International Congress on Environmental Modelling and Software

3rd International Congress on Environmental Modelling and Software - Burlington, Vermont, USA - July 2006

Jul 1st, 12:00 AM

Application of an Artificial Neural Network for Analysis of Subsurface Contamination at the Schuyler Falls Landfill, NY

Lance E. Besaw

Donna M. Rizzo

P. J. Mouser

Application of an Artificial Neural Network for Analysis of Subsurface Contamination at the Schuyler Falls Landfill, NY

L.E. Besaw[a], **D.M. Rizzo**[b] **and P.J. Mouser**[b]

[a] *Department of Civil and Environmental Engineering, University of Vermont, Burlington, VT, USA; email: lbesaw@cems.uvm.edu*

[b] *Department of Civil and Environmental Engineering, University of Vermont, Burlington, VT, USA*

Abstract: Current subsurface site characterization, plume delineation, remediation designs and monitoring network designs that rely on a limited, albeit large, number of sparsely collected data, tend to be expensive, cumbersome and frequently inadequate for solving multi-objective, long-term environmental management problems. We present a subsurface characterization methodology that integrates multiple types of data using a modified counterpropagation artificial neural network (ANN) to provide parameter estimates and delineate groundwater contamination at a leaking landfill. Apparent conductivity survey data and hydrochemistry data (*i.e.* heavy metals, $BOD_{5,20}$, chloride concentration, *etc.*) are used to estimate the extent of subsurface contamination at the Schuyler Falls Landfill, located in Clinton County NY. The results of this research illustrate the feasibility of combining principal component analysis (used to reduce data dimensionality) with the counterpropagation ANN and traditional geostatistical methods (kriging) to estimate subsurface contamination. The ANN methodology for obtaining parameter estimates is data-driven and can easily incorporate a large number of data types obtained from diverse measurement techniques. This technique is also flexible as it does not require the computation of large covariance matrices and, once the ANN is trained, can produce realizations for subsurface characterization and monitoring in real time.

Keywords: Artificial neural networks; Counterpropagation; Parameter estimation; Kriging

1. INTRODUCTION

Unlined, leaking landfills are a major source of groundwater pollution in the United States [USEPA, 2002]. Environmental concerns associated with subsurface contamination from landfill leachate include volatile organic carbons (*e.g.* benzene and trichloroethylene or TCE), heavy metals (*e.g.* mercury and lead) and diminished water quality (*e.g.* lowered pH and increased $BOD_{5,20}$). The extent of groundwater contamination from a leaking landfill in Schuyler Falls, N.Y. is estimated using a modified counterpropagation artificial neural network (ANN) that combines hydrochemistry data and apparent conductivity survey data.

In many earth science parameter estimation applications, the parameter of interest (primary data) is less abundant and more expensive to collect than the secondary data. The objective of this study is to delineate the leachate plume by estimating the more abundant apparent conductivity survey data while using the hydrochemistry as secondary data. To make the hydrochemistry data more manageable, the dimensionality of the original hydrochemistry variables is reduced using principal component analysis.

This study presents the application of a counterpropagation ANN parameter estimation technique to delineate subsurface contamination of a leachate plume. The counterpropagation ANN is data-driven, can incorporate many data types obtained from diverse measurement techniques and compares well with traditional geostatistical estimating methods (*e.g.* kriging methods). The feasibility for implementing this ANN estimation technique to assist the water quality assessment at a full scale landfill is demonstrated in this work.

2. BACKGROUND

2. 1 Artificial Neural Networks

Artificial neural networks (ANNs) are nonparametric statistical tools that can be viewed as universal approximators. ANNs specialize in identifying non-linear relationships given extremely large datasets and have a relatively simple mathematical architecture that makes them computationally efficient. This computational efficiency offers significant advantages for predictions using real time sensors or large data sets that would be unwieldy with other estimation methods.

ANNs were developed as large parallel-distributed information processing systems that attempt to model the learning procedure of the human brain [Rumelhart and McClelland, 1988]. Their architecture consists of layers of nodes with weighted arcs connecting the nodes within the different layers. The information passing structure, the number of layers, the number of nodes and the algorithms selected for adjusting the internal weights create alternative types of ANNs.

The counterpropagation ANN is a supervised learning algorithm that self-adapts to create statistical mappings of predictor and associated response vectors. It sequentially combines the Kohonen self-organizing learning algorithm and the Grossberg outstar structure [Hecht-Nielsen, 1987]. The execution of the counterpropagation algorithm is defined by two phases: an adaptation phase (training or calibration) and an operational phase (interpolation or mapping). Once trained, the network functions as a look-up table mapping a set of input predictor vectors $\mathbf{x} = (x_1, x_2\ldots)$ to an associated response $\mathbf{y} = (y_1, y_2\ldots)$ defined by some non-linear function $\mathbf{y} = \varphi(\mathbf{x})$.

Rizzo and Dougherty [1994] introduce the concept of applying a counterpropagation ANN to map discrete spatially distributed fields of log hydraulic conductivity. In another work, a comparison of traditional geostatistical estimation techniques (*e.g.* ordinary kriging and cokriging) with the counterpropagation ANN demonstrated that counterpropagation was a useful site characterization tool [Rizzo et al. 1996]. Although the application of counterpropagation ANNs to parameter estimation is not mainstream, other applications in earth related disciplines include its utilization to classify soil samples [Fidencio et al. 2001; Sullivan, 1999; Sullivan et al. 1998] and to model nonlinear pH-processes [Nie et al. 1996].

2. 2 Schuyler Falls Landfill

The Schuyler Falls Landfill is an unlined municipal waste landfill located in the town of Schuyler Falls, Clinton County, NY; see Figure 1 (a). The landfill is approximately 30 acres in size and is located roughly 0.5 miles south of the Saranac River and 7.5 miles west of Lake Champlain [Barton and Loguidice, 1996]. The site began accepting municipal waste in 1977 and continued operation until closure and capping in 1996. After this time, additional lined landfills at the site were undergoing permitting and construction to the south and west of the unlined site.

(a)

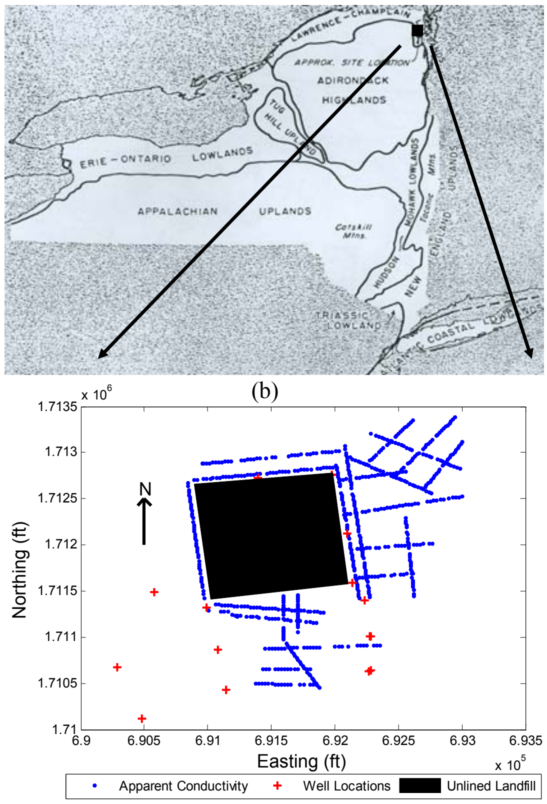

Figure 1. Location of (a) Schuyler Falls Landfill in upstate New York and (b) spatial representation of unlined landfill, available hydrochemistry well data and apparent conductivity data.

In a 1993 hydrologic investigation, groundwater contaminated by landfill leachate was found to be migrating from the southern and eastern regions of the unlined landfill (groundwater flow is in the north-east direction). The leachate plume consists of halogenated volatile organic compounds, including trichlorofluoromethane (Freon 11), trichloroethane (TCA), trichloroethylene (TCE), tetrachloroethene (PCE) and petroleum byproducts including benzene, toluene, alkanes, cylcloalkanes, and alkenes [Dunn Science Corp. 1986]. During 1993, the total VOCs detected on the southern boundary of the landfill exceeded 100,000 relative response units using a Petrex Soil Gas Survey; and an extensive closure investigation was ordered for the site [Barton and Loguidice, 1993]. Remediation activities, including two pumping wells and the installation of vertical barrier walls, began in 1997 in attempt to contain the leachate plume.

In 1996, apparent conductivity data was collected using an EM-34 electromagnetic survey along various transects around the landfill footprint to delineate the horizontal and vertical extent of the leachate plume [Mouser et al., 2005], see Figure 1 (b). Apparent conductivity (uS/cm) is often used to delineate contaminant plumes by measuring the strength of magnetic conductivity expressed as excess ions present in the contaminated groundwater and leachate. A total of four apparent conductivity datasets are used for plume delineation and the most conservative final estimation field is presented. In addition, from 1996 through 1997, hydrochemistry data was collected roughly quarterly at 17 wells throughout the landfill. More than 20 hydrochemistry variables were collected including pH, turbidity, $BOD_{5,20}$, chloride, and lead concentration. This study will use the hydrochemistry data to aid in the estimation of apparent conductivity near the landfill.

3. METHOD OF ANALYSES

Principal component analysis (PCA) is often used to reduce the dimensionality of datasets with intercorrelated variables. PCA was applied to the original Schuyler Falls hydrochemistry variables (more than 20) to generate a new (reduced) set of variables, called principal components, that explain a significant amount of variance found in the original data. The JMP statistical software (SAS Institute, Inc, V5.0.1) was used to generate the principal components from the analysis of the covariances given the original hydrochemistry data. The first principal component (PC1) describes 80% of the total variance and is thereby used to represent the more than 20 hydrochemistry data for each of the 17 wells. PC1 is therefore incorporated by the ANN as a predictor variable to estimate apparent conductivity.

3.1 Geostatistical Analysis

The field of geostatistics provides a methodology for describing and modeling the structure of a parameter of interest, in either space or time, based on some limited number of measurements. The spatial structure of a parameter of interest is described by generating an experimental semi-variogram given some limited number of measurements. A semi-variogram model is then best fit to the experimental semi-variogram (γ). This model describes how a parameter of interest exhibits dissimilarity as a function of separation distance and is computed as:

$$\gamma(\mathbf{h}) = \frac{1}{2}\left[u(\boldsymbol{\alpha}) - u(\boldsymbol{\alpha}+\mathbf{h})\right]^2, \qquad (1)$$

where u is the parameter of interest, $\boldsymbol{\alpha}$ represents the spatial coordinates and $\mathbf{h}$ is a separation distance. Once generated, the best fit semi-variogram is defined by a type of curve (*i.e.* exponential, spherical, Gaussian), a maximum distance of correlation (range), a parameter that accounts for measurement errors or variability among short scale measurements (nugget) and the semi-variance plateau (sill) [Isaaks and Srivastava, 1989], see Figure 2.

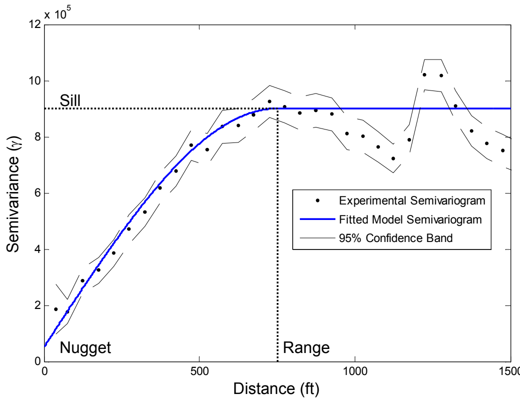

Figure 2. Apparent conductivity semi-variogram.

Ordinary kriging is an estimation method that is referred to as the "best linear unbiased estimator", or BLUE. This acronym is associated with ordinary kriging because the method generates parameter estimates using weighted linear combinations of available data while minimizing the variance of errors and fixing the mean residual to zero [Isaaks and Srivastava, 1989]. The estimates produced by the ordinary kriging technique are dependent on the spatial structure described by the model semi-variogram. The type of model curve, range of correlation, nugget and sill are all parameters that when used in conjunction with measurement data produce estimates of the parameter of interest by solving the kriging equations and inverting the covariance matrices, see [Deutsch and Journel, 1998].

3.2 Counterpropagation ANN

The counterpropagation ANN has many features that make it an excellent tool for parameter estimation. The ANN is data-driven and has the ability to directly map inherent non-linear relationships between the predictor and associated response vectors. The "data-driven" feature is advantageous in that the underlying process being mapped does not need to be fully understood. The complex relationship is extracted directly from the data. The disadvantage is that large amounts of data are required to ensure the mapping is statistically significant. For applications in which a parameter is collected nearly continuously in the vertical or horizontal direction (*i.e.* apparent conductivity) the counterpropagation ANN is an ideal estimation method.

The counterpropagation ANN is made up of three layers, an input layer, a hidden (or Kohonen) layer, and an output (or Grossberg) layer [Hecht-Nielsen, 1987]. Each layer contains a set of nodes that are fully connected to every node in the adjacent layer(s), by a set of weights. These weights, initially set to random values between 0 and 1, form a parallel information passing topology that allow for rapid data processing and prediction.

As preluded to earlier, the counterpropagation ANN is executed in a training phase (calibration) and an interpolation phase (estimation). During calibration, each training vector (made up of predictor variables, *i.e. Easting* and *Northing*) is presented sequentially to the input nodes of the network, see Figure 3 (a), and the corresponding measured output (***y***=*apparent conductivity*) is assigned as the target (desired output). The internal weights are then adjusted appropriately to make the ANN output for each training pattern match the known target.

The passage of all training vectors through the ANN concludes one iteration of calibration. The process is performed iteratively until the outputs from the ANN match the known targets to a pre-specified convergence criterion (in this study 10^{-6}). Once the ANN has successfully “learned” the inherent relationships from the collected measurements (*a.k.a.* calibrated), the internal weights become fixed and the ANN is used for estimation.

For the interpolation phase, data patterns with unknown target outputs are presented to the ANN sequentially and predictions are made for every point at which prediction is desired; see Figure 3 (b). In this study, modifications have been made to the original counterpropagation interpolation phase, which allow the ANN to average (based on the inverse distance) the outputs for multiple hidden nodes, producing a smoother estimate of groundwater contamination.

For the incorporation of secondary data, two counterpropagation ANNs are implemented in series. The first ANN is used to estimate apparent conductivity at the 17 locations where PC1 is known (wells). This ANN uses inputs of *Easting*, *Northing* and *PC1* to map associated apparent conductivity. The estimates of apparent conductivity mapped to all 17 well locations are then added to the apparent conductivity dataset and used to train the second ANN. The second ANN uses only inputs of *Easting* and *Northing*, see Figure 3 (a) and (b), to map apparent conductivity everywhere within the desired estimation domain.

4. RESULTS AND DISCUSSION

A geostatistical analysis of the four apparent conductivity datasets found the spatial structure of apparent conductivity measurements to exhibit a semi-variogram spherical model with a range of 750-850 ft and sills ranging from 575,000 to 1,250,000 and nuggets of 50,000, see Figure 2. These geostatistical parameters of spatial structure are used to estimate apparent conductivity using ordinary kriging, see Figure 4 (a). The ordinary kriging technique produces a smoothed parameter estimation field of apparent conductivity.

Figure 4 (b) shows the estimated apparent conductivity field using the sequential counterpropagation ANNs to combine hydrochemistry data (in the form of PC1) and apparent conductivity survey data. Both parameter fields respect the field data at the observation points to a root mean square error value of 10^{-6}. Large amounts of data near the landfill result in detailed estimates that mimic the kriging estimates. The blocky affect witnessed in the counterpropagation ANN estimation field is due to the inverse distance averaging of multiple (in this case 3) hidden node outputs. Such blocky or banded estimates are prevalent in regions where data is less abundant.

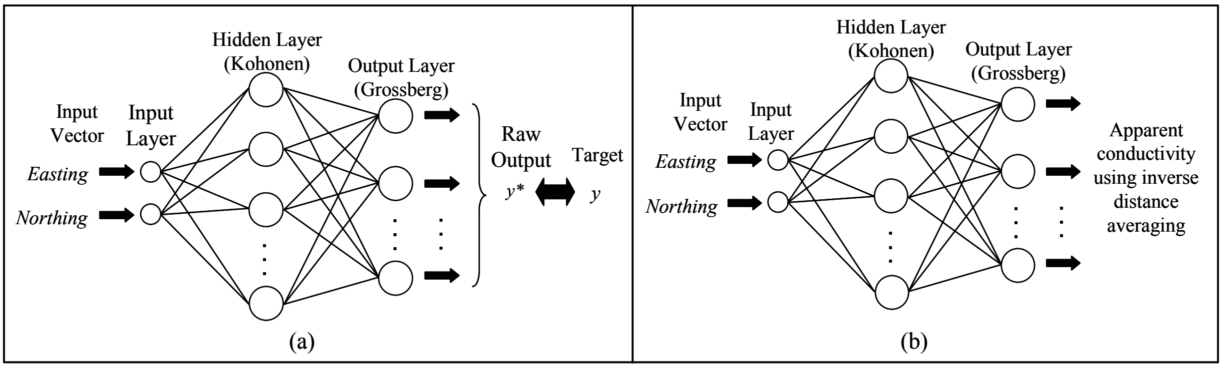

Figure 3. Architecture of counterpropagation ANN during the (a) training phase and (b) interpolation phase.

To determine whether or not the parameter fields are statistically different, both fields have been examined for measures of central tendency (means/medians) and dispersion (variance). The descriptive statistics for the kriged and ANN fields are summarized with means of 422.0 and 437.6 uS/cm, medians of 193.0 and 210.2 uS/cm and standard deviations of 614.7 and 637.3 uS^2/cm^2 respectively. Statistical parametric tests (a Two Sample t-test and analysis of variance) and non-parametric tests (Median test and O'Brien test) have been used to determine that the two parameter fields are not statistically different with respect to measures of central tendency (means/medians) and dispersion (variance). For these statistics a type I error rate of α=0.05 has been used to establish the level of significance. Without knowledge of the true apparent conductivity field, these statistics are the best method to quantitatively compare the two estimation techniques. Additional studies and comparisons of parameter estimation techniques have been performed on small slabs of Berea sandstone for which "reality" (air permeability) was considered known. However, due to page constraints, the results of these studies are not presented.

The magnitude of effort involved in the implementation of the two methods is an important consideration in this study. The computational demands and tasks associated with the kriging (and cokriging) methods (*i.e.* transforming data to accommodate different scales of measurements, generating numerous auto- and cross-semi-variograms, fitting appropriate variogram models and satisfying the criterion of coregionalization) are not required with the ANN. The data-driven nature of the ANN enables a non-linear mapping of the statistical relationships between the multiple variables without the need to model the parameter's spatial structure *a priori*. Research is currently underway to incorporate anisotropy and the uncertainty of parameter estimates into the network to better account for layering generally observed in the geologic setting. Increasing the number ANN inputs and/or running ANNs in series enables the incorporation of more predictor variables. Whereas, adding predictor variables to the kriging estimation technique, requires greater efforts to model auto- and cross semi-variograms.

This research demonstrates the potential for implementing this ANN estimation technique (as an alternative to kriging) to delineate the leachate contaminated groundwater and assess water quality associated with subsurface contamination at a full scale site. These results suggest the counterpropagation ANN is a promising parameter estimation method when incorporating multiple data types to enhance prediction accuracy and reduce uncertainty.

Future research will incorporate the hydrochemistry data using the method of ordinary cokriging to delineate the leachate plume. Ordinary cokriging is an estimation technique (similar to ordinary kriging) that exploits the auto- and cross-semivariance relationships between secondary and primary variables to produce better estimates (compared to ordinary kriging). However, the computational demands associated with cokriging (previously listed) were not achieved for this publication and, as a result the field of apparent conductivity was generated using the method of ordinary kriging.

5. ACKNOWLEDGEMENTS

The authors wish to thank Bernie Nadeau, Craig Squires Joe Gay and Dale Becker for providing access to Schuyler Falls landfill data, without which, this study would not have been possible. This research was funded in part by a Vermont NSF-EPSCoR Graduate Research Assistantship.

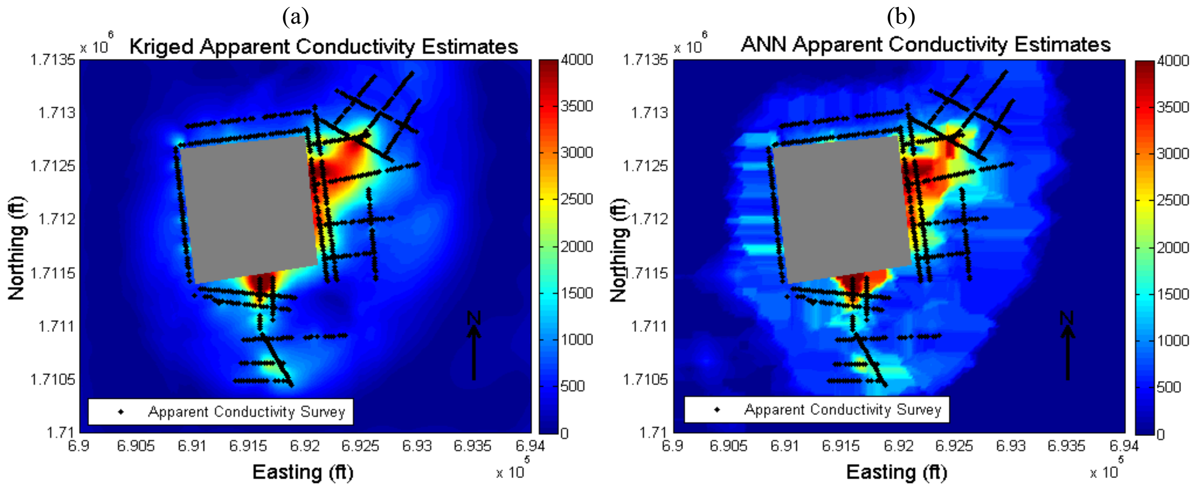

Figure 4. Estimation field of apparent conductivity (uS/cm) produced by (a) ordinary kriging and (b) sequential counterpropagation ANNs.

6. REFERENCES

Barton & Loguidice, P.C., Hydrogeologic Investigation Report, Schuyler Falls Sanitary Landfill Expansion, Syracuse, New York, 1993.

Barton & Loguidice, P.C., Clinton County Landfill Closure, Final Closure Investigation Report, Syracuse New York, 1996.

Deutsch, C.V., and Journel, A.G, *GSLIB, Geostatistical Software Library and User's Guide*, Oxford Press, 369 pp., New York, 1998

Dunn Science Corporation, Groundwater Contamination at the Schuyler Falls Sanitary Landfill, Latham, New York, 1986.

Fidencio, P.H. Ruisanchez, I., and Poppi, R.J., Application of artificial neural networks to the classification of soils from Sao Paulo state using near-infrared spectroscopy, *Analyst*, 126(12)4979-4984, 2001.

Hecht-Nielsen, R., Counterpropagation networks, *Applied Optics*, 26(23), 4979-4984, 1987.

Isaaks, E.H., and Srivastava, R.M., *Introduction to Applied Geostatistics,* Oxford Press, 561 pp., New York, 1989.

Mouser, P.J., Rizzo, D.M., and Besaw, L., Improving site characterization and classifying attenuation processes using microbiological profiles, geochemistry, and artificial neural networks from landfill-leachate contaminated groundwater, *ASCE 2005 World Water & Environmental Resources Congress*, Anchorage, Alaska, 2005.

Nie, J.H., Loh, A.P., and Hang, C.C., Modeling pH neutralization processes using fuzzy-neural approaches, *Fuzzy Sets and Systems*, 78(1), 5-22, 1996.

Rizzo, D.M., and Dougherty, D.E., Characterization of aquifer properties using artificial neural networks: Neural kriging, *Water Resources Research*, 30(2) 483-497, 1994.

Rizzo, D.M., Lillys, T.P., and Dougherty, D.E., Comparisons of site characterization methods using mixed data, *Uncertainty in Geologic Environment*, Madison, Wisconsin, 167-179, 1996.

Rumelhart, D.E., and McClelland, J.L., *Parallel Distributed Processing*, Massachusetts Institute of Technology Press, Massachusetts, 547 pp., 1988.

Sullivan, P.A., Hierarchical Artificial Neural Networks for Regionalized Cokriging, M.S. Thesis, University of Vermont, 1999.

Sullivan, P.A., Rizzo, D.M., and Dougherty, D.E., Hierarchical neural networks for regionalized cokriging, *Proc. XII International Conference on Computational Methods in Water Resources*, Crete, 1998.

USEPA, National Water Quality Inventory, 2000 Report, EPA-841-R-02-001, Office of Water, Washington DC, 2002