USING ARTIFICIAL NEURAL NETWORKS TO ASSESS MICROBIAL COMMUNITIES

Susan M. Pfiffner, (University of Tennessee, Knoxville, Tennessee)
Craig C. Brandt, Jack C. Schryver, and Anthony V. Palumbo, (Oak Ridge National Laboratory, Oak Ridge, Tennessee)
Jonas S. Almeida (New University of Lisbon, Oeiras, Portugal)

ABSTRACT: We are evaluating artificial neural networks (ANNs) as tools for assessing changes in soil microbial communities following exposure to metals. We analyzed signature lipid biomarker data collected from two soil microcosm experiments using an autoassociative ANN. In one experiment, the microcosms were exposed to 0, 100, or 250 ppm of metals, and in the other experiment the microcosms were exposed to 0 or 500 ppm of metals. The ANNs were able to distinguish between microcosms exposed and not exposed to metals in both experiments.

INTRODUCTION

A major challenge in the implementation of in situ bioremediation technologies is understanding the indigenous microbial community structure and how this structure is affected by environmental conditions. Understanding the microbial community structure and the environmental factors that control it would assist in optimizing those factors that offer the best opportunity for control of bioremediation. For example, assessing the changes in the community structure following the addition of nutrient amendments is critical in monitoring the effectiveness of bioremediation.

Microbial communities in soils have often been characterized by the analysis of signature lipid biomarkers (SLBs) (e.g., Tunlid and White, 1992). SLBs are known to respond to changes in environment conditions such as temperature and nutrients. However, these changes are often complex, nonlinear, and not readily amenable to traditional statistical analyses.

We are using artificial neural networks (ANNs) to investigate the complex relationships among SLBs. ANNs are nonlinear, nonparametric analysis methods that can learn from experience to improve their performance. These methods are tolerant of missing or noisy data and can be used for rapid analysis of new data sets.

An ANN consists of a large number of simple processing elements called nodes. Nodes are connected to other nodes by means of directed communication links, each with an associated weight. Each node has an internal state, called its activation level, which is a function of the inputs it receives. A node sends its activation as a signal to the other nodes with which it is connected (Fausett, 1994). Figure 1 illustrates the architecture of a simple ANN consisting of five input parameters, eight nodes, and five output parameters.

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

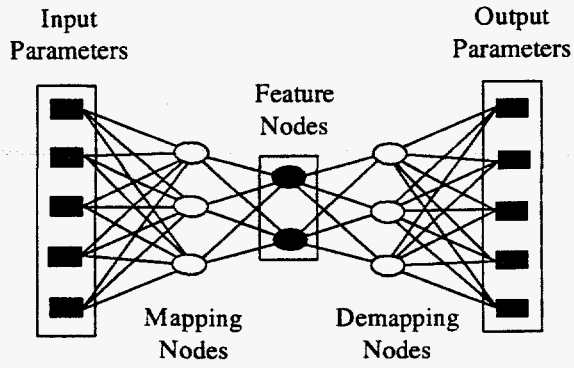

FIGURE 1. Example of a simple autoassociative ANN consisting of five input parameters, three mapping nodes, two feature nodes, three demapping nodes, and five output parameters. Solid lines represent directed communication links between components of the ANN.

The weights associated with the links in an ANN represent information used by the net to solve a problem. An ANN is trained by presenting it with a paired set of input and output patterns. The ANN adjusts its weights in order to minimize the difference between the output parameters calculated by the net and the actual output pattern.

MATERIAL AND METHODS

For this study, we analyzed SLB data obtained from two soil microcosm experiments that were designed to determine the effects of metal concentration of the microbial community structure. In one experiment, the microcosms were exposed to 0 and 500 ppm of a metal mixture that contained cadmium, strontium, cobalt, and cesium (Flemming et al., 1998). The other experiment used 0, 100, and 250 ppm concentrations of the same metal mixture that was used in the first experiment (Macnaughton et al., 1998). Triplicate microcosms from each metal treatment were sacrificed at 0, 7, 14, 28, and 56 days.

The sacrificed microcosms were analyzed for SLBs. A modified chloroform and methanol extraction, silicic acid column chromatography lipid class separation, alkaline methanolysis derivatization, and gas chromatography/mass spectrometry identification were used (White and Ringelberg, 1998). The resulting profiles contained 43 SLBs.

These SLBs from the second experiment were used to train an autoassociative ANN using the Stuttgart Neural Network Simulator (Zell et al., 1994). The architecture used to analyze the microcosm SLB data consisted of 43 input SLBs, four mapping nodes, two feature nodes, four demapping nodes, and 43 SLBs. To test for robustness the trained ANN was validated using the lipid parameters generated from the first experiment. The coordinates of the ANN's feature nodes were plotted and analyzed by hierarchical cluster analysis.

RESULTS

The ANN explained 85.7% of the variance in the data from the second microcosm experiment (Figure 2). The SLB profiles at day 7 (group 2) appear to be distinctly different from those at day 0 (group 1) regardless of the metal concentration in the microcosm. After day 7, the microcosms separated into two groups based on whether the microcosm was exposed (group 4) or not exposed (group 3) to metals. Within group 4, it appears that microcosms exposed to 250 ppm of metals are beginning to diverge from the microcosms treated with 100 ppm of metals.

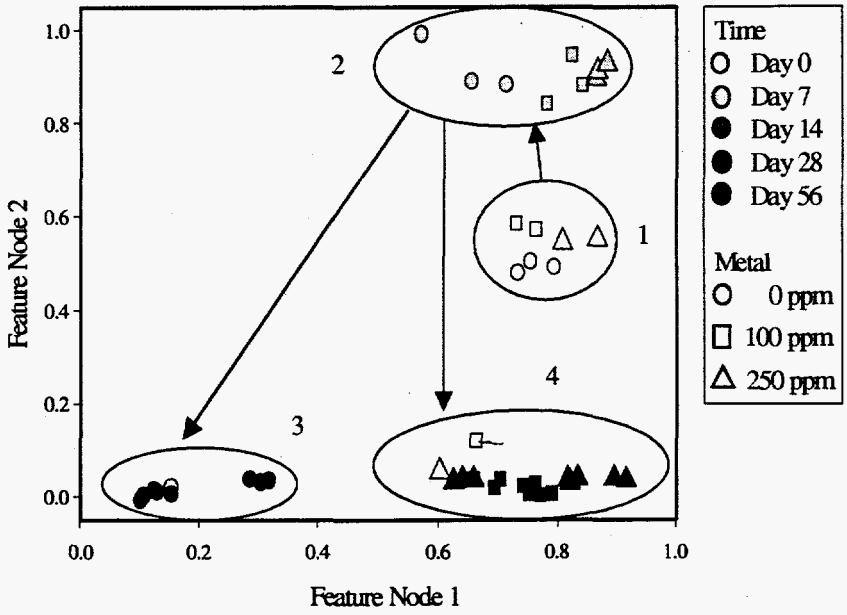

FIGURE 2. Plot of feature node coordinates from an autoassociative ANN analysis of the second (0, 100, and 250 ppm metal) microcosm experiment. Groupings were determined by a hierarchical cluster analysis of the node coordinates.

The SLB data from the first (0 and 500 ppm) experiment was used to test how well the ANN developed from the second (0, 100, and 250 ppm) experiment generalized. Figure 3 shows the results of the ANN analysis of SLB data from the first experiment using the network developed for the second experiment. The SLB profiles for the microcosms exposed to 500 ppm of metals are generally distinct from the SLB profiles for the unexposed (0 ppm) microcosms (group 1 versus group 2). Thus, the feature node plot from the first experiment was consistent with the feature plot from the second experiment.

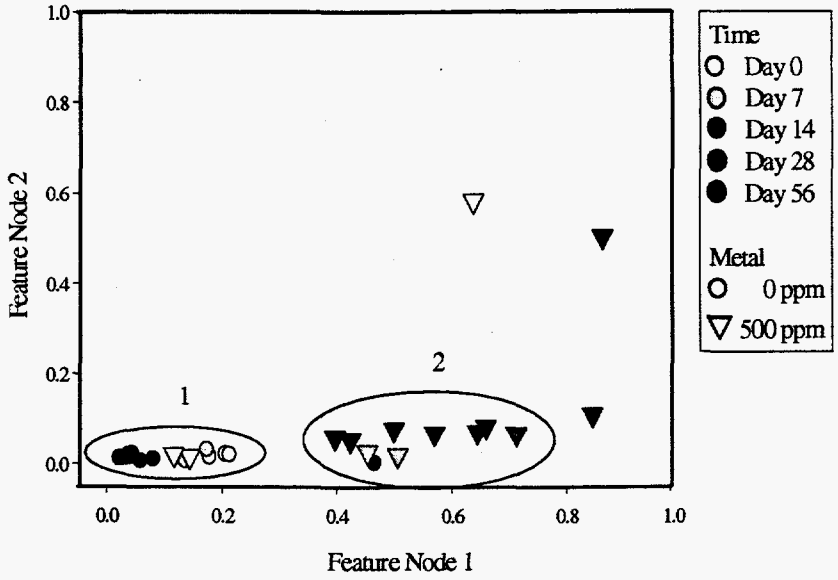

FIGURE 3. Plot of feature node coordinates from an autoassociative ANN analysis of the first (0 and 500 ppm metal) microcosm experiment using the ANN developed for the second (0, 100, and 250 ppm metal) experiment. Groupings were determined by a hierarchical cluster analysis of the node coordinates.

DISCUSSION

The microbial SLB profiles from the second experiment (0, 100, and 250 ppm metals) changed with time and exposure to metals. The differences found between the microcosms sacrificed at 0 and 7 days may be due to the microbial communities equilibrating to the experimental conditions. This effect, known as the disturbance artifact has been documented in other studies (Findlay et al., 1985). After 7 days, the ANN distinguished other temporal changes in the microbial community that appear to be related to metal exposure.

The ANN developed for the second experiment generalized well when applied to the data from the first experiment (0 and 500 ppm metals). There does not appear to be a disturbance artifact in this experiment. The ANN generally separated the microcosms exposed to metal from those that were not exposed. This corresponds to the trend seen in the second experiment.

We believe that ANNs offer a more robust approach for analyzing and interpreting complex data sets, and generalizing to new data sets. ANNs can be used as a tool for assessing microbial community structure changes as a result of pollution and is expected to be of particular relevance to the assessment of natural and accelerated bioremediation in soil.

ACKNOWLEDGMENTS

This research is supported by the U.S. Department of Energy Natural and Accelerated Bioremediation Research (NABIR) program. We would like to thank NABIR Science Team Leader, D.C. White, and personnel at the Center for Environmental Biotechnology at the University of Tennessee who designed, performed and analyzed the microcosm studies and provide the SLB data for this ANN research. Oak Ridge National Laboratory is managed by Lockheed Martin Energy Research, for the U.S. Department of Energy under contract DE-AC05-96OR22464.

REFERENCES

Fausett, L. V. 1994. *Fundamentals of Neural Networks.* Prentice Hall, NJ.

Findlay, R. V., P. C. Pollard, D. J. W. Moriarty, and D. C. White. 1985. Quantitative determination of microbial activity in community nutritional status in estuarine sediments: Evidence for a disturbance artifact." *Cand. J. Microbiol.* 31: 493-498.

Flemming, C. A., K. T. Leung, A. Peacock, S. Macnaughton, and D. C. White. 1999. "Changes in microbial community lipid, DNA and RNA biomarkers in soil contaminated with cadmium, cobalt, cesium, and strontium." (in preparation).

Macnaughton, S., J. R. Stephen, Y. J. Chang, A. Peacock, C. A. Flemming, K. T. Leung, and D. C. White. 1998. "Characterization of metal resistant soil eubacteria by PCR DGGE with isolation of resistant strains." *Cand. J. Microbiol* (submitted).

Pfiffner, S. M., A. V. Palumbo, T. Gibson, D. B. Ringelberg, and J. F. McCarthy. 1997. "Relating water and sediment chemistry to microbial characterization at a BTEX-contaminated site." *Appl. Biochem. and Biotech.* 63-65: 775-788.

Tunlid, A., and D. C. White. 1992. "Biochemical analysis of biomass, community structure, nutritional status, and metabolic activity of microbial communities in soils." In G. Stotzky and J. M. Bollag (Eds.) *Soil Biochemistry,* pp. 229-262. Marcel Dekker, Inc., New York, NY.

White, D. C. and D. B. Ringelberg. 1998. "Signature Lipid Biomarker Analysis." In R. S. Burlage, R. Atlas, D. Stahl, G. Geesey, and G. Sayler (Eds.), *Techniques in Microbial Ecology,* pp. 255-272. Oxford University Press, Inc., New York, NY.

Zell, A., N. Mache, R. Hubner, G. Mamier, M. Vogt, M. Schmalzl, and K. U. Herrmann. 1994. "SNNS (Stuttgart Neural Network Simulator)." In J. Skrzypek (Ed.), *Neural Network Simulation Environments,* pp. 165-186. Kluwer Academic Publishers, Assinippi Park, MA.