Augmenting Sparse Groundwater Level Data with Earth Observations vis Machine Learning

Norm Jones[1], Steven Evans[1], Gustavious Williams[1], Jim Nelson[1], and Daniel Ames[1]

[1]Brigham Young University

November 21, 2022

Abstract

Groundwater development will provide a more stable water source and enhance food security. Sustainable groundwater development requires collecting and analyzing data produced at global and national levels and disseminating that data and knowledge to end users such as States, NGOs, municipalities, businesses, and agropastoralists in a format that is useful for planning and decision-making. In developing countries, analyzing in situ measurements to de can be challenging due to sparsity of data and lack of tools and expertise. To address these problems we have developed a web-based geospatial tool that ingests in situ water level measurements and performs temporal and spatial interpolation to build interactive animated maps, time series plots, and long-term aquifer depletion curves. We use machine learning to find correlations among Earth observation data, such as precipitation or soil moisture, with water level data and perform more accurate interpolation. This approach ensures that scarce in situ data are used as effectively and accurately as possible. This tool helps water managers gain a better understanding of groundwater resources and determine how aquifers are responding to groundwater development, droughts, and climate change.

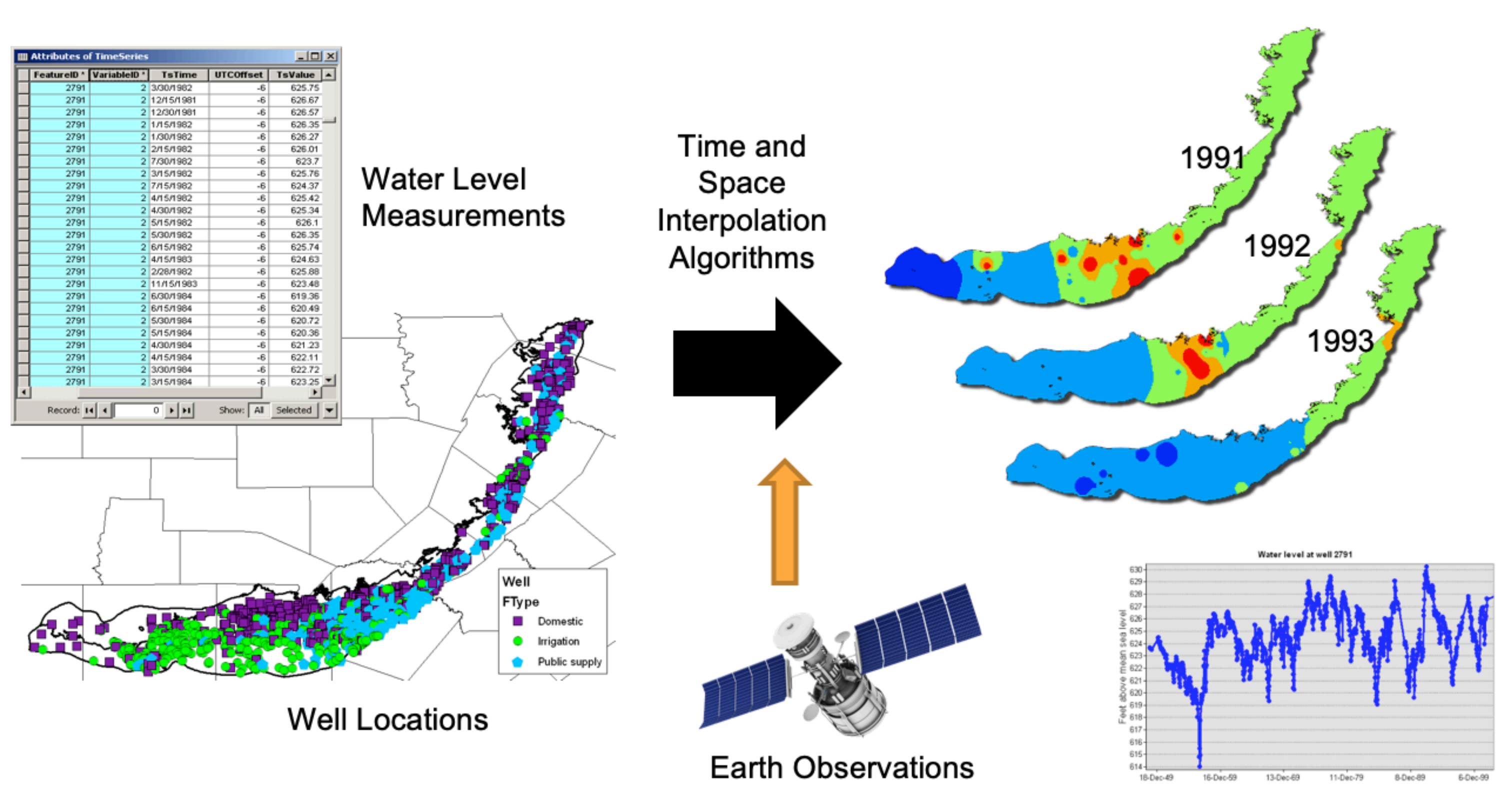
Augmenting Sparse Groundwater Level Data with Earth Observations via Machine Learning

Norm Jones , Steve Evans, Gus Williams, Jim Nelson, Dan Ames
Civil and Environmental Engineering, Brigham Young University, Provo, UT, USA

Abstract

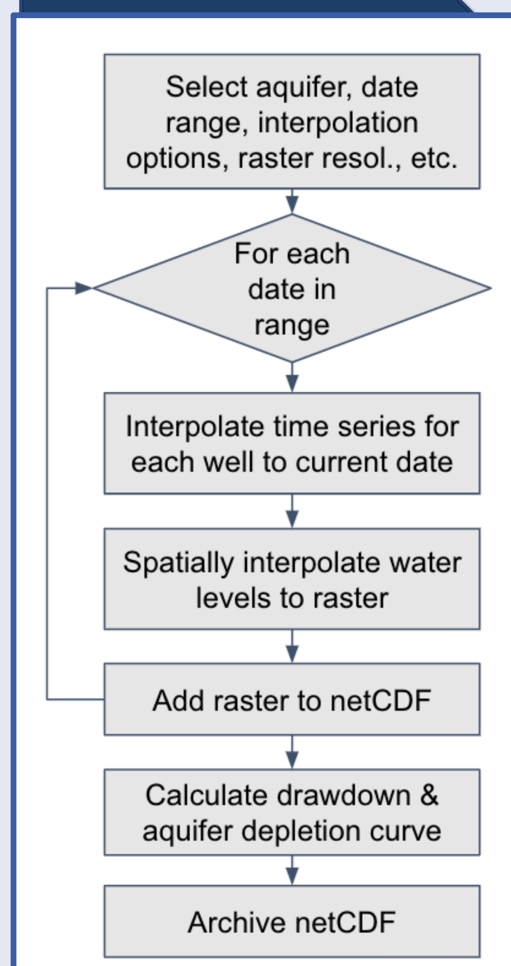
Sustainable groundwater development requires collecting and analyzing data produced at global and national levels and disseminating that data to end users in a format that is useful for planning and decision-making. In developing countries, analyzing in situ measurements to be can be challenging due to sparsity of data and lack of tools and expertise. To address these problems we have developed a web-based tool that ingests in situ water level measurements and performs temporal and spatial interpolation to build interactive animated maps, time series plots, and long-term aquifer depletion curves. We use machine learning to find correlations between Earth observation data with water level data to enhance interpolation accuracy. This tool helps water managers gain a better understanding of groundwater resources and determine how aquifers are responding to groundwater development, droughts, and climate change.

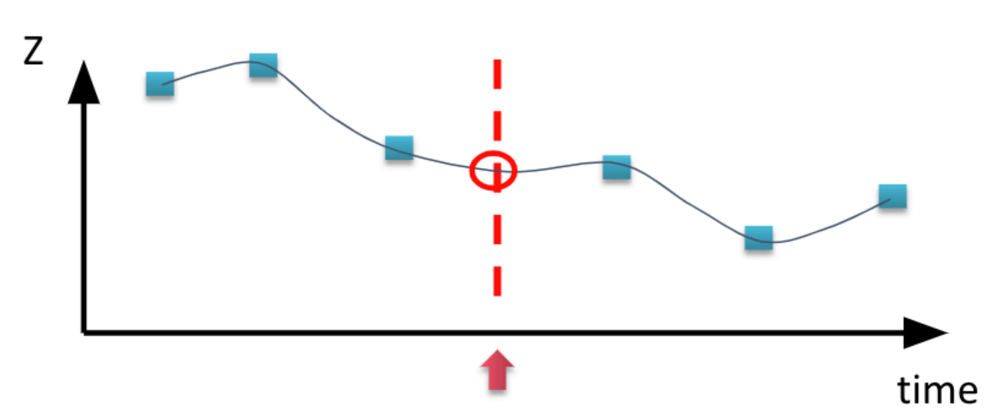
Methodology



The inputs to our tool include include well locations, historical water level measurements, and well information such as the wellhead elevation. The tool can directly ingest data from existing databases or for regions where such databases may not exist, data can be uploaded and updated using simple spreadsheets or comma separated value (CSV) text files. The in-situ water level measurements are combined with Earth observations to generate rasters and time series illustrating how groundwater resources are changing over time.

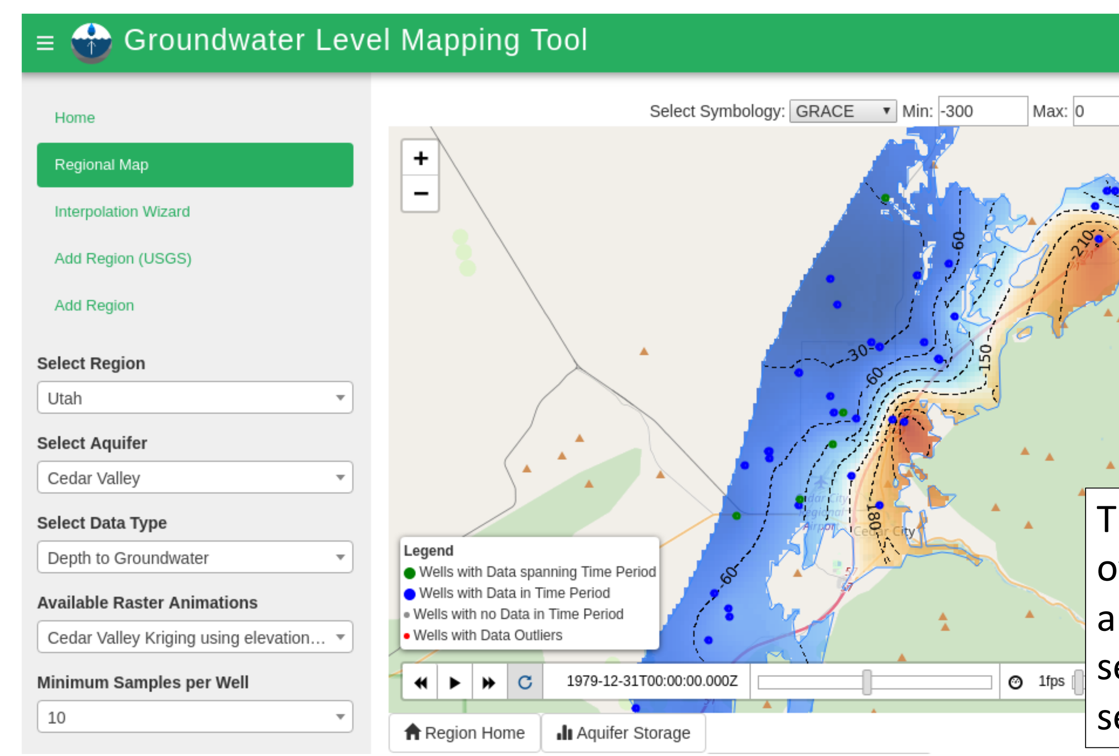
Workflow





To generate a set of maps for a selected time range, we use PCHIP or multi-linear regression/extreme learning machine to interpolate the well data in time to estimate water levels at the selected time intervals using correlated secondary information, such as precipitation or soil moisture from Earth observation data. After temporal interpolation, we spatially interpolate water level data to generate a water level map for each time step. The resulting raster maps are stored in netCDF files that are archived on a THREDDS server for distributed access.

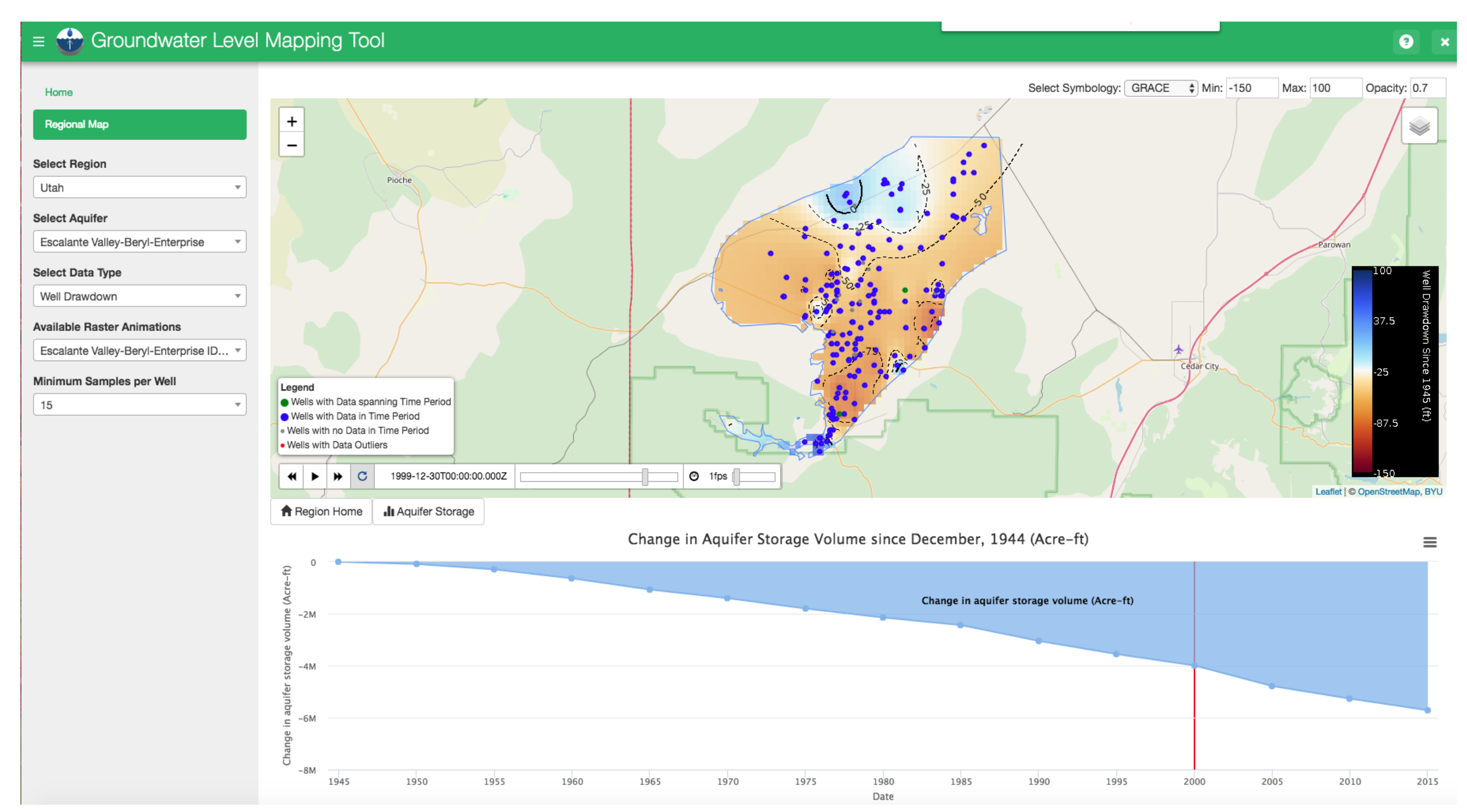
Web App



The results are visualized in a web app. The user first selects a region (state), then a subregion (aquifer) and then the data product to visualize (water level, depth to water table, or drawdown).

The rasters are animated over time using a Leaflet animation plug-in. Time series are also displayed for selected well locations.

Storage Depletion

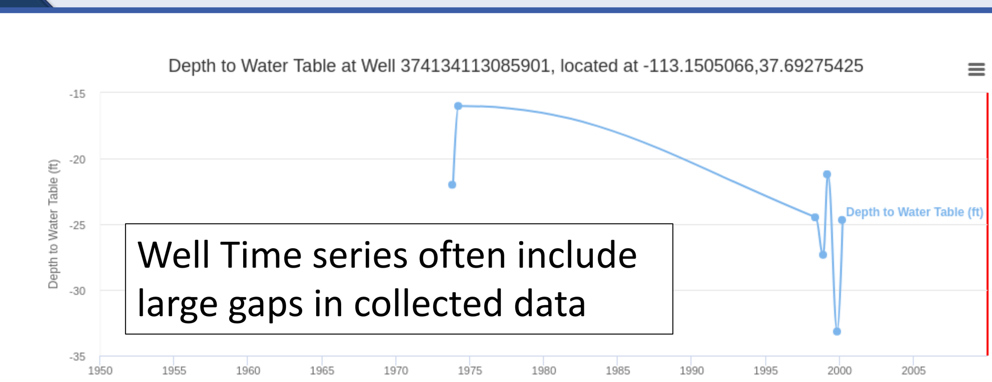


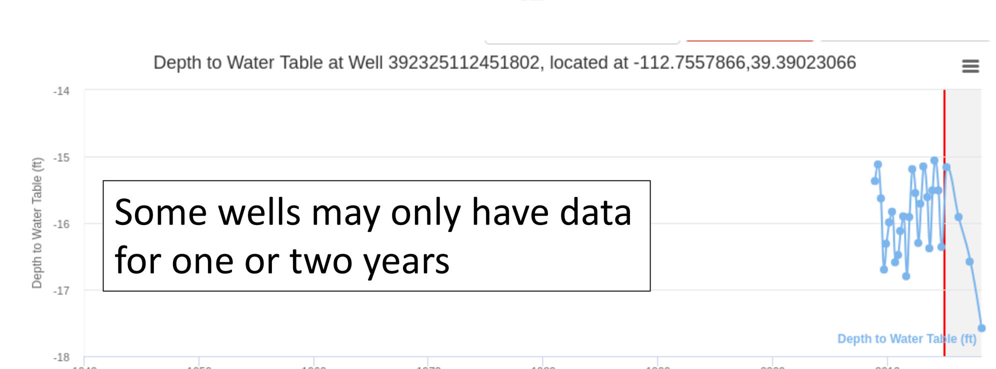
For a more holistic assessment of storage depletion for a given aquifer, we multiply the drawdown rasters for each time interval by the aquifer area and an average storage coefficient and aggregate the results over time to generate a storage depletion curve. This provides a useful tool for water managers to assess long-term conditions. The results can be incrementally updated as new in situ data is collected and imported to the system.

Data Gaps

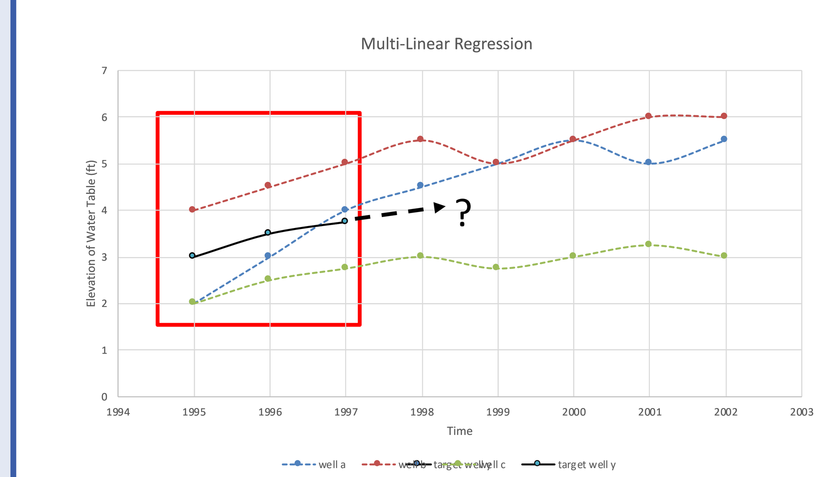
Monitoring wells are often samples at irregular or sporadic intervals. It is not uncommon for monitoring wells to be abandoned, or to have quite brief periods of record. We may have only one or two years of information from a well.

How can we use **machine learning** to best make use of what little data we have?

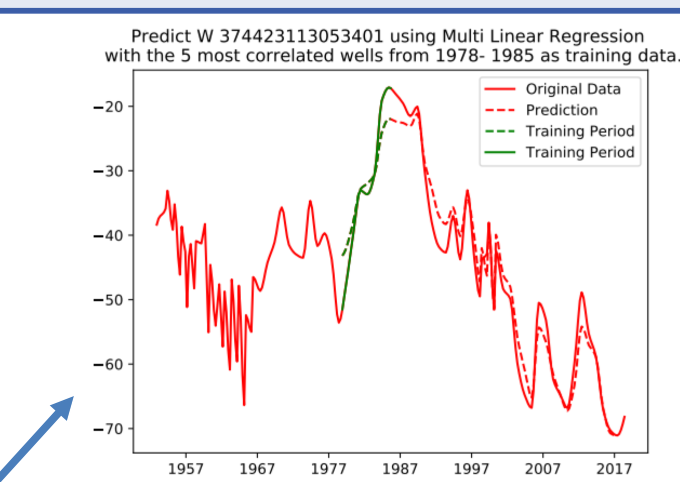


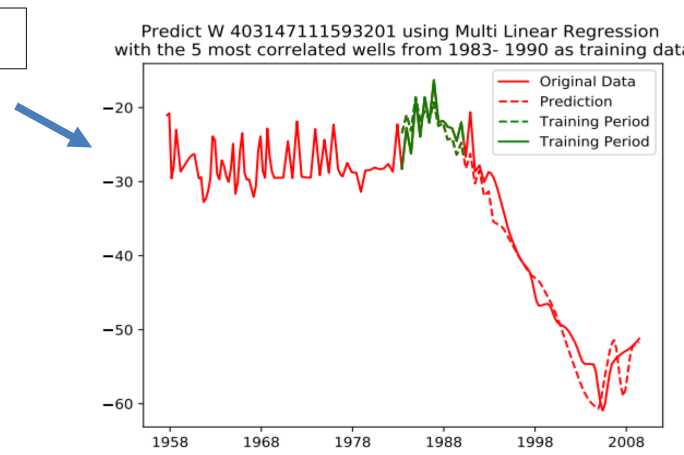


Multi-Linear Regression



RESULTS





$$\vec{Y} = A\vec{\beta} + \vec{e}$$

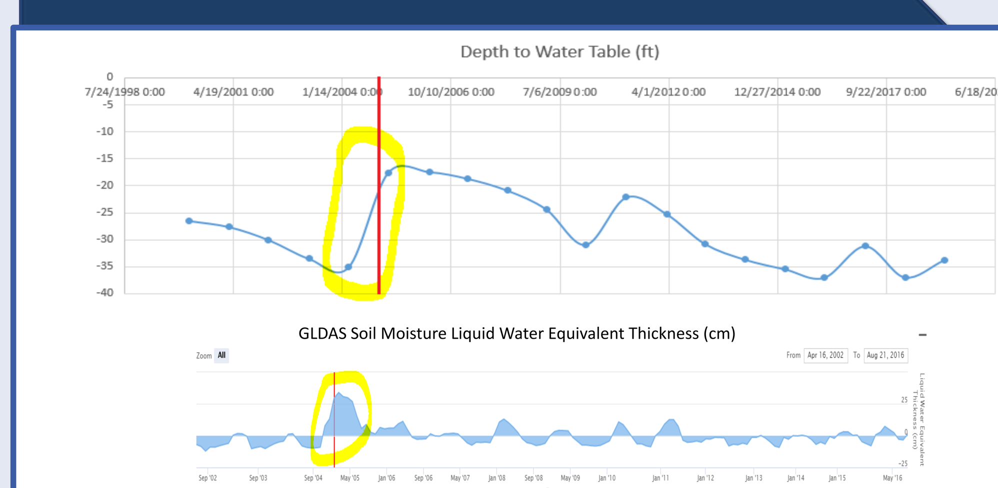
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \vec{e}$$

Normalize A and $\vec{Y}$

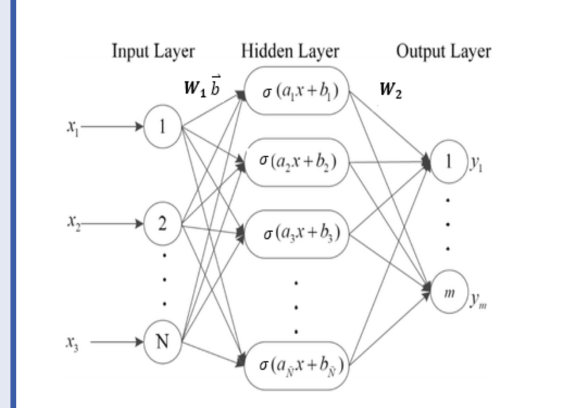
Determine $\vec{\beta}$ such that $\vec{e}$ is minimized using least squares

$$A^T\vec{Y} = [A^T A + \lambda I]\,\vec{\beta}$$

Extreme Learning Machine



Groundwater levels are correlated with soil moisture models, drought indicators, and other Earth observations (Palmer Drought Severity Index, GLDAS Root Zone Soil Moisture, GRACE Total Water Storage, etc.



$\vec{Y} = W2 * \sigma(W1\ X + \vec{b})$

X = input variables

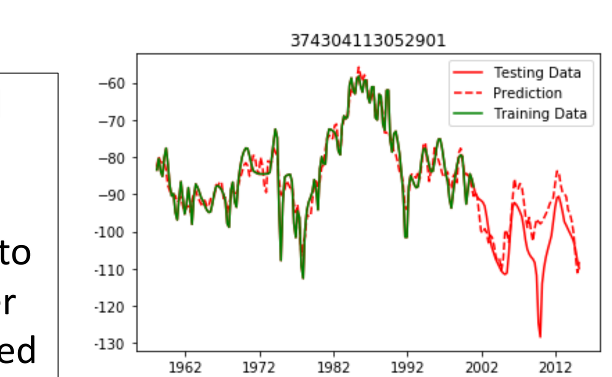
$W1$ = random values

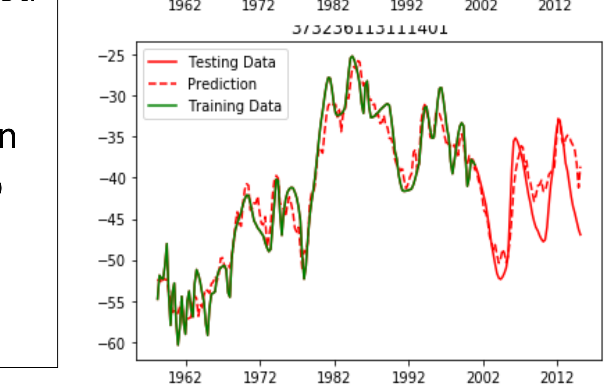
$W2$ = coef. fitted via least squares

σ =rectifier function

$\vec{b}$ = bias vector

We trained a neural network using an extreme learning machine algorithm to predict groundwater levels from correlated Earth observations. Results were relatively accurate in areas not subject to excessive groundwater pumping.





This work is being funded by the following NASA grants:

An AmeriGEOSS Cloud-based Platform for Rapid Deployment of GEOGLOWS Water and Food Security Nexus Decision Support Apps (80NSCC18K0440)"

Geospatial Information Tools That Use Machine-Learning to Enable Sustainable Groundwater Management in West Africa (80NSSC20K0155)

For more information:

Norm Jones
njones@byu.edu