# Predicting Longitudinal Dispersion Coefficient in Natural Streams by Artificial Neural Network

Gokmen Tayfur[1] and Vijay P. Singh, F.ASCE[2]

**Abstract:** An artificial neural network (ANN) model was developed to predict the longitudinal dispersion coefficient in natural streams and rivers. The hydraulic variables [flow discharge ($Q$), flow depth ($H$), flow velocity ($U$), shear velocity ($u*$), and relative shear velocity ($U/u*$)] and geometric characteristics [channel width ($B$), channel sinuosity ($\sigma$), and channel shape parameter ($\beta$)] constituted inputs to the ANN model, whereas the dispersion coefficient ($K_x$) was the target model output. The model was trained and tested using 71 data sets of hydraulic and geometric parameters and dispersion coefficients measured on 29 streams and rivers in the United States. The training of the ANN model was accomplished with an explained variance of 90% of the dispersion coefficient. The dispersion coefficient values predicted by the ANN model satisfactorily compared with the measured values corresponding to different hydraulic and geometric characteristics. The predicted values were also compared with those predicted using several equations that have been suggested in the literature and it was found that the ANN model was superior in predicting the dispersion coefficient. The results of sensitivity analysis indicated that the $Q$ data alone would be sufficient for predicting more frequently occurring low values of the dispersion coefficient ($K_x < 100$ m$^2$/s). For narrower channels ($B/H < 50$) using only $U/u*$ data would be sufficient to predict the coefficient. If $\beta$ and $\sigma$ were used along with the flow variables, the prediction capability of the ANN model would be significantly improved.

**CE Database subject headings:** Diffusion; Coefficients; Neural networks; River flow; Streamflow; Simulation; Physical properties.

## Introduction

Hazardous contaminants and effluent, when accidentally discharged into a river, undergo stages of mixing as they are transported downstream by the flowing water. The effluent is dispersed transversely, vertically, and longitudinally by advective and dispersive processes. Once the cross-sectional mixing is complete, the process of longitudinal dispersion becomes the most important mechanism (Seo and Cheong 1998). The intensity of longitudinal dispersion is measured by the longitudinal dispersion coefficient (Deng et al. 2002). Hence, the transport process and the consequent fate of pollutants are described, to a large extent, by the longitudinal dispersion coefficient. For that reason, the dispersion coefficient has been extensively investigated (Elder 1959; Sooky 1969; Fukuoko and Sayre 1973; McQuivey and Keefer 1976; Fischer et al. 1979; Sukhodolov et al. 1997; Deng et al. 2001; Seo and Baek 2004). Taylor (1953, 1954) first introduced the longitudinal dispersion coefficient as a measure of the one-dimensional dispersion for which Fischer (1967) developed the following integral expression:

[1] Professor, Dept. of Civil Engineering, Faculty of Engineering, Izmir Institute of Technology, Gulbahcekoyu, Urla, Izmir 35340, Turkey. E-mail: gokmentayfur@iyte.edu.tr.

[2] A. K. Barton Professor, Dept. of Civil and Environmental Engineering, Louisiana State Univ., Baton Rouge, LA 70803-6405. E-mail: cesing@lsu.edu.

$$K_x = -\frac{1}{A}\int_0^B hu' \int_0^y \frac{1}{\varepsilon_t h}\int_0^y hu' dy dy dy \qquad (1)$$

where $K_x$=longitudinal dispersion coefficient; $A$=cross-sectional area; $B$=channel width; $h$=local flow depth; $u'$=deviation of local depth mean flow velocity from cross-sectional mean; $y$=coordinate in the lateral direction; and $\varepsilon_t$=local (depth averaged) transverse mixing coefficient. The form of Eq. (1) is commonly reported in the literature in the above-mentioned form (Rutherford 1994; Seo and Cheong 1998; Deng et al. 2001; Kashefipour and Falconer 2002; Jirka 2004). The form of Eq. (1) given in Fischer et al. (1979) is slightly different, but it should not be difficult to see the similarities between the two forms.

Owing to the requirement for detailed transverse profiles of velocity and cross-sectional geometry, it is rather difficult to use Eq. (1). As a result, a number of investigators have proposed empirical equations based on experimental and field data for predicting the dispersion coefficient (McQuivey and Keefer 1974; Fischer 1975; Liu 1977; Magazine et al. 1988; Iwasa and Aya 1991; Seo and Cheong 1998; Kashefipour and Falconer 2002). Fischer (1975) developed the following simple method to predict the longitudinal dispersion coefficient, which is a simplified nonintegral form of Eq. (1):

$$K_x = 0.011\frac{U^2 B^2}{Hu_*} \qquad (2)$$

where $H$=cross-sectional average flow depth; $u*$=shear velocity, and $U$=cross-sectional average flow velocity. Eq. (2) has the advantage of simplicity in that it can predict the dispersion coefficient by using only the data of cross-sectional mean parameters.

Employing 59 hydraulic and geometric data sets measured in 26 rivers in the United States, Seo and Cheong (1998) used dimensional analysis and applied the one-step Huber method, which

is a nonlinear multiregression method, to derive the following equation to predict the dispersion coefficient in natural streams:

$$K_x = 5.915(Hu_*)\left(\frac{B}{H}\right)^{0.62}\left(\frac{U}{u_*}\right)^{1.428} \quad (3)$$

Using readily available hydraulic variables of $B/H$ and $U/u_*$, the dimensionless dispersion coefficient $K_x/Hu_*$ can be estimated from Eq. (3). Seo and Cheong (1998) used 35 of the 59 measured data sets for establishing Eq. (3) and the remaining 24 for verifying their model. They compared their model with those of previously developed models of Elder (1959), McQuivey and Keefer (1974), Liu (1977), Magazine et al. (1988), and Iwasa and Aya (1991) and concluded that their model more accurately predicted the dispersion coefficient in natural streams.

Deng et al. (2001) developed a more theoretically based approximation of Eq. (1). They first developed mathematical expressions for the lateral distribution of river flow depth ($h$), the deviation of the velocity from the mean flow velocity ($u'$), and the local transverse mixing coefficient ($\varepsilon_t$). Then, they substituted the developed expressions for $h$, $u'$, and $\varepsilon_t$ in Eq. (1). Finally, by also considering the effects of nonuniformities to a limited extent, they developed the following equation for predicting the dispersion coefficient in straight channels:

$$K_x = 0.15\left(\frac{Hu_*}{8\varepsilon_{t0}}\right)\left(\frac{B}{H}\right)^{5/3}\left(\frac{U}{u_*}\right)^{2} \quad (4)$$

where

$$\varepsilon_{t0} = 0.145 + \left(\frac{1}{3,520.0}\right)\left(\frac{B}{H}\right)^{1.38}\left(\frac{U}{u_*}\right) \quad (5)$$

The model of Deng et al. (2001) expressed by Eq. (4) not only includes the conventional parameters of ($B/H$) and ($U/u*$) but also the effects of the transverse mixing $\varepsilon_{t0}$. In their development, however, they made several assumptions. In principle, their model is limited to straight and uniform rivers. That is the reason that they attributed the differences between the measured and predicted values to the effects of dead zones, bends, secondary currents, and other irregular features that are not explicitly involved in Eq. (4). Further, their model is for the rivers where the width-to-depth ratio is greater than 10.

Using 81 sets of measured data from 30 rivers in the United States, Kashefipour and Falconer (2002) developed an equation for predicting the longitudinal dispersion coefficient in riverine flows. This equation relates the dispersion coefficient to the hydraulic and geometric parameters and was derived by using dimensional and regression analyses, and can be expressed as

$$K_x = 10.612(HU)\left(\frac{U}{u_*}\right) \quad (6)$$

Most of the studies reported so far have been carried out based on specific assumptions and channel conditions and therefore the performance of the equations varies widely for the same stream and flow conditions. Therefore, a model that would have general applicability is needed. To that end, a three layer feedforward artificial neural network (ANN) model was developed and is proposed as an alternative to the existing theoretical and empirical models to predict the dispersion coefficient in natural streams. The proposed ANN model has an ability to capture the relation between the hydraulic and geometric (input) parameters and the dispersion coefficient (output) of the system without making any assumptions with regard to stream geometry or flow dynamics.
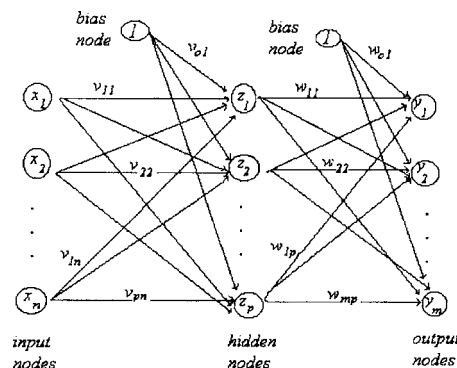


**Fig. 1.** Schematic representation of a three-layer feedforward artificial neural network

### Artificial Neural Networks

ANNs have an ability to capture relationships from given patterns and this ability has enabled ANNs to be employed in the solution of large-scale complex problems, such as pattern recognition, nonlinear modeling, classification, association, and control (ASCE Task Committee 2000). That is the reason that ANNs have, for over a decade, been employed in the solution of many hydraulic and hydrologic problems, such as rainfall-runoff modeling (Halff et al. 1993; Smith and Eli 1995; Tokar and Johnson 1999), streamflow forcasting (Markus et al. 1995; Chua and Holz 2005), sediment transport (Tayfur 2002; Sen et al. 2004), seepage (Tayfur et al. 2005), and groundwater flow (Aziz and Wong 1992; Morshed and Kaluarachchi 1998). The usefulness of ANNs has also been illustrated in predicting water quality parameters (Maier and Dandy 1996).

The common three layer-feedforward type of an artificial neural network is generally applied in practice. Fig. 1 shows a typical three-layer feedforward type ANN. In a feedforward network, the input quantities are fed into input layer neurons which, in turn, pass them on to the hidden layer neurons after multiplying by a weight. A hidden layer neuron adds up the weighted input received from each input neuron, associates it with a bias, and then passes the result on through a nonlinear transfer function (Fig. 2). The output neurons do the same operation as that of a hidden neuron.

Before applying to any problem, the network is first trained, whereby adjusting the weights and biases through some training algorithm minimizes the difference between target output and the calculated model output at each output neuron. Training of an ANN consists of three elements: (1) weights between neurons that define the relative importance of the inputs; (2) a transfer function that controls the generation of the output from a neuron; and (3) learning laws that describe how the adjustments of the weights are made during training.
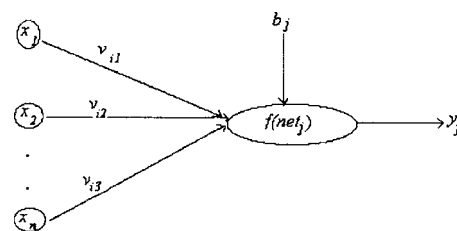


**Fig. 2.** Schematic of a node $j$

During training, a neuron receives inputs from previous layers, weighs each input with a prearranged value, and combines these weighted inputs. The combination of the weighted inputs is represented as (Tayfur 2002)

$$\text{net}_j = \sum x_i v_{ij} - b_j \tag{7}$$

where $net_j$=summation of the weighted input for the $j$th neuron; $x_i$=input from the $i$th neuron to the $j$th neuron; $v_{ij}$=weight from the $i$th neuron in the previous layer to the $j$th neuron in the current layer; and $b_j$=the threshold value, also called the bias, associated with node $j$. In Eq. (7), $\Sigma x_i v_{ij}$ should be greater than $b_j$ (i.e., $\Sigma x_i v_{ij} > b_j$) to produce a net information (net$_j$) for the activation function. Note that the use of bias is optional and may be applied either to the input signal or to the output signal.

The net$_j$ is passed through a transfer function to determine the level of activation. If the activation of a neuron is sufficiently strong, it produces an output that is sent as an input to other neurons in the successive layer. The sigmoid function is commonly employed as an activation function in the training of the network. The sigmoid function, which enables a network to map any nonlinear process and provides a graded nonlinear response, is a bounded, monotonic, nondecreasing function. One of the main reasons that the sigmoid function is employed is because of the simplicity of its derivative that is required during the training process (ASCE Task Committee 2000). The sigmoid function is expressed as (Tayfur 2002)

$$f(\text{net}_j) = \frac{1}{1 + e^{-\text{net}_j}} \tag{8}$$

The training of ANNs is widely accomplished by a backpropagation algorithm. Backpropagation is the most commonly used supervised training algorithm in the multilayered feedforward networks. In a backpropagation algorithm, information is processed in the forward direction from the input layer to the hidden layer and then to the output layer (Fig. 1). The objective of a backpropagation algorithm is, by minimizing a predetermined error function, to determine the optimal weights which would generate an output vector $Y=(y_1, y_2, \ldots, y_p)$ as close as possible to the target values of the output vector $T=(t_1, t_2, \ldots, t_p)$ with a selected accuracy. A form of the predetermined error function has the following form (ASCE Task Committee 2000)

$$E = \sum_p \sum_p (y_i - t_i)^2 \tag{9}$$

where $y_i$=component of an ANN output vector $Y$; $t_i$=component of a target output vector $T$; $p$=number of output neurons; and $P$=number of training patterns.

In the backpropagation algorithm, the least square error method along with a generalized delta rule is used to optimize the network weights. The gradient descent method, along with the chain rule of derivatives, is employed to modify network weights as (Tayfur 2002)

$$v_{ij}^{\text{new}} = v_{ij}^{\text{old}} - \delta \frac{\partial E}{\partial v_{ij}} \tag{10}$$

where $\delta$=learning rate which is used to increase the chance of avoiding the training process being trapped in a local minima instead of a global minima (ASCE Task Committee 2000).

As pointed out previously, a network learns by adjusting the biases and weights that link its neurons. However, before training can begin, a network's weights and biases must be set to small random values. Also, due to the nature of the sigmoid function

used in the backpropagation algorithm, it is prudent to standardize [i.e., convert to the range (0,1)] all external input and output values before passing them into the neural network. Without standardization, large values input into the ANN would require extremely small weighting factors to be applied and this can cause a number of problems (Dawson and Wilby 1998). There are no fixed rules as to which standardization approach should be used in a particular circumstance (Dawson and Wilby 1998). In this study, the following standardization method was used (Tayfur 2002):

$$z_i = \frac{x_i}{x_{\max} + 1} \tag{11}$$

where $z_i$=standardized value calculated for neuron $i$; $x_i$=real input or output value applied to neuron $i$; and $x_{\max}$=maximum input or output value of all values applied to neuron $i$.

More complete details of ANNs can be obtained from Somez (1998), Dawson and Wilby (1998), ASCE Task Committee (2000), amongst others.

### Experimental Data

Training and testing of the model was accomplished by employing 71 sets of measured data from 29 rivers in the United States (Table 1). The data sets were obtained from Deng et al. (2001). Tables 2 and 3 summarize the statistical information on the 71 measured data sets. The range for the dispersion coefficient ($K_x$) varies from 1.9 to 892 m$^2$/s (Table 3) and $K_x$ is greater than 100 m$^2$/s in 21 cases (Table 1), which represents about 30% of the total measured coefficient values (Table 2). The range for the width-to-depth ratio ($B/H$) of the data sets varies from 13.8 to 156.5 (Table 3) and $B/H$ is greater than 50 in 26 cases (Table 1), which is about 37% of the total number of 71 data sets (Table 2). The channel shape parameter ($\beta$), defined by Deng et al. (2001) in Eq. (12), varies from 2.62 to 5.05 (Table 3). According to Deng et al. (2001), this range corresponds to a natural channel shape with a flatbed region and two curving bank regions

$$\beta = \ln\left(\frac{B}{H}\right) \tag{12}$$

If $\beta=1.0$ then the channel has a triangular shape. Otherwise, it would have either a parabolic shape ($\beta=2$) or a channel shape with a flatbed region and two curving bank regions ($2<\beta<5$) or a rectangular shape ($\beta>5$) (Deng et al. 2001). The sinuosity ($\sigma$) of the streams and rivers varies from 1.08 to 2.54 (Table 3). The data of channel sinuosity in Table 1 were obtained from Deng et al. (2002). The channel sinuosity ($\sigma$) is defined as the ratio of the channel length to the valley length (Chang 1988, pp. 10–27 and 306–309). It can be calculated from 1:25,000 scale topographic maps based on specific stream reaches of the dye tests (Deng et al. 2002). For straight channels, the sinuosity is equal to unity.

Fifty one data sets of the 71 sets (72%) were used for training of the ANN model and the remaining sets (28%) were employed for testing the model. The data sets used for testing are marked by an asterisk (*) in Table 1. In choosing the data sets for training and testing, special attention was paid to have representative sets so as to avoid bias in model prediction. Tables 2 and 3 summarize statistical information for the training and testing data sets. Tables 2 and 3 show that the range and mean values of the variables are comparable for both data sets. Further, the percents of $B/H>50$ and $K_x>100$ m$^2$/s are also comparable for both sets (Table 2). In 25% of the test data sets, the measured value of the coefficient is greater than 100 m$^2$/s (this ratio is 31% in the training set) (Table

**Table 1.** Experimental Measurements of Longitudinal Dispersion Coefficient in Natural Streams

| Stream | $B$ (m) | $H$ (m) | $U$ (m/s) | $u*$ (m/s) | $B/H$ | $U/u*$ | $\beta$ | $\sigma$ | $K_x$ (m²/s) |
|---|---|---|---|---|---|---|---|---|---|
| Antietam Creek, Md. | 12.8 | 0.30 | 0.42 | 0.057 | 42.7 | 7.37 | 3.75 | 1.40 | 17.5 |
| Antietam Creek, Md. | 24.1 | 0.98 | 0.59 | 0.098 | 24.6 | 6.02 | 3.20 | 2.25 | 101.5 |
| Antietam Creek, Md.* | 11.9 | 0.66 | 0.43 | 0.085 | 18.0 | 5.06 | 2.89 | 2.25 | 20.9 |
| Antietam Creek, Md. | 21.0 | 0.48 | 0.62 | 0.069 | 43.8 | 8.99 | 3.78 | 1.26 | 25.9 |
| Monocacy River, Md.* | 48.7 | 0.55 | 0.26 | 0.052 | 88.5 | 5.00 | 4.48 | 1.28 | 37.8 |
| Monocacy River, Md.* | 93.0 | 0.71 | 0.16 | 0.046 | 131.0 | 3.48 | 4.88 | 1.28 | 41.4 |
| Monocacy River, Md. | 51.2 | 0.65 | 0.62 | 0.044 | 78.8 | 14.09 | 4.37 | 1.28 | 29.6 |
| Monocacy River, Md. | 97.5 | 1.15 | 0.32 | 0.058 | 84.8 | 5.52 | 4.44 | 1.61 | 119.8 |
| Monocacy River, Md. | 40.5 | 0.41 | 0.23 | 0.040 | 98.8 | 5.75 | 4.59 | 1.61 | 66.5 |
| Conococheague Creek, Md. | 42.2 | 0.69 | 0.23 | 0.064 | 61.2 | 3.59 | 4.11 | 2.25 | 40.8 |
| Conococheague Creek, Md. | 49.7 | 0.41 | 0.15 | 0.081 | 121.2 | 1.85 | 4.80 | 2.25 | 29.3 |
| Conococheague Creek, Md.* | 43.0 | 1.13 | 0.63 | 0.081 | 38.1 | 7.78 | 3.64 | 1.31 | 53.3 |
| Chattahoochee River, Ga.* | 75.6 | 1.95 | 0.74 | 0.138 | 38.8 | 5.36 | 3.66 | 1.27 | 88.9 |
| Chattahoochee River, Ga. | 91.9 | 2.44 | 0.52 | 0.094 | 37.7 | 5.53 | 3.63 | 1.57 | 166.9 |
| Salt Creek, Neb. | 32.0 | 0.50 | 0.24 | 0.038 | 64.0 | 6.32 | 4.16 | 1.38 | 52.2 |
| Diffcult Run, Va. | 14.5 | 0.31 | 0.25 | 0.062 | 46.8 | 4.03 | 3.85 | 1.09 | 1.9 |
| Bear Creek*, Colo. | 13.7 | 0.85 | 1.29 | 0.553 | 16.1 | 2.33 | 2.78 | 1.08 | 2.9 |
| Little Pincy Creek, Md. | 15.9 | 0.22 | 0.39 | 0.053 | 72.3 | 7.36 | 4.28 | 1.13 | 7.1 |
| Bayou Anacoco, La. | 17.5 | 0.45 | 0.32 | 0.024 | 38.9 | 13.33 | 3.66 | 1.41 | 5.8 |
| Bayou Anacoco, La. | 25.9 | 0.94 | 0.34 | 0.067 | 27.6 | 5.07 | 3.32 | 1.41 | 32.5 |
| Bayou Anacoco, La. | 36.6 | 0.91 | 0.40 | 0.067 | 40.2 | 5.97 | 3.69 | 1.41 | 39.5 |
| Comite River, La. | 15.7 | 0.23 | 0.36 | 0.039 | 68.3 | 9.23 | 4.22 | 1.31 | 69.0 |
| Bayou Bartholomew, La. | 33.4 | 1.40 | 0.20 | 0.031 | 23.9 | 6.45 | 3.17 | 2.46 | 54.7 |
| Tickfau River, La. | 15.0 | 0.59 | 0.27 | 0.080 | 25.4 | 3.38 | 3.23 | 1.75 | 10.3 |
| Tangipahoa River, La. | 31.4 | 0.81 | 0.48 | 0.072 | 38.8 | 6.67 | 3.66 | 1.46 | 45.1 |
| Tangipahoa River, La.* | 29.9 | 0.40 | 0.34 | 0.020 | 74.8 | 17.0 | 4.31 | 1.46 | 44.0 |
| Red River, La. | 253.6 | 1.62 | 0.61 | 0.032 | 156.5 | 19.06 | 5.05 | 1.20 | 143.8 |
| Red River, La. | 161.5 | 3.96 | 0.29 | 0.060 | 40.8 | 4.83 | 3.93 | 1.44 | 130.5 |
| Red River, La. | 152.4 | 3.66 | 0.45 | 0.057 | 41.6 | 7.89 | 3.73 | 1.44 | 227.6 |
| Red River, La. | 155.1 | 1.74 | 0.47 | 0.036 | 89.1 | 13.06 | 4.49 | 1.24 | 177.7 |
| Sabina River, La. | 116.4 | 1.65 | 0.58 | 0.054 | 70.5 | 10.74 | 4.26 | 1.19 | 131.3 |
| Sabina River, La.* | 160.3 | 2.32 | 1.06 | 0.054 | 69.1 | 19.63 | 4.24 | 1.17 | 308.9 |
| Sabina River*, Tex. | 14.2 | 0.50 | 0.13 | 0.037 | 28.4 | 3.51 | 3.35 | 2.53 | 12.8 |
| Sabina River*, Tex. | 12.2 | 0.51 | 0.23 | 0.030 | 23.9 | 7.67 | 3.17 | 2.05 | 14.7 |
| Sabina River*, Tex. | 21.3 | 0.93 | 0.36 | 0.035 | 22.9 | 10.29 | 3.13 | 1.47 | 24.2 |
| Mississippi River, La.* | 711.2 | 19.94 | 0.56 | 0.041 | 35.7 | 13.66 | 3.58 | 1.44 | 237.2 |
| Mississippi River, Mo.* | 533.4 | 4.94 | 1.05 | 0.069 | 108.0 | 15.22 | 4.68 | 1.38 | 457.7 |
| Mississippi River, Mo.* | 537.4 | 8.90 | 1.51 | 0.097 | 60.4 | 15.57 | 4.10 | 1.38 | 374.1 |
| Wind/Big. River, Wyo. | 44.2 | 1.37 | 0.99 | 0.142 | 32.3 | 6.97 | 3.48 | 1.56 | 184.6 |
| Wind/Big. River, Wyo. | 85.3 | 2.38 | 1.74 | 0.153 | 35.8 | 11.37 | 3.58 | 1.56 | 464.6 |
| Wind/Big. River, Wyo.* | 59.4 | 1.10 | 0.88 | 0.119 | 54.0 | 7.39 | 3.99 | 1.18 | 41.8 |
| Wind/Big. River, Wyo. | 68.6 | 2.16 | 1.55 | 0.168 | 31.8 | 9.23 | 3.46 | 1.18 | 162.6 |
| Copper Creep, Va. | 16.7 | 0.49 | 0.20 | 0.080 | 34.1 | 2.50 | 3.53 | 2.54 | 16.8 |
| Clinch River, Va. | 48.5 | 1.16 | 0.21 | 0.069 | 41.8 | 3.04 | 3.73 | 1.25 | 14.8 |
| Clinch River, Va.* | 28.7 | 0.61 | 0.35 | 0.069 | 47.0 | 5.07 | 3.85 | 1.14 | 10.7 |
| Clinch River, Va. | 57.9 | 2.45 | 0.75 | 0.104 | 23.6 | 7.21 | 3.16 | 1.14 | 40.5 |
| Clinch River, Va.* | 53.2 | 2.41 | 0.66 | 0.107 | 22.1 | 6.17 | 3.10 | 1.14 | 36.9 |
| Copper Creek, Va. | 18.3 | 0.38 | 0.15 | 0.116 | 48.2 | 1.29 | 3.88 | 2.54 | 20.7 |
| Copper Creek, Va. | 16.8 | 0.47 | 0.24 | 0.080 | 35.7 | 3.00 | 3.58 | 2.54 | 24.6 |
| Powell River, Tenn.* | 36.8 | 0.87 | 0.13 | 0.054 | 42.3 | 2.41 | 3.74 | 2.20 | 15.5 |
| Copper River, Va. | 19.6 | 0.84 | 0.49 | 0.101 | 23.3 | 4.85 | 3.15 | 1.26 | 20.8 |
| Nooksack River, Wash. | 64.0 | 0.76 | 0.67 | 0.268 | 84.2 | 2.50 | 4.43 | 1.30 | 34.8 |
| John Day River, Ore.* | 25.0 | 0.58 | 1.01 | 0.140 | 43.1 | 7.21 | 3.76 | 1.08 | 13.9 |
| John Day River, Ore.* | 34.1 | 2.47 | 0.82 | 0.180 | 13.8 | 4.56 | 2.62 | 1.89 | 65.0 |
| Yadkin River, N.C. | 70.1 | 2.35 | 0.43 | 0.101 | 29.8 | 4.26 | 3.39 | 2.17 | 111.5 |
| Yadkin River, N.C. | 71.6 | 3.84 | 0.76 | 0.128 | 18.6 | 5.94 | 2.92 | 2.17 | 260.1 |

**Table 1.** (*Continued.*)

| Stream | $B$ (m) | $H$ (m) | $U$ (m/s) | $u*$ (m/s) | $B/H$ | $U/u*$ | β | σ | $K_x$ (m²/s) |
|---|---|---|---|---|---|---|---|---|---|
| Minnesota River | 80.0 | 2.74 | 0.034 | 0.0024 | 29.2 | 14.17 | 3.37 | — | 22.3 |
| Minnesota River | 80.0 | 2.74 | 0.14 | 0.0097 | 29.2 | 14.43 | 3.37 | — | 34.9 |
| Amita River | 37.0 | 0.81 | 0.29 | 0.07 | 45.7 | 4.14 | 3.82 | — | 23.2 |
| Amita River | 42.0 | 0.80 | 0.42 | 0.069 | 52.5 | 6.09 | 3.96 | — | 30.2 |
| White River* | 67.0 | 0.59 | 0.35 | 0.044 | 113.6 | 7.95 | 4.73 | — | 30.2 |
| Nooksack River | 86.0 | 2.93 | 1.20 | 0.53 | 29.4 | 2.26 | 3.38 | 1.30 | 153.0 |
| Susquehanna River | 203.0 | 1.35 | 0.39 | 0.065 | 150.4 | 6.00 | 5.01 | 1.13 | 92.9 |
| Bayou Anacoco | 20.0 | 0.42 | 0.29 | 0.045 | 47.6 | 6.44 | 3.86 | 1.41 | 13.9 |
| Muddy River | 13.0 | 0.81 | 0.37 | 0.081 | 16.0 | 4.57 | 2.77 | — | 13.9 |
| Muddy River | 20.0 | 1.20 | 0.45 | 0.099 | 16.7 | 4.55 | 2.82 | — | 32.5 |
| Comite River | 13.0 | 0.26 | 0.31 | 0.044 | 50.0 | 7.05 | 3.91 | 1.31 | 7.0 |
| Comite River | 16.0 | 0.43 | 0.37 | 0.056 | 37.2 | 6.61 | 3.62 | 1.31 | 13.9 |
| Missouri River | 183.0 | 2.33 | 0.89 | 0.066 | 78.5 | 13.48 | 4.36 | 1.35 | 465.0 |
| Missouri River | 201.0 | 3.56 | 1.28 | 0.084 | 56.5 | 15.24 | 4.03 | 1.35 | 837.0 |
| Missouri River* | 197.0 | 3.11 | 1.53 | 0.078 | 63.3 | 19.62 | 4.15 | 1.35 | 892.0 |

Note: $B$=width; $H$=depth; $U$=velocity; $u*$=shear velocity; $U/u*$=relative shear velocity; β=shape parameter; σ=sinuosity; $K_x$=dispersion coefficient; and *=data sets used for testing of the ANN model.

2). In 40% of the test data sets, the width-to-depth ratio is greater than 50 (this ratio is 35% in the training set) (Table 2).

## Model Application

In this study, a three-layer feedforward artificial neural network model was constructed which had four neurons in the input layer, six neurons in the hidden layer, and one neuron in the output layer. The model was first trained and then verified.

### Training of the Artificial Neural Networks Model

For the number of neurons in the hidden layer, a trial-and-error procedure was used. The sigmoid function, defined by Eq. (8), was employed as an activation function in the training of the network and the backpropagation algorithm accomplished the training of the ANN model. Before starting the training process, random values of 0.2–0.4, and −1.0 were assigned for the network weights and biases, respectively. These assigned values are consistent with those used by Somez (1998). Also, before training and testing, all external input and output data were standardized using Eq. (11).

During training, 51 sets of channel width, depth, velocity, and shear velocity were input variables and the dispersion coefficient was the target output. The training of the network was accomplished with a 0.04 learning rate and after 20,000 iterations. Fig. 3 shows the measured dispersion data and the corresponding ANN model output data at the end of training stage. For a better illustration, predictions of three extreme values are not shown in Fig. 3. The model successfully predicted three extreme values of

837.0, 464.6, and 465.0 as 739.7, 467.9, and 313.2, respectively. Also in Fig. 3, the solid line is drawn with 45° to clearly show over- and underestimation. Although the data seem to be evenly distributed around the line, the number of overestimated values is slightly more than the number of underestimated values.

The coefficient of determination ($R^2$) between 51 measured and predicted values of the longitudinal diffusion coefficient data was 0.90 at the end of the training stage, implying that the ANN model was satisfactorily trained.

### Model Testing

The trained ANN model was then applied to predict the longitudinal dispersion coefficient for 20 measured data sets and its performance was compared with that of existing theoretical and empirical equations. Table 4 summarizes measured and predicted dispersion coefficient data. The predictions were made by the ANN model, the theoretical model of Deng et al. (2001) [Eq. (4)], and the empirical models of Fischer (1975) [Eq. (2)], Seo and Cheong (1998) [Eq. (3)], and Kashefipour and Falconer (2002) [Eq. (6)]. For the sake of brevity, the models of Deng et al. (2001), Fischer (1975), Seo and Cheong (1998), and Kashefipour and Falconer (2002) were designated as D-S-B, Fischer, S-C, and K-F models, respectively. As seen in Table 4, of the 20 data sets, 9 measured dispersion coefficients were closer to those predicted by the ANN model, as opposed to 4 by the D-S-B model, 2 by the Fischer model, 4 by the S-C model, and 5 by the K-F model. In other words, the ANN model gave better predictions for 38% of the measured data, as opposed to 18% by the D-S-B model, 8% by the Fischer model, 18% by the S-C model, and 21% by the K-F model.

**Table 2.** Statistical Information on the Whole, Training, and Testing Measured Data Sets

| Set | $\bar{U}$ (m/s) | $\bar{u}_*$ (m/s) | $(\bar{B}/\bar{H})$ | $\bar{β}$ | $\bar{σ}$ | $\bar{K}_x$ (m²/s) | % $B/H>50$ | % $K_x>100$ |
|---|---|---|---|---|---|---|---|---|
| Whole | 0.541 | 0.088 | 51.71 | 3.95 | 1.50 | 106.5 | 37 | 30 |
| Training | 0.501 | 0.082 | 50.30 | 3.92 | 1.54 | 96.4 | 35 | 31 |
| Testing | 0.644 | 0.103 | 55.30 | 4.01 | 1.40 | 132.2 | 40 | 25 |

Note: $\bar{U}$=mean velocity; $\bar{u}$=mean shear velocity; $\bar{B}/\bar{H}$=mean width/depth ratio; $\bar{β}$=mean shape parameter; $\bar{σ}$=mean sinuosity; and $\bar{K}_x$=mean dispersion coefficient.

**Table 3.** Range Values for the Whole, Training, and Testing Sets

| Set | $U_{min}-U_{max}$ (m/s) | $u_{*min}-u_{*max}$ (m/s) | $(B/H)_{min}-(B/H)_{max}$ | $\beta_{min}-\beta_{max}$ | $\sigma_{min}-\sigma_{max}$ | $K_{x_{min}}-K_{x_{max}}$ (m²/s) |
|---|---|---|---|---|---|---|
| Whole | 0.034–1.74 | 0.0024–0.553 | 13.8–156.5 | 2.62–5.05 | 1.08–2.54 | 1.9–892 |
| Training | 0.034–1.74 | 0.0024–0.268 | 16.0–156.5 | 2.77–5.05 | 1.09–2.54 | 1.9–837 |
| Testing | 0.130–1.53 | 0.0200–0.553 | 13.8–131.0 | 2.62–4.88 | 1.08–2.53 | 2.9–892 |

However, in order to perform the sensitivity analysis of the performance of the models, the values of the discrepancy ratio (DR) for each model was also computed. DR is commonly used as an error measure in the literature and was defined by Kashefi-pour and Falconer (2002) as

$$DR = \log \frac{K_{xp}}{K_{xm}} \qquad (13)$$

where $K_{xp}$=predicted dispersion coefficient, and $K_{xm}$=measured dispersion coefficient. It follows from Eq. (13) that if DR=0.0, then there is an exact prediction ($K_{xp}=K_{xm}$). Otherwise, there is either an overprediction (DR>0.0; and $K_{xp}>K_{xm}$) or underprediction (DR<0.0; and $K_{xp}<K_{xm}$).

The discrepancy ratios of each model for the data in Table 4 are shown in Fig. 4. Although there is a slight skewness of the results for the ANN model toward the positive zone, the distribution of error between −0.3 and 1.0 is symmetric. The DR values for the ANN model range from −0.22 to 1.13 (Fig. 4). These results imply that the ANN model slightly overestimated the measured data. However, according to Fig. 4, compared to the ANN model, all the other models, especially the Fischer model, either significantly underpredicted or overpredicted the measured data. The accuracy of each model is seen from Table 5. The accuracy of each model may be categorized by the number of DR values between −0.3 and 0.3 relative to the total number of data values (Seo and Cheong 1998; Kashefipour and Falconer 2002). As seen from Table 5, for this particular set of calibration and verification data, the accuracy of the ANN model is 70%, which is the highest of the models used. The D-S-B and K-F models follow the ANN model with 55% accuracy and the S-C model has an accuracy of 50% and the Fischer model has the lowest accuracy of 25%.

The root mean square error (RMSE), defined as (Dolling and Varas 2002)

$$RMSE = \sqrt{\frac{\sum_{i}^{N} (K_{xm_i} - K_{xp_i})^2}{N}} \qquad (14)$$

was computed for each model given in Table 5. In Eq. (14), $N$=number of observations, which is 20 for Table 4. In Table 5, three different RMSE values are presented for each model. The first one (column 3) is for the case where all the 20 measured data are included in the computation of the RMSE values; the second one (column 4) is for the case where the extreme measured dis-

**Table 4.** Comparison of Measured and Predicted Longitudinal Dispersion Coefficient ($K_x$)

| | Artificial neural networks | | Model | | | |
|---|---|---|---|---|---|---|
| River | Measured $K_x$ (m²/s) | ANNs | D-S-B | Fischer | S-C | K-F |
| Antietam Creek, Md. | 20.9 | 26.8 | 15.0 | 5.1 | 20.2* | 15.2 |
| Monocacy River, Md. | 37.8 | 27.1* | 28.2* | 61.7 | 27.1* | 7.6 |
| Monocacy River, Md. | 41.4 | 31.4* | 25.8 | 75.6 | 23.5 | 4.2 |
| Conococheague Creek, Md. | 53.3 | 43.0 | 93.1 | 88.2 | 96.7 | 58.8* |
| Chattahoochee River, Ga. | 88.9 | 77.6 | 168.6 | 127.9 | 169.1 | 82.1* |
| Bear Creek, Colo. | 2.9 | 39.2 | 28.1 | 7.3* | 52.3 | 27.1 |
| Tangipahoa River, La. | 44.0 | 26.5 | 28.7 | 143.0 | 39.2* | 24.5 |
| Sabina River, La. | 308.9 | 346.6* | 508.3 | 2,525.0 | 718.8 | 512.3 |
| Sabina River, Tex. | 12.8 | 21.9* | 4.6 | 2.0 | 5.2 | 2.4 |
| Mississippi River, La. | 237.2 | 838.0* | 1,618.0 | 2,133.0 | 1,855.0 | 1,619.0 |
| Mississippi River, Mo. | 457.7 | 838.0* | 1,245.0 | 1,0108.0 | 1,793.0 | 838.0* |
| Mississippi River, Mo. | 374.1 | 838.0* | 2,579.0 | 8,378.0 | 3,271.0 | 2,220.0 |
| Wind/Bigh. River, Wyo. | 41.8 | 59.7* | 156.6 | 226.5 | 160.0 | 75.9 |
| Clinch River, Va. | 10.7 | 26.9 | 28.4 | 25.1 | 27.5 | 11.5* |
| Clinch River, Va. | 36.9 | 76.6 | 118.3 | 52.2* | 139.7 | 104.1 |
| Powell River, Tenn. | 15.5 | 25.3 | 9.9* | 5.4 | 9.9* | 2.9 |
| John Day River, Ore. | 13.9 | 45.2 | 81.6 | 95.5 | 83.2 | 41.0* |
| John Day River, Ore. | 65.0 | 77.2 | 71.2* | 19.2 | 116.8 | 97.9 |
| White River | 30.2 | 31.7* | 46.4 | 229.0 | 54.3 | 17.7 |
| Missouri River | 892.0 | 763.4 | 950.8* | 4,120.0 | 1,317.0 | 990.5 |

Note: D-S-B=Deng et al. (2001); Fischer=Fischer (1975); S-C=Seo and Cheong (1998); K-F=Kashefipour and Falconer (2002); and *=close estimates of the measured data.
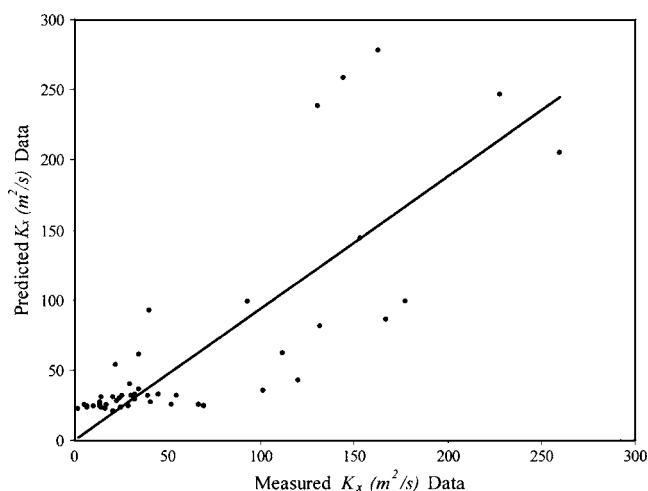


**Fig. 3.** Artificial neural network predicted dispersion coefficient data versus measured data at the end of the training stage
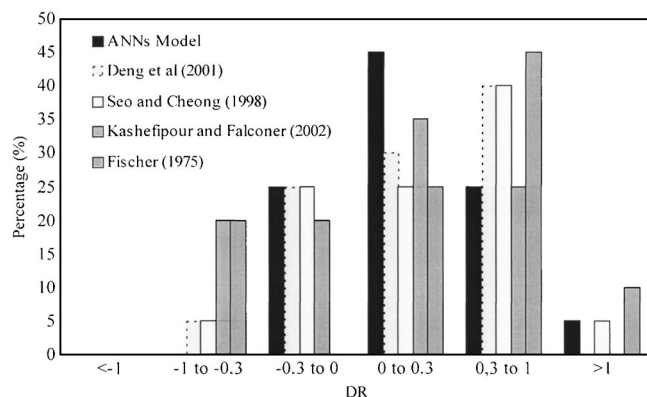
**Fig. 4.** Comparison of discrepancy ratio (DR) values for each model

**Table 5.** Comparison of Dispersion Coefficient Models

| Model | Accuracy (%) | RMSE $(m^2/s)$ | RMSE $(m^2/s)$ $(K_x > 100 \ m^2/s$ not included) | RMSE $(m^2/s)$ $(B/H > 50$ not included) |
|---|---|---|---|---|
| ANN | 70 | 193.0 | 19.3 | 183.0 |
| D-S-B | 55 | 610.0 | 46.7 | 416.0 |
| Fischer | 25 | 2,968.0 | 88.4 | 570.0 |
| S-C | 50 | 812.0 | 54.9 | 491.0 |
| K-F | 55 | 525.0 | 27.3 | 416.0 |

Note: RMSE=root mean square error; ANN=artificial neural network; D-S-B=Deng et al. (2001); Fischer=Fischer (1975); S-C=Seo and Cheong (1998); and K-F=Kashefipour and Falconer (2002).

persion coefficients $(K_x > 100 \ m^2/s)$ are not considered in the RMSE computation; and finally, the third one (column 5) is for the case where when the width-to-depth ratios are greater than 50 $(B/H > 50)$ are not included in the RMSE computation. As seen from Table 5, the ANN model has the lowest RMSE values in each case. This implies that the ANN model predicted the values of the dispersion coefficient in natural channels of different geometries better than did all other models. Although the D-S-B, S-C, and K-F models performed similarly, the Fischer model had the largest error in each case (Table 5). When the extreme values of the dispersion coefficient $(K_x > 100)$ were not considered in the error computation, the RMSE values decreased 10, 13, 15, 20, and 34 times for the ANN, D-S-B, S-C, K-F, and Fischer models, respectively. This implies that all the models, especially the K-F and Fischer models, significantly overpredicted the extreme values of the coefficient (see Table 4). When $B/H > 50$ data were not considered in the computation of the RMSE values, the decrease in the RMSE value was insignificant for the ANN model. This means that the ANN model can satisfactorily predict the values of the coefficient in wide and as well as in narrow channels. However, there was about a 1.5 times decrease in the RMSE values for the D-S-B, S-C, and K-F models and 5 times decrease for the Fischer model. This indicates that these models, especially the Fischer model, perform poorly in predicting the values of the coefficient in wide channels.

It follows from the previous comparison that the ANN model is capable of providing a superior prediction of the longitudinal dispersion coefficient in natural streams.

### Sensitivity Analysis

The sensitivity analysis of the ANN model was undertaken by predicting the measured longitudinal dispersion coefficient for the following seven cases. Table 6 summarizes the input variables and number of neurons in each input and hidden layers of the network for each case model. The number of neurons in each hidden layer was obtained by the trial and error procedure. Each case model was successfully trained with 20,000 iterations and 0.04 learning rate.

Case 1 considered channel width $(B)$, flow depth $(H)$, and flow velocity $(U)$ as inputs to the network and the dispersion coefficient $(K_x)$ as the target output. Note that shear velocity $(u^*)$ was not considered. The objective was to investigate the consequences of not utilizing the shear velocity as an input variable in the prediction of the dispersion coefficient.

Case 2 considered only flow discharge $(Q)$ as an input variable. In Case 1, there were three input variables, namely, the flow depth, flow velocity, and channel width. These variables in Case 1 were explicitly involved in the prediction of the dispersion coefficient. On the other hand, in Case 2, there was only one input variable, namely the flow discharge that is the product of flow depth, velocity, and channel width. Hence, in Case 2, the variables of flow depth, velocity, and channel width were implicitly involved in the prediction of the dispersion coefficient. The objective in Case 2 was to find out whether it would be sufficient to utilize only the discharge data to predict the dispersion coefficient.

Case 3 considered only flow velocity $(U)$ as the input variable. According to Jobson (2001), velocity plays an important role in

**Table 6.** Input Variables, Number of Neurons in Each Input and Hidden Layers, RMSE, and Accuracy Values for Each Case Model

| Artificial neural networks model | Input variables | Neurons input layer | Neurons hidden layer | RMSE $(m^2/s)$ | RMSE $(m^2/s)$ $(K_x > 100 \ m^2/s$ not included) | RMSE $(m^2/s)$ $(B/H > 50$ not included) | Accuracy (%) |
|---|---|---|---|---|---|---|---|
| Case 1 | $U, H, B$ | 3 | 5 | 193.0 | 21.2 | 183.0 | 75 |
| Case 2 | $Q$ | 1 | 3 | 191.0 | 20.0 | 179.0 | 65 |
| Case 3 | $U$ | 1 | 3 | 170.0 | 146.0 | 164.0 | 40 |
| Case 4 | $U, \beta$ | 2 | 4 | 137.0 | 105.0 | 118.0 | 70 |
| Case 5 | $U, \beta, \sigma$ | 3 | 5 | 119.0 | 77.7 | 99.0 | 58 |
| Case 6 | $U/u^*$ | 1 | 1 | 159.0 | 104.3 | 42.0 | 50 |
| Case 7 | $U/u^*, \beta, \sigma$ | 3 | 2 | 142.0 | 93.8 | 67.0 | 58 |

Note: ($U$=flow velocity; $H$=flow depth; $Q$=flow discharge; $u^*$=shear velocity; $U/u^*$=relative shear velocity; $B$=channel width; $\beta$=channel shape parameter; and $\sigma$=channel sinuosity.)
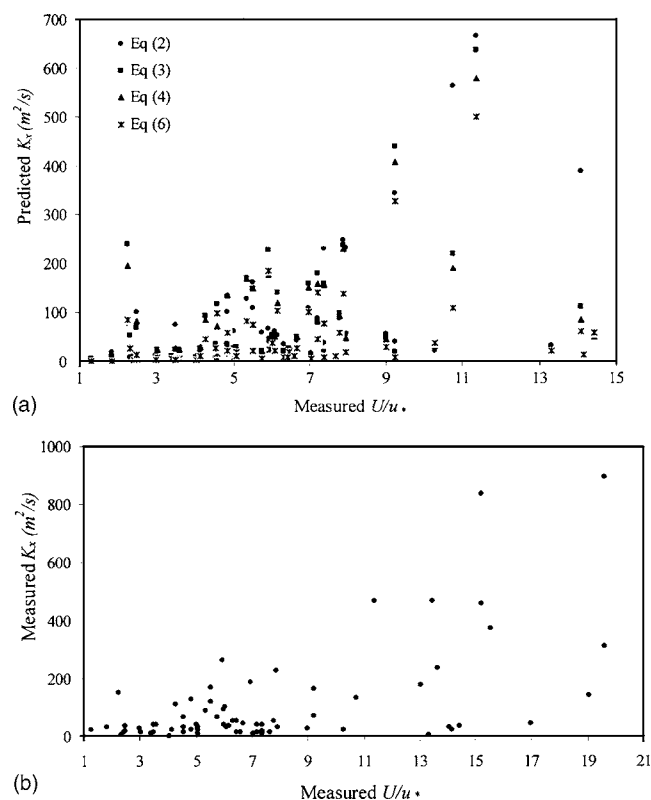
**Fig. 5.** (a) Predicted $K_x$ by different equations and (b) measured $K_x$ versus measured $U/u*$

the prediction of the effects of a pollutant spill. The objective here was to find out whether it would be sufficient to utilize only the velocity data in predicting the dispersion coefficient.

Case 4 considered flow velocity and channel shape parameter ($\beta$) as input variables. In this case, the objective was to find out whether the inclusion of the shape parameter along with velocity would improve the prediction.

Case 5 considered channel sinuosity ($\sigma$) in the input vector along with the channel shape parameter and flow velocity. The objective was to find out whether the inclusion of sinuosity would improve the predictions in Case 4.

Case 6 considered the relative shear velocity ($U/u*$) as the only input variable. In Eqs. (2)–(4) and (6), the dispersion coefficient ($K_x$) is expressed as a function of $U/u*$. Fig. 5(a) shows the relation of the $K_x$ predicted by Eqs. (2)–(4) and (6) to the measured $U/u*$. For clarity, the predicted extreme $K_x$ values (greater than $800 \text{ m}^2/\text{s}$) by the equations corresponding to 10 different measured $U/u*$ were not shown in Fig. 5(a). For these 10 different cases, the predictions by the equations varied from 800.0 to 10,000.0 $\text{m}^2/\text{s}$. Fig. 5(a) shows the remaining 61 different cases for which the equations yielded a wide range of variation in the prediction of $K_x$. Fig. 5(b), on the other hand, shows the measured $U/u*$ versus measured $K_x$. The linear relationship between the measured $K_x$ and the measured $U/u*$ is low with $R^2 = 0.34$ [Fig. 5(b)]. This implies that the behavior of the dispersion process is more nonlinear than linear and, therefore, it is more suited for a nonlinear model, such as ANN. The objective in this case was to find out whether it would be sufficient to predict $K_x$ from the $U/u*$ data.

Case 7 considered the channel shaper parameter ($\beta$) and channel sinuosity ($\sigma$) along with the relative shear velocity ($U/u*$) in the input vector. The objective here was to find out whether the

**Table 7.** Comparison of Measured and Predicted Longitudinal Dispersion Coefficient ($K_x$)

| River | Measured $K_x$ ($\text{m}^2/\text{s}$) | Case 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Antietam Creek, Md. | 20.9 | 29.2 | 38.9 | 64.5 | 25.6* | 42.3 | 55.0 | 100.2 |
| Monocacy River, Md. | 37.8 | 29.5 | 39.8* | 28.5 | 61.3 | 47.5 | 54.4 | 32.9 |
| Monocacy River, Md. | 41.4 | 33.7 | 40.6* | 16.4 | 61.9 | 52.4 | 41.8* | 20.6 |
| Conocochea. Creek, Md. | 53.3 | 41.7 | 45.9* | 134.9 | 96.5 | 70.0 | 92.2 | 89.6 |
| Chattahooch. River, Ga. | 88.9 | 74.8* | 77.3* | 182.3 | 132.0 | 98.7 | 58.2 | 47.3 |
| Bear Creek, Colo. | 2.9 | 59.3 | 41.7 | 384.2 | 458.8 | 175.6 | 35.0 | 27.6* |
| Tangipahoa River, La. | 44.0 | 28.2 | 39.1* | 42.7* | 68.6 | 64.4 | 264.0 | 366.8 |
| Sabina River, La. | 308.9 | 311.5* | 478.0 | 315.3* | 326.5* | 309.9* | 287.6 | 418.4 |
| Sabina River, Tex. | 12.8 | 25.0 | 38.4 | 13.8* | 12.8* | 33.1 | 42.0 | 68.2 |
| Mississippi River, La. | 237.2 | 838.0 | 826.0 | 107.2 | 74.1 | 58.6 | 313.2 | 291.0* |
| Mississippi River, Mo. | 457.7 | 838.0 | 826.0 | 311.7 | 371.0 | 410.6* | 240.1 | 287.0 |
| Mississippi River, Mo. | 374.1 | 838.0 | 826.0 | 428.0* | 481.0 | 523.9 | 245.3 | 326.5* |
| Wind/Bigh. River, Wyo. | 41.8 | 56.8* | 54.5* | 244.0 | 219.7 | 182.4 | 85.6 | 66.8 |
| Clinch River, Va. | 10.7 | 29.3* | 39.6 | 44.8 | 47.6 | 25.2 | 55.1 | 38.3 |
| Clinch River, Va. | 36.9 | 69.5 | 65.3 | 147.4 | 68.7 | 33.0* | 67.8 | 65.7 |
| Powell River, Tenn. | 15.5 | 28.1 | 39.1 | 13.8* | 19.3* | 36.2 | 35.4 | 37.4 |
| John Day River, Ore. | 13.9 | 44.6* | 41.6* | 297.0 | 248.0 | 192.8 | 82.8 | 65.4 |
| John Day River, Ore. | 65.0 | 74.7 | 58.7* | 218.0 | 76.9 | 83.3 | 50.3 | 79.6 |
| White River | 30.2 | 33.1* | 41.2 | 44.8 | 104.0 | — | 95.2 | — |
| Missouri River | 892.0 | 771.3* | 814.8* | 431.1 | 490.0 | 533.2 | 287.5 | 429.3 |

Note: *=Close estimates of the measured data.

**998** / JOURNAL OF HYDRAULIC ENGINEERING © ASCE / NOVEMBER 2005

inclusion of geometric parameters would improve the predictions in Case 6. Note that in Case 7, through the shape parameter the channel width and flow depth are implicitly involved in the input vector. Also, by the relative shear velocity, flow velocity and shear velocity are implicitly involved in the input vector. In a way, the flow variables (flow depth, velocity, and shear velocity) and geometric parameters (channel width and sinuosity) were all involved in the input vector of ANN to predict the dispersion coefficient.

The results of prediction of the measured dispersion coefficient data by the 7 case models are given in Table 7 and these show that Case 1–7 models, in general, yielded good estimates of measured data. The RMSE values and accuracy of each case model are listed in Table 6. From Tables 6 and 7, it can be deduced that Case 1 and 2 models, in general, exhibited similar performances and overpredicted the values of $K_x > 100$ m$^2$/s that constituted 25% of the testing data (Table 2). This is expected, because Case 1 model uses the flow velocity, depth and channel width data as inputs whereas Case 2 model uses the product of these variables, which is the flow discharge. In a way, both models use similar input information. The RMSE value in Case 2 is 191.0 m$^2$/s when all the predicted data are considered in the error computation. This value decreased to the lowest value of 20.0 m$^2$/s when $K_x > 100$ m$^2$/s data are not considered in the error computation (Table 6). This implies that flow discharge data can, in general, yield better predictions of the commonly encountered low values ($K_x < 100$ m$^2$/s) of the coefficient.

If flow velocity is used as the only input variable (Case 3) then the accuracy of prediction is the lowest$=40$% (Table 6). Using channel shape parameter along with the flow velocity (Case 4) improves the accuracy of Case 3 model from 40 to 70%. Using geometric characteristics (channel shape parameter and sinuosity) along with the flow velocity (Case 5) in the input vector yields satisfactory predictions with the lowest RMSE of 119.2 (Table 6).

The RMSE value of 159.0 in Case 6 model decreased to the lowest value of 42.3 when $B/H > 50$ data were not considered in the error computation (Table 6). This implies that using relative shear velocity ($U/u^*$) as the only input variable (Case 6) in the network yields better estimates of the dispersion coefficient in narrower ($B/H < 50$) channels that constitute 60% of the testing data (Table 2). On the other hand, the 50% accuracy of Case 6 model increased to 58% due to the inclusion of channel shape parameter and sinuosity along with the relative shear velocity (Case 7) in the input vector (Table 6).

## Summary and Conclusions

An artificial neural network model was developed for predicting the longitudinal dispersion coefficient in natural streams from flow variables and channel geometric characteristics. The satisfactory predictions of the measured data from streams having different geometric and flow characteristics revealed that the developed ANN model was superior to the existing theoretical and empirical equations.

When a one-dimensional dispersion model is applied to predict the concentration variation of pollutants in natural streams after the Fickian period is reached, the selection of a proper dispersion coefficient is the most important and also the most difficult task (Seo and Cheong 1998). It is a relatively simple task to use a measured dispersion coefficient, if it is known. However, for streams where mixing and dispersion characteristics are unknown, the dispersion coefficient can only be estimated using a theoretical or empirical equation. Because most studies have been carried out based on specific assumptions and channel conditions, the performance of existing equations varies widely for the same stream and flow condition [see Fig. 5(a)]. On the other hand, the ANN model, developed in this study, makes no assumption with regard to stream geometry or flow dynamics in the stream. It has proved to capture the relation between the input and output parameters of the system, resulting in satisfactory prediction of a wide range of values of the dispersion coefficient in natural streams of varying geometric and flow conditions. The developed ANN model overcomes the shortcomings of the existing models. It can be concluded that the ANN model can be a strong alternative modeling tool to the existing theoretical and empirical models to predict the dispersion coefficient in natural streams.

The sensitivity analysis results led to the following conclusions.

1. If the data on shear velocity, flow velocity, depth, and channel width are available, then the ANN model successfully predicts the wide ranging values of the dispersion coefficient of natural streams of different geometries. However, if one has only the discharge data then one can use that data in the ANN model to satisfactorily predict the more frequently encountered low values of the dispersion coefficient ($K_x < 100$ m$^2$/s).

2. If geometric characteristics of the channel shape parameter and sinuosity are used along with the flow velocity in the input vector, then the ANN model satisfactorily predicts the dispersion coefficient.

3. If the relative shear velocity is used as the only input variable, then the ANN model can yield satisfactory predictions of the dispersion coefficient in more frequently encountered narrower channels ($B/H < 50$).

4. The geometric characteristics, when used along with the relative shear velocity, can significantly improve the performance of the ANN model in predicting the longitudinal dispersion coefficient in natural streams.

## Notation

*The following symbols are used in this paper:*

$A$ = cross-sectional area;

$B$ = channel width;

$b_j$ = the threshold value, also called the bias, associated with node $j$;

$H$ = cross-sectional average flow depth;

$h$ = local flow depth;

$K_x$ = longitudinal dispersion coefficient;

$K_{xm}$ = measured dispersion coefficient;

$K_{xp}$ = predicted dispersion coefficient;

$N$ = number of observations

$net_j$ = summation of the weighted input for the $j$th neuron;

$P$ = number of training patterns;

$p$ = number of output neurons;

$t_i$ = component of a target output vector $T$;

$U$ = cross-sectional average flow velocity;

$u^*$ = shear velocity;

$u'$ = deviation of local depth mean flow velocity from cross-sectional mean;

$v_{ij}$ = weight from the ith neuron in the previous layer to the $j$th neuron in the current layer;

$x_i$ = input from the $i$th neuron to the $j$th neuron;

$x_{max}$ = maximum input value;
$y$ = coordinate in the lateral direction;
$y_i$ = component of an ANN output vector $\mathbf{Y}$;
$z_i$ = standardized value;
$\beta$ = shape parameter;
$\delta$ = learning rate; and
$\varepsilon_t$ = local transverse mixing coefficient.

## References

ASCE Task Committee on application of Artificial Neural Networks in Hydrology. (2000). "Artificial neural networks in hydrology. I: Preliminary concepts." *J. Hydrologic Eng.*, 5(2), 115–123.

Aziz, A. R. A., and Wong, K. F. V. (1992). "Neural network approach to the determination of aquifer parameters." *Ground Water*, 30(2), 164–166.

Chang, H. H. (1988). *Fluvial processes in river engineering*, Wiley, New York.

Chua, L. H. C., and Holz, K. P. (2005). "Hybrid neural network. Finite element river flow model." *J. Hydraul. Eng.*, 131(1), 52–59.

Dawson, W. C., and Wilby, R. (1998). "An artificial neural network approach to rainfall-runoff modeling." *Hydrol. Sci. J.*, 43(1), 47–66.

Deng, Z. Q., Singh, V. P., and Bengtsson, L. (2001). "Longitudinal dispersion coefficient in straight rivers." *J. Hydraul. Eng.*, 127(11), 919–927.

Deng, Z. Q., Bengtsson, L., Singh, V. P. and Adrian, D. D. (2002). "Longitudinal dispersion coefficient in single-channel streams." *J. Hydraul. Eng.*, 128(10), 901–916.

Dolling, O. R., and Varas, E. A. (2002). "Artificial neural networks for streamflow prediction." *J. Mater. Sci. Technol.*, 40(5), 547–554.

Elder, J. W. (1959). "The dispersion of a marked fluid in turbulent shear flow." *J. Fluid Mech.*, 5(4), 544–560.

Fischer, B. H. (1967). "The mechanics of dispersion in natural streams." *J. Hydraul. Div., Am. Soc. Civ. Eng.*, 93(6), 187–216.

Fischer, B. H. (1975). "Discussion of 'Simple method for predicting dispersion in streams,' by R. S. McQuivey and T. N. Keefer." *J. Environ. Eng. Div. (Am. Soc. Civ. Eng.)*, 101(3), 453–455.

Fischer, H. B., List, E. J., Koh, R. C. Y., Imberger, J., and Brooks, N. H. (1979). *Mixing in inland and coastal waters*, Academic, New York, 104–138.

Fukuoka, S., and Sayre, W. W. (1973). "Longitudinal dispersion in sinuous channels." *J. Hydraul. Div., Am. Soc. Civ. Eng.*, 99(1), 195–217.

Halff, A. H., Halff, H. M., and Azmoodeh, M. (1993). "Predicting runoff from rainfall using neural networks." *Proc., Engineering Hydrology*, ASCE, New York, 760–765.

Iwasa, Y., and Aya, S. (1991). "Predicting longitudinal dispersion coefficient in open-channel flows." *Proc., Int. Symp. on Environmental Hydrology*, Hong Kong, 505–510.

Jirka, G. H. (2004). "Mixing and dispersion in rivers." *River Flow 2004*, M. Greco, A. Carravetta, and R. D. Morte, eds., Taylor and Francis, London, 13–27.

Jobson, H. E. (2001). "Predicting river travel time from hydraulic characteristics." *J. Hydraul. Eng.*, 127(11), 911–918.

Kashefipour, M. S., and Falconer, R. A. (2002). "Longitudinal dispersion coefficients in natural channels." *Water Res.*, 36(6), 1596–1608.

Liu, H. (1977). "Predicting dispersion coefficient of streams." *J. Environ. Eng. Div. (Am. Soc. Civ. Eng.)*, 103(1), 59–69.

Magazine, M. K, Pathak, S. K., and Pande, P. K. (1988). "Effect of bed and side roughness on dispersion in open channels." *J. Hydraul. Eng.*, 114(7), 766–782.

Maier, H. R., and Dandy, G. C. (1996). "The use of artificial neural networks for the prediction of water quality parameters." *Water Resour. Res.*, 32(4), 1013–1022.

Markus, M., Salas, J. D., and Shin, H-S. (1995). "Predicting streamflows based on neural networks." *Proc., 1st Int., Conf. on Water Resources Engineering*, ASCE, New York, 1641–1646.

McQuivey, R. S., and Keefer, T. N. (1976). "Convective model of longitudinal dispersion." *J. Hydraul. Div., Am. Soc. Civ. Eng.*, 102(10), 1409–1424.

Morshed, J., and Kaluarachchi, J. J. (1998). "Parameter estimation using artificial neural network and genetic algorithm for free-product and recovery." *Water Resour. Res.*, 34(5), 1101–1113.

Rutherford, J. C. (1994). *River mixing*, Wiley, Chichester, U.K.

Sen, Z., Altunkaynak, A., and Ozger, M. (2004). "Sediment concentration and its prediction by perceptron Kalman filtering procedure." *J. Hydraul. Eng.*, 130(8), 816–826.

Seo, I. W., and Baek, K. O. (2004). "Estimation of the longitudinal dispersion coefficient using the velocity profile in natural streams." *J. Hydraul. Eng.*, 130(3), 227–236.

Seo, I. W., and Cheong, T. S. (1998). "Predicting longitudinal dispersion coefficient in natural streams." *J. Hydraul. Eng.*, 124(1), 25–32.

Smith, J., and Eli, R. N. (1995). "Neural-network models of rainfall-runoff process." *J. Water Resour. Plan. Manage.*, 121(6), 499–508.

Somez, I. (1998). "Meteorological applications of artificial neural networks." MSc thesis, Dept. of Meteorological Engineering, Istanbul Technical Univ., Istanbul, Turkey (in Turkish).

Sooky, A. A. (1969). "Longitudinal dispersion in open channels." *J. Hydraul. Div., Am. Soc. Civ. Eng.*, 95(4), 1327–1346.

Sukhodolov, A. N., Nikora, V. I., Rowinsky, P. M., and Czernuszenko, W. (1997). "A case study of longitudinal dispersion in small lowland rivers." *Cogn. Instruct.*, 69(7), 1246-1333.

Tayfur, G. (2002). "Artificial neural networks for sheet sediment transport." *Hydrol. Sci. J.*, 47(6), 879–892.

Tayfur, G., Swiatek, D., Wita, A., and Singh, V. P. (2005). "Case study: Finite element and artificial neural network models for flow through Jeziorsko earthfill dam in Poland." *J. Hydraul. Eng.*, 131(6), 431–440.

Taylor, G. I. (1953). "Dispersion of soluble matter in solvent flowing slowly through a tube." *Proc. R. Soc. London, Ser. A*, 219, 186–203.

Taylor, G. I. (1954). "The dispersion of matter in turbulent flow through a pipe." *Proc. R. Soc. London, Ser. A*, 223, 446–468.

Tokar, A. S., and Johnson, P. A. (1999). "Rainfall-runoff modeling using artificial neural networks." *J. Hydrologic Eng.*, 4(3), 232–239.