Temporal and Spatial prediction of groundwater level using Artificial Neural Networks, Fuzzy logic and Kriging interpolation

Evdokia Tapoglou¹, George P. Karatzas¹, Ioannis C. Trichakis¹,², Emmanouil Varouchakis¹

(1) School of Environmental Engineering, Technical University of Crete, Greece (2) Water Resources Unit, Institute for Environment and Sustainability, Joint Research Centre, European Commission
Scope and Objectives

Groundwater Hydraulic Head Simulation

Artificial Neural Networks (ANNs) → Temporal simulations

Kriging → Spatial Interpolation

Combination of ANN and Kriging

- Use of ANN for temporal simulation
- Use of Kriging for spatial interpolation
- Use of Fuzzy logic for the connection of these models
Artificial Neural Networks (ANN)

Alternative to conventional numerical modeling techniques

Artificial neurons connected through synaptic weights and organized in a network

Information processing is done in different levels:
- Input layer → Numerical input in the network
- Hidden layer(s) → Computational nodes
- Output layer → Final numerical result

Training Algorithm → Back propagation
Fuzzy logic

Extension of mathematical logic → degree of truth

Choose the appropriate neighbors for Kriging, for every prediction point

Takes into consideration:
- Distance between the prediction and the observation points
- ANN training and testing error

Observation points with the largest possibility of being optimal are used from kriging algorithm
Kriging

Spatial interpolation technique

Assumes that the closer input data are, the more positively correlated estimation errors are.

Simulated value $\bar{z}(x_k)$ is a linear combination of $\bar{z}(x_i)$ values in nearby sampling points with a weight $\lambda_i$

$$\bar{z}(x_k) = \sum_{i=1}^{\mu} \bar{z}(x_i) \cdot \lambda_i$$
The weights are determined from variogram, representing the variance of the difference between the measurements at two locations

$$\gamma(h) = \frac{1}{2M(h)} \sum_{i=1}^{M(h)} \{z(x_i) - z(x_i + h)\}^2$$

Experimental Variogram → Fitted in Theoretical model

### Variogram models

<table>
<thead>
<tr>
<th>Model</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$\gamma(h) = C \cdot h$</td>
</tr>
<tr>
<td>Exponential</td>
<td>$\gamma(h) = \begin{cases} C \cdot [1 - e^{-\frac{h}{\lambda}}], &amp; h &gt; 0 \ 0, &amp; h = 0 \end{cases}$</td>
</tr>
<tr>
<td>Power-law</td>
<td>$\gamma(h) = \theta \cdot h^{2H}$</td>
</tr>
</tbody>
</table>
Algorithm Development

Developed in Visual Studio 2010 → VB.NET

Visualization → MATLAB, using the .NET COM Interoperability function.

Data processing

Collection and evaluation of available data
Input parameters

- Timeseries statistical characteristics
- Various time lags considered
- Final input parameters → Combination of parameters and correlation coefficient
Algorithm Development

Developed in Visual Studio 2010 → VB.NET

Visualization → MATLAB, using the .NET COM Interoperability function.

Data processing → ANN Training (80% of dataset)

Training
- One ANN per well with available data
- Hydraulic head change simulation (output parameter)
- Training and evaluation RMSE
- Nash – Sutcliffe Efficiency coefficient
Algorithm Development

Developed in Visual Studio 2010 → VB.NET

Visualization → MATLAB, using the .NET COM Interoperability function.

Data processing → ANN Training (80% of dataset) → ANN Evaluation (100% of dataset)

**ANN Simulation:**
- Temporal simulation of hydraulic head in each well
- Use of already trained networks
Developed in Visual Studio 2010 → VB.NET

Visualization → MATLAB, using the .NET COM Interoperability function.

Fuzzy logic
- Connection between ANN and Kriging
- Choose the appropriate neighbors
Algorithm Development

Developed in Visual Studio 2010 → VB.NET

Visualization → MATLAB, using the .NET COM Interoperability function.

Data processing → ANN Training (80% of dataset) → ANN Evaluation (100% of dataset)

Use of fuzzy logic? → Fuzzy logic for neighbor selection

Kriging
- Spatial simulation
- Interpolates the ANN results in every time step

Kriging → Spatial – Temporal Model Maps
Study area

Near Munich, Germany – Along Isar river

Available data: 1/11/2008-31/10/2012

Area ~7800 km²

- Meteorological Data: 7 stations
- Surface water: 5 stations

Data source: © Bavarian State Office for the Environment, www.lfu.bayern.de
Study area

Near Munich, Germany – Along Isar river

Available data: 1/11/2008-31/10/2012

Area ~7800 km²

- 64 wells with daily hydraulic head measurements
- 2 wells with missing data (used for model validation)

Data source: © Bavarian State Office for the Environment, www.lfu.bayern.de
Correlation coefficient

\[ Correl(A, B) = \frac{\sum(a - \bar{a})(b - \bar{b})}{\sqrt{\sum(a - \bar{a})^2 \sum(b - \bar{b})^2}} \]

Meteorological data → 5 inputs

- 4 rainfall from the best correlated stations (consecutive days are preferred)
- 1 temperature measurement

Surface water level → 3 inputs

- 2 consecutive days input for the station with the best correlation
- 1 input for the station with the second best correlation
Training Algorithm → Back propagation

Number of hidden nodes

- "A factor of three times as many samples as network parameters training is adequate to achieve good generalization"

- Available data sets: 1460 days
- Number of weights that can be trained: 486
- Input Layer: 8 nodes
- Output layer: 1 node
- For 1 hidden layer: 54 nodes
- For 2 hidden levels: 20 and 15 nodes respectively
Performance index

- Training RMSE
- Testing RMSE

The architecture with 2 hidden layers (20,15) was chosen
ANNs/wells were divided into 5 categories (similar performance) → Results presented only from one well representative of its category

Additional performance index → Nash-Sutcliffe

ANN performance
Observed and simulated values vs. Time

ANN Training results
Temporal and Spatial prediction of groundwater level using Artificial Neural Networks, Fuzzy logic and Kriging interpolation.

ANN Training results

Observed vs. Simulated values
Trained ANNs are used to simulate temporally the hydraulic head change

Fuzzy logic → Optimal selection of neighbors for the Kriging to use

30 wells with a combination of good training/testing error and small distance from the prediction point are used
Computation of experimental variogram for every time step and every prediction point

Leave one out Cross-validation for points were ANN simulated data are available

Performance indicators

<table>
<thead>
<tr>
<th>Cross - Validation Error Types</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>( RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} [z^*(s_i) - z(s_i)]^2} )</td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
<td>( MAE = \frac{1}{N} \sum_{i=1}^{N}</td>
</tr>
<tr>
<td>Bias</td>
<td>( Bias = \frac{1}{N} \sum_{i=1}^{N} [z^*(s_i) - z(s_i)] )</td>
</tr>
</tbody>
</table>
## Evaluation – Variogram selection

<table>
<thead>
<tr>
<th>Variogram type</th>
<th>Cross-Validation RMSE</th>
<th>Cross-Validation MAE</th>
<th>Cross-Validation Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With Fuzzy</td>
<td>Without Fuzzy</td>
<td>With Fuzzy</td>
</tr>
<tr>
<td>Linear</td>
<td>2.18·10^{-2}</td>
<td>5.38·10^{-2}</td>
<td>1.41·10^{-2}</td>
</tr>
<tr>
<td>Exponential</td>
<td>5.57·10^{-2}</td>
<td>3.00·10^{-1}</td>
<td>1.94·10^{-2}</td>
</tr>
<tr>
<td>Power</td>
<td>1.95·10^{-2}</td>
<td>2.24·10^{-2}</td>
<td>1.33·10^{-2}</td>
</tr>
</tbody>
</table>
### Evaluation – Results

**Results for the 2 wells with incomplete data**

<table>
<thead>
<tr>
<th></th>
<th>Time Step (days)</th>
<th>Time Step (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Error</td>
<td>Variance</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>Max Difference</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>Min Difference</td>
</tr>
<tr>
<td>Average Error</td>
<td>$4.71 \times 10^{-4}$</td>
<td>0.206 m</td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td>1.66 \times 10^{-6}$</td>
</tr>
<tr>
<td>Max Difference</td>
<td>0.206 m</td>
<td></td>
</tr>
<tr>
<td>Min Difference</td>
<td>1.66 \times 10^{-6}$</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.04 \times 10^{-4}$ m</td>
<td></td>
</tr>
</tbody>
</table>

Temporal and Spatial prediction of groundwater level using Artificial Neural Networks, Fuzzy logic and Kriging interpolation.
### Evaluation - Results

<table>
<thead>
<tr>
<th>Category</th>
<th>Average</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Error</td>
<td>$1.4\cdot10^{-4}$</td>
<td>$7.4\cdot10^{-4}$</td>
<td>$1.6\cdot10^{-5}$</td>
</tr>
<tr>
<td>Testing Error</td>
<td>$1.8\cdot10^{-4}$</td>
<td>$7.2\cdot10^{-4}$</td>
<td>$4.2\cdot10^{-5}$</td>
</tr>
<tr>
<td>Kriging Error Variance</td>
<td>$2.4\cdot10^{-5}$</td>
<td>0.03</td>
<td>0</td>
</tr>
<tr>
<td>Cross Validation Error (RMSE)</td>
<td>$7.6\cdot10^{-3}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Spatial – temporal simulation of hydraulic head

Use of Fuzzy logic can improve the results

Small Simulation errors

Calculation of Kriging parameters uncertainty

Calculation of ANN uncertainty
Thank you!

Contact email: etapoglou@gmail.com