Groundwater Level Mapping Tool: An Open Source Web Application to Quantify Groundwater Resources

Steven W. Evansa, Norm L. Jonesa,\*, Daniel P. Amesa, Gustavious P. Williamsa

aCivil and Environmental Engineering, Brigham Young University, Provo, UT, USA

\*Corresponding Author. Email address: [njones@byu.edu](mailto:njones@byu.edu) N. Jones

Keywords: groundwater, aquifer storage, data imputation, time series, water resources management

Abstract

An accurate understanding of groundwater storage trends is essential for decision makers to effectively manage groundwater resources throughout an aquifer. Groundwater is expensive and difficult to monitor, and data collected from monitoring wells is often sporadic, available only at irregular of infrequent intervals. We present the design and development of an open source web application to visualize groundwater data and automatically calculate changes in aquifer storage volume. This application uses multi-linear regression (MLR) to infer missing data for poorly sampled wells, using correlated data from other wells in the same aquifer. The app then uses this MLR inferred data to spatially interpolate water levels throughout an aquifer at specified time steps using GSLIB Kriging code. Use of the inferred data increased the accuracy of the spatial interpolation over standard methods, and results of the automated aquifer storage calculation were comparable to those of several detailed USGS studies.

# Introduction

Groundwater is depended on worldwide as a major source for agricultural irrigation, industrial processes, mining, and drinking water. The USGS reports that 30.1% of the Earth’s fresh water consists of groundwater, while 1.2% consists of surface water in lakes, rivers, and streams (Gleick, 1993). Although fresh groundwater is so abundant, its responsible and sustainable use poses a significant challenge.

One of the great challenges of groundwater sustainability is the ability to accurately characterize the state of an aquifer, so as to implement practices, procedures, and regulations to promote its sustainable use. Although fresh groundwater is often abundant, and prevalently used in some areas, it is difficult and expensive to accurately quantify groundwater, compared to surface water resources. The state of surface water resources is readily visible to the naked eye, can be measured easily, and is straightforward to quantify. This is not the case for groundwater, which generally requires drilling a series of monitoring wells in order to measure the location of the phreatic surface. While surface storage in a waterbody reaches approximately the same elevation throughout the body, groundwater surface elevations may vary significantly, by hundreds of feet in some cases, throughout an aquifer, depending on the overlying land use, pumping of irrigation wells, aquifer recharge, and other factors. Groundwater levels are heavily influenced by climatic, geographic, lithological, and human factors. For these reasons, it is difficult to quantify and map aquifer water levels and storage volumes.

Even when data on water surface elevations is available from monitoring wells drilled into an aquifer, this data is generally not harnessed to its full potential to aid in decision-making. This data is available at point locations scattered in time and space throughout an aquifer, and it is difficult to piece these segments of data together into a complete picture of aquifer-wide behavior (Marchant & Bloomfield, 2018). Point data from monitoring wells are typically sparse and give only a limited sampling of the spatial distribution of water levels in the aquifer, and the data observations from these monitoring wells are often temporally sporadic, including large gaps in the time series data (Oikonomou, Alzraiee, Karavitis, & Waskom, 2018).

Because of these issues, quantifying groundwater resources requires spatial and/or temporal interpolation and/or extrapolation to at least some extent. One of the most widely used computer programs used for spatial interpolation in this and other fields is the Geostatistical Software Library (GSLIB), developed at Stanford University (Deutsch & Journel, 1992). This program performs spatial interpolation using the Kriging technique pioneered by the South African statistician and mining engineer, Danie G. Krige. This Kriging interpolation technique is used by many researchers in groundwater when attempting to interpolate spatial data. Kumar (2006) used Kriging interpolation to estimate unknown depths to water table in an aquifer in Rajasthan, India and found that kriged groundwater levels satisfactorily matched the observed groundwater levels. Other researchers demonstrated that the accuracy of groundwater surface elevation maps could be improved in some cases by introducing topography to the interpolation using Kriging with an external drift (Boezio, Costa, & Koppe, 2006). In most studies on spatial interpolation of groundwater, data are first lumped into temporal bins, and then interpolated spatially, assuming data within the bin is all from a specific time step (Ruybal, Hogue, & McCray, 2019).

Several techniques have also been developed and used in an effort to improve the temporal interpolation and estimation of well time series observations. Rouhani and Wackernagel (1990) used Kriging to perform temporal interpolation of depth to water table time series measurements in a basin south of Paris, France. Bidwell (2005) forecasted groundwater levels one month ahead in Canterbury, New Zealand using an ARMAX model based on the eigenstructure of aquifer dynamics. Others have used classical time series models including 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), and multiple linear regression to predict groundwater levels (Khorasani, Ehteshami, Ghadimi, & Salari, 2016; Mirzavand & Ghazavi, 2015; Sahoo & Jha, 2013). Oikonomou et al. (2018) employed an exogenous seasonal autoregressive integrated moving average (SARIMAX) stochastic model to describe the simulated groundwater level fluctuation process of a regional physical groundwater model and the Ensemble Smoother (ES) for predicting groundwater levels.

Lately, researchers have sought for increased accuracy in quantifying groundwater by using spatiotemporal interpolation. Ruybal et al. (2019) used the Arapahoe aquifer as a case study to demonstrate the benefits of spatiotemporal kriging over spatial kriging across as sparsely gauged and irregularly sampled aquifer. They found that spatiotemporal kriging allowed estimation of groundwater levels during times when data are not available, and avoided biases and anomalies caused by kriging with different data available in different time periods. Ahmadi and Sedghamiz (2007) conducted spatial and temporal analysis of an Iranian aquifer, and concluded that “spatial structure was a little bit stronger than temporal structure.” Combining spatial and temporal interpolation yields more accurate estimates by leveraging both temporal and spatial relationships between observations (Ruybal et al., 2019).

We have developed an open source Python-based web application to allow visualization and quantification of groundwater resources, leveraging both temporal and spatial interpolation. This web application, the Groundwater Level Mapping Tool, is built on Tethys Platform, an open source platform for lowering the barrier for environmental web app development (Swain et al., 2016). This application is generalized to allow its use world-wide, and allows decision makers to accomplish the following:

1. View time series and other data for each well within an aquifer.
2. View maps and animations of aquifer-wide groundwater levels at different time periods.
3. Calculate and view estimates of total aquifer storage change.

This paper details the temporal and spatial interpolation methods used by the application to map aquifer drawdown, and calculate aquifer storage change. Temporal interpolation is accomplished using multi-linear regression harnessing correlated observation wells within an aquifer. We present the results of this interpolation and compare its accuracy to other spatial and temporal interpolation methods. We also compare the results of our storage change calculated by the app to those of detailed USGS studies.

# Methods

## Multi-Linear Regression Harnessing Correlated Wells

To overcome the challenges of data gaps and to accurately extend data beyond its sampled range, we developed a new method of temporal interpolation, Multi-Linear Regression Harnessing Correlated Wells. This method was developed based on the assumption that wells within the same aquifer will likely exhibit similar characteristics in depth to water table fluctuations and trends. This similarity between wells located in the same aquifer is demonstrated in Figure 2‑1. The three wells shown are all located in the Cedar Valley Aquifer in southern Utah. Well 37342113100801 (green) is located near the southern end of the aquifer, while the two other wells are located near the center. There is a clear correlation between these wells; the water levels rise and fall following the same general pattern throughout the aquifer.

Figure 2‑1: Correlation between Wells in the Cedar Valley Aquifer

The Multi-Linear Regression method of time series interpolation utilizes this inner-aquifer correlation to yield accurate estimates of water surface elevation for a well at specific times outside the sampling range of the well. This method is useful when the times series data for a well does not thoroughly cover the time(s) of interest. Figure 2‑2 shows the time series data for Well 374210113044801 located in the Cedar Valley Aquifer, which has data from 1998 to 2008, which does not span the period of interest from 1985 to 2015. Using Multi-Linear Regression, this wells’ time series data can be extended to span the period of interest.

Period of Interest

Figure 2‑2: Example of Well Time Series that does not Cover Period of Interest

The first step in this method is to construct a single data frame, using the Pandas Python library, containing the time series information for every well in the aquifer (McKinney, 2010). This data frame includes an index of regular 3-month intervals with the time series observations of each well resampled to these regular 3-month intervals by means of a PCHIP interpolation. If a well does not have data within a 3-month interval, then the data frame contains a null value for that well at that time step. Next, those wells which contain data spanning the time period of interest are identified as “reference wells.” After these reference wells, which contain full time series data for the period of interest, are identified, the Pearson Correlation Coefficient is calculated between the well with missing time series data and each of the reference wells. This correlation coefficient between the target well and each reference well for data points is calculated by Equation (2‑1).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑1) |

Data from the reference wells with the five highest correlation values are used as inputs for the process of multi-linear regression. Figure 2‑3 shows the time series data for the target well with missing time series data along with the 5 most correlated wells in the aquifer. The area within the red rectangle shows the data that will be used to train the multi-linear regression model to predict the target well’s missing time series values.

Figure 2‑3: Well with Missing Data Plotted with Correlated Reference Wells

To effectively estimate the missing values of the target well, it is necessary to normalize the data. Notice from Figure 2‑3 the similar shape, but differing depths of each of the wells. This normalization transforms the data so that all the wells are in the same range, which allows prediction based on the similar shapes of the data rather than the differing magnitudes of the data. Data from the target and reference wells are normalized such that all values within the training data set are between 0 and 1. Let Y represent the target well, and R1, R2,R3,R4, and R5 represent the five reference wells. The training data subsets will be represented as Ytrain for the target well, and Rtrain1, Rtrain2, …, Rtrain5 for the five reference wells. The values of each well are normalized as shown in Equation (2‑2).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑2) |

These normalized data sets are plotted in Figure 2‑4, with the normalized training set again bounded in red. The reference wells are normalized using the minimum and maximum values of the training subset, rather than the entire data set, because this provides a more skilled prediction, especially when predicting time series values outside the scope of the training data.

Figure 2‑4: Normalized Target Well with Normalized Correlated Reference Wells

The normalized time series data of the target well is assumed to be a linear combination of the normalized time series data of the five reference wells. At each time (t), the normalized value of depth to groundwater at the target well (Yt) is approximated as a linear combination of the normalized depth to water table at time (t) of the reference wells (Rt). The equation for Yt is shown in Equation (2‑3), where is an error residual term.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | (2‑3) |  |

Equation (2‑3) can be rewritten in matrix form as Equation (2‑4).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑4) |

The term represents weights determined by regularized least-squares fit using the training data subset, such that the sum of the squared residuals is minimized. These weights are obtained by solving Equation (2‑5), where is a regularization term, and is the identity matrix. This process is carried out in the app by using the regularized least squares solver in the Statsmodels Python library (Seabold & Perktold, 2010).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑5) |

When using a high complexity estimator, such as multi-linear regression, the bias error is generally small, and the variance very large. Bias error is an error that causes the estimator algorithm to miss relevant relations between features and target outputs, thereby under fitting the data. Variance is an error from sensitivity to small fluctuations in the training set, which can cause the model to train to random noise in the training data, rather than actual correlations, thereby overfitting the data. The introduction of increases the bias of the estimate, but significantly decreases the variance, preventing overfitting and generally yielding a more accurate estimate. The method of regularization used here is Tikhonov or Ridge regularization (Tikhonov & Arsenin, 1977).

Once has been obtained by means of the regularized least squares model using the training data subset, the unknown Y terms may be estimated for each time step t by solving Equation (2‑3). Once the Y terms have been obtained, they must be unscaled to their original extent, yielding Yestimate. This is accomplished by applying Equation (2‑6).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑6) |

The process is now complete for estimating missing time series values for the target well. The estimated time series modelled using Multi-Linear Regression Harnessing Correlated Wells is shown in Figure 2‑5, together with the original recorded data.

Figure 2‑5: Well Time Series Modelled by Multi-Linear Regression

Once this process has been carried out for each well in the aquifer, we now have a data set with time series values of depth to groundwater at 3 month intervals, spanning our entire period of interest. The groundwater surface can now be estimated throughout the aquifer at any 3-month time step within the period of interest.

## Kriging Interpolation

With the extended time series information from the MLR process, we have groundwater surface data at specific points scattered throughout an aquifer available for each 3-month time step within our period of interest. We can now use these point data to map groundwater levels throughout the entire aquifer at any or all of the three-month time steps using Kriging spatial interpolation.

For each time step to be mapped, we develop a variogram for use in the Kriging interpolation. A model variogram is developed by first creating and plotting the experimental variogram from the available point data. The experimental variogram is a function of the Euclidean distance , and the semivariance between each pair of points in the dataset. The Euclidean distance between a pair of points with coordinates and is calculated for each pair of points in the dataset using Equation (2‑18).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑18) |

The semivariance between a pair of points with data values and is calculated for each pair of points in the dataset using Equation (2‑19).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑19) |

An ordered pair of distance and semivariance is created for each pair of points in the dataset. The ordered pairs of distance and semi-variance are sorted into ten bins of equal intervals based on distance . The range of these bins extends from the minimum distance between two observation points in the aquifer to the sum of the minimum distance and the half the maximum distance between observation points in the aquifer. This approach of eliminating extremely distant points from the bins produces a more accurate estimation because the smallest distances are the most important for developing a proper variogram for accurate estimation (Kitanidis, 1997). The use of the distant observations skews the automatic variogram fitting routine towards the large variances of the distant observations, which negatively impacts the estimation.

Once the ordered pairs are sorted into the ten bins spanning equal intervals, the mean distance and semivariance are computed within each bin. These ordered pairs compose the experimental variogram, and are used to fit a spherical variogram model, which was found to produce the best results by several researchers, as well as from initial inspection of preliminary test case results (Gundogdu & Guney, 2007; Nikroo, Kompani-Zare, Sepaskhah, & Shamsi, 2010). The spherical variogram model is a function of the distance of an observation point from the interpolation point and is given by Equation (2‑20), defined by the Nugget (), Range (), and Sill (), where the Partial Sill () is the difference between the sill and the nugget.

|  |  |  |
| --- | --- | --- |
| = |  | (2‑20) |

The optimal parameters and for the semivariance function (Equation (2‑20)), are determined using least squares optimization with a nugget value of 1. The model is fit by minimizing the residuals of the spherical variogram model compared to the ten ordered pairs of the experimental variogram. In this optimization, the residuals are weighted using a logistic function so that weights vary from a value approaching one at distance zero to a value approaching zero as the distance increases. This weighting allows the least squares optimization to fit the data better at closer distances, producing a better variogram for estimation (Kitanidis, 1997). Figure 2‑6 shows an example of a spherical variogram model fitted to an experimental variogram using this method.

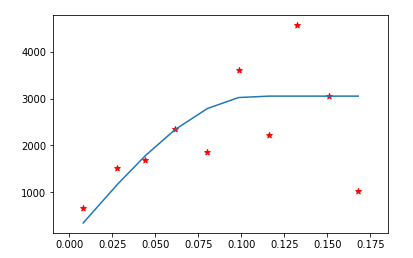


Figure 2‑6: Semivariance Function Fitted to Experimental Variogram

With the variogram parameters defined, the web application carries out the interpolation throughout the aquifer at a user-defined resolution by employing a Python-wrapping of the GSLIB Fortran Code (Deutsch & Journel, 1992). In cases where the data yields a singular matrix, the app is written to instead revert to an IDW interpolation for these singular grid points.

## Calculation of Aquifer Storage

With the results of the temporal and then spatial interpolation, it is possible to calculate changes in total aquifer storage volume based on an estimate of aquifer porosity. This is accomplished by performing mathematic operations on series of raster datasets of groundwater levels at specific times produced during the spatial interpolation phase. The first dataset (corresponding to the earliest time step in the series) is known as and serves as the baseline from which changes in aquifer storage are measured.

These changes are calculated by first calculating the drawdown from the bases case for each time step in the raster series. The drawdown is calculated on a cell-by-cell basis by applying Equation (2‑10) for each of the timesteps, resulting in a new set of raster datasets of drawdown.

|  |  |  |
| --- | --- | --- |
|  |  | (2‑10) |

Aquifer-wide storage changes are then calculated fore each time step by multiplying the drawdown at each grid cell by the average aquifer storage coefficient and the grid cell area , and summing over all grid cells in the aquifer, as shown in Equation (2‑11). The aquifer storage coefficient will be either the specific storage (for a confined aquifer) or the specific yield (for an unconfined aquifer), which is close to, but typically a little smaller than the porosity. These values may vary throughout an aquifer, and an aquifer may even be partially confined, partially unconfined. This tool is meant to produce a rough estimate of storage volume and so employs a rough average of the storage coefficient over the aquifer. To determine a more accurate estimate where storage coefficients are known in detail, the aquifer could be split into several sections based on storage coefficients, and the storage volume calculated separately and then summed together over the sections.

|  |  |  |
| --- | --- | --- |
|  |  | (2‑11) |

The area is not constant for each grid cell over the dataset, since the grids are defined at a specified latitude and longitude resolution. Each cell has constant height, but the cell width is dependent on the cell latitude. Cells closer to the equator will have larger widths than those nearer the poles. The area of each grid cell in the aquifer is calculated based on the resolution of the grid , the mean radius of the Earth , and the latitude of the center of each grid cell as shown in Equation (2‑12).

|  |  |  |
| --- | --- | --- |
|  |  | (2‑12) |

When the Groundwater Level Mapping Tool is in metric mode, the mean radius of the Earth is 6,371,000 meters (Moritz, 1980). The aquifer storage is calculated in cubic meters, since the depth to groundwater measurements and therefore drawdown is also in meters. In imperial units, the Groundwater Level Mapping Tool reports changes in aquifer storage volume in Acre-ft. The app uses a mean Radius of the Earth of 3,959 miles. In this case, the result of Equation (2‑11) is in square mile-ft, which is converted to acre-ft by multiplying by 640. An example of the output of this storage volume calculation procedure for the Cedar Valley, UT Aquifer is shown in Figure 2‑7.

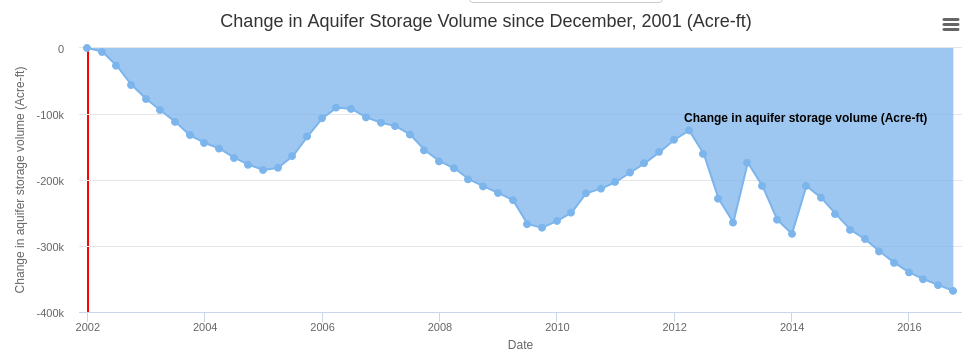


Figure 2‑7: Example Output of Aquifer Storage Volume

# Results and Discussion

## Testing of Multi-Linear Regression in Cedar Valley, UT

The method of time series extension by Multi-Linear Regression Harnessing Correlated Wells (MLR) detailed in Section 2.1 was tested using a set of ten wells from the Cedar Valley Aquifer each containing data from 1980 – 2015. The locations of these 10 wells are shown in Figure 3‑1. The Cedar Valley Aquifer was chosen as a test case because it contains several wells with a sufficiently long and detailed period of groundwater measurements and because the wells in this aquifer exhibit trends and patterns that pose significant difficulties to time series analysis. The time series in this aquifer are non-stationary, and neither increase or decrease at constant rates. The measurements taken from 1980-1995 were used to train the dataset, then the model predictions were compared to the actual data for measurements taken after 1995, thus testing the model. Both the training and testing datasets include significant peaks and valleys in the data, and cannot be predicted easily using classical time series analysis methods. The testing dataset also includes values well outside the scope of the training dataset. The MLR method of time series extension was used to make predictions at each well, and then compared against a naïve prediction, where the groundwater was assumed to remain constant after 1995, and a linear least squares prediction. Results typical of these model predictions are shown in Figure 3‑2 through Figure 3‑4, with the training data (green), the measured data (solid red), the MLR prediction (dashed red), the naïve prediction (dashed blue), and the least squares prediction (dashed yellow) for each well.

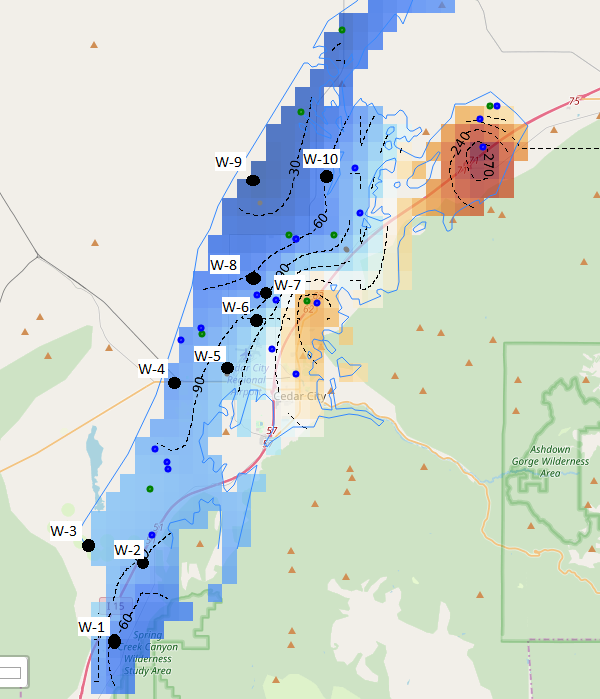


Figure 3‑1: Location of 10 Testing Wells for MLR in Cedar Valley Aquifer

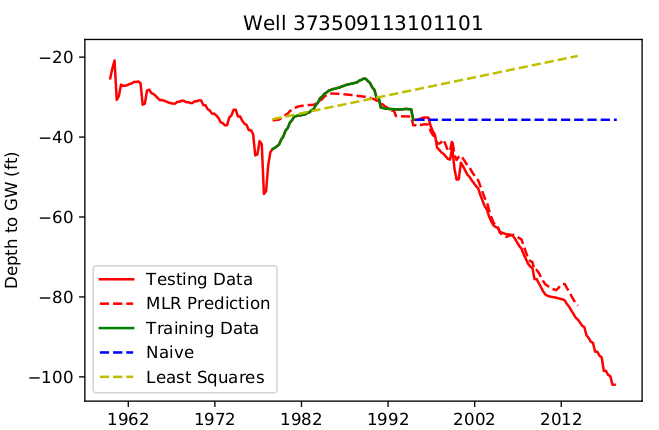


Figure 3‑2: Time Series Model for Well 373509113101101

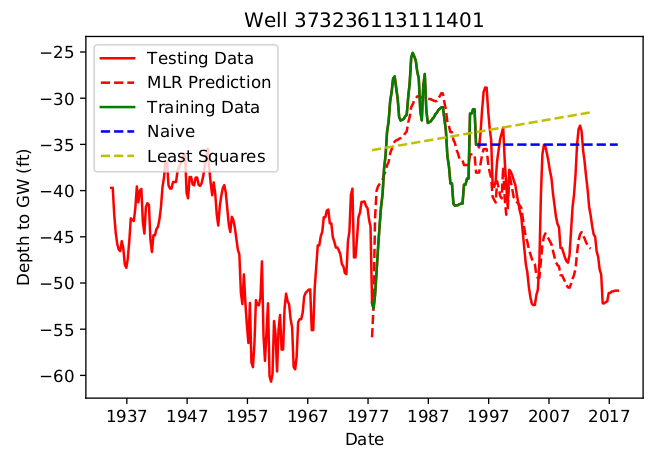


Figure 3‑3: Time Series Model for Well 373236113111401

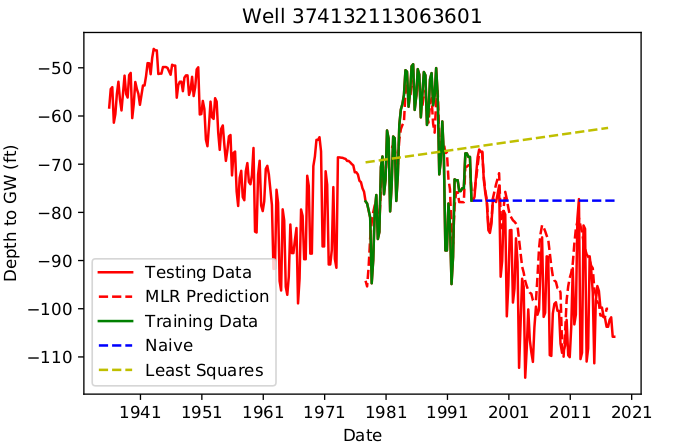


Figure 3‑4: Time Series Model for Well 374132113063601

As demonstrated by Figure 3‑2 through Figure 3‑4, the MLR method outperformed the naïve and least squares estimation methods for each well (with the exception of Well 374744113055001). Table 3‑1 shows the RMSE value for each of the ten wells in the Cedar Valley study area for the MLR, naïve, and least squares prediction methods. The MLR method decreased the RMSE value by an average of 57% from the naïve method, and 68% from the least squares method.

Table 3‑1: RMSE Values for Time Series Prediction for Ten Wells in Cedar Valley Aquifer

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **RMSE Value for Various Methods** | | |
| **Map ID** | **Well ID** | **MLR** | **Naïve** | **Least Squares** |
| W-1 | Well 373236113111401 | 5.67 | 9.63 | 9.89 |
| W-2 | Well 373509113101101 | 2.33 | 36.64 | 41.00 |
| W-3 | Well 373542113122401 | 5.29 | 34.58 | 23.45 |
| W-4 | Well 374105113085001 | 24.57 | 26.29 | 37.09 |
| W-5 | Well 374132113063601 | 10.33 | 21.16 | 32.59 |
| W-6 | Well 374304113052901 | 10.06 | 19.37 | 29.38 |
| W-7 | Well 374423113053301 | 11.33 | 17.43 | 45.88 |
| W-8 | Well 374423113053401 | 4.98 | 15.79 | 29.46 |
| W-9 | Well 374744113055001 | 4.14 | 2.73 | 9.20 |
| W-10 | Well 374745113022901 | 11.97 | 27.06 | 22.15 |

The method of time series extension using Multi-Linear Regression Harnessing Correlated Wells was also tested for accuracy against Kriging spatial interpolation, using a jackknife approach. This testing was carried out for ten wells in the Cedar Valley Aquifer. The depth to groundwater on December 31, 2014 was estimated at each testing well by implementing the PCHIP and then Kriging interpolation, omitting the measured depth at the testing well from the interpolation. The depth to groundwater on December 31, 2014 was then estimated at each testing well by implementing the MLR technique, using data from 1985-1995 as the training set, and then estimating twenty years of groundwater depths from 1995-2015. The results of both of these estimates, the actual measurement, and the % error for each estimate are displayed in Table 3‑2 for each of the ten tested wells.

Table 3‑2: Error Values for MLR Estimate Compared to Kriging Estimate for Ten Wells in Cedar Valley Aquifer

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **Measured Value** | **MLR Estimate** | | **Kriging Estimate** | |
| **Map ID** | **Well ID** | Depth to GW (ft) | Depth to GW (ft) | (% error) | Depth to GW (ft) | (% error) |
| W-1 | Well 373236113111401 | -47.6 | -49.9 | -(4.8%) | -66.8 | -(40.3%) |
| W-2 | Well 373509113101101 | -90.6 | -92.1 | -(1.7%) | -89.3 | (1.4%) |
| W-3 | Well 373542113122401 | -103.7 | -101.4 | (2.2%) | -88.6 | (14.6%) |
| W-4 | Well 374105113085001 | -74 | -79.1 | -(6.9%) | -88.6 | -(19.7%) |
| W-5 | Well 374132113063601 | -93.5 | -96.2 | -(2.9%) | -100.8 | -(7.8%) |
| W-6 | Well 374304113052901 | -109.7 | -113.8 | -(3.7%) | -77.4 | (29.4%) |
| W-7 | Well 374423113053301 | -75.6 | -60 | (20.6%) | -80.8 | -(6.9%) |
| W-8 | Well 374423113053401 | -67 | -67.2 | -(0.3%) | -61.8 | (7.8%) |
| W-9 | Well 374744113055001 | -18.6 | -20.2 | -(8.6%) | -30.3 | -(62.9%) |
| W-10 | Well 374745113022901 | -68.9 | -69.4 | -(0.7%) | -58.3 | (15.4%) |

The MLR produced relatively accurate results, with seven of the ten tested wells exhibiting less than 5% error, and nine of the wells exhibiting less than 10% error. Errors are significantly smaller for the MLR estimate than the Kriging estimate with the exception of W-2, where the error is practically the same, and W-7, where the MLR error is greater than Kriging. Kriging produces a better estimate at this well for two reasons: W-7 is quite close to W-6 and W-8 (see Figure 3‑5), which decreases the variance of the Kriging interpolation; and the time series for W-7 contains only four points in its training dataset from 1985-1995 as shown in Figure 3‑16, which decreases the accuracy of the MLR method. With these factors, it is unsurprising that Kriging outperformed MLR in this instance. These results demonstrate that in most cases where data is available from a different time period than desired, it is more accurate to interpolate temporally using MLR referencing other wells than to interpolate spatially from other wells. This is significant, since researchers using different temporal interpolation methods have previously concluded that “spatial structure was a little bit stronger than temporal structure”(Ahmadi & Sedghamiz, 2007).

## Testing of Multi-Linear Regression in Ogallala Aquifer, Texas

The method of time series extension by Multi-Linear Regression Harnessing Correlated Wells (MLR) detailed in Section 2.2 was tested further using 467 wells located in the Ogallala Aquifer in the Texas Panhandle, each containing data from 1960 – 2010. This area was chosen for further testing because of the large amount of data available for hundreds of wells in the aquifer. The time series observations from each of these wells were divided into a training set from 1960-1995, and a testing set from 1995-2010. The method of MLR was used to predict the values of depth to water table at each well from 1995-2010 and then compared against the actual values of the testing dataset. Some results of this prediction are shown in Figure 3‑5, Figure 3‑6, and Figure 3‑7, with the training data (green), the measured data (solid red), the MLR prediction (dashed red), the naïve prediction (dashed blue), and the least squares prediction (dashed yellow) for each well.

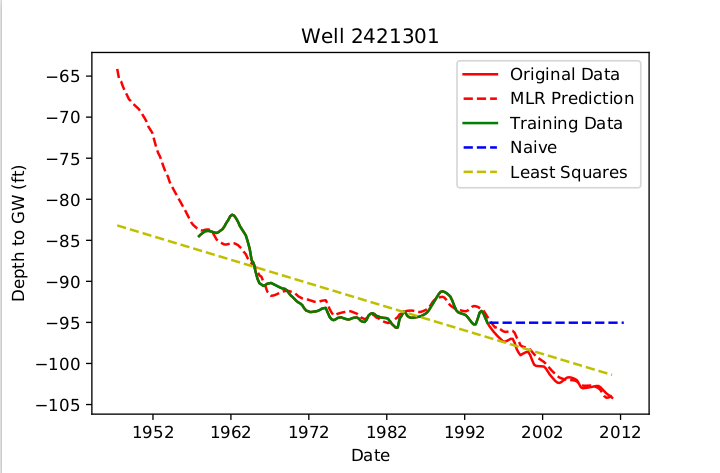


Figure 3‑5: Time Series Model for Well 2421301

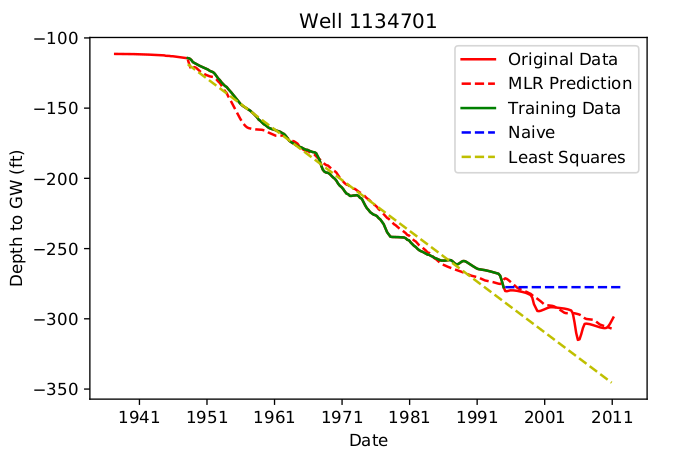


Figure 3‑6: Time Series Model for Well 1134701

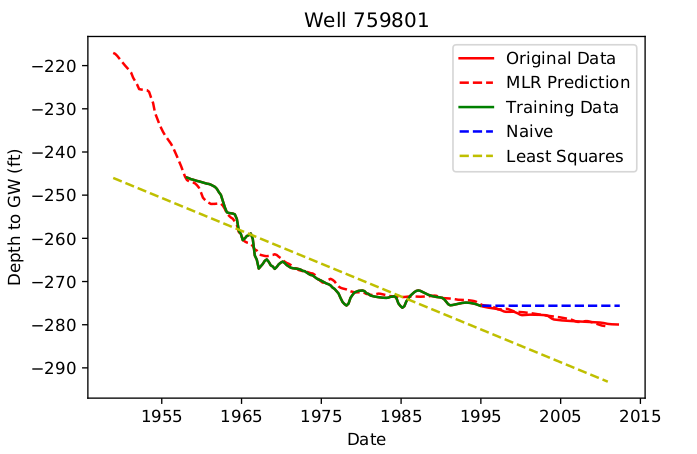


Figure 3‑7: Time Series Model for Well 759801

As shown in the preceding figures, the MLR method generally outperformed the naïve and least squares methods. As shown in Figure 3‑5, this MLR method was able to capture a period of decrease in water table elevation following a period of increase, which could not be accomplished with the naïve or least squares method. As shown in Figure 3‑6, the MLR model was able to correctly predict a variation from a basically constant decrease in water levels from 1950-1990, while the least-squares method simply continued along the same linear trend line. Figure 3‑7 also demonstrates the MLR method’s ability to correctly model changes in aquifer depletion rate.

The results of the time series models were compared using the Range Normalized RMSE method. Table 3‑3 shows the mean and median normalized RMSE values for the 467 tested wells.

Table 3‑3: NRMSE Values for Time Series Models  
in the Ogallala Aquifer

|  |  |  |  |
| --- | --- | --- | --- |
|  | **MLR Model** | **Naïve Prediction** | **Least Squares Prediction** |
| **Mean** | 0.108 | 0.181 | 0.250 |
| **Median** | 0.076 | 0.145 | 0.211 |

A box and whisker plot of the NRMSE values for the 467 wells is shown in Figure 3‑8, with the MLR Model in blue, the Naïve Prediction in orange, and the Least Squares Prediction shown in grey. Overall, the MLR Model exhibited the best results.

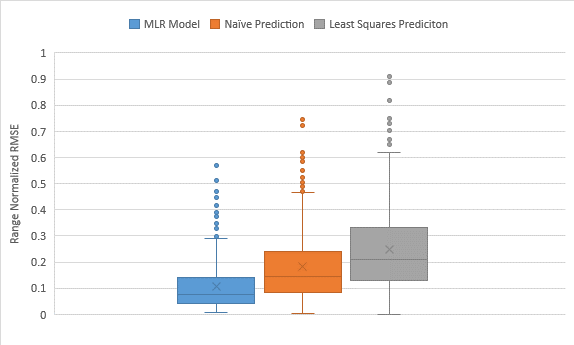


Figure 3‑8: Box and Whisker Plot of NRMSE Values in the Ogallala Aquifer

The MLR method outperformed both the naïve and least squares method in 314 of the 467 tested wells (67%). In those cases where the naïve prediction outperformed the MLR method, it was generally by a small margin, as shown in Figure 3‑9. In this case, the NRMSE value of the MLR model was 0.033, while that of the Naïve prediction was 0.026, both of which are acceptably small errors.

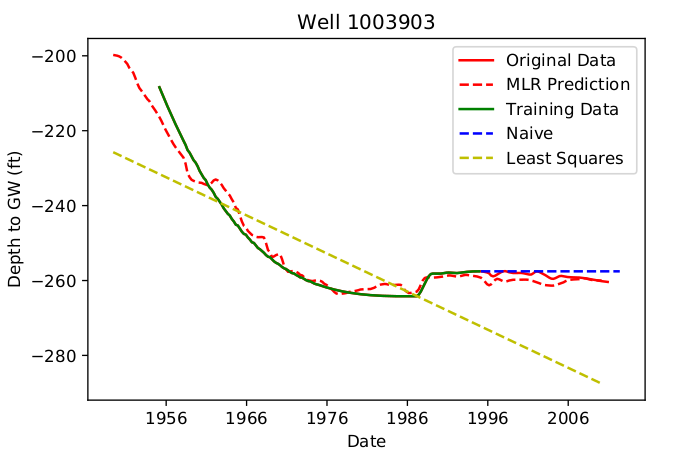


Figure 3‑9: Time Series Model for Well 1003903

The method of time series extension using Multi-Linear Regression Harnessing Correlated Wells was also tested for accuracy against Kriging spatial interpolation, using a jackknife approach. This testing was carried out for 407 wells in the Ogallala Aquifer in the Texas panhandle. These wells were selected because they each contained time series data from 1960 to 2010, which enabled comparison of the estimates to actual measured values. The depth to groundwater on December 31, 2009 was estimated at each testing well by implementing the PCHIP and then Kriging interpolation, omitting the measured depth at the testing well from the interpolation. The depth to groundwater on December 31, 2009 was then estimated at each testing well by implementing the MLR technique, using data from 1960-1995 as the training set, and then estimating fifteen years of groundwater depths from 1995-2010. The estimated depth to water table obtained from both of these methods were then compared to the measured value for the testing well, and the percent absolute error was measured. Figure 3‑10 shows a box and whisker plot of the percent absolute error for both the MLR (blue) and Kriging (orange) methods for these 407 testing wells.

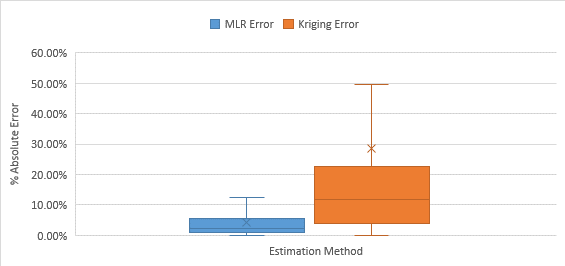


Figure 3‑10: Box and Whisker Plot of Percent Estimation Error in Ogallala Aquifer

Table 3‑4 shows the mean and median absolute percent error, and Table 3‑5 shows the mean and median absolute error (ft) for the Kriging and MLR estimates for the 407 wells in the Ogallala Aquifer. The MLR method outperformed the Kriging method significantly. Out of the 407 tested wells, 340 estimates (80%) yielded lower error using the MLR method, while 67 (20%) yielded lower error using Kriging. Over 80% of the MLR estimated depths were within ten feet of the actual measurements, which is quite accurate, considering the average depth to groundwater in this area is approximately 200 feet.

Table 3‑4: Mean and Median Absolute Percent Error for 417 Wells in the Ogallala Aquifer

|  |  |  |
| --- | --- | --- |
| **Test Statistic** | **MLR Method** | **Kriging Method** |
| **Mean Absolute Percent Error** | 4.35% | 29.35% |
| **Median Absolute Percent Error** | 2.11% | 12.82% |

Table 3‑5: Mean and Median Absolute Error for 417 Wells in the Ogallala Aquifer

|  |  |  |
| --- | --- | --- |
| **Test Statistic** | **MLR Method** | **Kriging Method** |
| **Mean Absolute Error** | 6.2 ft | 26.4 ft |
| **Median Absolute Error** | 3.5 ft | 19.1 ft |

## Testing of Aquifer Storage Volume Estimation in Cedar Valley, Utah

The method of aquifer storage calculation described in Section 2.3 was tested and compared to various study results for the Cedar Valley Aquifer in Southern Utah. This aquifer has recently experienced land subsidence, the opening of fissures, and some damage to infrastructure because of over pumping of the aquifer (Inkenbrandt, Lund, Lowe, Knudsen, & Bowman, 2014). The aquifer has been the subject of several studies. One difficulty of these studies is the development of an accurate water budget to estimate the storage change of the aquifer. For example, in the USGS conceptual water budget for the year 2000, aquifer recharge was estimated as 42,000 acre-ft/yr, while discharge was estimated at 38,000 acre-ft/yr, a 4,000 acre-ft/yr surplus. This estimated surplus is in direct conflict with observed drawdown of wells in the aquifer, as noted in the USGS report (Brooks & Mason, 2005). The USGS also estimated storage change in the aquifer using a groundwater model, which estimated annual recharge at 27,100 acre-ft/yr and discharge at 34,800 acre-ft/yr, a 7,700 acre-ft/yr deficit. This estimate seems more logical, as it matches the observed trends in lowering groundwater levels in the area. The USGS also estimated recharge in the aquifer using a Chloride mass-balance estimate, which yielded an estimated recharge of 20,800 acre-ft/yr. Last of all, the USGS employed a Basin Characterization Model (BCM) to estimate precipitation recharge throughout the basin. This model estimated recharge at 20,900 acre-ft/yr. The results of these studies, all carried out by the USGS, demonstrate the difficult nature of aquifer storage quantification (Heilweil & Brooks, 2010; Inkenbrandt et al., 2014; Thomas & Taylor, 1946). Based upon these reports and other studies, the Utah Division of Water Rights concluded that “the average annual groundwater deficit is probably about 7,600 acre-feet” over the last fifteen years for the Cedar Valley aquifer (D. P. E. Jones, 2016). Inkenbrandt et al. (2014) concluded that the deficit in 2000 was 10,700 acre-ft.

We used the Groundwater Level Mapping Tool to calculate aquifer storage change for the period from 2000 to 2015, corresponding to the same 15 year period studied by D. P. E. Jones (2016). In this calculation, I used a specific yield value of 0.1, which was used by Bjorklund, Sunsion, and Sandberg (1978) and by Inkenbrandt et al. (2014). Following the procedures outlined in Section 2.1 (MLR) for temporal interpolation, Section 2.2 (Kriging) for spatial interpolation, and Section 2.3 for the final calculation, we calculated the change in aquifer storage volume in the Cedar Valley aquifer between March, 2000 and March, 2015. A time series plot showing this aquifer depletion as calculated by the Groundwater Level Mapping Tool is displayed in Figure 3‑11.

Figure 3‑11: Storage Change since March, 2000 in Cedar Valley Aquifer

Over the 15-year period, the aquifer was depleted by 125,000 acre-feet, an average of 8,300 acre-feet per year. This value calculated using the Groundwater Level Mapping Tool is comparable to the USGS/Utah Division of Water Rights calculated value of 7,600 acre-feet per year. The Groundwater Level Mapping Tool also estimated a water budget deficit of 11,500 acre-feet per year for the year 2000, comparable to the Inkenbrandt et al. (2014) estimate of 10,700 acre-feet per year.

## Testing of Aquifer Storage Volume Estimation in Beryl Enterprise Area, Utah

The Groundwater Level Mapping Tool was also used to calculate changes in water storage in the Beryl Enterprise Aquifer in southern Utah, shown in Figure 3‑12.

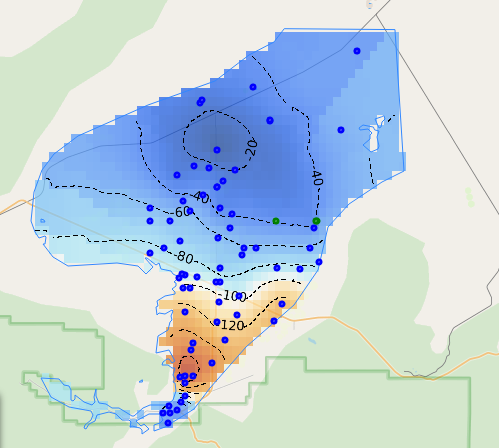


Figure 3‑12: Beryl Enterprise Aquifer in Southern Utah

The results were compared against a study prepared by the United States Geological Survey in cooperation with the Utah Department of Natural Resources, Division of Water Rights, which concluded that between 1937 and 1978, the aquifer lost between 1.3 and 1.5 million acre-feet of storage (Mower & Sandberg, 1982). Using the USGS estimate of 0.2 as the storage coefficient, the Groundwater Level Mapping Tool calculated a storage loss of 1.45 million acre-feet between 1937 and 1978. The aquifer storage change is shown in Figure 3‑13.

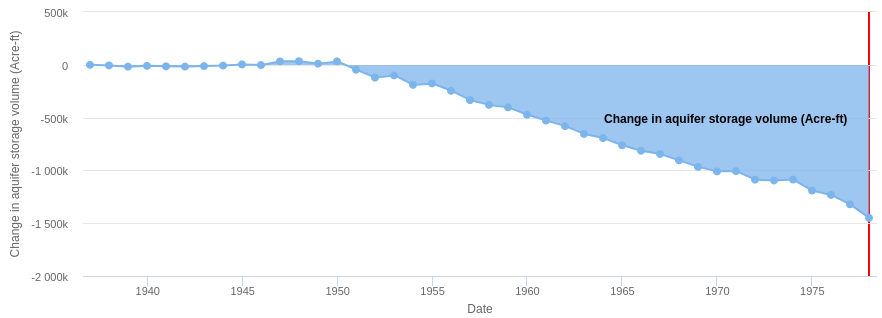


Figure 3‑13: Storage Change since March, 1937 in Beryl Enterprise Aquifer

The Utah Division of Water Rights estimated that the annual depletion rate of this aquifer around the year 2012 was approximately 65,000 acre-feet per year (K. L. Jones, 2012). The Groundwater Level Mapping Tool estimates this rate as 66,000 acre-feet per year, using a storage coefficient of 0.2, as shown in Figure 3‑14.

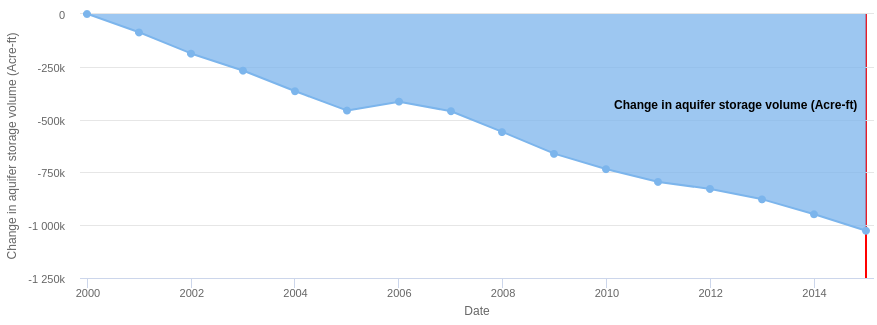


Figure 3‑14: Storage Loss since March, 2000 in Beryl Enterprise Aquifer

# Conclusions

The open source Groundwater Level Mapping Tool developed during this research has the potential for world-wide use, allowing water managers and other decision makers to quickly and easily view trends in aquifer storage levels. We developed the application to generate maps and animations of groundwater levels and drawdown which can be used to inform decision makers, enabling them to identify areas of concern and develop groundwater management plans to ensure the long term sustainability of aquifers. The application is also quite useful for calculating a rough estimate of aquifer storage change over time, typically a painstaking, laborious task. The Groundwater Level Mapping Tool’s automated method yielded results comparable to several detailed USGS studies in Utah’s Cedar Valley and Beryl-Enterprise area. These aquifer storage estimates can be used to develop a water budget and identify the safe yield of water withdrawal at which aquifers can continue sustainably.

To improve the accuracy of this application in mapping and quantifying groundwater, we developed a method of data imputation, Multi-Linear Regression Harnessing Correlated Wells. This method of temporally extrapolating recorded data to unsampled time periods, using correlated data from other wells, was used in conjunction with Kriging spatial interpolation to create maps of groundwater levels at specified time steps. This method outperformed the typical Kriging spatial interpolation method using only wells with measured data at the specified time steps, yielding more accurate maps of groundwater levels.

Overall, the Groundwater Level Mapping Tool developed during this research enables water managers to make informed decisions and implement wise management plans and regulations regarding the sustainable use of aquifers world-wide.

# Software Availability

The “Groundwater Level Mapping Tool,” which implements the methods described here, can be found at <https://tethys2.byu.edu/apps/gw/> (username: guest, password: guest). The web app was implemented on Tethys Platform (home page: https:// www.tethysplatform.org/; source code: https://github.com/ tethysplatform/tethys.git), which is an open source software framework for environmental web app development (Swain et al., 2016). The tools included in this system are all open source.

# Acknowledgements

This work was supported by NASA Grant xxxxxxxxx. Author contibutions to this article are as follows: Steven Evans conducted the bulk of the work and is the primary contributor; Norman L. Jones provided the funding, and coordinated and assisted with the research; Daniel P. Ames provided computational and technical support; and Gus P. Williams provided statistical support for this research.

# References

Ahmadi, S. H., & Sedghamiz, A. (2007). Geostatistical analysis of spatial and temporal variations of groundwater level. *Environmental monitoring and assessment, 129*(1-3), 277-294.

Bidwell, V. J. (2005). Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics. *Mathematics and Computers in Simulation, 69*(1-2), 12-20.

Bjorklund, L. J., Sunsion, C., & Sandberg, G. W. (1978). *Ground-water resources of the Parowan-Cedar City drainage basin, Iron County, Utah*. Retrieved from Utah Department of Natural Resources, Division of Water Rights:

Boezio, M., Costa, J., & Koppe, J. J. A. E. S. (2006). Kriging with an external drift versus collocated cokriging for water table mapping. *115*(3), 103-112.

Brooks, L. E., & Mason, J. L. (2005). *Hydrology and simulation of ground-water flow in Cedar Valley, Iron County, Utah* (2328-0328). Retrieved from US Geological Survey:

Deutsch, C. V., & Journel, A. (1992). GSLIB: geostatistical library and user’s guide. In: NY, Oxford University Press.

Gleick, P. H. (1993). Water in crisis. *Pacific Institute for Studies in Dev., Environment & Security. Stockholm Env. Institute, Oxford Univ. Press. 473p, 9*.

Gundogdu, K. S., & Guney, I. J. J. o. E. S. S. (2007). Spatial analyses of groundwater levels using universal kriging. *116*(1), 49-55.

Heilweil, V. M., & Brooks, L. E. (2010). Conceptual model of the Great Basin carbonate and alluvial aquifer system. *US Geological Survey Scientific Investigations Report, 5193*(2011), 191.

Inkenbrandt, P., Lund, W., Lowe, M., Knudsen, T., & Bowman, S. (2014). *Investigation of land subsidence and earth fissures in Cedar Valley, Iron County, Utah* (Vol. 150): Utah Geological Survey.

Jones, D. P. E. (2016). *Annual Recharge Estimate for Cedar City Valley*. Retrieved from Utah Department of Natural Resources: Division of Water Rights:

Jones, K. L. (2012). *Beryl Enterprise Groundwater Management Plan*. Retrieved from Utah Department of Natural Resources: Division of Water Rights:

Khorasani, M., Ehteshami, M., Ghadimi, H., & Salari, M. (2016). Simulation and analysis of temporal changes of groundwater depth using time series modeling. *Modeling Earth Systems and Environment, 2*(2), 90.

Kitanidis, P. K. (1997). *Introduction to geostatistics: applications in hydrogeology*: Cambridge University Press.

Kumar, V. (2006). Kriging of groundwater levels–a case study. *Journal of Spatial Hydrology, 6*(1).

Marchant, B., & Bloomfield, J. (2018). Spatio-temporal modelling of the status of groundwater droughts. *Journal of Hydrology, 564*, 397-413.

McKinney, W. (2010). *Data structures for statistical computing in python.* Paper presented at the Proceedings of the 9th Python in Science Conference.

Mirzavand, M., & Ghazavi, R. (2015). A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods. *Water Resources Management, 29*(4), 1315-1328.

Moritz, H. (1980). Geodetic reference system 1980. *Journal of Geodesy, 54*(3), 395-405.

Mower, R. W., & Sandberg, G. W. (1982). *Hydrology of the Beryl-Enterprise area, Escalante Desert, Utah, with emphasis on ground water; With a section on surface water*. Retrieved from

Nikroo, L., Kompani-Zare, M., Sepaskhah, A. R., & Shamsi, S. R. F. (2010). Groundwater depth and elevation interpolation by kriging methods in Mohr Basin of Fars province in Iran. *Environmental monitoring and assessment, 166*(1-4), 387-407.

Oikonomou, P. D., Alzraiee, A. H., Karavitis, C. A., & Waskom, R. M. (2018). A novel framework for filling data gaps in groundwater level observations. *Advances in water resources, 119*, 111-124.

Rouhani, S., & Wackernagel, H. J. W. R. R. (1990). Multivariate geostatistical approach to space‐time data analysis. *26*(4), 585-591.

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.

Sahoo, S., & Jha, M. K. J. H. J. (2013). Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. *21*(8), 1865-1887.

Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with python.* Paper presented at the Proceedings of the 9th Python in Science Conference.

Swain, N. R., Christensen, S. D., Snow, A. D., Dolder, H., Espinoza-Dávalos, G., Goharian, E., . . . Burian, S. J. (2016). A new open source platform for lowering the barrier for environmental web app development. *Environmental modelling & software, 85*, 11-26.

Thomas, H. E., & Taylor, G. H. (1946). *Geology and ground-water resources of Cedar City and Parowan Valleys, Iron County, Utah*. Retrieved from

Tikhonov, A. N., & Arsenin, V. I. (1977). *Solutions of ill-posed problems* (Vol. 14): Winston, Washington, DC.