Adaptive neuro-fuzzy inference system for real-time monitoring of integrated-constructed wetlands

Mawuli Dzakpasu, Miklas Scholz, Valerie McCarthy, Siobhán Jordan and Abdulkadir Sani

ABSTRACT

Monitoring large-scale treatment wetlands is costly and time-consuming, but required by regulators. Some analytical results are available only after 5 days or even longer. Thus, adaptive neuro-fuzzy inference system (ANFIS) models were developed to predict the effluent concentrations of 5-day biochemical oxygen demand (BOD_5) and NH_4-N from a full-scale integrated constructed wetland (ICW) treating domestic wastewater. The ANFIS models were developed and validated with a 4-year data set from the ICW system. Cost-effective, quicker and easier to measure variables were selected as the possible predictors based on their goodness of correlation with the outputs. A self-organizing neural network was applied to extract the most relevant input variables from all the possible input variables. Fuzzy subtractive clustering was used to identify the architecture of the ANFIS models and to optimize fuzzy rules, overall, improving the network performance. According to the findings, ANFIS could predict the effluent quality variation quite strongly. Effluent BOD_5 and NH_4-N concentrations were predicted relatively accurately by other effluent water quality parameters, which can be measured within a few hours. The simulated effluent BOD_5 and NH_4-N concentrations well fitted the measured concentrations, which was also supported by relatively low mean squared error. Thus, ANFIS can be useful for real-time monitoring and control of ICW systems.

Key words | ANFIS, constructed wetland, domestic wastewater, fuzzy models, neural network, prediction

Mawuli Dzakpasu
Dooge Centre for Water Resources Research,
School of Civil, Structural and Environmental Engineering,
Newstead Building,
University College Dublin,
Belfield,
Dublin 4,
Ireland

Miklas Scholz (corresponding author)
Abdulkadir Sani
Civil Engineering Research Group,
School of Computing, Science and Engineering,
Newton Building,
The University of Salford,
Greater Manchester M5 4WT,
UK
E-mail: *m.scholz@salford.ac.uk*

Mawuli Dzakpasu
Valerie McCarthy
Siobhán Jordan
Centre for Freshwater and Environmental Studies,
North Building,
Dundalk Institute of Technology,
Dundalk, Co. Louth,
Ireland

INTRODUCTION

The growing interest in the use of constructed wetlands (CWs) for wastewater treatment, along with stringent water quality standards, increasingly stresses the need to develop better process design and monitoring tools (Rousseau *et al*. 2004). More specifically, the biochemical oxygen demand (BOD) is an important parameter that characterizes the biodegradable constituent of organic matter in treatment plant and CW effluents. However, BOD is relatively time-consuming to measure and requires at least 5 days. For real-time water quality control purposes, CWs increasingly require the development of accurate inferential models for predicting the BOD and other parameters for process monitoring and control. Accordingly, multiple efforts have been dedicated to the modelling of CW processes, ranging from simple rules of thumb and regression equations to the well-known first-order k-C* models, Monod-type equations and more complex dynamic and process-based models (Rousseau *et al*. 2004; Kumar & Zhao 2011).

More recently, the use of artificial intelligence (AI) modelling techniques has increased with the recognition of its potential (Chen *et al*. 2008). AI mimics human perception, learning and reasoning to solve complex problems. Their ability to construct non-linear relationships that can explain complex relationships within a data set without the difficult task of dealing with deterministic non-linear mathematics (Hamed *et al*. 2004) makes AI models particularly suitable for modelling the processes within complex environmental systems. More significantly, AI models have the ability to generalize the input–output relationship to produce an output when presented with previously unseen inputs. Furthermore, AI models are data-driven and so there is no

doi: 10.2166/wst.2014.461

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

requirement to specify the mathematical form of the relationship being modelled.

Of the range of AI models (Chen *et al.* 2008), the commonly used techniques are artificial neural networks (ANNs) and fuzzy logic systems (FLS), both of which have played important roles in the development of models for complex environmental systems (Cinar 2005). The fundamental and complementary features of ANNs and FLS have been combined into integrated systems, termed fuzzy neural network (FNN). FNN combines the benefits of both ANN and FLS by bringing together the learning and the computation powers of ANNs and the high-level human-like thinking and reasoning ability of FLS. They are effective: fast, efficient and easily designed, implemented and understood. By combining them, learning in ANNs reduces the need to prime fuzzy systems. In addition, fuzzy systems attenuate 'noise', from which some ANNs suffer (Chen *et al.* 2008).

This paper employed a well-known and easily understood FNN technique, namely adaptive neuro-fuzzy inference system (ANFIS), to demonstrate its feasibility for modelling full-scale CW systems. The aim was to develop an ANFIS model for monitoring a full-scale integrated constructed wetland (ICW) treating domestic wastewater in Ireland. The specific objective was to develop a model for the rapid prediction of 5-day BOD (BOD_5) and NH_4-N, using other cost-effective, quicker and easier to measure water quality variables, to facilitate the use of these parameters for real-time monitoring and control of key water quality parameters within ICW systems.

MATERIALS AND METHODS

Study site description

The ICW treatment system at the centre of the study is located within the walls of the Castle Leslie Estate at Glaslough in County Monaghan, Ireland (06°53′37.94″W, 54°19′6.01″N). Dimensions of the wetland cells are presented in Table 1. The ICW system (Figure 1) was commissioned in October 2007, to treat sewage from Glaslough. The design capacity of the ICW system is 1750 population equivalent. The functional water area of the ICW cells is 3.25 ha within a curtilage area of 6.74 ha.

Influent primary domestic wastewater from the village is pumped directly into a receiving sedimentation pond. From there, the wastewater subsequently flows by gravity sequentially through the five earthen-lined cells. The effluent of the last cell discharges directly into the adjacent Mountain Water River.

Table 1 | Dimensions of the integrated constructed wetland cells at Glaslough, Ireland

ICW section	Area (m^2)	Depth (m)	Total volume (m^3)	Effective volume (m^3)
Sedimentation pond 1	285	0.45	128.3	85.5
Sedimentation pond 2	365	0.45	164.3	109.5
Cell 1	4664	0.42	1958.9	1399.2
Cell 2	4500	0.38	1710.0	1350
Cell 3	12660	0.32	4051.2	3798
Cell 4	9170	0.36	3301.2	2751
Cell 5	1460	0.29	423.4	423.4
Total wetland system	33104	–	11737.3	9916.6

The wetland cells were planted in a club pattern, and the main plants were *Carex riparia* Curtis, *Phragmites australis* (Cav.) Trin. ex Steud., *Typha latifolia* L., *Iris pseudacorus* L., and *Glyceria maxima* (Hartm.) Holmb. Other plants were *Glyceria fluitans* (L.) R.Br., *Juncus effusus* L., *Sparganium erectum* L. emend Rchb, *Elisma natans* (L.) Raf., and *Scirpus pendulus* Muhl.

Adaptive neuro-fuzzy inference system

ANFIS is a Takagi–Sugeno type (Takagi & Sugeno 1985; Jang 1993) fuzzy inference system implemented within the framework of adaptive neural networks. Based on a multilayer feed-forward network, ANFIS uses neural network learning algorithms and fuzzy reasoning to map a set of inputs to the output. In the network structure, part or all of the nodes are adaptive, which means that the output of each node depends on the parameter(s) pertaining to it. The parameter(s) should be changed to minimize a prescribed error measure according to the network's learning algorithm. Each node performs a particular function (node function) on incoming signals as well as providing a set of parameters pertaining to this node. The nature of node functions may vary from one node to the other, and the choice of each node function depends on the overall input–output function that the adaptive network is required to carry out (Jang 1993). To reflect different adaptive capabilities, an ANFIS uses both circle and square nodes, whereby a square (adaptive) node has parameters while a circle (fixed) node has none. The architecture of a typical ANFIS with two inputs, two fuzzy *if-then* rules and one output for the first-order

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

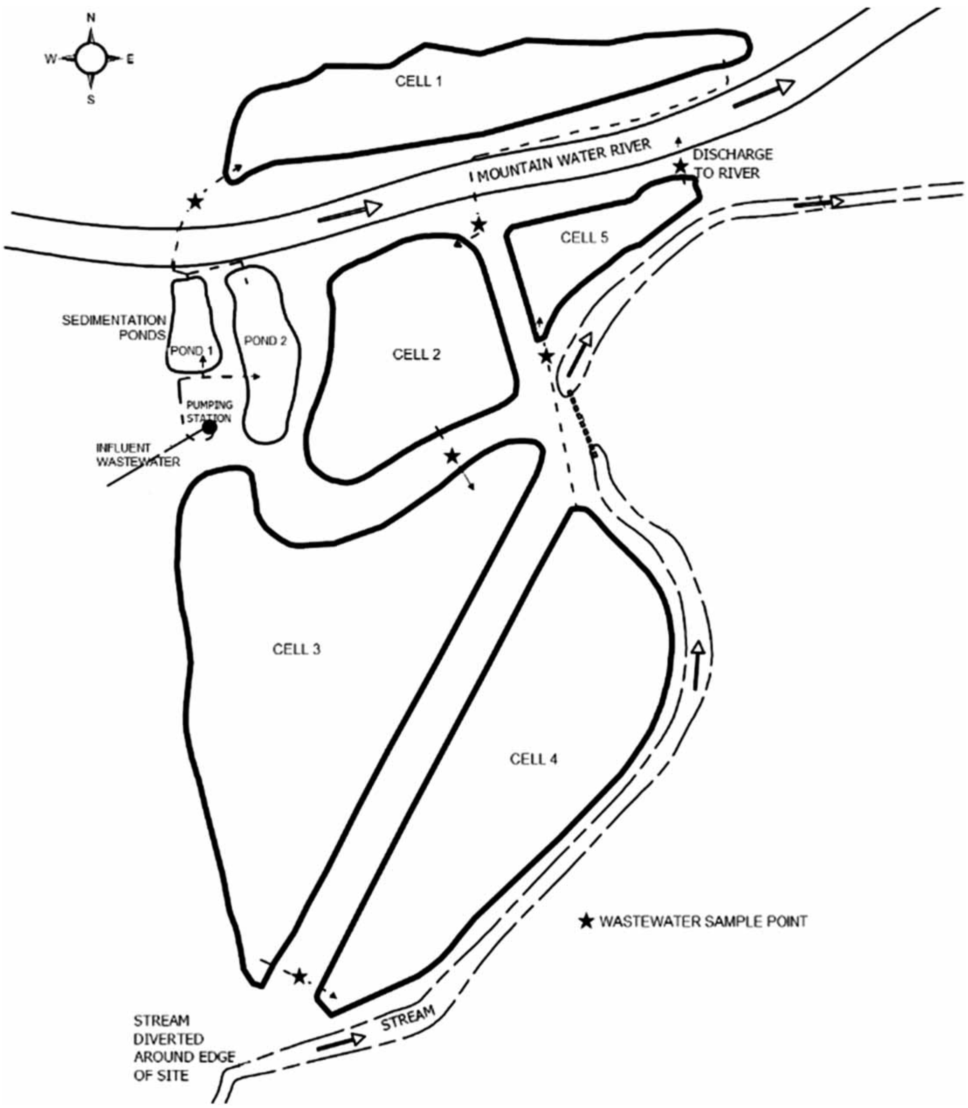

Figure 1 | Schematic diagram of integrated constructed wetland at Glaslough in Ireland.

Takagi and Sugeno type fuzzy model (Takagi & Sugeno 1985), where each input is assumed to have two associated membership functions (MFs), is shown in Figure 2.

For a first-order Takagi–Sugeno fuzzy model (Takagi & Sugeno 1985), a typical rule set with two fuzzy *if-then* rules can be expressed as (Jang 1993; Wang & Elhag 2008)

$$\text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \quad \text{then } f_1 = p_1x + q_1y + r_1 \quad (1)$$

$$\text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \quad \text{then } f_2 = p_2x + q_2y + r_2 \quad (2)$$

where A_1, A_2 and B_1, B_2 are the MFs for the inputs x and y, respectively; p_{ij}, q_{ij} and $r_{ij} = (i, j = 1, 2)$ are consequent parameters.

The architecture of a typical ANFIS consists of five layers (Figure 2), and every node in a given layer performs particular functions in the ANFIS such as *input*, *fuzzification*, *rule inference*, *normalization* and *defuzzification*. A detailed description is provided in Jang (1993).

Model development

Data variables selection

The general idea for the model development was to predict water quality parameters, which are time-consuming to measure, using other parameters, which are more cost-effective, quicker and easier to measure. Based on this

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

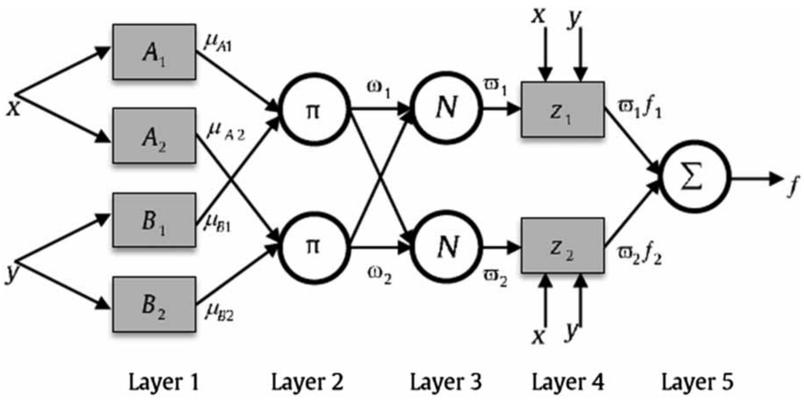

Figure 2 | ANFIS structure for a two-input first-order Sugeno model with two rules.

consideration, five variables which can be measured within a few hours, namely water temperature (*T*), pH, dissolved oxygen (DO), oxidation–reduction potential (ORP) and electrical conductivity (EC), were selected. The corresponding expensive and time-consuming to measure variables were the 5-days at 20 °C *N*-allylthiourea BOD_5 and ammonia-nitrogen (NH_4-N). Although NH_4-N may not always be expensive or time-consuming to measure in developed countries, there are potential cost, labour and time savings, if the analysis of many samples is required. Moreover, a predictive model can be particularly useful in situations where qualified technicians are unavailable to perform the analyses. Weekly data records for BOD_5 and NH_4-N, with no missing values, collected over the period of February 2008 to April 2012 from the ICW at the centre of this study were selected for modelling. These data records were, however, incomplete, as there were missing values for the other five parameters, which occurred completely at random within the data array. Details of the water quality sampling and analyses are presented in Dong *et al.* (2011) and Dzakpasu *et al.* (2011). In summary, *N*-allylthiourea BOD_5 was measured according to Method 5210 B published by APHA (1998). A respirometric (manometric) BOD OxiTop system (WTW GmbH, Weilheim, Germany) was used. NH_4-N was measured using test kits supplied by Hach Lange (Hach Method 8038, Hach Company, Loveland, CO, USA).

Data pre-treatment

For any modelling strategy, the quality of the outputs strongly depends on the quality of the selected input variables. Therefore, data pre-treatment provides essential techniques on how measured data sets can be validated, including how the quality of the data can be improved. This is necessary to obtain reliable analysis results. Preprocessing techniques to deal with the noise in data sets by replacing the missing values and omitting the outliers to improve the data quality and the performance of ANFIS is an option.

In the present study, all measurements were examined with respect to missing data, and missing values were replaced by values representing averages of adjacent data points. Furthermore, possible outliers and erroneous values were manually labelled. When identified, all outliers and erroneous values were subsequently removed from the data set and treated as missing values.

Selection of model input variables

In many previous models involving ANN and FLS, methods for input variable selection required generating different models while searching for the optimal combination of variables, which can be time-consuming and computationally expensive. In this study, a simple and effective method for selecting significant input variables on the basis of a class of sub-clusters created by a self-organizing network was used (Linkens & Chen 1999). Herein, a self-organizing map (SOM) network (Kohonen 1995) was applied to elucidate relationships and extract the most important input variables (i.e., *T*, pH, ORP, DO and EC), which independently and significantly influenced the effluent BOD_5 and NH_4-N. The SOM is an unsupervised competitive learning neural network model and algorithm that implements a characteristic non-linear projection method from the high-dimensional space of sensory or other input signals onto a low-dimensional regular lattice of neurons. The SOM can be visualized with the help of the unified distance matrix (*U*-matrix) and the component planes. The *U*-matrix visualizes distances between neighbouring map units, and thus

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

shows the cluster structure of the map. High values of the *U*-matrix point towards a cluster border. A uniform area of low values indicates a cluster. Each component plane indicates the values of one variable in each map unit (Vesanto *et al.* 1999).

Training and validation data sets

The selected pre-processed data records were divided into three subsets: a training set, a checking set to insure that models do not overfit the training data, and a validation set. During this step, care was taken to maintain the temporal coherence between data points.

Clustering of input and output variable values

Subtractive clustering is a fast and robust method for estimating the number and location of cluster centres present in a collection of data points. Initial fuzzy rules with rough estimates of MFs are obtained from the cluster centres; the MFs and other rule parameters are then optimized with respect to some output error criterion (Chiu 1997).

In this study, the training data were analysed by fuzzy subtractive clustering and were divided into seven and five clusters for *ANFIS(BOD₅)* and *ANFIS(NH₄-N)* models, respectively. Thus, each cluster represented a rule; the rules corresponding to Equation (3) may be stated as

$$R_{(r)}\text{: if } x_i \text{ is } A_r^1 \text{ and } x_2 \text{ is } A_r^2 \text{ and } \ldots \text{ and } x_m \text{ is } A_r^m \infty \quad \text{then } y_r \text{ is } f_r(x) \tag{3}$$

where R_i is the rth rule ($r = 1, 2, \ldots, n$), $x_j (1 \leq j \leq m)$ is the crisp input of rule r, y_r is the correspondent crisp output of rule r, A_r^m is the jth antecedent clause of rule r (fuzzy sets usually corresponding to linguistic labels), and $f_r(x) = \alpha_r^0 + \alpha_r^1 x_1 + \cdots + \alpha_r^m x_m$ is the consequent clause of rule r. The cluster centres represent the initial value of the premise parameter in the function.

Development of the ANFIS models

The ANFIS models were built by the *anfisedit* function in MATLAB®, where the data sets were initially analysed by fuzzy subtractive clustering to determine the cluster centres. The cluster centres represent the initial values of the premise parameters and, therefore, the initializing parameters were determined from here.

The types and number of MFs in each ANFIS, including *Gaussian curve, generalized bell curve, triangular* and *trapezoidal*-shaped functions, and the parameters were tested during clustering to determine an appropriate ANFIS model. The parameters for clustering were the following: squash factor, 1.25; aspect ratio, 0.5; rejection ratio, 0.15; range of influence, 0.5; error tolerance, 0. Table 2 presents the final architectures of the ANFIS models after clustering. With different input variables, the ANFIS models for BOD_5 and NH_4-N had *Gaussian curve* and *generalized bell curve* MFs, respectively, for each input variable giving the best result.

When the initial values of premise parameters and the architecture of the predictive models were defined, the hybrid-learning algorithm trained the networks. Subsequently, the premise and consequent parameters of the network were pruned, and MFs of the variables were optimized. The models were then used for predicting BOD_5 and NH_4-N.

Model performance evaluation criteria

The accuracy of the selected models was evaluated by the 'goodness of fit' between outputs of the model and the original measured system values, given the same input variables. The evaluation criteria considered in the present study were mean squared error (MSE) and time series plots, calculated as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \tag{4}$$

where $\hat{x}$ is a vector of predictions and x is the vector of the true values. An MSE of zero indicates that the estimator $\hat{x}$ predicts observations of the parameter x with perfect accuracy.

Table 2 | Optimal ANFIS architecture after fuzzy subtractive clustering

Basic structure	BOD	NH_4-N
Number of nodes	77	46
Number of linear parameters	35	20
Number of non-linear parameters	56	45
Number of inputs	3	3
Number of input MFs	7	5
Shape of input MFs	Gassmf	Gbellmf
Number of training data pairs	72	74
Number of fuzzy rules	7	5

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

RESULTS AND DISCUSSION

Relations between input and output variables

Table 3 summarizes the selected ICW effluent water quality variables between February 2008 and April 2012. Up to 200 measurements per variable were taken. The mean treatment efficiencies for NH_4-N and BOD_5 were 95 and 98%, respectively. According to the mean and median values, the ICW system achieved high treatment performance over the monitoring period.

SOM was applied to identify the relationships between the selected input data set. The *U*-matrix representation of the SOM visualizes the distances between the map nodes (Vesanto *et al*. 1999; Lee & Scholz 2006). The distances between the neighbouring map nodes were calculated and visualized by plotting different shades of grey between them. The component plane shows the value of the variable in each map unit (Lee & Scholz 2006), such that the lighter grey shades are associated with the high relative component value of the corresponding weight vector. This helps to identify and illustratively show clusters in the input data.

Figure 3 illustrates the SOM component planes visualization of the impact of the selected input variables on the concentrations of BOD_5 and NH_4-N in the ICW effluent. According to the SOM, high and values were replaced the average of adjacent data points effluent BOD_5 concentrations (>20 mg L^{-1}) were associated with relatively lower pH (<7), high T (>10 °C), and lower DO (<0.5 mg L^{-1}). The association of BOD_5 to EC and ORP was weak (Figure 3(a)), although EC appeared to exhibit a negative effect. This indicated that BOD_5 removal was probably influenced

Table 3 | Summary statistics of the selected effluent water quality variables from the integrated constructed wetland

	T	pH	ORP	DO	EC	NH_4-N	BOD_5
Mean	11.5	7.11	–13	2.37	456	1.27	8
Standard error	0.22	0.02	0.79	0.12	4.77	0.16	0.40
Median	12.6	7.09	–11	1.50	452	0.31	6
Mode	13.0	6.99	–12	0.4	400	0.15	4
Standard deviation	4.1	0.31	14.82	2.28	89.06	2.36	7.44
Range	18.5	1.47	111	11.37	747	12.75	46
Minimum	2.8	6.45	–60	0.12	128	0.05	4
Maximum	21.3	7.92	51	11.49	875	12.80	50

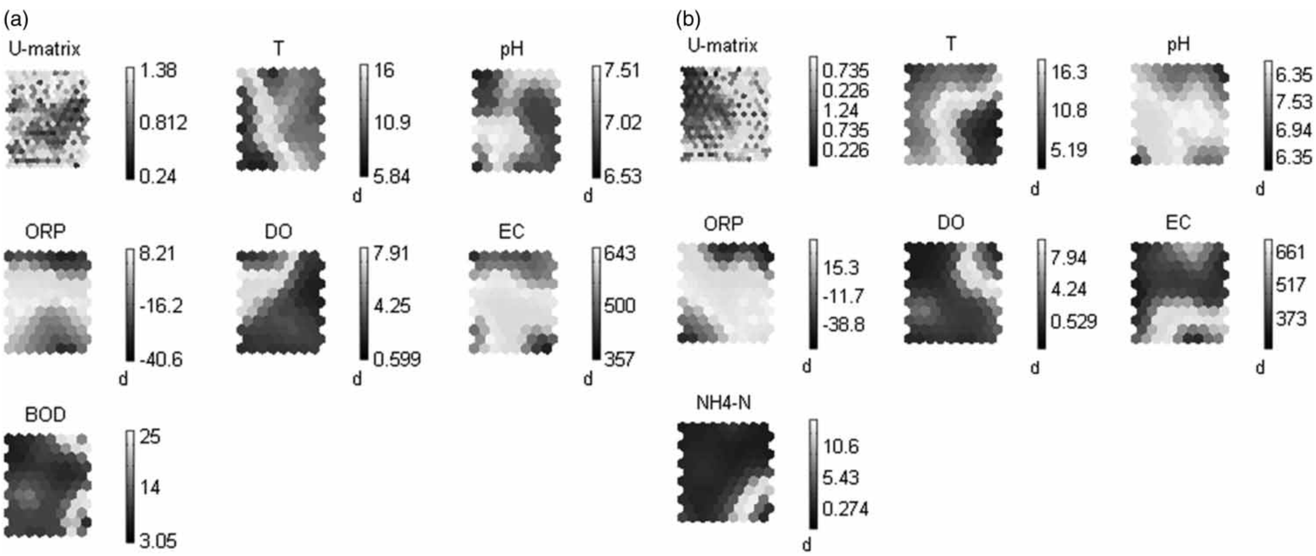

Figure 3 | SOM component planes visualization of the relationship between physicochemical parameters and effluent concentrations of (a) BOD_5 and (b) NH_4-N.

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

negatively by temperature and positively by DO and pH. Therefore, *T*, pH and DO, which showed stronger associations with BOD_5, were selected as the most important input variables, which independently and significantly influenced the effluent BOD_5. On the other hand, high effluent NH_4-N concentrations ($>10\ mg\ L^{-1}$) appeared to be associated with relatively lower *T* ($<5\ ^{\circ}C$), lower DO ($<1.0\ mg\ L^{-1}$) and higher EC ($>600\ \mu S\ cm^{-1}$) (Figure 3(b)). Thus, NH_4-N removal in the ICW was probably influenced negatively by temperature and DO, but positively by EC. There was no obvious association of pH and ORP to NH_4-N, and therefore, *T*, DO and EC were selected as the most important input variables, which independently and significantly influenced the effluent NH_4-N.

The relationships revealed by the SOM were found to conform well to the observed correlations reported by other previous studies. For example, it has been shown that organic matter and nutrient treatment efficiency were reduced by high salinity (measured as EC), which consequently, led to high effluent concentrations in CWs (Wu *et al*. 2008). Chapanova *et al*. (2007) demonstrated that a high salinity markedly reduced NH_4-N degradation and caused significant reductions in both nitrification and denitrification, resulting in higher effluent concentrations. Additionally, aerobic degradation of organic substances is noted to occur under high DO conditions, whereas most microbial processes occur under circum neutral pH conditions (Kadlec & Wallace 2009). Under high pH conditions, nevertheless, significant losses of N may occur in open water areas via ammonia gas volatilization, leading to lower NH_4-N effluent concentrations. Although the pH value measured in the ICW dropped to 6.45, there was no observed effect on NH_4-N removal, because the pH value remained within the circum neutral pH range between 6 and 8, which is near optimal for nitrification and denitrification processes to occur (Tchobanoglous *et al*. 2003; Kadlec & Wallace 2009). Moreover, water temperature modifies the rates of several key biological processes. For instance, microbial N processing is strongly influenced by water temperature, with doubling of rates over a temperature range of about $10\ ^{\circ}C$ (Kadlec & Reddy 2001). The microbial activities related to denitrification can decrease considerably at water temperatures below $15\ ^{\circ}C$ (Kuschk *et al*. 2003).

Simulation of effluent BOD_5 concentrations

Figure 4(a) depicts the prediction results of BOD_5 concentrations in the ICW effluent using ANFIS. The performance of the model during validation, when presented with previously unseen data, namely the checking and validation data (after training the model), is shown. At the training step, MSE between the predicted and observed concentrations of BOD_5 in the ICW effluent was 0.211, which was achieved after training over 60 epochs. A corresponding MSE of 0.876 was recorded for the checking data. During the validation, MSE recorded between the predicted and observed concentrations of BOD_5 in the ICW effluent was 0.914.

Overall, the ANFIS model performed very well in predicting the BOD concentrations in the ICW effluent with relatively high accuracy. The recorded MSE was found to be lower than other AI models applied for modelling BOD_5 concentrations in ICW effluents. For example, Zhang *et al*. (2009) recorded MSE of 6.351 for an ANN-based SOM model. Similarly, Dong *et al*. (2012) obtained MSE of 1.34 for predicting chemical oxygen demand using SOM. In addition, Tomenko *et al*. (2007) developed models with two different ANNs to predict the effluent BOD in a CW, in India, and recorded MSE of 8.34 and 15.80. In another study involving the use of ANN and partial least squares regression to predict the BOD, Basant *et al*. (2010) obtained MSE of 1.41 and 1.45, respectively.

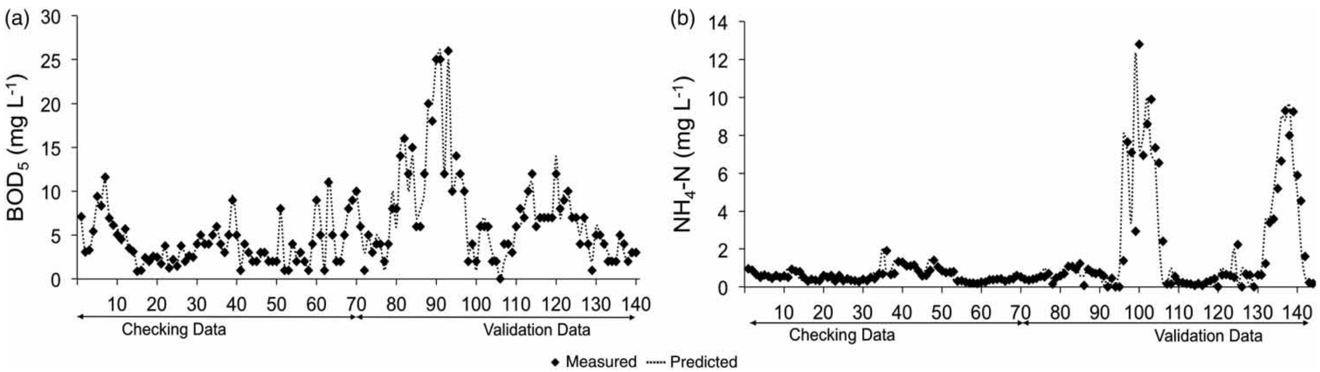

Figure 4 | Performance of ANFIS models in predicting (a) BOD_5 and (b) NH_4-N in ICW effluent.

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

Simulation of effluent NH_4-N concentrations

Figure 4(b) depicts the prediction results of NH_4-N concentrations in the ICW effluent using ANFIS. The performance of the model during validation, when presented with previously unseen data, namely the checking and validation data (after training the model), is shown. At the training step, MSE between the predicted and observed concentrations of NH_4-N in the ICW effluent reached a final value of 0.0916, which was achieved after training over 60 epochs. A corresponding MSE of 0.301 was recorded for the checking data. During the validation, MSE recorded between the predicted and observed concentrations of NH_4-N in the ICW effluent was 0.0570.

Overall, the ANFIS model performed very well in predicting the NH_4-N concentrations in the ICW effluent with relatively high accuracy. The recorded MSE was found to be similar to Zhang *et al.* (2009), who recorded MSE of 0.042 for an ANN-based SOM model employed for modelling NH_4-N concentrations in ICW effluents, but lower than Dong *et al.* (2012), who obtained MSE of 1.03 using the SOM model.

CONCLUSIONS

Findings indicate that ANFIS could predict the effluent quality variation quite strongly. Effluent concentrations of BOD_5 and NH_4-N, which are time-consuming and expensive to measure, were predicted relatively accurately by other effluent water quality parameters, which can be measured quickly within a few hours and cost effectively. The simulated effluent BOD_5 and NH_4-N concentrations well fitted the measured concentrations, which was also supported by relatively low MSE. Thus, by using ANFIS, it is possible to estimate the key effluent water quality parameters rapidly, which can assist the system operator to manage the ICW process operation and reduce the risk of failing regulatory requirements, while reducing man-hours, travel expenses and analytical costs.

ACKNOWLEDGEMENTS

The authors would like to thank Monaghan County Council for funding this research. Technical support and advice from Mark Johnston (Monaghan County Council), Rory Harrington and Susan Cook (Waterford County Council) and Aila Carty (VESI Environmental Ltd), all in Ireland, were also greatly appreciated.

REFERENCES

APHA 1998 *Standard Methods for the Examination of Water and Wastewater*, 20th edn. American Public Health Association/ American Water Works Association/Water Environment Federation, Washington, DC.

Basant, N., Gupta, S., Malik, A. & Singh, K. P. 2010 Linear and nonlinear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water – a case study. *Chemometrics and Intelligent Laboratory Systems* **104** (2), 172–180.

Chapanova, G., Jank, M., Schlegel, S. & Koeser, H. 2007 Effect of temperature and salinity on the wastewater treatment performance of aerobic submerged fixed bed biofilm reactors. *Water Science and Technology* **55** (8–9), 159–164.

Chen, S. H., Jakeman, A. J. & Norton, J. P. 2008 Artificial intelligence techniques: an introduction to their use for modelling environmental systems. *Mathematics and Computers in Simulation* **78** (2–3), 379–400.

Chiu, S. L. 1997 An efficient method for extracting fuzzy classification rules from high dimensional data. *Journal of Advanced Computational Intelligence and Intelligent Informatics* **1** (1), 31–36.

Cinar, O. 2005 New tool for evaluation of performance of wastewater treatment plant: artificial neural network. *Process Biochemistry* **40** (9), 2980–2984.

Dong, Y., Wiliński, P. R., Dzakpasu, M. & Scholz, M. 2011 Impact of hydraulic loading rate and season on water contaminant reductions within integrated constructed wetlands. *Wetlands* **31** (3), 499–509.

Dong, Y., Scholz, M. & Harrington, R. 2012 Statistical modeling of contaminants removal in mature integrated constructed wetland sediments. *Journal of Environmental Engineering* **138** (10), 1009–1017.

Dzakpasu, M., Hofmann, O., Scholz, M., Harrington, R., Jordan, S. N. & McCarthy, V. 2011 Nitrogen removal in an integrated constructed wetland treating domestic wastewater. *Journal of Environmental Science and Health, Part A* **46** (7), 742–750.

Hamed, M. M., Khalafallah, M. G. & Hassanien, E. A. 2004 Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling & Software* **19** (10), 919–928.

Jang, J. S. R. 1993 ANFIS: adaptive-network-based fuzzy inference systems. *IEEE Transactions on Systems, Man and Cybernetics* **23** (3), 665–684.

Kadlec, R. H. & Reddy, K. R. 2001 Temperature effects in treatment wetlands. *Water Environment Research* **73** (5), 543–557.

Kadlec, R. H. & Wallace, S. D. 2009 *Treatment Wetlands*, 2nd edn. CRC Press, Boca Raton, FL, USA.

Kohonen, T. 1995 *Self-Organizing Maps*. Springer Series in Information Sciences, Vol. 30. Springer, Berlin, Germany.

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest

Kumar, J. L. G. & Zhao, Y. Q. 2011 A review on numerous modeling approaches for effective, economical and ecological treatment wetlands. *Journal of Environmental Management* **92** (3), 400–406.

Kuschk, P., Wiebner, A., Kappelmeyer, U., Weibbrodt, E., Kästner, M. & Stottmeister, U. 2003 Annual cycle of nitrogen removal by a pilot-scale subsurface horizontal flow in a constructed wetland under moderate climate. *Water Research* **37** (17), 4236–4242.

Lee, B. H. & Scholz, M. 2006 Application of the self-organizing map (SOM) to assess the heavy metal removal performance in experimental constructed wetlands. *Water Research* **40** (18), 3367–3374.

Linkens, D. A. & Chen, M.-Y. 1999 Input selection and partition validation for fuzzy modelling using neural network. *Fuzzy Sets and Systems* **107** (3), 299–308.

Rousseau, D. P. L., Vanrolleghem, P. A. & De Pauw, N. 2004 Model-based design of horizontal subsurface flow constructed treatment wetlands: a review. *Water Research* **38** (6), 1484–1493.

Takagi, T. & Sugeno, M. 1985 Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics* **15** (1), 116–132.

Tchobanoglous, G., Burton, F. L. & Stensel, H. D. 2003 *Wastewater Engineering: Treatment and Reuse*, 4th edn. McGraw-Hill, New York.

Tomenko, V., Ahmed, S. & Popov, V. 2007 Modeling constructed wetland treatment system performance. *Ecological Modelling* **205** (3–4), 355–364.

Vesanto, J., Himberg, J., Alhoniemi, E. & Parhakangas, J. 1999 Self-organizing map in Matlab: the SOM toolbox. In: *Proceedings of the Matlab DSP Conference, Comsol, Helsinki, Finland*, pp. 35–40.

Wang, Y. M. & Elhag, T. M. S. 2008 An adaptive neuro-fuzzy inference system for bridge risk assessment. *Expert Systems with Applications* **34** (4), 3099–3106.

Wu, Y., Tam, N. F. Y. & Wong, M. H. 2008 Effects of salinity on treatment of municipal wastewater by constructed mangrove wetland microcosms. *Marine Pollution Bulletin* **57** (6), 727–734.

Zhang, L., Scholz, M., Mustafa, A. & Harrington, R. 2009 Application of the self-organizing map as a prediction tool for an integrated constructed wetland agroecosystem treating agricultural runoff. *Bioresource Technology* **100** (2), 559–565.

First received 29 August 2014; accepted in revised form 4 November 2014. Available online 17 November 2014

Downloaded from https://iwaponline.com/wst/article-pdf/71/1/22/470052/wst071010022.pdf
by guest