

PREDICTION OF DISPERSION COEFFICIENTS IN A SMALL STREAM USING ARTIFICIAL NEURAL NETWORKS

Wallis Steve G.⁽¹⁾, Piotrowski Adam⁽²⁾, Rowinski Pawel M.⁽²⁾, Napiorkowski Jaroslaw J.⁽²⁾

⁽¹⁾School of the Built Environment, Heriot-Watt University, Edinburgh, EH14 4AS, Edinburgh, UK
phone: +44 (0)131 451 3149; fax: +44 (0)131 451 4617; e-mail: s.g.wallis@hw.ac.uk

⁽²⁾Institute of Geophysics, Polish Academy of Sciences, ul. Ks. Janusza 64, 01-452 Warszawa, Poland
phone: +48 22 69 15 804; fax: +48 22 69 15 915; e-mail: adampp@igf.edu.pl

ABSTRACT

In this paper application of Multi-Layer Perceptron (MLP) Neural Networks to the prediction of dispersion coefficients in a small stream (the Murray Burn in Edinburgh, UK) is presented. Data from eighteen tracer experiments performed in that river are used to test the MLP networks, which were trained and validated on a database of information from other published work. Results from the MLP networks are compared with results from other techniques, such as: the method of moments applied to complete concentration distributions and to distributions with tails truncated at 10% of the peak value; an example empirical formulae and Fischer's routing procedure. Two different MLP networks are presented, one trained on all the data in the database and another one that used only data from "smaller rivers". The performance of the methods was assessed by comparing the results from each one with those from the routing method. The network trained on "smaller rivers" proved to be the most reliable. Results from the network trained on all the data in the database predicted smaller dispersion coefficients, and they were also smaller than those from Fischer's routing procedure. On the other hand, the method of moments showed the poorest correlation with the routing method. This is caused by a large degree of scatter in the method of moments results, which emanates from it being heavily influenced by the tails of the concentration-time profiles measured in the experimental programme.

Keywords: neural network, dispersion coefficient, stream

1 INTRODUCTION

The task of predicting the fate of pollutants that enter watercourses is undertaken by the water industry throughout the world. Although new ideas and models are proposed from time to time, few have made any real impact, and the majority of current best practice relies on the pioneering work of Taylor (1954) and Fischer (1967). In this approach, pollutant transport predictions are based on the Advection-Dispersion Equation, satisfactory application of which rests on an ability to estimate dispersion coefficients for rivers and streams. Despite much work that has improved our understanding of the physical processes that cause dispersion, the prospect of robustly evaluating dispersion coefficients across a wide range of hydraulic conditions still evades us. A recent development that shows some promise for improving this situation, however, lies in the application of Multi-Layer Perceptron (MLP) Neural Networks (Kashefipour et al., 2002; Rowinski et al., 2005; Tayfur & Singh, 2005).

This paper contributes to the above topic by describing the application of MLP networks to the prediction of dispersion coefficients in a small stream, for which data from a programme of tracer experiments were available. The aim of the paper is to examine the performance of the MLP networks by comparing their predictions against those derived from several other popular methods. Section 2 reviews existing methods for predicting dispersion coefficients, Section 3 provides background on MLP models and Section 4 presents details of the experimental programme. The results are presented and discussed in Section 5, where a

new equation for predicting dispersion coefficients in small streams is proposed, and conclusions are drawn in Section 6.

2 BACKGROUND

The transport of a conservative soluble pollutant along a uniform channel may be described by the following well-known Advection-Dispersion Equation:

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} = D \frac{\partial^2 C}{\partial x^2} \quad (1)$$

where C is the cross-sectional average solute concentration, U is the cross-sectional average longitudinal flow velocity, D is the longitudinal dispersion coefficient, x is the longitudinal co-ordinate and t is time. In this equation U quantifies the downstream movement of solute and D quantifies the longitudinal spreading (dispersion) of solute. Several assumptions underlie the derivation of this equation, principally that the solute should be cross-sectionally well mixed and that the cause of the longitudinal spreading is a shear flow dispersion mechanism. In regard to the latter, it is also assumed that the tendency of cross-sectional gradients of longitudinal velocity to create cross-sectional gradients of solute concentration is in equilibrium with the tendency of cross-sectional mixing (due to turbulence and secondary currents) to erode those concentration gradients (Taylor, 1954). In rivers and streams Fischer (1967) demonstrated that dispersion occurs predominantly by the interaction of transverse velocity gradients and transverse mixing, the contribution from the equivalent vertical flow structure being small.

Clearly, to apply equation (1) to a practical scenario the dispersion coefficient for the reach is required. Since dispersion coefficients cannot be measured in-situ directly from a simple individual measurement, several techniques have been devised for estimating them. The four most common approaches are the method of moments; Fischer's routing procedure; empirical equations and numerical integration of the flow structure. Details of these are commonplace in the literature (appearing, for example, in Fischer et al., 1979; Rutherford, 1994; Singh & Beck, 2003; Wallis & Manson, 2004). The first three of these techniques rely directly or indirectly on observations of the dispersion process that are provided by solute tracer experiments. The experiments provide solute concentration-time profiles at the ends of the reach under study, the analysis of which enables the dispersion coefficient to be estimated. Since these three methods are used later in the paper, the essential ingredients of them are given below. The flow structure integration method is less popular because it requires detailed measurements of, or estimates of, the transverse profiles of (a) (depth averaged) longitudinal velocity and (b) transverse mixing rates. The method is not used in this paper, because the required data were not available.

2.1 METHOD OF MOMENTS

This is the traditional method of estimating dispersion coefficients from a pair of observed concentration-time profiles. It involves calculating the variances and centroids of the profiles using their first two temporal moments. However, it is well known that the method can be unreliable because the results are very sensitive to the information in the tails of the concentration-time profiles. In an attempt to reduce the resulting uncertainty on the moments, truncation of the profiles at the point where the concentration drops below 3% of the peak concentration has been reported (Mazijk & Veling, 2005). Here, truncation at the point where the concentration dropped below 10% of the peak concentration was used, referred to below as the truncated method of moments. Jobson (1996) reports the use of 10% truncation, but for other reasons.

2.2 FISCHER'S ROUTING PROCEDURE

This method involves optimising the dispersion coefficient to obtain a best fit between a predicted downstream concentration-time profile and an observed downstream concentration-time profile. The predicted downstream profile is computed by routing the observed upstream concentration-time profile through the reach using an analytical solution of equation (1). The velocity in equation (1) can be optimised at the same time or it can be assumed that the velocity is known from the value given by dividing the reach length by the centroid travel time of the solute cloud. The latter method was used in the current work.

2.3 EMPIRICAL EQUATIONS

Several workers have proposed empirical equations that have been derived by correlating observed values of dispersion coefficients to bulk flow hydraulic parameters. In the majority of cases the observed dispersion coefficients used were estimated from tracer experiments and the method of moments, but recently some workers (for example, Seo & Cheong (1998)) have preferred the use of Fischer's routing procedure for estimating the dispersion coefficients because this reduces the uncertainty in the values caused by the difficulties of dealing with the tails of the concentration-time profiles. Most of these empirical equations take the following non-dimensional form:

$$\frac{D}{HU_*} = F \left(\frac{U}{U_*} \right)^a \left(\frac{B}{H} \right)^b \quad (2)$$

where B is the channel width, H is the mean depth of flow, U_* is the shear velocity, a and b are constants (>1) and F is a function that ranges in complexity from a simple constant (for example, Fischer (1975); Seo & Cheong (1998)) to an expression involving measures of one or more of channel shape, transverse mixing and channel sinuosity (for example, Seo & Baek (2002); Deng et al. (2002)). The constants a and b are both frequently 2, but not in every case. The nature of equation (2) exposes the roles of channel aspect ratio and friction (through the shear velocity). Increases in the former tend to increase dispersion (Rutherford, 1994), while rougher channels tend to create more turbulence, which by encouraging more cross-sectional mixing, tends to reduce dispersion. One of the latest and most promising of these equations has been used in this work, namely that due to Deng et al. (2001). In this, $a = 2$, $b = 5/3$ and $F = (0.15/(8k))$, where k is the transverse mixing coefficient that is evaluated using the following equation.

$$k = 0.145 + \left(\frac{1}{3520} \right) \left(\frac{U}{U_*} \right) \left(\frac{B}{H} \right)^{1.38} \quad (3)$$

3 ARTIFICIAL NEURAL NETWORKS

Multi-Layer Perceptron (MLP) Neural Networks are well known “universal approximators” and have been successfully applied for dispersion coefficient assessment in recent years (Kashefipour et al., 2002; Rowinski et al., 2005; Tayfur & Singh, 2005; Piotrowski et al., 2006). In the current work MLP networks, with one hidden layer and with a sigmoidal activation function applied only to hidden nodes, were employed. The networks used the following data: reach mean values of channel width (B), channel depth (H), flow velocity (U), shear velocity (U_*) and river sinuosity index (sin), arranged in the form of B/H , U/U_* , $3UB$, sin . A database composed of measured dispersion coefficients and various bulk flow hydraulic parameters from 81 river reaches published in Deng et al. (2001) and

Sukhodolov et al. (1997) was available. These 81 data were divided into training (50) and validation (31) sets, as in Rowinski et al. (2005) and Piotrowski et al. (2006). The networks were trained by means of the Levenberg-Marquardt algorithm (Haykin, 1994) with a multi-start approach.

The first neural network (MLPA) was trained and verified on all the data in the database, while the second neural network (MLPS) was trained and verified on a sub-set of the database that excluded large rivers, i.e. those with widths $> 100\text{m}$ or dispersion coefficients $> 100\text{m}^2/\text{s}$. Thus this second network was trained on "small" rivers only. Clearly, these numerical boundaries for width and dispersion coefficient are rather arbitrary, and in view of the size of the stream used in this study (see Section 4), many of the rivers in the truncated database are still "large". However, there were not enough "small" rivers in the database to optimise the network effectively if the values defining these boundaries were reduced.

Both MLP networks, trained and verified on the Deng et al. (2001) and Sukhodolov et al. (1997) data base, were then applied to a set of tracer experiments performed on the Murray Burn, Edinburgh. Since these data had not been used during the training of the MLP networks, this use of the networks was a truly independent test of their ability to predict dispersion coefficients.

4 APPLICATION OF DISPERSION COEFFICIENT PREDICTORS

Data for this study were derived from a programme of twenty-six tracer experiments undertaken in the Murray Burn, which is a stream that flows through the Heriot-Watt University Campus at Riccarton in Edinburgh. The experiments were conducted in a 0.5 km reach of the stream, and each experiment consisted of the (gulp) injection of a known mass of Rhodamine WT dye followed by the measurement of solute concentration-time profiles at four sampling sites further downstream (Burke, 2002). The profiles were obtained by collecting water samples from the central part of the cross-section. Tracer concentrations in the samples were determined under laboratory conditions using a single calibrated Turner Designs fluorometer and allowing for temperature effects. The sampling interval was matched to the flow rate of each experiment, with the aim of capturing well-resolved profiles. Typically, sampling intervals of 30s, 60s and 120s were used for high, medium and low flows, respectively.

Dispersion coefficients were derived from the six methods referred to earlier: the two artificial neural networks; the two methods of moments; Fischer's routing procedure; and the empirical equation given by Deng et al. (2001). Predictions were obtained for the first sub-reach of the Murray Burn, i.e. the reach between the first and the second sampling sites. This reach is 137m long with a mean width of 3.7m and a mean longitudinal slope of 0.025. It contains two bends and has a sinuosity of 1.08. The stream bed is covered with cobbles and boulders of nominal sizes ranging between the order of 1cm and the order of 15cm. Eighteen sets of tracer concentration-time data were available in the first sub-reach from the experimental programme (these covered the flow range $0 - 1\text{m}^3/\text{s}$). Only cases containing complete (or very nearly complete) concentration-time profiles were included in the analysis. For each experiment, the flow rate was estimated from the tracer data using dilution gauging applied to the concentration-time curve at the first sampling site, by dividing the mass of tracer used by the area under the concentration-time profile.

Other data required for the application of the methods described in Sections 2 and 3 were obtained as follows. The channel width was derived from measurements of the stream channel. These were undertaken at one (medium) flow rate only, and so no information was available on any variations of width with flow rate. The average width of the channel was used for all experiments in all the analyses. The flow velocity for each experiment was estimated by dividing the reach length by the travel time of the centroid of the tracer cloud

between sampling sites 1 and 2. For each experiment, the channel depth was estimated by dividing the cross-sectional area of the flow by the mean channel width, with the cross-sectional area of the flow being calculated from the flow rate and the flow velocity.

5 RESULTS AND DISCUSSION

Results from the six methods described above are presented in Table 1 for each of the eighteen tracer experiments referred to earlier. Generally, the dispersion coefficients are very small (being of the order of $1 \text{ m}^2/\text{s}$), which befits the size of the stream, and they increase with flow rate, which is consistent with other studies (Rutherford, 1994). In order to compare the results from the different methods it was assumed that the values from Fischer's routing procedure were the most reliable ones. This was based on a number of considerations. Firstly, the method gave consistently good fits to the downstream data: an example is shown in Figure 1. Secondly, the nature of the method is that dispersion coefficients derived from it are representative of the change in shape of the entire concentration-time profile as it passes through the reach: it is not, for example, excessively influenced by the tails. Finally, several other workers have adopted the method as their preferred way of estimating dispersion coefficients from tracer data of the type being used here (for example, Seo & Cheong, (1998)). It is interesting to note that Figure 1 illustrates a general trend found in most of the routing procedure results, namely that the peak concentration and the tail were reproduced well, but the routed peak showed a small phase lag compared to the observed peak.

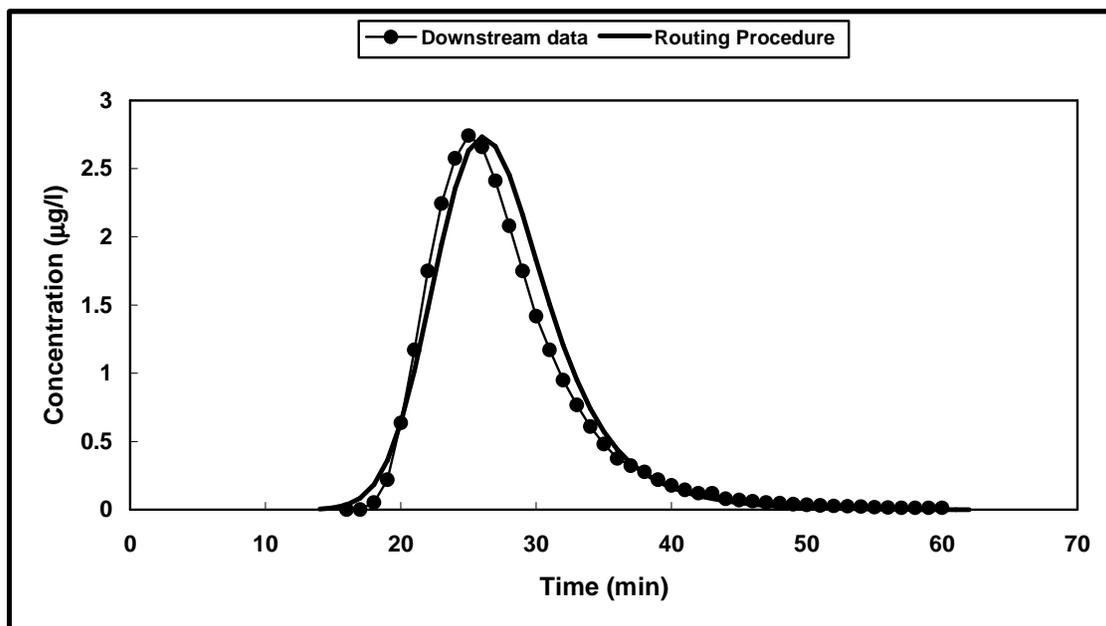


Figure 1. Result of Fischer's routing procedure for Experiment 10.

The performance of the other five methods is now assessed by comparing the results from each method with those from the routing method. Table 2 presents some statistics that summarise these comparisons. The statistics, which were evaluated over the eighteen experiments, were derived as follows:

- Mean of ratio – this is the average value of the ratio of the dispersion coefficient from a method to the corresponding dispersion coefficient from the routing method

- Standard deviation of ratio - this is the standard deviation of the ratio of the dispersion coefficient from a method to the corresponding dispersion coefficient from the routing method
- RMSE – this is the root mean square value of the error (or difference) between the dispersion coefficient from a method to the corresponding dispersion coefficient from the routing method
- R – this is the correlation coefficient between the dispersion coefficients from a method and those from the routing procedure.

Table 1. Predicted dispersion coefficients from six methods: MLPA = neural network using complete database; MLPS = neural network using database with large rivers excluded; Deng = Deng et al. (2001); MOM = method of moments; TMOM = truncated method of moments; RP = Fischer's routing procedure.

		Dispersion Coefficient (m ² /s)					
Experiment	Flow (l/s)	MLPA	MLPS	Deng	MOM	TMOM	RP
2	68.0	0.43	0.72	1.01	0.82	0.58	0.59
4	43.6	0.38	0.59	0.78	0.91	0.43	0.43
5	47.5	0.39	0.58	0.75	1.16	0.46	0.43
6	128.5	0.52	0.97	1.56	1.62	0.80	0.83
7	134.2	0.53	1.04	1.75	1.02	0.67	0.88
8	45.9	0.39	0.63	0.84	0.28	0.60	0.38
9	35.3	0.36	0.62	0.83	0.06	0.75	0.52
10	56.3	0.41	0.68	0.93	1.07	0.76	0.58
15	49.5	0.40	0.73	1.03	0.41	0.61	0.56
16	15.6	0.30	0.29	0.34	0.78	0.34	0.29
17	13.9	0.29	0.31	0.37	0.65	0.52	0.38
18	33.0	0.36	0.47	0.58	0.74	0.51	0.50
20	261.0	0.69	1.36	2.86	3.84	2.24	2.13
21	162.1	0.57	1.13	2.03	1.98	0.58	0.72
22	257.5	0.67	1.28	2.52	3.08	1.45	1.75
23	62.1	0.42	0.66	0.88	0.65	0.55	0.45
24	535.4	0.91	1.45	2.97	2.99	1.81	1.92
26	952.4	2.84	1.97	6.36	0.48	2.20	2.74

Table 2. Summary statistics of the performance of five methods in relation to the results of Fischer's routing procedure.

Statistic	MLPA	MLPS	Deng	MOM	TMOM
Mean of ratio	0.74	1.13	1.73	1.56	1.06
Standard deviation of ratio	0.21	0.31	0.45	0.78	0.25
RMSE	0.51	0.34	1.04	0.91	0.19
R	0.80	0.94	0.94	0.59	0.97

Some clear trends are apparent in the results. For example, the empirical equation of Deng et al. (2001) and the method of moments significantly over predict the dispersion coefficient,

while the first artificial neural network (MLPA) significantly under predicts it. The truncated method of moments and the second artificial neural network (MLPS) both over predict the coefficient, but to much smaller extents than the empirical equation of Deng et al. (2001) or the method of moments. The two former methods also show the highest correlation with, and the smallest root mean square error with, the routing method results.

It is worth noting that the method of moments shows the poorest correlation with the routing method. This is caused by a large degree of scatter in the method of moments results, which emanates from the heavy influence of the tails on this method. Even when the tails are complete, defining where they finish is far from easy and can have a surprisingly large influence on the dispersion coefficients. In comparison with the method of moments, the results from the truncated method of moments are much more consistent, showing a much reduced scatter and, interestingly, much closer agreement with the routing procedure results.

A useful way of portraying the dispersion coefficients is in the non-dimensional form of equation (2). Hence, Figure 2 shows the method of moments results, and Figure 3 shows the artificial neural network results plotted in this way, assuming that the constants a and b are both 2. Results from Fischer's routing method are shown on both figures. Figure 2 reinforces the earlier comments on the reduced scatter in the method of moments results when the tails of the concentration-time profiles are truncated. Figure 3 shows the closer agreement with Fischer's routing method of the MLPS network compared to the MLPA network. It is also noticeable that the scatter is also much reduced. Indeed, the results from the MLPS network show a good linear trend, and the linear trend line (constrained to pass through the origin) is shown on the figure. The equation of this line ($R^2 = 0.83$) is proposed as a new predictor of dispersion coefficient in small streams, and is given by:

$$\frac{D}{HU_*} = 0.017 \left(\frac{U}{U_*} \right)^2 \left(\frac{B}{H} \right)^2 \quad (4)$$

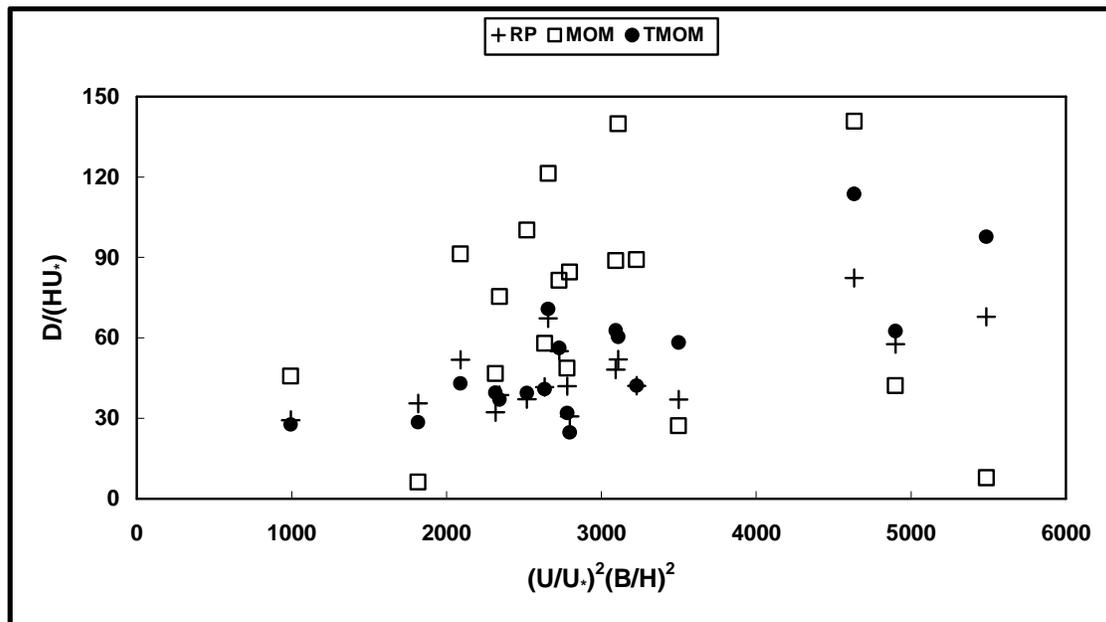


Figure 2. Non-dimensional dispersion coefficients: method of moments and routing procedure.

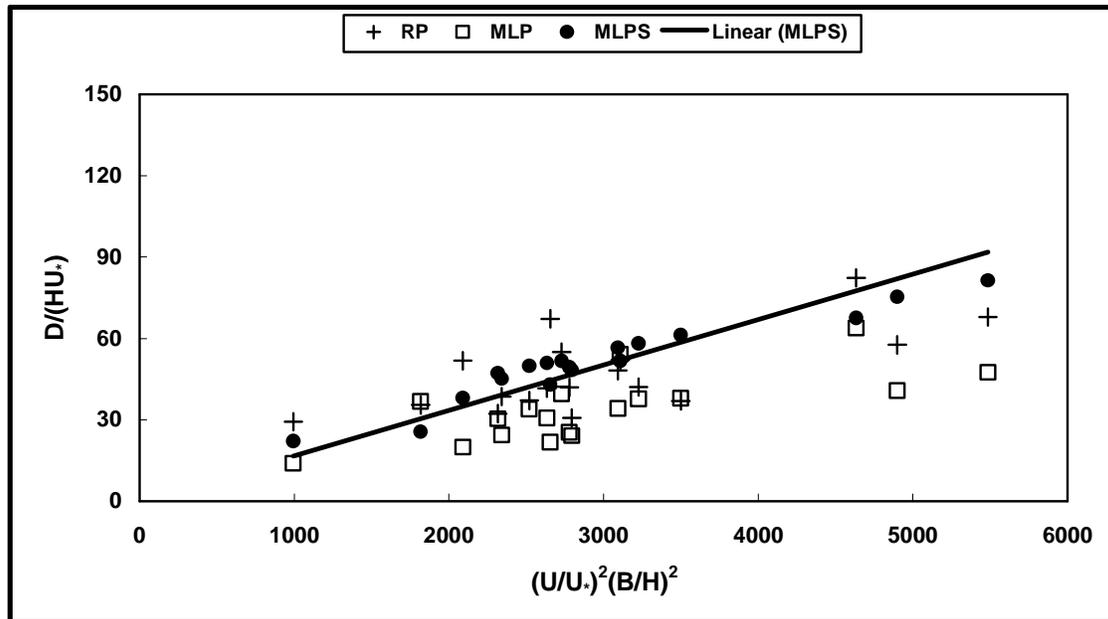


Figure 3. Non-dimensional dispersion coefficients: artificial neural networks and routing procedure.

It is interesting to note that equation (4) is quite similar to the first empirical equation proposed for predicting dispersion coefficients (Fischer, 1975) in which the numerical constant is 0.011.

6 CONCLUSIONS

Dispersion coefficients for a reach of a small stream (the Murray Burn) have been estimated using six methods. Comparisons of the results lead to the following conclusions:

- A neural network trained on "small" rivers (MLPS) from a database of observed cases of dispersion yields a robust predictor of dispersion coefficients for the Murray Burn
- Results from a neural network trained on all the data in the database (MLPA) predicted coefficients that were smaller than those from MLPS and which were significantly smaller than those from Fischer's routing procedure
- The truncated method of moments gave dispersion coefficients that were significantly more robust than those from the traditional method of moments
- The empirical equation of Deng et al. (2001) significantly over predicted the dispersion coefficients in the Murray Burn.

ACKNOWLEDGMENTS

The work described in this paper was initiated while the second author visited the Institute of Geophysics, Warsaw, in July 2006 under the bilateral International Exchange Programme between the Royal Society of Edinburgh and the Polish Academy of Sciences. Further work was carried out during the 26th International School of Hydraulics in Poland during September 2006. Due acknowledgments for financial support are made to the Royal Society of Edinburgh, the Polish Academy of Sciences, the Institute of Geophysics, Warsaw and the Royal Academy of Engineering, UK. This work was financially supported in part by The Polish Ministry of Higher Education and Science: Grant No. 2 P04D 026 29.

REFERENCES

- Burke, N.J. (2002), Travel time and flow characteristics of a small stream system, PhD Thesis, Heriot-Watt University, Edinburgh (unpublished).
- Deng, Z.-Q., Singh, V.P. and Bengtsson, L. (2001), Longitudinal dispersion coefficient in straight rivers, *Journal of Hydraulic Engineering, American Society of Civil Engineers*, 127(11), 919-927.
- Deng, Z.-Q., Bengtsson L., Singh, V.P. and Adrian, D.D. (2002), Longitudinal dispersion in single channel streams, *Journal of Hydraulic Engineering, American Society of Civil Engineers*, 128(10), 901-916.
- Fischer, H.B. (1967), The mechanics of dispersion in natural streams, *Journal of the Hydraulics Division, American Society of Civil Engineers*, 93(6), 187-215.
- Fischer, H.B. (1975), Discussion of 'Simple method for predicting dispersion in streams' by McQuifer, R.S. and Keefer, T.N., *Journal of the Environmental Engineering Division, American Society of Civil Engineers*, 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 Press, New York.
- Haykin S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Co., New York, USA.
- Jobson, H.E. (1996), Prediction of Traveltime and Longitudinal Dispersion in Rivers and Streams, U.S. Geological Survey Water-Resources Investigations Report, 96-4013.
- Kashefipour, S.M., Falconer, R.A. and Lin, B. (2002), Modelling longitudinal dispersion in natural flows using ANNs, in *Proceedings of River Flow 2002*, pp. 111-116, Louvain-la-Neuve, Belgium.
- Mazijk, A van and Veling, E.J.M. (2005), Tracer experiments in the Rhine Basin: evaluation of the skewness of observed concentration distributions, *Journal of Hydrology*, 307, 60-78.
- Piotrowski, A., Rowinski, P.M. and Napiorkowski, J.J. (2006), Assessment of longitudinal dispersion coefficients by means of different neural networks, in *Proceedings of 7th International Conference on Hydroinformatics*, Nice, France.
- Rowinski, P.M., Piotrowski, A. and Napiorkowski, J.J. (2005), Are artificial neural network techniques relevant for the estimation of longitudinal dispersion coefficient in rivers?, *Hydrological Sciences Journal*, 50(1), 175-187.
- Rutherford, J.C. (1994), *River Mixing*, Wiley, Chichester, UK.
- Seo, I.W. and Cheong, T.S. (1998), Predicting longitudinal dispersion coefficient in natural streams, *Journal of Hydraulic Engineering, American Society of Civil Engineers*, 124(1), 25-32.
- Seo, I.W. and Baek, K.O. (2002), Estimation of longitudinal dispersion coefficient for streams, in *Proceedings of River Flow 2002*, pp. 143-150, Louvain-la-Neuve, Belgium.
- Singh, S.K. and Beck, M.B. (2003), Dispersion coefficient of streams from tracer experiment data, *Journal of Environmental Engineering, American Society of Civil Engineers*, 129(6), 539-546.
- Sukhodolov, A.N., Nikora, V.I., Rowinski, P.M. and Czernuszenko, W. (1997), A case study of longitudinal dispersion in small lowland rivers, *Water Environment Research*, 69(7), 1246-1253.
- Tayfur, G. and Singh, V.P. (2005), Predicting longitudinal dispersion coefficient in natural streams by artificial neural network, *Journal of Hydraulic Engineering, American Society of Civil Engineers*, 131(11), 991-1000.
- Taylor, G.I. (1954), The dispersion of matter in turbulent flow through a pipe, *Proceedings of the Royal Society of London*, A223, 446-468.
- Wallis, S.G. and Manson, J.R. (2004), Methods for predicting dispersion coefficients in rivers, *Proceedings of Institution of Civil Engineers, Water Management*, 157(3), 131-141.