한국지하수토양환경학회 추계학술발표회
2004년 9월 9~10일 전 주 대 학 교

Artificial neural network application to solute transport through unsaturated zone

Hee-Sung Yoon, Kang-Kun Lee

School of Earth and Environmental Sciences, Seoul National University (oolahee1@snu.ac.kr)

Abstract

The unsaturated zone is a significant pathway of the surface contaminant movement and is a highly heterogeneous medium. Therefore, there are limitations in applying conventional convection-dispersion equation(CDE). Artificial neural network(ANN) is considered to be a versatile tool for approximating complex functions. For evaluating the applicability of ANN, numerical tests using ANN were conducted with training set generated by HYDRUS-2D which is based on CDE. The results represent that ANN can estimate the solute transport and the choice of network parameters and generation of training set patterns are important for efficient estimation.

key words : unsaturated zone, convection-dispersion equation, artificial neural network

1. Introduction

The unsaturated zone, or vadose zone, expands from land surface to groundwater table, and it contains the capillary fringe (Fetter, 1999). Groundwater may begin to be polluted as the surface contaminants infiltrate into subsurface and reach the groundwater table. Thus the flow path through the unsaturated zone is the significant pathway of surface contaminants. The unsaturated zone is a highly heterogeneous medium, however, there have been attempts to describe the mechanisms of water and solute movements through it. The conventional convection-dispersion equation(CDE) for solute transport in vadose zone is based on a set of assumptions about dispersion process, so that it may be valid under certain conditions in soil (Jury et al., 1991). Therefore it is difficult to describe the solute transport using CDE in the situations, where the thickness of the unsaturated zone is not large enough to satisfy the CDE's assumptions. Because of these limitations, alternative approaches, using black box model, has been studied. The transfer function model represents the solute transport in soil without explicit description of physical transport process, using a travel time probability density function (Skaggs et al., 1998). But the transfer function model is based on the linearity, and the travel time pdf should be fitted to a specified distribution function such as log-normal distribution function.

Recently, researches of artificial neural network(ANN) application in groundwater flow and transport have been emerged. ANN is a flexible mathematical structure patterned after the biological functioning of the nervous system. In general, ANN is considered to be a versatile tool for approximating complex functions that are difficult to model mathematically or evaluate numerically (Morshed et al., 1998).

2. Methodology

2.1 Feedforward Network Architecture

Feedforward network is composed of input layer, hidden layers and output layer. Each layer has its own nodes. The number of hidden layers and nodes in hidden layers can only be determined by trial and error. Figure 1 shows a feedforward network with one hidden layer.

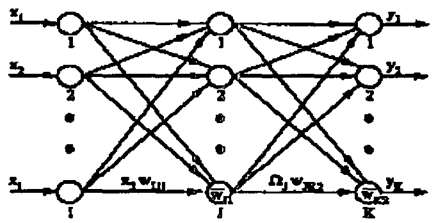


Figure 1. A feedforward network

2.2 Back-propagation Algorithm

Back-propagation algorithm (BPA) is a gradient based algorithm, and the weights at each node is updated using gradient of error by weights. BPA proceeds as follows ;

1) compute hidden node inputs -> hidden node outputs -> inputs to hidden nodes
2) compute the network outputs
3) compute error between network output and target value
4) Modify the weights between hidden and output nodes
5) Modify the weights between input and hidden nodes
6) repeat 1) ~ 5) until all training patterns are presented
7) repeat 6) until error is satisfactory(Mehrotra et al, 1997)

2.3 Training and Testing Set Generation

ANN is an alternative method for better estimation than CDE. However, for evaluating applicability of ANN to solute transport through unsaturated zone, numerical tests was conducted using CDE based simulations. Training and testing set for solute transport through unsaturated zone is obtained using HYDRUS-2D. HYDRUS-2D is the software package for simulating water flow and solute transport in variably saturated media.

Simulation domain was sandy unsaturated zone of 1 meter depth with various initial water contents and water flux at the surface. Conservative solute injection of 1000 ppm concentration was assumed and injection duration was 1 day.

3. Results and Discussions

3.1 Test I

Inputs in Test I were simulation time, constant water flux at the surface and initial water content.

Total 1250 training sets are obtained by HYDRUS-2D. Table 1 shows the inputs and their number and range.

Table 1. Information of inputs in Test I

Inputs	Number	Range
Time	50	0 ~ 30 (days)
Water flux	5	3 ~ 15 (cm/day)
Wate content	5	0.06 ~ 0.35

The number of nodes in hidden layer and learning rate were determined as 11 and 0.03 respectively by trial and error. After training, computed RMSE was 3.4×10^{-3}.

Figure 2 shows the testing results. The results were on the whole satisfactory except the front part of the concentration curve, where the solute was not detected. x axis represents the simulation time by day and y axis represents the normalized concentration (C/Co).

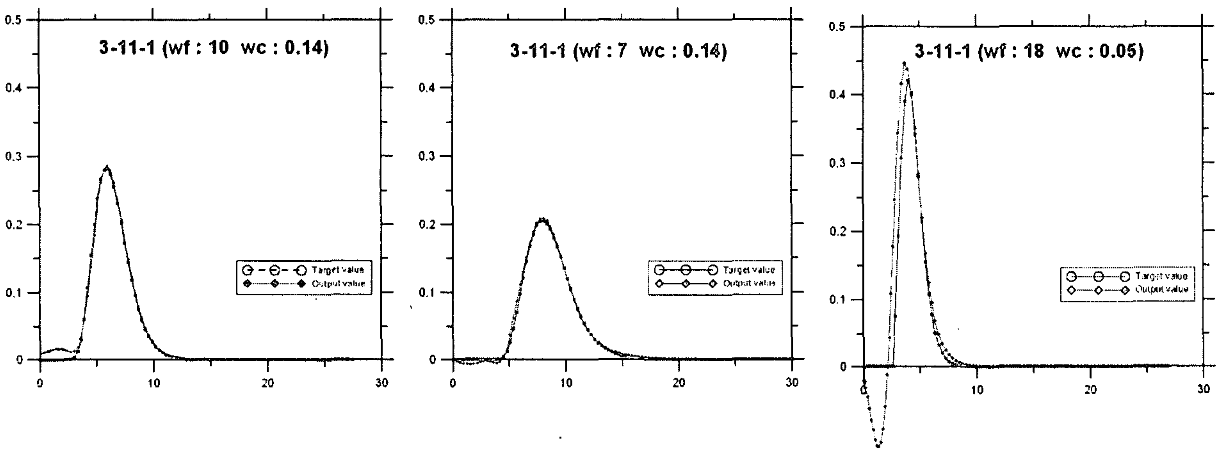


Figure 2. Testing results (Test I)

3.2 Test II

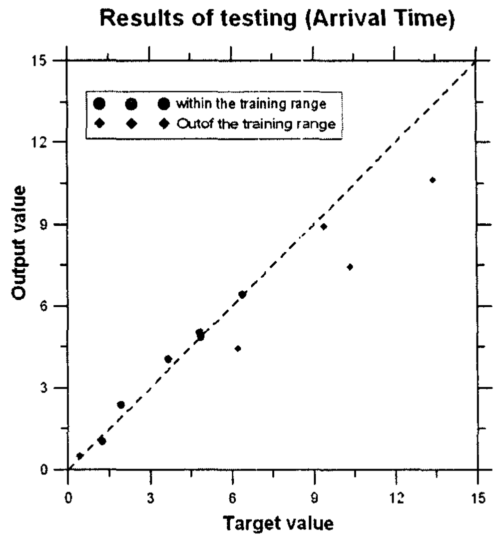
For raising the degree of efficiency in testing, training procedure was divided into 2 sets. The one was for arrival time of the solute, and the other was for the normalized cumulative mass curve starting at the solute arrival time. Inputs in Test II were same as in Test I. Total 25 training sets were obtained for arrival time, and 1025 for mass curve. Input number and range were presented at Table 2.

Table 2. Information of inputs in Test II

Inputs	Number	Range
Time	41	0 ~ 40 (days)
Water flux	5	1.5 ~ 4 (cm/day)
Water content	5	0.08 ~ 0.35

3.2.1 *Solute arrival time*

For solute arrival time training the number of nodes in hidden layer and learning rate were determined as 11 and 0.01, respectively. Testing results of which the range of input values are within the training's, showed good agreement between target and output values. Figure 3 represents the results of testing.



Fugure 3. Testing results of arrival time (Test II)

3.2.2 Normalized cumulative mass curve

For mass curve training the number of nodes in hidden layer and learning rate were determined as 13 and 0.005, respectively. Calculated training RMSE was 4.4×10^{-3}. Training RMSE was slightly higher than Test I, however, in testing results, there was not large error at the front of the curve and the shape of curve fitted well. Figure 4 shows the results of testing. x axis represents the simulation time by day and y axis normalized cumulative mass.

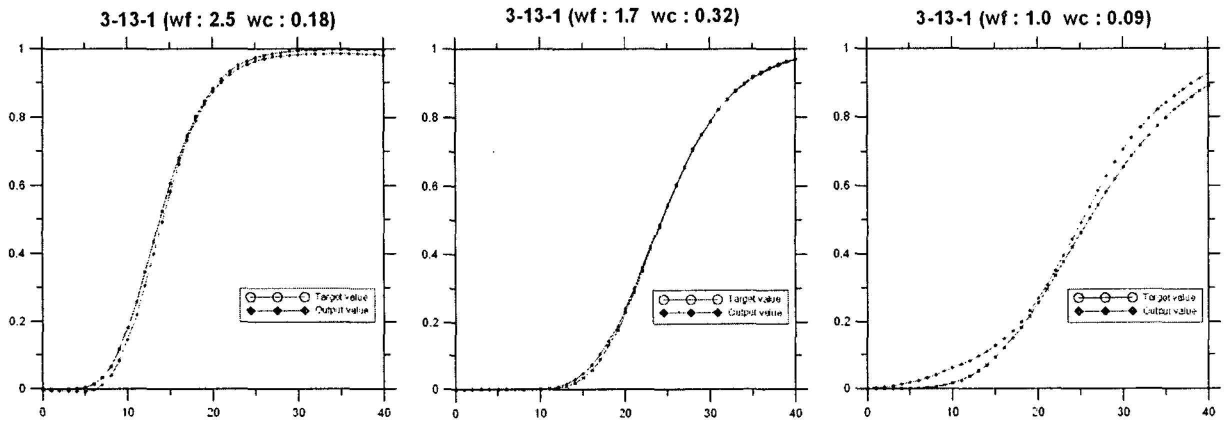


Figure 4. Testing results of mass curve (Test II)

These results show that solute transport through unsaturated zone can be estimated efficiently by using artificial neural network. It is important to determine the node numbers, learning rate and so on, and make training set in order that the network is trained effectively. If sufficient training sets are obtained, it is possible that artificial neural network application to laboratory and field tests can make better estimation than CDE based simulation.

References

1. Fetter, C.W., 1999, Contaminant hydrogeology, Second edition, Prentice Hall, pp170-207
2. Jury, William A., Gardner, Wilford R., and Gardner, Walter H., 1991, Soil Physics, John Wiley & Sons, Inc, pp 243-254

3. Skaggs, Todd H., Kabala, Z.J., and Jury, William A., 1998, Deconvolution of a nonparametric transfer function for solute transport in soils, Journal of Hydrology, 207, 170-178
4. Morshed, Jahangir & Kaluarachchi, Jagath J. , 1998, Application of artificial neural network and genetic algorithm in flow and trasport simulations, Advances in Water Resources Vol. 22, No.2 pp. 145-158.
5. Mehrotra, Kishan , Mohan, Chilukuri K. and Ranka, Sanjay, 1997, Elements of Artificial Neural Networks, Messachusetts Institute of Technology, pp 65-88

Acknowledgement

This research was supported by a grant from Sustainable Water Resource Research Center of 21th Century Frontier Research Program and AEBRC at POSTECH.