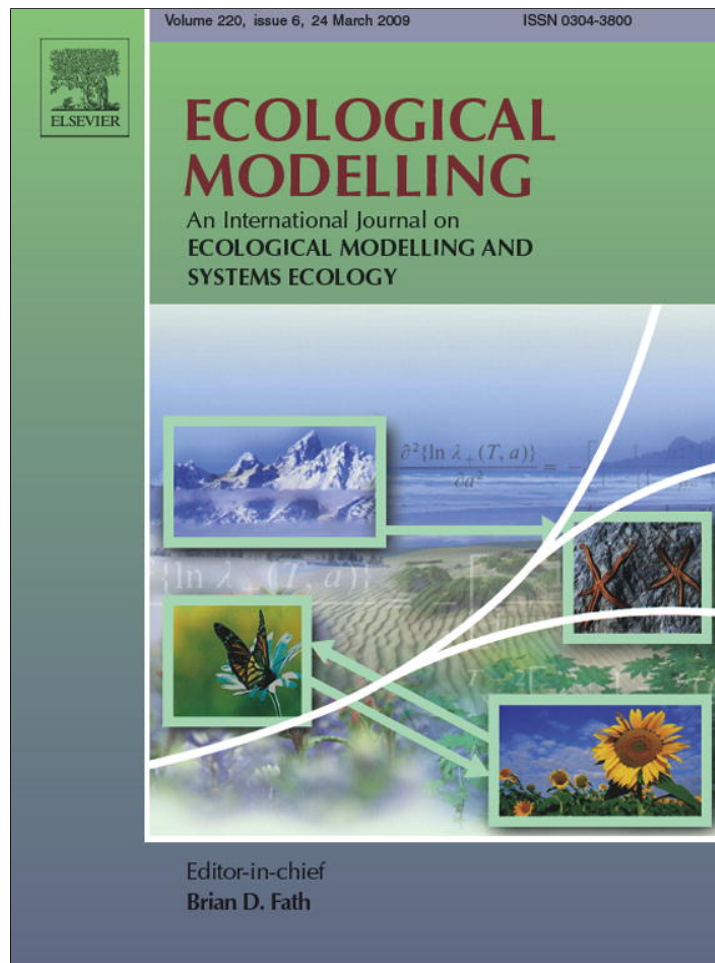


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

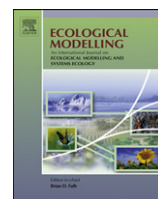
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

A neural network experiment on the simulation of daily nitrate-nitrogen and suspended sediment fluxes from a small agricultural catchment

François Anctil^{a,*}, Mélanie Filion^a, Julien Tournéville^b

^a Chaire de recherche EDS en prévisions et actions hydrologiques, Département de génie civil, Université Laval, Pavillon Pouliot, Québec G1K 7P4, Canada

^b Cemagref, Water Quality and Hydrology Research Unit, P.B. 44, 92163 Antony Cedex, France

ARTICLE INFO

Article history:

Received 11 July 2008

Received in revised form 4 December 2008

Accepted 12 December 2008

Available online 30 January 2009

Keywords:

Neural networks

Surface water quality

Nitrate-nitrogen

Suspended sediment

ABSTRACT

In this paper, we report an application of neural networks to simulate daily nitrate-nitrogen and suspended sediment fluxes from a small 7.1 km² agricultural catchment (Melarchez), 70 km east of Paris, France. Nitrate-nitrogen and sediment losses are only a few possible consequences of soil erosion and biochemical applications associated to human activities such as intensive agriculture. Stacked multilayer perceptrons models (MLPs) like the ones explored here are based on commonly available inputs and yet are reasonably accurate considering their simplicity and ease of implementation. Note that the simulation does not resort on water quality flux observations at previous time steps as model inputs, which would be appropriate, for example, to predict the water chemistry of a drinking water plant a few time steps ahead. The water quality fluxes are strictly mapped to historical mean flux values and to hydro-climatic variables such as stream flow, rainfall, and soil moisture index (12 model input candidates in total), allowing its usage even when no flux observations are available. Self-organizing feature maps based on the network structure established by Kohonen were employed first to produce the training and the testing data sets, with the intent to produce statistically close subsets so that any difference in model performance between validation and testing has to be associated to the model and not to the data subsets. The stacked MLPs reached different levels of performance simulating the nitrate-nitrogen flux and the suspended sediment flux. In the first instance, 2-input stacked MLP nitrate-nitrogen simulations, based on the same-day stream flow and on the 80-cm soil moisture index, have a performance of almost 90% according to the efficiency index. On the other hand, the performance of 3-input stacked MLPs (same-day stream flow, same-day historical flux, and same-day stream flow increment) reached a little more than 75% according to the same criterion. The results presented here are deemed already promising enough, and should encourage water resources managers to implement simple models whenever appropriate.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Human activities such as intensive agriculture, often alter surface water chemistry. Nitrate-nitrogen and sediment losses are only a few possible consequences of soil erosion and biochemical applications. Such contaminants are the prime food supply of filter feeders, regulate the water turbidity, limit light penetration and photosynthesis, accelerate the degradation of lakes and rivers, affect biodiversity, and promote the occurrence of water-related diseases. Owing to the complexity of the processes controlling water chemistry in a catchment, simple tools are needed to support decision-making and to develop best-management practices.

Catchment stakeholders and managers increasingly seek information on water quality at catchment scales. This information

is required to limit the consequence of human activities on surface water chemistry and in many instances to test ideas on how to restore water usages. Over the years, many models have been devised to achieve this goal. For instance, Gassman et al. (2007) state that “the Soil and Water Assessment Tool (SWAT) model (Arnold et al., 1998; Arnold and Fohrer, 2005) has proven to be an effective tool for assessing water resource and nonpoint source pollution problems for a wide range of scales and environmental conditions across the globe.” However, despite the tremendous efforts and ingenuities devoted to the development of water quality models such as SWAT and many others, lots of improvements are still needed. For example, Merritt et al. (2003) conclude an exhaustive review of erosion and sediment transport models noting that “physics-based models and the more complex conceptual models are not particularly appropriate for estimating catchment exports for the following reasons: lack of sufficient spatially distributed input data to drive the models, paucity of calibration data in space and time to define an appropriate parameter set for the models and hence reliable output, over-dependency of

* Corresponding author. Tel.: +1 418 656 3653; fax: +1 418 656 2928.
E-mail address: francois.anctil@gci.ulaval.ca (F. Anctil).

model results on the experience of the user." Similar conclusions have pushed many, like Nour et al. (2006), to seek simpler solutions.

Neural networks (NNs) have evolved from intelligent computer-based systems mimicking the human brain to non-linear mathematical algorithms that map a set of input variables to a set of output variables. As such, NNs need no explicit information on the processes causing the response. They are completely data driven, which often allow reducing the magnitude of the information needed to build them, in opposition to standard physically based models (Schultz and Wieland, 1997). The disadvantage of NNs is that they provide no physical or chemical information about the occurring processes—it is very difficult to decipher neural weights and biases. Nonetheless, there has been a growing trend for the experimentation of NNs in the hydrologic and water quality domains (Chau, 2006). Reports of NN applications simulating nitrate-nitrogen losses are given by Lek et al. (1996), Clair and Ehrman (1998), Lischeid (2001), Sharma et al. (2003), Suen and Eheart (2003), Yu et al. (2004), Holmberg et al. (2006), and May and Sivakumar (2008), while reports of NN applications simulating suspended sediment losses are given by Salehi et al. (2000), Sarangi and Bhattacharya (2005), Nour et al. (2006), and Zhu et al. (2007).

In this paper, we report an application of NN to simulate daily nitrate-nitrogen and suspended sediment fluxes from a small agricultural catchment that is artificially drained, a situation not often simulated by ANN. The simulation does not resort on water quality flux observations at previous time steps as model inputs, which would be appropriate, for example, to predict the water chemistry of a drinking water plant a few time steps ahead. The proposed NN models are strictly mapped to hydro-climatic variables such as stream flow, rainfall, and soil moisture index, and to historical mean flux values allowing its usage even when flux observations have ceased. First, self-organizing feature maps based on the network structure established by Kohonen (1997) were employed to produce the training and the testing data sets. Second, multilayer perceptrons, a category of NN models with efficient and practical applicability comparable to that of usual black box and conceptual hydrologic models, were employed for the simulations.

The context of application is presented in the next section, with a description of the Melarchez catchment, the data collected, and the neural networks employed. The subsequent section presents the protocol of experiment, the modelling performance criteria, and the water quality model developments. Conclusions on the relevant findings of this work are provided in the last section, with an emphasis comparing differences in NN architecture and performance for nitrate-nitrogen and suspended sediment fluxes.

2. Context of application

2.1. The Melarchez experimental basin

With an area of 7.1 km², the Melarchez experimental basin is entirely dedicated to agricultural activities, 70 km east of Paris, France. The Melarchez basin is part of the 104 km² Orgeval experimental catchment (Fig. 1) where agriculture takes place on 80% of its surface while the remaining 20% is forested. The average annual air temperature is 9.7 °C, the annual mean rainfall is 706 mm and the annual mean potential evaporation is 592 mm. Most of the basin, covered with table-land loess up to 10 m thick, is relatively flat with slopes increasing near the small valley at the river mouth (80% of the territory spans between 130 and 170 m above mean sea level).

The hydrological behaviour of the Melarchez basin is influenced by its regional geology. The unconsolidated deposits, about 2 m in thickness, are essentially composed of sand and loam lenses of low permeability. The impermeable lenses allow a temporary perched water table that can reach the soil surface in wet periods. Conse-

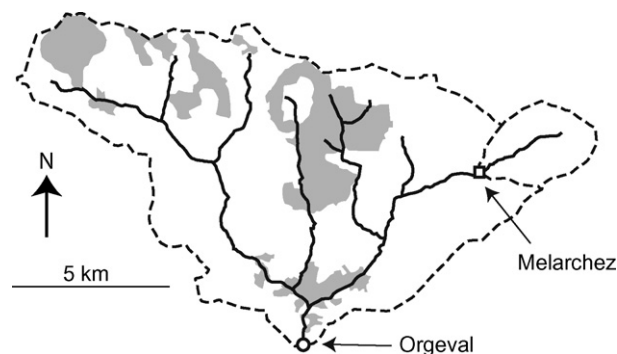


Fig. 1. The Orgeval experimental watershed near Paris, France, including the eastern sub-watershed of Melarchez. Shaded areas are forested.

quently, two water tables may coexist: one in the limestone Brie Formation aquifer and the other in the perched aquifer.

The Melarchez basin is drained by a first order 2.3-km stream whose width ranges from 3.5 m upstream to 7.5 m downstream, and depth ranges from 1 m upstream to 2 m downstream. Its surface flow mostly combines groundwater releases and discharges from the anthropogenic subsurface network that drains 75% of the basin.

The Melarchez basin is affected by non-point source pollution caused by intensive agricultural soil use and a high level of chemical fertilizers usage. In addition to agricultural sources, nitrates could result from the organic matter mineralization producing nitrites and nitrates through nitrification process in an oxidizing medium. The unassimilated nitrogen stays in transit in the soil and may be carried out by runoff, contaminating surface water, or be transported deeper into the soil, contaminating groundwater. Nitrate-nitrogen (NO₃⁻-N) concentration in ground and surface water varies as a function of time. High concentrations occur in spring, after spreading, when nitrates are in the superficial soil layer and are easier to carry by runoff. Peaks also occur in autumn and at the beginning of winter (Molénat et al., 2000), when there is a dominance of organic matter mineralization and crop residues. Lower concentrations follow flood events that leach the soil of its nutrients: soil leaching is a dilution phenomenon. In the Melarchez basin, the deep aquifer interacts with the surface water, bringing buried nitrates to the surface.

In the context of Brie Plateau, Sogon et al. (1999) showed that the two most common erosive processes in the fields are leaching of the finest soil particles from the ploughed horizon to the underground drainage network and detachment and transport of topsoil particles by overland flow. Rill and gully erosion are rare and take place on bare soils only in extreme rainfall events. Meybeck et al. (1999) compiled the spatial and temporal variability of suspended sediment in the Seine basin, which includes the Melarchez sub-basin. General levels in the Seine River are very low due to a gentle relief and to an even distribution of rainfall throughout the year. However, there is a marked increase of the suspended sediment range from Strahler stream order 8–1: highest quantiles increase and lowest quantiles decrease. For instance, suspended sediment peaks at Melarchez are produced by local rainstorms, while such summer flood are not observed at higher stream orders because rainstorms are locally distributed. The suspended sediment level at Melarchez is thus more dependent of the agricultural practices at the time of the storm. Any first and/or fast increase of river velocity is also more prone to reactivate river bed material deposited since the previous flood event.

2.2. The database

The hydrometric and water quality data available to this work consists of daily rainfall (mm), stream flow (m³), nitrate-nitrogen

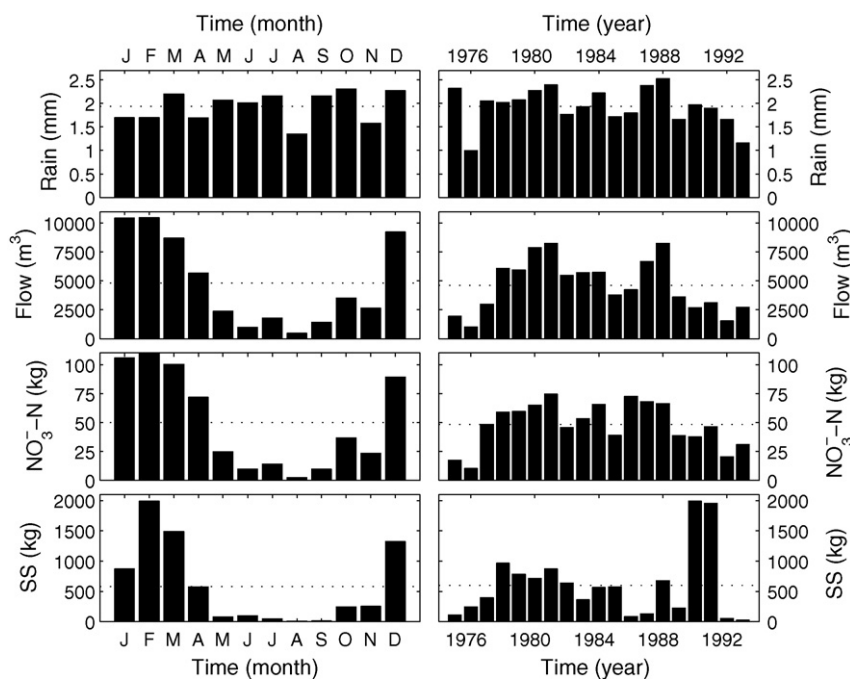


Fig. 2. Monthly (left panel) and yearly (right panel) means of the observed daily rainfall (mm), stream flow (m³), nitrogen-nitrate flux (kg), and suspended sediment flux (kg).

flux (kg), and suspended sediment flux (kg), from 1975 to 1993. All data was collected at the gauging station. Water level was transformed into stream flow by a rating curve; water and sediment were sampled twice a day, using an automatic sampler. Approximately 16% of the data is missing from the daily nitrate-nitrogen record and as much as 20% of the suspended sediment fluxes, which leaves 5605 daily observations in the first instance, and 5271 in the second.

Annual and monthly means for each daily parameter are drawn in Fig. 2. Snow fall is negligible on the catchment, where average daily temperatures persist below 0°C only for short periods of time and never consistently. The water inputs of significance are liquid precipitation, which are relatively constant on a monthly basis over the years (Lauzon et al., 2004). However, the proportion of precipitation reaching the streams shows a strong seasonality (Fig. 2). At times, very little or no stream flow is observed at the hydrometric station for relatively long periods of time despite sustained precipitation. Stream flow is mainly influenced by the soil moisture content of the perched aquifer and the intensity of the evapotranspiration fluxes (Filiz, 1973; Gomendy, 1996; Billy et al., 2008). There are three distinct stream flow phases each year. During summer (April to September), the hydrological balance is negative, exhorting the superficial aquifer to disappear and the deep aquifer to drop. During this period, flows occur only following intense precipitation, which temporarily saturate the soil. During fall (October to December), lower evapotranspiration fluxes allow precipitation to gradually recharge the deep and superficial aquifers. Flows in fall are generally higher than those of the summer, but high flows, if occurring, remain sporadic and short. During winter (December to March), the soil is usually saturated and flows are high and quick. The highest flows typically occur during winter. The monthly patterns of the stream flow and of the nitrate-nitrogen flux are almost identical, suggesting a strong link between those two parameters (Fig. 2). However, the suspended sediment flux behaves differently. The same situation prevails for the inter-annual fluctuations (Fig. 2), where moderate variations in the annual rainfall leads to larger, but similar, stream flow and nitrate-nitrogen flux inter-annual fluctuations, but different and even larger suspended sediment inter-annual fluctuations.

Small catchments like the Melarchez basin are subjected to strong daily hydrological fluctuations that are difficult to model, as illustrated in Fig. 3, 1982 is the year with the lowest number of missing observations. Even if the Melarchez climatology cannot be identified as violent with a maximum daily rainfall of only 57 mm from 1975 to 1993, daily stream flow, nitrate-nitrogen flux, and suspended sediment flux ranged from about nil to 35.7 times (September 1987), 37.5 times (February 1990), and 327 times (February 1990) their respective average value.

2.3. Neural networks as descriptive and predictive tools

The database is limited, with only a few types of inputs (i.e., rainfall, stream flow, nitrate-nitrogen flux, and suspended sediment flux), which consequently reduces the choice of models to

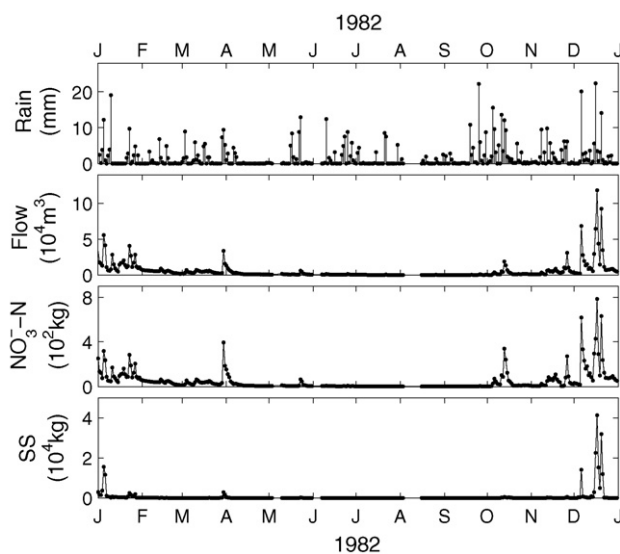


Fig. 3. Observed daily precipitation (mm), stream flow (m³), nitrogen-nitrate flux (kg), and suspended sediment flux (kg), in 1982.

simple conceptual ones or more empirical tools such as neural networks (NNs). In a hydrologic context, the ASCE Task Committee on Application of Artificial Neural Networks in Hydrology provides a well documented description of both types of network structure employed here: the self-organizing feature maps and the multilayer perceptron (ASCE, 2000a), and gives several examples of applications demonstrating the merits of such techniques (ASCE, 2000b). NNs are data driven approaches that do not depend on a formal physical or chemical reasoning to provide an answer to a given set of inputs.

Self-organizing feature maps (SOFMs), which are based on the structure established by Kohonen (1997), are a descriptive tool equivalent to clustering techniques that discriminate input vectors with respect to patterns present in them. This network is made of an input layer that receives the data and an output layer composed of several neurons often structured as a flat plane. The weights associated to a given output neuron is similar to a mass center, which can be compared against input vectors. The weights define the data patterns, and an output neuron, among all the others, is said to be activated if its weight vector most closely matches the input vector fed to the network. The calibration process, which determines the values of the weights of the network, ensures that the output layer is spread over the entire data domain and defines the range of patterns in the data in some meaningful coordinate system, which is why the Kohonen network is called a self-organizing map.

As a predictive tool, the multilayer perceptrons (MLPs) employed here for the simulation of nitrate-nitrogen and suspended sediment fluxes are typical and by far the most commonly used structure in the field of water resources (Coulibaly et al., 1999; Maier and Dandy, 2000). They consisted of one layer of inputs, one hidden layer of neurons with a sigmoid activation function, and one output neuron with a linear activation function. To ensure that all types of inputs fed to the network are on the same scale, they were linearly standardized so that their mean is zero and their standard deviation is one. Bayesian regularization, as described by MacKay (1992), Foresee and Hagan (1997), and Anctil et al. (2004b) was employed with the Levenberg–Marquardt back-propagation algorithm as the calibration procedure, where the objective function is a weighted sum of the mean of squared errors between observed and estimated water quality parameter, and the sum of the squares of the NN weights. The use of this weighted sum attempts to ensure that the NNs provide accurate output values for input values not represented in the data set employed for the calibration of the network. The second element of this weighted sum imposes a constraint that results in smaller weights, which produces a smoother network response. Another advantage of the Bayesian regularization procedure is that generalisation of the network may be achieved using only two sub data sets, instead of the three sub data sets necessary for the more commonly used stop training procedure (Anctil and Lauzon, 2004). Furthermore, the stacking method (Wolpert, 1992) is performed here. In the context of this application, it is a method by which several networks are calibrated, and the simulation of a nitrate-nitrogen flux and suspended sediment flux is obtained by calculating the mean of the responses of all these networks.

3. Protocol of experiment

3.1. Data partitioning

This section addresses several technical issues related to the development of neural network models. It starts with a clear departure from the common practice in the development of models to choose calibration and validation data sets continuous in time. In common practice, one has to be careful in the selection of time

Table 1

Summary of characteristics for the training and testing data sets.

	Nitrate-nitrogen		Suspended sediment	
	Training	Testing	Training	Testing
P_{mean} (mm)	1.98	1.84	1.95	1.94
P_{max} (mm)	56.5	57.0	56.5	57.0
P_{min} (mm)	0	0	0	0
Q_{mean} (m ³)	4,614	4,732	4,807	4,970
Q_{max} (m ³)	173,664	135,648	173,664	152,064
Q_{min} (m ³)	0	0	0	0
F_{mean} (kg)	48.4	48.8	556	662
F_{max} (kg)	1,898	1,373	168,924	193,425
F_{min} (kg)	0	0	0	0
Number	3,723	1,882	3,480	1,791

periods to ensure that both the calibration and validation data sets contain events that represent all possible conditions of flux production (Klemes, 1986). Neural networks do not need continuous streams of data, for they require only that the input vectors fed to them be complete.

The SOFM network was applied for the subdivision of the time series into a training subset and a testing subset, following recommendations by Bowden et al. (2002). The clustering was performed on vectors of concomitant: rainfall, stream flow, and flux records (the clustering is performed separately for the nitrate-nitrogen flux and for the suspended sediment flux). A few output layer configurations have been tested (3×3 , 3×4 and 4×4), and the smallest one have been kept to (1) ensure that a large number of input vectors is associated to each class of behaviours (i.e., output neuron); and (2) provide information on the behaviours of the system that is as meaningful as that provided by the larger configuration. Once the clustering is complete, the input vectors in each class are randomly divided into two subsets. Two thirds of the data was employed for training and one third for testing. Statistics of the resulting training and testing data sets are given in Table 1 confirming that the SOFM clustering was successful in leading to the creation of statistically equivalent data sets for training and testing. The SOFM procedure is applied with the intent to produce statistically close subsets so that any difference in model performance between validation and testing has to be associated to the model and not to the data subsets.

The characteristics of the water quality behaviours of the Melarchez basin are portrayed by the SOFM clustering in Tables 2 and 3, where the clusters are sorted in descending order of cluster mean daily rainfall intensity. From Table 2, it is quite obvious that the nitrate-nitrogen fluxes are more correlated to stream flow than to rainfall. For instance, cluster #3 groups the largest flows and nitrate-nitrogen fluxes but only the third largest rainfall intensities. Furthermore, in cluster #9, the highest rainfalls are associated to only the sixth highest flows and nitrate-nitrogen fluxes out of nine. On the other hand, low rainfall intensities, low flows, and low nitrate-nitrogen fluxes (cluster #4) are by far the most frequent behaviour with 2577 occurrences out of 5605 daily observations.

Table 2

Summary of characteristics for each nitrate-nitrogen flux SOFM cluster.

Cluster #	P_{mean} (mm)	Q_{mean} (m ³)	F_{mean} (kg)	Number of observations
9	12.3	2,920	28	292
6	11.7	18,913	220	119
3	6.6	49,610	528	233
8	4.7	1,666	16	483
5	1.8	3,658	36	259
7	1.4	505	4	488
2	1.2	12,889	140	374
1	0.1	4,031	40	780
4	0.1	566	4	2,577
Mean	1.9	4,654	48	5,605

Table 3
Summary of characteristics for each suspended sediment flux SOFM cluster.

Cluster #	P_{mean} (mm)	Q_{mean} (m ³)	F_{mean} (kg)	Number of observations
9	19.7	27,780	6499	143
6	9.2	3,442	256	299
8	4.7	44,949	7382	205
3	4.3	1,336	28	433
5	2.9	8,479	437	175
2	1.5	955	20	517
7	0.7	14,472	1214	271
4	0.2	4,365	178	785
1	0.1	590	13	2443
Mean	1.9	4,862	592	5271

The SOFM clustering of the suspended sediment fluxes (Table 3) leads to many features identical to the nitrate-nitrogen clustering, namely a strong stream flow influence and a large representation of low rainfall, low flow, and low suspended sediment flux events (cluster #1). However, the highest rainfalls (cluster #9) are now associated to the second highest flows and suspended sediment fluxes, which is a noticeable departure from the nitrate-nitrogen clustering.

3.2. Model input candidates

The successful development of MLP models, simulating either the nitrate-nitrogen daily fluxes or the SM daily fluxes exiting the Melarchez basin, depends largely on the availability of pertinent model input parameters. Twelve model input candidates are considered in the present study (Table 4). The first three model input candidates are the same-day stream flow (Q), the previous-day stream flow (Q_{-1}), and the same-day stream flow increment (ΔQ). Stream flow is often intimately linked to water quality fluxes, as revealed by the SOFM clustering in Tables 2 and 3. The time lag allows accounting for delays in water flowing out of the catchment—a 1-day delay is probably enough for a small catchment such as the Melarchez basin, as confirmed by hydrological analyses (Molénat et al., 2000), while the incremental value serves a similar purpose and sometimes lead to a better model (Anctil and Rat, 2005). The fourth and fifth model input candidates are the same-day and previous-day precipitation (P and P_{-1}). Precipitation is often a very important element of runoff generation because it is the only climate variable, aside from snowmelt at higher latitude or in mountainous areas that can explain rapid flow increases (Anctil et al., 2004a). The sixth and seventh model input parameters are the same-day historical mean flux (F^*) and historical mean flux increment (ΔF^*). The former is inspired by Shamseldin (1997) who showed that consideration of the mean value of the historical observation for each day of the year is often beneficial to simulation MLP models. This requires that some historical observations are available at site. For instance, in the present study, the same-

Table 4
List of model input candidates.

#	Input	Description
1	Q	Same-day stream flow (m ³)
2	Q_{-1}	Previous-day stream flow (m ³)
3	ΔQ	Same-day stream flow increment (m ³): $Q - Q_{-1}$
4	P	Same-day precipitation (mm)
5	P_{-1}	Previous-day precipitation (mm)
6	F^*	Same-day historical mean flux (kg)
7	ΔF^*	Same-day historical mean flux increment (kg): $F^* - F^*_{-1}$
8	SMI_{10}	Same-day 10-cm depth soil moisture index (mm)
9	SMI_{20}	Same-day 20-cm depth soil moisture index (mm)
10	SMI_{40}	Same-day 40-cm depth soil moisture index (mm)
11	SMI_{80}	Same-day 80-cm depth soil moisture index (mm)
12	SMI_{120}	Same-day 120-cm depth soil moisture index (mm)

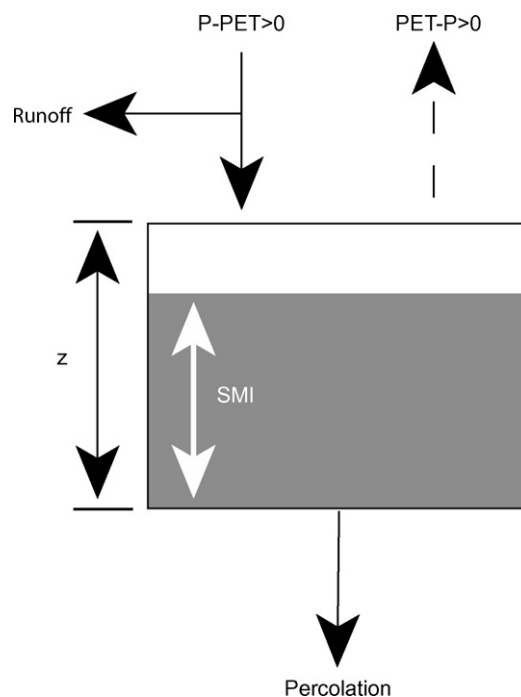


Fig. 4. Schematic of the conceptual soil reservoir model.

day historical means were based on up to 21 observations. The last five model input candidates are different variants of a soil moisture index. As an alternative to soil moisture content observations, Anctil et al. (2004a) explored a soil moisture index (SMI) derived from the GR4J hydrological model (Perrin et al., 2003), which is a conceptual formulation of the impact of precipitation P and potential evapotranspiration PET on the water balance within a reservoir of fixed depth z (Fig. 4). The SMI is the water level within the reservoir. Perrin et al. (2003) showed that the reservoir depth ranged between 10 and 120 cm for 80% of 429 selected catchments across the world. In the present study, five depths z were used to construct different soil moisture index records, that is, 10, 20, 40, 80 and 120 cm. A 1-year warm-up period was used to alleviate the problem of the initial conditions. All five calculated SMI_z time series were thus also considered as potential supplemental inputs to the MLP models.

3.3. Evaluation of performance

The evaluation of the performance is primarily based on the mean average error (MAE), a linear scoring rule, and on the root mean squared error (RMSE), a quadratic scoring rule. They both describe the average magnitude of the deviation from the observed values, range from 0 to ∞ (lower values are better), and show the errors in the same unit and scale as the parameter itself, which excludes comparisons of the performance of forecasts for basins of different sizes or with different stream flow regimes. This drawback is overcome by using a skill score, which is a simple standardization made by comparing the performance of the forecast with the performance of a reference forecast. The Nash and Sutcliffe (1970) efficiency index (EI) is a common standardization of the MSE, which ranges from $-\infty$ to 1. It reaches 1 for a perfect fit between predicted and observed values, and 0 when the hydrological model is no better than a one-parameter 'no-knowledge' model that gives the mean of the observations as simulation for all time steps. Finally, since all of the above scoring rules are averages over a data set, scatter plots are also drawn to visually assess the concordance between the simulated and the observed water quality parameters.

As this study involved a comparison of the performance of various combinations of model inputs, the r^2 criterion of Nash and Sutcliffe (1970) was applied. This criterion expresses the proportion of the initial variance unaccounted for by the base MLP that may subsequently be accounted for by the base MLP with a supplemental moisture input. The r^2 criterion is expressed by the following equation:

$$r^2 = \frac{SS_{II} - SS_I}{1 - SS_I}, \quad (1)$$

where SS is a skill score, for instance EI, I denotes a reference MLP, and II denotes the MLP with a supplemental input. Negative values of the r^2 signify that the modification to the input vectors has negative effects on the MLP performance. Senbeta et al. (1999) suggest that r^2 values greater than 10% may be considered as an indication of the significance of a supplemental model input in the overall performance of the water quality model.

4. Results

The selection of the model inputs and of the number of neurons in the hidden layer is important to neural networks. As stressed by Chau (2006), it is necessary to determine the optimal choice between available environmental parameters, in opposition to using them all. Cases involving a large number of potential model inputs imply a very large number of combinations to test, and the number of combinations increases if the number of neurons in the hidden layer is a factor to consider in the decision process. In this application, the goal was to produce separate MLPs for the simulation of the nitrate-nitrogen and suspended sediment daily fluxes F .

A preliminary analysis for this application concluded that the choice of the model inputs is more significant in terms of performance gains than that of the number of neurons in the hidden layer, as for Anctil et al. (2004a). Consequently, the calibration strategy applied here consisted of a first phase to find the most adequate combination of model inputs for the MLP model, fixing the number of hidden nodes to 5. The second phase involves finding the number of neurons in the hidden layer that provides the best performance with the chosen combination of variables.

The choice of the combination of model inputs consisted initially in testing all of them individually, and to choose the one that yield the best model performance. This best model input is combined individually with each of the remaining one, and the pair that produces the best modelling performance is retained and combined with each of the remaining model inputs individually. This process can continue until all model inputs are employed, but it was stopped in this application when the performance benefit of adding an extra

Table 5

Performance of MLPs with five hidden neurons in simulating the nitrate-nitrogen fluxes in the testing data set.

Model inputs	1-input MLPs	2-input MLPs	
	EI	EI	r^2 (%)
Q	0.857		ref
Q_{-1}	0.440	0.850	-5.1
ΔQ	0.684	0.861	2.9
P	0.106	0.845	-8.6
P_{-1}	0.224	0.873	11.2
F^*	0.138	0.876	13.3
ΔF^*	0.057	0.876	13.2
SMI_{10}	0.113	0.872	10.5
SMI_{20}	0.116	0.877	14.1
SMI_{40}	0.125	0.879	15.2
SMI_{80}	0.134	0.880	15.9
SMI_{120}	0.134	0.876	13.3

Bold values are best results.

model input was considered small when compared against the parsimony and generality of the base model. For any given combination of model inputs and number of neurons in the hidden layer tested, stacking was always employed. It implies that a model is composed of a group of MLPs. For any given combination of model inputs and number of neurons in the hidden layer, 50 MLPs are calibrated, and the global water quality estimate of the model is the mean of the estimates from all the networks. Only then the model performance is calculated.

From the twelve model input candidates in Table 4, and 5-hidden neurons MLPs, the most pertinent input for nitrate-nitrogen simulation is the same-day stream flow Q , which alone leads to an efficiency index of 0.857 in the testing data (Table 5), no other model input candidate stands as a potential alternative to same-day stream flow. The remarkable link between stream flow and nitrate-nitrogen flux was already evidenced by the SOFM clustering in Table 2. This very simple model is improved by 15.9% according to Nash–Sutcliffe r^2 (EI = 0.880) when the 80-cm soil moisture index is used as a second model input, revealing that soil moisture plays a role in the level of the nitrate-nitrogen flux exiting the catchment. All 3-input MLPs failed in improving the performance over the selected 2-input MLP. After optimisation of the number of hidden neurons, ranging from 2 to 20, the best EI performance reached 0.888 when 12 hidden neurons are used, a further Nash–Sutcliffe r^2 gain in performance of 7.2% when compared to the 5-hidden neurons MLPs. The final 2-input stacked MLPs have a MAE of 14 kg and a RMSE of 41 kg. Such a large difference between MAE and RMSE reveals that the error is non-uniform, as illustrated by the scatter plot of the observed versus the simulated nitrate-nitrogen flux (Fig. 5). This graph shows the limits of the stacked MLPs to

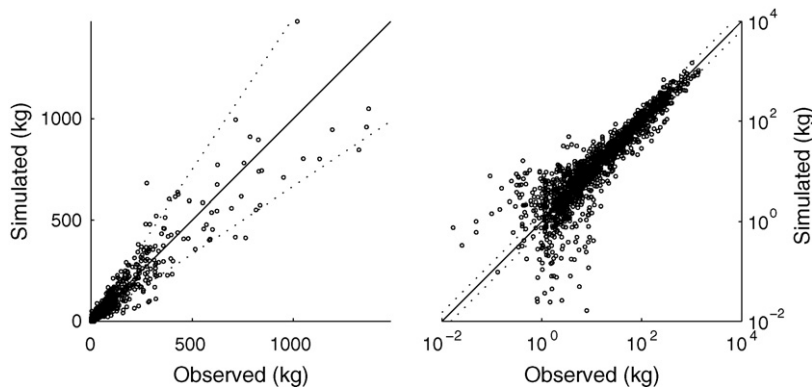


Fig. 5. Observed versus simulated nitrogen-nitrate fluxes on standard (left panel) and logarithmic (right panel) scales with the 2-12-1 MLP using Q and SMI_{80} as inputs. The dashed lines delimit errors smaller than 50%.

hal-00584129, version 1 - 7 Apr 2011

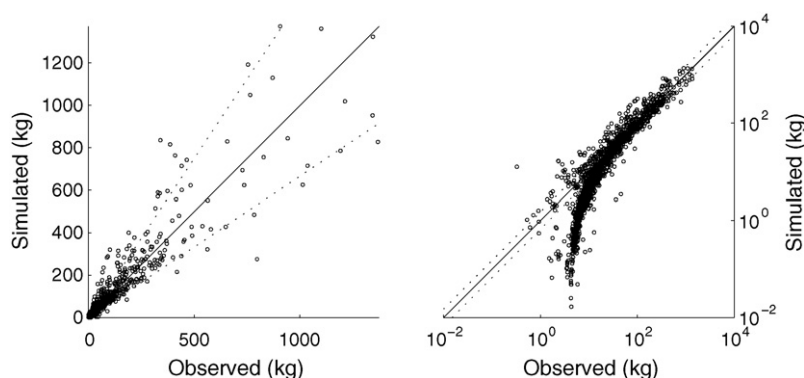


Fig. 6. Observed versus simulated nitrogen-nitrate fluxes on standard (left panel) and logarithmic (right panel) scales with the multi-linear model using Q and SMI_{80} as inputs. The dashed lines delimit errors smaller than 50%.

Table 6
Performance of MLPs with five hidden neurons in simulating the suspended sediment fluxes in the testing data set.

Model inputs	1-input MLPs		2-input MLPs		3-input MLPs	
	EI		EI	r^2 (%)	EI	r^2 (%)
Q	0.173			ref		
Q_{-1}	0.059	0.173	0.0		0.715	-0.6
ΔQ	0.125	0.170	-0.4		0.769	18.5
P	0.019	0.174	0.0		0.752	12.4
P_{-1}	0.042	0.197	2.8		0.751	12.2
F	0.011	0.716	65.7			ref
ΔF	0.009	0.625	54.7		0.691	-9.0
SMI_{10}	0.021	0.207	4.0		0.573	-50.5
SMI_{20}	0.020	0.195	2.6		0.738	7.5
SMI_{40}	0.021	0.208	4.2		0.736	7.1
SMI_{80}	0.023	0.216	5.2		0.705	-3.9
SMI_{120}	0.023	0.237	7.7		0.734	6.3

Bold values are best results.

simulate nitrate-nitrogen of the small Melarchez catchment which is characterized by fast reaction to winter rainfalls (see Fig. 3). Nonetheless, most large nitrate-nitrogen fluxes are simulated with an error smaller than 50%. For comparison, a multi-linear model was built using the same two input parameters as for the proposed NN model. The performance of this simpler model is lesser (testing data: $EI = 0.849$, $MAE = 16$ kg, and $RMSE = 47$ kg), but it may be seen in Fig. 6 that discrepancies between simulated and observed fluxes mostly occur for lower flux values.

Linkage between stream flow and suspended sediment flux is not as strong as for the nitrate-nitrogen flux simulations (see Table 3). Nonetheless, stream flow also produced the best 1-input 5-hidden neurons stacked MLPs performance of the available 12 model input candidates. However, in this case, the efficiency index

only reaches 0.173 (Table 6). A second model input, namely the same-day historical flux, is needed for performance to increase to an EI of 0.716: a Nash–Sutcliffe r^2 performance leap of 65.7%. This model is further improved by 18.5% after the inclusion of the same-day stream flow increment as a third model input. All 4-input MLPs failed in improving the performance over the selected 3-input MLP. The number of hidden input of the resulting 3-input stacked MLPs is then optimized. A further gain of only 1.1% is achieved using 7 hidden neurons. The final 3-input 7-hidden neurons stacked MLPs has EI of 0.769, MAE of 600 kg and $RMSE$ of 2700 kg. The scatter plots in Fig. 7 show that simulations are biased for low values of the stream flow and that errors are often larger than 100%, especially for smaller suspended sediment flux events. In this case, a multi-linear model exploiting the same three inputs than the proposed NN model (Fig. 8) fails to simulated the target fluxes (testing data: $EI = -3.88$, $MAE = 1000$ kg, and $RMSE = 4830$ kg).

5. Discussion and conclusion

Even if simulation of water quality fluxes from a small agricultural catchment is the major objective of this work, issues related specifically to the development of models based on multilayer perceptron neural networks must be addressed at the same time. Such issues affect modelling performance and consequently influence the appreciation of the results. The construction of calibration and validation data sets, the selection of input variables and the number of neurons in the hidden layer, the choice of employing Levenberg–Marquardt with Bayesian regularization as calibration procedure and the consideration of using stacking to obtain response distributions have been performed as a way to build a rigorous model development process that highlights MLPs potential to simulate water quality fluxes.

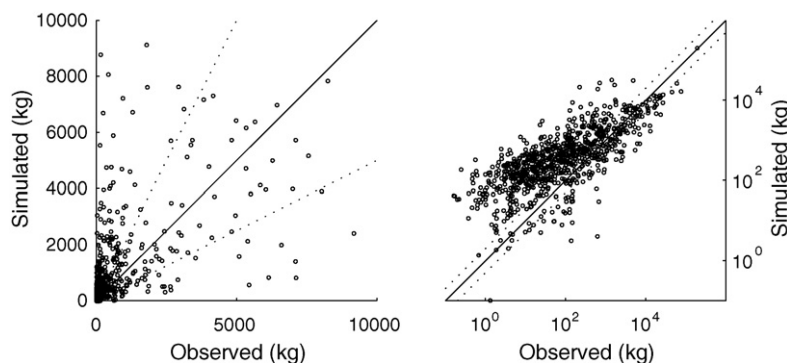


Fig. 7. Observed versus simulated suspended sediment fluxes on standard (left panel—for values up to 10,000 kg) and logarithmic (right panel) scales with the 3-7-1 MLP using Q , F and ΔQ as inputs. The dashed lines delimit errors smaller than 100%.

hal-00584129, version 1 - 7 Apr 2011

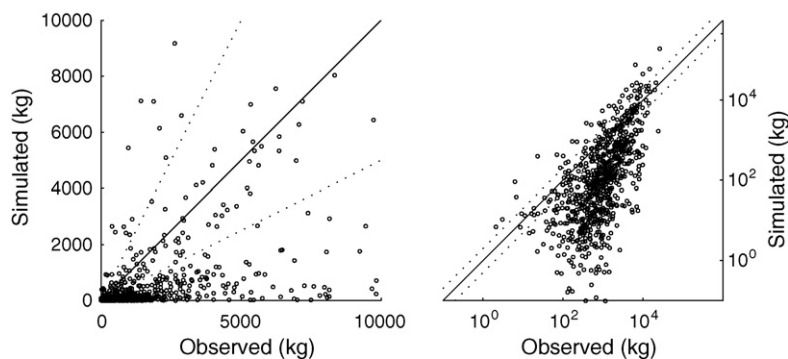


Fig. 8. Observed versus simulated suspended sediment fluxes on standard (left panel—for values up to 10,000 kg) and logarithmic (right panel) scales with the multi-linear model using Q , F and ΔQ as inputs. The dashed lines delimit errors smaller than 100%.

Tile drainage present on this catchment short-circuits the natural pathway of water. This favours pollutants that circulate through the ground at the expense of those associated to surface runoff. Tile drainage is thus more favourable to soluble pollutant losses like nitrate than to surface particulate pollutant losses. However, as shown by Sogon et al. (1999), the widespread usage of tile drainage in the Brie Plateau also favours the leaching of the finest soil particles from the ploughed horizon. Incidentally, the fact that the same-day historical flux is a suspended sediment model input shows that these exports are seasonally marked, while their level is a function of the soil moisture content and cover state, which are influenced by the agricultural practices.

The stacked MLPs reached different levels of performance simulating the nitrate-nitrogen flux and the suspended sediment flux. In the first instance, the nitrate-nitrogen simulations exhibit a quasi-linear relation and have a performance of almost 90% according to the efficiency index. On the other hand, the performance of the stacked MLPs reached only a little more than 75% according to the same criterion, leading to a MAE score that is about 3–4 times larger than for the nitrate-nitrogen flux, taking into account the scale difference between the two fluxes. This second model fully benefited from the non-linear capacity of the NNs, even if simulations are biased for low values of the stream flow, while its multi-linear counterpart failed modelling the suspended sediment flux. However, it should be stressed that the suspended sediment flux observations are spread over about six orders of magnitude (see Fig. 7) which is twice as much than the nitrate-nitrogen flux (see Fig. 5).

Even though both contaminants were sampled at the same location and using a similar experimental protocol, it is obvious from the model input selection and from the final model performance that their behaviours are quite distinct. No prior scientific judgement was imposed when constructing the MLPs. Both fluxes were confronted to the same twelve potential inputs that allowed the consideration of three hydrological processes: rain events potentially generating stream flow, soil water content governing the flow path within the catchment, and the stream flow itself, by which the contaminants are transported. Actually, the suspended sediment, because of its particulate form, is governed by mechanisms in addition to the ones already mentioned. The particulate flux may originate from two sources: land processes and fluvial processes. As such, in opposition to the nitrate-nitrogen flux, it is much more likely to vary from location to location. Indeed, the suspended sediment flux may be reduced locally by a temporary or permanent deposition of the suspended particles or be increased by a handing-over of particles locally available.

Stacked MLP models like the ones explored here are based on commonly available inputs, yet are reasonably accurate considering their simplicity and ease of implementation. The results presented

here are deemed already promising enough, and should encourage water resources managers to implement simple models whenever appropriate. All concepts presented in this study can easily be duplicated for other similar catchments.

Acknowledgment

Financial support for the undertaking of this work has been provided by the Natural Science and Engineering Research Council of Canada.

References

- Ancil, F., Rat, A., 2005. Evaluation of neural network streamflow forecasting on 47 watersheds. *J. Hydrol. Eng.* 10 (1), 85–88.
- Ancil, F., Lauzon, N., 2004. Generalisation for neural networks through data sampling and training procedures, with applications to streamflow predictions. *Hydrol. Earth Syst. Sci.* 8 (5), 940–958.
- Ancil, F., Michel, C., Perrin, C., Andréassian, V., 2004a. A soil moisture index as an auxiliary ANN input for stream flow forecasting. *J. Hydrol.* 286, 155–167.
- Ancil, F., Perrin, C., Andréassian, V., 2004b. Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environ. Model. Softw.* 19, 357–368.
- Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large-area hydrologic modeling and assessment: part I. Model development. *J. Am. Water Resour. Assoc.* 34 (1), 73–89.
- Arnold, J.G., Fohrer, N., 2005. SWAT2000: current capabilities and research opportunities in applied watershed modeling. *Hydrol. Process.* 19 (3), 563–572.
- ASCE, 2000a. Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydrol. Eng.* 5 (2), 115–123.
- ASCE, 2000b. Artificial neural networks in hydrology. II: Hydrologic applications. *J. Hydrol. Eng.* 5 (2), 124–137.
- Billy, C., Birgand, F., Sebilo, M., Billen, G., Tournebise, J., Kao, C., 2008. Nitrate dynamics in artificially drained nested watersheds. *Phys. Chem. Earth Parts A/B/C*, doi:10.1016/j.pce.2008.09.007.
- Bowden, G.J., Maier, H.R., Dandy, G.C., 2002. Optimal division of data for neural network models in water resources applications. *Water Resour. Res.* 38 (2), 1010, doi:10.1029/2001WR000266.
- Chau, K.W., 2006. A review on integration of artificial intelligence onto water quality modelling. *Mar. Pollut. Bull.* 52, 726–733.
- Clair, T., Ehrman, J.M., 1998. Using neural networks to assess the influence of changing seasonal climates in modifying discharge, dissolved organic carbon, and nitrogen export in eastern Canadian rivers. *Water Resour. Res.* 34 (3), 447–455.
- Coulibaly, P., Ancil, F., Bobée, B., 1999. Prévion hydrologique par réseaux de neurones artificiels: État de l'art. *Can. J. Civil Eng.* 26 (3), 293–304.
- Filiz, S., 1973. Étude du ruissellement et de l'infiltration sur le bassin versant expérimental de l'Orgeval à l'aide de l'oxygène 18. Ph.D. thesis. Dynamic geology. Université Paris VI, Paris.
- Foresee, F.D., Hagan, M.T., 1997. Gauss-Newton approximation to Bayesian learning. Proceedings, 1997 IEEE International Conference on Neural Networks, vol. 3. Houston, TX, pp. 1930–1935.
- Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. *Trans. ASABE* 50 (4), 1211–1250.
- Gomendy, V., 1996. Variabilités spatiales et temporelles des propriétés structurales et hydriques des horizons de surface de la couverture limoneuse du bassin versant de l'Orgeval (Brie). Ph.D. thesis. Pedology, Université Nancy I, Nancy, France.
- Holmberg, M., Forsius, M., Starr, M., Huttunen, M., 2006. An application of artificial neural networks to carbon, nitrogen and phosphorus concentrations in three boreal streams and impacts of climate change. *Ecol. Model.* 195, 51–60.

- Klemes, V., 1986. Operational testing of hydrological simulation models. *Hydrol. Sci. J.* 31 (1), 13–24.
- Kohonen, T., 1997. *Self-organizing Maps*, second ed. Springer-Verlag, Berlin, p. 426.
- Lauzon, N., Anctil, F., Petrinovic, J., 2004. Characterization of soil moisture conditions at temporal scales from a few days to annual. *Hydrol. Process.* 18, 3235–3254, doi:10.1002/hyp.5656.
- Lek, S., Dimopoulos, I., Fabre, A., 1996. Predicting phosphorus and phosphorus load from watershed characteristics using back-propagation neural networks. *Acta Ecol.* 17 (1), 43–53.
- Lischeid, G., 2001. Investigating trends of hydrochemical time series of small catchments by artificial neural networks. *Phys. Chem. Earth (B)* 26, 15–18.
- MacKay, D.J.C., 1992. Bayesian interpolation. *Neural Comput.* 4 (3), 415–447.
- Maier, H.R., Dandy, G.C., 2000. Neural networks for prediction and forecasting of water resources variables: review of modelling issues and applications. *Environ. Model. Softw.* 15 (1), 101–124.
- May, D., Sivakumar, M., 2008. Comparison of artificial neural network and regression models in the prediction of urban stormwater quality. *Water Environ. Res.* 80 (1), doi:10.2175/106143007X184591.
- Merritt, W.S., Letcher, R.A., Jakeman, A.J., 2003. A review of erosion and sediment transport models. *Environ. Model. Softw.* 18, 761–799.
- Meybeck, M., Idlafkih, Z., Fauchon, N., Andreassian, V., 1999. Spatial and temporal variability of total suspended solids in the Seine basin. *Hydrobiologia* 410, 295–306.
- Molénat, J., Davy, P., Gascuel-Oudou, C., Durand, P., 2000. Spectral and cross-spectral analysis of three hydrological systems. *Phys. Chem. Earth (B)* 25, 391–397.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: part I—a discussion of principles. *J. Hydrol.* 10, 282–290.
- Nour, M.H., Smith, D.W., El-Din, M.G., Prepas, E.E., 2006. Neural networks modelling of streamflow, phosphorus, and suspended solids: application to the Canadian Boreal forest. *Water Sci. Technol.* 53 (10), 91–99.
- Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. *J. Hydrol.* 279, 275–289.
- Salehi, F., Prasher, S.O., Amin, S., Madani, A., Jebelli, S.J., Ramaswamy, H.S., Tan, C., Drury, C.F., 2000. Prediction of annual nitrate-N losses in drain outflows with artificial neural networks. *Trans. ASAE* 43, 1137–1143.
- Sarangi, A., Bhattacharya, A.K., 2005. Comparison of artificial neural network and regression models for sediment loss prediction from Banha watershed in India. *Agric. Water Manage.* 78, 195–208.
- Schultz, A., Wieland, R., 1997. The use of neural networks in agroecological modelling. *Comput. Electron. Agric.* 18, 73–90.
- Senbeta, D.A., Shamseldin, A.Y., O'Connor, K.M., 1999. Modification of the probability-distributed interacting storage capacity model. *J. Hydrol.* 224, 149–168.
- Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modelling. *J. Hydrol.* 199, 272–294.
- Sharma, V., Negi, S.C., Rudra, R.P., Yang, S., 2003. Neural networks for predicting nitrate-nitrogen in drainage water. *Agric. Water Manage.* 63, 169–183.
- Sogon, S., Penven, M.-J., Bonte, P., Muxart, T., 1999. Estimation of sediment yield and soil loss using suspended sediment load and ¹³⁷Cs measurements on agricultural land, Brie Plateau, France. *Hydrobiologia* 410, 251–261.
- Suen, J.P., Eheart, J.W., 2003. Evaluation of neural networks for modeling nitrate concentrations in rivers. *J. Water Res. Plan. Manage. ASCE* 129 (6), 505–510.
- Wolpert, D.H., 1992. Stacked generalisation. *Neural Netw.* 5, 241–259.
- Yu, C., Northcott, W.J., McIsaac, G.F., 2004. Development of an artificial neural network for hydrologic and water quality modeling of agricultural watersheds. *Trans. ASAE* 47, 285–290.
- Zhu, Y.M., Lua, X.X., Zhou, Y., 2007. Suspended sediment flux modeling with artificial neural network: an example of the Longchuanjiang River in the Upper Yangtze Catchment, China. *Geomorphology* 84, 111–125.