Are artificial neural network techniques relevant for the estimation of longitudinal dispersion coefficient in rivers?

PAWEŁ M. ROWIŃSKI, ADAM PIOTROWSKI & JAROSŁAW J. NAPIÓRKOWSKI

Institute of Geophysics, Polish Academy of Sciences, Ks. Janusza 64, 01-452 Warsaw, Poland
pawelr@igf.edu.pl

Abstract Accurate application of the longitudinal dispersion model requires that specially designed experimental studies are performed in the river reach under consideration. Such studies are usually very expensive, so in order to quantify the longitudinal dispersion coefficient, as an alternative approach, various researchers have proposed numerous empirical formulae based on hydraulic and morphometric characteristics. The results are presented of the application of artificial neural networks as a parameter estimation technique. Five different cases were considered with the network trained for different arrangements of input nodes, such as channel depth, channel width, cross-sectionally averaged water velocity, shear velocity and sinuosity index. In the case where the sinuosity index is included as an input node, the results turned out to be better than those presented by other authors.

Key words artificial neural networks; longitudinal dispersion; pollutant transport; rivers

Les techniques de réseaux de neurones artificiels sont-elles pertinentes pour estimer le coefficient de dispersion longitudinale en rivières?

Résumé L'application précise du modèle de dispersion longitudinale nécessite la mise en œuvre d'études expérimentales spécifiques dans le cours d'eau considéré. De telles études sont en général très coûteuses, si bien que, dans le but de quantifier de manière alternative le coefficient de dispersion longitudinale, plusieurs chercheurs ont proposé de nombreuses formules empiriques basées sur des caractéristiques hydrauliques et morphométriques. Nous présentons les résultats de l'application de réseaux de neurones artificiels comme technique d'estimation des paramètres. Cinq cas différents ont été considérés, avec, pour l'apprentissage, des combinaisons variées de nœuds d'entrée, dont la profondeur du chenal, la largeur du chenal, la vitesse moyenne de l'eau à travers la section, la vitesse de cisaillement et l'indice de sinuosité. Lorsque l'indice de sinuosité est pris en compte comme nœud d'entrée, les résultats sont améliorés par rapport à ceux d'autres auteurs.

Mots clefs réseaux de neurones artificiels; dispersion longitudinale; transport de polluant; rivières

INTRODUCTION

Despite extensive studies, quantitative description of longitudinal dispersion in rivers still constitutes a key question lying at the heart of considerations in environmental fluid mechanics. The simplest, but still most often used, model in engineering practice is that based on the advection–dispersion equation. Questions about its applicability can be raised and alternative approaches proposed (e.g. Czernuszenko & Rowiński, 1997; Lees *et al.*, 2000; Wörman, 2000; Rowiński, 2002). However, a large amount of existing data is based on the implicit assumption of the correctness of such a formulation. To treat the Fickian model as a predictive tool, one needs to know how to relate the usually unknown longitudinal dispersion coefficient to basic hydraulic and

Open for discussion until 1 August 2005

morphometric characteristics of the natural stream under consideration. Numerous empirical and semi-empirical formulae have been found and various researchers dispute which expression is the most useful. The advantage of one expression over another is often just a matter of the selection of data and the manner of their presentation. Moreover, regardless of the expression applied, one may easily find an outlier in the data which definitely does not support the applicability of a particular formula. An expectation that, in spite of the complexity of the river reach, the dispersion coefficient may be represented by one of the empirical formulae seems to be exaggerated; nevertheless, it is quite a common practice in hydraulic engineering. However, there is a tendency to differentiate the approaches depending on the complexity of the case. Therefore Deng *et al*. (2001, 2002) and Guymer (1998) devoted separate studies to straight and meandering streams. Holley & Jirka (1986) claimed that all of the methods in the literature were based on subsets of the available data, and that independent verification or evaluation of the methods had rarely been done by disinterested persons. This statement seems to remain valid. It turns out that when a longitudinal dispersion model is planned to be used, identification of the dispersion parameter should be based on specially designed experimental studies performed in the reach under consideration. On the other hand, managers and decision makers, who have just a modelling tool at their disposal and the basic information about the stream, are supposed to derive some conclusions about the admixture pattern in the stream after its release at some location. To what extent may they rely on the simulations when the dispersion coefficient is taken from a completely different location or from an empirical formula obtained for a different set of data? In other words, one may ask whether the spread of pollution can be evaluated based on the easily available bulk channel parameters.

The picture is not very optimistic in this respect and one may pose a question about what lesson one can derive from previous experience. An inspiration for this study was a paper by Kashefipour *et al*. (2002) in which, to predict the longitudinal dispersion coefficient, an artificial neural network (ANN) was trained for different arrangements of channel depth, channel width, cross-sectionally averaged water velocity and shear velocity. The authors followed a similar track to answer the questions posed above and obtained essential improvement in predicting the dispersion coefficient when an additional parameter reflecting the meandering of the river, namely the sinuosity index (*S*), was used for training the network. Generally speaking, this study is concerned with a trial to extract some knowledge from a relatively large database that results from a series of independent experiments carried out all over the world. Can such knowledge be put into practical use through relatively good physically-based mathematical models of pollution transport? The present study is aimed at stimulating a discussion on this vital topic.

LONGITUDINAL DISPERSION

In the situation where the concentration distribution in a lateral direction more or less equalizes, the main interest is paid to concentrations of constituents averaged over the cross-section. This usually occurs at long distances from the release point, particularly in rivers of relatively simple geometry. It is assumed that any quantity φ may be decomposed into the area-averaged value $\overline{\varphi}^a$ and a deviation from the area-averaged quantity φ^a. Further it is assumed that the created dispersive flux is proportional to the gradient of the area-averaged concentration (Fick's law):

$$\overline{q}_L^a = -E_L \frac{\partial \overline{C}^a}{\partial x} \tag{1}$$

where x is the longitudinal axis, $\overline{q}_L^a$ is the dispersion flux, $\overline{C}^a$ is the admixture concentration averaged over the cross-section A, and E_L is the longitudinal dispersion coefficient. Integration of the advection-diffusion in three dimensional form and some algebraic manipulations lead to the following one-dimensional (1-D) mass conservation equation:

$$\frac{\partial \overline{C}^a}{\partial t} + \overline{U}^a \frac{\partial \overline{C}^a}{\partial x} = \frac{1}{A}\frac{\partial}{\partial x}\left(A E_L \frac{\partial \overline{C}^a}{\partial x} \right) \tag{2}$$

where $\overline{U}^a$ is cross-sectional averaged velocity. The longitudinal dispersion coefficient is hence a key parameter for the description of the longitudinal transport of a constituent in a river. As boundary conditions are considered, concentrations should be given for inflows, while the zero gradient condition is used otherwise, since no mass transport takes place along river banks. For channels in which an assumption about constant cross-sectional area and a constant mean longitudinal velocity is a reasonable approximation, equation (2) may be simplified to:

$$\frac{\partial \overline{C}^a}{\partial t} + \overline{U}^a \frac{\partial \overline{C}^a}{\partial x} = E_L \frac{\partial^2 \overline{C}^a}{\partial x^2} \tag{3}$$

Equation (3) has a relatively simple form and some boundary value problems associated with this equation for semi-infinite channels have analytical solutions. Assume that at time $t = 0$ an instantaneous release of mass M occurs at the origin x and the concentration tends to zero at infinite time, i.e.:

$$\overline{C}^a(\pm\infty, t) = 0 \qquad \overline{C}^a(x, 0) = M\delta(x) \tag{4}$$

where $\delta(x)$ is a Dirac delta function. Then the solution of equation (3) reads:

$$\overline{C}^a(x,t) = \frac{M}{\sqrt{4\pi E_L t}} \exp\left(-\frac{\left(x - \overline{U}^a t\right)^2}{4 E_L t} \right) \qquad t > 0 \tag{5}$$

To simplify the notation further, in this paper the overbar representing the averaging procedure will be omitted. It can be shown that the variance of the solute fulfilling equation (3) increases linearly with time and the longitudinal dispersion coefficient satisfies the following relationship allowing for its experimental determination:

$$E_L = \frac{1}{2}\frac{\mathrm{d}\sigma_x^2}{\mathrm{d}t} \tag{6}$$

where σ_x^2 is the spatial variance in the longitudinal direction.

The literature contains many different forms and methods of the evaluation of dispersion coefficients dependent on various hydraulic conditions (e.g. Fischer, 1967; Holley & Jirka, 1986; Sukhodolov *et al.*, 1997; Guymer, 1998; Deng *et al.*, 2001, 2002).

A long tradition exists in the application of the longitudinal dispersion model. Such a model is very useful for designing outfalls or water intakes and, above all, for evaluating risks from, for example, accidental releases of hazardous contaminants. one-dimensional models are the easiest in application and less demanding as to the details of the hydraulic characteristics of the considered reach compared to 2-D and 3-D approaches. However, one should be aware of the restrictions that burden 1-D models. A 1-D model in the form given above applies only in relatively simple river reaches (in terms of the geometry) and only after some initial mixing period (the solute concentration should be well distributed over the channel width). No universal criterion for the validity of the equation exists. For example, Jirka (2004) shows that complete lateral mixing in cases of large rivers (e.g. the River Rhine) may require a distance of 160 km. The advection–dispersion equation has been applied successfully to many real cases; nevertheless, questions about its applicability frequently arise. The tail of a solute tracer pulse is often more pronounced than can be accounted for by the traditional advection–dispersion model. A common method for simulating such long tails has been to allow for storage zones along the stream channel. These storage zones are assumed to be stagnant relative to the longitudinal flow of the stream and to obey a first-order mass transfer type of exchange relationship. Very often a quicker decrease of the concentration maximum is observed than follows from equation (5). Also, a nonlinear growth of the concentration distribution variance and dependence of the dispersion coefficient on time have often been manifested in experimental studies. The discrepancy between the Fickian solution and the experimental data, especially in the lower range of concentration distributions ($C < 0.5C_{max}$) may even be shown for relatively simple geometries (Sukhodolov *et al.*, 1997).

All the above features mean that the longitudinal dispersion coefficient cannot be uniquely identified when the Fickian model is applied to a natural stream, and each applied method can lead to slightly different results. The necessary conditions that have to be fulfilled are usually not fully satisfied and, therefore, what is labelled as a measured coefficient is loaded with a significant error in itself. One has to look at any error analysis with caution.

The objective of this brief discussion is to emphasize the uncertainty in the estimates of the dispersion coefficient.

PROBLEM STATEMENT

Estimation of the longitudinal dispersion coefficient constitutes a basic difficulty in the application of the so-called Fick model. Several estimation methods have been elaborated in the literature, such as physically-based empirical methods, fitting of the theoretical slope of the Laplace transformed solution for the concentration of the flow zone to the observed slope, moments matching procedures, or even visual determination of the set of parameters yielding the best fit to the concentration data. An obvious element is the relation of the computed solute concentrations to some experimentally obtained curves. Recently, a relevant optimization problem was solved by means of the global random search procedure applied to a longitudinal dispersion model that takes into account dead zones (Rowiński *et al*., 2004). The classical Fickian model may be treated as a simpler case of the dead-zone model (Czernuszenko *et al*., 1997). All these methods work when one knows the breakthrough curves for a particular river reach, but the question remains whether one is able to predict the value of the dispersion coefficient based on previous experience. As mentioned in the preceding section, there are numerous, more or less empirical forms allowing for the evaluation of the dispersion coefficient—all of them are, however, of disputable value. Assuming a good quality of the historical data in this respect, one may try to use the technique based on artificial neural networks which is the basis for all further considerations herein.

ARTIFICIAL NEURAL NETWORKS AS PARAMETER ESTIMATION TECHNIQUE

Artificial neural networks have been developed by looking for analogies to the behaviour and functioning of the brain and nervous system of living organisms. The most important feature imitating the brain is the ability to learn from examples and to utilize the gained knowledge to solve new problems. It is a kind of ability for generalization which the authors hoped could be useful for identification of the dispersion coefficient in water quality models.

The multi-layer perceptron network, i.e. one of supervised feed-forward networks, was selected for this study from a large variety of neural networks (e.g. Korbicz *et al*., 1994; Haykin, 1994). Such networks are able to approximate the values of output variable y dependent on the set of input variables $x_1, x_2, ..., x_N$ and, hence, $y = f(x_1, x_2, ..., x_N)$. Artificial neural networks comprise of several simple units—nodes or neurons—arranged in a parallel and cascade fashion, i.e. in input, hidden and output layers. The topology of the neural network used in this study is presented in Fig. 1. The number of input nodes is the same as the number of input variables (three or four in this study), the number of hidden neurons should be found "as optimal" for the

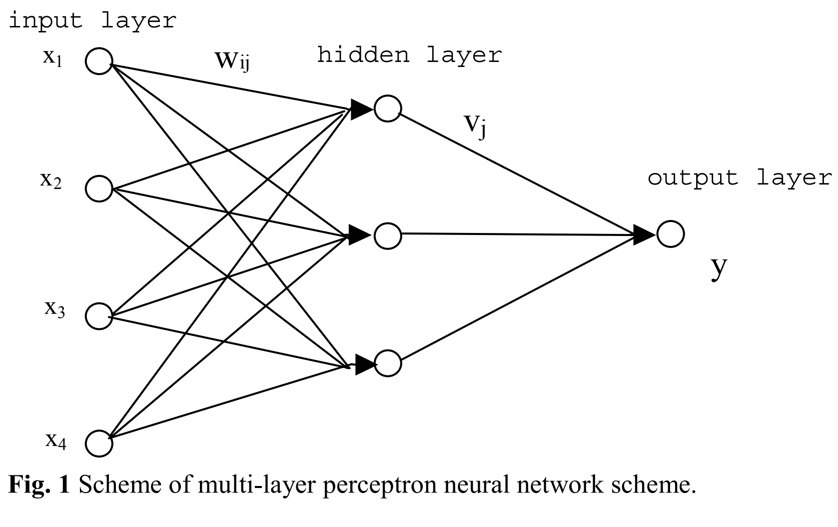


Fig. 1 Scheme of multi-layer perceptron neural network scheme.

solution of a given problem, and the number of output nodes is equal to the number of output variables (one output variable, namely longitudinal dispersion coefficient was considered in this study). The applied network consists of only one hidden layer, due to the general observation that neural networks with only one hidden layer and a finite number of nodes are able to approximate every continuous, bounded, everywhere differentiable function (e.g. Haykin, 1994; Hsu *et al.*, 1995). Simple computational elements (the nodes) are linked via weighted connections. The values of these connections are adaptively modified during the process of training the network. Each node performs a weighted sum of its inputs and filters it through a given, so-called activation function. Following a number of other authors, a sigmoidal function was used for this purpose, i.e.:

$$f(z) = \frac{1}{1 + \mathrm{e}^{-z}} \tag{7}$$

where z is the weighted sum

$$z_j = f\left(\sum_{i=1}^{N} w_{ji} x_i\right) \tag{8}$$

Afterwards the weighted signals z_j (multiplied by proper weights v_j), are transferred to the neuron of the third layer, where the new weighted sum is computed:

$$y = \sum_{j=1}^{K} v_j f\left(\sum_{i=1}^{N} w_{ji} x_i\right) \tag{9}$$

The learning process is accomplished by means of a training algorithm that lets the network modify its weights according to the training examples. The weights are modified by applying some algorithm that minimizes the mean square error between the network's output and the desired target. The objective function J is defined as:

$$J(w,v) = \min_{w,v} \sum_{i=1}^{m} (y_i(w,v) - d_i)^2 \tag{10}$$

where N is the number of inputs; K is the number of hidden neurons; w and v are vectors of weights of the neural network; m is the number of training vectors; $y_i(\ )$ is the output value calculated from the neural network for the ith training vector; and d_i is the measured value for the ith training vector.

In the present study, a supervised learning method was adopted, so each teaching example was composed of an input to the network as well as the corresponding output. Training an ANN may require a lot of time and work and, therefore, sufficiently fast optimization techniques are sought. Due to the fact that the network of interest is relatively small with a single output unit, a Lavenberg-Marquardt nonlinear optimization algorithm could be used (Press *et al.*, 1990).

The ANN has learned from different sets of the available data. The reasoning and the method of selection of those data will be presented further in the paper. The data sets were divided randomly into training and verification parts. The verification set of independent data (not used in the training of the network) was used to evaluate the true performance of the network.

COMPUTATIONAL RESULTS

Following a strong tendency in the application of ANNs to work on combinations of the channel measurements, all networks (see Table 1) were trained for specific arrangements of channel depth, *H* (m), channel width, *B* (m), cross-sectionally averaged water velocity, *U* (m s^{-1}) and shear velocity, $U*$ (m s^{-1}), namely for nodes described as $x_1 = B/H$, $x_2 = U/U*$, $x_3 = 3UB$. Moreover, the ratio of the length of the main river bed to the length of the valley, i.e. the sinuosity index, *S*, was considered as an additional input node ($x_4 = S$) in Cases 1 and 3. Note that the sought output variable *y* is the dispersion coefficient, E_L.

The *UB* product is preceded by a proportionality factor of 3, which does not influence the results, but enables some consistency with the results of Deng *et al.* (2002) to be achieved. This proportionality factor comes from the observation of those authors that the longitudinal dispersion coefficient may be approximated to some extent by $E_L \approx 3UB$. This value 3 should by no means be treated as universal; in other studies different values were obtained, for example Rochusaar & Paal (1970) obtained a value of 1.5 for small Estonian rivers, while Sukhodolov *et al.* (1997) obtained a value of 0.83 for the experimental results for Moldovian rivers. Sukhodolov *et al.* (1997) showed that this proportionality factor should be a function of the degree of heterogeneity of the velocity distribution and the Lagrangian spatial scales, such as alternate bars in rivers where those bars can be treated as major channel forms. Those factors undoubtedly depend on planform channel geometry and therefore it is expected

Table 1 Basic parameters of computational cases.

	Case 1	Case 2	Case 3	Case 4	Case 5
Number of data	70	70	81	81	99
Number of training data	60	60	50	50	79
Number of verification data	10	10	31	31	20
Sources of data	(1)		(1), (2)		(2), (3), (4), (5)
Input variables	B/H U/U_* $3UB$ S	B/H U/U_* $3UB$	B/H U/U_* $3UB$ S	B/H U/U_* $3UB$	B/H U/U_* $3UB$
Input nodes	4	3	4	3	3
Hidden nodes	3	4	3	3	4
Output nodes	1	1	1	1	1
Number of all computed E_{lp} beyond $-0.3 < DR < 0.3$	4/70	14/70	7/81	31/81	19/99
Number of computed E_{lp} from verification data beyond $-0.3 < DR < 0.3$	1/10	5/10	3/31	18/31	5/20
Percentage of training data mean error	7.02%	9.89%	7.12%	13.97%	12.52%
Percentage of verification data mean error	7.33%	25.95%	10.27%	28.02%	17.79%
Performance (Pf) of training set	0.248	0.350	0.212	0.367	0.317
Performance (Pf) of verification set	0.356	0.735	0.207	0.509	0.300

B: channel width; *H*: channel depth; *U*: mean velocity; U^*: bed shear velocity; *S*: sinuosity index; *DR*: discrepancy ratio.

Sources of data:

(1) Deng *et al.* (2002); (2) Sukhodolov *et al.* (1997); (3) Kashefipour *et al.* (2002); (4) Rowiński *et al.* (2003); (5) Deng *et al.* (2001).

that some universal parameter representing the channel geometry should also be incorporated in the considerations; hence, the sinuosity index was treated as an additional parameter in Cases 1 and 3. When a channel meanders, it influences the dynamics of water flow. The onset of turbulent flow deflects some of the water towards the channels sides. As it reaches the side of the channel, it is reflected back toward the opposite side of the channel. As the water changes sides, it obviously also flows downstream, resulting in a zigzag flow line pattern. Meanders are an essential large-scale feature of river topography and, as such, must influence the mixing process. On the other hand, the treatment of sinuous natural channels by means of the longitudinal advection–dispersion model may incorporate large errors. Especially in strongly curved channels, transverse mixing becomes extremely important (Boxall *et al.*, 2003).

Three different sets of data described in Table 1 were mostly taken from the literature. The works from which the data were adapted are cited so that the original sources can be identified. The first data set (Cases 1 and 2) was taken from Deng *et al.* (2002) and includes 70 subsets of field data measured on 30 streams in the United States. Owing to scarcity of data, only 10 measurements were put aside as a verification set. The second data set (Cases 3 and 4) was formed by the first data set extended over 11 river reaches from the territory of Poland and Moldova (Sukhodolov *et al.*, 1997), which also broadened the geographical area for the considerations. The streams added are rather small with relatively small values of observed longitudinal dispersion coefficients. In these cases, more data were omitted during training, but because of the large range of E_L to be covered by the learning set data, the training set was again kept bigger than the verification one. The third set of data (Case 5), taken from Kashefipour *et al.* (2002), Rowiński (2002), Sukhodolov *et al.* (1997) and Deng *et al.* (2001), covers a wider range of dispersion coefficients (99 sets of field data). However, the information about sinuosity index described below was not available for this particular data set.

All networks described in this section were trained upon logarithms of the appropriate data. The results for particular cases were compared by means of:

(a) the discrepancy ratio:

$$DR = \log\left(\frac{E_{lp}}{E_{lm}}\right) \tag{11}$$

where E_{lp} is the predicted longitudinal dispersion coefficient, and E_{lm} is the measured longitudinal dispersion coefficient; and

(b) the percentage of mean error calculated according to the formula used by Kashefipour *et al.* (2002):

$$\delta = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|DR_i\right|}{\frac{1}{N}\sum_{i=1}^{N}\log E_{lmi}}\cdot 100 \tag{12}$$

where N is the number of computed river reaches (see Table 1).

As suggested above and as follows from a variety of studies (Guymer, 1998; Deng *et al.*, 2002; Rowiński *et al.*, 2003), some characterization of planform geometry

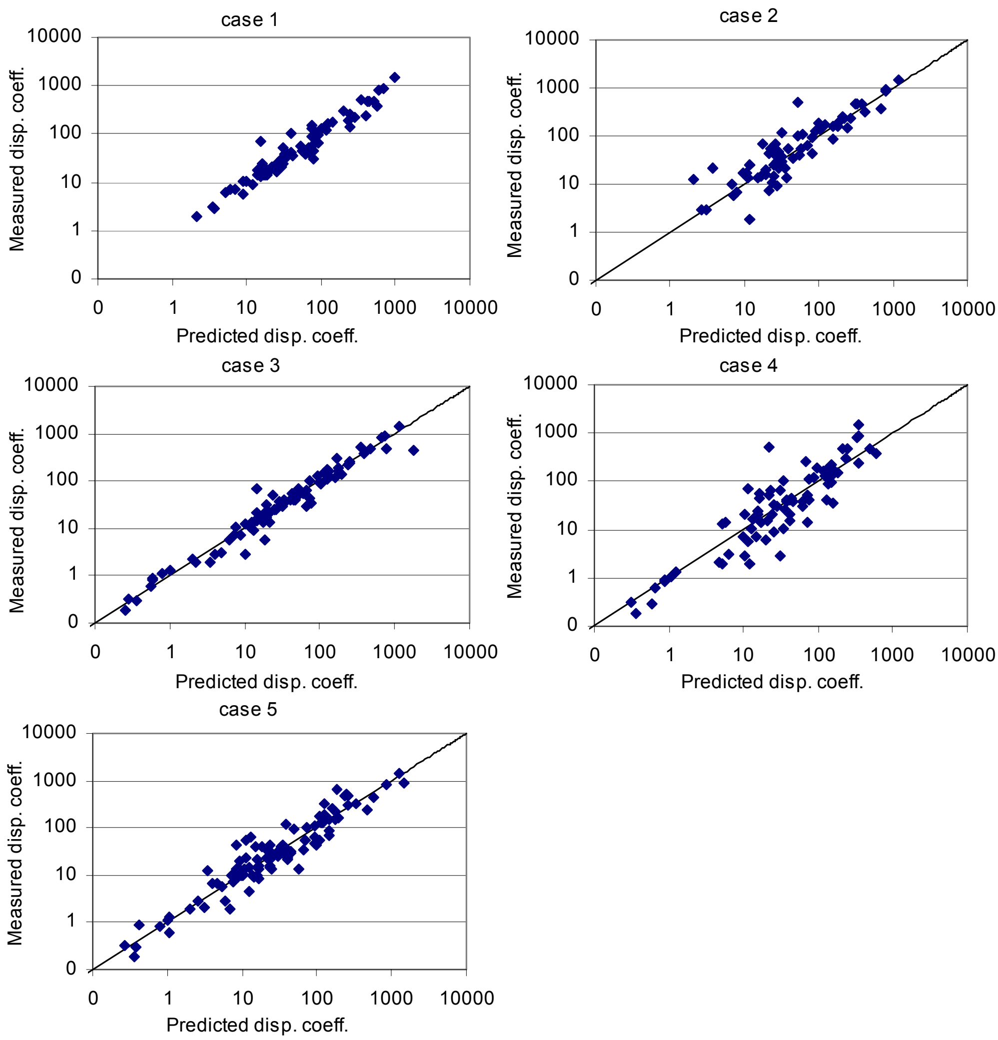


Fig. 2 Comparison of predicted and measured E_L (m^2 s^{-1})—logarithmic scale.

should influence the results significantly. The two first cases (Cases 1 and 2) are based on Deng *et al*. (2002) data. Those authors developed a new analytical method for prediction of longitudinal dispersion coefficient and applied it to a set of field data from 70 river reaches. The results obtained achieved an accuracy in which 91.4% of the calculated values ranged from 0.5 to 2 times the observed values, which corresponds to discrepancy ratio of –0.3 to +0.3. This is the best among all other literature results, but its main drawback is its complexity coming from an application of approximation methods for triple numerical integration with a set of regression equations. This complexity creates large uncertainty in applications. An ANN working as a kind of a black-box tool proved to provide results not worse than those of Deng *et al*. (2002). In cases based solely on Deng *et al*. (2002) data, the outliers lying beyond the discrepancy ratio (–0.3 to +0.3) are fewer than in the study of Deng *et al*. (2002) when taking into account the verification set. The percentage of mean error computed by means of equation (12) for the training and verification sets dropped to 7.06% in Case 1. For Case 1 also, the frequency of occurrence of discrepancy ratios for predicted to measured longitudinal dispersion (Fig. 3) are not worse and are mostly in

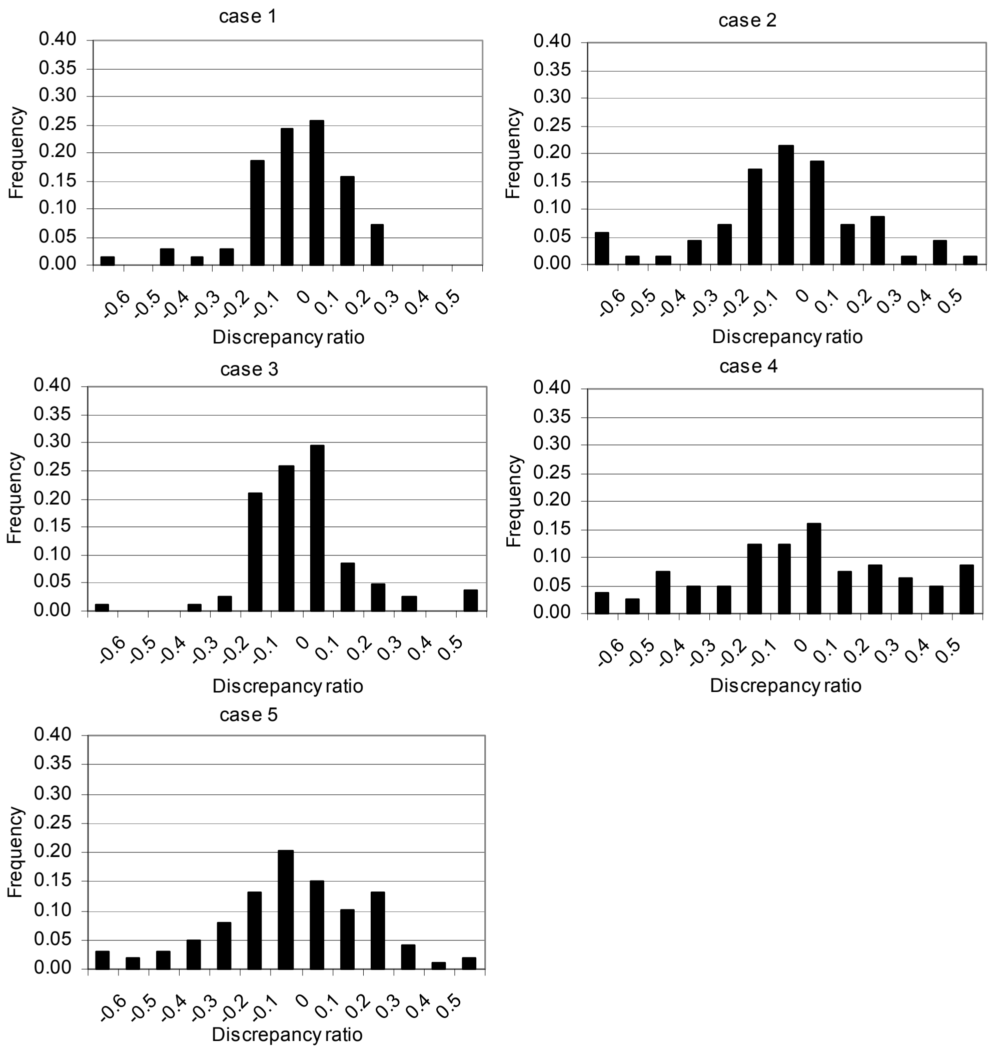
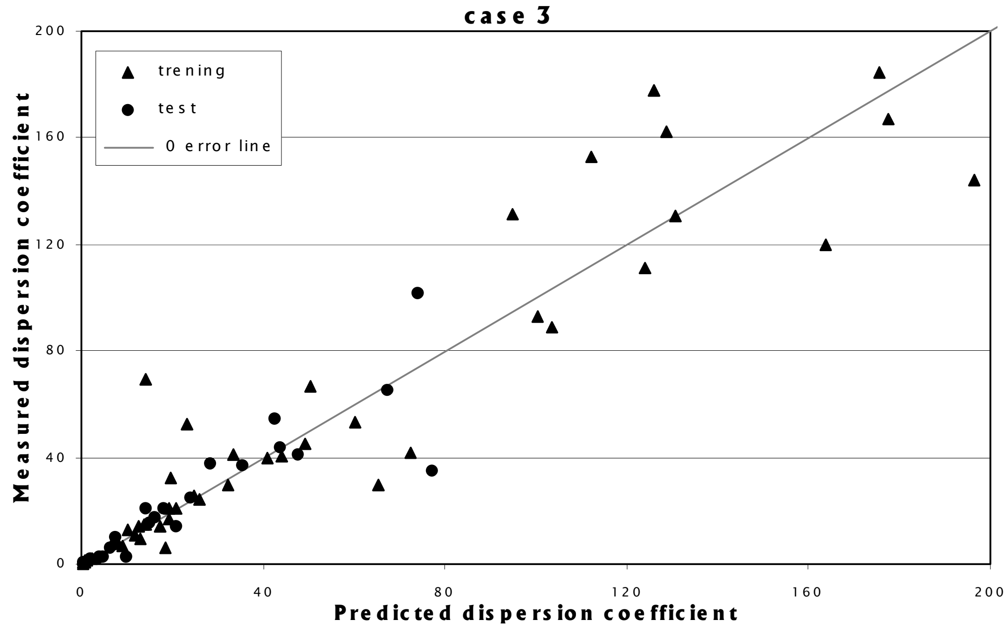


Fig. 3 Frequency of logarithms of discrepancy ratios, predicted to measured E_L ($m^2\ s^{-1}$).

the range –0.2 to 0.2. Note that for Case 2, where an additional parameter reflecting the meandering of the river is not taken into account, the results obtained are much worse (see Table 1, Figs 2 and 3).

Note that, as depicted in Fig. 3, the frequency of occurrence of discrepancy ratios of predicted to measured longitudinal dispersion (obtained for logarithms of the dispersion coefficients ratios (Deng *et al.*, 2002)) is a good representation of the model results.

For Case 3, where the additional data were taken from small streams (Sukhodolov *et al.*, 1997), the results deteriorated slightly but still were much better than in cases without sinuosity index (see Table 1, Figs 2 and 3). They were also at the level achieved by Deng *et al.* (2002), which may be considered as the best at the present time. This clearly shows that the given network works reasonably well and is not overtrained. It therefore seems that such a network can be used for predictions of longitudinal dispersion coefficient for independent data collected in different streams. A good performance of such a network is also shown in Fig. 4, where the comparisons between predicted and computed coefficients E_L are shown in a linear coordinate



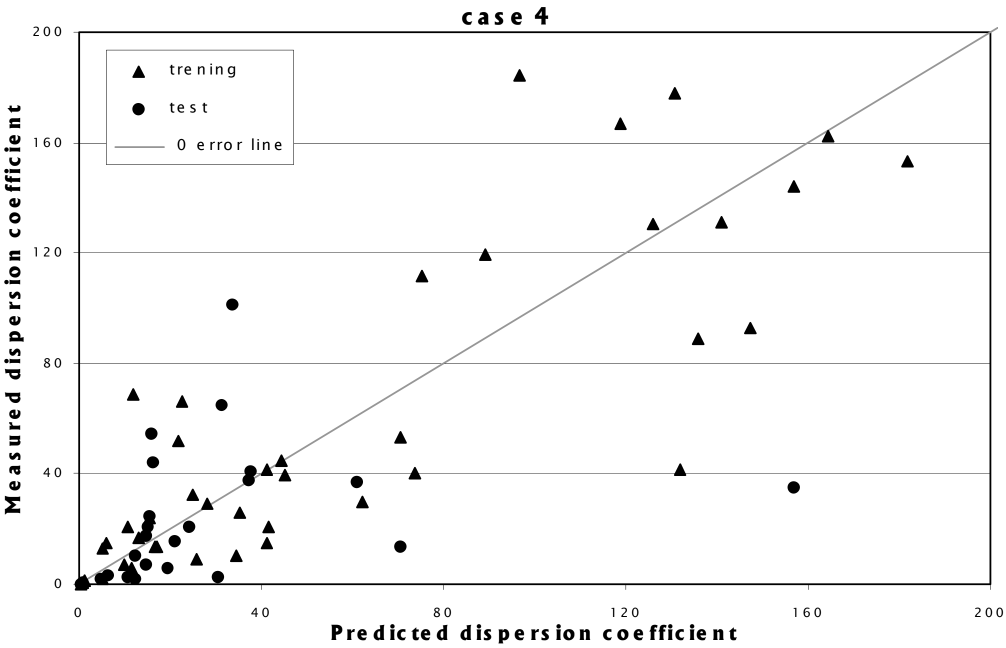


Fig. 4 Comparison of predicted and measured E_L (m^2 s^{-1})—linear scale.

system for Cases 3 and 4 (with dispersion coefficients less than 200 taken into account, only because of the size of figure). In Case 5, where the data cover a wider range of dispersion coefficient (99 sets of field data), the network performs better than in Case 4, but still worse than in cases with sinuosity index.

To summarize, all the cases varied in terms of the balance between the number of training and verification data. The usual practice is to accept quite a large training data set due to a scarcity of longitudinal dispersion coefficients of acceptable quality. But all of the results performed well for the verification data, particularly in cases when the sinuosity index (*S*) was taken into account. Analysing the results in Table 1 and Fig. 2,

Table 2 Comparison of importance of different input variables.

Omitted input variable in data set	Training sets: Mean rank of 10 networks	Range of percentage of mean error (%)	Verification sets: Mean rank of 10 networks	Range of percentage of mean error (%)
B/*H*	9.4	7.6–11.4	9.7	11.4–17.9
U/*U*∗	11.6	8.8–12.2	11.3	13.9–19.4
S	25.5	13.4–17.2	28.7	28.0–33.9
3*UB*	35.5	22.0–38.6	32.3	29.7–53.1

it is clear that performance of ANN models for Cases 1 and 3 is much better than that for Cases 2 and 4. This shows that the sinuosity index plays an important role as the input variable and should be taken into account in predicting E_L by means of empirical models.

Additional computations were performed to evaluate the significance of all input variables. The input data and the method of division into training and verification parts was made as in Case 3 (Table 1). Excluding consecutively one out of four input variables in each of 10 new tests, 40 networks were trained (Table 2). These 40 networks were ranked from the best (rank 1) to the poorest one (rank 40) based on the percentage of mean error (equation (12)). This ranking was performed separately for training and verification sets. The mean rank of each 10 networks indicates the importance of the excluded input variable. The bigger the mean rank is, the poorer the model, so the omitted variable is more important. From Table 2 it is clear that 3*UB* and sinuosity index (*S*) are the most important input variables.

CONCLUSIONS

Although the results obtained with the use of artificial neural networks are not fully satisfying, they are more accurate and far less costly than physically-based models allowing for the prediction of longitudinal dispersion coefficient and, consequently, the pattern of pollution spread in rivers. The authors are by no means recommending eliminating the physically-based models, which allow one to gain more understanding of the behaviour of a system. However, the neural networks may be very useful in situations where the local data cannot be easily provided. The ANNs can be readily integrated with, for example, decision support systems and, as such, can constitute a very useful tool for decision makers. The performance of neural networks methodology was very much improved when the input data were extended by river sinuosity index and the results turned out to be better than those based on any other method.

The authors believe that the present study may bring new points to the discussion about the applicability of artificial neural networks in water-related problems and may add an argument in weakening the reluctance of the wider hydrological community to the apply this methodology. The similar aim, though from another perspective, has been a guiding principle of recent studies of, for example, Wilby *et al*. (2003).

Acknowledgements Funding for this research was provided by The Sendzimir Foundation, USA and by Polish Committee for Scientific Research (KBN) (Grant no. 6P04D02020).

REFERENCES

Boxall, J. B., Guymer, I. & Marion, A. (2003) Transverse mixing in sinuous natural open channel flows. *J. Hydraul. Res*. **41**(2), 153–165.

Czernuszenko, W. & Rowiński, P. M. (1997) Properties of the dead-zone model of longitudinal dispersion in rivers. *J. Hydraul. Res*. **35**(4), 491–504.

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

Deng, Z. Q., Bengtsson, L. & Singh, V. P. (2002) Longitudinal dispersion coefficient in single-channel streams. *J. Hydraul. Engng ASCE* **128**(10), 901–916.

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

Guymer, I. (1998) Longitudinal dispersion in a sinuous channel with changes in shape. *J. Hydraul. Engng ASCE* **124**(1), 33–40.

Haykin, S. (1994) *Neural Networks, a Comprehensive Foundation*. Macmillan College Publishing Co., New York, USA.

Holley, E. R. & Jirka, G. H. (1986) Mixing in rivers. Technical Report E-86-11, Environmental and water quality operational studies. US Army Corps of Engineers, Washington DC, USA.

Hsu, K., Gupta, H. V. & Sorooshian, S. (1995) Artificial neural network modeling of the rainfall–runoff process. *Water Resour. Res*. **31**(10), 2517–2530.

Jirka, G. H. (2004) Mixing and dispersion in rivers. In: *River Flow 2004* (ed. by M. Greco, A. Carravetta & R. Della Morte), 13–27. A.A. Balkema Publishers, a member of Taylor & Francis Group plc, London.

Kashefipour, S. M., Falconer, R. A. & Lin, B. (2002) Modeling longitudinal dispersion in natural channel flows using ANNs. In: *River Flow 2002* (ed. By D. Bousmar & Y. Zech), 111–116. A.A. Balkema/Swets & Zeitlinger, Lisse, The Netherlands.

Korbicz, J., Obuchowicz, A. & Umiński, D. (1994) *Artificial Neural Networks, Foundations and Applications*. Akademicka Oficyna Wydawnicza PLJ, Warszawa (in Polish).

Lees, J. M., Camacho, L. A. & Chapra, S. (2000) On the relationship of transient-storage and aggregated dead zone models to solute transport in streams. *Water Resour. Res.* **36**(1), 213–224.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1990) *Numerical Recipes in C. The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK.

Rochusaar, L. & Paal, L. (1970) The results of the experimental investigation of the longitudinal dispersion coefficient in open channels. *Proc. Tallin Politech. Inst.* A **298**(6).

Rowiński, P. M. (2002) "Constituent transport" in fresh surface waters. In: *Encyclopaedia of Life Support Systems* (ed. by J. C. I. Dooge). EOLSS Publishers Co., Oxford, UK. http://www.eolss.net.

Rowiński, P. M., Napiórkowski, J. J. & Owczarczyk, A. (2003) Transport of passive admixture in a multi-channel river system—the Upper Narew case study. Part 2. Application of dye tracer method. *Ecohydrol. Hydrobiol.* **3**(4), 381–388.

Rowiński, P. M., Dysarz, T. & Napiórkowski, J. J. (2004) Estimation of longitudinal dispersion and storage zone parameters. In: *River Flow 2004* (ed. by M. Greco, A. Carravetta, R. Della Morte), 1201–1210. A.A. Balkema Publishers, a member of Taylor & Francis Group plc, London.

Sukhodolov, A. N., Nikora, V. I., Rowiński, P. M. & Czernuszenko, W. (1997) A case study of longitudinal dispersion in small lowland rivers. *Water Environ. Res.* **69**(7), 1246–1253.

Wilby, R. L., Abrahart, R. J. & Dawson, C. W. (2003) Detection of conceptual model rainfall–runoff processes inside an artificial neural network. *Hydrol. Sci. J.* **48**(2), 163–181.

Wörman, A. (2000) Comparison of models for transient storage in small streams. *Water Resour. Res.* **36**(2), 455–468.

Received 20 July 2004; accepted 26 October 2004