Creating a Spatial Groundwater Database from Historical Records: Adventures in Interpolation

Sinead Anderson, Samuel Blanchard, Maggi Kelly

Environmental Leadership Pathway

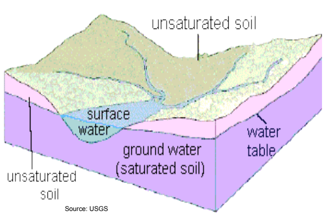
College of Natural Resources, University of California, Berkeley

Introduction & Background

Abstract

Groundwater is an invaluable resource that is extremely important to the vitality of California. Over 30 million people, including large industries and agriculture use this natural resource and may have the potential to contaminate or deplete California's groundwater resources. Groundwater is difficult to monitor from the surface and historical data on groundwater conditions and use is scarce and difficult to locate for the public and regulatory organizations. This study uses advanced Geographical Information System (GIS) tools to compile and digitize California's groundwater data in a pilot study located in central California. Historic 1950s groundwater data from the United States Geological Survey (USGS) from wells near Fresno, CA were digitized from paper and put into a spatial database. Spatial interpolations of the 1950s digitized data were created to compare the data to current groundwater data. Interpolations use information from known data points to create a continuous surface of information where data points do not exist. Using three methods of interpolation: Kriging, Spline, IDW visual trends can be seen throughout the data for various attributes of groundwater such as sodium or pH concentration. Research proved Kriging to be the most accurate method providing the least error in prediction rates throughout the interpolated surface. Through interpolation it is possible to visually identify which wells are contaminated and which wells are not. The data allows us to see pollution, contaminates, and concentrates in the historic study area for comparison with the current data to see changes over time. This project is important because it helps us understand what the historic trends in our groundwater resources are, in order to address the potential problems we might face today.

Background:



Groundwater is water that is located beneath the surface held in permeable rock, and soil that feeds into rivers and lakes.

Over one half of the U.S. population relies on groundwater for its drinking water supply. Even more groundwater is used for irrigating agriculture, and its industrial use is growing every day. Groundwater is valuable because it's plentiful and clean. There is about 50 times more water underground, than in all the lakes and rivers on the Earth's surface combined. And in many areas, especially those with dry climates, groundwater is the most abundant and economical source of water available. Because it is filtered as it passes through the soil, groundwater tends to be less polluted than surface water; however this valuable resource could potentially be threated.

References

Davis, G.H. and J.F. Poland, 1957. Ground-Water Conditions in the Mendota-Huron Fresno and Kings Counties, California. USGS Water-Supply Paper 1360-G. U.S Government Printing Office, Washington D.C.

Interpolating the Difference

Methods:

Historic groundwater data was digitized from a analog paper map to a spatial digital database using ESRI ArcGIS™.

ESRI ArcGIS™ Geostatistical Analyst: Statistical Tool for Data Exploration, Modeling, and Advanced Surface Generation was used to interpolate of the digitized data. Geostatistical analyst allows for the interpolation of known point values over a continuous surface to estimate values where data does not exist. Using ArcGIS™ Geostatistical Analyst three methods of interpolation were chosen: Kriging, Spline and Inverse Distance Weighted (IDW) to visually assess and determine which method would best interpolate; sodium, pH, and depth of the groundwater data.

- Kriging- creates an estimated surface using a scattered set of points with z-values Formula for Kriging:

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i)$$

- where:
  - $Z(s_i)$ = the measured value at the ith location
  - λ_i = an unknown weight for the measured value at the ith location
  - s_0 = the prediction location
  - N = the number of measured values

- Spline- estimates values through a mathematical function which decreases overall surface

The algorithm used for the Spline tool uses the following formula for the surface interpolation:

$$S(x,y) = T(x,y) + \sum_{j=1}^{N} \lambda_j R(r_j)$$

- where:
  - $j = 1, 2, \ldots, N$
  - N is the number of points.
  - λ_j are coefficients found by the solution of a system of linear equations.
  - r_j is the distance from the point (x,y) to the j^{th} point.

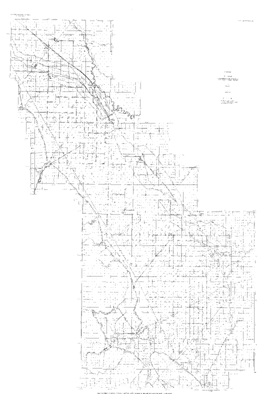
$T(x,y)$ and $R(r)$ are defined differently, depending on the selected option.

- Inverse Distance Weighted (IDW)- decides cell values using a linearly weighted amalgam of a set of sample points.

Data

Historic 1957 analog groundwater data from the United States Geological Survey (USGS) from wells near Fresno, CA were acquired. The data was used to generate interpolations of groundwater attributes.

Data source examples:



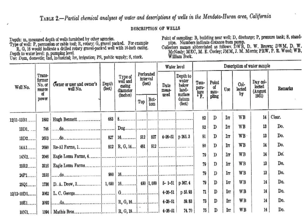
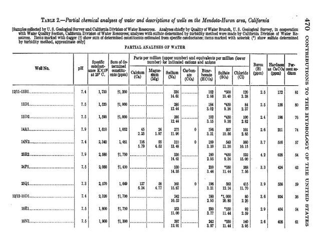
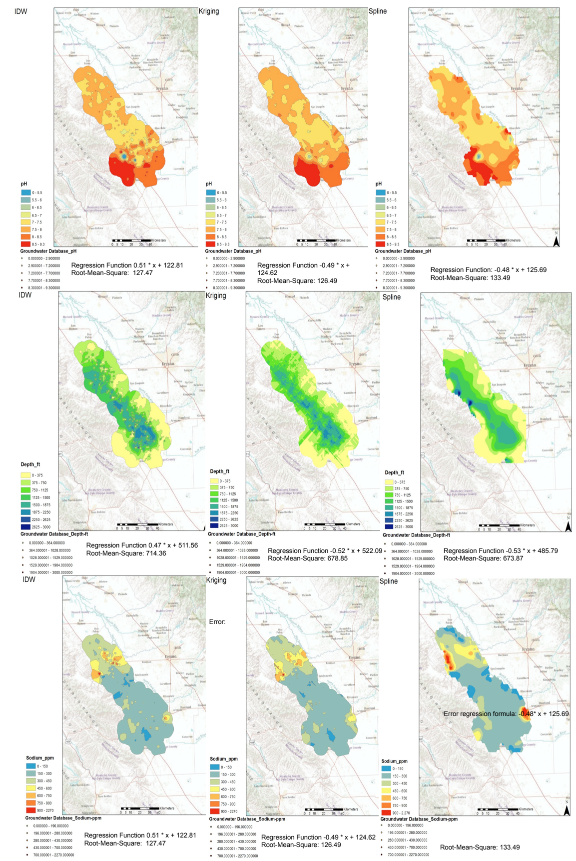


TABLE 2.—Partial chemical analyses of water and descriptions of wells in the Mendota-Huron area, California



Example of 1957 analog spatial and tabular ground water quality Information from wells near Fresno, CA (Davis and Poland 1957)

Results



Images of Interpolated attributes: sodium, pH and depth by author

Conclusion

Based on our results we found that Kriging produced lowest prediction error. Although Kriging is similar to Inverse Distance Weighted(IDW) it shows to be the best deterministic method for displaying and querying groundwater attributes such as: sodium, pH and depth. The significance of this research is beneficial to the public and will be added to an online database for public display and query. Furthermore, the 1957 tabular ground water quality information from wells near Fresno, CA has been digitized for comparison with the USGS current groundwater data in order to see and compare changes over time.