and 0.64, respectively (Figs. 4a-4b). These values indicate a good fit of the models to the datasets. However, those results make clear that not all the dependent variables were explained with precision. On the other hand, low RMSE and MAE values show that the prediction errors presented for large areas are much lower than those derived from ΔGWS_{CLSM}. Also, the correlation between RIMAS and GRACE-based TWS demonstrates the mismatch between their time series (see Supporting Table S3). Despite the great variability of the GRACE signal in both experiments E1 and E2, which used data from all regions of Brazil, the average results show the great approximation of the simulated values ΔGWS_{SIM} and ΔGWS_{OBS} (Figs. 5a-5h), demonstrating the low variance of the presented results achieved by the constructed model. The validation results of E1 and E2 (Figs. 5c, 5d, 5g and 5h), show that the averages observed by the model have a good approximation with the *in situ* observations. To measure the influence of the decompositions on the TWS data, experiment E2 was performed using only the TWS data and HMB data, the metrics are RMSE=4.0cm, MAE=2.5cm, NSE=0.02 and KGE=0.5. The validation results are RMSE=3.8cm, MAE=2.3cm, NSE=0.11 and KGE=-1.1.

Artificial intelligence models tend to perform better with larger batches of training and testing (Lecun et al., 2015). Due to the large concentration of data in the central and northeast regions of Brazil, the model may have presented a biased result (Figs. 4a-4b), that is, with a tendency to present better results in areas with greater amount of data, demonstrated validation wells. In addition, there is a greater variability of GRACE signals in the north of the country.

As for the variability of the GRACE signal, it is expected that groundwater will present a different participation in each region of Brazil and in each aquifer studied. Signal variability is related to the hydrological processes of each region, as well as types, cover and land use. The northern region of Brazil has large bodies of surface water such as the Tocantins, Solimões, Negro and Amazon rivers, in addition to extensive floodable areas in the Amazon region, indicating a large contribution of surface water to the region's TWS signal (A. Getirana et al., 2017; Melo & Getirana, 2019). Unlike the northern region and the swamps of the Brazilian Pantanal, the other regions of the country have a smaller amount of large surface water bodies. Even with the water filled up reservoirs built for the hydroelectric plants in these regions, a smaller component of surface water in the TWS signal is expected. Complementarily, these areas other than the northern regions have a higher groundwater extraction rate compared to the northern region of Brazil (IBGE, 2020). These factors may help to explain greater errors in the northern region of Brazil.

4.3. Experiments E3 and E4

Experiment E4 denotes that the highest values for RMSE are in Alter do Chão (Fig. 6b), Parecis (Fig. 6h) and Guarani (Fig. 6f) aquifers. Although featuring RMSE values of 3.1cm and MAE of 2.4cm, Alter do Chão aquifer has an NSE value of 0.89 and KGE 0.35. This result may occur due to the proximity of the wells to the Amazon River, which would affect the static water level fluctuations according to the water level variation of the river. In addition, many of these wells are inserted in an urban context, where land use and land cover jointly with underground water exploitation water can interfere in the results achieved by the models. In the case of the Parecis aquifer, calibration results are RMSE=2.4cm, MAE=1.4cm, NSE=0.71 and KGE=0.80, and 2.9cm, 2.3cm, 0.41 and 0.6, respectively, for validation.

For the Guarani aquifer, the variability of the GRACE signal associated with extraction processes might have hindered the best convergence of the models. The aquifer extends over thousands of kilometers and has experienced increasing groundwater pumping in its recharge areas for many years, as described by Takahashi (2012).

Areas with a higher concentration of monitoring wells return lower error values for each aquifer. Such a relationship is more visible in the results for the Cabeças/Serra Grande aquifers (Fig. 6c) and Urucuia (Fig. 6e). Errors of the test wells present lower values in these areas. Içá and Alter do Chão aquifers do not have as much data for training, resulting in higher model errors. However, results of all the aquifer models still return ΔGWS_{SIM} values better than ΔGWS_{CLSM}.

As noted by Brookfield et al. (2018), linear correlation analyses have limited ability to derive relationships between the TWS and *in situ* groundwater measurements in areas with deep vadose zones, hence deep static levels. Such a limitation is not observed in non-linear models, such as the one adopted here. This is highlighted in the responses of the proposed ensemble model, mainly for the Urucuia aquifer (Fig. 6e), an area of significant groundwater extraction and deep static water levels. Our model also has the ability to provide accurate estimates over areas with low thickness of the vadose zone and with great influence of surface waters, as observed in the wells in the Içá (Fig. 6a) and Alter do Chão (Fig. 6b) aquifers.

The difference in scale between satellite data and *in situ* measurements is addressed in many works, which may use statistical, dynamic methods (Gaur & Simonovic, 2019; Sehgal et al., 2021; Yin et al., 2018) and more recently artificial intelligence (Ali et al., 2021; Liu et al., 2020; Miro & Famiglietti, 2018). This issue was well solved by our model, as demonstrated in the results. Calibration results are above expectations in the Araripe aquifer (Fig. 6d), with RMSE, MAE, NSE and KGE of 0.27cm, 1.6cm, 0.92 and 0.86, respectively, and 2.1cm, 1.7cm, 0.11 and -0.11 for validation. This shows that small aquifers and wells located within the same GRACE pixel were not a problem for the approximations made by the model.

Experiment E4 (Fig. 7) shows that, when the input data is concentrated in the same aquifer (e.g., Guarani, Urucuia and Bauru aquifers), ΔGWS_{SIM} provides more stable results. This is expected due to the lower TWS variability, as well as to the constant geological characteristics of those aquifers. However, there is a small improvement in simulations which can be explained by the reduced amount of training data for the models.

Fig. 7 shows averaged ΔGWS_{OBS} , ΔGWS_{SIM} and ΔGWS_{CLSM} over the Guarani, Bauru-Caiuá and Urucuia aquifers. For Guarani, E4 shows that the model can predict the averaged behavior of the aquifer with good accuracy. Validation wells have better results than ΔGWS_{CLSM}. However, these wells depict a deviation from the expected response obtained for the wells included in the validation group for testing. This could indicate that the smallest number of values for training the series might have directly interfered in the result. In Bauru-Caiuá, the model presents results very close to *in situ* measurements, and validation wells average RMSE=0.23cm, MAE=0.11cm, NSE=0.98 and KGE=0.43. The model's response could be associated with the spatial distribution of *in situ* data within the aquifer (Fig. 6g), which covers almost the entire aquifer with a relatively constant spacing between the wells. In that aquifer, the model can reproduce observations with good precision for E4. For experiment E3, our model outperforms GLDAS in all metrics. This may have occurred because the wells included in this experiment are concentrated in an area of intense underground water extraction.

Differences between test and validation GWS_{CLSM} values in Figure 7 are attributed to the selection of validation wells in each experiment. These wells may represent very distant areas within the same aquifer, as the aquifers under investigation cover vast areas. CLSM simulates different processes in such distant areas, leading to different storage results. These variations in the simulated processes can generate differences between test and validation results if the wells selected in each set are in very distant regions within the study area.

For the Urucuia aquifer, results were above expectations. However, the great variability of the GRACE-based TWS estimates at each studied point, associated with different responses in each monitoring well to the groundwater extraction processes in the region, might contribute to a small departure of the predicted data with respect to *in situ* data in the validation, even though this area presented the best results in the study. Small variations in ΔGWS_{OBS} and ΔGWS_{SIM} could indicate a constant loss of water along the column in the aquifer, as reported by the work of (Gonçalves et al., 2020).

4.4. Statistic tests and storage

Using two decades of AI-based groundwater estimates, we attempted to quantify the spatiotemporal variability of the Urucuia and Bauru-Caiuá aquifers. Besides their socioeconomic importance, these two aquifers were selected for their good results in the model fit and data availability (70 and 60 wells are distributed across their respective domains). First, trends at individual wells were computed using the Mann-Kendall and Sen tests. The tests were applied to both model outputs and *in situ* measurements, adopting a confidence level of 0.95%. Trend slopes were then interpolated using the ordinary kriging approach for each GRACE grid within the aquifers. Groundwater volume change was determined by integrating grid-based trends over aquifers.

Both model estimates and *in situ* measurements show decreasing trends across the Urucuia aquifer (see Figs. 8a-8b). Sen's test shows a decreasing trend for the simulation and RIMAS observations averages (Fig. 8a), -0.36cm/year and -0.1cm/year for ΔGWS_{SIM} and ΔGWS_{OBS}, respectively. The longer time series derived from the model allows us to observe the behavior of the aquifer even before the installation of the monitoring network. The simulation shows a continuous water loss in the aquifer (Fig. 8a), with an intensification in 2006. Based on these trends, we estimate that the Urucuia aquifer has lost about $36km^3$ of water during 2002-2021, which is about the regulatory reserves of the aquifer (Gaspar, 2006). Such a water loss can be explained by an extended drought that has been impacting the region for over a decade (Getirana, 2016; Rodell et al., 2018; Getirana et al., 2021) combined with groundwater overexploitation (Vieira, 2021).

In Bauru-Caiuá, groundwater loss is concentrated in the northern portion of the aquifer (Fig. 8c), also impacted by the extended drought. It is worth noting that the area with the greatest water gain, in the southern part of the aquifer. The result of the Mann-Kendall test shows no trend and the Sen's test indicates a very smooth slope of both *in situ* and simulated groundwater time series (Fig. 8d), -0.02cm/year and -0.03cm/year for ΔGWS_{OBS} and ΔGWS_{SIM}, respectively. For the RIMAS wells, the results of the Mann-Kendall test show a good relationship with the process of gain and loss of water column (Fig. 8c) with decreasing results concentrated in the northern portion of the aquifer and increasing concentrations in the southern portion. The approximate water change across the aquifer is $-6km^3$, indicating a small change during the study period.

5. Conclusions and recommendations

Here, we demonstrate the viability of satellite-based monitoring of Brazilian aquifers. A novel artificial intelligence model was conceived and built by employing GRACE-based TWS data and its decomposition using wavelet and seasonal techniques, jointly with point-based *in situ* hydrogeological data. As a benchmark for our results, we used GLDAS outputs, specifically, groundwater change derived from CLSM with GRACE data assimilation (ΔGWS_{CLSM}). The validation results of the proposed methodology over all selected aquifers showed good groundwater estimates, outperforming GLDAS.

The TWS signal decomposition process proved to be very useful for the model, which adequately approximates the variations in groundwater storage in the different experiments. The proposed methodology can be applied in areas with a short history of groundwater monitoring and discontinuous time series, since aquifers such as Parecis, Iça and Cabeças/Serra Grande have less than ten years of static water level measurements and all wells in the RIMAS have gaps in their time series.

It is important to emphasize that the proposed models are representing the sample space inserted in the dataset. Differences between scales of GRACE data with a resolution of 0.25° and *in situ* measurements collected at an approximate point scale were overcome by the models. Furthermore, more than one well per pixel was also not a problem, as demonstrated by the results for the aquifers Içá, Parecis, Araripe and Urucuia. It is expected that the proposed model can be applied in areas with physical and geological characteristics similar to the training region, since the response of GRACE-based TWS signals tend to be similar. This feature can help to spatialize the storage of groundwater to unmonitored areas, being very useful for large aquifers. We demonstrate that it is possible to extend the groundwater monitoring period back to 2002, when GRACE was launched. As a result, we obtained groundwater information over eight additional years, which resulted in a more complete picture of water loss across the Urucuia and Bauru-Caiuá aquifers. This proves that the proposed methodology can provide important information for the management of groundwater resources.

As a model constraint, aquifers with short measurement periods are difficult to approximate. Regions with great variation in physical characteristics, such as soil type, geology, land use and occupation, precipitation rates or groundwater extraction, can create situations in which the proposed model does not respond as expected. Another limitation of the model is the spatial resolution for unmonitored areas, which initially depends on the resolution of the GRACE data. In experiment E2, which contains all monitored wells, some wells are inserted in an environment of crystalline rocks. Despite the good results, the model has not yet been properly adjusted for such a geological environment, as well as for karstic aquifers, being the subject of future investigation. Additionally, the proposed methodology can be applied to other aquifers, assuming that the aforementioned limitations are respected and the input data is sufficient for an adequate adjustment of the model.

Groundwater monitoring using satellite data and artificial intelligence can be a solution to spatialize groundwater storage values with good accuracy. Additionally, it is possible to make predictions for storage in different scenarios and with low computational costs, modifying only TWS values. This approach can also help in understanding aquifer dynamics, since, after the initial adjustments, the model can evaluate the past groundwater behavior using the GRACE data that started in 2002. The proposed methodology can be replicated in other aquifers globally with sufficient data for adequate model adjustment.

Acknowledgements

This work was funded by the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES) and the Brazilian Council for Scientific and Technological Development (CNPq).

Open Research

Monthly *in situ* groundwater measurements and the software developed in this study to process datasets are open access and available through Camacho et al. (2023). GRACE data is available on Save (2020). GLDAS data is available through the Goddard Earth Sciences Data and Information Services Center (https://disc.gsfc.nasa.gov/).

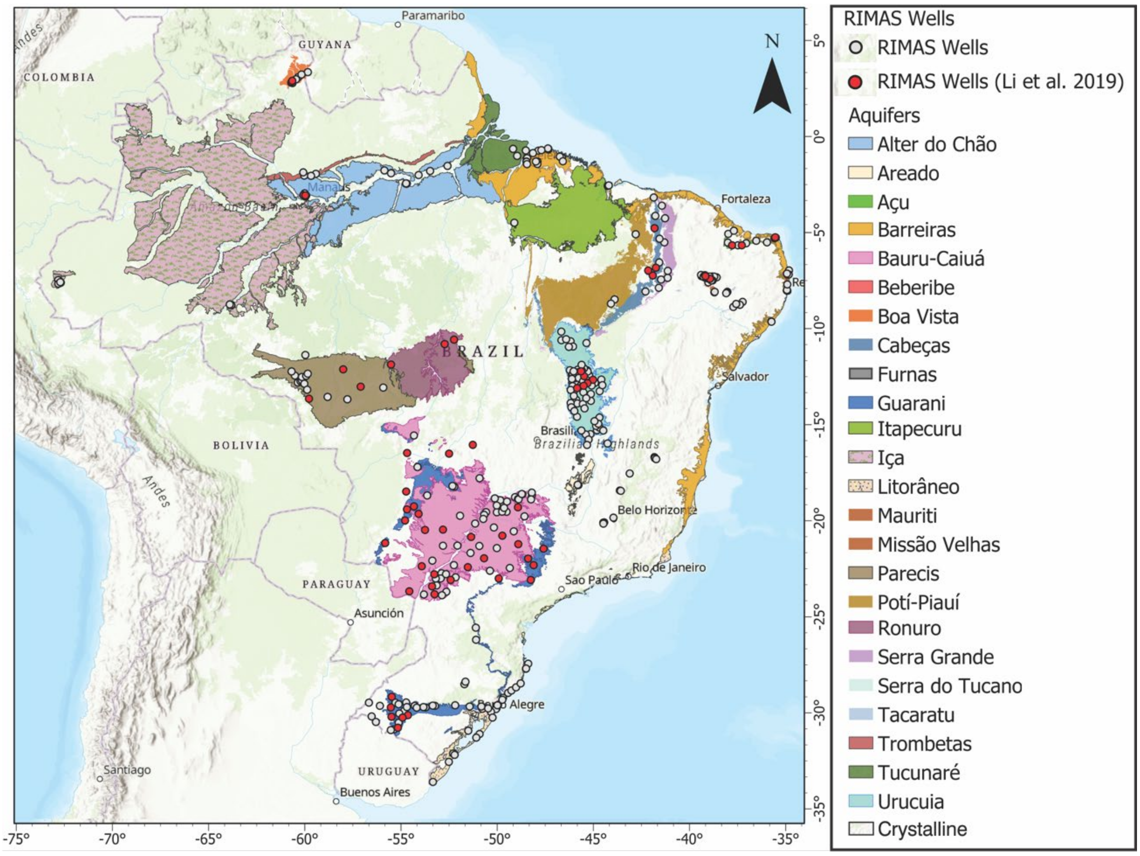

Figure 1. Geographical location of Brazilian aquifers and spatial distribution of the RIMAS groundwater monitoring network. RIMAS wells colored in red are those used in Li et al. (2019).

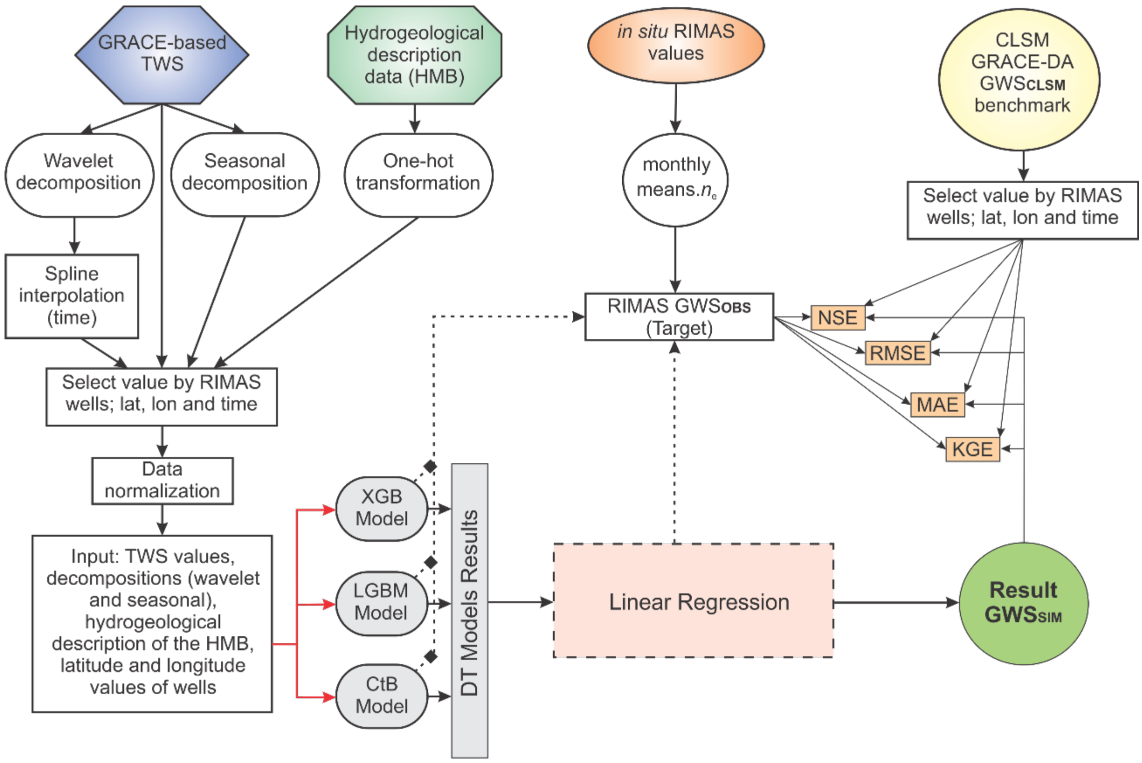

Figure 2. Processing flow diagram and ensemble model architecture. Both DT models and Linear Regression model use observed groundwater storage change (ΔGWS_{OBS}) as the target.

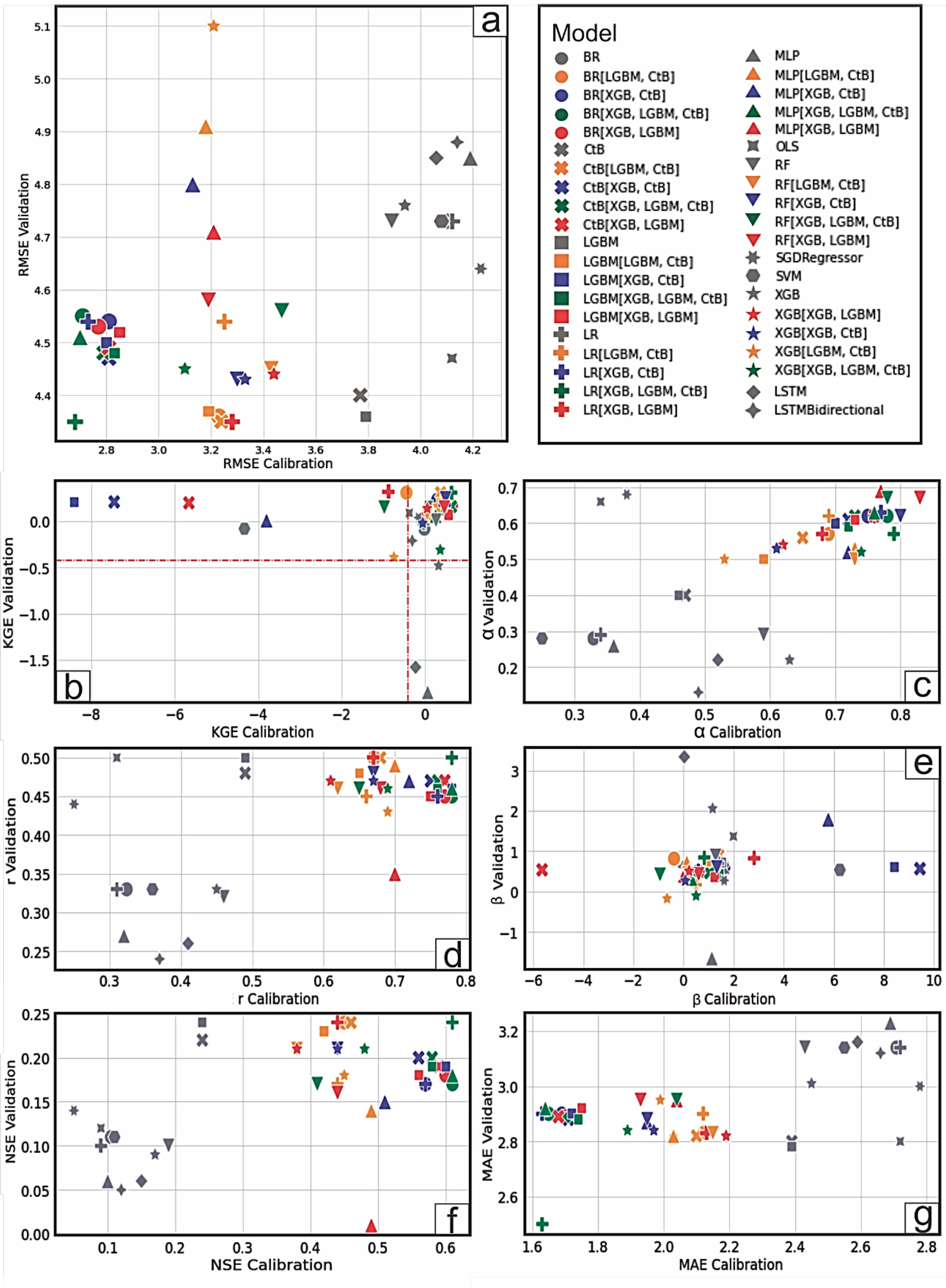

Figure 3. Sensitivity test of models in experiment E2: (a) root mean square error (RMSE) [cm], (b) Kling-Gupta efficiency (KGE) and its three components (c) α [-], (d) r [-] and (e) β [-], (f) Nash-Sutcliffe efficiency (NSE) [-] and mean absolute error (MAE) [cm]. Models are: Extreme Gradient Boosting (XGB), Light Gradient Boosting Model (LGBM), CatBoost Model (CtB), Random Forest (RF), Ordinary Least Squares Model (OLS), Linear Regression (LR), Bayesian Ridge Model (BR), Stochastic Gradient Descent (SGDRegressor), Support-vector Machine (SVM), Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM). The acronyms in parentheses in the upper right table indicate that the results of the models were used as input to

the external model. Red ticked lines in (b) indicate KGE=-0.41. Values above these lines indicate models with better fit than long-term averaged observations.

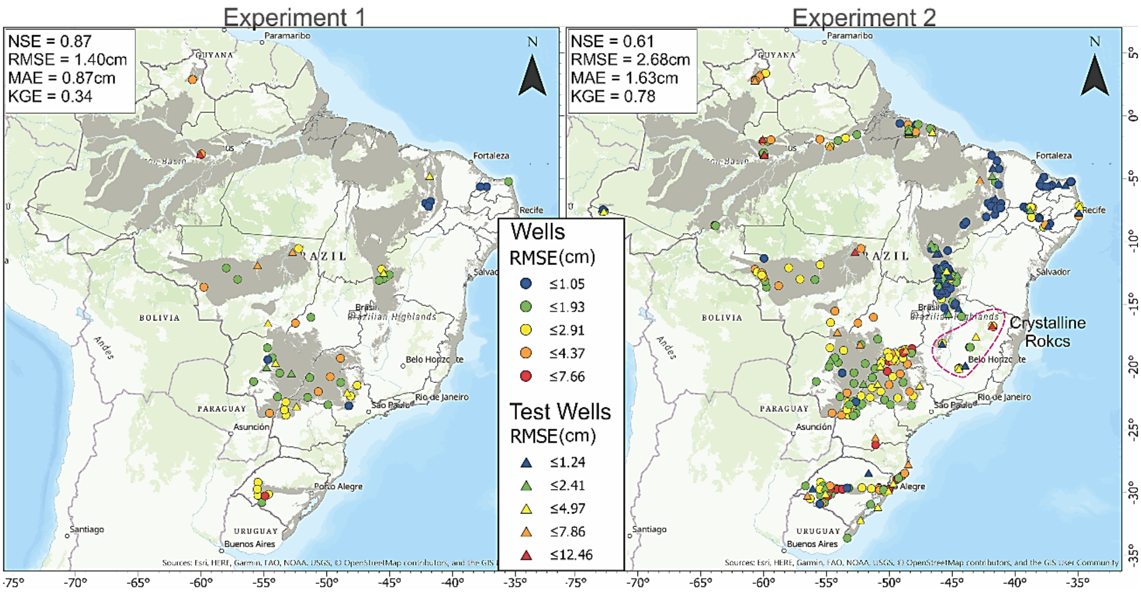

Figure 4. Results for experiments E1 and E2 using the ensemble model Linear Regression [XGB, LGBM, CtB]: spatial distribution of root mean square error (RMSE) [cm] for model calibration using (a) wells considered in Li et al. (2019) and (b) all RIMAS wells. Averages of metrics considered in this study are provided in the upper left boxes.

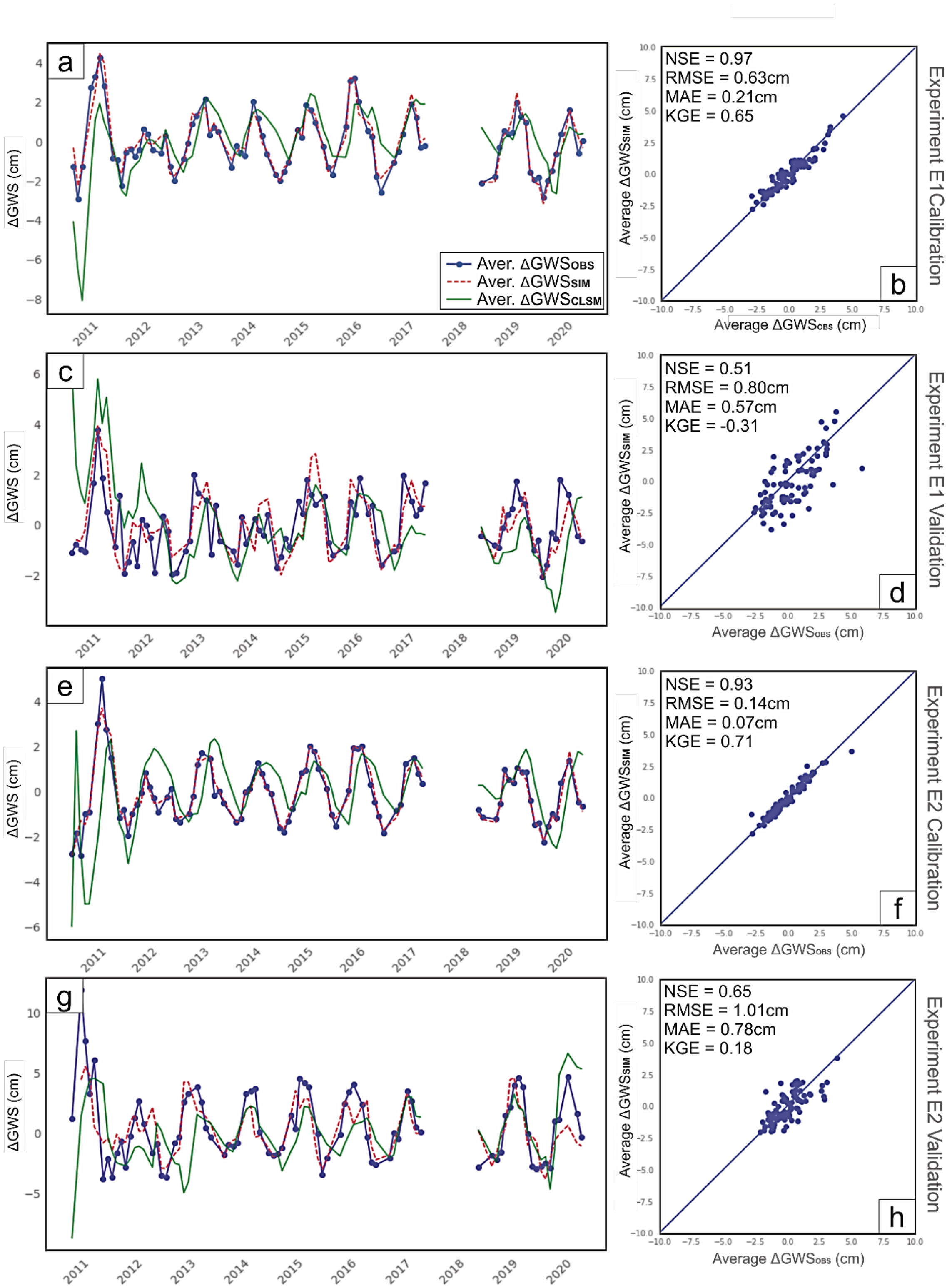

Figure 5. On the left side, averaged groundwater storage change time series derived from observations (ΔGWS_{OBS}), AI simulations (ΔGWS_{SIM}) and GLDAS (ΔGWS_{CLSM}) for experiments E1 and E2. On the right side, scatter plots between monthly ΔGWS_{OBS} and ΔGWS_{SIM} and metrics considered in the model evaluation: Nash-Sutcliffe efficiency (NSE), root mean square error (RMSE), mean absolute error (MAE) and Kling-Gupta efficiency (KGE).

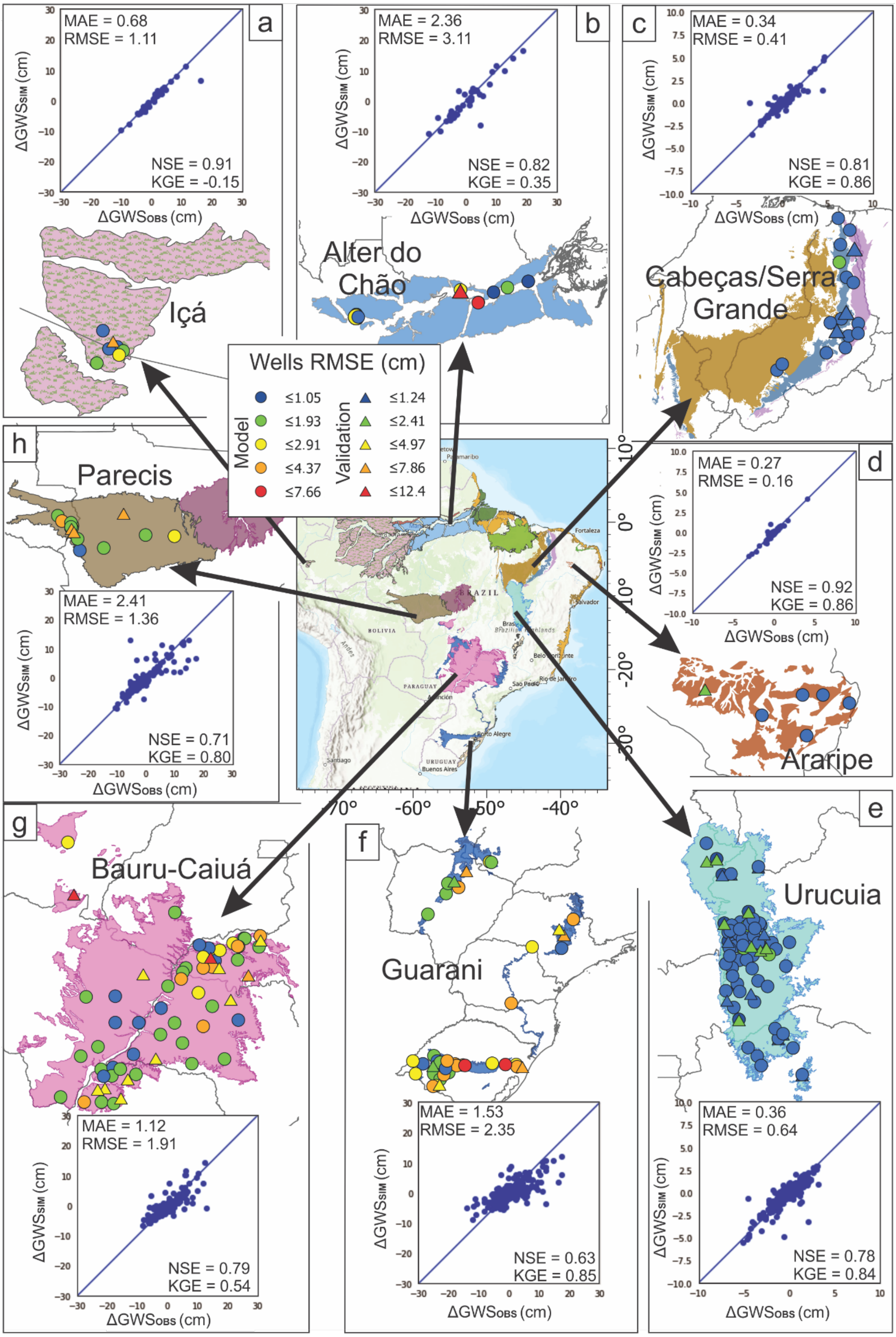

Figure 6. Results for experiment E4. Maps show the spatial distribution of root mean square errors (RMSE) [cm] over different aquifers. The scatter plots represent the 20% simulated test values (ΔGWS_{SIM}) and observed values (ΔGWS_{OBS}) in each aquifer. Averages of metrics considered in this study are also provided.

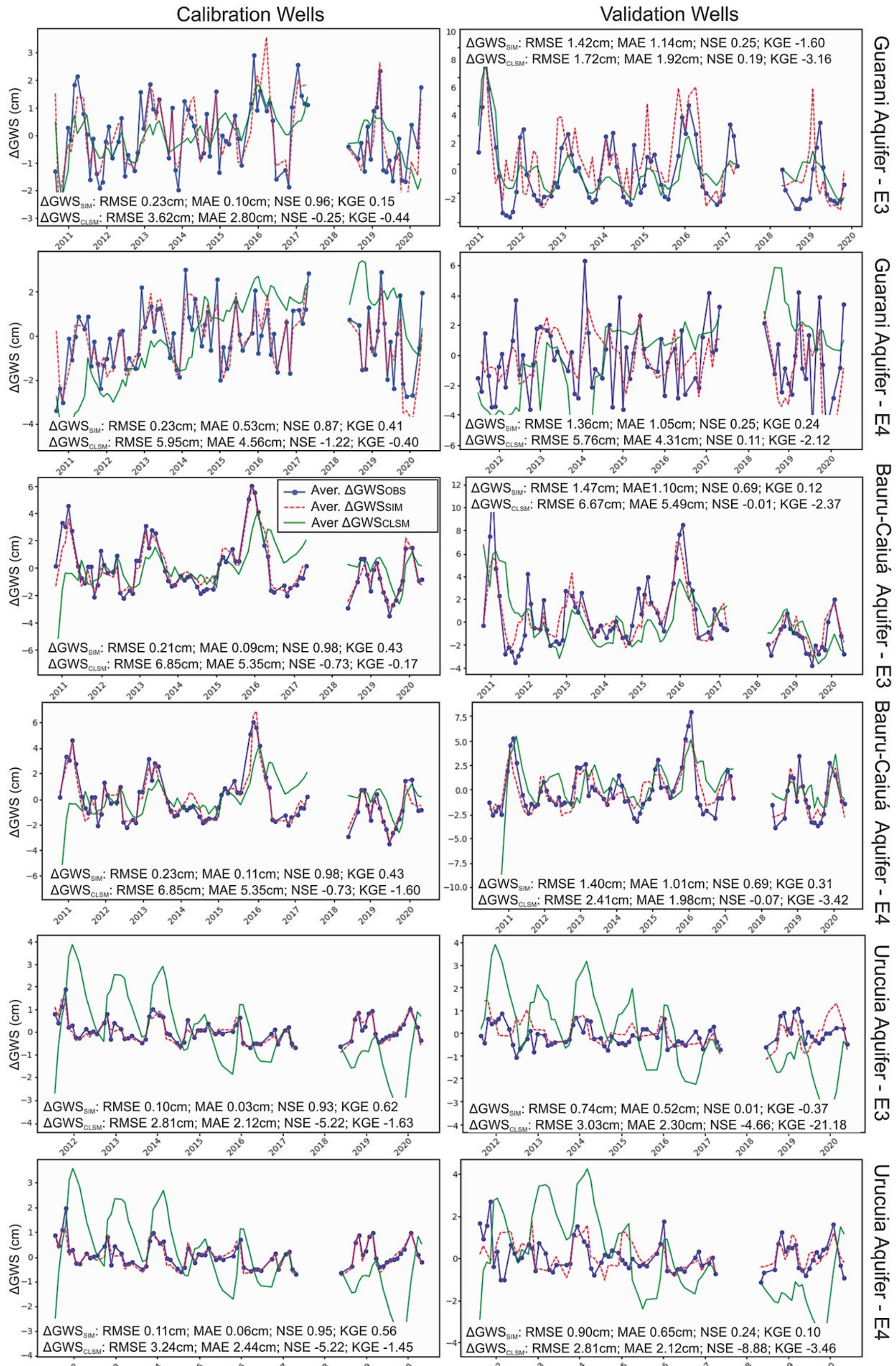

Figure 7. Aquifer-averaged time series of the groundwater storage change (**Δ**GWS) derived from experiments E3 and E4, over Guarani, Bauru-Caiuá and Urucuia aquifers, considering in situ values (ΔGWS_{OBS}). Time series correspond to observations (ΔGWS_{OBS}), simulations (ΔGWS_{SIM}) and GLDAS outputs (ΔGWS_{CLSM}). Metrics considered in this study are provided.

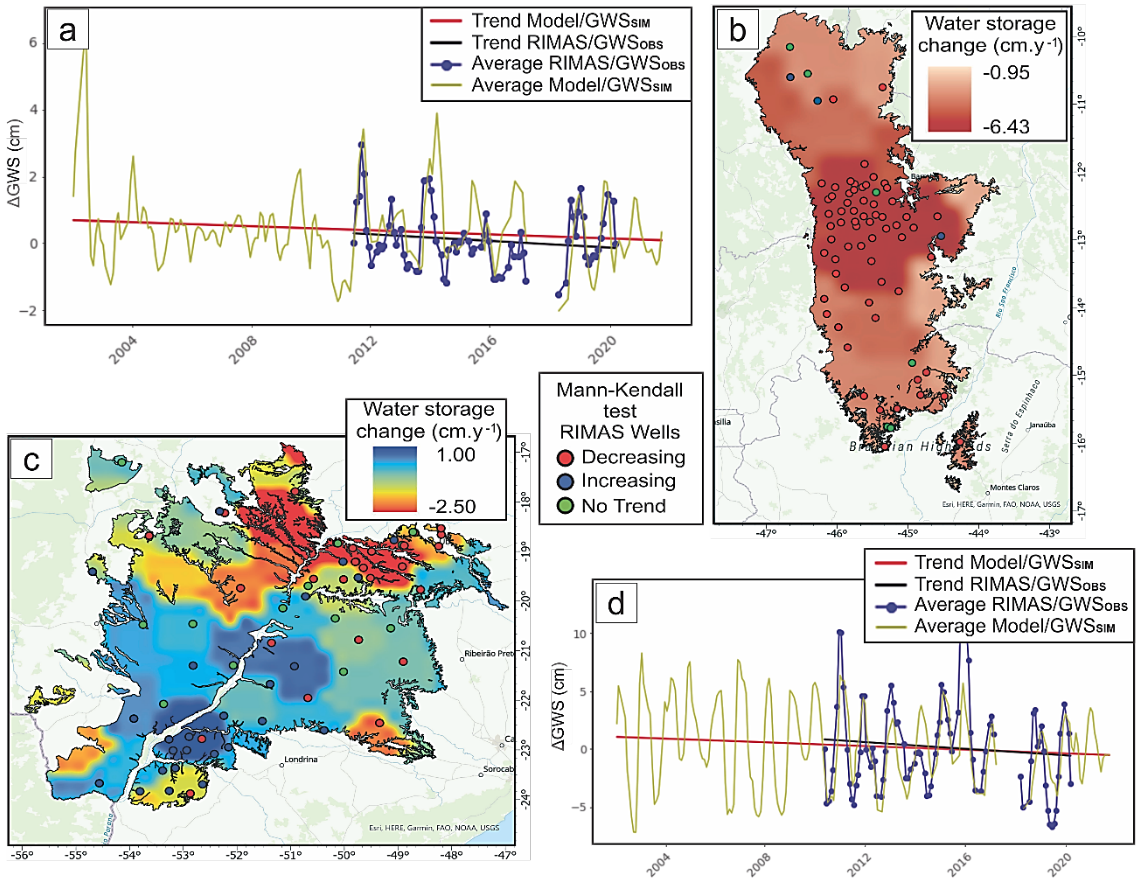

Figure 8. Spatiotemporal groundwater storage change (**Δ**GWS) in the Urucuia and Bauru-Caiuá aquifers: (a) aquifer-averaged **Δ**GWS time series and (b) spatially distributed trends over the Urucuia aquifer; (c) spatially distributed trends over the Bauru-Caiuá aquifer and (d) aquifer-averaged **Δ**GWS time series. Time series show experiment E4 outputs over 2002-2021 and RIMAS measurements over 2011-2020. **Δ**GWS trends over the Urucuia aquifer are -0.08cm/year from experiment E4 and 0.10cm/year from RIMAS. Over the Bauru-Caiuá aquifer, values are -0.03cm/year and -0.05cm/year, respectively. The dots on maps indicate results of the Mann-Kendall test for the RIMAS wells.

References

Afzaal, H., Farooque, A. A., Abbas, F., Acharya, B., & Esau, T. (2020). Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning. *Water (Switzerland)*, *12*(1). https://doi.org/10.3390/w12010005

Ahmed, S., Jayakumar, R., & Salih, A. (Eds.). (2008). *Groundwater Dynamics in Hard Rock Aquifers* (1st ed.). Springer, New Delhi. Retrieved from https://link.springer.com/book/10.1007/978-1-4020-6540-8

Ali, S., Liu, D., Fu, Q., Cheema, M. J. M., Pham, Q. B., Rahaman, M. M., et al. (2021). Improving the Resolution of GRACE Data for Spatio-Temporal Groundwater Storage Assessment. *Remote Sensing*, *13*(17), 3513. https://doi.org/10.3390/rs13173513

Alley, W. M., Healy, R. W., LaBaugh, J. W., & Reilly, T. E. (2002). Flow and storage in groundwater systems. *Science*, *296*(5575), 1985–1990. https://doi.org/10.1126/science.1067123

Andrew, R., Guan, H., & Batelaan, O. (2017). Estimation of GRACE water storage components by temporal decomposition. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2017.06.016

Ashraf, F. Bin, Haghighi, A. T., Riml, J., Mathias Kondolf, G., Kløve, B., & Marttila, H. (2022). A Method for Assessment of Sub-Daily Flow Alterations Using Wavelet Analysis for Regulated Rivers. *Water Resources Research*, *58*(1). https://doi.org/10.1029/2021WR030421

Barzegar, R., Fijani, E., Moghaddam, A. A., & Tziritis, E. (2017). Forecasting of groundwater level fluctuations using ensemble hybrid multi-wavelet neural network-based models. https://doi.org/10.1016/j.scitotenv.2017.04.189

Basu, B., Morrissey, P., & Gill, L. W. (2022). Application of nonlinear time series and machine learning algorithms for forecasting groundwater flooding in a lowland karst area. *Water Resources Research*. https://doi.org/10.1029/2021WR029576

Bechtold, M., De Lannoy, G. J. M., Koster, R. D., Reichle, R. H., Mahanama, S. P., Bleuten, W., et al. (2019). PEAT-CLSM: A Specific Treatment of Peatland Hydrology in the NASA Catchment Land Surface Model. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2130–2162. https://doi.org/10.1029/2018MS001574

Blanke, S. (2021). Hyperactive: An optimization and data collection toolbox for convenient and fast prototyping of computationally expensive models. Retrieved November 11, 2021, from https://github.com/SimonBlanke/Hyperactive

Brookfield, A. E., Hill, M. C., Rodell, M., Loomis, B. D., Stotler, R. L., & Porter, M. E. (2018). In-Situ and GRACE-based Groundwater Observations : Similarities , Discrepancies , and Evaluation in the High Plains Aquifer in Kansas. *Remote Sensing of Environment*, *205*(February), 408–418. https://doi.org/10.1029/2018WR023836

Camacho, C., Getirana, A., Rotunno, O. C., & Mourão, M. A. A. (2023). ML-based groundwater processing in Brazil. figshare. *Figshare*, *[Software]*. https://doi.org/10.6084/m9.figshare.22009562.v3

Chagas, V. B. P., Chaffe, P. L. B., & Blöschl, G. (2022). Climate and land management accelerate the Brazilian water cycle. *Nature Communications 2022 13:1*, *13*(1), 1–10. https://doi.org/10.1038/s41467-022-32580-x

Chen, T., & Guestrin, C. (2016). XGBoost. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Vol. 13-17-Augu, pp. 785–794). New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785

Condon, L. E., Kollet, S., Bierkens, M. F. P., Fogg, G. E., Maxwell, R. M., Hill, M. C., et al. (2021). Global Groundwater Modeling and Monitoring: Opportunities and Challenges. *Water Resources Research*, *57*(12), 1–27. https://doi.org/10.1029/2020WR029500

Daubechies, I. (1992). Ten Lectures on Wavelets. In D. H. Griffel (Ed.), *The Mathematical Gazette* (Vol. 79, p. 19). https://doi.org/10.2307/3620105

Dierckx, P., & Schumaker, L. L. (1994). Curve and Surface Fitting with Splines. *Mathematics of Computation*, *63*(207), 427. https://doi.org/10.2307/2153590

Diniz, J. A. O., Bomfim, L. F. C., Freitas, M. A. de, Monteiro, A. B., Cardoso, A. de C., Franzini, A. S., et al. (2014). Mapa Hidrogeológico do Brasil. Retrieved August 18, 2022, from http://www.cprm.gov.br/publique/Hidrologia/Estudos-Hidrologicos-e-Hidrogeologicos/Mapa-Hidrogeologico-do-Brasil-ao-Milionesimo-756.html

Ditmar, P. (2018). Conversion of time-varying Stokes coefficients into mass anomalies at the Earth's surface considering the Earth's oblateness. *Journal of Geodesy*, *92*(12), 1401–1412. https://doi.org/10.1007/s00190-018-1128-0

Ebrahimi, H., & Rajaee, T. (2017). Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Global and Planetary Change*, *148*, 181–191. https://doi.org/10.1016/j.gloplacha.2016.11.014

Erkyihun, S. T., Rajagopalan, B., Zagona, E., Lall, U., & Nowak, K. (2016). Wavelet-based time series bootstrap model for multidecadal streamflow simulation using climate indicators. *Water Resources Research*, *52*(5), 4061–4077. https://doi.org/10.1002/2016WR018696

Frappart, F., & Ramillien, G. (2018). Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. *Remote Sensing*, *10*(6). https://doi.org/10.3390/rs10060829

Gaur, A., & Simonovic, S. P. (2019). Introduction to Physical Scaling. In *Trends and Changes in Hydroclimatic Variables* (pp. 199–273). Elsevier. https://doi.org/10.1016/B978-0-12-810985-4.00004-9

Gemitzi, A., & Lakshmi, V. (2017). Downscaling GRACE data to estimate groundwater use at the aquifer scale. In *Proceedings of the 15th*. Retrieved from https://pdfs.semanticscholar.org/0d1c/ee42bbe11b34ba88215c22064fe2ab8d215c.pdf

Géron, A. (2019). *Hands-on Machine Learning whith Scikit-Learing, Keras and Tensorfow*. (N. Tache, Ed.), *O'Reilly Media, Inc* (1st ed.). Boston: O'Reilly Media, Inc.

Getirana, A., Kumar, S., Girotto, M., & Rodell, M. (2017). Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability. *Geophysical Research Letters*, *44*(20), 10,359-10,368. https://doi.org/10.1002/2017GL074684

Getirana, Augusto. (2016). Extreme Water Deficit in Brazil Detected from Space. *Journal of Hydrometeorology*. https://doi.org/10.1175/JHM-D-15-0096.1

Getirana, Augusto, Rodell, M., Kumar, S., Beaudoing, H. K., Arsenault, K., Zaitchik, B., et al. (2020). GRACE Improves Seasonal Groundwater Forecast Initialization over the United States. *Journal of Hydrometeorology*, *21*(1), 59–71. https://doi.org/10.1175/JHM-D-19-0096.1

Getirana, Augusto, Jung, H. C., Arsenault, K., Shukla, S., Kumar, S., Peters-Lidard, C., et al. (2020). Satellite Gravimetry Improves Seasonal Streamflow Forecast Initialization in Africa. *Water Resources Research*, *56*(2), 2019WR026259. https://doi.org/10.1029/2019WR026259

Getirana, Augusto, Libonati, R., & Cataldi, M. (2021). Brazil is in water crisis — it needs a drought plan. *Nature*, *600*(7888), 218–220. https://doi.org/10.1038/d41586-021-03625-w

Getirana, Augusto, Biswas, N. K., Qureshi, A. S., Rajib, A., Kumar, S., Rahman, M., & Biswas, R. K. (2022). Avert Bangladesh's looming water crisis through open science and better data. *Nature*, *610*(7933), 626–629. https://doi.org/10.1038/d41586-022-03373-5

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., Rodell, M., Draper, C., Bhanja, S. N., & Mukherjee, A. (2017). Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in India. *Geophysical Research Letters*, *44*(9), 4107–4115. https://doi.org/10.1002/2017GL072994

Gleeson, T., Wagener, T., Döll, P., Zipper, S. C., West, C., Wada, Y., et al. (2021). GMD perspective: The quest to improve the evaluation of groundwater representation in continental- to global-scale models. *Geoscientific Model Development*, *14*(12), 7545–7571. https://doi.org/10.5194/gmd-14-7545-2021

Gonçalves, R. D., Stollberg, R., Weiss, H., & Chang, H. K. (2020). Using GRACE to quantify the depletion of terrestrial water storage in Northeastern Brazil: The Urucuia Aquifer System. *Science of the Total Environment*, *705*(xxxx), 135845. https://doi.org/10.1016/j.scitotenv.2019.135845

de Graaf, I. E. M., Sutanudjaja, E. H., van Beek, L. P. H., & Bierkens, M. F. P. (2015). A high-resolution global-scale groundwater model. *Hydrology and Earth System Sciences*, *19*(2), 823–837. https://doi.org/10.5194/hess-19-823-2015

de Graaf, I. E. M., van Beek, R. L. P. H., Gleeson, T., Moosdorf, N., Schmitz, O., Sutanudjaja, E. H., & Bierkens, M. F. P. (2017). A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. *Advances in Water Resources*, *102*, 53–67. https://doi.org/10.1016/j.advwatres.2017.01.011

Hu, K., Awange, J. L., Khandu, Forootan, E., Goncalves, R. M., & Fleming, K. (2017). Hydrogeological characterisation of groundwater over Brazil using remotely sensed and model products. *Science of the Total Environment*, *599–600*, 372–386. https://doi.org/10.1016/j.scitotenv.2017.04.188

Huang, X., Gao, L., Crosbie, R. S., Zhang, N., Fu, G., & Doble, R. (2019). Groundwater recharge prediction using linear regression, multi-layer perception network, and deep learning. *Water (Switzerland)*, *11*(9). https://doi.org/10.3390/w11091879

IBGE. (2020). *Pesquisa Nacional de Saneamento Básico*. Rio de Janeiro.

Iqbal, N., Khan, A., Rizwan, A., Ahmad, R., Kim, B. W. A. N., Kim, K., & Kim, D. (2021). Groundwater Level Prediction Model Using Correlation and Difference Mechanisms Based on Boreholes Data for Sustainable Hydraulic Resource Management. *IEEE Access*, *9*, 96092–96113. https://doi.org/10.1109/ACCESS.2021.3094735

Jung, H. C., Getirana, A., Arsenault, K. R., Kumar, S., & Maigary, I. (2019). Improving surface soil moisture estimates in West Africa through GRACE data assimilation. *Journal of Hydrology*, *575*(January), 192–201. https://doi.org/10.1016/j.jhydrol.2019.05.042

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing*

Systems, *2017-Decem*(Nips), 3147–3155. Retrieved from https://papers.nips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

Kendall, M. (1948). Rank correlation methods. Retrieved from https://psycnet.apa.org/record/1948-15040-000

Khalil, B., Broda, S., Adamowski, J., Ozga-Zielinski, B., & Donohoe, A. (2015). Short-term forecasting of groundwater levels under conditions of mine-tailings recharge using wavelet ensemble neural network models. *Hydrogeology Journal*, *23*(1), 121–141. https://doi.org/10.1007/s10040-014-1204-3

Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011). Comprehensive Review of Neural Network-Based Prediction Intervals and New Advances. *IEEE Transactions on Neural Networks*, *22*(9), 1341–1356. https://doi.org/10.1109/TNN.2011.2162110

Kollet, S., Gasper, F., Brdar, S., Goergen, K., Hendricks-Franssen, H.-J., Keune, J., et al. (2018). Introduction of an Experimental Terrestrial Forecasting/Monitoring System at Regional to Continental Scales Based on the Terrestrial Systems Modeling Platform (v1.1.0). *Water*, *10*(11), 1697. https://doi.org/10.3390/w10111697

Kumar, S. V., Zaitchik, B. F., Peters-Lidard, C. D., Rodell, M., Reichle, R., Li, B., et al. (2016). Assimilation of Gridded GRACE terrestrial water storage estimates in the North American land data assimilation system. *Journal of Hydrometeorology*, *17*(7), 1951–1972. https://doi.org/10.1175/JHM-D-15-0157.1

Lähivaara, T., Malehmir, A., Pasanen, A., Kärkkäinen, L., Huttunen, J. M. J., & Hesthaven, J. S. (2019). Estimation of groundwater storage from seismic data using deep learning. *Geophysical Prospecting*, *67*(8), 2115–2126. https://doi.org/10.1111/1365-2478.12831

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

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

Liu, Y., Xia, X., Yao, L., Jing, W., Zhou, C., Huang, W., et al. (2020). Downscaling Satellite Retrieved Soil Moisture Using Regression Tree-Based Machine Learning Algorithms Over Southwest France. *Earth and Space Science*, *7*(10). https://doi.org/10.1029/2020EA001267

Malakar, P., Mukherjee, A., Bhanja, S. N., Ray, R. K., Sarkar, S., & Zahid, A. (2021). Machine-learning-based regional-scale groundwater level prediction using GRACE. *Hydrogeology Journal*, *29*(3), 1027–1042. https://doi.org/10.1007/s10040-021-02306-2

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245. https://doi.org/10.2307/1907187

Marcuzzo, F., Andrade, L., & Melo, D. (2012). Métodos de Interpolação Matemática no Mapeamento de Chuvas do Estado do Mato Grosso (Interpolation Methods in Mathematics of Rainfall Mapping of the State of Mato Grosso). *Revista Brasileira de Geografia Física*, *4*(4), 793. https://doi.org/10.26848/rbgf.v4i4.232714

Maxwell, R. M., Condon, L. E., & Kollet, S. J. (2015). A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geoscientific Model Development*, *8*(3), 923–937. https://doi.org/10.5194/gmd-8-923-2015

Melati, M. D., Fleischmann, A. S., Fan, F. M., Paiva, R. C. D., & Athayde, G. B. (2019). Estimates of groundwater depletion under extreme drought in the Brazilian semi-arid region using GRACE satellite data: application for a small-scale aquifer. *Hydrogeology Journal*, *27*(8), 2789–2802. https://doi.org/10.1007/s10040-019-02065-1

Melo, D. de C. D., & Getirana, A. (2019). Radar altimetry as a proxy for determining terrestrial water storage variability in tropical basins. *Remote Sensing*, *11*(21), 1–15. https://doi.org/10.3390/rs11212487

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

Moosavi, V., Vafakhah, M., Shirmohammadi, B., & Behnia, N. (2013). A Wavelet-ANFIS Hybrid Model for Groundwater Level Forecasting for Different Prediction Periods. *Water Resources Management*, *27*(5), 1301–1321. https://doi.org/10.1007/s11269-012-0239-2

Moosavi, V., Vafakhah, M., Shirmohammadi, B., & Ranjbar, M. (2014). Optimization of Wavelet-ANFIS and Wavelet-ANN Hybrid Models by Taguchi Method for Groundwater Level Forecasting. *Arabian Journal for Science and Engineering*, *39*(3), 1785–1796. https://doi.org/10.1007/s13369-013-0762-3

Mourão, M. A. A. (2009). *Projeto - Implantação de Rede Integrada de Monitoramento das Águas Subterrâneas - Proposta Técnica*. Belo Horisonte. Retrieved from http://rimasweb.cprm.gov.br/layout/pdf/proposta_monitoramento_CPRM_2009.pdf

Negnevitsky, M. (2002). *Artificial Intelligence-A Guide to Intelligence Systems*. (Addison-Wesley, Ed.), *Pearson Education Limited* (2nd ed., Vol. 1). Harlow: Pearson Education.

Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Arsenault, K. R., Li, B., & Getirana, A. (2019). Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend. *Water Resources Research*, *55*(12), 11274–11294. https://doi.org/10.1029/2019WR025363

Peeters, L., Fasbender, D., Batelaan, O., & Dassargues, A. (2010). Bayesian data fusion for water table interpolation: Incorporating a hydrogeological conceptual model in kriging. *Water Resources Research*, *46*(8), 1–11. https://doi.org/10.1029/2009WR008353

Perktold, J., Seabold, S., & Taylor, J. (2022). Statsmodels seasonal Decompose. Retrieved January 14, 2022, from https://www.statsmodels.org/dev/generated/statsmodels.tsa.seasonal.seasonal_decompose.html

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2017). CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, *2018-Decem*(Section 4), 6638–6648. Retrieved from http://arxiv.org/abs/1706.09516

PyKrige. (2022). PyKrige — PyKrige 1.7.0 documentation. Retrieved January 13, 2023, from https://geostat-framework.readthedocs.io/projects/pykrige/en/stable/

PyWavelets. (2022). Signal extension modes — PyWavelets Documentation. Retrieved August 18, 2022, from https://pywavelets.readthedocs.io/en/latest/ref/signal-extension-modes.html

Qi, X., & Neupauer, R. M. (2008). Wavelet analysis of dominant scales of heterogeneous porous media, *44*, 9406. https://doi.org/10.1029/2006WR005720

Razavi, S., Tolson, B. A., & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, *48*(7). https://doi.org/10.1029/2011WR011527

Reinecke, R., Foglia, L., Mehl, S., Trautmann, T., Cáceres, D., & Döll, P. (2019). Challenges in developing a global gradient-based groundwater model (G³M v1.0) for the integration into a global hydrological model. *Geoscientific Model Development*, *12*(6), 2401–2418. https://doi.org/10.5194/gmd-12-2401-2019

Ren, H., Song, X., Fang, Y., Hou, Z. J., & Scheibe, T. D. (2021). Machine Learning Analysis of Hydrologic Exchange Flows and Transit Time Distributions in a Large Regulated River. *Frontiers in Artificial Intelligence*, *4*(April), 1–18. https://doi.org/10.3389/frai.2021.648071

Rhif, M., Abbes, A. Ben, Farah, I. R., Martínez, B., & Sang, Y. (2019). Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review. https://doi.org/10.3390/app9071345

Richey, A. S., Thomas, B. F., Lo, M.-H., Reager, J. T., Famiglietti, J. S., Voss, K., et al. (2015). Quantifying renewable groundwater stress with <scp>GRACE</scp>. *Water Resources Research*, *51*(7), 5217–5238. https://doi.org/10.1002/2015WR017349

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., et al. (2003). THE GLOBAL LAND DATA ASSIMILATION SYSTEM. https://doi.org/10.1

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M.-H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(water), 651–659. https://doi.org/10.1038/s41586-018-0123-1

Rowlands, D. D., Luthcke, S. B., McCarthy, J. J., Klosko, S. M., Chinn, D. S., Lemoine, F. G., et al. (2010). Global mass flux solutions from GRACE: A comparison of parameter estimation strategies—Mass concentrations versus Stokes coefficients. *Journal of Geophysical Research*, *115*(B1), B01403. https://doi.org/10.1029/2009JB006546

Ruybal, C. J., Hogue, T. S., & McCray, J. E. (2019). Evaluation of Groundwater Levels in the Arapahoe Aquifer Using Spatiotemporal Regression Kriging. *Water Resources Research*, *55*(4), 2820–2837. https://doi.org/10.1029/2018WR023437

Save, H. (2020). CSR GRACE and GRACE-FO RL06 Mascon Solutions v02. https://doi.org/10.15781/cgq9-nh24

Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, *121*(10), 7547–7569. https://doi.org/10.1002/2016JB013007

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*(4). https://doi.org/10.1029/2011WR011312

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

Scikit-learn. (2021). Scikit-learn: machine learning in Python — scikit-learn 1.0.2 documentation. Retrieved January 11, 2022, from https://scikit-learn.org/stable/index.html

Sehgal, V., Gaur, N., & Mohanty, B. P. (2021). Global Flash Drought Monitoring Using Surface Soil Moisture. *Water Resources Research*, *57*(9), 1–32.

https://doi.org/10.1029/2021WR029901

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's Tau. *Journal of the American Statistical Association*, *63*(324), 1379–1389. Retrieved from https://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1968.10480934?needAccess=true#.YBhG2TFKjIU

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

Sklearn. (2023). Linear Regression — scikit-learn 1.2.2 documentation. Retrieved March 28, 2023, from https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

Sneyers, R. (1991). *On the statistical analysis of series of observations.* (Meteorological Organization, Ed.), *On the statistical analysis of series of observations.* (2nd ed.). Geneva: Meteorological Organization.

Stollnitz, E. J., DeRose, A. D., & Salesin, D. H. (1995). Wavelets for computer graphics: a primer.1. *IEEE Computer Graphics and Applications*, *15*(3), 76–84. https://doi.org/10.1109/38.376616

Sun, A. Y. (2013). Predicting groundwater level changes using GRACE data. *Water Resources Research*, *49*(9), 5900–5912. https://doi.org/10.1002/wrcr.20421

Sun, A. Y., Scanlon, B. R., Zhang, Z., Walling, D., Bhanja, S. N., Mukherjee, A., & Zhong, Z. (2019). Combining Physically Based Modeling and Deep Learning for Fusing GRACE Satellite Data: Can We Learn From Mismatch? *Water Resources Research*, *55*(2), 1179–1195. https://doi.org/10.1029/2018WR023333

Swenson, S., Wahr, J., & Milly, P. C. D. (2003). Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*. https://doi.org/10.1029/2002WR001808

Takahashi, A. T. (2012). *Projeto Rede Integrada de Monitoramento das Águas Subterrâneas: Sistema Aquífero Guarani Nos Estados De São Paulo, Mato Grosso Do Sul E Paraná [Takahashi, A.T. e Mourão, M.A.A. Coord.]. Levantamento De Recursos Hídricos Subterrâneos* (Vol. 15). São Paulo. Retrieved from http://rimasweb.cprm.gov.br/layout/apresentacao.php

Tao, H., Hameed, M. M., Marhoon, H. A., Zounemat-Kermani, M., Heddam, S., Kim, S., et al. (2022). Groundwater level prediction using machine learning models: A comprehensive review. *Neurocomputing*, *489*, 271–308. https://doi.org/10.1016/j.neucom.2022.03.014

Tapley, B., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, *31*(9), 1–4. https://doi.org/10.1029/2004GL019920

Torrence, C., & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

UNESCO. (2022). Intergovernmental Hydrological Programme: Groundwater.

Verdin, A., Rajagopalan, B., Kleiber, W., & Funk, C. (2015). A Bayesian kriging approach for blending satellite and ground precipitation observations. *Water Resources Research*, *51*(2), 908–921. https://doi.org/10.1002/2014WR015963

Vieira, M. B. de S. (2021). *Estudo das vazões do sistema aquífero Urucuia em períodos de recessão hídrica*. *UNB*. Universidade de Brasilia. Retrieved from https://rigeo.cprm.gov.br/handle/doc/22169

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

Yin, W., Hu, L., Zhang, M., Wang, J., & Han, S. C. (2018). Statistical Downscaling of GRACE-Derived Groundwater Storage Using ET Data in the North China Plain. *Journal of Geophysical Research: Atmospheres*, *123*(11). https://doi.org/10.1029/2017JD027468

Yosefvand, F., & Shabanlou, S. (2020). Forecasting of Groundwater Level Using Ensemble Hybrid Wavelet–Self-adaptive Extreme Learning Machine-Based Models. *Natural Resources Research*, *29*(5), 3215–3232. https://doi.org/10.1007/s11053-020-09642-2

Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008). Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, *9*(3), 535–548. https://doi.org/10.1175/2007JHM951.1

Zare, M., & Koch, M. (2018). Groundwater level fluctuations simulation and prediction by ANFIS- and hybrid Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models: Application to the Miandarband plain. *Journal of Hydro-Environment Research*, *18*, 63–76. https://doi.org/10.1016/j.jher.2017.11.004

Zhang, G., Zheng, W., Yin, W., & Lei, W. (2020). Improving the Resolution and Accuracy of Groundwater Level Anomalies Using the Machine Learning-Based Fusion Model in the North China Plain. *Sensors*, *21*(1), 46. https://doi.org/10.3390/s21010046