Article

Exploiting Earth Observation Data to Impute Groundwater Level Measurements with an Extreme Learning Machine

Steven Evans, Gustavious P. Williams *[ID]**, Norman L. Jones**[ID]**, Daniel P. Ames**[ID] **and E. James Nelson**

Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT 84602, USA; stevenwevans2@gmail.com (S.E.); njones@byu.edu (N.L.J.); dan.ames@byu.edu (D.P.A.); jimn@byu.edu (E.J.N.)
* Correspondence: gus.williams@byu.edu

check for
updates

Abstract: Groundwater resources are expensive to develop and use; they are difficult to monitor and data collected from monitoring wells are often sporadic, often only available at irregular, infrequent, or brief intervals. Groundwater managers require an accurate understanding of historic groundwater storage trends to effectively manage groundwater resources, however, most if not all well records contain periods of missing data. To understand long-term trends, these missing data need to be imputed before trend analysis. We present a method to impute missing data at single wells, by exploiting data generated from Earth observations that are available globally. We use two soil moisture models, the Global Land Data Assimilation System (GLDAS) model and National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) soil moisture model to impute the missing data. Our imputation method uses a machine learning technique called Extreme Learning Machine (ELM). Our implementation uses 11 input data-streams, all based on Earth observation data. We train and apply the model one well at a time. We selected ELM because it is a single hidden layer feedforward model that can be trained quickly on minimal data. We tested the ELM method using data from monitoring wells in the Cedar Valley and Beryl-Enterprise areas in southwest Utah, USA. We compute error estimates for the imputed data and show that ELM-computed estimates were more accurate than Kriging estimates. This ELM-based data imputation method can be used to impute missing data at wells. These complete time series can be used improve the accuracy of aquifer groundwater elevation maps in areas where in-situ well measurements are sparse, resulting in more accurate spatial estimates of the groundwater surface. The data we use are available globally from 1950 to the present, so this method can be used anywhere in the world.

Keywords: GLDAS; PDSI; machine learning; Extreme Learning Machine; groundwater levels

1. Introduction

1.1. Challenges

One challenge facing groundwater managers is that the limited amount of available data makes it difficult to understand long-term aquifer behavior and trends. For example, missing data, or data gaps, make it difficult to develop groundwater elevation maps, as any well with missing data at a given time cannot be used to create the map. Groundwater data are only available at wells or piezometers that are costly to install and monitor [1]. Data collection can be difficult and time-consuming, requiring visits to each well, and often requires manual sampling. These challenges limit the temporal resolution of the data

and frequently results in large time gaps in water level records at individual wells, making it difficult to understand long-term aquifer behavior [2]. The challenges associated with groundwater data availability are intensified in developing countries because of resource limitations [3]. The ability to impute or estimate these missing data at individual wells helps address this challenge, especially if the method can use data available globally and in areas with limited well data.

1.2. Method Background and Overview

This paper presents a method to impute missing water surface measurements for a well. Our approach uses historical information at any given well to determine local behaviors and trends, then uses correlations with Earth observations to extend these trends across measurement gaps. This approach generates a complete time series of measurements at every well in an aquifer. These data can then be used to generate water surface elevation maps over time, which can be used to evaluate storage changes in an aquifer and determine if the use is sustainable. We show that this approach provides complete data series at each well, allowing more accurate groundwater level maps to be generated.

For data imputation, we use existing well data to train the model on aquifer and well characteristics. Climatic, geographic, lithological, and human factors influence aquifer water levels and storage and can be used to guide data imputation. The anthropogenic processes are generally local and captured by historic well data, while geographic and lithological aspects can have both local and regional impacts, and climatic processes typically affect even larger areas; these processes may be reflected in Earth observation data. For data imputation, historic well data provides information on local and regional trends and processes, while Earth observation data provides information on regional weather and climate impacts on aquifers. Earth observations can also provide information on land use, which influences how wells are used. For example, pumping patterns and rates in an agricultural area are typically different than those for wells used for potable water near housing areas; Earth observation data can capture these changes in land use over time.

We use a machine learning algorithm called an Extreme Learning Machine (ELM), to train a model using water level measurements at single wells and Earth observations. We then use the trained model with the available Earth observation data to impute missing data across measurement gaps at a well. In general, these type of models are used to impute data over relatively small gaps and are less accurate for data imputation over long time periods. This model, if successful, captures local aquifer properties, trends, and responses and regional impacts of climate, weather, and land use. We show that our imputation method for a single well is generally more accurate than using spatial interpolation methods for that location. After detailing our computational methods, we present two case studies where we compare the accuracy of the imputed data points at a well versus using spatial interpolation without the imputed data to estimate data at the well location, a method often done to create water surface elevation maps. The Earth observation data we use are available globally, supporting use in developed and developing countries, allowing them to extend the use of wells with missing data.

1.3. Overview of Imputation Methods

Data imputation methods for groundwater levels use temporal patterns in historic data at a well combined with other data that are correlated and related to groundwater use or levels, to impute missing data values at a well. There is a wide body of literature on various data imputation methods that includes very simple methods such as linear interpolation between missing data points, various spline computational methods, and more complex time series analysis and statistical methods. In many cases, these same methods can be used for either data imputation across gaps or data extrapolation (e.g., prediction) with much of the literature focused on the latter. Traditionally, environmental data

imputation used classical statistics-based time series modeling approaches—an excellent overview is presented in [4,5]. Examples of these methods include auto-regressive (AR), moving-average (MA), auto-regressive moving-average (ARMA), auto-regressive integrated moving-average (ARIMA), and seasonal auto-regressive integrated moving-average (SARIMA) methods, often combined with multiple linear regression to estimate model parameters [6–8]. Recently, other statistical methods have been reported in the literature, including decomposition approaches combined with Bayesian and other statistical models [9–11], other machine learning methods [12,13], kriging temporal data [14], and an ARMAX model based on the eigenstructure of aquifer dynamics [15]. Rado, et al. [16] provides a recent review that focuses on machine learning methods for time series data imputation and extrapolation. These methods all treat water level measurements at a well as a time series with a non-random structure that can be modeled and used to impute or extrapolate the time series at a given well. These methods are not used to estimate data at locations where measurements have not been taken.

In this work, we build on machine learning methods that leverage correlation patterns between groundwater levels and surface conditions that can be measured using Earth observation sensors. Though literature exists demonstrating the use of correlations between point data and spatial data, such as Earth observations, to impute missing values for other environmental time series data, to our knowledge, there is little published literature on using machine learning approaches to exploit correlations between well measurement point data and Earth observation spatial data for groundwater time-series data imputation. There are a large number of studies that predict groundwater levels at a well based on measured data at that well (e.g., time series analysis) but none that exploit other spatial information to inform this imputation. Reported methods for groundwater level data imputation at a well include, for example, using genetic programing [17], artificial neural networks [18], and hybrid soft-computing [19] methods. Sahoo and Jha [8] present a comparison of groundwater-level predictions using multiple linear regression and artificial neural network methods, which provides a good overview of this field, while Gong, et al. [20] present a comparison of support vector machines and adaptive neuro fuzzy inference systems for well data imputation, and this paper includes references to other literature related to machine learning approaches. However, these studies do not include Earth observation data and mostly focus on extrapolation or prediction, not imputation issues, which while similar, present different challenges and opportunities.

Groundwater data imputation is somewhat unique, as groundwater levels are influenced by both regional weather and climate processes which affect water levels on a basin or aquifer scale, this implies that relatively coarse Earth observation data may be useful. Earth observation data can be used to infer these weather and climate processes, as these trends in weather and climate conditions affect groundwater demand and extraction. Water levels are influenced by pumping, which has significant local impacts on the water level surface. While pumping is based on water demand, this local demand is influenced by water use requirements, which are driven by processes such as agriculture and weather patterns such as drought or monsoons. Both water demand and use can be correlated to Earth observations, which can measure land use and indications of soil moisture.

1.4. Remote Sensing Data

Since groundwater use is influenced by weather, climate, and types of land use in the region or area, remote sensing data can be used to provide insights into these processes. A number of different remote sensors and products may provide insight into water use and potentially be correlated to groundwater levels [21]. These remote sensing measurements include computed indexes related to water supply or demand such as precipitation [22,23], soil moisture [24], evapotranspiration [25], temperature [26], plant cover and growth [27,28], and computed indexes such as the Palmer Drought Severity Index (PDSI) [29] or other drought indexes, such as the Scaled Drought Condition Index (SDCI) [30], which could

correlate to either groundwater demand or groundwater recharge. There are also integrated computed datasets of soil moisture based on Earth observation data, such as the Global Land Data Assimilation System (GLDAS) model or the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) soil moisture model. GLDAS incorporates observations such as vegetation, land cover, soil, elevation, precipitation, and temperature information derived from or measured by various satellites (AVHRR aboard *NOAA-15* satellite, MODIS aboard NASA's *Terra* satellite, EOS aboard NASA's *Aqua* satellite) and computes soil moisture on a global scale [31]. The CPC model estimates soil moisture using a one-layer hydrological model, which uses precipitation and temperature data from Earth observations as inputs [29]. Monthly precipitation and temperature data are available from the CPC PRECipitation REConstruction over Land [23] and the CPC Global Land Surface Air Temperature Analysis [26], respectively.

These remote sensing data can provide information on regional and longer-term trends in an aquifer, if aquifer processes such as recharge or extraction are correlated with trends in weather or climate. They can also be used to estimate local trends in an aquifer as groundwater extraction rates may be correlated to rainfall frequency and drought conditions. Earth observations can be used to determine local land use by inferring land use categories such as urban, agricultural, pastoral, forest, or other types, which can be correlated with how a well is managed [32]. Soil moisture models, such as GLDAS and the CPC model, have been used to estimate groundwater levels and recharge [33], though not for data imputation.

In this work, we exploit remote sensing-derived products to improve the accuracy of groundwater level imputation at a single well. We use soil moisture data from both the GLDAS and CPC soil moisture models, combined with in-situ well data, to develop a model of the time series of groundwater level measurements at a given well. While these remote sensing data can only infer changes over larger regions, these data provide a useful, direct measure of how regional climate, weather, and land use affect water demand at a given well. Correlations with historic data can show how water levels at a given well respond to weather or climate forcings, such as drought or wet seasons. Remotely-sensed data can be used to separate basin-level processes, such as water storage changes due to weather, population, or climatic change, from local changes such as those from well pumping. By using both historic data and Earth observation data to build a model of a well, this model can include all these processes. Our approach uses both of these data sources to build a model of groundwater elevation changes at a given well, exploiting the correlations between water demands with local uses, such as agricultural demands, which can be inferred from soil moisture measured by Earth observation methods. In our model, this correlation is not explicit as we do not use land cover as an input. However, the land cover influences the soil moisture data from the GLDAS and CPC data sets. Soil moisture estimates from GLDAS and CPC are also indications of groundwater demand, as lower levels of soil moisture indicate the need for more irrigation, while higher levels indicate less demand. At some wells, this change in demand may be correlated with pumping rates and thus groundwater levels; for other wells, (e.g., potable water use), this correlation might be less. Our model creates a model for each individual well to exploit these correlations. For this study, we will use Earth observation, or remote sensing, data to provide information on regional forcings and combine this with historic information from wells that characterize local aquifer and well processes to impute missing data.

1.5. Extreme Learning Machine

Data imputation for groundwater is a problem that is not well suited to machine learning approaches such as artificial neural networks or other popular methods, as the available data are limited, and most machine learning methods perform better with large amounts of data. For example, 10 years of data taken quarterly at a well is only 40 data points, a very small sample for most machine learning methods. Since we

need to train a model for each individual well with missing data, the amount of data available and the training time for a model are relevant decision parameters when selecting a machine learning approach.

Challenges related to limited data and the need for fast training are traditionally approached using feed-forward methods such as least square support vector machines or proximal support vector machines [34]. However, these methods are designed for classification problems and cannot be applied to regression problems, such as data imputation. Huang, et al. [34–36] present an algorithm called an ELM that simplifies these support vector approaches into a method that can be used for regression with minimal training data. ELM does not need to train iteratively as it is a single hidden layer feed-forward network without backpropagation. Since it does not use gradient descent optimization methods associated with many neural network methods, it has faster learning speeds, up to 1000 ×, over traditional machine learning approaches and can be trained on significantly less data [34]. ELM has better scalability than support vector machine methods and similar or better performance [34,36,37]. We selected an ELM model to use in this research, as ELM performs well on small datasets and is very efficient in terms of training time. For our problem, we may only have 10 to 20 years' worth of monthly data, or on the order of a 100–200 data points for a given well or model, a very small dataset for most machine learning applications.

An ELM has the form of Equations (1) and (2), where W_1 is the matrix of input-to-hidden-layer weights, X represents the input data, σ is an activation function, $\vec{b}$ is a bias vector, and W_2 is the matrix of hidden-to-output layer weights. The rectifier activation function, Equation (2), is used to compute σ. This process is summarized in Figure 1, where $a_1 \ldots a_N$ represent the rows of the matrix W_1.

$$\vec{Y} = W_2 * \sigma\left(W_1 X + \vec{b}\right) \tag{1}$$

$$\sigma(x) = \max(0, x) \tag{2}$$
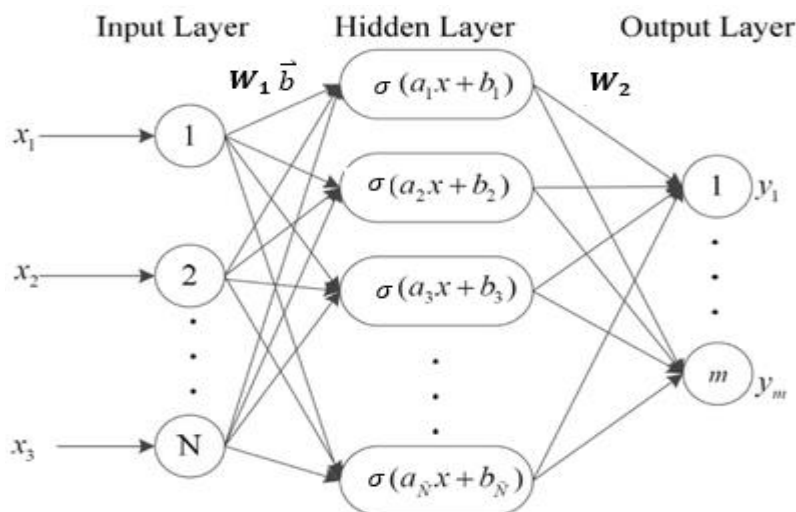
Figure 1. Schematic of Extreme Learning Machine showing the relationship between the input data (*x*) and the output data (*y*). This is a feed-forward network, without backpropagation, giving very fast training times and requiring less data for accurate computations.

The ELM maps non-linear relationships between the input data and the training data by means of the random weight matrix W_1 and the rectifier activation function, σ. For our applications, we use ELM to map the non-linear relationship between Earth observation-based input and groundwater levels.

Once we compute this mapping using a feed-forward pass with a least-squares fit, we use the model to infer missing data using only Earth observation data as inputs. Since the ELM uses a least squares fit with only one computational step, and does not require the iterative steps of a back-propagation network using gradient descent algorithms, it is efficient and works well with small datasets.

The algorithm to create an ELM for a given well has three general steps:

1. Fill W_1 and $\vec{b}$ with random values
2. Compute σ, using the rectifier activation function, Equation (2)
3. Calculate the output weights, W_2, by performing a least-squares fit to a vector of the response variables, $\vec{Y}$.

For data imputation, the $\vec{b}$ vector and both the W_1 and W_2 matrixes are saved and used to impute the missing values using Equation (3).

$$\vec{Y}_{test} = W_2 \sigma\left(W_1 X + \vec{b}\right) \tag{3}$$

where X represents the Earth observation data over the imputation period.

Our use of an ELM method for data imputation assumes that groundwater levels are correlated with climatic factors such as drought and soil moisture, both of which can be estimated using satellite Earth observations combined with hydrological models and have trends that can be characterized using the data measured at a well. Figure 2 shows an example of this correlation, with depth to water table data shown in the upper panel, while the GLDAS soil moisture estimates at the pixel nearest the well shown in the bottom panel. In Figure 2, the water table decreases and increases in lagged correlation with periods of decreased or increased soil moisture estimated by the GLDAS model. The figure shows that groundwater levels generally decrease during periods of drought (i.e., negative soil moisture anomalies shown as red areas below the line), and increase during wet periods (i.e., positive soil moisture anomalies shown as blue areas above the line). Figure 2 shows an apparent correlation between soil moisture and groundwater levels, but this correlation is highly non-linear and appears to be lagged in time. The ELM estimates these non-linear correlations and exploits them to impute missing data at a well.
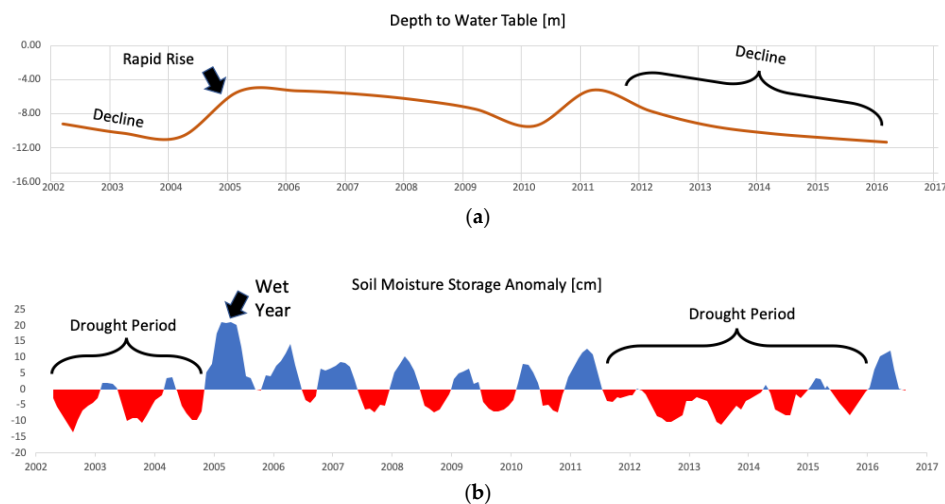
Figure 2. Correlation between soil moisture and groundwater levels near Beryl, Utah: (**a**) Depth to Groundwater at Well 37349113434201 near Beryl, Utah; (**b**) Soil Moisture Liquid Water Equivalent Thickness Simulated by GLDAS near Beryl, Utah showing the lag between high soil moister in late 2004–2005 and a corresponding increase in the groundwater level in early 2005.

1.6. Research Overview

The purpose of this research was to develop a method of imputing missing groundwater level observations at a monitoring well by merging existing well data and global Earth observations. This method incorporates easily available global gridded datasets and in-situ well observations, allowing its potential use throughout the world. The missing groundwater level observations are imputed using an ELM with PDSI and the GLDAS and CPC soil moisture models as inputs. We use ELM because it is a single hidden layer feedforward network that can be rapidly trained with minimal data while providing very good accuracy. We present case studies where we test the method in the Cedar Valley and Beryl Enterprise areas of southern Utah, USA. In our case study areas, groundwater level estimates obtained using ELM were more accurate than estimates obtained from Kriging spatial interpolation which only used in-situ well data.
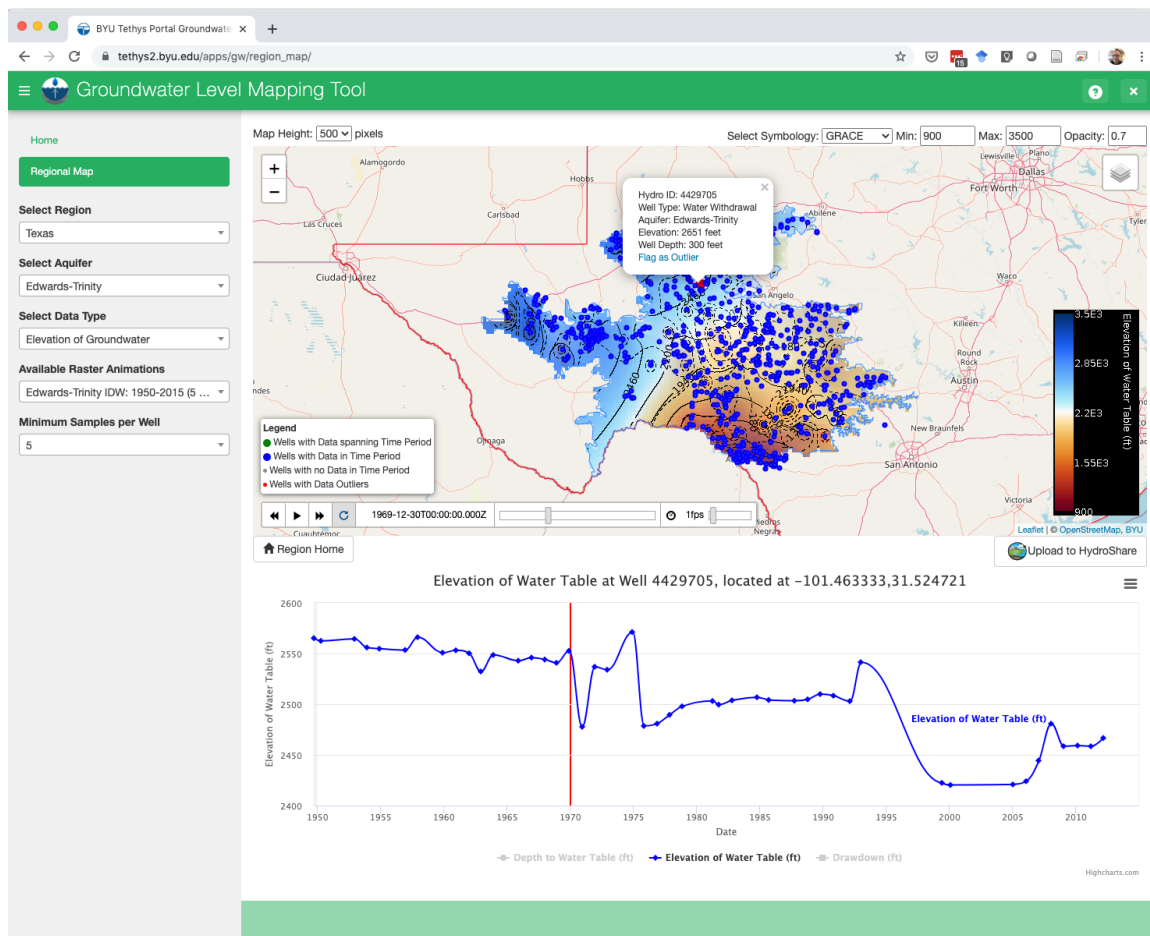
2. Methods

2.1. Overview and Data Sources

The imputation approach we present requires three broad steps: (1) interpolate available historic data at a well to a regular time step, typically quarterly; (2) combine these measured water surface elevation data with Earth observation data to train a machine learning algorithm to identify patterns and correlations in these data; and (3) use the trained model to impute the missing surface elevation data at that well. This process is then repeated for each well, with each well having its own model. The size of the data gap or length of the prediction that can be accurately addressed depends on the behavior of a given well. If trends and patterns are well defined and repeat, then this approach is relatively accurate over relatively long periods and can be used for data extrapolation or out-of-sample estimation for any period in which the Earth observation data are available. For well data with less well-defined patterns, the periods where the data imputation results are accurate are shorter.

We obtained well data from the United State Geologic Survey National Water Information System (NWIS) web service [38]. Individual groundwater level measurements were reported as depth to groundwater for each well, while each well record includes information including well head elevation and the aquifer in which the well is completed. Using this well information, we converted depth-to-groundwater data into groundwater surface elevation data.

We obtained the PDSI and CPC datasets from a THREDDS server hosted by NOAA [29]. The PDSI global dataset is available at 0.5-degree resolution from 1850–2015 at monthly frequency [39]. CPC soil moisture data are available globally at 0.5-degree resolution from 1948 onwards at monthly frequency. We obtained the GLDAS soil moisture data from the National Aeronautics and Space Administration (NASA) web server operated by the California Institute of Technology Jet Propulsion Laboratory [31]. GLDAS data are available monthly from 2001 onward at a resolution of 0.25°. While initially we downloaded and used the GLDAS data, the results reported in the case studies do not use these data because as they are only available from 2001 on.

We implemented all the methods reported here in an open-source, web-based application called the Groundwater Level Mapping Tool (GLMT) [40–42]. The interface for the GLMT is shown in Figure 3. GLMT is a visual tool and can be used to display data for any number of regions with a number of aquifers in each region. GLMT is generalized to allow its use world-wide. GLMT helps decision makers: view time series and other data for each well within an aquifer, view maps of aquifer-wide groundwater levels at different time periods, and calculate and view estimates of changes in total aquifer storage.

Figure 3. The user interface for the Groundwater Level Mapping Tool which was used to package the algorithms and methods described in this manuscript. This figure shows an aquifer in Texas along with the time history of water levels for a selected well.

GLMT ingests well data and retrieves the required remote sensing information. Then, using the algorithms and approaches presented here, generates a time series of groundwater level maps and uses these maps to compute the change in aquifer storage over time. For this work, we extended GLMT by providing an algorithm to retrieve and ingest the Earth observation data from web services; resample the well data and remote sensing data to a common, regular time step; identify wells with missing data; generate the input data for the ML model; train a ML model for each well; and impute the missing data at a well. GLMT performs this data imputation algorithm for each well in the aquifer, then uses the time series generated at each well location to compute water level maps through time. We implemented the computational routines using the Python programing language with extensive use of the NumPy, SciPy, Hydrostats, and HydroErr python libraries [43–46], while the web interface was developed using Tethys Platform, a library for developing environmental web applications [40,47].

2.2. Data Preparation and Interpolation

As input to the ELM, we used groundwater level measurements for a given well, and two soil moisture datasets based on Earth observations, PDSI, and CPC. We used values taken from the grid cell closest to the location of the target well for the gridded monthly Earth observation data. We extracted these data

directly from the NOAA THREDDS server where the data from these datasets are hosted. We resampled or interpolated these data to a three-month interval using a cubic hermite interpolating polynomial (PCHIP) method [48].

Well data are often irregularly sampled, with nominal sampling times being monthly, quarterly, semi-annually, or annually. For our algorithm, we re-sampled these measurements to a common time step. During the course of this work we used data from various aquifers in Utah, Texas, Colombia, South Africa, and the Dominican Republic. In all cases, the well records were quite irregular and inconsistent, even for the same aquifer. While we reviewed and analyzed data from numerous aquifers, in this paper we only present case studies for two Utah aquifers. We resampled these data using PCHIP algorithms, and only resampled data where the time interval between data points was short—less than 90 days.

PCHIP is often selected to resample environmental data over other methods, such as cubic splines or linear interpolation, as it provides a continuous first derivative and does not exceed the local physical data bounds—in other words, no overshoot or undershoot occurs, which is common with splines [9,11]. We used the PCHIP implementation from the Python Pandas library [45,49]. The PCHIP method interpolates data using a piecewise cubic polynomial with continuous first derivatives. PCHIP interpolation preserves the shape of the data and respects monotonicity. At intervals where the data are monotonic (i.e., neither increasing nor decreasing), so is the interpolating function, and at points where the data have local minima or maxima, so does the interpolating function, thus eliminating over- or under-shoot. The second derivative is typically not continuous, though the interpolating function is smooth with a continuous first derivative [48].

We used a total of eleven time series datasets for input to the ELM model, the two monthly Earth observation sets, eight moving average datasets of the monthly data (4 averages for each observation set), and one linear dummy dataset.

We computed the eight rolling average datasets using the interpolated monthly data from the Earth observation datasets, PDSI and CPC. For each of these datasets, we computed 1-, 3-, 5-, and 10-year moving averages, resulting in four additional time series for each dataset. For the rolling average computations, we used the "rolling" method from the Python Pandas Python library [49]. This computes the rolling average at the right side of the window, meaning all the data used to compute the average at a given time step were from the past.

These rolling averages improve the accuracy of our imputation method as the varying window lengths allow the model to capture the effects of both long- and short-term climatic events, which contribute to changes in groundwater levels. Often, groundwater levels increase fairly quickly following an extremely wet month or year. These changes are captured in the model by incorporating the 1-year average and the original monthly data. When moisture levels are slightly below normal for extended periods of time, this generally causes a period of slow decline in groundwater levels. This phenomenon is captured well by the 5 and 10-year rolling averages.

For model input, we included a dummy linear dataset in the model to capture any constant underlying linear trends in groundwater levels not present in the Earth observation data [34]. This pseudo-linear dataset is especially important in cases where the groundwater is being constantly over pumped because it allows the model to capture this near linear decrease in the water table cause by continuous over pumping. We generated the pseudo-linear data set using the decimal year value of the measurement and normalized these values from 0–1 over the time span of the model.

ELM requires data normalized to a range between 0 and 1. We normalized all input data streams and the training data (i.e., measured data) stream using Equation (4), which generates normalized

data R_{norm} from the raw data, R. For ELM development, the model is trained on the available data, R_{train}. R_{train} represents the available well data used to train the model.

$$R_{norm} = \frac{R - \min(R_{train})}{\max(R_{train}) - \min(R_{train})} \tag{4}$$

The minimum and maximum values for each dataset are saved and used to normalize the values used in the imputation step so the model will use the same range as used in the training dataset. This means that the model input and output values may exceed the 0–1 limits. After imputation, the imputed water level estimates, R_{norm}, are transformed back for analysis and presentation using Equation (5) and the maximum and minimum water level values saved from the training dataset, $\max(R_{train})$ and $\min(R_{train})$, respectively.

$$R = R_{norm}(\max(R_{train}) - \min(R_{train}) + \min(R_{train}) \tag{5}$$

2.3. Data Imputation Using Extreme Learning Machines

For development and demonstration, we selected groundwater wells with long, complete time histories and separated these data into training and testing datasets. The testing datasets are periods for which we have data, but we did not use these data to train the model. We used the testing data to represent missing data that require imputation and allow us to evaluate the accuracy of the imputation. We trained an ELM model for a given well using historical well data and Earth observation data over the same period. We then used the resulting model to impute data over the testing period using only Earth observation data over this testing period. We then compared these imputed data to testing data measured during this period to evaluate accuracy.

For illustration purposes, the training and use of an ELM model can be presented as a series of steps as follows:

Let:

- M be the number of samples in the training dataset (i.e., measured time steps),
- N be the number of input data time series (for our case, $N = 11$),
- $\vec{Y}$ be a vector of length M, the number of the groundwater level measurements over the training period,
- X be an MxN matrix of the input data over the training period,
- h be the number of nodes in the hidden layer (for our case, $h = 500$),
- $\vec{b}$ be vector of length h,
- W_1 be an Nxh matrix,
- W_2 be a matrix of hidden-to-output weights, and
- σ is the rectifier function (Equation (2)).

First, we populate the matrix W_1 and the vector $\vec{b}$ with random values. The W_1 matrix and the $\vec{b}$ vector are saved for use later in the imputation step. Next, we solve for W_2, beginning by computing A using Equation (6).

$$A = \sigma\left(W_1 X + \vec{b}\right) \tag{6}$$

Then we use the measured values $\vec{Y}$ and the computed matrix A to estimate the matrix W_2, which minimizes the error, e, shown in Equation (7).

$$\vec{Y} = W_2 A + e \tag{7}$$

We estimate the matrix W_2 using a regularized least squares method shown in Equation (8), where λ is a regularization parameter and I is an identity matrix.

$$A^T\vec{Y} = \left[A^TA + \lambda I\right]W_2 \tag{8}$$

Now that W_1 has been computed, we have a fitted ELM model that can be used solve for $\vec{Y}$ as shown in Equation (9). This step uses only Earth observation data to impute missing measured well data at a given well. This process is repeated and an ELM model is generated for each well.

We compute W_2 by solving Equation (8) using the least squares function in the Numpy Python library [46]. For our work, we use the Python library, Statsmodels, and the OLS.fit routine [50], which uses the underlying Numpy library. The OLS.fit routine uses the Moore–Penrose pseudoinverse to invert the matrix [51]. The Moore–Penrose pseudoinverse can act as a replacement matrix inverse and is often used to solve a system of linear equations when the system does not have a unique solution or has many solutions. It is commonly used in machine learning methods where the matrix to be inverted is often ill-conditioned. Figure 4 provides pseudo code that demonstrates computation of the model and data imputation for a single well.

To impute missing data, the W_1 and W_2 matrices and $\vec{b}$ vector saved from the previous computations are used. Imputation is computed as follows:

- $\vec{Y}$ is a vector of length M, that will contain estimated groundwater levels for a well
○ M is the number of time steps to impute,
- X is a MxN matrix of data used to infer groundwater depth
○ N is the number of input data series (17 for this paper) for imputation,

$$\vec{Y} = W_2\sigma\left(W_1X + \vec{b}\right) \tag{9}$$

After the ELM imputation step (Equation (9)), then $\vec{Y}$ must be rescaled, yielding the depth to groundwater estimates $\vec{Y}_{estimate}$ in correct units using Equation (10). The minimum and maximum values of $\vec{Y}$ were saved from the training data and used to re-scale the imputed data. In Equation (10), $\vec{Y}_{estimate}$ represents the rescaled imputed data, $\vec{Y}$ represents the output of the ELM model or the normalized imputed data, and $\vec{Y}_{train}$ represents the data used for training the ELM model used to determine the minimum and maximum used to rescale the imputed data.

$$\vec{Y}_{estimate} = \vec{Y} * \left[\max\left(\vec{Y}_{train}\right) - \min\left(\vec{Y}_{train}\right)\right] + \min\left(\vec{Y}_{train}\right) \tag{10}$$

2.4. Accuracy

To assess the accuracy of the imputed data, we report four different error metrics, These are: the mean error (ME), the mean absolute percentage error (MAPE), the root mean squared error (RMSE), and the coefficient of variation (R^2) [44]. To compute these error metrics, we used the python package HydroErr [43].

```python
# generate random input weights, W1
W1=np.random.normal(size=[N, h]) # np is the numpy library
 # np is the numpy library
 # X is a matrix of input training data, size MxN
 # M is number of input series (17 in our case)
 # N is the number of input time steps
 # h is the number of neurons (500 in our case)
# generate bias vector, b
b=np.random.normal(size=[h])


# generate the A matrix (Eq 6)
a=np.dot(X,W1)+b
theta=np.maximum(a,0,a) # basis function


# compute the output weights, W2
I=np.identity(N)      # generate an identity matrix, NxN


# fit using numpy library, uses a Moore-Penrose pseudoinverse of the matrix
W2=np.linalg.lstsq(theta.T.dot(theta) + lamb * I, theta.T.dot(Y)[0]
 # lamb is the regularization parameter (in our case 100)
 # Y are the observed well data
 # this is from Eq 8


# Predict – X is a matrix of Earth observation data, W1, W2, and b are saved from above
a=np.dot(X,W1)+b
theta=np.maximum(a,0,a) # basis function
Y_imputed= np.dot(theta,W2)


# Y_imputed needs to be re-scaled before use
```

Figure 4. Pseudo code for Extreme Learning Machine (ELM) imputation following python syntax, where **X** is a matrix of Earth observation data, **Y** is a vector of training data, **W1**, **theta**, and **W2** are variables, **X1** is a matrix of each observation data over the imputation target period, and **Y_imputed** are the unscaled imputed data. **X**, **X1**, and **Y** are scaled to the limits of the training data, **Y_imputed** needs to be re-scaled to obtain actual values.

We use the ME to evaluate if the imputed data are biased. If the ME is positive, then the imputation is systematically over predicting the values, and conversely under predicting values if it is negative. The ME is defined as:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) \tag{11}$$

where y_i is the imputed data value at time step i, and x_i is the measured value at the time step.

We use MAPE and RMSE to evaluate the accuracy of the imputed data. MAPE ranges from 0 to infinity, $0 \le RMSE < \infty$, is unit-less and presents the mean absolute error as a percentage of the observed values. MAPE allows data sets with different magnitudes to be compared; smaller is better. The MAPE is defined as:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - x_i}{x_i}\right| \tag{12}$$

The RMSE ranges from 0 to infinity, $0 \le RMSE < \infty$, in the same units as the data set. For RMSI, smaller is better and this metric highlights larger errors. The RMSE is defined as:

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}|y_i - x_i|^2\right)^{1/2} \tag{13}$$

R^2, the coefficient of variation ranges from 0 to infinity, $0 \le RMSE < \infty$, and is unit-less. It does not indicate bias, a number closer to 1 indicates a better match. The RMSE is defined as:

$$\begin{aligned}
SS_{xx} &= \sum_{i=1}^{n}(x_i - \bar{x})^2 \\
SS_{yy} &= \sum_{i=1}^{n}(y_i - \bar{y})^2 \\
SS_{xy} &= \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
R^2 &= \frac{SS_{xy}^2}{SS_{xx}SS_{yy}}
\end{aligned} \tag{14}$$

Each of these error metrics help characterize the accuracy of the imputed data and allow us to better understand how well the method works.

3. Application and Case Studies

3.1. Application

Figure 5 shows an example of the results of our data imputation method. The imputed data are generated using only the trained model with Earth observation data as input. In Figure 5, the orange dots represent measured water level data at the well, the light blue crosses represent data generated by the PCHIP interpolation algorithm, and the solid blue line represents the data values estimated using the ELM model. We used the PCHIP algorithm to interpolate values at the start of each month, so all data are on the same time step. We used PCHIP to interpolate over gaps of up to 365 days, any larger gaps are filled using the ELM model. At the scale of the plot, these small gaps interpolated using PCHIP cannot be seen. In the original data, many months had two or more measured points, many had none, and few of the measurements occurred on the first day of the month. We used the PCHIP interpolation to generate a data set with one measurement per month, with all the measurements on the first of the month—these are the blue crosses in Figure 5. The Earth observation data were generated with values at the beginning of

each month. The ELM model also imputes data on the first of each month. In Figure 5, the imputed data, over both the training data and the gaps, demonstrates how the penalty function that the ELM model uses approximates, but does not exactly match, the measured data used for training. The next step, which is not shown, would be to replace the data gaps in the measured data with the imputed data, resulting in a time series with no missing values. The process shown in Figure 5 is repeated for each well.
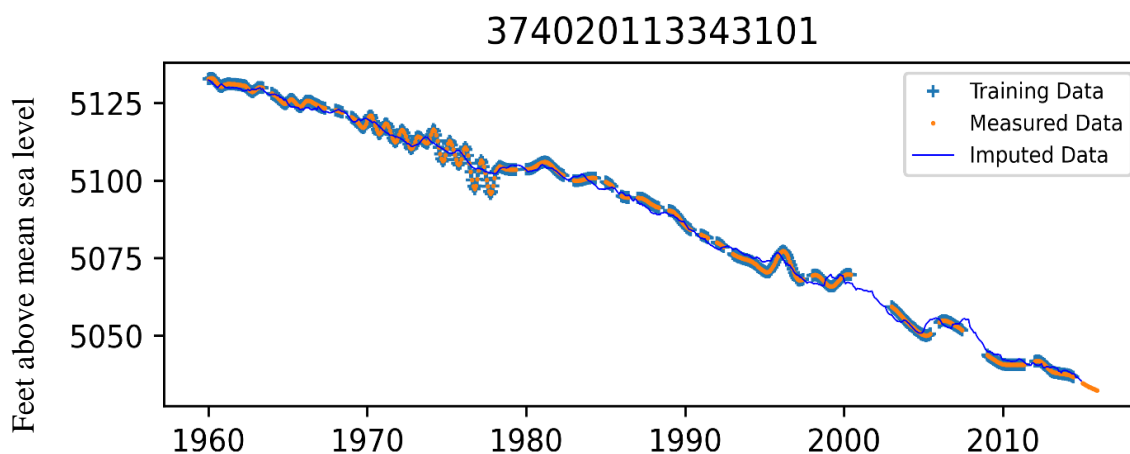
Figure 5. Example data showing measured data (orange dots), training data (blue crosses), and imputed data (blue lines) over the entire period. All measured data were used to train the model used to impute data for the gaps. This figure includes imputed data over the entire period to show how the ELM optimal fit minimizes the error but does not exactly match the measured data.

When gap filling in practice, we will always use all the available measured data to train the model, as was done in Figure 5. In the case studies in the next section, we only used data from the period from 1958–1999 to train the model; we imputed data not only over gaps in the training data set, but over the entire 15 year period from 2000–2015, a very long period. This was done so that measured data could be used to compute the accuracy of the imputed data. This is a conservative estimate of the accuracy of the method for data gap imputation for two reasons: the gap is very large and it is an extrapolation. Most gaps have data on either side of the gap, which, if used, result in a more accurate model.

In the two case studies that follow, we present a "worst case" scenario for data imputation and a more typical case. For the worst case, we create missing data over a 15-year time period from 2000–2015. This is worst case in two respects: it is a very large gap, and it is at the end of the measured data—that is it requires extrapolation. For most gaps, you use data on either side of the gap to anchor the imputation; this is not available for these gaps. For the more typical case, we create missing data over a 7-year time period from 1995–2002; this period was randomly selected. This period, while more typical, still represents a very large gap, much larger that gaps that commonly exist. Even with these large gaps, our case studies show that our ELM method provides an accurate imputation, with MAPE values, for the 7-year gap scenario, of less than 0.1 and most significantly smaller.

The worst case scenario shows that the ELM method can be used for extrapolation, though with lower accuracy. However, in cases where no other data are available, this extrapolation can result in reasonable estimates. The case studies show that interpolation is more accurate than extrapolation, and interpolating over small gaps is more accurate than over larger gaps.

3.2. Cedar Valley, Utah Case Study

We present a case study using the ELM method in the Cedar Valley Aquifer in Utah, USA which is located just west of Kanarraville, Utah in the Cedar Valley. We used five wells to perform this test, the results from three of these wells are shown in Figures 6 and 7, while the accuracy results are shown in Table 1. We selected these wells because they had data available over the entire test period of 1958–2015, which allowed us to evaluate the accuracy of the method by comparing imputed data values to actual measured values, while retaining data for model training. We imputed or predicted data over the entire data record, which includes both the training and testing periods. We show the imputed data over the training period to demonstrate the fit between the model and the available data. We show the imputed data over the testing period to characterize the model fit using only Earth observation data for imputation. ELM is a least-squares method that minimizes the error of the imputed data in the training period. It does not match these data exactly, as can be seen in the figures.
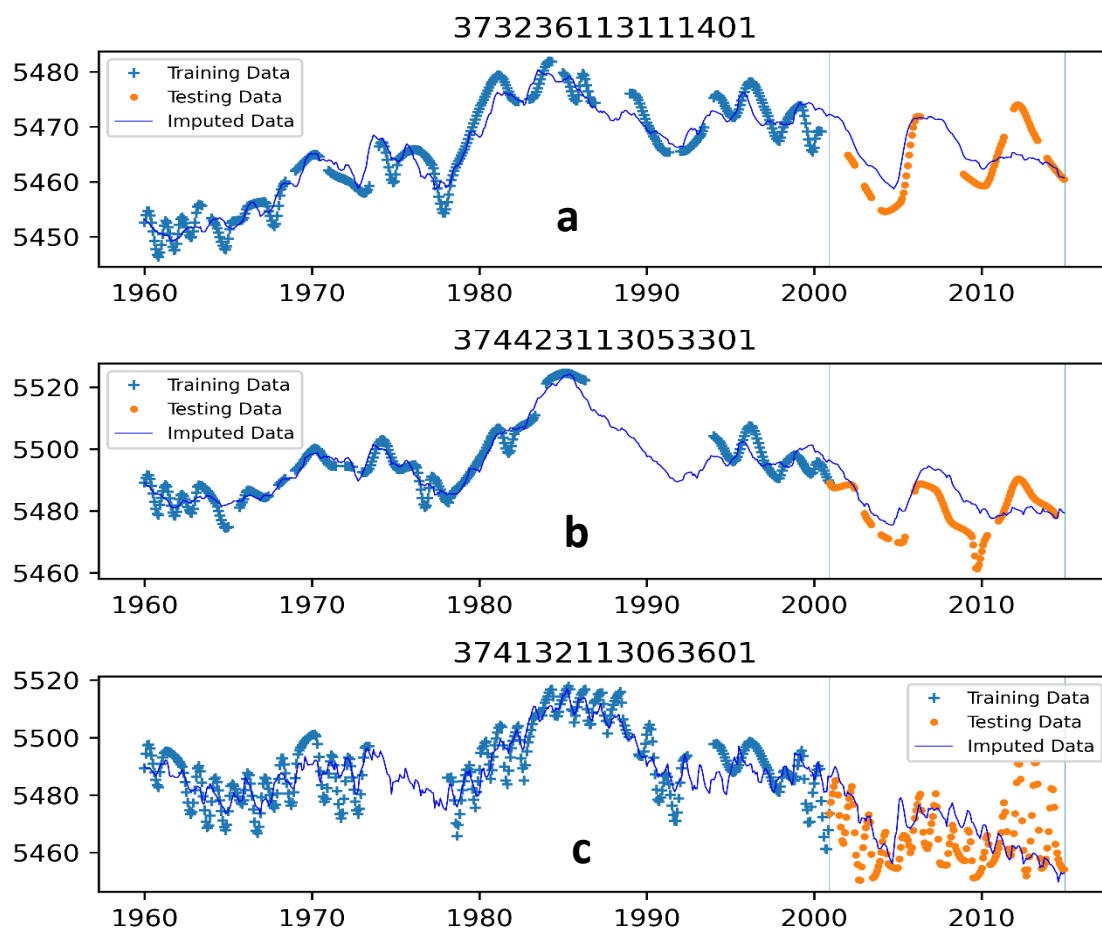
Figure 6. ELM results over a 15 year testing period from 2000–2015 for Wells 37236113111401, 374423113053301, and 374132113063601 as subfigures (**a**–**c**), respectively. Imputed data are solid blue lines, measured data used for training are blue crosses, and measured data used for testing are orange dots.
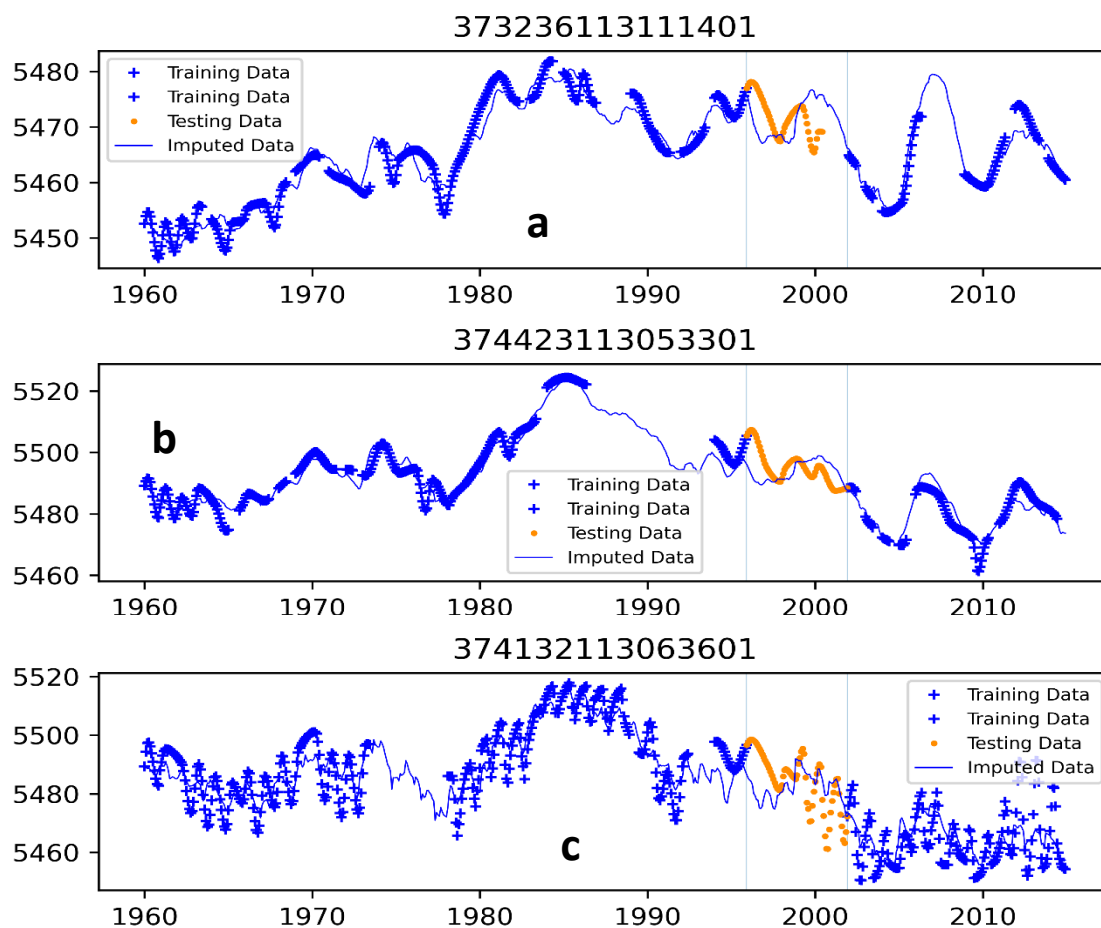
Figure 7. ELM results over a 7 year testing period from 1997–2002 for Wells 37236113111401, 374423113053301, and 374132113063601 as subfigures (**a**–**c**), respectively. Imputed data are solid blue lines, measured data used for training are blue crosses, and measured data used for testing are orange dots.

In Figure 6, we used data from 1958–1999 as training data for the model, and data from 2000–2015 to test the algorithm. In Figure 7, we used data from 1995–2002 to test the model and the remainder of the data to train the model. Figure 6 shows data extrapolation beyond the training period; a more common use is shown in Figure 7, where data are imputed across gaps that have measured data on each side of the gap. We do not expect this to be used for gaps as large as those used in these examples—15 years and 7 years for Figures 6 and 7, respectively. Imputation methods, including our ELM model, are more accurate over smaller gaps with supporting data on both sides of the gap. However, the large gaps provide a rigorous test of the algorithm.

Figures 6 and 7 shows the imputed data (solid blue line), the measured data used for training (blue crosses), and measured data used for testing (orange dots) for Wells 37236113111401, 374423113053301, and 374132113063601, for Figure 6a–c, respectively. In Figure 7, the subfigures are labeled Figure 7a–c. We selected these three wells from the five used in the case study as each demonstrates a specific issue.

For the well in subfigures a and aa, the depth to groundwater varies between 25 and 60 ft and the imputed depth to groundwater values match the measured data well for both testing periods. This is true even though the testing period shows fluctuations and is not a linear trend. The imputed data in both figures shows a structure with a rising and falling patterns repeated during the testing period.

Table 1. ME, MAPE, RMSE and R^2 values for data imputed using the ELM model for Five Wells in Cedar Valley for both the 2000–2015 and 1995–2002 testing periods. Imputed data were compared to measured data in each period.

Well ID	ME	MAPE	RMSE	R^2
Error metrics for the 2000–2015 Testing Period				
Well 373236113111401	0.392	0.073	4.754	0.313
Well 374132113063601	−12.253	0.276	20.657	0.045
Well 374304113052901	9.701	0.199	17.75	0.181
Well 374423113053301	4.458	0.13	8.555	0.255
Well 374927113033401	6.187	0.129	8.221	0.37
Error metrics for the 1995–2002 Testing Period				
Well 373236113111401	−0.101	0.047	3.615	0.701
Well 374132113063601	−0.116	0.125	8.724	0.584
Well 374304113052901	1.363	0.097	9.451	0.672
Well 374423113053301	0.002	0.068	4.981	0.746
Well 374927113033401	0.224	0.06	4.332	0.763

Figures 6b and 7b demonstrate a good fit, but the imputed data in Figure 6b overestimates the groundwater levels, a trend present in Figure 6a, but to a smaller extent. In Figure 7b, this overestimate is not present as there are training data on both sides of the gap. This consistent overestimation in all the wells for the 2000–2015 testing period possibly occurs because of increased human pumping activity, which may have changed in the last 15 years. This change would not be explicitly represented in the training data used to create the model for this well. In a more common approach, such as the data from the 1995–2002 testing period where we fill smaller data gaps, we have training data in the later periods on either side of a gap being filled, better capturing this trend. The correlation between soil moisture variables and pumping rates assumes that the relationship present in the training data is continued in the testing period. However, if, as in the case of the 2000–2015 testing period, we attempt to impute data over a long time period where pumping behaviors may have changed, perhaps because of land-use or population changes, these correlations will not be reflected in the model and result in less accurate imputation.

Figures 6c and 7c capture the general trends of the groundwater fluctuations. However, the model for this well fails to properly capture the magnitude of the seasonal variation of the water table throughout the year. This most likely occurred because of the inconsistent sampling rate of the well and perhaps issues with measurement as demonstrated by the variability of the measured data (orange dots). This well was sampled at a monthly frequency during part of the training period, but not for the entire duration. This mixture of higher and lower sampling rates in the training data seemed to negatively impact the model as the seasonal correlations for this well model were not fully characterized in the training data. In addition, the large seasonal variations present in the testing period are not represented in the training data.

Table 1 presents the computed accuracy of the imputed data over both testing periods, 2000–2015 and 1997–2002. We present the four different metrics described above: ME, MAPE, RMSE and R^2 [43,44].

The ME values for the two different time periods show that the bias in the imputed data is significantly less for the shorter time period with training data on both sides of the gap. The MAPE is approximately half the value for the shorter time period, with significant improvements in the RMSE as well. R^2 values for the longer period are not good, but are acceptable for the shorter period.

As we discussed in the background section, our major use for data imputation of well measurements is to develop a series of groundwater elevation maps to characterize the change in storage for an aquifer. Traditionally, if a well has no data for a given water level map time step, that well is left out of the data used by the spatial interpolation code; in our case, we use Kriging. This means that the water surface elevation for that time step, at the location of the well with missing data, is estimated (or imputed) by the spatial interpolation scheme. We compared the ELM imputation results to results computed using Kriging spatial interpolation, using a jackknife approach [52]. For jackknifing, we remove one well from the set of wells and use the remaining wells to perform the Kriging interpolation at the missing well location. Then we put this well back into the dataset and remove the next well for prediction. We compared the Kriged results at the well location with the value ELM imputed value. We computed this test for the same five wells in the Cedar Valley Aquifer. We only performed this comparison for a single time step, December 1, 2014, nearly at the end of the imputed period. Table 2 presents these results along with the percent error of each method.

Table 2. Comparison of data imputed with the ELM model and data estimated using Kriging on 1 December 2014.

| Well ID	Measured Value	ELM Estimate		Kriging Estimate	
	Depth to GW (ft)	Depth to GW (ft)	(% error)	Depth to GW (ft)	(% error)
Well 373236113111401	−47	−41	(12.9%)	−65	− (39.0%)
Well 374132113063601	−99	−96	(2.8%)	−109	− (10.1%)
Well 374304113052901	−110	−106	(3.1%)	−98	(10.9%)
Well 374423113053301	−76	−65	(14.6%)	−86	− (14.0%)
Well 374927113033401	−26	−30	− (14.1%)	−33	− (28.1%)
Average Error			9.5%		20.4%

3.3. Beryl-Enterprise Area, Case Study

We also tested our method on the Beryl Enterprise area aquifer in southern Utah. This region has experienced significant drawdown over the last century, especially in the southern portion of the aquifer. This area has undergone significant recent population growth and land use changes. We do not expect the current groundwater use patterns to be similar to those in the past. We selected ten wells in the aquifer for testing with accuracy results presented in Tables 3 and 4. Figure 8 presented plots of three wells we selected to highlight various issues. Figure 8 presents Wells 373338113431502, 373419113434201, and 373527113415101, as Figure 8a–c, respectively. While Figure 9 presents the same wells as Figure 9a–c, respectively. The data are presented in the same manner as previous figures, measured data in the testing and training periods as blue and orange dots, respectively, with the blue solid line representing the imputed data.

Table 3. ME, MAPE, RMSE and R^2 values for ELM Time Series Prediction for Ten Wells in Beryl Enterprise Aquifer for both the 2000 – 2015 and 1995 – 2002 testing periods.

Well ID.	ME	MAPE	RMSE	R^2
Error metrics for the 2000–2015 Testing Period				
373338113431502	1.57	0.06	4.00	0.71
373419113434201	−0.23	0.05	3.02	0.74
373527113415101	18.53	0.36	20.29	0.62
373644113411501	18.24	0.36	19.28	0.70
373735113393801	16.25	0.33	17.96	0.69
373854113411501	14.81	0.30	16.29	0.79
374020113343101	14.25	0.28	14.86	0.80
374041113373501	11.30	0.23	12.66	0.75
374053113415101	19.95	0.40	21.39	0.65
374319113415201	3.02	0.07	3.70	0.89
Error metrics for the 1995–2002 Testing Period				
373338113431502	0.23	0.04	2.38	0.81
373419113434201	0.57	0.04	2.74	0.77
373527113415101	−2.44	0.11	7.07	0.80
373644113411501	−1.19	0.06	3.93	0.97
373735113393801	−0.71	0.08	4.56	0.94
373854113411501	−0.98	0.05	3.12	0.99
374020113343101	0.07	0.03	1.92	0.98
374041113373501	−0.08	0.02	1.40	0.99
374053113415101	−1.02	0.05	3.30	0.98
374319113415201	0.28	0.03	1.68	0.97

Table 4. Comparison of ELM Error to Kriging Error for Wells in the Beryl-Enterprise Area. The comparison is made using data from 1 December 2014.

| Well ID	Measured Value	ELM Estimate		Kriging Estimate	
	Depth to GW (ft)	Depth to GW (ft)	(% error)	Depth to GW (ft)	(% error)
Well 373338113431502	−39	−47	(21.8%)	−58	(48.9%)
Well 373419113434201	−35	−40	(13.6%)	−64	(82.6%)
Well 373644113411501	−246	−211	(14.2%)	−172	(30.1%)
Well 373735113393801	−230	−199	(13.7%)	−222	(3.6%)
Well 373854113411501	−216	−182	(15.7%)	−217	− (0.6%)
Well 374020113343101	−192	−172	(10.8%)	−187	(2.6%)
Well 374041113373501	−166	−148	(10.7%)	−176	− (6.4%)
Well 374053113415101	−198	−173	(12.5%)	−192	(2.9%)
Average Error			(14.1%)		(22.2%)
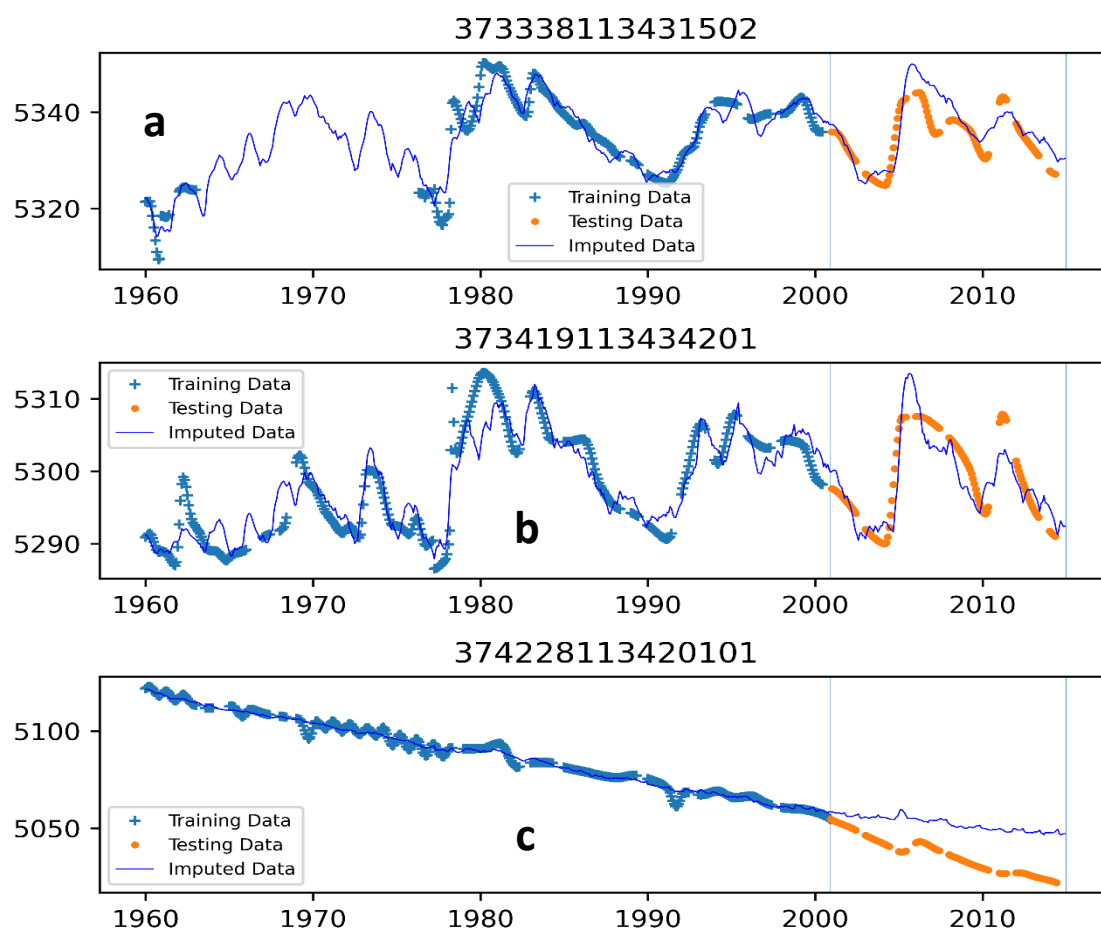
Figure 8. ELM results over a 15 year testing period from 2000–2015 for Wells 373338113431502, 373419113434201, and 373527113415101, as subfigures (**a**–**c**), respectively. Imputed data are solid blue lines, measured data used for training are blue crosses, and measured data used for testing are orange dots.

Results for the seven wells that are not shown were not as good as those for the wells shown in the plot. Most of the wells not shown had patterns that were similar to those shown in subfigure f for the longer testing period, 2000–2015. For this testing period, the measured data show a significant downward trend not captured in the model training on the earlier period. For this time period in all ten wells, the ELM model correctly identified periods of increase and decrease in water levels, but generally failed to properly estimate the magnitude of these changes, resulting in an underestimation of drawdown for each of the tested wells.

Figure 9 shows these same wells for the testing period 1995–2002; in this case, there are data on each side of the gap and the model clearly captures this increasing downward trend in the latter part of the data set. Table 3 shows the ME, MAPE, RMSE, and R^2 values for all ten wells in each testing period. For the longer period (2000–2015), the ME shows a clear over prediction bias with all but 3 values near 20 feet, these same wells for the shorter period show almost no bias with the largest ME value being only −2.44 feet. The MAPE, RMSE, and R^2 values also all show significant improvement. These data show that if aquifer processes are significantly different during the imputation period—in this case, significantly increased pumping—imputation will not be accurate. However, for gap filling—even for gaps as long as 7 years—if the training data capture the aquifer behavior, imputation is quite accurate. Figures 8c and

9c show this behavior the most clearly. As noted above, most of the wells not plotted showed a similar pattern, with drawdown rates increasing towards the end of the period.
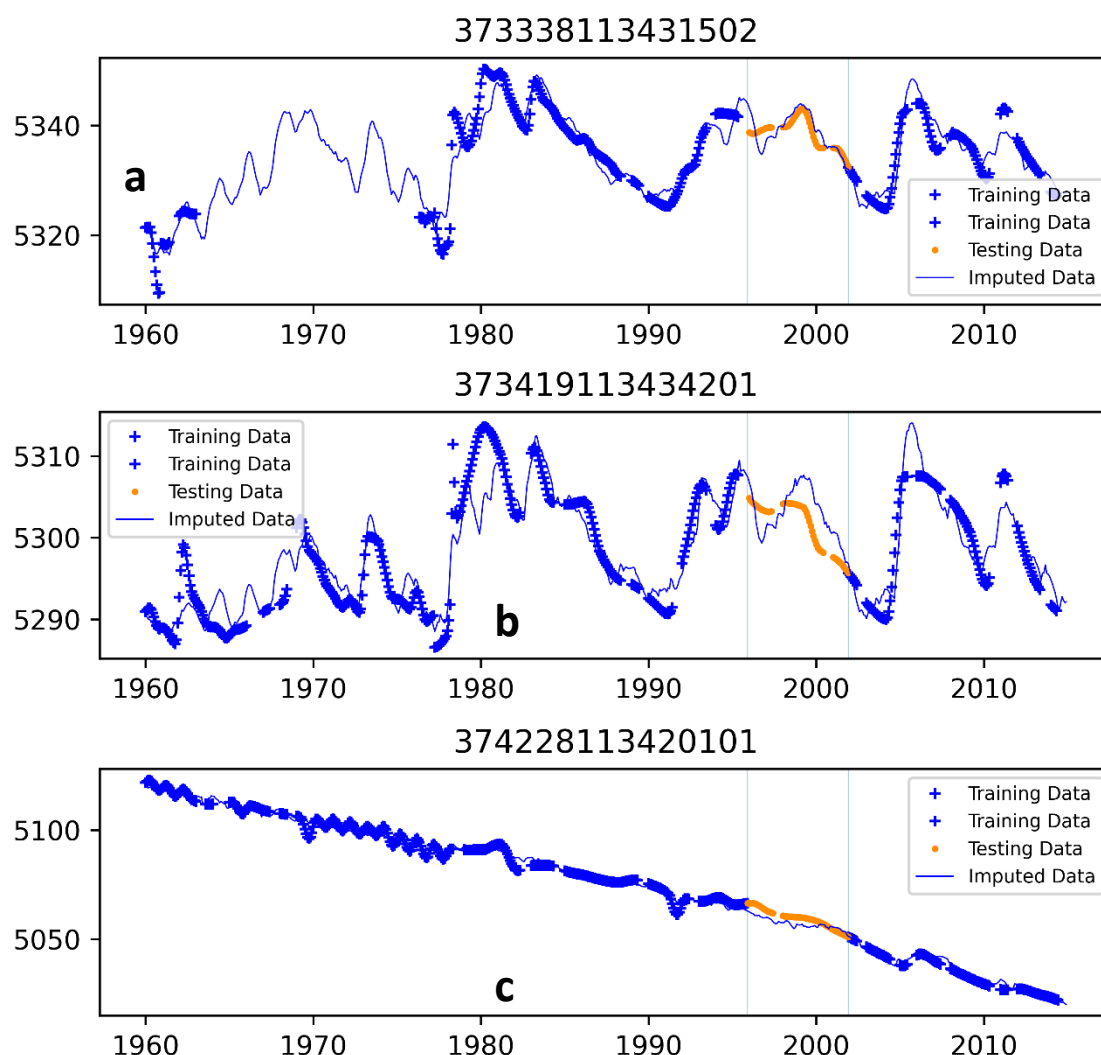
Figure 9. ELM results over a 7-year testing period from 1997–2002 for Wells 373338113431502, 373419113434201, and 373527113415101, as subfigures (**a**–**c**), respectively. Imputed data are solid blue lines, measured data used for training are blue crosses, and measured data used for testing are orange dots.

As above, we computed a value at each well location using Kriging interpolation for a single date, December 1, 2014, and compared these results to those of the ELM model. Results are shown in Table 4. In the first three wells, the ELM method significantly outperformed Kriging, while Kriging outperformed the ELM in each of the remaining wells.

The area around the last seven wells has seen significant increased pumping over this 15 year testing period for irrigation; this increased pumping is not captured by the ELM model. However, since Kriging used measured data in the other wells for this same time step, it did capture this change in behavior. For all these wells, the ELM model underestimated the drawdown, overestimating the groundwater levels and Kriging outperformed ELM significantly. Again, if the ELM model were trained on data that captured these trends, for example a gap that had data on either side, we expect it would perform better.

4. Discussion

Based on the results from the Cedar Valley and the Beryl Enterprise Aquifer, we demonstrate that global soil moisture models and PDSI computed from Earth observation data may be used to impute groundwater levels at a well with relative accuracy when aquifer processes in the training period are similar to or an extension of uses in the imputation period. The accuracy of this method decreases if well pumping or land use changes from the training time period. This decrease in accuracy occurs because the inputs do not explicitly account for this change in human activity as it relates to weather and climate conditions. Imputation is more accurate if the training period is similar to the imputation period, for example, if there are training data before and after the missing data, we expect the imputed data to be more accurate.

This method could be improved by feature engineering. Feature engineering is when specific time series data that may be correlated to aquifer behavior are incorporated into the model. This could include data such as land use changes over time, population change, electric or other power use for the wells, or any other data that might be correlated. The only requirement is that these time series data are available over both the training period, and the imputation periods (i.e., data gaps). This could include incorporating additional Earth observation data such as those from the GRACE mission, which measures water storage volumes directly, but at a very coarse spatial resolution and only over a limited recent time period.

5. Conclusions

We developed a method for imputing groundwater level data using ELM with PDSI and soil moisture models, derived from satellite Earth observations as inputs. We show that over long periods, groundwater levels and soil moisture are sufficiently correlated. We demonstrated that because of this correlation, groundwater levels can be inferred from Earth observations. This method was most accurate in areas where current groundwater uses are similar to those of the training period—there has not been significant population or land use changes. This method outperformed Kriging spatial interpolation in areas where the land use during the training period was similar, but underperformed if there had been significant changes.

References

1. Oikonomou, P.D.; Alzraiee, A.H.; Karavitis, C.A.; Waskom, R.M. A novel framework for filling data gaps in groundwater level observations. *Adv. Water Resour.* **2018**, *119*, 111–124. [CrossRef]
2. Marchant, B.; Bloomfield, J. Spatio-temporal modelling of the status of groundwater droughts. *J. Hydrol.* **2018**, *564*, 397–413. [CrossRef]
3. Mogheir, Y.; De Lima, J.; Singh, V. Assessment of informativeness of groundwater monitoring in developing regions (gaza strip case study). *Water Resour. Manag.* **2005**, *19*, 737–757. [CrossRef]
4. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*; Springer: Berlin/Heidelberg, Germany, 2017.
5. Shumway, R.H.; Stoffer, D.S. Time series analysis and its applications. *Stud. Inform. Control* **2000**, *9*, 375–376.
6. Khorasani, M.; Ehteshami, M.; Ghadimi, H.; Salari, M. Simulation and analysis of temporal changes of groundwater depth using time series modeling. *Modeling Earth Syst. Environ.* **2016**, *2*, 90. [CrossRef]

7. Mirzavand, M.; Ghazavi, R. A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods. *Water Resour. Manag.* **2015**, *29*, 1315–1328. [CrossRef]

8. Sahoo, S.; Jha, M.K. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: A comparative assessment. *Hydrogeol. J.* **2013**, *21*, 1865–1887. [CrossRef]

9. Williams, D.A.; Nelsen, B.; Berrett, C.; Williams, G.P.; Moon, T.K. A comparison of data imputation methods using bayesian compressive sensing and empirical mode decomposition for environmental temperature data. *Environ. Model. Softw.* **2018**, *102*, 172–184. [CrossRef]

10. Al-Alawi, S.M.; Abdul-Wahab, S.A.; Bakheit, C.S. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environ. Model. Softw.* **2008**, *23*, 396–403. [CrossRef]

11. Nelsen, B.; Williams, D.A.; Williams, G.P.; Berrett, C. An empirical mode-spatial model for environmental data imputation. *Hydrology* **2018**, *5*, 63. [CrossRef]

12. Cabaneros, S.M.S.; Calautit, J.K.; Hughes, B.R. A review of artificial neural network models for ambient air pollution prediction. *Environ. Model. Softw.* **2019**, *119*, 285–304. [CrossRef]

13. Bashir, F.; Wei, H.-L. Handling missing data in multivariate time series using a vector autoregressive model-imputation (var-im) algorithm. *Neurocomputing* **2018**, *276*, 23–30. [CrossRef]

14. Rouhani, S.; Wackernagel, H.J.W.R.R. Multivariate geostatistical approach to space-time data analysis. *Water Resour. Manag.* **1990**, *26*, 585–591. [CrossRef]

15. Bidwell, V.J.J.M. Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics. *Math. Comput. Simul.* **2005**, *69*, 12–20. [CrossRef]

16. Rado, O.; Al Fanah, M.; Taktek, E. Performance analysis of missing values imputation methods using machine learning techniques. In *Intelligent Computing, Proceedings of the 2019 Computing Conference, London, UK, 16–17 July 2019*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 738–750.

17. Fallah-Mehdipour, E.; Haddad, O.B.; Mariño, M.A. Prediction and simulation of monthly groundwater levels by genetic programming. *J. Hydro Environ. Res.* **2013**, *7*, 253–260. [CrossRef]

18. Yoon, H.; Jun, S.-C.; Hyun, Y.; Bae, G.-O.; Lee, K.-K. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.* **2011**, *396*, 128–138. [CrossRef]

19. Chang, F.-J.; Chang, L.-C.; Huang, C.-W.; Kao, I.F. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. *J. Hydrol.* **2016**, *541*, 965–976. [CrossRef]

20. Gong, Y.; Zhang, Y.; Lan, S.; Wang, H. A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near lake okeechobee, florida. *Water Resour. Manag.* **2016**, *30*, 375–391. [CrossRef]

21. Lettenmaier, D.P.; Alsdorf, D.; Dozier, J.; Huffman, G.J.; Pan, M.; Wood, E.F. Inroads of remote sensing into hydrologic science during the wrr era. *Water Resour. Manag.* **2015**, *51*, 7309–7342. [CrossRef]

22. Yang, Y.; Luo, Y. Evaluating the performance of remote sensing precipitation products CMORPH, PERSIANN, and TMPA, in the arid region of northwest China. *Theor. Appl. Climatol.* **2014**, *118*, 429–445. [CrossRef]

23. Chen, M.; Xie, P.; Janowiak, J.E.; Arkin, P.A. Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrol.* **2002**, *3*, 249–266. [CrossRef]

24. Wagner, W.; Scipal, K.; Pathe, C.; Gerten, D.; Lucht, W.; Rudolf, B. Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data. *J. Geophys. Res. Atmos.* **2003**, *108*. [CrossRef]

25. Boegh, E.; Soegaard, H.; Thomsen, A. Evaluating evapotranspiration rates and surface conditions using landsat tm to estimate atmospheric resistance and surface resistance. *Remote Sens. Environ.* **2002**, *79*, 329–343. [CrossRef]

26. Fan, Y.; Van den Dool, H. A global monthly land surface air temperature analysis for 1948–present. *J. Geophys. Res. Atmos.* **2008**, *113*. [CrossRef]

27. Ramsey, R.D.; Wright, D.L., Jr.; McGinty, C. Evaluating the use of landsat 30m enhanced thematic mapper to monitor vegetation cover in shrub-steppe environments. *Geocarto Int.* **2004**, *19*, 39–47. [CrossRef]

28. Sonnenschein, R.; Kuemmerle, T.; Udelhoven, T.; Stellmes, M.; Hostert, P. Differences in landsat-based trend analyses in drylands due to the choice of vegetation estimate. *Remote Sens. Environ.* **2011**, *115*, 1408–1420. [CrossRef]

29. Huang, J.; van den Dool, H.M.; Georgarakos, K.P. Analysis of model-calculated soil moisture over the united states (1931–1993) and applications to long-range temperature forecasts. *J. Clim.* **1996**, *9*, 1350–1362. [CrossRef]

30. Rhee, J.; Im, J.; Carbone, G.J. Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sens. Environ.* **2010**, *114*, 2875–2887. [CrossRef]

31. Rodell, M.; Houser, P.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.-J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394. [CrossRef]

32. Rogan, J.; Chen, D. Remote sensing technology for mapping and monitoring land-cover and land-use change. *Prog. Plan.* **2004**, *61*, 301–325. [CrossRef]

33. Jackson, T.J. Remote sensing of soil moisture: Implications for groundwater recharge. *Hydrogeol. J.* **2002**, *10*, 40–51. [CrossRef]

34. Huang, G.-B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. ManCybern. Part B (Cybern.)* **2011**, *42*, 513–529. [CrossRef] [PubMed]

35. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]

36. Huang, G.-B. What are extreme learning machines? Filling the gap between frank rosenblatt's dream and john von neumann's puzzle. *Cogn. Comput.* **2015**, *7*, 263–278. [CrossRef]

37. Zhu, W.; Miao, J.; Qing, L. Constrained extreme learning machine: A novel highly discriminative random feedforward neural network. In Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, China, 6–11 July 2014; pp. 800–807.

38. Beran, B.; Piasecki, M. Availability and coverage of hydrologic data in the us geological survey national water information system (nwis) and us environmental protection agency storage and retrieval system (storet). *Earth Sci. Inform.* **2008**, *1*, 119–129. [CrossRef]

39. Dai, A.; National Center for Atmospheric Research Staff (Eds.) The Climate Data Guide: Palmer Drought Severity Index (pdsi). Available online: https://climatedataguide.ucar.edu/climate-data/palmer-drought-severity-index-pdsi (accessed on 9 May 2019).

40. Qiao, X.; Li, Z.; Ames, D.P.; Nelson, E.J.; Swain, N.R. Simplifying the deployment of ogc web processing services (wps) for environmental modelling–introducing tethys wps server. *Environ. Model. Softw.* **2019**, *115*, 38–50. [CrossRef]

41. Swain, N.R.; Christensen, S.D.; Jones, N.; Nelson, E. Tethys: A platform for water resources modeling and decision support apps. In AGU Fall Meeting Abstracts, Proceedings of the AGU Fall Meeting, San Francisco, CA, USA, 15–19 December 2014.

42. Evans, S.W. Groundwater Level Mapping Tool: Development of a Web Application to Effectively Characterize Groundwater Resources. Master's Thesis, Brigham Young University, Provo, UT, USA, April 2019.

43. Roberts, W.; Williams, G.P.; Jackson, E.; Nelson, E.J.; Ames, D.P. Hydrostats: A python package for characterizing errors between observed and predicted time series. *Hydrology* **2018**, *5*, 66. [CrossRef]

44. Jackson, E.K.; Roberts, W.; Nelsen, B.; Williams, G.P.; Nelson, E.J.; Ames, D.P. Introductory overview: Error metrics for hydrologic modelling–a review of common practices and an open source library to facilitate use and adoption. *Environ. Model. Softw.* **2019**, *119*, 32–48. [CrossRef]

45. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. Scipy 1.0: Fundamental algorithms for scientific computing in python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]

46. Walt, S.v.d.; Colbert, S.C.; Varoquaux, G. The numpy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30. [CrossRef]

47. Swain, N.R.; Christensen, S.D.; Snow, A.D.; Dolder, H.; Espinoza-Dávalos, G.; Goharian, E.; Jones, N.L.; Nelson, E.J.; Ames, D.P.; Burian, S.J. A new open source platform for lowering the barrier for environmental web app development. *Environ. Model. Softw.* **2016**, *85*, 11–26. [CrossRef]

48. Fritsch, F.N.; Carlson, R.E. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal.* **1980**, *17*, 238–246. [CrossRef]

49. McKinney, W. Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; pp. 51–56.

50. Seabold, S.; Perktold, J. Statsmodels: Econometric and statistical modeling with python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; p. 61.

51. Lu, S.; Wang, X.; Zhang, G.; Zhou, X. Effective algorithms of the moore-penrose inverse matrices for extreme learning machine. *Intell. Data Anal.* **2015**, *19*, 743–760. [CrossRef]

52. Quenouille, M.H. Approximate tests of correlation in time-series 3. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1949; pp. 483–484.