Environmental Technology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t791546829

Neural networks modelling of nitrogen export: model development and application to unmonitored boreal forest watersheds

X. Li[a]; M. H. Nour[b]; D. W. Smith[a]; E. E. Prepas[cd]
[a] Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada [b] ISL Engineering and Land Services Ltd., Edmonton, AB, Canada [c] Faculty of Forestry and the Forest Environment, Lakehead University, Thunder Bay, ON, Canada [d] Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

Online publication date: 31 March 2010

To cite this Article Li, X. , Nour, M. H. , Smith, D. W. and Prepas, E. E.(2010) 'Neural networks modelling of nitrogen export: model development and application to unmonitored boreal forest watersheds', Environmental Technology, 31: 5, 495 — 510
To link to this Article: DOI: 10.1080/09593330903527880
URL: http://dx.doi.org/10.1080/09593330903527880

Environmental Technology
Vol. 31, No. 5, 14 April 2010, 495–510

Taylor & Francis
Taylor & Francis Group

Neural networks modelling of nitrogen export: model development and application to unmonitored boreal forest watersheds

X. Li*[a], M.H. Nour[b], D.W. Smith[a] and E.E. Prepas[c]

[a]*Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada, T6G 2W2;* [b]*ISL Engineering and Land Services Ltd., Edmonton, AB, Canada, T6E 5L9;* [c]*Faculty of Forestry and the Forest Environment, Lakehead University, Thunder Bay, ON, Canada, P7B 5E1, and Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada, T6G 2E1*

(*Received 6 May 2009; Accepted 1 December 2009*)

In remotely located boreal forest watersheds, monitoring nitrogen (N) export in stream discharge often is not feasible because of high costs and site inaccessibility. Therefore, modelling tools that can predict N export in unmonitored watersheds are urgently needed to support management decisions for these watersheds. The hydrological and biogeochemical processes that regulate N export in streams draining watersheds are complex and not fully understood, which makes artificial neural network (ANN) modelling suitable for such an application. This study developed ANN models to predict N export from watersheds relying only on easily accessible climate data and remote sensing (RS) data from the public domain. The models were able to predict the daily N export ($g/km^2/d$) in five watersheds ranging in size from 5–130 km^2 with reasonable accuracy. Similarity indices were developed between any two studied watersheds to quantify watershed similarity and guide the transferability of models from monitored watersheds to unmonitored ones. To demonstrate the applicability of the ANN models to unmonitored watersheds, the calibrated ANN models were used to predict N export in different watersheds (unmonitored watersheds in this perspective) without further calibration. The similarity index based upon a rainfall index, a peatland index and a RS normalized difference water index showed the best correlation with the transferability of the models. This study represents an important first step towards transferring ANN models developed for one watershed to unmonitored watersheds using similarity indices that rely on freely available climate and RS data.

Keywords: artificial neural networks; nitrogen; modelling; watershed similarity; remote sensing

Introduction

Water quality modelling in lakes and streams involves understanding physical, geochemical and biological processes in the surrounding watershed, which in turn are regulated by interrelated factors, such as vegetation, soils, geology, weather conditions and anthropogenic disturbances in the watershed. However, relationships between water quality parameters and watershed features and processes are complex, not deterministic, and currently are not fully understood. Artificial neural network (ANN) models are capable of modelling complicated and non-linear processes; therefore they have gained popularity in water quality modelling applications over the past decade [1]. Additional features of ANN models that contribute to their utility for surface water quality modelling are: (1) they are capable of identifying relationships between inputs and outputs without fully understanding the mechanistic principles behind them; (2) they can work well even when the training data set contains noisy data; and (3) they are relatively easy to learn and use.

ANN models have demonstrated strengths over conventional statistical and mechanistic models, especially when only sparse, gapped data are available for model training [2,3]. Conventional statistical time series modelling approaches (e.g., autoregressive moving average model) assume that the time series is stationary and develop equations to describe the problem to achieve statistical optimality. In contrast, ANN practitioners do not need to make such assumptions [4]. ANN models have also proven to be superior to conventional statistical time series models by a number of case studies [5–8] because they can handle non-linear problems and non-stationary data sets [5,7]. Compared with mechanistic models, ANN models can provide comparable modelling accuracy but are more applicable in practice when professional expertise and data are limited. Case studies selected ANN models because

*Corresponding author. Email: xiangfei@ualberta.ca

ISSN 0959-3330 print/ISSN 1479-487X online

DOI: 10.1080/09593330903527880
http://www.informaworld.com

they can satisfy the modelling objectives by using only routine monitoring data [9–11].

Nitrogen (N) is one of the most significant water quality parameters that affect ecological health. Factors that change the availability and cycling of N in forest ecosystems include climate gradients and variations, atmospheric N deposition, soil types, vegetation cover, hydrologic pathways and landscape disturbance [12]. The export of N from a watershed to its drainage streams is very complex due to the interaction between N export and the factors above. It is very difficult to mathematically represent these factors, because they are non-linearly related, spatially distributed on a watershed scale and exhibit temporal variation. The complexity of the system makes ANN modelling a suitable alternative for N export modelling.

Feed-forward multilayer perceptron (MLP) trained with the error back-propagation (BP) algorithm is the most widely used neural network for water quality modelling [13,14]. MLP has been used to model a variety of water quality parameters such as sediment loads [15–17], phosphorus concentrations [7,10], microbial contamination [9,18,19], and phytoplankton communities [20–22]. In terms of N modelling, MLP has been applied to simulate nitrate leaching in agricultural drainage effluent [23], forecast nitrate loads on an agricultural watershed based on historical data [24], predict total and inorganic N concentrations in 927 streams in the United States from features in mixed-use (forested, agricultural and urban) watersheds [25] and simulate annual total N export from 15 predominantly forested river basins in Atlantic Canada over a 10-year period using climate data and watershed features [26].

However, there are limited applications of ANNs in modelling N export in surface waters draining forest – and particularly boreal forest – watersheds at a daily time scale. Monitoring N export in these watersheds is often logistically unfeasible due to the associated high costs and the relative inaccessibility of many of these sites. Therefore, the use of modelling tools that can predict N in unmonitored watersheds is urgently needed to support decision making in watershed management in the boreal forest. ANN modelling of N in unmonitored watersheds relies on transferring calibrated models based on other watersheds to the unmonitored ones of interest. The more similar each pair of watersheds is, the higher the probability of success in model transferability. Thus, it is important to develop indices that can quantify watershed similarity, which can potentially guide the transferability of models from one watershed to the other. There is limited literature available for water quality model transfer indices. For hydrological model transfer, the median Euclidean distance worked out from annual water budget, greenness fractions, and physical distances have been used to measure watershed hydrologic similarity [27].

The Forest Watershed and Riparian Disturbance (FORWARD) project has been monitoring streamflow, water quality and weather in boreal forest watersheds in the Swan Hills of Alberta, Canada since 2001 [28,29]. The objectives of this study were to use N export data collected as part of the FORWARD project from five of these watersheds to: (1) develop ANN models that can predict daily N export in stream channels draining each of the studied watersheds based on easily accessible climate and remote sensing (RS) data; (2) develop watershed similarity indices for these watersheds using the same climate and RS data; (3) apply the five developed models from one watershed to the other watersheds, without further calibration, and evaluate the performance of model transfer; and (4) relate indices of watershed similarity to model performance to determine the optimal similarity index, which can then be used to guide model transfer to unmonitored watersheds. This research provides an important first step toward using climate and RS data to model N export in unmonitored watersheds.

Study area and database

The study area is located in the Swan Hills, northwest of Edmonton, Alberta, Canada. The five watersheds under study range in size from 5.1–129.4 km^2 (Table 1, Figure 1). The peatland and riparian cover data summarized in Table 1 have been previously documented [30].

Rainfall and temperature data were acquired from seven public-domain weather stations and fire towers close to the study area (Figure 1). Satellite-derived vegetation indices (e.g., enhanced vegetation index (EVI) and reflectance values at certain desired wavelengths) were acquired from Moderate Resolution Imaging Spectroradiometer (MODIS) through the National Aeronautics and Space Administration (NASA) from the years 2001 through 2005. Data acquired from MODIS were exported using the software Geomatica V9.1 [31]. The exported MODIS images were then loaded into ArcGIS 9.2 [32], overlaid by the watershed shape files, and the correspond-

Table 1. Area and soil coverage in the studied watersheds.

Watershed	Peatland, %	Riparian, %	Area, km^2
1A	25.2	0	5.1
Cassidy	4.8	0.6	5.9
Thistle	10.5	4.6	8.5
Two Creek	17	2.4	129.4
Willow	10	3.4	15.6

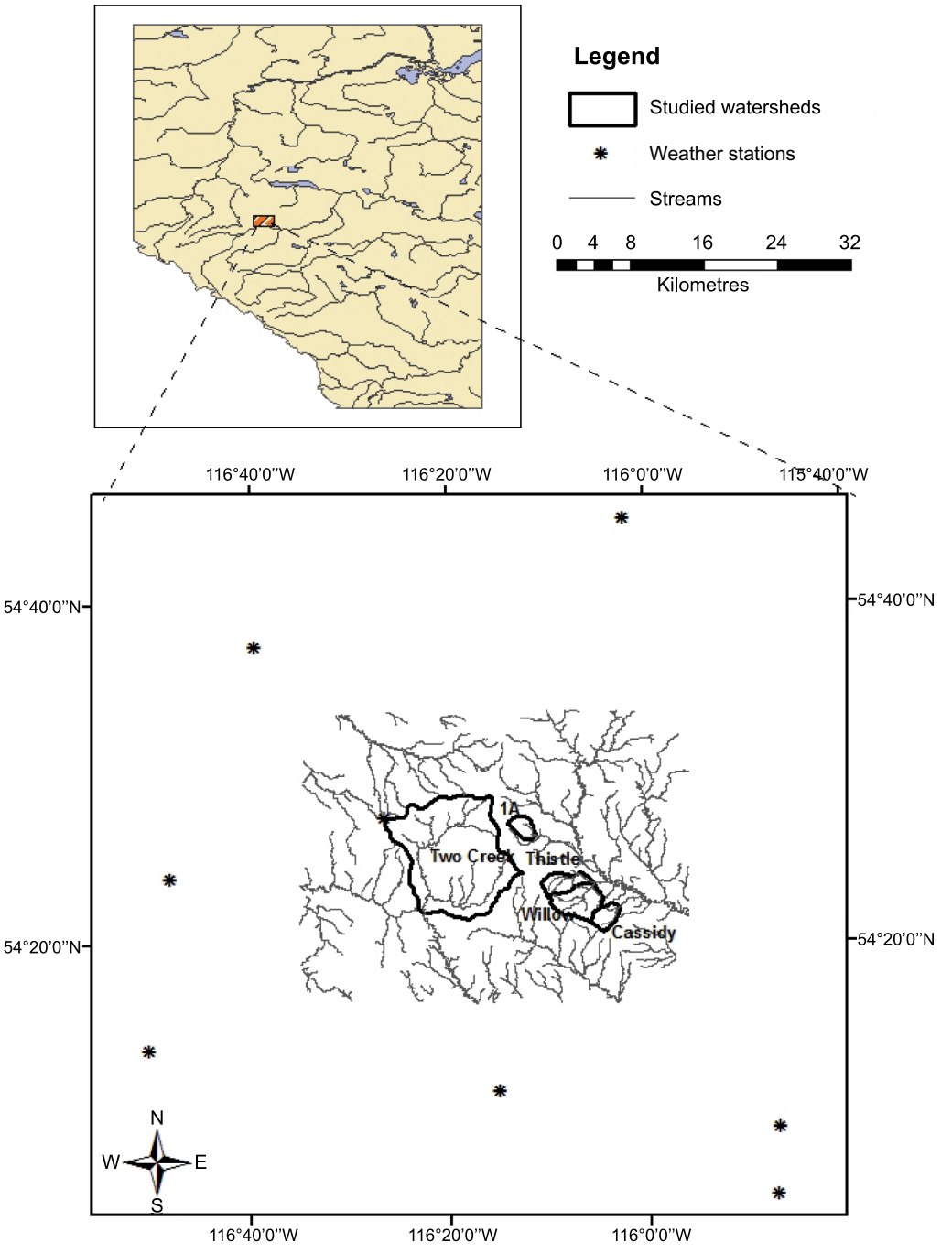


Figure 1. Study area: the watersheds under study and the weather stations.

ing data for each watershed were extracted and averaged over the watershed area using ArcGIS spatial analyst tools. Streamflow (m^3/s) and total dissolved N (TDN) concentration (μg/L) data were obtained from the FORWARD database for monitoring stations situated at each watershed outlet during the open water season (typically from April to October). For details of streamflow gauging and water sample collection in FORWARD streams, see Prepas *et al.* [30]. The areal TDN export (TDN_E) ($g/km^2/d$) for each watershed was calculated as indicated by the following equation:

$$TDN_E = \frac{Q \times TDN_C}{A} \times \left(\frac{24 \times 60 \times 60}{1000} \right) \quad (1)$$

where Q is the daily average streamflow in m^3/s, $TDNc$ is the daily average TDN concentration in μg/L and A is the area of watershed in km^2.

MLP model development

To ensure adequate model performance, all models were systematically developed following procedures that are commonly used for ANN model building in environmental science and engineering: input determination, data pre-processing, data division, determination of model internal parameters, selection of the model training algorithm and stopping criterion, and model evaluation [1,14].

Input determination

To develop a robust ANN model, it is critical to carefully select a representative and significant set of input variables. Inclusion of correlated, noisy or non-significant inputs only serves to increase the computational complexity, make training more difficult, and affecting the generality of the developed models [8,33]. This study focused on developing a data-driven modelling approach for daily TDN export predictions. The mechanisms governing N export from watersheds can be found in detail in the literature [34,35]. Generally speaking, TDN export is highly correlated with time and has seasonal fluctuation. Also, the influence of factors such as rainfall on TDN export exhibits a time delay effect. Therefore, model inputs were divided into cause/effect inputs, inputs reflecting the seasonal cyclic nature of the modelled variable and time-lagged inputs. These inputs were determined based on a combination of *a priori* knowledge of the system being modelled and trial-and-error screening by ANNs. Since this study aims at developing models for unmonitored watersheds, the selected model inputs should be easily accessible and obtainable without on-site measurements.

In general, daily TDN export is influenced by the available sources of N in the watershed and the momentum for N release from the watershed. The latter is correlated to streamflow, which is mainly controlled by rainfall and snowmelt [36–38]. In North America, rainfall information is widely available from public domain weather stations and fire towers. Daily snowmelt can be estimated by the temperature-index approach, because a linear function of daily snowmelt and average air temperature exists, given that the air temperature exceeds a base temperature. The cumulative degree days (*dd*), as represented by the following equation, can serve as an integrated measure of heat energy available to melt snow and can act as a surrogate to the temperature-index snowmelt approach:

$$dd = \sum_{i=0}^{i=N-1} \left(T_{avg(i)} - T_{b(i)}\right) \cdot \left(t_{i+1} - t_i\right) \qquad (2)$$

where dd is the total degree days at time t in °C day, T_{avg} is the daily average air temperature in °C, T_b is a base temperature typically set at 0°C, N is the number of days during which $T_{avg} \geq T_b$ and $(t_{i+1} - t_i)$ is typically taken as 1 day.

Air temperature is another climatic parameter that can be accessed from public weather stations and is correlated with N export. Increases in air temperature with decreases in precipitation lead to large decreases in runoff and hence N export [26], particularly, in the study area [30].

The recent rapid development of RS technology and the reduced cost of acquiring RS data now make it possible to take into consideration vegetation phenology, one of the most important factors affecting the N cycle. Forest vegetation can affect N cycling in a watershed and act as a temporary sink for N taken up as a nutrient [36]. In many cases, a major portion of the annual N export occurs during spring snowmelt because N mineralization has occurred under the snowpack during winter [39] and the uptake of N moving with snowmelt runoff by forest vegetation is minimal [40]. The MODIS sensor on board Terra launched by NASA in December 1999 has greatly improved scientists' ability to measure plant growth on a global scale, with moderate spatial (250m × 250m pixel size) and temporal resolution. The RS EVI provided by the MODIS Land Group has shown a high correlation with vegetation conditions [41–43]. EVI is a relatively new data product developed by the MODIS Science Team to improve upon the quality of its predecessor, the normalized difference vegetation index (NDVI), for forested ecosystems. The successful application of EVI to vegetation dynamics indicates that it has the potential for reflecting vegetation dynamics and soil/vegetation interactions during N model construction. The EVI makes use of an atmospheric resistance term by adding information from the blue wavelength and two constants, C_1 and C_2. In addition, it uses a canopy adjustment term to minimize the effect of the changes of optical properties of soil background by introducing a constant, L. EVI is formulated by:

$$EVI = G \times \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + C_1\rho_{Red} - C_2\rho_{Blue} + L} \qquad (3)$$

where $G = 2.5$, $C_1 = 6$, $C_2 = 7.5$, and $L = 1$. The terms ρ_{Blue}, ρ_{Red}, and ρ_{NIR} represent the reflectance at the blue (0.45–0.52 μm), red (0.6–0.7 μm) and near-infrared (0.7–1.1 μm) wavelengths, respectively.

Based on an understanding of the processes involved in N modelling, the most significant cause/effect factors that can be obtained from public-domain databases for unmonitored watersheds are rainfall (R), a

snowmelt indicator (*dd*), average air temperature (T_{mean}), and a vegetation growth indicator (*EVI*). Among those factors, rainfall and snowmelt has a time delay impact on N export [11]. In turn, the time lags of rainfall and snowmelt were determined for the N export model.

The seasonal periodicity of the modelled parameter was accounted for by assigning Julian day of the year to each daily record. To account for the long-term cyclic nature, a year index taking the value of either '−1' or '+1' was added to the vector of inputs. A year was given a value of '−1' if the total rainfall in the open water season of that year was lower than the 30-year-average rainfall sum; otherwise, the year index was given a value of '+1'. Assigning a time index to each data record has proven to be successful in helping the ANN to identify the periodicity of data series in other applications [44–46].

Data pre-processing

Daily EVI values were calculated from the original MODIS 16-day interval EVI data using linear interpolation. The daily rainfall values at the location of the studied watersheds were interpolated from the daily rainfall data acquired at the surrounding Environment Canada weather stations and fire towers using inverse distance weighted (IDW) interpolation. The IDW interpolation assumes that the closer objects are more alike than those far apart, which mean that rainfall values at closer weather stations to the modelled location have greater impact on the predicted rainfall at that location. The IDW rainfall calculations were carried out using the following equations:

$$r_i = \sqrt{\left(x - x_i\right)^2 + \left(y - y_i\right)^2} \tag{4}$$

$$w_i = \frac{1/r_i^2}{\sum_{j=1}^{n} 1/r_j^2} \tag{5}$$

$$\hat{R}_{(x,y)} = \sum_{i=1}^{n} w_i \cdot R_i \tag{6}$$

where (x, y) is the coordinate of the location where the rainfall is to be estimated, i and j is the weather station number, n is the number of weather stations to be used in estimation, w_i is the weight of rainfall at the weather station i, R_i is the measured rainfall at station i and $\hat{R}_{(x,y)}$ is the estimated rainfall.

Data division

Data division is also an important step in ANN model development, because model performance can be significantly affected by the representativeness of subsets. A Self-organizing Mapping (SOM) network was implemented using NeuroShell 2 [47] to divide the available data into five clusters, from each of which data sets for training, testing and validation were sampled at a ratio of 3:1:1. The data set development using SOM has been demonstrated to effectively improve a model's performance and generality [48,49]. After data division, Kolmogorov–Smirnov tests were performed using MATLAB R2007a [50] to verify that the three data subsets represented the same population, since TDN export in the study streams was not normally distributed. The training data set was used to calibrate the model by updating the network connection weights. The testing data set was used to determine when to stop training attempting to avoid model overfitting. The validation data set, which the model had never seen, was then used to test model generality.

Model architecture

At this step, it is important to determine the number of hidden layers, the number of nodes in the hidden layers, the type of activation function(s), learning rate, momentum, and initial weights. The developed models have one hidden layer because a typical MLP ANN with a single hidden layer can approximate any functions given sufficient degrees of freedom [14]. The optimal number of nodes in the hidden layer was determined through a constructive algorithm [14,51]. This algorithm is based on the fact that the performance of the ANNs on training and testing data sets increases with the addition of more nodes until a critical point, where there is no improvement on the training data whereas the prediction accuracy on testing data decreases. The optimal number of hidden nodes is determined at the point where the trained networks' performance on testing data started to decrease. However, the single hidden layer with only one activation function did not produce acceptable N export results during preliminary model. This is because N export is highly correlated to streamflow, which is influenced by snowmelt in spring and by rainfall in summer and autumn. A previous study indicated that one hidden layer with three activation functions was able to capture the different driving forces for streamflow in the study area [7]. Thus, the same architecture was used in this application as well. The designed BP architecture has three slabs in the hidden layer. Different activation functions can be applied to hidden layer slabs to detect different features in a pattern processed through a network [47]. The designed network used a

Gaussian function on one hidden slab to detect features in the mid-range of the data and used a Gaussian complement in another hidden slab to detect features from the upper and lower extremes of the data. Using the logistic function in the third slab was helpful to map the irregularity in the data and simulate the patterns that are not captured by the other Gaussian functions.

The weight modification during calibration is the learning rate times the error plus the momentum times the previous weight change. Thus, the larger the learning rate is, the larger the weight changes and the faster the calibration will proceed. However, too large a learning rate may result in oscillation or non-convergence. To allow faster training without oscillation, a portion of previous weight change that is determined by momentum is added into the current weight change. The default values of learning rate, momentum as well as initial weights provided in the Neuroshell 2 package were used and were demonstrated to be effective.

Model training

During ANN model training, the connection weights are initially assigned arbitrary small values. As training progresses, the mean squared error (MSE) between the target output and the network output is calculated, and the weights are updated systematically. Weight adjustments are made based on an objective function that reduces MSE, attempting to reach a global minimum in the error surface. The training process stops when a prescribed stopping criterion is reached. The NeuroShell 2 software package was used to train the models [47]. The important principle for model training is to find the balance between convergence and generalization. Therefore, a test data set that represents the system being modelled, but that does not contain the same patterns as the training data set, is used to determine when to stop training (typically termed 'the early stopping technique' in the literature). The MSE for the training data set typically gets smaller as the network weights are updated based on model's prediction accuracy on training data set. As the training proceeds, the model reads the test data set and computes its prediction MSE. The MSE for the test data set gets smaller as model training progresses until an optimal point is reached, after which MSE starts to increase, reflecting a state when the model is starting to memorize the training data set. Thus, in all developed models, training was stopped when the model performed best on the test data set, that is, when the MSE for the test data set was smallest.

Model evaluation

There is no single statistical measure that can evaluate the performance of all models. Correlation-based measures have been widely used to evaluate model performance, but they are oversensitive to extreme values and insensitive to additive and proportional differences between observations and model predictions [52]. Therefore, in this study, correlation-based measures (Equations (7) and (8)) were supplemented with other error measures including mean absolute error (MAE) (Equation (9)), root MSE (RMSE) (Equation (10)) and graphing of observations and predictions to provide better evaluation of model prediction ability. The correlation-based measures that were used are the coefficient of determination (r^2) (Equation (7)) and Nash Sutcliffe coefficient (E) (Equation (8)). Higher r^2 and E values indicate better agreement between the observations and model predictions.

$$r^2 = \left\{ \frac{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)\left(P_i - \overline{P}\right)}{\left[\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2\right]^{0.5}\left[\sum\left(P_i - \overline{P}\right)^2\right]^{0.5}} \right\}^2 \quad <0.0, 1.0> \quad (7)$$

$$E = 1 - \frac{\sum_{i=1}^{N}\left(O_i - P_i\right)^2}{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2} \quad <-\infty, 1.0> \quad (8)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|O_i - P_i\right| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2} \quad (10)$$

where P_i and O_i are the predicted and the measured TDN export at time i, respectively; $\overline{O}$ is the mean of the measured TDN export for the entire time period and; N is the number of data points for the study period. The r^2 is one of the most commonly used measures and only evaluates linear relationships between observations and predictions, whereas E provides an improvement over r^2 because it is sensitive to proportional and additive differences between the observed and predicted means and variances. Generally, RMSE is equal to or greater than MAE and the degree to which RMSE exceeds MAE can indicate the extent to which outliers exist in the data [52]. MAE is preferred over RMSE in this study because of the existence of extreme values and

the intrinsic large variation in the magnitude of the range of the modelled parameter. The r^2 and RMSE were still used as they are commonly used to measure model performance in the literature. In addition to these four measures calculated for each training, testing and validation data set, the time series of observed and modelled profiles along the modelling period were plotted to examine when poor predictions occurred.

Model performance

The daily TDN export of five watersheds was modelled using the above noted ANN modelling algorithm. Table 2 summarizes the incorporated model inputs for each modelled watershed. The inputs generally include causal inputs (*R*, *Tmean*, *dd* and *EVI*), time-lagged inputs (R_{t-1}, R_{t-2}, etc., and dd_{t-1}, dd_{t-2}, etc.) and the inputs reflecting seasonal (Julian day) and annual (Year index) cycles of TDN export. The optimal ANN model architecture and internal parameters for all modelled watersheds are presented in Table 3. A single hidden layer with three activation functions (logistic, Gaussian and Gaussian complete) having the same number of nodes produced the best simulation for the studied watersheds. The training algorithm for Cassidy and Two Creek models was back-propagation and that for Willow and Thistle was back-propagation with a batch update.

The performance of the developed models was evaluated by statistical measures of goodness-of-fit (Table 4) and by examining the time series plot of the measured and modelled TDN export profiles. The *E* and r^2 values of the validation data set exceeded 0.72 for all watersheds, except Thistle. For all studied watersheds, the values of TDN export were in the ranges of zero g/km^2/d to thousands of g/km^2/d (Figure 2). The MAE values for all data subsets were small compared with the peaks of TDN export. The peaks of TDN export occurred in spring during snowmelt and in summer during heavy rainfall events (Figure 2). The time series plot of measured and predicted results showed that the models successfully predicted the seasonal and annual variation of TDN export for the studied watersheds. Most of the peaks were predicted with a reasonable accuracy. The peaks for the Willow, Cassidy, Two Creek and Thistle watersheds in the year 2002 were poorly predicted. This is likely because the

Table 2. Summary of all model inputs.

Model	Inputs
1A	R_t, R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, R_{t-5}, T_{mean}, dd_t, dd_{t-1}, dd_{t-2}, *EVI*, Julian Day, Year index
Cassidy	R_t, R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, T_{mean}, dd_t, dd_{t-1}, dd_{t-2}, *EVI*, Julian Day, Year index
Thistle	R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, T_{mean}, dd_{t-1}, dd_{t-2}, dd_{t-3}, *EVI*, Julian Day, Year index
Two Creek	R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4},T_{mean}, dd_t, dd_{t-1}, dd_{t-2}, *EVI*, Julian Day, Year index
Willow	R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, R_{t-5}, T_{mean}, dd_{t-2}, dd_{t-3}, *EVI*, Julian Day, Year index

Note: R_t, R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, R_{t-5} is rainfall in mm at lags of 0, 1, 2, 3, 4 and 5 days, respectively; T_{mean} is mean daily air temperature in degree C; dd_t, dd_{t-1}, dd_{t-2} and dd_{t-3} are cumulative degree days at lags of 0, 1, 2, and 3 days, respectively; EVI is the MODIS enhanced vegetation index; Year index is assigned value at either −1 or +1 (if the total rainfall of a year from April to October is lower than the 30-year average, that year is assigned −1; if the total rainfall of a year from April to October is higher than the 30-year average, that year is assigned +1).

Table 3. Summary table showing optimum ANN models' architecture and ANN internal parameters.

	Willow Model	Cassidy Model	Two Creek Model	Thistle Model	1A
Data division (TS:SS:VS)	3: 1: 1	3: 1: 1	3: 1: 1	3: 1: 1	3: 1: 1
Scaling function	<–1, 1>	<–1, 1>	<–1, 1>	<–1, 1>	<–1, 1>
Optimum network (I-[H-H-H]-O)	11L-[4LO-4GC-4G]-1LO	12L-[4LO-4GC-4G]-1LO	11L-[4LO-4GC-4G]-1LO	11L-[4LO-4GC-4G]-1LO	13L-[4LO-4GC-4G]-1LO
Training algorithm	BP-BU	BP	BP	BP-BU	BP
Learning rate	0.1	0.1	0.1	0.1	0.1
Momentum coefficient	0.1	0.1	0.1	0.1	0.1
Initial weight range	[– 0.3, + 0.3]	[– 0.3, + 0.3]	[– 0.3, + 0.3]	[– 0.3, + 0.3]	[– 0.3, + 0.3]
Epoch size	500	500	500	500	500
Stopping criterion	On best test set	On best test set	On best test set	On best test set	On best test set

Note: I and O are input and output layers, respectively; [H-H-H] represents a single hidden layer with different activation functions; L, is the linear scaling function; G, GC, and LO are the Gaussian, Gaussian complement, and logistic activation functions, respectively; TS, SS, and VS are the training, testing and validation data sets, respectively; and < > means a open interval; BP is a typical gradient descent back-propagation algorithm; BP-BU is a back-propagation algorithm with a batch update.

Table 4. Statistical measures of performance for the calibrated models.

	Willow Model			Cassidy Model			Two Creek Model			Thistle Model			1A		
Measures	TS	SS	VS	TS	SS	VS	TS	SS	VS	TS	SS	VS	TS	SS	VS
E	0.89	0.74	0.73	0.85	0.81	0.72	0.71	0.80	0.75	0.87	0.67	0.59	0.71	0.77	0.70
r^2	0.89	0.75	0.74	0.85	0.82	0.72	0.75	0.81	0.75	0.87	0.67	0.66	0.72	0.79	0.72
MAE	112	143	134	144	141	165	221	225	243	116	183	181	251	312	287
RMSE	231	326	317	279	244	330	341	359	378	238	447	419	522	852	562

Note: TS, SS and VS are training data set, testing data set, and validation data set, respectively; MAE is mean absolute error in g/d/km^2; RMSE is root mean squared error in g/d/km^2.

training data set did not contain enough similar data patterns as the ones to be predicted (2002 was the driest of the five study years) to let the models learn and identify the occurrence of these peaks. Thus, the models did not predict the peaks well due to the nature of these data-driven models. The overall performance of the models was fairly good, given that the models were constructed with only readily available public domain input data.

Modelling N export in unmonitored watersheds

The devised TDN export models produced reasonable prediction accuracy, highlighting the possibility of predicting TDN export in unmonitored watersheds where such information is available at no cost. A critical step for water quality modelling in unmonitored watersheds is to determine how to transfer calibrated models to unmonitored watersheds.

Watershed similarity measurement

Model transferability from one watershed to other watersheds is a very hard task due to the inherent variability in many watershed factors, including climate and watershed characteristics (e.g., topography, vegetation, land use and surficial geology). This study used the composite information on soil types, rainfall and vegetation conditions to investigate watershed similarity. The soil types of concern to this case study were peatland and riparian, which can be identified through RS imagery [53,54]. Peatlands store precipitation and surface water. The chemical and biological processes of nitrification, denitrification and anaerobic ammonium oxidation occur in peatland water and soils, such that the amount and form of N differs between water entering and water leaving the peatland. In general, forested peatlands contribute to high N export [55] and peatland cover was positively associated with ammonium (a fraction of TDN) export in the study watersheds [30]. Riparian soils also theoretically play an important role in regulating N export to surface waters. The role of riparian systems in Canada's boreal forest is complex due to the spatial variation in weather, soils, vegetation cover, slope, accumulation of organic matter, geographic location and relief [56]. In general, riparian areas have the potential to reduce excess N export into surface water [57,58]. Rainfall was also considered because of its known contribution to N export.

Watershed vegetation dynamics, coverage and disturbance can be easily monitored using satellite data. In addition to the abovementioned EVI, the RS normalized difference water index (NDWI) is a surrogate for vegetation health in terms of leaf water content and chlorophyll content [59]. The NDWI was calculated using the near- (λ_{NIR}) and mid-infrared (λ_{MIR}) frequency ranges downloaded from MODIS:

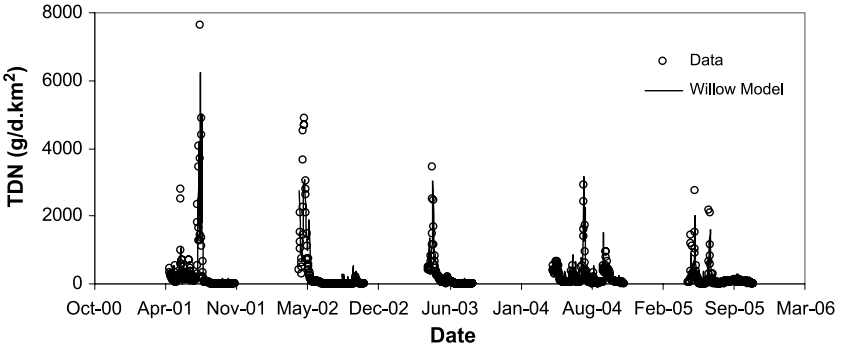
$$NDWI = \frac{\lambda_{NIR} - \lambda_{MIR}}{\lambda_{NIR} + \lambda_{MIR}} \quad (11)$$

To develop a representative index of watershed similarity, the usefulness of the following indices (individual or combined) in reflecting the success of model transferability was examined: peatland index, riparian index, rainfall index, EVI and NDWI. These indices were calculated and plotted against the prediction accuracy of transferred models in terms of *E*. The relationships between the *E* values and the calculated indices were analysed, attempting to develop an index that can correlate to the success of model transferability.

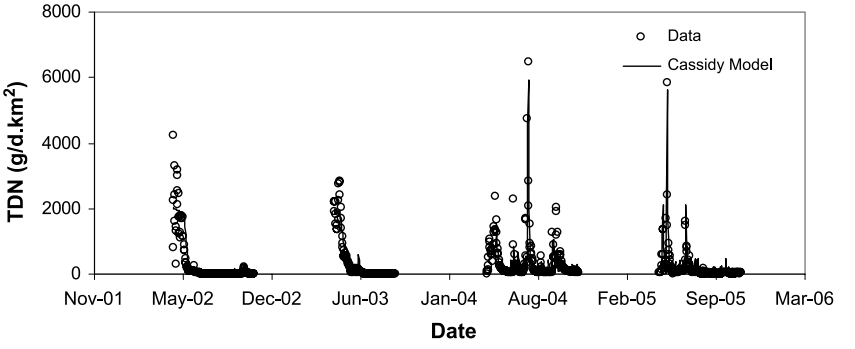
To normalize the variation of the indicators used (not to bias the higher magnitude parameters), the original data were first standardized by applying the following relation: (original data – minimum value) / (maximum value – minimum value). The standardized data were then used to compute the indices based on the Euclidean distance principle as follows:

$$Peatland\_Index = \sqrt{\left(Peatland_i - Peatland_j\right)^2} \quad (12)$$

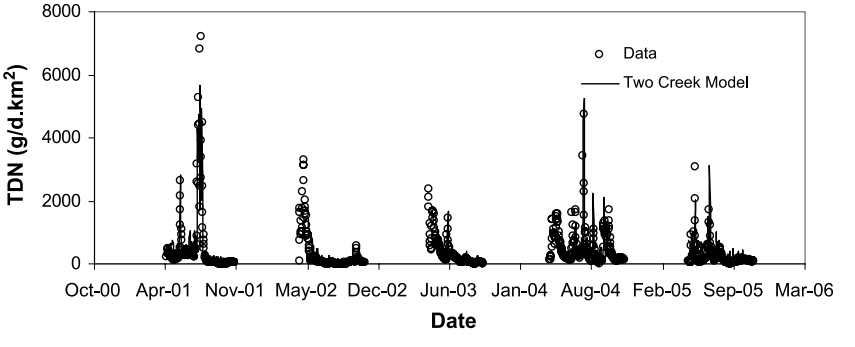
$$Riparian\_Index = \sqrt{\left(Riparian_i - Riparian_j\right)^2} \quad (13)$$



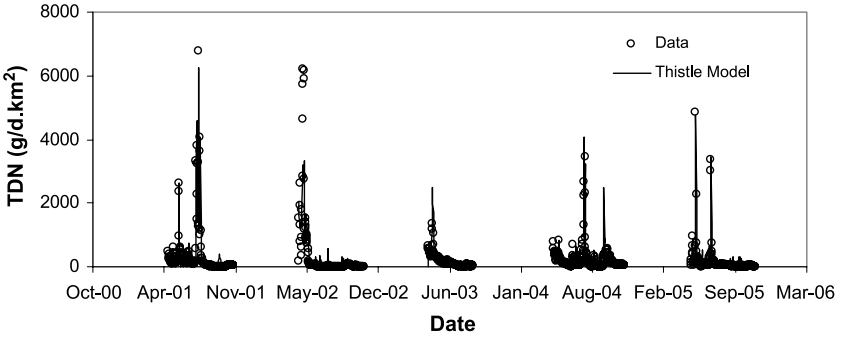
(A)



(B)



(C)



(D)

Figure 2. Time series plot of measured versus modelled daily TDN export for (A) Willow, (B) Cassidy, (C) Two Creek, and (D) Thistle watershed.

$$Rain\_Index = \sqrt{\sum_{k=1}^{n}\left(Rain_{i,k} - Rain_{j,k}\right)^2} \quad (14)$$

$$EVI\_Index = \sqrt{\sum_{k=1}^{n}\left(Evi_{i,k} - EVI_{j,k}\right)^2} \quad (15)$$

$$NDWI\_Index = \sqrt{\sum_{k=1}^{n}\left(NDWI_{i,k} - NDWI_{j,k}\right)^2} \quad (16)$$

where i and j represent two different watersheds and k represents a day within the studied n days. The calculated indices between any two studied watersheds are presented in Table 5.

Application of calibrated models to unmonitored watersheds

The calibrated models on Willow, Thistle, Cassidy and Two Creek watersheds were applied to other watersheds (unmonitored watersheds in this context) from which the models have never seen the data. The statistical measures of model performance when the calibrated models were transferred to different watersheds were calculated (Table 6). Among all the model transfer cases, applying the Willow model to Thistle watershed resulted in the best performance, with $E = 0.62$ and $r^2 = 0.63$. The seasonal and annual periodicity of TDN export in Thistle watershed was simulated well using the Willow TDN export model (Figure 3A). These two watersheds are adjacent to one another and have the most similar soil properties in terms of peatland and riparian cover in the watershed (Table 1). They are also very similar in terms of rainfall and vegetation dynamics (Table 5).

Applying the Willow model to watershed 1A generated the poorest performance, with $E = 0.42$ and $r^2 = 0.44$. The 1A watershed has the largest range of TDN export (0 to >10,000 g/km^2/d TDN) (Figure 3B), probably because of its high percentage of peatland coverage (Table 1), which leads to retention of N in normal weather conditions but to excess releases during spring snowmelt or large storm events. The differences of TDN export regimes between these two watersheds result in the poor performance on transferring Willow model to 1A watershed. However, the overall model transferring results are very promising, since five years of data were predicted fairly well, without being trained with any watershed-specific data points.

Relationships between watershed similarity indices and model transfer performance in terms of the E are summarized in Figure 4. Among the individual watershed indices, rainfall_index explained the most variation in E ($r^2 = 0.71$, $P \leq 0.05$), followed by NDWI ($r^2 = 0.69$,

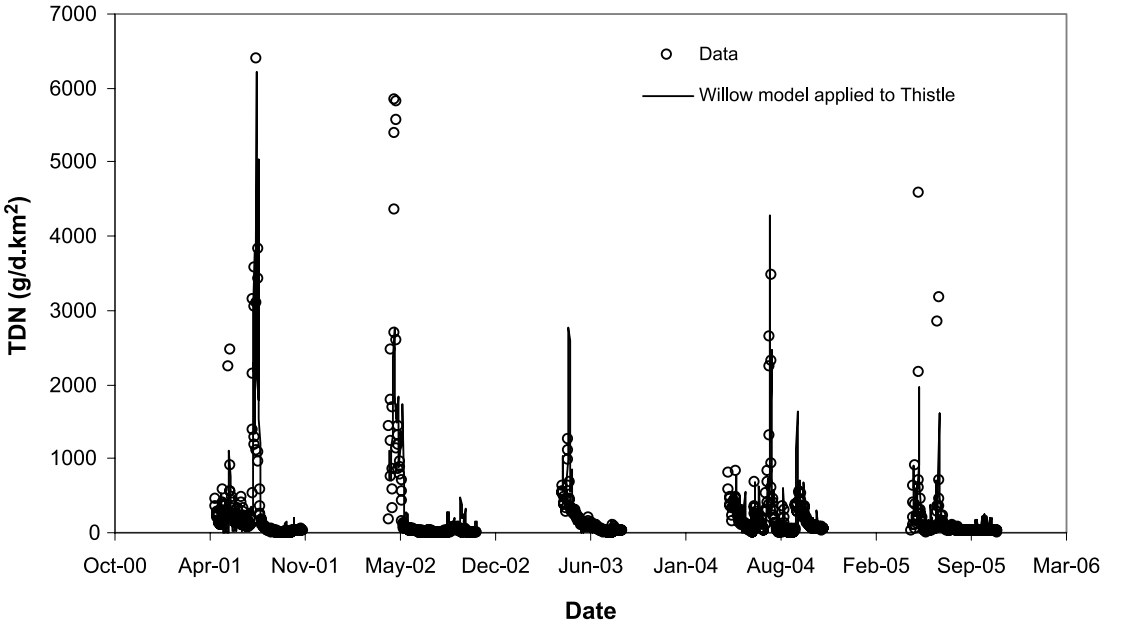
Table 5. Summary of watersheds similarity indices.

Watersheds	Peatland_Index	Riparian_Index	Rain_Index	EVI_Index	NDWI_Index
Willow – Cassidy	0.21	0.61	0.16	1.30	0.81
Willow – Thistle	0.02	0.26	0.01	0.58	0.64
Willow – 1A	0.61	0.74	0.39	2.84	2.79
Willow – Two Creek	0.28	0.22	0.55	1.69	2.73
Thistle – Cassidy	0.23	0.87	0.16	1.58	0.99
Thistle – 1A	0.59	1.00	0.39	2.58	2.75
Thistle – Two Creek	0.26	0.48	0.55	1.50	2.65
Cassidy – 1A	0.82	0.13	0.54	3.84	2.66
Cassidy – Two Creek	0.49	0.39	0.70	2.63	2.40
Two Creek – 1A	0.33	0.52	0.21	1.79	1.22

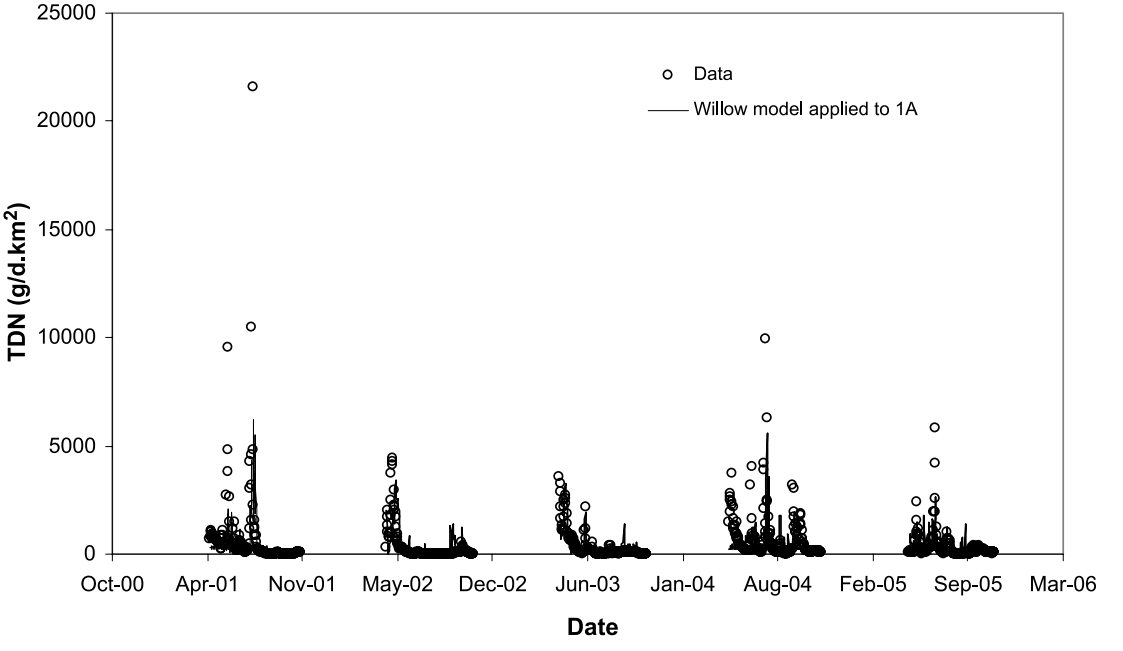
Note: the lower the similarity index, the more similar are the watersheds.

Table 6. Statistical measures of the model performance when the calibrated models were applied to other watersheds.

Measures	Willow applied to Cassidy	Willow applied to Thistle	Willow applied to 1A	Willow applied to Two Creek	Thistle applied to Cassidy	Thistle applied to 1A	Thistle applied to Two Creek	Cassidy applied to 1A	Cassidy applied to Two Creek	Two Creek applied to 1A
E	0.50	0.62	0.42	0.43	0.52	0.43	0.43	0.45	0.40	0.46
r^2	0.52	0.63	0.44	0.52	0.54	0.45	0.57	0.48	0.60	0.49
MAE	187.26	150.49	366.90	303.89	215.41	360.03	272.97	303.77	250.13	321.08
RMSE	446.03	403.82	901.60	526.45	465.28	896.45	528.69	630.59	411.59	872.08



(A)



(B)

Figure 3. Time series plot of measured versus modelled daily TDN export (A) when Willow model applied to Thistle, and (B) Willow model applied to 1A.

$P \leq 0.05$). Peatland_index and EVI were related to E, with r^2 values of 0.45 ($P \leq 0.05$) and 0.44 ($P = 0.06$), respectively. The riparian index was examined here because the conducted literature review demonstrated that the riparian systems generally have impact on N export from watersheds. However, the riparian index did not provide an important measure of watershed similarity in this study. One possible reason is that the riparian coverage in the selected watersheds is in the range of 0%–4.6%, which may be not large enough to address the differences among the watersheds in terms of riparian variations.

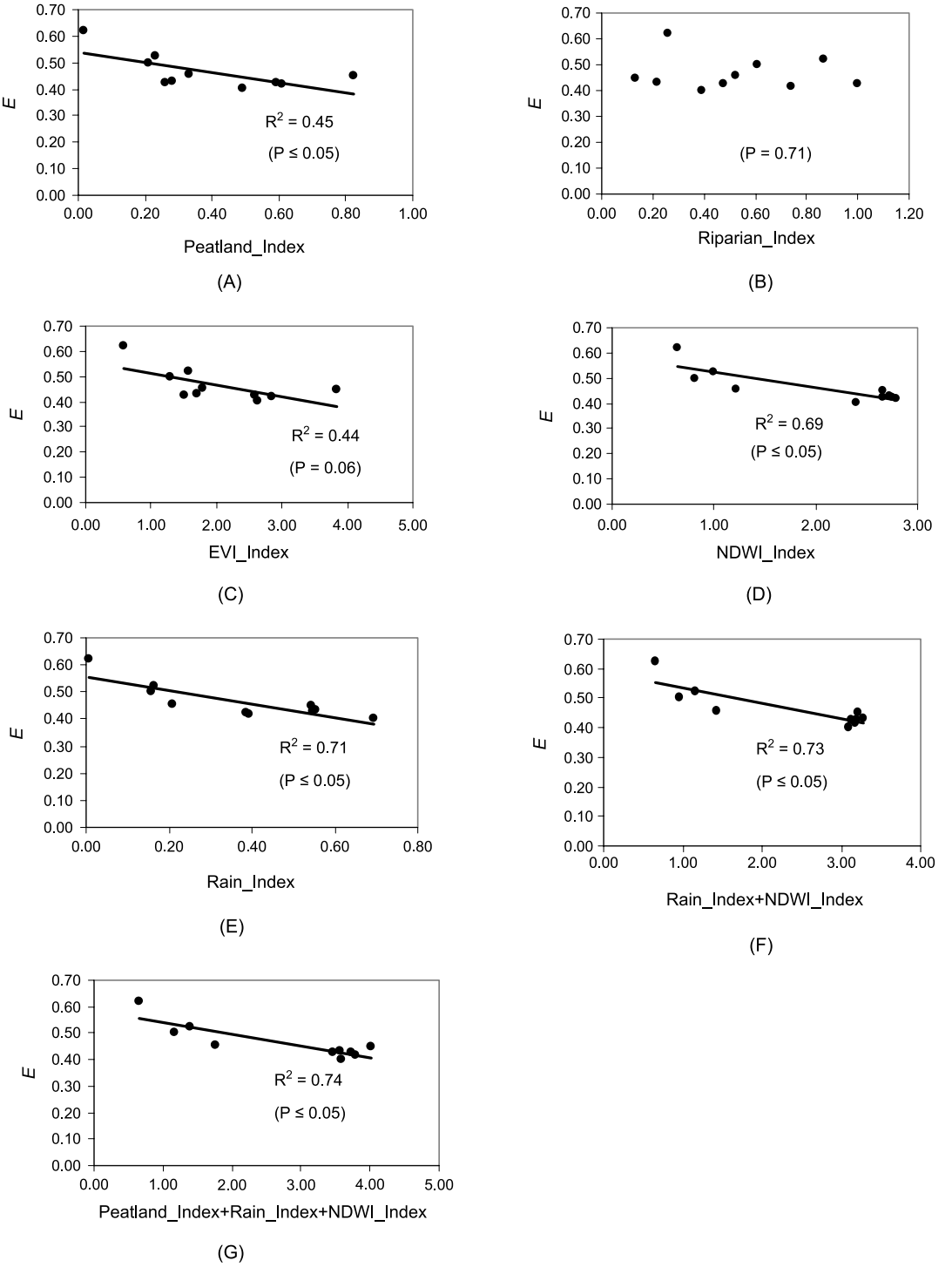


Figure 4. Plot of model transferability performance measure E versus watershed similarity indices.

The combined effects of two most important individual indices, rainfall and vegetation conditions, displayed a stronger relationship than either of them alone with E ($r^2 = 0.73$, $P \leq 0.05$) (Figure 4F). If the three most important factors were considered, their contribution to the variation in E was marginally improved to $r^2 = 0.74$ ($P \leq 0.05$) (Figure 4G).

The proposed measures of watershed similarity make use of public domain rainfall and RS information that can be easily calculated for unmonitored boreal forest watersheds. Initial results from the current study can be used to predict the expected success of model transferability as follows: (1) to predict N export in an unmonitored watershed j, select several monitored

watersheds; (2) calculate watershed similarity index between any monitored watershed *i* and the unmonitored watershed *j*: $Peatland\_Index_{i-j}$+ $Rain\_Index_{i-j}$ + $NDWI\_Index_{i-j}$; (3) select the most similar monitored watershed *k* to the unmonitored watershed, which has the lowest similarity index (the more similar are two watersheds, the lower the value of similarity index between them); and (4) apply the calibrated model *k* to the unmonitored watershed.

In regards to predicting N export from watersheds, there are mechanistic alternatives to ANNs such as Soil and Water Assessment Tool (SWAT) and Generalized Watershed Loading Functions (GWLF). The SWAT model is a physically based distributed watershed model and operates on a daily time step. It was developed to predict the impact of watershed management on water, sediment, nutrients (N and phosphorus), and agricultural yields in large basins for a long simulation period [60]. GWLF is considered to be a GIS-based distributed and lumped parameter watershed model and has the ability of simulating monthly sediment and nutrient loadings [61]. For execution, GWLF requires transport-, nutrient-, and weather-related data [61]. The transport data define the necessary parameters for each source area to be considered (e.g., area size, curve number, etc.) and global parameters (e.g., initial storage, sediment delivery ratio, etc.) that apply to all source areas. The nutrient data specify the various loading parameters for the different source areas identified (e.g., number of septic systems, urban source area accumulation rates, manure concentrations, etc.). The weather data contain daily average temperature and total precipitation values. Compared with ANN models, SWAT and GWLF models take considerations of spatial variations and the physical processes of nutrient exports in further details. However, ANN models have less data requirements and do not need specific transport-related data such as sediment delivery ratio, which make them practically applicable in physically inaccessible watersheds.

The results obtained suggest that the key to obtaining good model predictions on unseen data is the availability of representative data for model training (including wet, dry and normal conditions), and the key for successful model transferability is watershed similarity. Further investigations are needed to rigorously test and to expand on the proposed watershed similarity indices.

Conclusions

The current study proposed a MLP algorithm that uses low-cost, readily available meteorological and satellite data to model TDN export in boreal forest watersheds. The IDW interpolation technique was used to generate the rainfall data at studied watersheds by using surrounding Environment Canada weather station data. The temperature index approach was used to account for snowmelt. The MLP algorithm was applied to five watersheds to model N export. The performance of the models was evaluated using statistical measures of model performance, as well as examining the time series plots of measured versus modelled TDN profiles. Although the modelled parameter had a wide range of values (i.e., the peak values were over thousands of the low values), it was simulated fairly well. The best MLP architecture for all the developed models had a single hidden layer with three activation functions. The application of the devised algorithm to five watersheds ranging from around 5–130 km^2 in area demonstrated the success of the ANN modelling approach in predicting daily TDN using only public-domain and readily available data, indicating its potential application to unmonitored watersheds.

To demonstrate the applicability of the developed models to unmonitored watersheds, the calibrated models were used to predict TDN export in other watersheds (unmonitored watersheds in this perspective) without further calibration. The results of transferring the calibrated models to other unmonitored watersheds were promising, with *E* and r^2 values in the range of 0.40–0.62 and 0.44–0.63, respectively. The transferred models managed to predict the seasonal and annual periodicity of N export, even though some peak values were not well predicted.

In an effort to quantify watershed similarity to potentially guide the transferability of models from one watershed to the other, five similarity indices were developed and tested. The relationship (in terms of r^2) between the proposed indices and model transfer performance (*E*) was then calculated for all proposed indices. The best watershed similarity index was found to be the combined (*Rainfall_Index*+ *Peatland_Index*+ *NDWI_Index*), with $r^2 = 0.74$. Initial results from the current study can be used to predict the expected success of model transferability in unmonitored boreal forest watersheds. Although the proposed indices are not mature enough to have a meaningful threshold above which the models should not be transferred from one watershed to the other, this initiative is innovative in expanding and refining watershed similarity indices relies on free-of-cost RS information.

The advantage of using ANNs in watershed modelling is that they can model complex N transfer and storage processes without extensive data requirements and are applicable in unmonitored watersheds. On the other hand, the users of ANN models should be aware of their limitations, which result from the nature of data-driven models. The ANN models should be re-trained if there are significant changes of the watershed characteristics based on which they are originally calibrated.

Acknowledgements

The authors wish to thank Janice Burke for helpful comments on earlier drafts of this manuscript. This study is a part of the Forest Watershed and Riparian Disturbance (FORWARD) project. The FORWARD project is funded by the Natural Sciences and Engineering Research Council of Canada (CRD and Discovery Grant Programs) and Millar Western Forest Products Ltd., as well as the Canada Foundation for Innovation, Blue Ridge Lumber Inc. (a Division of West Fraser Timber Company Ltd.), Alberta Newsprint Company (ANC Timber), Vanderwell Contractors (1971) Ltd., the Living Legacy Research Program, Buchanan Forest Products Ltd., Buchanan Lumber (a division of Gordon Buchanan Enterprises), Alberta Forestry Research Institute, the Forest Resource Improvement Association of Alberta, and the Ontario Innovation Trust.

References

[1] ASCE Task Committee, *Artificial Neural Networks in Hydrology. I: Preliminary Concepts,* J. Hydrol. Eng. 5 (2000), pp. 115–123.

[2] R.J. Abrahart and L. See, *Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments,* Hydrol. Process. 14 (2000), pp. 2157–2172.

[3] P. Srivastava, J.N. McVair, and T.E. Johnson, *Comparison of process-based and artificial neural network approaches for streamflow modeling in an agricultural watershed,* J. Am. Water. Resour. Assoc. 42 (2006), pp. 545–563.

[4] S. Amari, A.R. Barron, E. Bienenstock, S. Geman, L. Breiman, J.L. McClelland, B.D. Ripley, R. Tibshirani, B. Cheng, and D.M. Titterington, *Neural networks – a review from statistical perspective – comments and rejoinders,* Statist. Sci. 9 (1994), pp. 31–54.

[5] L.C. Chang, F.J. Chang, and Y.M. Chiang, *A two-step-ahead recurrent neural network for stream-flow forecasting,* Hydrol. Process. 18 (2004), pp. 81–92.

[6] O. Kisi, *Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation,* Hydrol. Sci. J. 49 (2004), pp. 1025–1040.

[7] M.H. Nour, D.W. Smith, and M. Gamal El-Din, *Artificial neural networks and time series modelling of TP concentration in boreal streams: A comparative approach,* J. Environ. Eng. Sci. 5 (2006), pp. 39–52.

[8] G.J. Bowden, J.B. Nixon, G.C. Dandy, H.R. Maier, and M. Holmes, *Forecasting chlorine residuals in a water distribution system using a general regression neural network,* Math. Comput. Model. 44 (2006), pp. 469–484.

[9] G.M. Brion and S. Lingireddy, *Artificial neural network modelling: a summary of successful applications relative to microbial water quality,* Water Sci. Technol. 47 (2003), pp. 235–240.

[10] M.H. Nour, D.W. Smith, M. Gamal El-Din, and E.E. Prepas, *The application of artificial neural networks to flow and phosphorus dynamics in small streams on the Boreal Plain, with emphasis on the role of wetlands,* Ecol. Model. 191 (2006), pp. 19–32.

[11] X. Li, M.H. Nour, D.W. Smith, and E.E. Prepas, *Modelling nitrogen composition in streams on the Boreal Plain using genetic adaptive general regression neural networks,* J. Environ. Eng. Sci. 7 (2008), pp. S109–S125.

[12] J.D. Aber, C.L. Goodale, S.V. Ollinger, M.L. Smith, A.H. Magill, M.E. Martin, R.A. Hallett, and J.L Stoddard, *Is nitrogen deposition altering the nitrogen status of northeastern forests,* BioSci. 53 (2003), pp. 375–389.

[13] C.M. Zealand, D.H. Burn, and S.P. Simonovic, *Short streamflow forecasting using artificial neural networks,* J. Hydrol. 214 (1999), pp. 32–48.

[14] H.R. Maier, and G.C. Dandy, *Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications,* Environ. Model. Softw. 15 (2000), pp. 101–124.

[15] H.K. Cigizoglu, *Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons.* Adv. Water Resour. 27 (2004), pp. 185–195.

[16] H.K. Cigizoglu, and O. Kisi, *Methods to improve the neural network performance in suspended sediment estimation,* J. Hydrol. 317 (2006), pp. 221–238.

[17] G. Tayfur, and V. Guldal, *Artificial neural networks for estimating daily total suspended sediment in natural streams,* Nord. Hydrol. 37 (2006), pp. 69–79.

[18] G.M. Brion, T.R. Neelakantan, and S. Lingireddy, *Using neural networks to predict peak* Cryptosporidium *concentrations,* J. Amer. Water Works Assoc. 93 (2001), pp. 99–105.

[19] T.R. Neelakantan, G.M. Brion, and S. Lingireddy, *Neural network modelling of* Cryptosporidium *and Giardia concentrations in the Delaware River, USA,* Water Sci. Technol. 43 (2001), pp. 125–132.

[20] H.R. Maier, G.C. Dandy, and M.D. Burch, *Use of artificial neural networks for modelling* cyanobacteria Anabaena *spp. in the River Murray, South Australia,* Ecol. Model. 105 (1998), pp. 257–272.

[21] H.R. Maier, T. Sayed, and B.J. Lence, *Forecasting cyanobacterium Anabaena spp. in the River Murray, South Australia, using B-spline neurofuzzy models,* Ecol. Model. 146 (2001), pp. 85–96.

[22] L.O. Teles, V. Vasconcelos, E. Pereira, and M. Saker, *Time series forecasting of* cyanobacteria *blooms in the Crestuma Reservoir (Douro River, Portugal) using artificial neural networks,* Environ. Manage. 38 (2006), pp. 227–237.

[23] J.W. Kaluli, C.A. Madramootoo, and Y. Djebbar, *Modeling nitrate leaching using neural networks,* Water Sci. Technol. 38 (1998), pp. 127–134.

[24] C. Yu, W.J. Northcott, and G.F. McIsaac, *Development of an artificial neural network for hydrologic and water quality modeling of agricultural watersheds,* Trans. ASAE. 47 (2004), pp. 285–290.

[25] S. Lek, M. Guiresse, and J.-L. Giraudel, *Predicting stream nitrogen concentration from watershed features using neural networks,* Water Res. 33 (1999), pp. 3469–3478.

[26] T.A. Clair, and J.M. Ehrman, *Variations in discharge and dissolved organic carbon and nitrogen export from terrestrial basins with changes in climate: A neural network approach,* Limnol. Oceanogr. 41 (1996), pp. 921–927.

[27] T.Y. Gan, and S.J. Burges, *Assessment of soil-based and calibrated parameters of the Sacramento model and parameter transferability,* J. Hydrol. 320 (2006), pp. 117–131.

[28] D.W. Smith, E.E. Prepas, G. Putz, J.M. Burke, W.L. Meyer, and I. Whitson, *The Forest Watershed and Riparian Disturbance study: a multi-discipline initiative*

to evaluate and manage watershed disturbance on the Boreal Plain of Canada, J. Environ. Eng. Sci. 2 (2003), pp. S1–S13.

[29] E.E. Prepas, G. Putz, D.W. Smith, J.M. Burke, and J.D. MacDonald, *The FORWARD Project: Objectives, framework and initial integration into a Detailed Forest Management Plan in Alberta,* For. Chron. 84 (2008), pp. 330–337.

[30] E.E. Prepas, J.M. Burke, I.R. Whitson, G. Putz, and D.W. Smith, *Associations between watershed characteristics, runoff, and stream water quality: hypothesis development for watershed disturbance experiments and modelling in the Forest Watershed and Riparian Disturbance (FORWARD) project,* J. Environ. Eng. Sci. 5 (2006), pp. S27–S37.

[31] *Geomatica V9.1*, PCI Geomatics, Richmond Hill, Ontario, Canada, 2007. Available at http://www.pcigeomatics.com/.

[32] *ArcGIS 9.2*, ESRI, Redlands, California, USA, 2007; software available at http://www.esri.com/software/arcgis/

[33] G.J. Bowden, G.C. Dandy, and H.R. Maier, *Input determination for neural network models in water resources applications. Part 1 – background and methodology,* J. Hydrol. 301 (2005), pp. 75–92.

[34] I.F. Creed and L.E. Band, *Export of nitrogen from catchments within a temperate forest: Evidence for a unifying mechanism regulated by variable source area dynamics,* Water Resour. Res. 34 (1998), pp. 3105–3120.

[35] M. de Wit and H. Behrendt, *Nitrogen and phosphorus emissions from soil to surface water in the Rhine and Elbe basins,* Water Sci. Technol. 39 (1999), pp. 109–116.

[36] R. Hatano, T. Nagumo, H. Hata, and K. Kuramochi, *Impact of nitrogen cycling on stream water quality in a basin associated with forest, grassland, and animal husbandry, Hokkaido, Japan,* Ecol. Eng. 24 (2005), pp. 509–515.

[37] J. Ide, O. Nagafuchi, M. Chiwa, A. Kume, K. Otsuki, and S. Ogawa, *Effects of discharge level on the load of dissolved and particulate components of stream nitrogen and phosphorus from a small afforested watershed of Japanese cypress* (Chamaecyparis obtusa), J. For. Res. 12 (2007), pp. 45–56.

[38] S.P. Inamdar, N. O'Leary, M.J. Mitchell, and J.T. Riley, *The impact of storm events on solute exports from a glaciated forested watershed in western New York, USA,* Hydrol. Process. 20 (2006), pp. 3423–3439.

[39] D. Pelster, J.M. Burke, and E.E. Prepas, *Runoff and inorganic nitrogen export from Boreal Plain watersheds six years after wildfire and one year after harvest,* J. Environ. Eng. Sci. 7 (2008), pp. S51–S61.

[40] A. Lauren, L. Finer, H. Koivusalo, T. Kokkonen, T. Karvonen, S. Kellomadki, H. Mannerkoski, and M. Ahtiainen, *Water and nitrogen processes along a typical water flowpath and streamwater exports from a forested catchment and changes after clear-cutting: a modelling study,* Hydrol. Earth Syst. Sci. 9 (2005), pp. 657–674.

[41] L.G. Ferreira and A.R. Huete, *Assessing the seasonal dynamics of the Brazilian Cerrado vegetation through the use of spectral vegetation indices,* Int. J. Remote Sens. 25 (2004), pp. 1837–1860.

[42] A. Khan, *Satellite Characterization of Vegetation Dynamics in the Western Canadian Boreal Plain,* M. Sci. Dissertation, University of Alberta, 2005.

[43] E.M.D. Silveira, L.M.T. de Carvalho, F.W. Acerbi-Junior, and J.M. de Mello, *The assessment of vegetation seasonal dynamics using multitemporal NDVI and EVI images derived from MODIS,* Cerne 14 (2008), pp. 177–184.

[44] S.V. Gregory, F.J. Swanson, W.A. McKee, and K.W. Cummins, *An ecosystem perspective of riparian zones,* BioSci. 41 (1991), pp. 540–551.

[45] V. Sharma, S.C. Negi, R.P. Rudra, and S. Yang, *Neural networks for predicting nitrate-nitrogen in drainage water,* Agric. Water Manage. 63 (2003), pp. 169–183.

[46] Q. Zhang and S.J. Stanley, *Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modelling,* Water Res. 31 (1997), pp. 2340–2350.

[47] *Neuroshell 2*, Ward Systems Group Inc, Frederick, MD, USA, 2005. Available at http://www.wardsystems. com/

[48] G.J. Bowden, H.R. Maier, and G.C. Dandy, *Optimal division of data for neural network models in water resources applications,* Water Resour. Res. 38 (2002), pp. 1–11.

[49] M.A. Shabin, H.R. Maier, and M.B. Jaksa, *Data division for developing neural networks applied to geotechnical engineering,* J. Comput. Civil. Eng. 18 (2004), pp. 105–114.

[50] *MATLAB R2007a*, The MathWorks Inc., Natick, MA, USA; software available at http://www.mathworks. com/

[51] M.H. Nour, D.W. Smith, M. Gamal El-Din, and E.E. Prepas, *Towards a generic neural network model for the prediction of daily streamflow in ungauged boreal plain watersheds,* J. Environ. Eng. Sci. 7, (2008), pp. S79–S93.

[52] D.R. Legates and G.J. McCabe, *Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation,* Water Resour. Res. 35 (1999), pp. 233–241.

[53] J.H. Everitt, C. Yang, D.E. Escobar, R.I. Lonard, and M.R. Davis, *Reflectance characteristics and remote sensing of a riparian zone in south Texa,* Southw. Natural. 47 (2002), pp. 433–439.

[54] P. Shanmugam, Y.H. Ahn, and S. Sanjeevi, *A comparison of the classification of wetland characteristics by linear spectral mixture modelling and traditional hard classifiers on multispectral remotely sensed imagery in southern India,* Ecol. Model. 194 (2006), 379–394.

[55] J.Z. Zhu and A. Mazumder, *Estimating nitrogen exports in response to forest vegetation, age and soil types in two coastal-forested watersheds in British Columbia,* For. Ecol. Manage. 255 (2008), pp. 1945–1959.

[56] S.H. Luke, N.J. Luckai, J.M. Burke, and E.E. Prepas, *Riparian areas in the Canadian boreal forest and linkages with water quality in streams,* Environ. Rev. 15 (2007), pp. 79–97.

[57] M.E. Baker, M.J. Wiley, and P.W. Seelbach, *GIS-based hydrologic modeling of riparian areas: Implications for stream water quality,* J. Amer. Water Resour. Assoc. 37 (2001), pp. 1615–1628.

[58] E.E. Cey, D.L. Rudolph, R. Aravena, and G. Parkin, *Role of the riparian zone in controlling the distribution and fate of agricultural nitrogen near a small stream in southern Ontario,* J. Contam. Hydrol. 37 (1999), pp. 45–67.

[59] D.Y. Chen, J.F. Huang, and T.J. Jackson, *Vegetation water content estimation for corn and soybeans using*

spectral indices derived from MODIS near- and short-wave infrared bands, Remote Sens. Environ. 98 (2005), pp. 225–236.
[60] J.G. Arnold, R. Srinivasan, R.S. Muttiah, and J.R. Williams, *Large area hydrologic modeling and assessment part I: Model development,* J. Am. Water Resour. Assoc. 34 (1998), pp. 73–89.
[61] *AVGWLF*, Penn State Institute Energy & The Environment, PA, USA, 2009. Available at http://www.avgwlf.psu.edu/.