

Turbidity Prediction in a River Basin by Using Artificial Neural Networks: A Case Study in Northern Spain

C. Iglesias · J. Martínez Torres · P. J. García Nieto ·
J. R. Alonso Fernández · C. Díaz Muñoz · J. I. Piñeiro ·
J. Taboada

Received: 14 July 2013 / Accepted: 2 October 2013 /
Published online: 20 December 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Chemical and physical-chemical parameters define water quality and are involved in water body type and habitat determination. They support a biological community of a certain ecological status. Water quality controls involve a large number of measurements of variables and observations according to the European Water Framework Directive (Directive 2000/60/EC). In some cases, such as areas with especially critical uses or points in which potential pollution episodes are expected, the automatic monitoring is recommended. However, the chemical and physical-chemical measurements are costly and time consuming. Turbidity is shown as a key variable for the water quality control and it is also an integrative parameter. For this reason, the aim of this work is focused on this main parameter through the study of the influence of several water quality parameters on it. The artificial neural networks (ANNs) have been used in a wide range of biological problems with promising results. Bearing this in mind, turbidity values have been predicted here by using artificial neural networks (ANNs) from the remaining measured water quality parameters with success taking into account the synergistic interactions between the input variables in the Nalón river basin (Northern Spain). Finally, the main conclusions of this study are exposed.

Keywords Artificial neural networks (ANNs) · Water quality monitoring · Water Framework Directive (WFD) · Water pollution

C. Iglesias · J. I. Piñeiro · J. Taboada
Department of Natural Resources and Environmental Engineering, University of Vigo, 36310 Vigo, Spain

J. Martínez Torres
Centro Universitario de la Defensa, Academia Militar, 50090 Zaragoza, Spain

P. J. García Nieto (✉)
Department of Mathematics, Faculty of Sciences, University of Oviedo, 33007 Oviedo, Spain
e-mail: pauli@constru.uniovi.es

J. R. Alonso Fernández · C. Díaz Muñoz
Cantabrian Basin Authority, Spanish Ministry of Agriculture, Food and Environment, 33071 Oviedo, Spain

1 Introduction

The chemical and physical-chemical parameters, in addition to hydromorphological aspects, have determined fundamentally the types of water body and habitat and, consequently, the type of the specific biological community. Water quality controls, based on the previous parameters, are necessary to avoid or minimize potential risks (García-Barcina et al. 2002; Vigil 2003; Rabalais et al. 2009; Li and Migliaccio 2010). Furthermore, they will support the interpretation, assessment and classification of the results arising from the monitoring of the biological quality elements. In this sense, the monitoring and assessment of the chemical and physical-chemical quality elements are a requirement in the European Water Framework Directive (EWF) (Directive 2000/60/EC) and they are necessary for the final ecological status assessment. For instance, primary chemical and physical-chemical parameters such as nutrients, oxygen, temperature, conductivity, pH, etc. support the interpretation of biological data. Other parameters such as the dissolved organic carbon (DOC) or solid content are required for reliable interpretation of the chemical results.

Automated Water Quality Information System (AWQIS), implemented by Spanish Government between September 1993 and November 1995, can be considered a very useful tool for the water quality monitoring (Andrews 1984; Bartram and Rees 2000; Aslan-Yilmaz et al. 2004). It consists of several stations strategically placed to acquire data on a near real time. The chemical and physical-chemical data gathered from this regular monitoring are essential since: (a) they allow the early detection of changes and trends in water quality; (b) they provide the basis for calibrating both the predictive water models and ecological models; and (c) they allow to assess alternative remediation strategies.

The dynamic nature of large water bodies is the main reason for the variability of the physical, chemical and biological processes as well as their parameters, which define the water quality. Therefore, the automatic monitoring is an useful tool to control water quality, particularly in some areas with especially critical uses that need preventive actions, and in points where potential pollution episodes are expected (areas with high pressures and impacts). These automatic alert stations control some water quality parameters such as pH, conductivity, dissolved oxygen, ammonium, organic matter and turbidity. All of them are necessary to assess the status of water bodies. However, each measurement involves high costs and time consuming. Therefore, minimize the number of parameters to be measured should be a goal to consider.

Turbidity can be considered the most important parameter, among all the required ones to ascertain water status, since it can be regarded as an integrative parameter: turbidity is largely affected by the other parameters so that it is indicative of the global effect of the remaining parameters. Furthermore, it is an integrative parameter because high values of turbidity normally indicate high values of other parameters associated with water quality, such as the chemical oxygen demand (COD), or the concentrations of different substances related to pollution (ammonium, sulfate, nitrate, etc.) (Vigil 2003; Li and Migliaccio 2010; Díaz Muñoz et al. 2012).

In automatic monitoring, turbidity is generally measured by a turbidimeter (especially for very low values). However, turbidimeters have some disadvantages: high cost, expensive maintenance and an easy damage. Furthermore, it requires a power supply and the use of carcinogenic substances as the formazine. In this sense, it would be very interesting to estimate turbidity values with sufficient accurateness from the remaining measured parameters.

The aim of this work was to obtain a predictive model to calculate turbidity values from the other measured parameters (Barnes and Chu 2010). Furthermore, the contribution degree of the remaining parameters as well as the kind of pollution can be determined from this information, at the studied site.

For the above-mentioned purpose, artificial neural networks (ANNs) (Haykin 2008; Babel and Shinde 2011; Abghari et al. 2012) were used as automated learning tools, training them in order to predict turbidity from other parameters such as conductivity, dissolved oxygen, pH, ammonium content and temperature of the water bodies. According to previous research, ANNs have been proved to be an effective tool to predict natural parameters, being successfully used in a wide range of environmental fields: forest modeling (Ingram et al. 2005; García Nieto et al. 2012a), solar radiation estimation (Behrang et al. 2010; Rahimikhoob 2010), ozone levels prediction (Elkamel et al. 2001; Abdul-Wahab and Al-Alawi 2002; Al-Alawi et al. 2008), study of water properties (Sengorur et al. 2006; Palani et al. 2008; Hea et al. 2011) and so on.

In summary, the present study is structured as follows: firstly, the materials, methods and dataset used are explained; secondly, the results of the different ANNs trained are presented and discussed; and finally, the main conclusions of this research work are drawn.

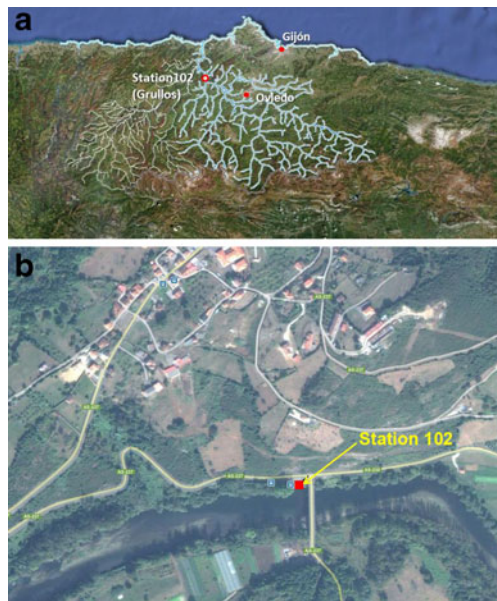
2 Materials and methods

Data collected in the station located in Grullós (Principality of Asturias) (AWQIS station number 102) from the Nalón river were analysed. This river is the most important of the Principality of Asturias (Northern Spain), because of both its length and its flow and also because of its basin with 4,900 km² (see Fig. 1a and b).

2.1 Study Area and Dataset

The Nalón river basin covers about half of the total area of the Principality of Asturias (Northern Spain) (see Fig. 1a). Just before flowing into the Cantabrian Sea, it forms the San Esteban estuary (Díaz Muñiz et al. 2012). The Nalón river has an annual mean river flow of 43,41 m³/s, and the daily maximum and minimum river flows are 475 and 0,36 m³/s, respectively.

Fig. 1 **a** The Nalón river basin (branched region) with the location of the automated monitoring station number 102; **b** an aerial photograph of the automated monitoring station 102 (Grullós station) in detail (the red square)



The Nalón river is also one of the most polluted rivers of the autonomous community of Principality of Asturias since it receives discharged wastewater of the major cities and industries in the region. Although cleaning-up (drainage) works have been carried out in the entire Nalón river basin in the last 20 years, it is mandatory to intensify efforts in order to fulfil the requirements of the WFD (Directive 2000/60/EC). In this sense, the automated water quality monitoring of this river basin (Andrews 1984; Bartram and Rees 2000; Aslan-Yilmaz et al. 2004; Nagesh Kumar et al. 2004) is shown as an interesting tool for controlling its water quality, especially if we have also taken into account that Nalón is an intensive river for water supply, industrial use and agricultural use.

It is very important to consider the presence of *synergy* among the independent variables that defining the water turbidity in the calculations. Synergy is the interaction of multiple elements in a system to produce an effect different from or greater than the sum of their individual effects. In the natural world, synergistic phenomena are ubiquitous, ranging from physics (for example, the different combinations of quarks that produce protons and neutrons) to chemistry (a popular example is water, a compound of hydrogen and oxygen), to the cooperative interactions among the genes in genomes, the division of labour in bacterial colonies, the synergies of scale in multi-cellular organisms, etc.

To study the existing synergy among the above-mentioned parameters on the water turbidity and to improve the performance of the neural network, a new variable resulting from the product of two of the original (measured) input variables was introduced. It was calculated in ten different ways to take into account all different synergistic effects: conductivity*ammonium, conductivity*pH, conductivity*dissolved oxygen, conductivity*temperature, ammonium*dissolved oxygen, ammonium*pH, ammonium*temperature, dissolved oxygen*pH, dissolved oxygen*temperature and pH*temperature.

Regarding climate, it is possible to distinguish two basic types in the Iberian Peninsula (Spain and Portugal): the Mediterranean and oceanic climates (Pausas 2004). The Principality of Asturias has a clearly oceanic climate with abundant rainfalls throughout the year, a moderate solar radiation and the temperature contrast are always moderate, especially on the coast. Although it increases inland, both in the interior valleys and in the mountains, never reach the extreme values observed in the Mediterranean climate (Pausas 2004).

The dataset analysed was collected along 2010. These data included measurements of turbidity, ammonium, conductivity, dissolved oxygen, pH and temperature every 5 min. The automated water quality monitoring station collects data 24 h a day and 7 days a week. Measurements of water quality parameters were recorded at 15-minute intervals using data loggers in situ and converted into hourly averages.

The sampling was carried out at 75 cm below the water surface. The sensors used for measuring the studied parameters at the automated monitoring stations were:

- For turbidity: the Zellewewger Turbidity immersion sensor WP240 in combination with the two-channel TxProTM-2 Transmitter, both for drinking water and industrial water. Its measurement principle is an optical sensor 90 ° from scattered light to near infrared light (880 nm).
- For conductivity: the Conducell 4USF ARC sensor made up of two electrodes with a fixed cable. Its measurement principle is a conductive cell with graphite electrodes for midrange applications.
- For ammonium: the AMTAX Inter2 process photometers. The ammonium concentration measurement is based on the indophenol blue method.

- For dissolved oxygen: the oxygen probe LDO Hamilton. The oxygen sensor is made up of an oxygen impermeable hard substrate. An oxygen sensitive luminescent dye, along with a scattering agent, is pad-printed on the substrate. The duration of the luminescence is proportional to the concentration of dissolved oxygen in the sample.
- For pH: the pH probe Hamilton pH ARC made up of a sensor of electrodes and a transmitter (pH E-H liquisys S CPM 223).
- For temperature: the temperature probe 771217-0005.

2.2 Artificial Neural Network (ANN)

Artificial neural networks (ANNs) are mathematical models that are inspired by the structure and functional aspects of biological neural networks (Fausset 1993). A neural network consists of an interconnected group of artificial neurons, and it processes information using a connection approach to computation. The number of types of ANNs and their uses are very high. Since the first neural model, proposed by MacCulloch and Pitts (1943), there have been developed another ones considered as ANNs.

A particular case of neuronal network is the multilayer perceptron (MLP) with a layered structure where each neuron is a perceptron (Bishop 2008). The functional model of neural networks, focusing on feedforward networks (MLP), whose architecture can be represented in an acyclic diagram so that each node is not backpropagated (see Fig. 2), with specific activation functions and weights with fixed values. The network implements the function $f: \mathbf{X} \subset \mathcal{R}^d \rightarrow \mathbf{Y} \subset \mathcal{R}^c$ (Nagesh Kumar et al. 2004; Heaton 2012):

$$\mathbf{f}(\mathbf{x}) = \phi(\psi(\mathbf{x})) \tag{1}$$

$$\phi : \mathbf{X} \subset \mathcal{R}^d \rightarrow \mathbf{T} \subset \mathcal{R}^p$$

$$\psi : \mathbf{T} \subset \mathcal{R}^p \rightarrow \mathbf{Y} \subset \mathcal{R}^c$$

where \mathbf{T} is the hidden space, ϕ is the activation function of the hidden layer and ψ is the activation function of the input layer and, finally, the function implemented by the MLP is (Bishop 2008; Heaton 2012; Mustafa et al. 2012):

$$\mathbf{f}(\mathbf{x}) = \sum_{j=1}^p \phi_j \left(c_j \psi(\mathbf{w}_j^T \mathbf{x} + \mathbf{w}_0) \right) + c_0 \tag{2}$$

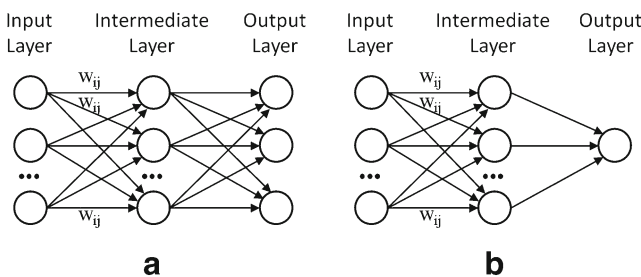


Fig. 2 Schematic diagram of the architecture of a typical artificial neuronal network (ANN): **a** ANN with multiple outputs; and **b** ANN with one output

where p is the number of neurons of the hidden layer, \mathbf{w} and \mathbf{w}_0 are the weights of the input layer and c and c_0 are the weights of the hidden layer. It can find algorithms which can adjust the weights of the ANN in order to obtain the desired output from the network. This process of adjusting the weights is called learning or training.

The back-propagation algorithm is used in layered feed-forward ANNs and it means that the artificial neurons are organized in layers, and send their signals forward, and then the errors are propagated backwards. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer (see Fig. 2). The idea of the back-propagation algorithm is to reduce this error, until the ANN learns the training data.

To fix ideas, an artificial neural network (ANN) is typically defined by three types of parameters (Bishop 2008; Babel and Shinde 2011; Heaton 2012): the interconnection pattern between different layers of neurons, the learning process for updating the weights of the interconnections and the activation function that converts a neuron's weighted input to its output activation.

3 Analysis of Results and Discussion

The pollutant controls are necessary to avoid or minimize potential risks (García-Barcina et al. 2002; Rabalais et al. 2009; García Nieto et al. 2013). The Grullas automated quality monitoring station was used to measure different parameters related to water pollution in order to control water quality of the Nalón river basin in 2010, with a 15-minute time interval between samplings. Turbidity was chosen as the most important parameter for this study because of two reasons: (1) it can be considered an integrative water quality parameter (high turbidity values are often associated to high values of other parameters such as chemical oxygen demand, ammonium ion, nitrates, etc.) (Díaz Muñiz et al. 2012); and (2) turbidity measurements represent a key test for water quality assessment since pure water has a zero turbidity value. In other words, this parameter is a good water quality indicator. Turbidity is a measurement of the suspended particulate matter and it is usually produced by anthropogenic sources as forest harvesting, road building, agriculture, urban developments, sewage treatment plant effluents, mining and industrial effluents (France and Peters 1995; García Nieto et al. 2012b; Alonso Fernández et al. 2013). Turbidity parameter is critical for water quality assessment since water quality decreases as water turbidity increases and vice versa.

Turbidity is also an important parameter in defining the *water status* since a good surface water status requires a rich, balanced and sustainable ecosystem and that the established environmental quality standards for pollutants are respected (Directive 2000/60/EC). Indeed, the Water Framework Directive (WFD) (Directive 2000/60/EC) establishes in order to reach a good condition of surface waters that this good condition is defined as the status achieved by a surface water body when both its ecological status and its chemical status are at least good. The ecological status is defined by the WFD as an expression of the quality of the structure and functioning of aquatic ecosystems associated with surface waters. On one hand, it is important to take into account that the turbidity increase is caused by both the human activities and phytoplankton growth (Verity et al. 1993). On the other hand, large turbidity levels can reduce the amount of light reaching lower depths in the water bodies, inhibiting the growth of submerged aquatic plants and affecting consequently, species that dependent on them (e.g., fish) (Clark 2001; Alonso Fernández et al. 2013).

To estimate turbidity from other chemical and physical-chemical parameters it is important to select the model that best fits the experimental data (Barnes and Chu, 2010). The criterion considered in this research to measure the goodness-of-fit was the coefficient of determination

R^2 (Freedman et al. 2007). This ratio indicates the proportion of total variation in the dependent variable explained by the model (inside-bark volume in our case). A dataset takes values t_i , each of which has an associated modelled value y_i . The former are called the observed values and the latter are often referred to as the predicted values. Variability in the dataset is measured through different sums of squares:

- $SS_{tot} = \sum_{i=1}^n (t_i - \bar{t})^2$: the total sum of squares, proportional to the sample variance.
- $SS_{reg} = \sum_{i=1}^n (y_i - \bar{t})^2$: the regression sum of squares, also called the explained sum of squares.
- $SS_{err} = \sum_{i=1}^n (t_i - y_i)^2$: the residual sum of squares.

In the previous sums, \bar{t} is the mean of the n observed data:

$$\bar{t} = \frac{1}{n} \sum_{i=1}^n t_i \quad (3)$$

Bearing in mind the above sums, the general definition of the coefficient of determination is:

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \quad (4)$$

A coefficient of determination value of 1.0 indicates that the regression curve fits the data perfectly.

In this way, neural networks were applied to predict the turbidity from the other remaining variables, studying their influence in order to optimize its calculation through the analysis of the coefficient of determination R^2 in test of the neural networks. The coefficient of determination is a statistical measure of how well a regression curve approximates real data points. Furthermore, it is a descriptive measure between zero and one, indicating how good one term is predicted by another one, being $R^2=1$ the best approximation and $R^2=0$ the worst approximation.

The addition of synergistic variables, resulted in an increase of R^2 from 0.70 to 0.80 and therefore an increase of the correlation coefficient from 0.84 to 0.89. This additional synergistic variable was calculated in ten different ways: Conductivity*Ammonium, Conductivity*pH, Conductivity*Dissolved oxygen, Conductivity*Temperature, Ammonium*Dissolved oxygen, Ammonium*pH, Ammonium*Temperature, Dissolved oxygen*pH, Dissolved oxygen*Temperature and pH*Temperature. This interaction is known as synergy or synergistic behaviour. This synergistic behaviour is the result of joint action of two or more causes, but characterized by having a greater effect than that resulting from the sum of these causes. Synergy has been advanced as a hypothesis on how complex systems operate. Environmental systems may react in a nonlinear way to perturbations, so that the outcome may be greater than the sum of the individual component alterations. Synergistic responses are a complicating factor in environmental modelling.

The accuracy was calculated training the model with the 90 % of the sample and testing with the 10 % implementing a 10-fold cross validation process. This algorithm allows the model to be trained with a set of sample and to be tested with any element not used in training.

The initial neural network was as follows. The output of the ANN is the turbidity and inputs of the ANN are ammonium, conductivity, dissolved oxygen, pH and temperature. It was observed that the coefficient of determination (R^2) is stable when the number of neurons is greater than 60. Thus, in subsequent tests, neural networks were trained with up to 70 neurons.

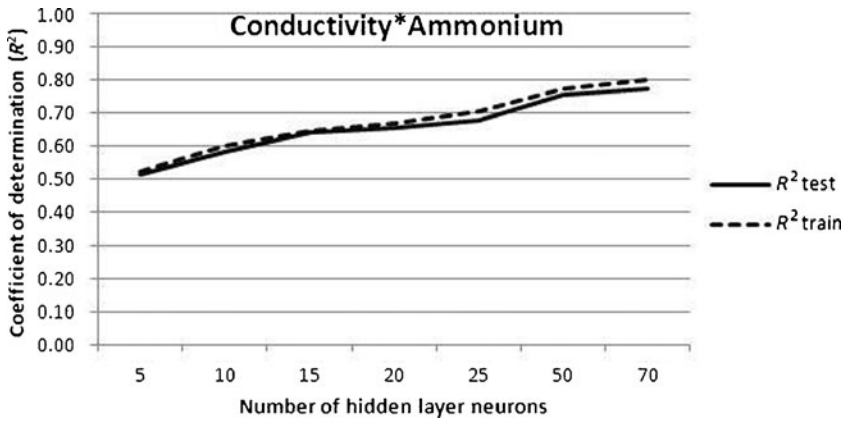


Fig. 3 Coefficient of determination (R^2) of the neural network predicting turbidity with the synergistic variable Conductivity*Ammonium

However, the results have shown that the turbidity is poorly predicted despite the considerable amount of input data, since the R^2 obtained in testing with 70 neurons in the hidden layer was 0.7. Consequently, a new variable was introduced in order to improve the performance of the neural network, taking into account the existing relationship among the measured variables.

Bearing in mind the different synergistic variables that can be calculated, the neural network was trained with up to 70 neurons. The highest R^2 in testing was obtained in the neural networks where the synergistic variable was defined as Conductivity*Ammonium (see Fig. 3), Dissolved oxygen*Temperature (see Fig. 4) and pH*Temperature (see Fig. 5). This fact indicates the importance of those parameters, focusing on the R^2 in the test set.

The analysis of the weights of each input variable of the network permits to study their importance and determine those ones that influence more in the value of the turbidity. For the 70 neurons network, the mean of the absolute value of the weights of the variables are shown

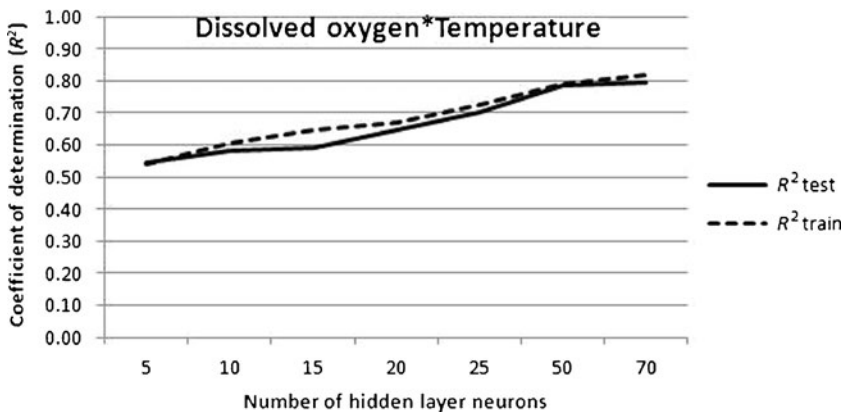


Fig. 4 Coefficient of determination (R^2) of the neural network predicting turbidity with the synergistic variable Dissolved oxygen*Temperature

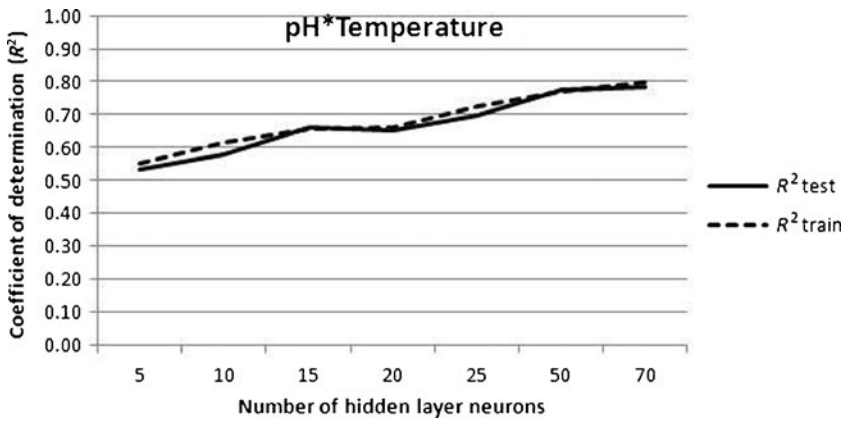


Fig. 5 Coefficient of determination (R^2) of the neural network predicting turbidity with the synergistic variable pH*Temperature

in Table 1. These weights allow us to distinguish clearly the parameters which condition the most the turbidity.

These results indicate that the temperature variable is the most important parameter for the estimation of turbidity. In order to verify this point, a new neural network was trained taking ammonium, conductivity, dissolved oxygen and pH (without temperature) as the input variables, and turbidity as the output variable. The R^2 obtained were less than the observed one for the initial neural network, proving that the network for the prediction of the turbidity performs better when the temperature is taken into account.

A further analysis was carried out to confirm the importance of temperature in the prediction of turbidity, studying in the same way the weights of the variables of the neural networks with the synergistic variables. Focusing on Table 2, the highest weights belong to the temperature and to the synergistic variable when it includes the variable temperature.

Table 3 shows the Nalón’s average monthly river flow in cubic meters per second along 2010. From these Nalón’s flow rates indicated in Table 3, it is possible to observe that the turbidity is well predicted by temperature since flow rates are typically larger in winter-spring than in summer. Indeed, during rainfall events or floods, turbidity would increase as a result of large flow rates leading to increased erosion power of water in channels and hill slopes. Additionally, when the flow rates are very small, turbidity increases due to the decreased ability of the river dilution. Normally this second effect is smaller than the first one (Díaz Muñoz et al., 2012).

Therefore, turbidity means a loss of transparency of the water due to the presence of suspended particles in it. Furthermore, it is indicative of the quality of the water (Vigil 2003). The turbidity-causing particles can come from suspended sediments, either from erosion or

Table 1 Mean of the absolute value of the weights of the neural network with 70 neurons for the estimation of turbidity from ammonium, conductivity, dissolved oxygen, pH and temperature measurements

Ammonium	Conductivity	Dissolved oxygen	pH	Temperature
15.4479	32.7060	36.5091	31.7491	115.9802

Table 2 Mean of the absolute value of the weights of the neural network with 70 neurons for the prediction of turbidity from ammonium (Ammo), conductivity (Cond), dissolved oxygen (Diss oxyg), pH, temperature (Temp) along with one synergistic variable

Synergistic variable	Ammo	Cond	Diss oxyg	pH	Temp	Synergistic variable
Cond*Ammo	29.1245	26.4071	27.0781	23.8135	77.8780	43.7151
Cond*pH	13.5192	84.1893	27.9682	26.1739	84.3741	95.1441
Cond*Diss oxyg	15.8590	57.9184	61.7641	42.1263	94.4955	94.9102
Cond*Temp	28.4593	59.9611	49.0624	43.0544	241.5228	231.0599
Ammo*pH	55.2808	42.0877	41.5254	40.6314	98.1919	56.5763
Ammo*Diss oxyg	68.7679	37.5281	34.9272	30.9129	102.5687	51.9604
Ammo*Temp	25.9110	33.0091	23.8117	23.5218	64.9713	36.9868
Diss oxyg*pH	14.5125	37.0513	49.4464	31.3699	81.4540	51.8707
Diss oxyg*Temp	16.1457	41.8914	35.4589	46.3892	136.7359	163.4885
pH*Temp	18.9091	44.4897	30.7531	32.7276	78.0850	71.4100

have been removed from the bottom, but can also come from municipal waste, which carry small particles of sand and dirt or due to an increase in phytoplankton.

Additionally, turbidity causes various alterations in the waters (Vigil 2003; Li and Migliaccio 2010):

- The suspended particles absorb infrared radiation, causing an increase in temperature. Warming affects the aquatic flora and fauna by altering its ecological equilibrium.
- The gases are more soluble in cold water than in hot water. Increased water temperature decreases the amount of dissolved oxygen and can cause death of aquatic animals.
- The turbid waters do not allow the passage of visible light, so that the phytoplankton cannot perform photosynthesis causing dissolved oxygen depletion.

Finally, this research work was able to predict the Nalón river's turbidity values along 2010 at the automated monitoring station 102 (Grullos station) in agreement with the real values of turbidity observed with great accurateness and success with and without synergistic variables (see Fig. 6).

Table 3 Nalón average monthly river flow in cubic meters per second as a function of time from January 2010 to December 2010

Month	Average monthly river flow (m ³ /s)
January	250
February	200
March	150
April	90
May	190
June	280
July	80
August	60
September	40
October	55
November	280
December	210

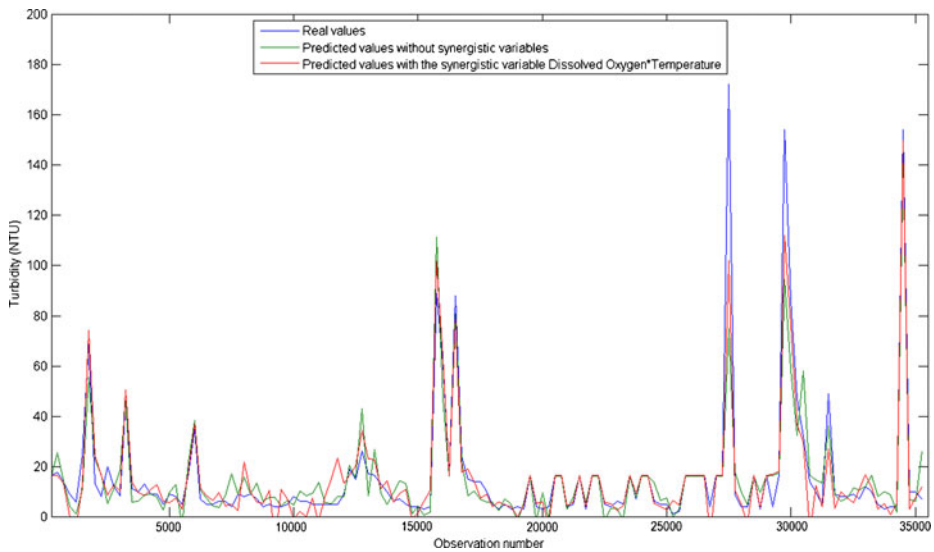


Fig. 6 Comparison among the three curves for the turbidity: **a** real values of turbidity (*blue curve*); **b** predicted values of turbidity without synergistic variables by using ANNs at the automated monitoring station 102 (Grullas station) along 2010 ($R^2=0.7$) (*green curve*); and **c** predicted values of turbidity with the synergistic variable Dissolved oxygen*Temperature by using ANNs at the automated monitoring station 102 (Grullas station) along 2010 ($R^2=0.8$) (*red curve*)

4 Conclusions

Several neural networks were trained for turbidity prediction from the other measured quality variables, in order to lower costs in the quality assessment of water bodies. A first attempt consisted on predicting the turbidity from ammonium, conductivity, dissolved oxygen, pH and temperature values. The first calculation of R^2 shows that this neural network has a low accuracy.

Therefore, a new additional variable resulting from the product of two of the original (measured) parameters was introduced in the network. This synergistic variable improved considerably the results of the network. Additionally, the weights of the ANN variables show the influence of the temperature. This subject was also studied with a neural network where the temperature was not taken into account.

Hence, two main conclusions were obtained from this work: firstly, the temperature could be considered the most influential parameter in the turbidity values; and finally, the neural network for the turbidity prediction performs better when a synergistic variable is included.

Finally, our methodology can be applied to other rivers with similar or different sources of pollutants with success, bearing in mind that the specificities of each location must be taken into account.

Acknowledgments This study was possible thanks to the experimental dataset from the Automated Water Quality Information System (AWQIS) located in the basin of the Nalón river collected by the Cantabrian Basin Authority (Ministry of Agriculture, Food and Environment of the Government of Spain). At the same time, the authors wish to acknowledge the computational support provided by the Department of Mathematics at University of Oviedo and 'Centro Universitario de la Defensa' at University of Zaragoza. Additionally, Carla Iglesias acknowledges her grant FPU 12/02283 from the Spanish Ministry of Education. English grammar and spelling of the manuscript have been revised by a native person.

References

- Abdul-Wahab SA, Al-Alawi SM (2002) Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. *Environ Modell Softw* 17:219–228
- Abghari H, Ahmadi H, Besharat S, Rezaverdinejad V (2012) Prediction of daily pan evaporation using wavelet neural networks. *Water Resour Manag* 26(12):3639–3652
- Al-Alawi SM, Abdul-Wahab SA, Bakheit CS (2008) Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. *Environ Modell Softw* 23(4):396–403
- Alonso Fernández JR, Díaz Muñiz C, García Nieto PJ, de Cos Juez FJ, Sánchez Lasheras F, Roqueñí MN (2013) Forecasting the cyanotoxins presence in fresh waters: A new model based on genetic algorithms combined with the MARS technique. *Ecol Eng* 53:68–78
- Andrews MJ (1984) Thames estuary: pollution and recovery. In: Sheehan PJ, Miller DR, Butler GC, Bourdeau PH (eds) *Effects of pollutants at the ecosystem level*, Scope 22. John Wiley & Sons, New York, pp 195–227
- Aslan-Yilmaz A, Okus E, Övez S (2004) Bacteriological indicators of anthropogenic impact prior to and during the recovery of water quality in an extremely polluted estuary, Golden Horn, Turkey. *Mar Pollut Bull* 49: 951–958
- Babel MS, Shinde VR (2011) Identifying prominent explanatory variables for water demand prediction using artificial neural networks: a case study of Bangkok. *Water Resour Manag* 25(6):1653–1656
- Barnes DJ, Chu D (2010) *Introduction to modeling for biosciences*. Springer, New York
- Bartram J, Rees G (2000) *Monitoring bathing waters: a practical guide to the design and implementation of assessments and monitoring programmes*. E & FN SPON, London
- Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR (2010) The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. *Sol Energy* 84(8):1468–1480
- Bishop CM (2008) *Neural networks for pattern recognition*. Oxford University Press, New York
- Clark RB (2001) *Marine pollution*. Oxford University Press, New York
- Díaz Muñiz C, García Nieto PJ, Alonso Fernández JR, Martínez Torres J, Taboada J (2012) Detection of outliers in water quality monitoring samples using functional data analysis in San Esteban estuary (Northern Spain). *Sci Total Environ* 439:54–61
- Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, L-327 Luxembourg
- Elkamel A, Abdul-Wahab S, Bouhamra W, Alper E (2001) Measurement and prediction of ozone levels around a heavily industrialized area: a neural network approach. *Adv Environ Res* 5(1):47–59
- Fausset LV (1993) *Fundamentals of neural networks: architectures, algorithms and applications*. Pearson, New York
- France RL, Peters RH (1995) Predictive model of the effects on lake metabolism of decreased airborne litterfall through riparian deforestation. *Conserv Biol* 9(6):1578–1586
- Freedman D, Pisani R, Purves R (2007) *Statistics*. W. W. Norton & Company, New York
- García Nieto PJ, Martínez Torres J, Araújo Fernández M, Ordóñez Galán C (2012a) Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*. *Appl Math Model* 36: 6137–6145
- García Nieto PJ, Alonso Fernández JR, Sánchez Lasheras F, de Cos Juez FJ, Díaz Muñiz C (2012b) A new improved study of cyanotoxins presence from experimental cyanobacteria concentrations in the Trasona reservoir (Northern Spain) using the MARS technique. *Sci Total Environ* 430:88–92
- García Nieto PJ, Alonso Fernández JR, de Cos Juez FJ, Sánchez Lasheras F, Díaz Muñiz C (2013) Hybrid modelling based on support vector regression with genetic algorithms in forecasting the cyanotoxins presence in the Trasona reservoir (Northern Spain). *Environ Res* 122:1–10
- García-Barcina JM, Oteiza M, de la Sota A (2002) Modelling the faecal coliform concentrations in the Bilbao estuary. *Hydrobiologia* 475(476):213–219
- Haykin SO (2008) *Neural networks and learning machines*. Prentice Hall, New York
- Hea B, Oki T, Sun F, Komori D, Kanae S, Wang Y, Kim H, Yamazaki D (2011) Estimating monthly total nitrogen concentration in streams by using artificial neural network. *J Environ Manage* 92(1):172–177
- Heaton J (2012) *Introduction to the math of neural networks*. Heaton Research, New York
- Ingram JC, Dawson TP, Whittaker RJ (2005) Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks. *Remote Sens Environ* 94(4):491–507
- Li Y, Migliaccio K (2010) *Water quality concepts, sampling, and analyses*. CRC Press, Boca Raton (FL)
- MacCulloch WS, Pitts WS (1943) A logical calculus of the ideas immanent in nervous activity. *B Math Biophys* 5:115–133
- Mustafa MR, Rezaur RB, Saiedi S, Isa MH (2012) River suspended sediment prediction using various multilayer perceptron neural network training algorithms—a case study in Malaysia. *Water Resour Manag* 26(7):1879–1897

- Nagesh Kumar D, Srinivasa Raju K, Sathist T (2004) River flow forecasting using recurrent neural networks. *Water Resour Manag* 18(2):143–161
- Palani S, Liong SY, Tkalich P (2008) An ANN application for water quality forecasting. *Mar Pollut Bull* 56: 1586–1597
- Pausas JG (2004) Changes in fire and climate in the eastern Iberian Peninsula (Mediterranean basin). *Climatic Change* 63(3):337–350
- Rabalais NN, Turner RE, Justic D, Díaz RJ (2009) Global change and eutrophication of coastal waters. *ICES J Mar Sci* 66:1528–1537
- Rahimikhoob A (2010) Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. *Renew Energ* 35(9):2131–2135
- Sengorur B, Dogan E, Koklu R, Samandar A (2006) Dissolved oxygen estimation using artificial neural network for water quality control. *Fresen Environ Bull* 15(9a):1064–1067
- Verity PG, Yoder JA, Bishop SS, Nelson JR, Craven DB, Blanton JO, Robertson CY, Tronzo CR (1993) Composition, productivity and nutrient chemistry of a coastal ocean planktonic food web. *Cont Shelf Res* 13:741–776
- Vigil KJ (2003) *Clean water: an introduction to water quality and water pollution control*. Oregon State University Press, Portland