University of Alberta

Neural networks modelling of stream nitrogen using remote sensing information: model development and application

by

Xiangfei Li

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Environmental Engineering

Department of Civil and Environmental Engineering

Fall 2009
Edmonton, Alberta

EXAMINING COMMITTEE

Dr. Daniel W. Smith, Civil and Environmental Engineering

Dr. Thian Gan, Civil and Environmental Engineering

Dr. Mohamed Gamal El-Din, Civil and Environmental Engineering

Dr. Ian D. Buchanan, Civil and Environmental Engineering

Dr. Phillip M. Fedorak, Biological Sciences

Dr. Patrick Hettiaratchi, Civil Engineering, University of Calgary

DEDICATION

to

My Husband

My Children

&

My Parents

ABSTRACT

In remotely located forest watersheds, monitoring nitrogen (N) in streams often is not feasible because of the high costs and site inaccessibility. Therefore, modelling tools that can predict N in unmonitored watersheds are urgently needed to support management decisions for these watersheds. Recently, remote sensing (RS) has become a cost-efficient way to evaluate watershed characteristics and obtain model input variables. This study was to develop an artificial neural network (ANN) modelling tool relying solely on public domain climate data and satellite data without ground-based measurements.

ANN was successfully applied to simulate N compositions in streams at studied watersheds by using easily accessible input variables, relevant time-lagged inputs and inputs reflecting seasonal cycles. This study was the first effort to take the consideration of vegetation dynamics into N modelling by using RS-derived enhanced vegetation index (EVI) that was capable of describing the differences of vegetation canopy and vegetation dynamics among watersheds. As a further study to demonstrate the applicability of the ANN models to unmonitored watersheds, the calibrated ANN models were used to predict N in other different watersheds (unmonitored watersheds in this perspective) without further calibration. A watershed similarity index was found to show high correlation with the transferability of the models and can potentially guide transferring the trained models into similar unmonitored watersheds. Finally, a framework to incorporate water quantity/quality modelling into forestry management was proposed to demonstrate the application of the developed

models to support decision making. The major components of the framework include watershed delineation and classification, database and model development, and scenario-based analysis. The results of scenario analysis can be used to translate vegetation cut into values of EVI that can be fed to the models to predict changes in water quality (e.g. N) in response to harvesting scenarios.

The results from this research demonstrated the applicability of ANNs for stream N modelling using easily accessible data, the effectiveness of RS-derived EVI in N model construction, and the transferability of the ANN models. The presented models have high potential to be used to predict N in streams in the real-world and serve forestry management.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor Dr. Daniel W. Smith for his guidance, encouragement, support, patience and understanding during the study. Dr. Daniel W. Smith devoted his time, interest, ideas and efforts to make this study possible. He also provided a critical and constructive manuscript review. My study at the University of Alberta was a great experience for me to learn how to conduct high quality research and contribute new knowledge to the society because of Dr. Daniel W. Smith. He gave me the opportunity and freedom to develop critical literature reviews and research proposals, and helped me to analyze the problems and develop better solutions. The experience was challenging, impressive and beneficial for my whole life. I am also greatly thankful to my co-supervisor Dr. Thian Y. Gan for his time, encouragement, and support in the area of hydrology and remote sensing. Special thanks to my committee members Dr. Mohamed Gamal El-Din, Dr. Ian D. Buchanan, Dr. Phillip M. Fedorak, and Dr. Patrick Hettiaratchi.

I wish to acknowledge the assistance from Dr. Mohamed Nour for sharing his experience and ideas in modelling and Janice Burke for her helpful comments to several paper manuscripts. I also like to thank Grant Burkel for distributing the data, Nicole Fraser for clarifying the maps, and Mark Serediak for arranging the field trip.

Many thanks to the Forest Watershed and Riparian Disturbance (FORWARD) project for its financial support. The FORWARD project is funded by the Natural Sciences and Engineering Research Council of Canada (CRD and Discovery Grant Programs), Millar Western Forest Products Ltd., 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.

Finally, I am deeply thankful to my husband, who has sacrificed a lot, worked hard to support the family, looked after and educated our children, and continually encouraged me to complete the study. I also like to acknowledge my two lovely children, Alicia and Catherine, for reminding me to be persistent. I truly thank my parents who cared my children, allowing me to complete my thesis. Their help was essential to make possible my achievement in both study and family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A/C-NNs	Aggregated/Compound Neural Networks
AMN	Associative Memory Network
ANN	Artificial Neural Network
ANSWERS	Aerial Non-point Source Watershed Environment Simulation
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARMAX	Autoregressive Moving Average with Exogenous input
AVHRR	Advanced Very High Resolution Radiometer
CCF	Cross Correlation Function
DEM	Digital Elevation Model
DFMP	Detailed Forest Management Plan
EOS	Earth Observing System
EVI	Enhanced Vegetation Index
FMA	Forest Management Agreement
FORWARD	Forest Watershed and Riparian Disturbance
GA-GRNN	Genetic Adaptive General Regression Neural Network
GIS	Geographic Information System
GRNN	General Regression Neural Network
HRU	Hydrologic Response Unit
HSPF	Hydrologic Simulation Program Fortran
KSOM	Kohonen's Self-Organizing Map
LAI	Leaf Area Index
LULC	Land Use Land Cover
MAE	Mean Absolute Error
MLP-BP	Feed-forward Multilayer Perceptron trained with the Back-Propagation
MLR	Multi-Linear Regression
MODIS	Moderate-resolution Imaging Spectroradiometer
MNNs	Modular Neural Networks

MSE	Mean Squared Error
MWFP	Millar Western Forest Products Ltd.
N	Nitrogen
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
PMI	Partial Mutual Information
RBFN	Radial Basis Function Network
RDNN	Range Dependent Neural Network
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
RS	Remote Sensing
SOM	Self Organizing Map
SWAT	Soil and Water Assessment Tool
SWAT-BF	called SWAT-Boreal Forest
TDN	Total Dissolved Nitrogen
TP	Total Phosphorus
TSS	Total Suspended Solids
VIs	Vegetation Indices

CHAPTER 1. BACKGROUND AND INTRODUCTION

1.1 Background

Natural (e.g., wildfires) and anthropogenic (e.g., forest harvesting) watershed disturbance can change watershed features such as soil properties and vegetation cover and affect nutrient cycling and budget in the forestry system (Chanasyk et al. 2003). Wildfires and harvesting have been associated with increased water yield and release of nitrogen (N), phosphorus and sediments from watersheds into streams several years after the disturbances (Prepas et al. 2003; Pelster et al. 2008). In Swan Hills, Alberta, the increased nitrate-N and ammonium-N concentrations, combined with increased runoff one year after harvest resulted in impact ratios for areal nitrate-N and ammonium-N exports that were 170% and 130%, respectively, higher in harvested than reference watersheds (Pelster et al. 2008). The increase of nutrient in surface water can potentially cause severe environmental problems such as dissolved oxygen depletion, algal blooms, cyanobacterial toxin production, chlorophyll a and biodiversity disruption. Therefore, as an important land-based resource activity in Alberta, forestry must be scientifically planned to ensure sustainable fibre production and minimal adverse environmental impact.

Currently, the harvesting activities in Alberta are governed by Forest Management Agreements (FMA) with the Province. The obligations of an FMA permit holder are to harvest no more than the amount of timber stated in the FMA, and to promptly regenerate and maintain the harvested areas in a forested condition. In addition, the FMA holder should also plan its harvest strategy to prevent detrimental effects to other interests in its FMA area. A set of harvesting control policies are currently used as best practices, attempting to minimize the adverse impacts of forest harvesting on biodiversity, ecological integrity, water quantity and quality and timber supply. In an era of increased land use and resource development, forest management planning processes will require increasingly more sophisticated modelling tools and science-based evidence to identify and avoid significant impacts on the environment (Smith et al. 2003).

Therefore, the Forest Watershed and Riparian Disturbance (FORWARD) project was initiated in 2001. It is a long-term project to integrate aquatic and soil science, hydrology, and forestry into models that link water quantity and quality and disturbance indicators with watershed management on the Boreal Plain of western Canada. The modelling component of the FORWARD project is to develop modelling procedures that can be used for predicting the impacts of forest harvesting on water quantity and quality in streams.

1.2 Watershed Modelling

1.2.1 Watershed Models Classification

Watershed models can be classified based on the degree of spatial resolution into: (1) lumped models that use average values of input variables over the entire watershed area, and thus have minimal data requirements; (2) semi-distributed models that divide the watershed into sub-watersheds, in which each sub-watershed carries a distinct set of input variables; and (3) distributed models that are made up of very small areas that may be represented by imagery based pixels in terms of input representations and parameter routing, and therefore having huge data requirements.

Watershed models can also be classified according to physical conceptualization into: (1) empirical (also called data-driven); (2) physically based (also called mechanistic); and (3) conceptual (also called parametric) models. Empirical models use time-series of input and output records to derive both the model structure and the corresponding parameter values. They do not need a complete understanding of the physical, chemical, hydro-morphological and biological processes controlling flow processes and contaminant transport mechanisms. Physically based models mathematically describe a process using a set of principles, based on the conservation of mass, momentum and energy. They are distributed models and have intensive data requirements. Conceptual models include both simplified physically based components and empirical components. That means, based on a conceptualization of the watershed, the structure of these models is specified in advance, then the observations of the watershed response are used to find appropriate values for the model parameters through empirical relations. Conceptual models form the large majority of models used in practice. Conceptual watershed-scale water quantity and quality models include, but are not limited to, the Soil and Water Assessment Tool (SWAT) (Arnold et al. 1998), aerial non-point source watershed environment simulation-2000 (ANSWERS-2000) (Beasley et al. 1980; Bouraoui and Dillaha 1996), Hydrologic Simulation Program Fortran (HSPF) (Johanson et al. 1984), erosion productivity impact calculator (Sharpley and Williams 1990), annualized agriculture non-point source pollutant loading model (AnnAGNPS) (Bingneer et al. 2001), and the Guelph model for evaluating the effects of agricultural management systems on erosion and sedimentation (GAMES) (Cook et al. 1985). The use of this class of models presents the challenge of estimating or calibrating a large number of model parameters from information with limited availability. Obtaining the information necessary for model calibration is time-consuming and expensive. To use the models for forested watersheds, certain modifications have to be completed first as most of the models were originally developed for agricultural dominated watersheds.

Data-driven models have been successful in capturing the relationship between the input and output data with less knowledge of the modelled system in terms of the interaction of the biological, geological, chemical and physical processes. Consequently, they are attractive alternatives to traditional conceptual models. Among those techniques, artificial neural network (ANN) models hold promise for water quantity and quality modelling. They can often capture data patterns without extensive knowledge of the particular site-related problems and can model complicated and nonlinear processes with fewer input variables than mechanistic models. Because they are capable of handling large-scale and complex problems, ANN models provide great advantages in a wide range of water quality applications, such as modelling sediment concentrations (Cigizoglu and Alp 2006; Cigizoglu and Kisi 2006; Tayfur and Guldal 2006; Alp and Cigizoglu 2007), phosphorus concentrations (Nour et al. 2006a; Nour et al. 2006b), cyanobacteria blooms (Maier et al. 2001b; Teles et al. 2006) in surface waters, and N loads and concentrations from agricultural settings into drainage water (Salehi et al. 2000; Sharma et al. 2003).

1.2.2 Artificial Neural Networks

ANNs consist of a large number of simple, highly interconnected processing elements (neurons) in an architecture inspired by the structure of the cerebral cortex of the brain (Tsoukalas and Uhrig 1997). ANN models are developed in an attempt to mimic the learning of human brains. The processing neurons are generally organized in layers. A feedforward networks has an input layer, one or more hidden layers and an output layer. The nodes in one layer are connected to each node in the next layer, but not to those in the same layer. The input from each node in the previous layer is multiplied by a connection weight. At each node, the weighted input signals are summed and then added a threshold value. Then a transfer function is performed on the combined input to produce the output of the node. The types of ANNs that have been used in water quantity and quality modelling include feed-forward multilayer perceptron trained with the back-propagation algorithm (MLP-BP), recurrent neural networks (RNNs), radial basis function networks (RBFNs), general regression neural networks (GRNNs), modular neural networks (MNNs), Kohonen's self-organizing map (KSOM), and associative memory network (AMN). So far, MLP-BP is the most widely used ANNs.

ANN models have become an attractive modelling tool in solving complex problems because (ASCE Task Committee 2000) ANNs: (1) are able to identify the relationship between inputs and outputs without fully understanding the mechanistic principles behind them, (2) can work well even when the training sets contain noise and measurement errors, (3) can be adapted to solutions to compensate for changing circumstances, and (4) are easy to use once trained. In addition, ANNs have demonstrated strengths over the other two alternative models − conventional statistical models and numerical models − from several perspectives. The traditional statistical modelling approaches (e.g., multivariate linear regression (MLR)) involve making assumptions about the system under study and developing equations describing the problem to achieve statistical optimality. In contrast, the ANN practitioners do not need to make such assumptions and figure out equations for the problems. Their primary objective is to achieve high prediction accuracy and develop models that work in practice (Breiman 1994; Tibshirani 1994). The modelling approach differences between ANN models and traditional statistical models, coupled with a lack of strict rules charging the ANN model development lead to the tendency of choosing ANNs over MLR when predicting surface water quality parameters. ANNs can achieve higher prediction accuracy than MLR as shown through a number of case studies (e.g., Kisi 2004; Bowden et al. 2006). Compared to numerical models, ANN models can provide comparable modelling accuracy but are more applicable in practice when professional expertise and data are limited.

The development of ANN models generally includes the following steps: (1) input determination; (2) data division into training, testing and validation datasets; (3) determination of model architecture (e.g., number of hidden layers, number of neurons in each layer, activation function, and learning rate); (4) model calibration; and (5) model evaluation. The determination of model inputs is one of the most important steps in developing a successful ANN model. Input variables should be carefully selected to closely describe the physical system being modelled and not to include noisy or correlated variables because inclusion of these kinds of inputs can increase computational complexity and deteriorate model performance. Ideally, the available input/output data pairs should be divided into three data sets (training, testing and validating data sets) that can represent the same population to assure model generality. The training data set is used for model training and for the optimization of the model connection weights. The testing set is used to decide when to stop training to avoid model overfitting. The cross-validation data set is used to evaluate the model against a totally independent data set. Model architecture determination generally relies upon modellers' experience; however, guidelines have been proposed in recent studies (e.g., Maier and Dandy 2000; Nour et al. 2006b). For an ANN to generate output predictions that are as close as possible to the objective values, model calibration is an essential step to find optimal weights by minimizing a predetermined error function. After model calibration, the power of the models in terms of prediction accuracy, robustness and generality should be verified before the models are put in use. The models usually are tested through several criteria that include both correlation-based measures (e.g., coefficient of determination, coefficient of efficiency) and additional statistical measures (e.g., mean absolute measure, summary statistics) (Legates and McCabe 1999).

ANN is an efficient tool for modelling complex systems at desirable accuracy if well developed. The popularity of ANNs, on the other hand, reveals its limitations as well as further expectations from model practitioners. In turn, advanced studies have been conducted in the following perspectives to make ANN a more reliable modelling tool: (1) combine ANNs and other alternative models with emphasis on numerical models to couple the strengths of ANNs in simplifying the description of the problem and reducing data requirement and computational time, and the strengths of numerical models in describing the physical principles; (2) find the ways of extracting information from ANNs because they are looked at as 'grey-box', which means that ANNs contain physical information about the system being modelled; (3) improve the generality of ANNs models because the models developed based on experimental data are expected to be applicable in similar situations; and (4) establish systematic procedures and guidelines for ANNs model development, which is crucial to develop successful ANNs models.

1.3 Watershed N Modelling Approach

1.3.1 Model Selection

It is recommended to consider 11 criteria to select the appropriate watershed N models (Putz et al. 2003). Generally speaking, first, the model should be able to achieve the overall modelling objectives, that is, to predict water quality at satisfactory accuracy, link watershed disturbance indicators with water quality and can be potentially incorporated into forestry management. Second, the models should be applicable in practice. Specifically, the following issues including the spatial and temporal resolutions of inputs and outputs, the data requirements and availability for the models, the requirements of computational hardware and personal skills should be considered.

Modelling N composition in streams is very complex because of the difficulty in mathematically representing factors like land use and land cover, soil and vegetation N dynamics, in-stream nitrification/denitrification and meteorological parameters. These factors are complex, nonlinearly related, spatially distributed on a watershed scale and temporally variable. Also, hydrological, biological and chemical processes underlying N storage, transformation and release are not well understood. Hence, the application of mechanistic models is costly and impractical, because large amounts of data are required to establish parameters and verify model performance.

In contrast, ANN models can often capture data patterns without extensive knowledge of the particular site-related problems and can handle large-scale and complex problems. Therefore, ANN models are suitable to predict N in watersheds. In addition, the successful application of ANNs in water quantity and quality modelling indicates that ANN models are useful in supporting environmental decision-making (Maier et al. 2001a; Rudra et al. 2005; Dakou et al. 2006; Diamantopoulou et al. 2007; Elhatip and Komur 2008).

Application of ANNs for N modelling has mostly relied on MLP-BP and focused on nitrate (NO_3^-) in agricultural settings, such as predicting annual NO_3^- loss into drainage (Salehi et al. 2000), simulating NO_3^- leaching in drainage effluent (Kaluli et al. 1998), forecasting NO_3^- loads on a watershed based on historical data (Yu et al. 2004) and predicting NO_3^- concentration in drainage water after application of fertilizers and manure (Sharma et al. 2003). Lek et al. (1999) applied MLP-BP to predict total and inorganic N concentrations in streams with correlation coefficients of 0.82 and 0.80, respectively, from watershed features. In another study, using weather station data, daily streamflow and the Julian day as model inputs, Suen and Eheart (2003) developed a MLP-BP model with overall accuracy of 0.8 to predict if NO_3^- concentrations in a river were greater or less than 10 mg L^{-1}.

1.3.2 Data Requirements

The availability of data is always a concern to develop practically-applicable watershed models. Recently, the development of Earth observing satellite system and the advancements in computer and software technologies have made it possible to evaluate watershed characteristics and obtain model input variables through remote sensing (RS). RS is a cost-efficient way to improve the spatial and temporal coverage of surface water and watershed monitoring (Koponen et al. 2004). RS technology develops quickly and includes most of the electromagnetic spectrum to provide a variety of information about the Earth. To acquire RS data efficient and appropriate for its specific application, spatial resolution, temporal resolution, spectral resolution and spectral wavelength need to be considered.

The Earth observing system (EOS), launched December of 1999, is the first observing system to offer integrated measurements of the Earth's processes. The moderate-resolution imaging spectroradiometer (MODIS) is the key EOS instrument and its mission is to study global/local interactions. MODIS was designed to improve monitoring of land, ocean and atmosphere, particularly, to provide innovative land products data that are especially designed to support modelling applications (Reed et al. 2002). The MODIS collects reflected and emitted radiation from the earth in 36 bands from 0.405 to 14.385 μm and provides spatial resolution of 250 m, 500 m and 1 km (Barnes et al. 1998). The MODIS Land Group provide not only satellite data but also high level data products (e.g., vegetation indexes (VIs)) that are specifically designed to support the global to regional monitoring, modelling, and assessment. Furthermore, MODIS data are freely available, thus providing a means of acquiring time series representations of vegetation dynamics at an affordable cost. For instance, a successful nutrient model requires information regarding soil and vegetation nutrient status. RS VIs can represent vegetation health and stress in terms of the vegetation chlorophyll content and the leaf water content, which can be linked to soil/vegetation nutrient interactions and thus can aid in formulating relatively accurate and usable nutrient watershed models (Cheng et al. 2006). Such information can potentially act as a surrogate for soil/vegetation nutrient transport and therefore can potentially represent vegetation dynamics in nutrient models.

The direct use of RS digital data to estimate hydrological state parameters is an important application to watershed modelling and is normally achieved through electro-optical or statistical modelling of known hydrometric data with satellite data. Although there has been some success in the application of RS data in hydrology, the incorporation of RS information into nutrient modelling still requires more effort. To develop low-cost N models that are applicable to a number of watersheds (e.g., unmonitored watersheds, watersheds at remote areas), RS data is an essential component of the data requirements in this study.

1.4 Research Objectives

In order to develop ANN N models that only rely on weather station measures and RS data, and have high applicability in practice, the following objectives have to be achieved:

1. Develop ANN models that can simulate ammonium, nitrate and total dissolved N in streams draining undisturbed and disturbed watersheds;
2. Develop ANN models that are capable of modelling aquatic N concentrations and loadings in unmonitored watersheds only relying on climate data and RS information;
3. Investigate the applicability and transferability of the above developed models (2) in unmonitored watersheds:
    a. Develop individual and combined watershed similarity indices between any two watersheds ($i - j$), which can guide model transferability,

b. Apply the calibrated ANN models based on watershed *i* to a different watershed *j* without further calibration, and

c. Find a watershed similarity index that can reasonably guide model transferability by investigating the relationship between similarity indices and the performance of model transfer; and

4. Propose a framework to incorporate ANNs models into forest management.

1.5 Thesis Organization

A paper format has been employed to this thesis to preserve the diversity of the models and achieve the research objectives.

Chapter 2 provides comprehensive reviews on: the application of ANNs to water quality modelling, which is very beneficial to understand the ANN modelling work in this thesis; the application of geographic information system (GIS) and RS to water quality modelling as they have become important tools for watershed modelling and watershed management; and how snowpacks in winter affect N in streams because in north region the existence of microbial activities in snowpacks can influence the N export into streams during snowmelt.

The first part presents the general procedure and methodology of ANN model development; summarizes and compares different types of ANNs that have even been used in water quality modelling; provides the advanced studies to overcome the challenges for ANN models; reviews and discusses the applications of ANNs to modelling surface water quality parameters; and suggests future research recommendations to improve ANNs applications in water quality modelling.

The second part provides essential technical information of RS and GIS that are important to development and application of simulation models to watershed management. RS data provides valuable information about watershed characteristics for water quality modelling. Generally, RS data can be used to estimate the input parameters (e.g., surface temperature, soil moisture) for watershed models, delineate watershed and streams, classify land-use and land-application, monitor water quality parameters, and provide vegetation indices data for vegetation dynamics study. RS data can be best utilized if they are integrated into a GIS that is designated to manage large volumes of data. The use of GIS ranges from just display and visualization of results to storage and retrieval of remotely sensed data and environmental data, from spatial analysis of landscape and preparation of model parameters for integration with watershed modelling. Three of the most common strategies for linking a GIS to a simulation model are loose coupling, tight coupling and an embedded system approach.

The third part helps to understand the seasonal variation of N in streams and the determination of ANN N model input variables. A number of in-situ and lab works have confirmed the active N processes under snow cover, and demonstrated that overwinter N processes are important and should be considered when modelling N leaching and concentrations in surface water. Snowpack insulates soil from the very low atmospheric temperature in winter and thus enables microbial activities including mineralization and nitrification/denitrification. Consequently, this leads to increase of N in soil under certain conditions of snow accumulation, snow depth and snow consistence. The major sources of N into streams after winter break are snowpacks and soils. It is commonly recognized that inorganic N concentrations peak in surface water during snowmelt or winter breaks.

Chapter 3 is the modelling of N compositions in streams on the Boreal Plain using GRNN. GRNN was selected because there is only one parameter to be optimized and it is very fast to train. GRNN models were developed following strict procedures and applied to simulate daily mean NO_3^-, ammonium (NH_4^+) and total dissolved N (TDN) concentrations in streams at three watersheds in the Swan Hills of Alberta, Canada. The optimal inputs were derived from five major variables: rainfall, daily mean air temperature, cumulative degree-days, enhanced vegetation index (EVI) and Julian day of the year. The consistent performance of GA-GRNN models for two relatively undisturbed watersheds, as well as a burned watershed, was obtained with the inclusion of the RS-derived EVI as one of the model inputs. This index was capable of describing vegetation canopy differences among watersheds, as well as vegetation phenology. In terms of model architecture, the developed models were not sensitive to the initial smoothing factor and training with a genetic algorithm improved model performance on testing data sets. The developed models successfully simulated NO_3^-, NH_4^+ and TDN concentrations for three streams, with r^2 values exceeding 0.83 for all data sets. This study distinguished itself from other N modelling studies (Kaluli et al. 1998; Lek et al. 1999; Sharma et al. 2003; Yu et al. 2004; Khalil et al. 2005; Almasri and Kaluarachchi 2005) in that it: (1) explored the water quality modelling capability of GRNNs trained with a genetic algorithm; (2) took into consideration the dynamics of vegetation phenology on N modelling by using RS data and; (3) developed GA-GRNN models that successfully predicted not only NO_3^-, but also NH_4^+ and TDN concentrations. More importantly, it implies the high potential of applying GA-GRNN models for predicting other surface water quality parameters on other similar or different watersheds.

As further study from Chapter 3, in Chapter 4 a model of N export in unmonitored watersheds was developed by transferring trained models to new watersheds that the models have never seen during their calibration and find out an index that can measure models' transferability. MLP-BP N export models were developed using low-cost, readily available meteorological data and satellite data in forested watersheds. The performance of the models was evaluated using correlation-based measure, absolute error measures and time series plot of measured and modelled values. Although the modelled parameter varied over a big range (i.e., the peak values were thousands of times the low values), it was simulated fairly well. The best MLP-BP architecture for all the models had a single hidden layer with three activation functions. The networks were trained using either typical gradient BP or BP with batch update. Modelling N export only using readily available data with reasonable accuracy indicates its potential application to unmonitored watershed. To demonstrate the applicability of the developed models to unmonitored watersheds, the calibrated models were used to predict N export in other different watersheds without further calibration. The Nash Sutcliffe coefficient E was greater than 0, which means that the models produced better estimates than the mean of the observed values. The correlation coefficient r^2 and index of agreement d were in the range of 0.44 to 0.63 and 0.73 to 0.88, respectively. The transferred models could catch the seasonal and annual periodicity of N export even though some peak values were not well predicted. The overall results indicated that transferring the calibrated models developed using the proposed algorithm to other different watersheds is promising. To transfer models to unmonitored watersheds, it is very important to find a similarity index that can effectively measure watershed similarity because the more similar the watersheds are the greater success in the model transfer. The indexes representing watershed soil types, rainfall and watershed vegetation conditions were computed then their correlation with model transfer performance was tested. For each single index, Rainfall_Index (r^2 = 0.71) and NDWI_Index (normalized difference water index) (r^2 = 0.69) had the highest correlation with model transferability. Peatland_Index, Riparian_Index and EVI_Index were not significantly correlated with model transferability. The best watershed similarity index was (Rainfall_Index+ Peatland_Index+ NDWI_Index) (r^2 = 0.74) when all of the factors of watershed soil types, rainfall and vegetation conditions were considered.

Chapter 5 provides a framework to incorporate water quantity/quality modelling into forestry management. This chapter summarizes the modelling results from two approaches, SWAT and ANN, applied to FORWARD project. The framework is: (1) Delineate the digital elevation model (DEM) of the MWFP FMA area into first order watersheds (~ 5 km^2) using the eight-direction pour-point algorithm and a reasonable threshold for flow accumulation. (2) Use rainfall interpolation techniques like kriging and inverse distance weighted interpolation techniques to estimate daily rainfall intensity in the centroid of each watershed using data from surrounding weather stations (fire towers, Environment Canada, and FORWARD project stations). (3) Formulate streamflow (Q), Total suspended solids (TSS) and nutrient models for several of the 16 FORWARD experimental watersheds and validate with the remaining watersheds. (4) The previously delineated watersheds, including the FORWARD study watersheds, will be grouped into different categories according to hydrologic homogeneity in term of VIs, average slope, % wetland composition, yearly precipitation, and basin area. (5) Each calibrated model is then run for all the watersheds falling into its group of similar watersheds. Upon successful implementation of models to the whole FMA area, scenario-based analysis that forces harvesting disturbance on the landbase can be fed to the models to identify the impact of different land use activities on water quality and quantity. To design these scenarios, a relation has to be established between VIs and typically used vegetation metrics (e.g., timber volume, average age, height, and diameter at breast height). This relation can be used to translate vegetation cut into values of VIs that can be fed to the models to predict changes in streamflow, water-phase solids, and nutrients in response to harvesting scenarios.

Chapter 6 presents overall conclusions and recommendations for further research in this area. The contribution of this research to academia is that it provides an approach to simulate N in unmonitored watersheds at daily scale. Specifically, it improves N modeling through incorporating vegetation dynamics, which significantly affect N cycle and leaching into streams. This research contributes to industry by developing effective and cost-efficient models that produce satisfying accuracy but only require public domain data. They have high practical values for lumbering industry because the developed models can link lumbering activities with water quality parameters through VIs. The models can be potentially applied to support decision-making on planning lumber harvesting and meet environmental regulatory requirements.

1.6 References

Alp, M., and Cigizoglu, H.K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ. Modell. Softw., **22**: 2-13.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., and Williams, J.R. 1998. Large area hydrologic modeling and assessment part I: Model development. J. Am. Water Resour. Assoc., **34**: 73-89.

Barnes, W.L., Pagano, T.S., and Salomonson, V.V. 1998. Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. IEEE Trans. Geosci. Remote Sens., **36**: 1088-1100.

Beasley, D.B., Huggins, L.F., and Monke, E.J. 1980. ANSWERS - A MODEL FOR WATERSHED PLANNING. Trans. Asae, **23**: 938-944.

Bingneer, R.L., Theurer, F.D., Cronshey, R.G., and Darden, R.W. 2001. AGNPS 2001. Available from http://msa.ars.usda.gov/ms/oxford/nsl/AGNPS.html.

Bouraoui, F., and Dillaha, T.A. 1996. ANSWERS-2000: Runoff and sediment transport model. J. Environ. Eng.-ASCE, **122**: 493-502.

Bowden, G.J., Nixon, J.B., Dandy, G.C., Maier, H.R., and Holmes, M. 2006. Forecasting chlorine residuals in a water distribution system using a general regression neural network. Math. Comput. Modell., **44**: 469-484.

Breiman, L. 1994. Neural Networks: A Review from Statistical Perspective]: Comment. Stat. Sci., **9**: 38-42.

Chanasyk, D.S., Whitson, I.R., Mapfumo, E., Burke, J.M., and Prepas, E.E. 2003. The impacts of forest harvest and wildfire on soils and hydrology in temperate forests: A baseline to develop hypotheses for the Boreal Plain. J. Environ. Eng. Sci., **2**: S51-S62.

Cheng, Y.B., Zarco-Tejada, P.J., Riano, D., Rueda, C.A., and Ustin, S.L. 2006. Estimating vegetation water content with hyperspectral data for different canopy scenarios: Relationships between AVIRIS and MODIS indexes. Remote Sens. Environ., **105**: 354-366.

Cigizoglu, H.K., and Alp, M. 2006. Generalized regression neural network in modelling river sediment yield. Adv. Eng. Softw., **37**: 63-68.

Cigizoglu, H.K., and Kisi, O. 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol., **317**: 221-238.

Cook, D.J., Dickinson, W.J., and Rudra, R.P. 1985. GAMES-the Guelph Model for Evaluating the Effects of Agricultural Management Systems in Erosion and Sedimentation. User's Manual. Guelph, Ont.

Dakou, E., Goethals, P.L.M., D'Heygere, T., Dedecker, A.P., Gabriels, W., De Pauw, N., and Lazaridou-Dimitriadou, M. 2006. Development of artificial neural network models predicting macroinvertebrate taxa in the river Axios (Northern Greece). Anna. Limnol.-Int. J. Limnol., **42**: 241-250.

Diamantopoulou, M.J., Georgiou, P.E., and Papatnichail, D.M. 2007. Performance of neural network models with Kalman learning rule for flow routing in a river system. Fresenius Environ. Bull., **16**: 1474-1484.

Elhatip, H., and Komur, M.A. 2008. Evaluation of water quality parameters for the Mamasin dam in Aksaray City in the central Anatolian part of Turkey by means of artificial neural networks. Environ. Geology, **53**: 1157-1164.

Johanson, R.C., Imhoff, J.C., Little, J.L., and Donigian, A.S. 1984. Hydrological Simulation Program-Fortran (HSPF) User's Manual, Athens, GA.

Kaluli, J.W., Madramootoo, C.A., and Djebbar, Y. 1998. Modeling nitrate leaching using neural networks. Water Sci. Technol., **38 (7)**: 127-134.

Kisi, O. 2004. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. Hydrol. Sci. J.-J. Sci. Hydrol., **49**: 1025-1040.

Koponen, S., Kallio, K., Pulliainen, J., Vepsalainen, J., Pyhalahti, T., and Hallikainen, M. 2004. Water quality classification of lakes using 250-m MODIS data. IEEE Geosci. Remote Sens. Lett., **1**: 287 - 291.

Legates, D.R., and McCabe, G.J. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res., **35**: 233-241.

Lek, S., Guiresse, M., and Giraudel, J.-L. 1999. Predicting stream N concentration from watershed features using neural networks. Water Res., **33**: 3469-3478.

Maier, H.R., Burch, M.D., and Bormans, M. 2001a. Flow management strategies to control blooms of the cyanobacterium, Anabaena circinalis, in the River Murray at Morgan, South Australia. Regulated Rivers-Res. Manage., **17**: 637-650.

Maier, H.R., Sayed, T., and Lence, B.J. 2001b. Forecasting cyanobacterium Anabaena spp. in the River Murray, South Australia, using B-spline neurofuzzy models. Ecol. Modell., **146**: 85-96.

Maier, H.R., and Dandy, G.C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Modell. Softw., **15**: 101-124.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006a. 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. Modell., **191**: 19-32.

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

Pelster, D.E., Burke, J.M., and Prepas, E.E. 2008. Runoff and inorganic N export from Boreal Plain watersheds six years after wildfire and one year after harvest. J. Environ. Eng. Sci., **7**: S51-S61.

Prepas, E.E., Burke, J.M., Chanasyk, D.S., Smith, D.W., Putz, G., Gabos, S., Chen, W., Millions, D., and Serediak, M. 2003. Impact of wildfire on discharge and phosphorus export from the Sakwatamau watershed in the Swan Hills, Alberta, during the first two years. J. Environ. Eng. Sci., **2**: S63-S72.

Putz, G., Burke, J.M., Smith, D.W., Chanasyk, D.S., Prepas, E.E., and Mapfumo, E. 2003. Modelling the effects of boreal forest landscape management upon streamflow and water quality: Basic concepts and considerations. J. Environ. Eng. Sci., **2**: S87-S101.

Reed, B.C., Brown, J.F., and Loveland, T.R., 2002. Geographic data for environmental modelling and assessment. In: Skidmore, A. ed., Environmental Modelling with GIS and Remote Sensing. Taylor & Francis, London, pp. 52-69.

Rudra, R.P., Negi, S.C., and Gupta, N. 2005. Modelling approaches for subsurface drainage water quality management. Water Quality Res. J. Canada, 40: 71-81.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. 1986. Learning representations by back-propagating errors. Nature, **323**: 533-536.

Salehi, F., Prasher, S.O., Amin, S., Madani, A., Jebelli, S.J., Ramaswamy, H.S., Tan, C., and Drury, C.F. 2000. Prediction of annual nitrate-N losses in drain outflows with artificial neural networks. Trans. ASAE, **43**: 1137-1143.

Sharma, V., Negi, S.C., Rudra, R.P., and Yang, S. 2003. Neural networks for predicting nitrate-N in drainage water. Agirc. Water Manage., **63**: 169-183.

Sharpley, A.N., and Williams, J.R. 1990. EPIC-Erosion/Productivity Impact Calculator: I. Model Documentation, Tech. Bull,.

Smith, D.W., Russell, J.S., Burke, J.M., and Prepas, E.E. 2003. Expanding the forest management framework in the province of Alberta to include landscape-based research. J. Environ. Eng. Sci., **2**: S15-S22.

Suen, J.P., and Eheart, J.W. 2003. Evaluation of neural networks for modeling nitrate concentrations in rivers. J. Water Resour. Plann. Manage.-Asce, **129**: 505-510.

Tayfur, G., and Guldal, V. 2006. Artificial neural networks for estimating daily total suspended sediment in natural streams. Nordic Hydrol., **37**: 69-79.

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

the ASCE Task Committee. 2000. Artificial Neural Networks in Hydrology. I: Preliminary Concepts. J. Hydrol. Eng., **5**: 115-123.

Tibshirani, R. 1994. [Neural Networks: A Review from Statistical Perspective]: Comment. Stat. Sci., **9**: 48-49.

Tsoukalas, L.H., and Uhrig, R.E., 1997. Fuzzy and Neural Approaches in Engineering. A Wiley-Interscience publication, New York : Wiley.

Yu, C., Northcott, W.J., and McIsaac, G.F. 2004. Development of an artificial neural network for hydrologic and water quality modeling of agricultural watersheds. Trans. ASAE, **47**: 285-290.

CHAPTER 2. STREAM WATER QUALITY MODELLING FOR WATERSHED MANAGEMENT USING ARTIFICIAL NEURAL NETWORKS: A REVIEW

2.1 Introduction

Stream water quality modelling involves understanding physical, geochemical and biological processes that includes a number of inter-related factors such as watershed features, meteorological factors, geological factors and anthropogenic effects. However, the relationships between these processes and water quality parameters are complex, not deterministic and currently even not fully understood to our knowledge. Therefore, artificial neural networks (ANNs), capable of modelling complicated and non-linear processes, have been gaining popularity in water quality modelling over the past decade. As for large scale areas especially where measurements are not feasible or cost prohibitive, remote sensing (RS) data provide valuable information required for watershed modelling and management. In this context, geographic information system (GIS) technology, as an essential and functional tool for handling and utilizing a large volume of spatial data, plays an important role on dealing with RS data and supporting watershed management.

In this research, to successfully model nitrogen (N) using ANNs in ungauged watersheds with the aid of remotely-sensed information and apply the developed ANNs models to watershed management, it is important to: (1) follow a systematic procedure for ANN model development, (2) determine the model inputs and internal parameters, (3) understand the advantages and limitations of ANNs in modelling water quality, (4) find out the periodicity of N in streams in north region, (5) get knowledge of the physical processes affecting N in streams, in particular, how does snowmelt affect N leaching from soil to streams in north region, (6) know the use of RS data for water quality modelling in general and what type of RS data can serve N modelling, (7) understand in which ways GIS can contribute to RS data processing and water quality modelling. Therefore, this review is prepared to summarize knowledge of the above concerns to fulfill the research objectives.

2.2 The Modelling Tool - Artificial Neural Networks

2.2.1 ANNs

ANNs are an artificial intelligence (AI) approach that mimics the operation of human brain to solve problems. From a statistical perspective, ANNs essentially employ a similar modelling philosophy as that of conventional statistical models to predict environmental variables (Maier and Dandy 2001), where the objective is to identify the relationships between model input variables and corresponding output variables. This objective is achieved with ANNs by employing enough data sets representing the input-output relationship for adjusting the model internal parameters (e.g., the connection weights) and minimizing an error function between the predicted and observed outputs (Maier and Dandy 2001). ANN models have become an attractive modelling tool in solving complex problems. The ASCE Task Committee (2000) observed that ANNs: (1) are able to identify the relationship between inputs and outputs without requiring the full understanding of the mechanistic principles behind them; (2) can work well even when the training sets contain noise and measurement errors; (3) can be adapted to compensate for changing circumstances; and (4) are easy to use once trained.

An ANN mode is a data processing system consisting of a large number of simple and highly interconnected processing neurons (or called elements) in an architecture inspired by the structure of the cerebral cortex of the brain (Tsoukalas and Uhrig 1997). There are seven major components to an ANN architecture: (1) processing neurons, (2) a state of activation, (3) an output function for each neuron, (4) a pattern of connectivity or weights between neurons, (5) a propagation rule, (6) an activation function to combine the inputs impinging on a neuron with the current state of that neuron to produce a new level of activation for that unit, (7) a learning rule whereby weights are adjusted (Rumelhart et al. 1986). The processing neurons are generally organized in layers. ANNs are generally categorized into two groups that are feed-forward networks and feed-backward networks based on the direction of information flow and processing. Feed-forward neural networks, the most popular category, have model inputs processed forward through the network in sequence and do not contain the feed-back connections necessary to provide a dynamic mode. Feed-backward networks have recurrent loops within the architecture that make possible for the network to retain a short memory with respect to the previous input information. Such incorporation of information makes feed-backward networks particularly suitable for modelling time dependent systems. In order for an ANN to generate output predictions that are as close as possible to the objective values, a training process, also called learning, is employed to find optimal ANN parameters (e.g., weights, smoothing factors) minimizing a predetermined error function.

The types of ANN models that have been used for surface water quality modelling include feed-forward multilayer perceptron trained with the back-propagation algorithm (MLP-BP), recurrent neural networks (RNN), radial basis function networks (RBFNs), general regression neural networks (GRNNs), modular neural networks (MNNs), Kohonen's self-organizing map (KSOM), and associative memory network (AMN). The architecture of these ANNs is summarized in Figure 2-1. Table 2-1 describes these ANNs and compares the advantages and limitations of each type of ANNs. MLP-BP is the most popular type of ANNs for prediction and forecasting applications even though other ANNs sometimes performed better than MLP-BP.

2.2.2 Advanced Studies on ANN Modelling

Although ANN models have been considered as successful tools for complex problems, it is necessary to be aware of their limitations to ensure their proper implementation and an adequate interpolation of the modelling results. The major concerns about ANNs have resulted from the fact that they are looked as 'black-box', which means that they do not represent physical mechanisms. This 'black-box' recognition of ANNs leads to the difficulty of interpreting and extrapolating the modelling results. However, modellers commonly expect that proper interpretation of modelling results and extraction of physical knowledge about the system being modelled is possible. The most recent research is conducted in the following perspectives to make ANNs a more desirable modelling tool.

2.2.2.1 Hybrid ANN Models

The combination of ANNs with numerical models has gained considerable interest as the hybrid neural models with a degree of deterministic components integrated into them, can achieve the best of both the neural networks and numerical models. Within the hybrid model structure, the numerical model specifies the basic dynamics of the relevant process variables, and the ANNs model properly accounts for the unknown and nonlinear parts of numerical models (Lee et al. 2002). Typically, there are two types of hybrid models: the serial hybrid neural model and the parallel hybrid model. In the serial hybrid model, an ANN model is put in series with a numerical model. The numerical model mathematically describes the well understood system whereas the ANNs model describes the unknowns (Chen and Adams 2006). In the parallel hybrid model, an ANNs model is placed in parallel with a numerical model. The ANNs compensates for the differences between the numerical model predictions and the process data (Zhao et al. 1997; Lee et al. 2002). In addition, ANNs have been combined with fuzzy logic and genetic algorithms to take account of uncertainties (Chen et al. 2001; Choi and Park 2001; Chen et al. 2003a; Nayak et al. 2007) and combined with traditional regression models (e.g., autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA)) to improve the forecasting accuracy for time series data (Jain and Kumar 2007; Sallehuddin et al. 2007; Valenzuelaa et al. 2008).

2.2.2.2 Knowledge Extraction

ANNs was thought to explain very little about the influences of the independent input variables in the prediction process (Olden and Jackson 2002; Olden et al. 2004). However, recent studies have confirmed that physical information can be extracted from the internal structures of ANN models (Jain et al. 2004). The contributions of input variables can be evaluated by the 'partial derivatives' method, the 'weights' method, the 'pertub' method, the 'profile' method, the ‘classical stepwise’ method, the ‘improved stepwise a’ method, and the ‘improved stepwise b’ method (Gevrey et al. 2003). Studies are ongoing to develop the methodology to extract knowledge from trained ANNs to make ANNs a valuable quantitative tool to evaluate, understand and predict environmental phenomena (Olden and Jackson 2002; Wilby et al. 2003; Jain et al. 2004; Olden et al. 2004; Sudheer and Jain 2004; Sudheer 2005; Gevrey et al. 2006; Kingston et al. 2006). So far, the commonly used methods to determine the contribution of input variables to output are based on sensitivity analysis and optimal weights. Knowledge from ANNs models is usually extracted by interpreting information contained in a single optimal weight vector (Gevrey et al. 2003; Olden et al. 2004). However, there exist many different weight vectors that result in similar network performance, which means that there may be a range of ‘optimal’ weight vectors. Therefore, a probabilistic knowledge extraction from the weights of an ANN to reveal the range of relationships between the model inputs and outputs should be considered (Kingston et al. 2006).

2.2.2.3 Model Generalization

Besides simulation accuracy, model generalization is a major criterion to evaluate the performance and power of ANNs models. Model generalization is primarily related to the learning or training methods and the amount and representativeness of training data (Bishop 1995). Studies have been carried out to investigate how to alleviate over-fitting and improve generalization through improving learning algorithm (e.g., early stopping) (Anctil and Lauzon 2004), developing simpler ANNs without reducing its prediction accuracy (Gaume and Gosset 2003), and selecting representative training data (Anctil and Lauzon 2004; Shabin et al. 2004). The quantity and quality of data, the type of noise and the mathematical properties of the algorithm for estimating the usual large number of parameters are crucial for the generalization ability of ANNs (Giustolisi and Laucelli 2005).

In addition to optimizing a single network, combining multiple ANNs is another way to improve the generalization ability of ANNs because it is difficult to develop a single neural network with high modelling quality for a complex task. The combination of several ANNs, namely aggregated/compound neural networks (A/C-NN), appeared to be a promising approach (Mani and Omkar 2002; Versace et al. 2004; Wang and Yan 2004; Ahmad and Zhang 2005a, b; Samanta et al. 2005; Wang et al. 2006). Combination of multiple neural networks can work in two ways: (1) Training-A/C-NN: The individual models simulating the same relationship are developed from different data sets and/or different training algorithms and combined to provide the outputs through methods such as stacking, bootstrapping and Bayesian selective combination; (2) Subtasks-A/C-NN: A whole task is divided into specific situations (or subtasks) based on a certain criteria. The overall forecast is a weighted sum of the individual model output for each situation. To test the effectiveness of training-A/C-NN, multiple neural networks were combined using data fusion techniques by which the combination weight of each neural network was determined according to the model input data (Ahmad and Zhang 2005b).

2.2.2.4 Guidelines of ANNs Model Development

The basic theory of ANNs and the case studies are described very well in most publications. However, the development of ANNs models in a systematic way, which is crucial to find the optimal model, is not properly described in general. It is urgently needed to establish guidelines that assist with the development and applications of ANN models (Wilby et al. 2003; Al-Yemni and Yang 2005; Dawson et al. 2006). Guidelines that can effectively assist with the development of ANN models need to include at least the following components: select suitable network algorithm such as MLP-BP, GRNN, RBF, estimate neural network parameters (e.g., activation functions, learning rate), determine adequate model inputs, divide the whole data into representative subsets, find proper training algorithm and stopping criteria, evaluate model's performance properly, and interpret modelling results. Although a comprehensive guide covering each step of developing an ANNs model has not been presented, it is promising that the consideration of individual components has been addressed in studies, such as determination of input parameters (Bowden et al. 2005a; Bowden et al. 2005b) and data division (Bowden et al. 2002; Shabin et al. 2004).

2.2.3 Application of ANNs for Water Quality Modelling

Over the past decade, ANNs have been successfully used to predict and model a variety of water quality parameters such as sediment, nutrients, microbial concentration, and salinity. The information that is useful and of interest to modelers and researchers is summarized in Table 2-2.

According to Table 2-2, MLP-BP is the most widely used ANNs for water quality modelling because MLP-BP is capable of handling a wide range of problems (Table 2-1). Comparison studies indicate that ANNs are superior to linear regression to model time-correlated variables. MLP-BP sediment models were significantly superior to conventional sediment rating curve and MLR method (Cigizoglu and Alp 2006; Alp and Cigizoglu 2007). In modelling total phosphorus concentration, MLP-BP models outperformed autoregressive moving average with exogenous input (ARMAX) (Nour et al. 2006b). The MLP-BP model was able to predict several strains of indicator bacterial concentrations in a river from other related physical, chemical, and bacteriological variables and was superior to conventional imputation and MLR models (Chandramouli et al. 2007).

MLP-BP models have been compared with other ANNs (e.g., RBF, GRNN) models to explore the advantages of different ANNs (Table 2-2). GRNN and RBF performed better than MLP in the estimation of total sediment load (Kisi 2004). Realizing the limitation of MLP-BP approach to interpret the relationship between the inputs and the outputs, B-spline AMNs were developed to forecast *Anabaena* spp. (Maier et al. 2000, 2001). The weights of AMNs can be interpreted as a set of fuzzy membership functions and then the relationship between the model inputs and outputs are written as a set of fuzzy rules. AMNs and MLP models obtained comparable forecasting accuracy, but the B-spline AMNs model provided more explicit information, for example, indicated that the incidences of *Anabaena* spp. generally occurred after passing of the flood hydrograph and when water temperature were high (Maier et al. 2000, 2001). The surface water quality parameters generally vary in a big range and have a high standard deviation. Thus, the ANNs models sometimes fail to predict the peaks with satisfied accuracy. Accordingly, range dependent neural networks (RDNN) (also called A/C-NN) were developed to improve the prediction accuracy for extreme values (Cigizoglu and Kisi 2006; Jain 2008). The RDNNs trained with three subsets of training data were found to have strengths over conventional ANN in predicting both low and high sediment values more precisely. Similarly, a A/C-NN consisting of more than one ANNs, each of which was trained with a subset of data, modelled the integrated stage-discharge-suspended sediment rating relationship more precisely than a single ANN (Jain 2008). The MNN model that worked through dividing the whole task into subtasks showed better simulation results than GRNN model to describe the water and nutrient mass dynamics (Kim and Kim 2007).

Table 2-2 shows that the surface water quality parameters are modelled in daily, weekly and monthly time steps. To use the ANN models for watershed management, it is important that ANN models are able to forecast water quality in advance. Actually, the most ANN models only predict the modelled parameters at current time except for sediment prediction (1 day ahead) (Cigizoglu 2004), nitrate (1 week ahead) (Markus et al. 2003), Cynobacteria (4 weeks ahead) (Maier et al. 1998, 2000, 2001), and salinity (14 days ahead) (Maier and Dandy 1999; Kingston et al. 2005; Bowden et al. 2002, 2005a, 2005b). Input determination plays a significant role on the success of the ANNs forecasting models. The above studies used not only a prior knowledge about the modelled system, but correlation analysis, partial mutual information (PMI) and trial-and-error to find the antecedent variables highly related to the modelled parameters.

The database required for the ANN models is very important as data availability may present a restriction for model selection and performance in practice. It is critical to carefully select a representative and significant set of input variables to develop a robust ANN model. Table 2-3 summarizes the input variables for modelling sediment, nutrients, microbial concentration, and salinity, which is a good reference for input determination to anybody who is interested in developing ANNs models for these parameters.

2.3 Water Quality Modelling with the Aid of GIS and RS

2.3.1 GIS and Its Applications

GIS is composed of four basic components: (1) data input and editing, (2) storage of geographic databases, (3) data analysis and data modelling, and (4) data visualization and presentation (Mattikalli 2000). The uses of GIS include display and visualization of results (Chen et al. 2003b; Yacobi and Schlichter 2003), storage and retrieval of remotely sensed data and environmental data (Ballester et al. 2003; Xu and Zeng 2003), preparation of model parameters (Almasri and Kaluarachchi 2005), spatial analysis of landscape (Cedfeldt et al. 2000; Brown 2006), and integration with environmental modelling (Bhaduri et al. 2000; Baker et al. 2001). In most of the cases, GIS is often used to combine information from different sources such as cartography (i.e. maps), remote sensing (i.e. aerial and satellite imagery), hydrology surveys, environmental monitoring, and create overlapping layers that can be accessed, transformed, and manipulated interactively in one spatial structure. With regard to water quality modelling, GIS has contributed to solve the following problems:

2.3.1.1 Estimating and Assessing Water Quality

GIS is often applied for water quality and runoff modelling at a watershed scale as they are increasingly affected by spatially distributed human activities such as urbanization, agricultural development and logging. In this context, GIS has been shown to be useful for handling spatial data and mapping the results. In addition, GIS plays a key role in assessing and analyzing the contributed factors and the results. For example, it is used to locate high risk of contamination by pesticides (Sinkevich et al. 2005), assess pollutants distribution (Davies and Neal 2004) and lake eutrophication (Xu et al. 2001), evaluate pollution risks to water supply intakes (Foster and McDonald 2000), analyze land use/land cover (LULC) change impact on water quality (Basnyat et al. 2000; Bhaduri et al. 2000; Choi et al. 2003; Haverkamp et al. 2005) and agricultural pressures and impacts on water quality (Giupponi and Vladimirova 2006).

2.3.1.2 Identification and Quantification of both Point and Non-point Sources of Pollution

GIS has traditionally been used to identify and quantify both point and non-point sources of river pollution such as identify the main contamination sources of heavy metals, organic compounds and other physicochemical parameters in river waters and describe their temporal and spatial distributions (Terrado et al. 2006), develop nitrate leaching maps for different N fertilization rates in an agricultural watershed (de Paz and Ramos 2004), estimate the spatial and temporal N-loading patterns to a watershed (Fernandez et al. 2002, 2006), identify potential pollution resulting from manure spread on agricultural lands (Giasson et al. 2002), and identify the sources of nitrate pollution (Matejicev et al. 2003) and nutrient load to river (Mourad et al. 2005).

2.3.1.3 Other Applications

GIS is also an important tool for watershed management. It has been used for sustainable management of water resources (Belmonte et al. 1999), watershed management (Clark 2000; Blum 2001; Choi et al. 2005; Lant et al. 2005), GIS-based decision support system for water assessment (Chowdary et al. 2003).

2.3.2 Development of Earth Observing System

RS is one of the most successful GIS-related technologies, which are defined as the technologies that are commonly used in conjunction with GIS. In practice, RS gathers reflected radiation from the atmosphere. This information can be related to other conditions and then used to understand the Earth's environment. RS also provides important coverage, mapping and classification of landcover features, such as vegetation, soil, water, forests and urban. The spaceborne RS is the only cost-efficient way to improve the spatial and temporal coverage of surface water and watershed monitoring (Koponen et al. 2004). RS is an integral part of GIS because it provides substantial amount of spatial data to the GIS databases. Without the data from RS, GIS cannot explore its full functionality (Skidmore 2002)

The first Earth Observing System (EOS) satellites, as part of the Earth Science Enterprise (ESE) comprehensive program by National Aeronautics and Space Administration (NASA) were launched in December of 1999. NASA's EOS and its mission were initiated to contribute to decision making and allow the development of policies to preserve and protect the Earth's environment based on a better understanding of the biosphere on Earth (Robinson 1995). A moderate-resolution imaging spectroradiometer (MODIS) is the key EOS instrument and its mission is to study global/local interactions. MODIS was designed to improve monitoring of land, ocean and atmosphere, in particular, to provide innovative land products data that are especially designed to support modelling applications (Reed et al. 2002).

Focused on environmental monitoring in general and the estimation of vegetation indices (VIs) in particular, MODIS provides spectral and spatial resolution superior to that of advanced very high resolution radiometer (AVHRR) sensor (Fensholt and Sandholt 2005). The MODIS collects reflected and emitted radiation from the earth in 36 bands from 0.405 to 14.385 μm and provides spatial resolution of 250 m, 500 m and 1 km (Barnes et al. 1998). The upgraded spectral resolution improved cloud and atmosphere characterization which enabled the removal of atmospheric effect on surface observation and the provision of atmospheric measurement (Justice et al. 1998). The combination of coarse and high resolution data are needed to realize systematic global monitoring of land surface (Skole et al. 1997).

The MODIS science team developed software to generate data products that meet the requirements of global change research. The land discipline group provided a combination of basic surface variables of spectral reflectance, albedo, and land surface temperature as well as higher order variables such as VIs, leaf area index (LAI), fraction of absorbed photosynthetically active radiation, active fires, burned area, and snow and ice cover (Justice et al. 1998). The MODIS Land

Group provide not only satellite data but also high level data products that are specifically designed to support the global to regional monitoring, modelling, and assessment. Specially, the MODIS Science Team came out with a new data product that is enhanced vegetation index (EVI). Satellite measurements of leaf area, leaf duration and net primary productivity provide important inputs to parameterize or validate ecosystem process models.

2.3.3 The Integration of RS and GIS

RS data can be best utilized if they are integrated into a GIS that is designated to manage large volumes of data. An important feature of GIS is that it is able to overlay different layers of spatially geo-referenced data. Thus, integrating remote sensed data into GIS enables the user to graphically and analytically determine how spatial structures and objects interact with each other (Mattikalli 2000). The information stored in a GIS is only static while a basin hydrological system is dynamic. Thus, data in a GIS need to be updated to represent the temporal change. In this context, remotely sensed satellite data offer excellent input to the GIS to provide repetitive, synoptic, and accurate environmental information of the changes in a wide range of spatial scales, and offer the potential to monitor these dynamic changes. Remotely sensed data can be used to create a permanent geographically located database to provide a baseline for future comparisons (Ritchie et al. 2003). Moreover, successful applications of RS have contributed to modification of existing water quality models and development of new types of models to incorporate widely available spatial data. RS data enable us to map the variation in terrain properties spatially and temporally, such as vegetation, water and geology. In many cases, to serve the modelling purpose RS data have to be merged with ancillary georeferenced information such as soil, geology, and elevation. At this point, GIS offers an appropriate technology for merging these various layers of spatial data.

Technically, information is extracted from satellite imagery by image processing methods and then combined with other data layer in a GIS. Satellite imagery is stored in a raster format that makes it ideally suited for incorporation into a GIS. Therefore, RS imagery can be easily imported into a GIS project as an image theme/layer. Integration of remotely sensed data with a GIS can greatly enhance modelling and analyzing capability of the GIS, and its potential has been demonstrated in many areas of hydrology and water management.

2.3.4 The Applications of GIS and RS to Water Quality Modelling

Watershed-scale water quality models typically require a considerable amount of data (e.g., topography, vegetation cover, soil characteristics, stream channel characteristics and subsurface infiltration) for model calibration, which is sometime cost-prohibitive, especially for forested watershed. The RS technology provides a cost-effective way to evaluate and quantify large numbers of watershed physical characteristics and state variables. RS techniques have expanded widely, to the point that they now include most of the electromagnetic spectrum and can provide unique information about properties of the surface or shallow layers of the Earth.

To acquire RS data efficiently and appropriately for its specific application, spatial resolution, temporal resolution, spectral resolution and spectral wavelength need to be considered. The choice of spatial resolution depends on the nature of the problems and the details needed in the watershed-scale models. Thus, it is often a trade-off between cost and the sufficient details needed. RS data are acquired with a given resolution in time. In some cases of dynamic processes and small basins, the data may be needed daily or more often. Whereas in some less dynamic processes and large basins, data on two weeks or longer may be satisfactory (Schultz and Engman 2000). In addition, the spectral resolution and wavelength should be selected because the reflectances from different spectral bands are used specifically to identify the different features (vegetation, clear water, soil, etc.) on the Earth. Table 2-4 shows remote sensing applications of different spectral bands.

RS data in various spectral bands provides information on watershed characteristics (e.g., landcover, land use, vegetation, etc.) (Table 2-4). Some of water quality model parameters can be estimated using the comprehensive RS information. Another important aspect of RS is that it provides data in remote areas, where measurements are not feasible or can not be carried out due to prohibitive cost. The limitation of data is the main reason that results in the unsatisfactory performance of so many water quality models. Hence, in many fields of water quality modelling, it is expected that the availability of RS information will lead to the development of much more efficient models.

The rapid development of RS technology, the reduced cost of acquiring RS data, and the capability of RS to assist in describing watershed characteristics have lead to the incorporation of RS information into water quality models. GIS and RS technologies have become indispensible tools for watershed hydrological analysis, modelling and management (Jones et al. 2002). The use of RS to provide data and GIS to process spatial data has been used to support water quality modelling as follows:

Develop Watershed Database: Multi-spectral satellite image data are processed to generate thematic maps. Data derived from various sensors can be integrated to improved maps with higher detail of cartographic information than an individual image. Two or more data layers can be overlayed to merge spatial information.

Watershed Delineation: In general, delineation of watershed and subcatchment is an essential step for most of watershed scale studies. Delineation is often achieved through digital elevation model (DEM) data that can be used to develop flow direction, flow accumulation, and pour point coordinates information.

Integrated Use of DEM: A DEM is a numerical representation of topography in a raster format and each cell is given a value of elevation. DEM data are very important to derive a variety of information including slope, aspect, curvature etc. DEM became a popular tool for land characterization due to its simple data structure and wide availability. In conjunction with satellite reflectance data, a DEM can also be used to simulate hydrological processes.

Land Use/Land Cover Change Detection: LULC play important roles in characterizing our environment at different scales-local, regional and global. Providing multi-temporal repetitive data to identify and quantify land surface changes, RS greatly enhances the capability of GIS in updating map information, which is very useful for decision makers and environmental managers.

With the aid of RS and GIS technologies, the watersheds can be subdivided into different categories of LULC and the contributions of each type of LULC to water quality can be monitored, analyzed and determined. In this context, RS imagery is used to obtain the necessary spatial coverage and classification schemes such as the normalized difference vegetation index (NDVI). The integrated analysis of landscape characteristics based on RS and GIS is a comprehensive tool to enable us to understand the complex environmental questions such as the effects of land use change on the biogeochemistry of riverine systems (Ballester et al. 2003).

Watershed Runoff Modelling: The advent of distributed runoff modelling requires different spatial parameters that are collected from each grid rather than the whole basin, which is time consuming, laborious and relatively expensive. Integrating a GIS with the model can make chores easier and often transparent to the user as well as make available the calculation and display of runoff flow depths across watershed sub-basins. In urban watersheds, the spatial analysis capabilities of a GIS can be used to analyze hydrological processes. Watershed attributes such as surface characteristics (pervious, impervious, slope, roughness), geometry and dimensions of flow planes, routing lengths, geometry and characteristics of routing segments, and soil information (infiltration rates, hydraulic conductivity, and storage capacities) can be efficiently stored in a GIS and utilized for urban runoff calculations (Mattikalli 2000).

Monitoring and Modelling Water Quality: RS data can be used to monitor water quality parameter because the presence of pollutants can affect the reflectance light at the visible/thermal bands from surface water in different ways. Hence, reflectance from certain bands has correlation with specific pollutants. RS imagery have been applied to estimate the values of several water quality parameters (e.g., chlorophyll-a, turbidity, TSS, Secchi depth) (Yin et al. 2005; Tyler et al. 2006). The monitoring of cyanobacterial blooms (Kutser et al. 2006), lake water quality characteristics, including chlorophyll and colored dissolved organic matter (Brezonik et al. 2005; Kutser et al. 2005), and contaminated stream water (Vignolo et al. 2006) was successful using Landsat-based RS.

Estimate Input Parameters for Water Quality Models: RS data from different spectral range provide specific information commonly used as water quality model inputs. For instance, surface temperature data can be produced from thermal-infrared imagery; soil moisture data can be produced from microwave reflectance. These two types of data can be combined to estimate evaporation and evapotranspiration rates. Using the data derived from RS spectral reflectance is particularly beneficial for some remote areas where on-site measurements are not applicable.

RS can provide data for global to regional monitoring, modelling and assessment such as VIs (e.g., NDVI, EVI). VIs have been successfully used to detect forest disturbance (Jin and Sader 2005), monitor vegetation dynamics (Beck et al. 2006), vegetation cover and condition (Fensholt 2004; Ben-Ze'Ev et al. 2006) and forage condition (Kawamura et al. 2005). A number of comparative studies have linked forest disturbance to the change of water quality (Prepas et al. 2001; Ensign and Mallin 2001; Swank et al. 2001). Also, the vegetation cover, such as in riparian areas, was proven to impact water quality in streams (Luke et al. 2007). The incorporation of VIs into model inputs for nutrient modelling improved the model’s performance (Li et al. 2008).

2.3.5 Integration of GIS, RS with Water Quality Models

The studies that combine the power of water quality modelling with the GIS to deal with spatial variations have been successful in environmental assessment and watershed management. The principal benefit of coupling GIS with environmental models is to enable the models to deal with large amount of spatial data that are essential to describe many environmental processes geographically. The primary GIS data provided to the models are the DEM, the LULC map, vegetation cover, and the soil map. Here, GIS can provide modelers with new platforms for data management and visualization. Data are preprocessed into a form suitable for analysis which includes data import, scale, coordinate transformation, data conversion, data structure and spatial analysis of input data. In addition to standard functions (e.g., coordinate transformation, conversion between spatial formats, raster algebra, vector operations, network analysis, display and visualization) GIS supports modelling directly so that some tasks such as analysis, calibration and prediction can be carried out by using GIS itself. The postprocessing mostly analyzes the spatial information, and visualizes exports data through reformatting, tabulation, and mapping. To date, no GIS system has the spatial and temporal data representation flexibility as well as the algorithmic capability needed to construct process-based models internally; consequently, environmental models and GIS will be coupled (Loague and Corwin 2000).

Another reason that drives the integration of water models with GIS is the need to provide a scientific land management and decision-making tool by incorporating the simulation models. Also, it is becoming recognized that land management and decision-making require integrated assessment of environmental (air, land, water, and the interactions and fluxes between these), ecological, social and economic systems. Consequently, the multiple uses of resources and multiple goals requires integrating information and analysis to further decision-making, which can be addressed using the data management facilities of GIS (Aspinall and Pearson 2000).

2.3.5.1 Coupling GIS and Simulation Models Theory

Simulation models are very useful tools to analyze watershed processes, and develop and assess watershed management scenarios. Implementation of these models often requires the integration of GIS, RS and multiple databases for determination of the model input parameters and for analysis and visualization of the simulation results (He 2003). The mechanisms and approaches to link GIS and water quality models have received much attention in both geographic information sciences and simulation modelling. Three of the most common strategies for linking a GIS to a simulation model are loose coupling, tight coupling and an embedded system approach (Corwin et al. 1997) (Figure 2-2).

In a loose coupling, data are stored in one system and transferred to another and subsequently read by the other system. This approach usually involves a standard GIS package (e.g., Arc/Info) and hydrological modelling programs (e.g., HEC-1, HEC-2, STORM) or a statistical package (e.g., SAS or SPSS) (Sui and Maggio 1999). The important characteristic of the loose coupling is that GIS and environmental models are implemented separately without common interface. The simulation models and GIS are integrated via data exchange using either ASCII or binary data format. A majority of GIS applications for modelling represent this approach because it requires little software modification.

In a tight coupling, the simulation model and GIS share the same database and the data management is integrated into the system. Characteristically, a tight coupling provides a common user interface for both the GIS and the model i.e. the information sharing between the respective components is transparent.

The model becomes embedded models as the degree of coupling between the GIS and the model increases to the point where the model's functions are essentially part of the built-in functionality of the GIS. In embedded systems, the coupling of software components occurs within a single application with shared memory rather than sharing the database (loose coupling) and a common interface (tight coupling). Thus, embedded systems require a substantial amount of time to develop and may be difficult to modify when changes are needed. Also, the modelling capabilities are usually simplistic and calibrations must take place outside of the package. These models tend not to be industry standard (Sui and Maggio 1999).

The types of coupling stated above only represent how GIS and environmental models can share the same data technically. This type of coupling does not represent the integration in terms of achieving compatible views of the world and in turn has not necessarily improved in the scientific foundation of either GIS or environmental modelling. The main difficulty for the integrated use of GIS and environmental simulation models is their different data models because GIS is a static representation and simulation models is concerned with dynamic processes (Hellweger and Maidment 1999; Aspinall and Pearson 2000). This means that their data models will be quite different and result in different database structures. It should be the databases where GIS and environmental simulation models are related to each other (Brimicombe 2003). In the future, these databases will be increasingly networked and so GIS and environmental modelling will be more greatly integrated and interoperative.

The successful use of GIS bridges the gaps between scientific research modelling and management applications and encourages the conceptual developments about the ways of utilizing spatial data in environmental modelling (Goodchild et al. 1996). On the negative side, the tremendous success of modelling using GIS has resulted in the problems of framework inflexibility, in that GIS are often treated as the technical and conceptual framework into which environmental modelling must fit, rather than it provides of services (primarily spatial data analysis and management) to environmental modelling (Argent 2004).

Apart from GIS-centric view of environmental model integration, the structural context of the environmental modelling situation and an appropriate higher level of data model for framing the problem were investigated (Livingstone and Raper 1994; Raper and Livingstone 1995). Similarly, as an alternative to describing linkage between GIS and environmental models based on the database, the practical implementation and operation of GIS and environmental modelling from a programming and interface perspective should also be considered (Aspinall and Pearson 2000). The tools for analysis can be provided in menus or in additional programs accessed through menus added to the GIS interface. Through linking GIS and environmental models in the manner of a toolbox with the environmental models being developed and implemented as additional tools or suites of tools for the GIS (Aspinall and Pearson 2000). The two different views of integration of GIS and environmental models is illustrated in Figure 2-3 (Argent 2004). In one view, the model is built into or accessed from a GIS, whereas in the other view, the GIS is a spatial data collection and other services are accessed from the environmental model.

2.3.5.2 Applications of Integrating GIS, RS and Water Quality Models

Reviews on the integration of GIS with hydrological modelling and environmental modelling can be found in literature (Lam and Pupp 1996; Sui and Maggio 1999; Yang et al. 1999; He 2003; Argent 2004). The integration of GIS with hydrological modelling is applied for abundant empirical studies in various regions, most of which rely on a combination of loose and tight coupling (Hao et al. 2003; de Jong van Lier et al. 2005). The integration of GIS with hydrological modelling enables GIS users to go beyond data management and thematic mapping to conduct sophisticated analysis and simulation for scientific research and policy analysis. GIS provided hydrologists and hydraulic engineers with the ideal computing platform for data inventory, parameter estimation, mapping and visualizing the modelling results. Thus, GIS greatly facilitates the design, calibration, and implementation of hydrological/hydraulic modelling. However, the loose coupling of hydrological modelling and GIS does not improve the scientific foundation of these models (Sui 1999). It is highly recommended that the integration of GIS with hydrological modelling should involve the development of a high level common ontology that is compatible with both GIS and hydrological models (Sui and Maggio 1999). The common ontological framework should incorporate multi-dimensional concepts of space, time and scale.

The interfaces between GIS and environmental models often carry out data processing and visualization, tool coordination, and improve the applicability of models. These interfaces are also user friendly. The integration of GIS and the distributed continuous time, non-point source pollution model soil and water assessment tool (SWAT) was effective and efficient for data collection and to visualize and analyze the input and output of simulation models (Arnold et al. 1999). GIS has been used by combination with SWAT for many cases (Di Luzio et al. 2005; Qi and Grunwald 2005; Santhi et al. 2005; Grunwald and Qi 2006; Olivera et al. 2006). The main motivation of integrating GIS with water quality models at watershed scale is to discover and evaluate the crucial areas within a watershed so that the land use and management practices can be optimized to control the loading of nutrients and sediment into water bodies (Grunwald and Qi 2006).

In addition to the powerful and comprehensive models such as SWAT and BASINS, a variety of other models have been integrated with GIS and formed GIS-based models, which have been used for water quality assessment upon human activities (Xu et al. 2001; Giupponi and Vladimirova 2006), water contamination risk assessment (Sinkevich et al. 2005), identification of contamination sources (Terrado et al. 2006), and water quality modelling (Fernandez et al. 2002; Vivoni and Richards 2005). Here, GIS is to incorporate spatial data as water quality is often affected by spatially distributed natural or human factors.

The use of multi-temporal remote sensing images in support of environmental modelling analysis in a GIS environment contributes to identify a variety of long-term interactions between resources and land use, and the built environment has been a highly successful approach in recent years (Ning et al. 2006). The spatial and time variation of the nitrate in the basin of a small river was simulated by combining nitrate dynamic modelling and GIS (Matejicek et al. 2003), where the NDVI is used to estimate the level of denitrification.

2.4 N Export to Surface Water in Forested Watersheds

To determine ANN N model input variables, it is important to understand the export of N into streams in forested watersheds.

2.4.1 Sources of N in Streams

The N is undergoing a complex cycle and its leaching into streams is affected by many interacted factors (Figure 2-4). The two primary theories developed to predict long-term changes in nitrate loss rates are the N saturation theory (Aber et al. 1989; Aber et al. 1997; Currie et al. 1999; Venterea et al. 2004) and the nutrient retention hypothesis (Vitousek and Reiners 1975; Vitousek and Matson 1984). The common emphasis shared by these to theories is the interaction between plant nutrient demand and soil nutrient supply over time (Aber et al.

2002). The N saturation theory suggests that cumulative N deposition adds to soil N pools and rates of N mineralization, eventually leading to excess availability, induced or increased nitrification, and elevated nitrate leaching. The nutrient retention hypothesis suggests that both the accumulation of limiting nutrients through deposition and reductions in net plant demand due to decreasing biomass accretion leads to the alleviation of nutrient limitations and increasing leaching loss rates. The most important participating factors - plant nutrient demand and soil nutrient supply - are seasonally variable, and thus export of N to surface water display a seasonal cycle.

Inter-annual variability in the timing and depth of snowpack accumulation may explain the year to year variability in inorganic N concentrations in surface water in these ecosystems (Brooks et al. 1998). The winter snow pack is the major control both on hydrologic N export and on soil source/sink relationships for N concurrent with this transport mechanism (Brooks et al. 1999). The effect of winter snow cover on the fate of both atmospheric and soil N needs to be considered when evaluating the potential effects of increased N deposition on either terrestrial or aquatic ecosystems in seasonally snow-covered watersheds. Sometimes, nitrification in the forest floor-mineral soil contributes more than the atmospheric NO_3^- in snowpack to the NO_3^- pulses in streams during later winter/early spring (Piatek et al. 2005).

Generally, the major sources of N leaching into streams after winter break are snowpacks and soils. Inorganic N released from winter snowpacks provides a large pulse of mobile, potentially available N each spring. The soil storage capacity is responsible for the variability of NO_3^- at different sites, while the potential snowmelt and the flow paths account for the variation of NO_3^- from year to year within a site. In forest ecosystem, carbon and N cycles interact at several points between plants and microorganisms. The degree to which nitrification occurs depends upon the competition of nitrifier with the strength of plant demand on N. Nitrate leaching is proportional to the amount of NO_3^- remaining in the soil after plant uptake and water drainage rates (Aber et al. 1997).

2.4.2 The Effect of Snow Cover on Microbial N Transformation and N leachate

Traditionally, N mineralization in soil is assumed to take place within a temperature range of 5 to 35 °C. On the other hand, it is also found that microbial activities exist at low temperatures, even at sub-zero °C (Clein and Schimel 1995; Rankinen et al. 2004). Although the bulk of soil water freezes just below 0 °C, there still exist liquid water films surrounding soil particles down to temperature at least or below -10°C (Romanovsky and Osterkamp 2000), and the presence of unfrozen water in soil allows microbial activity to continue (Rivkina et al. 1998; Mikan et al. 2002; Rivkina et al. 2004). The accumulation of snow and snow depth play an important role on controlling soil temperature and ensuring the existence of unfrozen water.

The presence of snow is imperative for microbial processes in winter with very low atmospheric temperature. Snow is an effective insulator and thus the depth of snow regulates soil temperature. In the case of snow-covered soil, the soil temperature is typically only a few degrees below zero even though the surrounding air temperature is much lower. The reduction of snowpack accumulation will induce soil freezing (Groffman et al. 2006), consequently regulate soil biogeochemical processes and solute delivery to streams in forested watershed (Groffman et al. 1999, 2006).

A field incubation study carried out in a sub-arctic region observed mineralization and immobilization of N in soil during winter (Schmidt et al. 1999). Other studies even found that 40% of net N mineralization in boreal forest soil occurred during the winter (Stottlemyer and Toczydlowski 1999; Kielland et al. 2006). During dormant season, positive nitrification was observed in soils from willow and white spruce stands prior to the mid-winter (January) sampling, whereas significant nitrification was found in spring (May) in soil from stands of alder, poplar, and particularly black spruce (Kielland et al. 2006). It was surmised that the first flush of net N mineralization occurs in spring due to soil thaw and then the net N release declines as plants start competing with soil microbes in early May-June. The second flush of net N release in August is attributed principally to root mortality, which is fairly synchronous across successional stages (Ruess et al. 2003).

The effect of decreases in snowpack accumulation on microbial biomass and activity was quantified and the results showed that a relatively mild freezing event induced significant increases in N mineralization and nitrification rates, solute leaching and soil nitrous oxide production (Groffman et al. 1999). Low soil temperatures limit soil N mineralization under ambient snow conditions, but the deeper snow conditions with the associated warmer winter soil temperature dramatically increase over-winter N mineralization (Schimel et al. 2004). Mild soil freezing (temperature > -5 °C) increased soil NO_3^- concentration by physical disruption of the soil ecosystem rather than by direct stimulation of mineralization and nitrification (Groffman et al. 2001). A study on how snow cover controls subnivial (below snowpack) microbial processes and N leachate from the snow-soil interface to surface waters in high-elevation catchments indicates that a portion of the spatial and temporal variability in N export from these seasonally snow-covered systems was attributed to variability in winter snow cover across landscape types and inter-annually within a landscape type (Brooks and Williams 1999). Soils remain frozen and there is little microbial activity and N leachate is high in shallow-short duration snowpacks; total microbial activity is highly variable and the amount of N leachate is highly variable in high interannual variability in snow depth and duration; total microbial activity is high and there is little N leachate in continuous snow cover; and microbial activity is reduced because of carbon limitation and N leachate is high in deep long-duration snow cover verging on perennial snowpacks (Brooks and Williams 1999).

Mineralization of organic matter was the dominant source of soil inorganic N before and during the spring thaw. Nitrification and denitrification does occur under snow cover. In a word, studies indicate that microbial activity under seasonal snowpacks plays an important role in controlling N export in surface water.

2.4.3 Seasonal N Concentrations in Stream Water

The accumulation of N in snowpacks and the existence of nitrification during winter obviously lead us to think that snowmelt will result in leaching of N from these sources and consequently increase of N in streams. In fact, there is evidence that inorganic N concentration peaks occur in later winter/early spring. The snow pack is able to store large amounts of different substances, which are released during the first snow melt. The observed high N concentrations in late winter/early spring often resulted from the release of N from the snow pack and/or soil (Arheimer et al. 1996; Williams et al. 1996b; Williams et al. 2001; Rankinen et al. 2004). In comparison, inorganic N concentrations are often lower in summer than in the dormant seasons in non-polluted and undisturbed northern rivers (Williams et al. 1996a; Williams et al. 2001; Kaste and Skjelkvale 2002). The great difference of N concentrations between summer and winter indicates that soil water contains higher N in the dormant season than in the growing season, when inorganic N is usually retained effectively (Rankinen et al. 2004). The study of N leaching in northern latitudes indicated that NO_3^- concentration decreased during the growing season to an almost negligible level, and NH_4^+ concentration had the similar seasonal pattern except that the concentration levels were even lower (Rankinen et al. 2004). The data series from ten years of measurements in 20 small Swedish and Finnish catchments showed that enhanced inorganic-N concentrations normally occurred during the spring, and low concentrations occurred during the growing season (June-August) (Arheimer et al. 1996).

2.4.4 N Modelling

Modelling of N forms in streams is very complex due to the difficulty of mathematically representing factors like the land use and land cover, soil and vegetation N dynamics, in-stream nitrification/denitrification, and meteorological parameters. Also, hydrological, biological and chemical processes underlying N storage, transformation and release are not well understood in these environments. The concentration of N fluctuates with significant seasonal cycles, which should be captured by effective models. The largest NO_3^- fluxes in streams from forested watershed occur with large runoff events during early spring snowmelt when vegetation and microbial uptake of inorganic nitrogen is low.

ANNs have been mostly used in agricultural watersheds to simulate nitrate leaching in agricultural drainage effluent (Kaluli et al. 1998), forecast nitrate loads on an agricultural watersheds based on historical data (Yu et al. 2004), simulate nitrate-N concentrations in drainage water after the use of manure and/or fertilizer (Sharma et al. 2003), and predict total and inorganic nitrogen concentrations in streams from watershed features (Lek et al. 1999). A GRNN model to predict the nutrient loading into the neighboring water from large-plotted paddy rice fields with hydro-meteorological factors and nutrient contents from water sources tell that both environmental inputs (i.e., nutrient contents in irrigation streams) and hydrological processes (i.e., rainfall, surface discharge) have significant impact on N leaching into water body (Kim et al. 2007). These ANN N models have considered only nitrogen load from external sources, nitrogen leaching and seasonal/annual cycles, but not vegetation dynamics, one of the most important factors impacting nitrogen cycle. Recently, the rapidly developing RS technology and the reduced cost of requiring remotely sensed VIs data has made this possible. Li et al. (2008) successfully applied GRNN to predict nitrate, ammonium and total dissolved nitrogen in streams within undisturbed and disturbed watersheds using weather parameters, time index reflecting seasonal cycles and EVI as model inputs. The inclusion of VIs reflecting the watershed disturbance and vegetation dynamics highly improved the accuracy of model prediction (Li et al. 2008).

The mineralization and nitrification happening under snow cover in winter should be considered when modelling N in forest watersheds. A number of studies have observed microbial activities under snow cover at the temperature of below 0 °C (Rivkina et al. 2004; Panikova et al. 2006). From the 1990s, it has been recognized that overwinter and snowmelt processes play important role on controlling N cycling and retention. Recent research indicates that the largest N fluxes from forested watersheds (sharply increased N concentrations in streams) occur with large runoff events, in particular, during early spring snowmelt when vegetation and microbial uptake of inorganic N is low (Aber et al. 2003). The possible sources of N in streamwater in snowmelt come from atmospheric N, snow, mineralization in soils under the snowpack, groundwater during early phases of the melt, premelt stored water and nitrification, and a combination of these factors (Piatek et al. 2005). Besides, other factors including the presence of wetlands and vegetation types are also important in controlling N generation and loss rates (Campbell et al. 2002; Ito et al. 2005).

The ultimate goal of water quality modelling is to support watershed management and decision making on watershed activities. The success of ANNs modelling highlight that the ANNs models can provide real time predictions, can answer questions related to the impact of climate change and watershed activities on water quality, and can be implemented in water resources management. So far, most of the research uses current meteorological information and watershed properties to predict current or recent (within several days) N concentrations. More recent research indicates that incorporation of information on vegetation dynamics and vegetation cover obtained from RS data has improved model performance. To apply the developed modelling tools to watershed management, further studies on forecast N weekly, monthly or yearly in advance are needed. This can be achieved in two ways: (1) use previous information (i.e., weather information at t-1…t-10) to predict the modelled parameter at time t; (2) the model inputs at time t are first predicted from other sources and then incorporated into ANNs models to forecast the interested water quality parameter at time t. This is a promising method as most of the ANNs water quality models include weather information that is predictable using RS.

2.5 Summary and Recommendations

Over the past decade, ANNs have been proven through a number of case studies to be a reliable tool capable of modelling a variety of water quality parameters. The performance of ANNs is evaluated not only by their own absolute performance but also by comparison with traditional MLR models and numerical models based on a series of model evaluation criteria. In most studies, ANN models outperform traditional MLR probably because the water quality parameters being predicted are non-linearly interacted with model input variables in a complex non-linear way, which is already beyond the scope of MLR. Compared with numerical models, ANNs can provide comparable or even better modelling results in some cases as the capability of numerical models may be a trade off due to the our limited knowledge in certain situations and the non-availability of reliable data for model calibration. Not limited to the most popular MLP-BP, a number of comparison studies are carried out to search for the best ANN for certain problems recognizing the advantage and limitations of different networks such as RBF, AMN, GRNN and MNN. Generally, MNN have higher prediction accuracy than a single ANN for complex problems; GRNN may work better for time series parameters than MLP-BP; and AMN has strengths in providing more explicit information about the system being modelled. However, there have no criteria for selecting a suitable or optimal ANN algorithm for different types of problems.

The popularity and successful applications of ANNs, at the same time, revealed the modelers' concerns and encouraged advanced studies to make ANNs a better modelling tool. Great efforts have been put into (1) developing hybrid ANN models, (2) extracting knowledge from trained ANN models, (3) improving model generalization ability, and (4) developing guidelines of ANN model development.

GIS and RS technologies have become important components for water quality modelling and watershed management. The use of GIS ranges from just display and visualization of results to storage and retrieval of remotely sensed data and environmental data; from spatial analysis of landscapes to integration with environmental modelling. RS is one of the most successful GIS-related technologies and RS data can be best utilized if they are integrated into a GIS that is designated to manage large volumes of data. Regarding water quality modelling, RS data are mainly used to: (1) estimate the input parameters for both lumped and distributed watershed models, (2) delineate watershed and streams, (3) classify land-use and land-application, (4) monitor water quality parameters (e.g., particles, turbidity, chlorophyll), (5) estimate vegetation indices and leaf area index to study vegetation dynamics, and (6) estimate meteorological parameter (e.g., surface temperature, snow) and watershed properties (e.g., soil moisture).

To integrate GIS with environmental models, in one view, the model is built into or accessed from a GIS, whereas in another view, the GIS is a spatial data collection and other services are accessed from the environmental model. From data exchange perspective, three of the most common strategies for linking a GIS to a simulation model are loose coupling, tight coupling and an embedded system approach. The research on integration of GIS and environmental models, either from conceptual perspectives or case studies, is basically to combine GIS's power in data process and management and environmental model's ability to simulate the studied dynamic system. Ultimately, the integration is to serve decision-making for watershed management.

A number of field and lab studies have confirmed the active N processes under snow cover, and demonstrated that over-winter N processes are important and must be considered when modelling N leaching and its concentration in surface water. Snow pack insulates soil from the very low atmospheric temperature in winter and thus enables microbial activities. Consequently, this leads to increase of N in the soil under certain conditions of snow accumulation, snow depth and snow consistence. The major sources of N in streams after winter break are the snowpacks and the soils. The relative contribution of each source is determined by the combined effects of watershed characteristics, air pollution, vegetation, climate condition and disturbances. It is commonly recognized that inorganic N concentrations peak in surface water during snowmelt or winter breaks. Therefore, when simulating N concentrations in surface water using ANNs, it is important to include model inputs like stream flow, vegetation cover and time index that can represent this seasonality.

To develop desirable ANN models for monitoring and forecasting water quality parameter and further to assist in decision making for watershed management, further studies are recommended:

1. Most studies use current meteorological information and watershed properties to predict current or recent (within several days) water quality parameters. However, to apply ANNs models to watershed management, further studies need to focus on forecasting water quality probably weekly or monthly in advance to give the decision-maker enough response time.
2. Support decision making: Water quality modelling is a crucial component of environmental decision support system. Most studies on ANN modelling of water quality suggest that ANN models can be employed to support the decision making for watershed management and environmental issues. So far, there are few studies on incorporating ANN models into decision making systems. Therefore, further studies need to be conducted on employing ANN water quality models to direct decision making.
3. Studies on water quality modelling need to consider incorporation of RS information, which can be used to monitor watershed properties and provide significant input variables for water quality models.
4. N modelling using ANNs needs to include VIs as model inputs as VIs can describe vegetation dynamics and vegetation cover, which significantly affecting N retention in soil and its leaching into surface water.

Table 2-1. Comparison of different ANNs applied to water quality modelling.

ANN algorithm	Description	Advantages	Limitations	References
MLP-BP	− Has an input layer, one or more hidden layers and an output layer (Figure 2-1a) − Each hidden layer neuron receives the weighted output of the previous neurons. At each neuron, the weighted input signals are summed and a threshold value is added − An activation function is performed on the combined input to produce the output of the node	− Work well on a wide range of problems	− Calibration is tedious because many internal parameters to be determined	(ASCE Committee 2000)
RNN	− Consist of both feed-forward and feed-backward connections between the layers and neurons as the inputs to the neurons come from both external input and internal neurons (Figure 2-1b).	− Suitable to model dynamic relationships	− Many internal parameters to be determined − The data should be in time-series with the same gap.	(Mandic 2001; Chiang et al. 2004)
RBFN	− Consists of three layers (Figure 2-1c). − The individual hidden neurons compute their activation using a radial basis function, typically the Gaussian kernel function − The output units simply sum the weighted activations of individual hidden neurons.	− Can be trained much faster − Can be easily optimized because the number of hidden neurons the only parameter need to be adjusted	− May only work well on certain problems	(Meireles et al. 2003)
GRNN	− Consists of an input layer, a pattern layer, a summation layer and an output layer (Figure 2-1d) − The output of GRNN is a simple division of the results from the summation unit by the result from the division unit	− Converge to the optimal regression surface and deal with sparse data efficiently − Can be easily optimized − Does not provide wild estimation.	− The estimation is limited by the minimum and maximum boundary	(Specht 1991)

Table 2-1. ***Continued***

ANN algorithm	Description	Advantages	Limitations	References
MNN	− Consist of an input layer, a processing layer comprising the expert networks and the gating network, and an output layer (Figure 2-1e). − To develop a MNN model, the whole task is divided into several simpler substasks, then these subtasks are solved through expert networks and finally the subsolutions are combined through a gating network to produce the desired solution of the original complex task.	− Suitable for complex problems	− Many internal parameters to be determined	(Tsoukalas and Uhrig 1997)
KSOM	− Consists of two layers − In competitive learning, the neurons of the networks can recognize groups of similar input vectors (Figure 2-1f).	− Able to order multivariate data and preserve the topological structure of the data Suitable for classification	− Cannot work on continuous functions	(Kohonen 1982)
AMN	− The input space of AMNs is normalized by a p-dimensional lattice, and each cell of the lattice represents similar regions of the input space (Figure 2-1g). − The AMN has only one hidden layer, which consists of basis functions that are defined on the lattice formed by normalizing the input space.	− Provide more explicit information about the relationship between the inputs and the outputs	− Not suitable for problems with a large number of inputs	(Brown and Harris 1994), (Brown and Harris 1995)

Table 2-2. Summary of water quality modelling using ANNs.

#	Modelled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparison	Reference
1	Sediment (mg/L)	A variety of locations	MLP	Prior knowledge	Tractive shear stress Velocity ratio Suspension parameter Longitudinal slope Water depth ratio	0	T:V=1:1 random	% of data in range, plot2	other conventional models	(Nagy et al. 2002)
2	Sediment (mg/L)	Tongue River in Montana, USA	GRNN, RBF, MLP	Correlation analysis	Lagged Q and sediment	0	Time range*	MRSE, MAE, plot1, and plot2	MLP RBF, GRNN and MLR	(Kisi 2004)
3	Sediment (mg/L)	Tongue River in Montana, USA	Neuro-Fuzzy	Prior knowledge, trial-and-error	Lagged Q and sediment	0	Time range *	MRSE, R^2	MLP, sediment rating curve, and MLR	(Kisi 2005)
4	Sediment (mg/L)	Banha watershed in the Upper Damodar Valley, Jharkhand state, India	MLP	Prior knowledge	Geomorphol ogical parameters and runoff rate	0	T:S:V = 3: 2: 4, N/A	E, MRSE, R^2, AAD, Plot1	regression models	(Sarangi and Bhattacharya 2005)
5	Sediment (mg/L)	St. Esprit watershed, Quebec, Canada.	MLP	MARS	Geomorphol ogical parameters and runoff rate	0	T:S:V = 2: 1: 1, N/A	MRSE, R^2, plot1, and plot2	regression equations	(Sarangi et al. 2005)
6	Sediment, (tons/d)	Schuylkill river, Philadelphia. USA	MLP	Correlation analysis	Lagged upstream and downstream sediment	1 day	Time range *	MSE, R^2, Plot1, Plot2	MLR, sediment rating curve	(Cigizoglu 2004)

Table 2-2. ***Continued***

#	Modeled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparisons	Reference
7	Sediment (tons/d)	Juniata River of Pennslyvania	FFBP and GRNN	Correlation analysis	Lagged Q and -1d sediment	0	N/A	MSE, R^2, Plot1, Plot2	MLR, sediment rating curve	(Cigizoglu and Alp 2006)
8	Sediment (tons/day)	Schuylkill River. USA	FFBP-RDNN	Correlation analysis	Lagged Q and –d sediment	0	Time range *	MAE, R2, Plot1, Plot2	MLP	(Cigizoglu and Kisi 2006)
9	sediment (kg/s)	Vamsadhara River basin of India	MLP	Prior knowledge, trend analysis	R and Q	0	Time range *	MRSE, R^2, E	linear regression	(Agarwal et al. 2005)
10	Sediment (tons/d)	Juniata Catchment, USA	FFBP, RBF	Correlation analysis	Lagged R and Q	0	Time range *	MSE, R^2, Plot1, Plot2	FFBP and RBF	(Alp and Cigizoglu 2007)
11	Sediment (tons/ha)	Upper Siwane River, India	MLP	Prior knowledge, trial-error	Lagged R, T	0	Time range *	RMSE, R^2	LR	(Raghuwanshi et al. 2006)
12	Sediment (mg/L)	N/A	MLP	Correlation analysis, trial-error	Lagged R	0	Time range *	MAE, R^2, Plot1, Plot2	two-dimensional unit sediment graph theory	(Tayfur and Guldal 2006)
13	Sediment (mg/L)	watersheds in the Canadian Boreal Plain	MLP	Correlation analysis	Lagged Q and R, T, ddt, snowfall, periodicity index	0	KNN	MRSE, R^2, Plot1, Plot2	N/A	(Nour et al. 2006a)

Table 2-2. ***Continued***

#	Modeled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparison	Reference
14	Sediment (kg/s)	Longchuanjiang River, China	MLP	Correlation analysis	Lagged R and T	0	T:S:V=16:5:15 Time range *	MRSE, R^2, Plot1, Plot2	MLR	(Zhu et al. 2007)
15	Sediment (mg/L)	Mississippi River and Conococheague Creek, USA	Compound MLP	Correlation analysis, trial-error	Lagged Q and S, river stage	0	Time range *	R^2, SSE, Plot1	compound rating curve	(Jain 2008)
16	Sediment flow rate (N/s)	W-2 of Treynor Catchment and W7 of Goodwin Creek watershed, USA	MLP	Correlation analysis, trial-error	Lagged R, Sediment flow rate, and runoff	0	T:V=7:3	RMSE, R^2, E, plot1	the linear transfer function model	(Rai and Mathur 2008)
17	Sediment (tons/year)	Tigris River, Turkey	MLP	Correlation analysis	R, T, Q	0	T:S:V=165:82:82	Plot1, plot2, %error	Linear regression	(Hamidi and Kayaalp 2008)
18	Sediment (mg/L)	Quebrada Blanca and Rio Valenciano Station, USA	MLP	Correlation analysis	Lagged Q and sediment	0	Time range *	R^2, plot1, plot2	Three training algorithms	(Kisi 2008)
19	Nitrate leaching in agriculture drainage effluent (mg/L)	Soulanges, Quebec	MLP	Prior knowledge	Julian day, denitrification, cropping system, water table depth, N application, precipitation, N con. in drainage, drain flow	0	Time range *	R^2, Plot1, plot2	N/A	(Kaluli et al. 1998)

Table 2-2. ***Continued***

#	Modeled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparison	Reference
20	Nitrate loss to drain flow from agricultural farm (Kg N/ha.)	Eugene F. Whelan Experimental Farm Woodslee, Ontario, Canada	MLP	Prior knowledge	Julian day, year, T, soil condition, evapotranspiration, snow melt, drain	0	random	RMSE, plot2	N/A	(Salehi et al. 2000)
21	Nitrate in drainage of agriculture watershed (mg/L)	Greenbelt Research Farm of Agriculture Canada, Ottawa, Ont.	MLP	Prior knowledge	Julian day, treatment, nitrogen applied, R, snow fall, T	0	N/A	R^2, E	RBF	(Sharma et al. 2003)
22	Nitrate (mg/L)	Sangamon River near Decatur, Illinois,	MLP	Correlation analysis, trial-error	Lagged R, T, N concentration and streamflow	1 week	Time range *	RMSE, plot1	linear regression	(Markus et al. 2003)
23	Nitrate in river (mg/L)	Upper Sangamon River in Illinois	MLP and RBFNNs	Prior knowledge	R, T, Streamflow, Julian day	0	Time range *	RMSE, plot2	traditional regression model and SWAT	(Suen and Eheart 2003)
24	nitrate in stream (mg/L)	Vermilion River in Illinois	MLP	Prior knowledge	Q, R, nitrate load	0	Time range *	RMSE, R^2, E	N/A	(Yu et al. 2004)
25	Nitrate and total dissolved nitrogen (ug/L)	Watersheds in Canadian Boreal Plain	GRNN	Correlation analysis, trial-error	Lagged ddt, R, T, EVI, Julian day	0	T:S:V=3:1:1, KNN	RMSE, R^2, Plot1	N/A	(Li et al. 2008)
26	Phosphorus (mg/L)	watershed in the Canadian Boreal Plain	MLP	Correlation analysis	Q, T, TSS, ddt, snowmelt, periodicity index	0	KNN	RMSE, R^2, plot1, plot2	N/A	(Nour et al. 2006a)

Table 2-2. ***Continued***

#	Modeled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparison against	Reference
27	Phosphorus (mg/L)	watershed in the Canadian Boreal Plain	MLP	Correlation analysis	Lagged Q and P, T, periodicity index	0	T:S:V=3:1:1	RMSE, R^2, plot1, plot2	ARMAX	(Nour et al. 2006b)
28	Phosphorus (mg/L)	Incheon area	MNN	Prior knowledge, trial-error	Lagged R, surface discharge	0	T:S=4:1	RMSE, R^2	GRNN	(Kim and Kim 2007)
29	Phosphorus (mg/L)	Odra River, Poland	MLP	Prior knowledge	(1)Previous P or (2) other water quality parameters: T, pH, TKN, BOD, COD etc.	0	T:S:V=54%:23%:23%, random	R^2,MAE, Error S.D., S.D. Ratio	GRNN, MLR, RBF	(Mozejko and Gniot 2008)
30	Cynobacteria (*Anabaena. spp.*) (cells/mL)	River Murray at Morgan, Australia	MLP	Sensitivity analysis, trial-error	Lagged Q, turbidity, color, T, , total P, soluble P, total iron, and oxidize N	4-week	Time range *	RMSE, plot1	N/A	(Maier et al. 1998)
31	Cynobacteria (*Anabaena. spp.*) (cells/mL)	River Murray at Morgan, Australia	B-spline AMN	Prior knowledge, trial-error	Lagged Q and T	4-week	Time range *	Plot1	MLP-BP	(Maier et al. 2000)
32	Cynobacteria (*Anabaena. spp.*) (cells/mL)	River Murray at Morgan, Australia	B-spline AMN	Prior knowledge, trial-error	Lagged Q and T	4-week	Temporally *	Plot1	MLP-BP	(Maier et al. 2001)

Table 2-2. ***Continued***

#	Modeled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparison against	Reference
33	Peak *Cryptosporidium*	Delaware River, USA	MLP	Correlation analysis	Turbidity, pH, TC, FC, R, Q, *Clostridium perfringens*, *Escherichia Coli*	0	T:V=51:17	Plot2, %error	N/A	(Brion et al. 2001)
34	Classification of sources and ages of fecal contamination	Inlet to Town Branch WWTP of Lexington, KY, USA	MLP	ANOVA	R, total, atypical and background colonies, FC, fecal streptococci	0	T:V=2:1	%error	N/A	(Brion et al. 2002)
35	*Cryptosporidium* and *Giardia* (cells/mL)	Delaware River, USA	MLP	trial-error	pH, Q, R, *C. perfringens*, *E. coli*, FC, turbidity, TC, alkalinity	0	T:V=51:17, random	%error	N/A	(Neelakantan et al. 2001)
36	*Cryptosporidium* and *Giardia*(cells/mL)	Delaware River, USA	MLP	Prior knowledge	pH, Q, R, *C. perfringens*, *E. coli*, FC, turbidity, TC, alkalinity, suspended solids, dissolved solids, coliphage, F-specific coliphage, rcoliphage : F-specific coliphage	0	T:V=51:17, random	%error	Compared different training methods	(Neelakantan et al. 2002)

Table 2-2. ***Continued***

#	Modeled parameter	Location	ANNs	Input determination	Input parameters	Forecast length	Data division	Evaluation	Comparison against	Reference
37	*E. coli* (cfu/mL)	the Agricultural Research Center, Univ. of Ariz. AZ	MLP	Prior knowledge	pH, turbidity, conductivity	0	T:S:V=72%:20%:8%	Plot1, %accuracy	GRNN	(Kim et al. 2008)
38	Cynobacteria bloom	Crestuma Reservoir, Portugal	GRNN	Cluster analysis, sensitivity analysis	COND Fe NO_3^- TURB O taxa pH DISCH CHLOR Tmax DO	Two-week	Time range	Statistical analysis (mean, error S.D., RMSE, R^2, etc.)	NA	(Teles et al. 2006)
39	*salinity*	River Murray at river bridge, Australia	MLP	Prior knowledge	Lagged Salinity at other locations and lagged Q	14 days	T:V=1469:365	RMSE, %error	compared six training rules	(Maier and Dandy 1999)
40	*salinity*	Same as the above	MLP	PMI	Lagged Salinity at other locations, and lagged Q and river level	14 days	Time range *	RMSE, AIC, Plot1	Compared training algorithms	(Kingston et al. 2005)
41	*Salinity*	Same as the above	MLP	Prior knowledge	Lagged Salinity at other locations, and lagged Q and river level	14 days	Time range *, GA, KNN	RMSE, AIC, Plot1	Presented optimal data division GA-SOM	(Bowden et al. 2002)
42	*Salinity*	Same as the above	MLP	PMI, KNN-GAGRNN	Lagged Salinity at other locations, and lagged Q and river level	14 days	T:S:V=3.2:0.8:1, GA	RMSE, Plot1	Compared input determination methods	(Bowden et al. 2005a; Bowden et al. 2005b)

Note:
1. the ANN models in all cases were trained with BP.
2. In "data division" column, "time range" means: a continuous set of data were chosen from the whole data set for calibration and the other data left are used for validation.

3. Plot1 is a time series plot of measured and predicted values.
4. Plot2 is a plot of measured vs. predicted.
5. AAD: Absolute average deviation
6. MARS: Multivariate adaptive regression spline
7. S.D.: Standard deviation
8. RMSE: Root mean square error
9. AIC: Akaike's information criterion
10. TC is total coliform and FC is fecal coliform.

Table 2-3. Summary of ANN model inputs for water quality modelling.

Variables modelled	Examples of possible Inputs
Sediment	Meteorological data: rainfall, precipitation Stream data: (time-lagged) streamflow, antecedent sediment data, runoff Morphological parameters: bifurcation ratio, area ratio, channel length ratio, slope drainage factor and relief ratio Data reflecting seasonal/annual cycle: seasonal index Remote sensed data: Remote sensed optical data and microwave data Other data: fuzzy set membership
Nitrate	Meteorological data: rainfall, cumulative rainfall, snowfall, air temperature (maximum, minimum), precipitation, snow melt condition, Geographical and physical factors: treatment, total nitrogen applied, on-ground nitrogen loading and recharge data, past nitrate concentrations, soil condition (frozen or non-frozen surface), streamflow, and nitrate load Land cover: RS VIs Data reflecting seasonal/annual cycle: Julian day of the year, year of experimentation
Phosphorus	Meteorological data: rainfall, air temperature, Geographical and physical factors: (time-lagged) streamflow, antecedent phosphorus, phosphorus load, suspended solids Land cover: RS VIs Data reflecting seasonal/annual cycle: seasonal index,
Cynobacteria (*Anabaena.* spp.) (cells/mL)	Turbidity, color, temperature, flow, the concentrations of total nitrogen, soluble and total phosphorus, conductivity, dissolved oxygen
Peak *Cryptosporidium* spp.	Turbidity, *Clostridium perfringens*, pH, total coliforms, fecal coliforms, river flow, *Escherichia coli*, precipitation
Classification of sources and ages of fecal contamination	Commonly measured indicator bacteria: background colonies, atypical coliforms, total coliforms, fecal coliforms, fecal streptococci Weather conditions: rain Others: turbidity, sources
Cryptosporidium spp. and *Giardia* spp. (cells/mL)	Inputs for *Crytosporidium* spp.: precipitation, river flow, pH, turbidity, alkalinity, suspended solids, dissolved solids, *Clostridium perfringens*, fecal coloform, *E. coli*; and total coliform Inputs for *Giardia* spp.: precipitation, river flow, pH, turbidity, alkalinity, suspended solids, dissolved solids, *Clostridium perfringens*, *E. coli*, fecal coliform, coliphage, F-specific coliphage, ratio of coliphage to F-specific coliphage, total coliform
Salinity	Antecedent flow and salinity, water level

Table 2-4. Remote sensing applications for different spectral bands after (Schultz and Engman 2000).

Spectral band	Wavelength	Examples of applications
Blue	0.45 to 0.50 μm	Land use, vegetation characteristics, sediment
Green	0.50 to 0.60 μm	Green reflectance of healthy vegetation
Red	0.60 to 0.70 μm	Vegetation discrimination due to red chlorophyll absorption
Panchromatic	0.50 to 0.75 μm	Mapping, land use, stereo pairs
Near - infrared	0.75 to 0.90 μm	Biomass, crop identification, soil-crop, land-water boundary
Mid-infrared	1.5 to 1.75 μm	Plant turgidity, droughts, clouds, snow-ice discrimination
Mid-infrared	2.0 to 2.35 μm	Geology, rock formations
Thermal infrared	10 to 12.5 μm	Relative temperature, thermal discharges, vegetation classification, moisture
Microwave – short wave	0.1 to 5 cm	Snow cover, depth, vegetation water content
Microwave – long wave	5 to 24 cm	Melting snow, soil moisture, water-land boundaries, penetrate vegetation

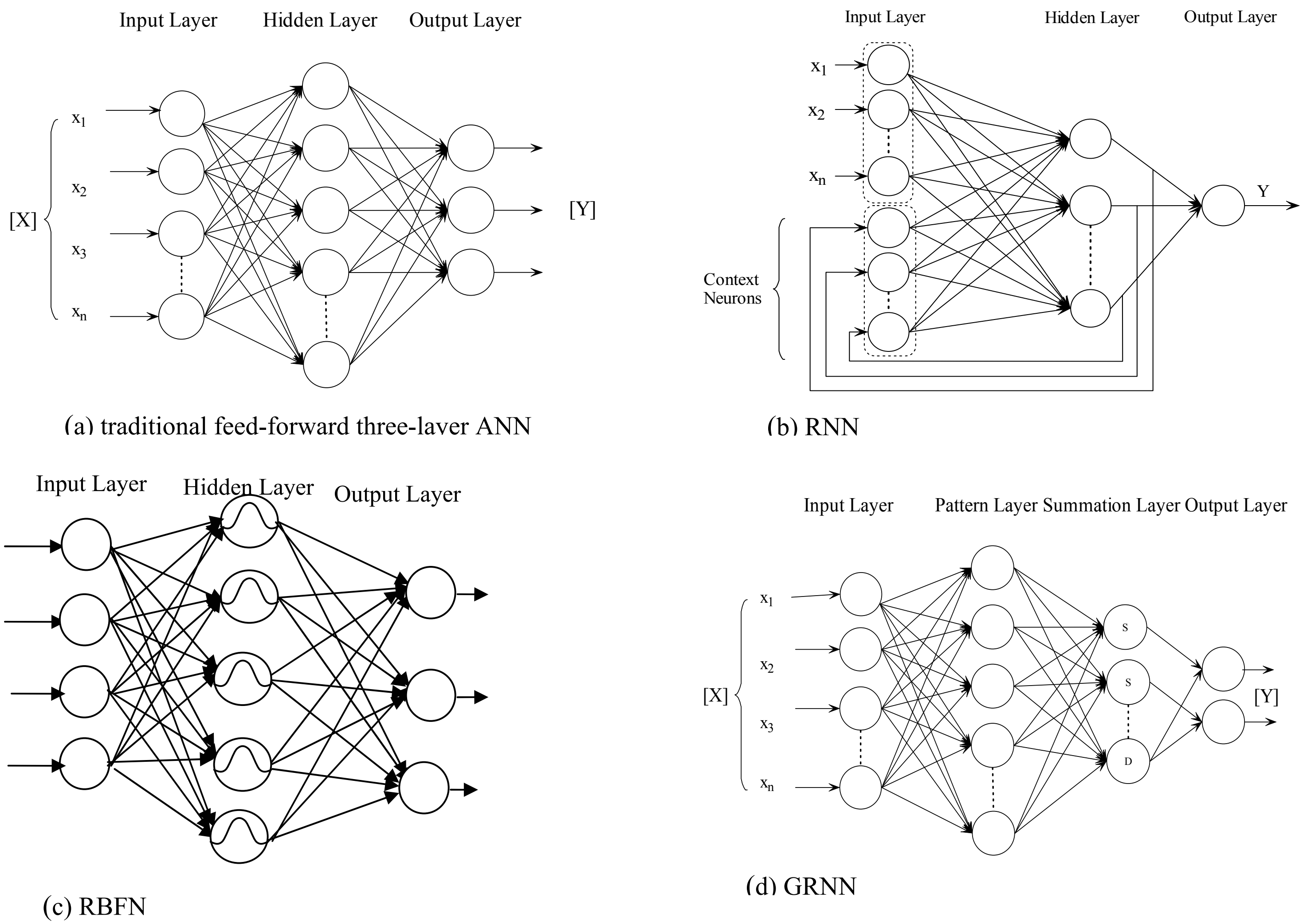
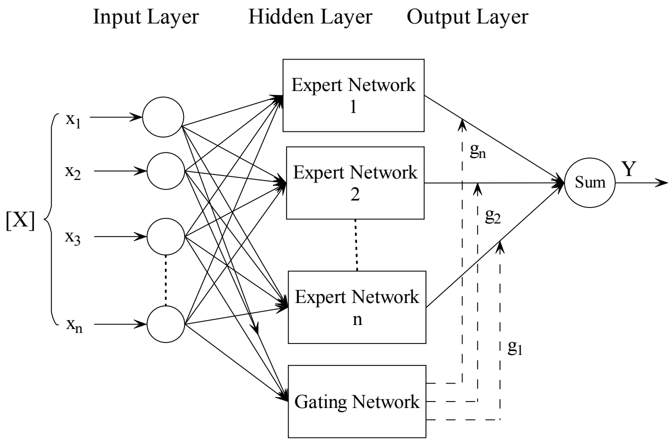
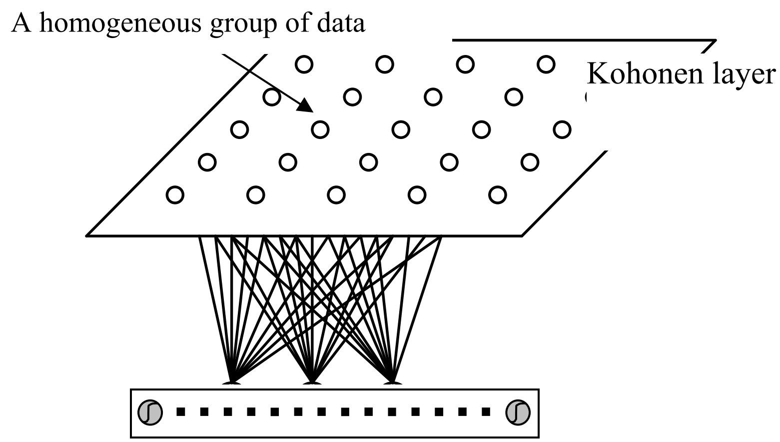


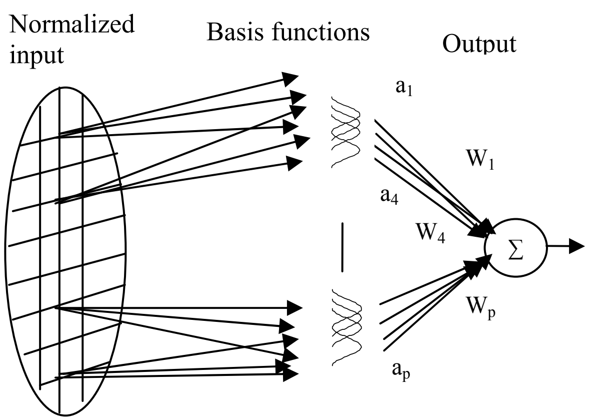
Figure 2-1. Configurations of ANNs.



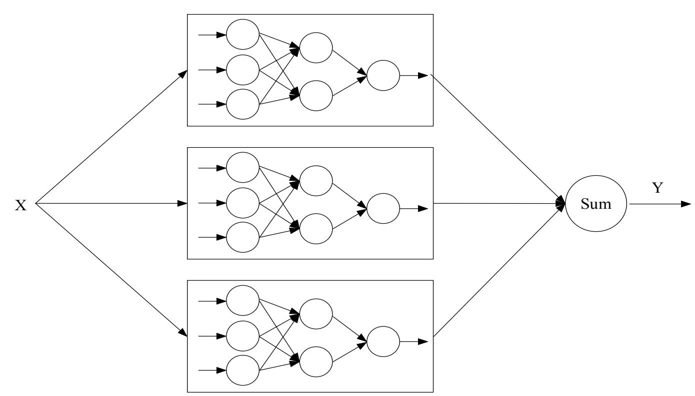
(e) MNN



(f) KSOM



(g) AMN



(h) A/C-NN

Figure 2-1. ***Continued***

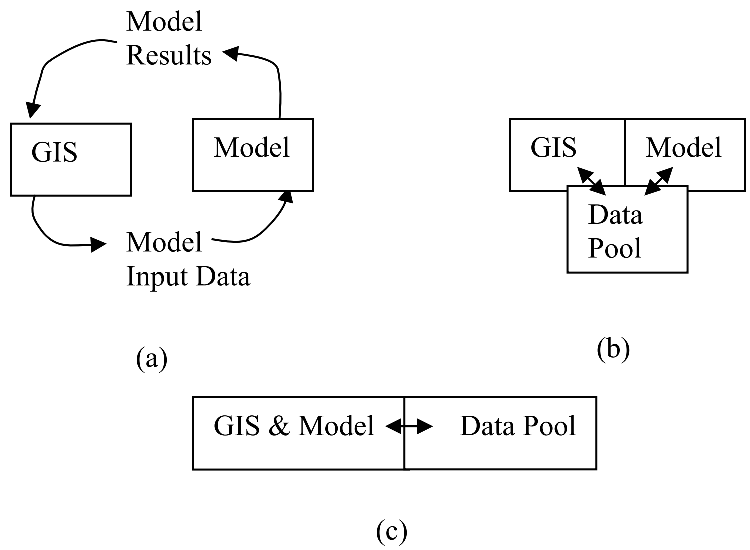


Figure 2-2. The most common ways of coupling a GIS with an environmental model (a) loose coupling, (b) tight coupling, and (c) embedded system (Corwin et al. 1997).

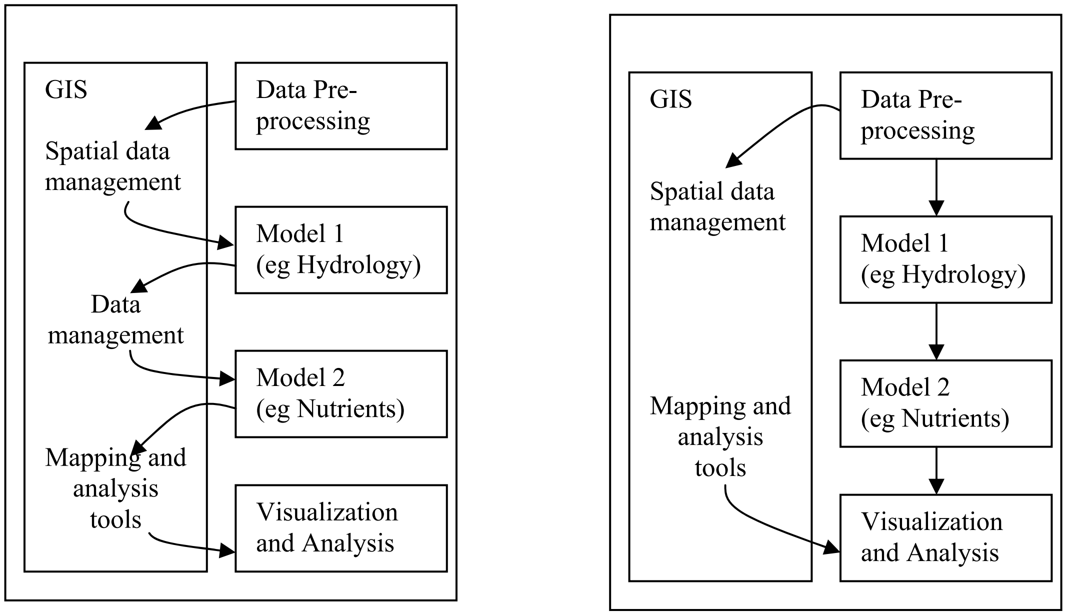


Figure 2-3. Two different views of integrating GIS and environmental models (Argent 2004).

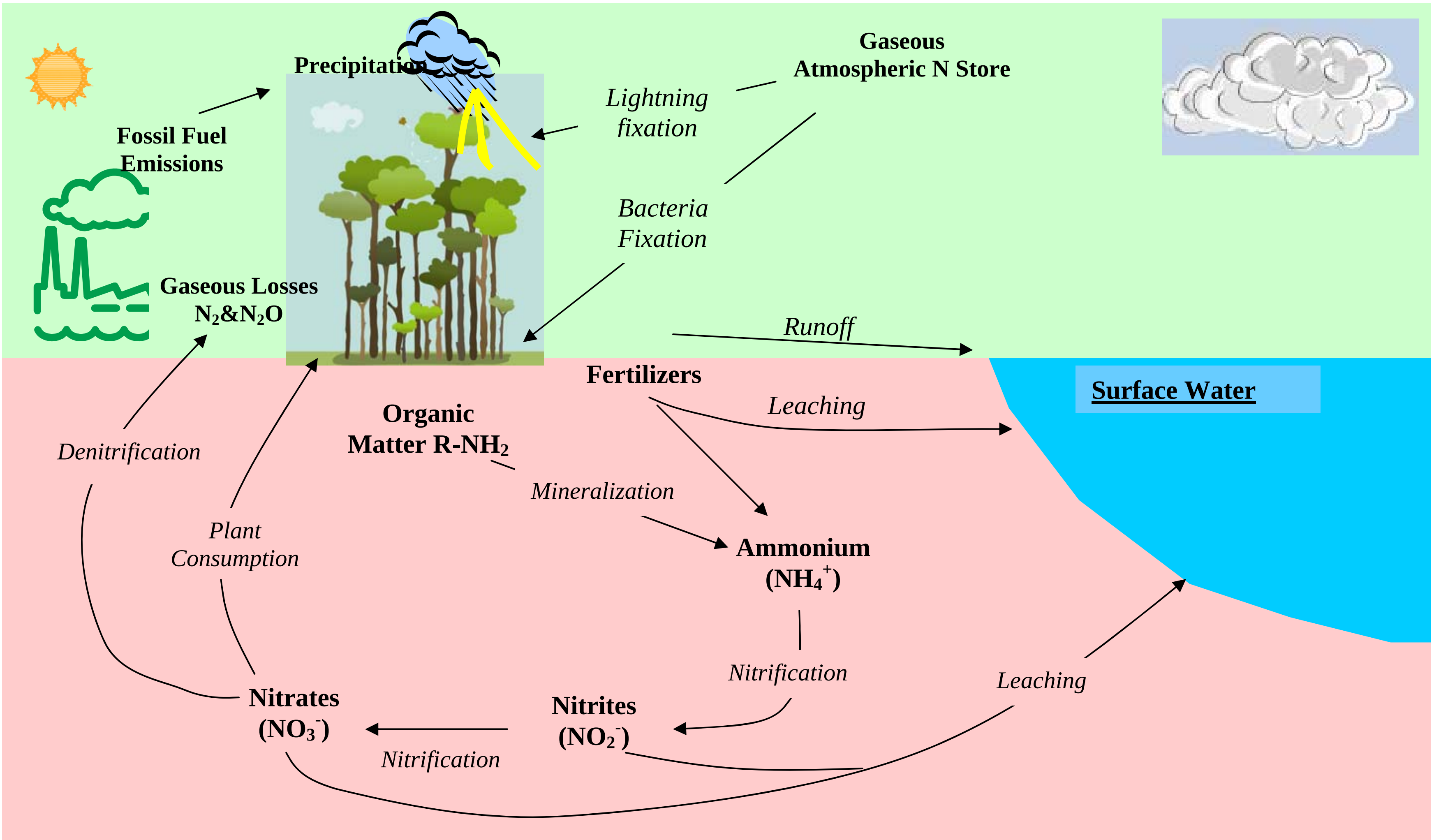


Figure 2-4. Nitrogen cycle

2.6 References

Aber, J.D., Goodale, C.L., Ollinger, S.V., Smith, M.L., Magill, A.H., Martin, M.E., Hallett, R.A., and Stoddard, J.L. 2003. Is nitrogen deposition altering the nitrogen status of northeastern forests? Bioscience, **53**: 375-389.

Aber, J.D., and Magill, A.H. 2004. Chronic nitrogen additions at the Harvard Forest (USA): the first 15 years of a nitrogen saturation experiment. Forest Ecol. Manage., **196**: 1-5.

Aber, J.D., Nadelhoffer, K.J., Steudler, P., and Melillo, J.M. 1989. Nitrogen saturation in northern forest ecosystems. Bioscience, **39**: 378-386.

Aber, J.D., Ollinger, S.V., and Driscoll, C.T. 1997. Modeling nitrogen saturation in forest ecosystems in response to land use and atmospheric deposition. Ecol. Model., **101**: 61-78.

Aber, J.D., Ollinger, S.V., Driscoll, C.T., Likens, G.E., Holmes, R.T., Freuder, R.J., and Goodale, C.L. 2002. Inorganic nitrogen losses from a forested ecosystem in response to physical, chemical, biotic, and climatic perturbations. Ecosystems, **5**: 648-658.

Abrahart, R.J., and White, S.M. 2001. Modelling sediment transfer in Malawi: Comparing backpropagation neural network solutions against a multiple linear regression benchmark using small data sets. Phys. Chem. Earth Pt. B, **26**: 19-24.

Agarwal, A., Singh, R.D., Mishra, S.K., and Bhunya, P.K. 2005. ANN-based sediment yield river basin models for Vamsadhara (India). Water SA, **31**: 95-100.

Ahmad, Z., and Zhang, J. 2005a. Bayesian selective combination of multiple neural networks for improving long-range predictions in nonlinear process modelling. Neural Comput. Appl., **14**: 78-87.

Ahmad, Z., and Zhang, J. 2005b. Combination of multiple neural networks using data fusion techniques for enhanced nonlinear process modelling. Comput. Chem. Eng., **30**: 295-308.

Al-Yemni, M., and Yang, R.Y.K. 2005. Hybrid neural-networks modeling of an enzymatic membrane reactor. J. Chinese Inst. Eng., **28**: 1061-1067.

Almasri, M.N., and Kaluarachchi, J.J. 2005. Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. Environ. Modell. Softw., **20**: 851-871.

Alp, M., and Cigizoglu, H.K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ. Modell. & Softw., **22**: 2-13.

Anctil, F., and Lauzon, N. 2004. Generalisation for neural networks through data sampling and training procedures, with applications to streamflow predictions. Hydrol. Earth Sys. Sci., **8**: 940-958.

Argent, R.M. 2004. An overview of model integration for environmental application - components, frameworks and semantics. Environ. Modell. Softw., **19**: 219-234.

Arheimer, B., Andersson, L., and Lepisto, A. 1996. Variation of nitrogen concentration in forest streams influences of flow, seasonality and catchment characteristics. J. Hydrol., **179**: 281-304.

Arnold, J.G., Srinivasan, R., Ramanarayanan, T.S., and DiLuzio, M. 1999. Water resources of the Texas Gulf Basin. Water Sci. Technol., **39 (3)**: 121-133.

Aspinall, R., and Pearson, D. 2000. Integrated geographical assessment of environmental condition in water catchments: Linking landscape ecology, environmental modelling and GIS. J. Environ. Manage., **59**: 299-319.

Baker, M.E., Wiley, M.J., and Seelbach, P.W. 2001. GIS-based hydrologic modeling of riparian areas: Implications for stream water quality. J. Am. Water Resour. Assoc., **37**: 1615-1628.

Ballester, M.V.R., Victoria, D.D., Krusche, A.V., Coburn, R., Victoria, R.L., Richey, J.E., Logsdon, M.G., Mayorga, E., and Matricardi, E. 2003. A remote sensing/GIS-based physical template to understand the biogeochemistry of the Ji-Parana river basin (Western Amazonia). Remote Sens. Environ., **87**: 429-445.

Barnes, W.L., Pagano, T.S., and Salomonson, V.V. 1998. Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. IEEE Trans. Geosci. Remote Sens., **36**: 1088 - 1100.

Basnyat, P., Teeter, L.D., Lockaby, B.G., and Flynn, K.M. 2000. The use of remote sensing and GIS in watershed level analyses of non-point source pollution problems. Forest Ecol. Manage., **128**: 65-73.

Belmonte, A.C., Gonzalez, J.M., Mayorga, A.V., and Fernandez, S.C. 1999. GIS tools applied to the sustainable management of water resources - Application to the aquifer system 08-29. Agric. Water Manage., **40**: 207-220.

Berntson, G.M., and Aber, J.D. 2000. Fast nitrate immobilization in N saturated temperate forest soils. Soil Biol. Biochem., **32**: 151-156.

Bhaduri, B., Harbor, J., Engel, B., and Grove, M. 2000. Assessing watershed-scale, long-term hydrologic impacts of land-use change using a GIS-NPS model. Environ. Manage., **26**: 643-658.

Bishop, C., 1995. Neural networks for pattern recognition. Clarendon Press, Oxford.

Blum, M.L., 2001. Using geographic information systems (GIS) to aid in watershed management and stream restoration: Steamboat Creek, Nevada. University of Nevada, Reno, United States -- Nevada, p. 119.

Bowden, G.J., Dandy, G.C., and Maier, H.R. 2005a. Input determination for neural network models in water resources applications. Part 1 - background and methodology. J. Hydrol., **301**: 75-92.

Bowden, G.J., Maier, H.R., and Dandy, G.C. 2005b. Input determination for neural network models in water resources applications. Part 2. Case study: forecasting salinity in a river. J. Hydrol., **301**: 93-107.

Bowden, G.J., Maier, H.R., and Dandy, G.C. 2002. Optimal division of data for neural network models in water resources applications. Water Resour. Res., **38**: 1010.

Brezonik, P., Menken, K.D., and Bauer, M. 2005. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). Lake Reserv. Manage., **21**: 373-382.

Brimicombe, A., 2003. GIS, environmental modelling and engineering. New York : Taylor & Francis, London.

Brion, G.M., Neelakantan, T.R., and Lingireddy, S. 2001. Using neural networks to predict peak *Cryptosporidium* concentrations. J. Am. Water Works Assoc., **93 (1)**: 99-105.

Brion, G.M., Neelakantan, T.R., and Lingireddy, S. 2002. A neural-network-based classification scheme for sorting sources and ages of fecal contamination in water. Water Res., **36**: 3765-3774.

Brooks, P.D., Campbell, D.H., Tonnessen, K.A., and Heuer, K. 1999. Natural variability in N export from headwater catchments: snow cover controls on ecosystem N retention. Hydrol. Processes, **13**: 2191-2201.

Brooks, P.D., and Williams, M.W. 1999. Snowpack controls on nitrogen cycling and export in seasonally snow-covered catchments. Hydrol. Processes, **13**: 2177-2190.

Brooks, P.D., Williams, M.W., and Schmidt, S.K. 1998. Inorganic nitrogen and microbial biomass dynamics before and during spring snowmelt. Biogeochemistry, **43**: 1-15.

Brown, I. 2006. Modelling future landscape change on coastal floodplains using a rule-based GIS. Environ. Modell. Softw., **21**: 1479-1490.

Brown, M., and Harris, C., 1994. Neurofuzzy adaptive modelling and control. Prentice Hall International (UK) Ltd.

Brown, M., and Harris, C.J. 1995. A perspective and critique of adaptive neurofuzzy systems used for modeling and control applications. Int. J. Neural Sys., **6**: 197-220.

Campbell, D.H., Kendall, C., Chang, C.C.Y., Silva, S.R., and Tonnessen, K.A. 2002. Pathways for nitrate release from an alpine watershed: Determination using delta N-15 and delta O-18. Water Resources Res., **38**: 1-10.

Cedfeldt, P.T., Watzin, M.C., and Richardson, B.D. 2000. Using GIS to identify functionally significant wetlands in the Northeastern United States. Environ. Manage., **26**: 13-24.

Chandramouli, V., Brion, G., Neelakantan, T.R., and Lingireddy, S. 2007. Backfilling missing microbial concentrations in a riverine database using artificial neural networks. Water Res., **41**: 217-227.

Chen, J., and Adams, B.J. 2006. Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. J. Hydrol., **318**: 232-249.

Chen, W.C., Chang, N.B., and Chen, J.C. 2003a. Rough set-based hybrid fuzzy-neural controller design for industrial wastewater treatment. Water Res., **37**: 95-107.

Chen, W.C., Chang, N.B., and Shieh, W.K. 2001. Advanced hybrid fuzzy-neural controller for industrial wastewater treatment. J. Environ. Eng.-ASCE, **127**: 1048-1059.

Chen, X., Ding, X., and Chen, S. 2003b. The study of chlorophyll detection in coastal waters based on environmental variables. International conference of GIS and remote sensing in hydrology, water resources and environment (ICGRSHWE), China, pp. 316-321.

Chiang, Y.M., Chang, L.C., and Chang, F.J. 2004. Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling. J. Hydrol., **290**: 297-311.

Choi, D.-J., and Park, H. 2001. A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process. Water Res., **35**: 3959-3967.

Choi, J.Y., Engel, B.A., and Farnsworth, R.L. 2005. Web-based GIS and spatial decision support system for watershed management. J. Hydroinformatics, **7**: 165-174.

Choi, J.Y., Engel, B.A., Muthukrishnan, S., and Harbor, J. 2003. GIS based long term hydrologic impact evaluation for watershed urbanization. J. Am. Water Resour. Assoc., **39**: 623-635.

Chowdary, V.M., Rao, N.H., and Sarma, P.B.S. 2003. GIS-based decision support system for groundwater assessment in large irrigation project areas. Agric. Water Manage., **62**: 229-252.

Cigizoglu, H.K. 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. Adv. Water Resour., **27**: 185-195.

Cigizoglu, H.K., and Alp, M. 2006. Generalized regression neural network in modelling river sediment yield. Adv. Eng. Softw., **37**: 63-68.

Cigizoglu, H.K., and Kisi, O. 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol., **317**: 221-238.

Clark, M.J., 2000. Putting water in its place: a perspective on GIS in hydrology and water management. In: Gurnel, A.M., and Montgomery, D.R. eds., Hydrological Applications of GIS. John Wiley & Sons, pp. 3-14.

Clein, J.S., and Schimel, J.P. 1995. Microbial activity of tundra and taiga soils at subzero temperatures. Soil Biol. Biochem., **27**: 1231-1234.

Corwin, D.L., Vaughan, P.J., and Loague, K. 1997. Modeling nonpoint source pollutants in the vadose zone with GIS. Environ. Sci. Technol., **31**: 2157 - 2175.

Currie, W.S., Aber, J.D., and Driscoll, C.T. 1999. Leaching of nutrient cations from the forest floor: effects of nitrogen saturation in two long-term manipulations. Can. J. Forest Res., **29**: 609-620.

Davies, H., and Neal, C. 2004. GIS-based methodologies for assessing nitrate, nitrite and ammonium distributions across a major UK basin, the Humber. Hydrol. Earth Sys. Sci., **8**: 823-833.

Dawson, C.W., See, L.M., Abrahart, R.J., and Heppenstall, A.J. 2006. Symbiotic adaptive neuro-evolution applied to rainfall-runoff modelling in northern England. Neural Networks, **19**: 236-247.

de Jong van Lier, Q., Sparovek, G., Flanagan, D.C., Bloem, E.M., and Schnug, E. 2005. Runoff mapping using WEPP erosion model and GIS tools. Comput. Geosci., **31**: 1270-1276.

de Paz, J.M., and Ramos, C. 2004. Simulation of nitrate leaching for different nitrogen fertilization rates in a region of Valencia (Spain) using a GIS-GLEAMS system. Agric. Ecosys. Environ., **103**: 59-73.

Di Luzio, M., Arnold, J.G., and Srinivasan, R. 2005. Effect of GIS data quality on small watershed stream flow and sediment simulations. Hydrol. Processes, **19**: 629-650.

Fensholt, R., and Sandholt, I. 2005. Evaluation of MODIS and NOAA AVHRR vegetation indices with in situ measurements in a semi-arid environment. Int. J. Remote Sens., **26**: 2561-2594.

Fernandez, G.P., Chescheir, G.M., Skaggs, R.W., and Amatya, D.M. 2002. WATGIS: A GIS-based lumped parameter water quality model. Trans. ASAE, **45**: 593-600.

Fernandez, G.P., Chescheir, G.M., Skaggs, R.W., and Amatya, D.M. 2006. DRAINMOD-GIS: A lumped parameter watershed scale drainage and water quality model. Agric. Water Manage., **81**: 77-97.

Foster, J.A., and McDonald, A.T. 2000. Assessing pollution risks to water supply intakes using geographical information systems (GIS). Environ. Modell. Softw., **15**: 225-234.

Gaume, E., and Gosset, R. 2003. Over-parameterisation,a major obstacle to the use of artificial neural networks in hydrology? Hydrol. Earth Sys. Sci., **7**: 693-706.

Gevrey, M., Dimopoulos, I., and Lek, S. 2006. Two-way interaction of input variables in the sensitivity analysis of neural network models. Ecol. Modell., **195**: 43-50.

Gevrey, M., Dimopoulos, L., and Lek, S. 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol. Modell., **160**: 249-264.

Giasson, E., Bryant, R.B., and DeGloria, S.D. 2002. GIS-based spatial indices for identification of potential phosphorous export at watershed scale. J. Soil Water Cons., **57**: 373-381.

Giupponi, C., and Vladimirova, I. 2006. Ag-PIE: A GIS-based screening model for assessing agricultural pressures and impacts on water quality on a European scale. Sci. Total Environ., **359**: 57-75.

Giustolisi, O., and Laucelli, D. 2005. Improving generalization of artificial neural networks in rainfall-runoff modelling. Hydrol. Sci. J.-J. Sci. Hydrol., **50**: 439-457.

Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., and Maidment, D.C., 1996. GIS and Environmental Modeling: Process and Research Issues. GIS World Books, Fort Collins.

Groffman, P.M., Driscoll, C.T., Fahey, T.J., Hardy, J.P., Fitzhugh, R.D., and Tierney, G.L. 2001. Effects of mild winter freezing on soil nitrogen and carbon dynamics in a northern hardwood forest. Biogeochemistry, **56**: 191-213.

Groffman, P.M., Hardy, J.P., Driscoll, C.T., and Fahey, T.J. 2006. Snow depth, soil freezing, and fluxes of carbon dioxide, nitrous oxide and methane in a northern hardwood forest. Global Change Bio., **12**: 1748-1760.

Groffman, P.M., Hardy, J.P., Nolan, S., Fitzhugh, R.D., Driscoll, C.T., and Fahey, T.J. 1999. Snow depth, soil frost and nutrient loss in a northern hardwood forest. Hydrol. Processes, **13**: 2275-2286.

Grunwald, S., and Qi, C. 2006. GIS-based water quality modeling in the Sandusky Watershed, Ohio, USA. J. Am. Water Resour. Assoc., **42**: 957-973.

Hamidi, N., and Kayaalp, N. 2008. Estimation of the amount of suspended sediment in the Tigris River using artificial neural networks. Clean-Soil Air Water, **36**: 380-386.

Hao, F., Zhang, X., Cheng, H., Liu, C., and Yang, Z. 2003. Runoff and sement yield simulation in a large basin using GIS and a distributed hydrological model. International Conference of GIS and Remote Sensing in Hydrology, Water Resour. Environ., China, pp. 157-166.

Haverkamp, S., Fohrer, N., and Frede, H.G. 2005. Assessment of the effect of land use patterns on hydrologic landscape functions: a comprehensive GIS-based tool to minimize model uncertainty resulting from spatial aggregation. Hydrol. Processes, **19**: 715-727.

He, C. 2003. Integration of geographic information systems and simulation model for watershed management. Environ. Modell. Softw., **18**: 809-813.

Hellweger, F.L., and Maidment, D.R. 1999. Definition and connection of hydrologic elements using geographic data. J. Hydrol. Eng., **4**: 10-18.

Ito, M., Mitchell, M.J., Driscoll, C.T., and Roy, K.M. 2005. Nitrogen input-output budgets for lake-containing watersheds in the Adirondack region of New York. Biogeochemistry, **72**: 283-314.

Jain, A., and Kumar, A.M. 2007. Hybrid neural network models for hydrologic time series forecasting. Appl. Soft. Comput., **7**: 585-592.

Jain, A., Sudheer, K.P., and Srinivasulu, S. 2004. Identification of physical processes inherent in artificial neural network rainfall runoff models. Hydrol. Processes, **18**: 571-581.

Jain, S.K. 2008. Development of integrated discharge and sediment rating relation using a compound neural network. J. Hydrol. Eng., **13**: 124-131.

Jones, P.L., Khairy, W.M., and Coleman, T.L. 2002. Advances in watersheds modeling using remote sensing and geographical information systems technologies. IEEE Geosci. Remote Sens. Lett.: 2927-2929.

Justice, C.O., Vermote, E., Townshend, J.R.G., Defries, R., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.L., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Wan, Z., Huete, A.R., van Leeuwen, W., Wolfe, R.E., Giglio, L., Muller, J., Lewis, P., and Barnsley, M.J. 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. IEEE Trans. Geosci. Remote Sens., **36**: 1228 - 1249.

Kaluli, J.W., Madramootoo, C.A., and Djebbar, Y. 1998. Modeling nitrate leaching using neural networks. Water Sci. Technol., **38 (7)**: 127-134.

Kaste, O., and Skjelkvale, B.L. 2002. Nitrogen dynamics in runoff from two small heathland catchments representing opposite extremes with respect to climate and N deposition in Norway. Hydrol. Earth Sys. Sci., **6**: 351-362.

Kielland, K., Olson, K., Ruess, R.W., and Boone, R.D. 2006. Contribution of winter processes to soil nitrogen flux in taiga forest ecosystems. Biogeochemistry, **81**: 349-360.

Kim, M., Choi, C.Y., and Gerba, C.R. 2008. Source tracking of microbial intrusion in water systems using artificial neural networks. Water Res., **42**: 1308-1314.

Kim, M.Y., and Kim, M.K. 2007. Dynamics of surface runoff and its influence on the water quality using competitive algorithms in artificial neural networks. J. Environ. Sci. Health Part A–Toxic/ Hazardous Substances Environ. Eng., **42**: 1057-1064.

Kim, M.Y., Seo, M.C., and Kim, M.K. 2007. Linking hydro-meteorological factors to the assessment of nutrient loadings to streams from large-plotted paddy rice fields. Agric. Water Manage., **87**: 223-228.

Kingston, G.B., Lambert, M.F., and Maier, H.R. 2005. Bayesian training of artificial neural networks used for water resources modeling. Water Resour. Res., **41**: W12409.

Kingston, G.B., Maier, H.R., and Lambert, M.F. 2006. A probabilistic method for assisting knowledge extraction from artificial neural networks used for hydrological prediction. Math. Comput. Modell., **44**: 499-512.

Kisi, O. 2004. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. Hydrol. Sci. J.-J. Sci. Hydrol., **49**: 1025-1040.

Kisi, O. 2005. Suspended sediment estimation using neuro-fuzzy and neural network approaches. Hydrol. Sci. J.-J. Sci. Hydrol., **50**: 683-696.

Kisi, O. 2008. Constructing neural network sediment estimation models using a data-driven algorithm. Math. Comput. Simul., **79**: 94-103.

Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern., **43**: 59-69.

Koponen, S., Kallio, K., Pulliainen, J., Vepsalainen, J., Pyhalahti, T., and Hallikainen, M. 2004. Water quality classification of lakes using 250-m MODIS data. IEEE Geosci. Remote Sens. Lett., **1**: 287 - 291.

Kutser, T., Metsamaa, L., Strombeck, N., and Vahtmae, E. 2006. Monitoring cyanobacterial blooms by satellite remote sensing. Estuar. Coast. Shelf Sci., **67**: 303-312.

Kutser, T., Pierson, D., Tranvik, L., Reinart, A., Sobek, S., and Kallio, K. 2005. Using satellite remote sensing to estimate the colored dissolved organic matter absorption coefficient in lakes. Ecosystems, **8**: 709-720.

Lam, D., and Pupp, C., 1996. Integration of GIS, expert systems, and modeling for state-of-environment reporting. In: Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., Crane, M., and Glendinning, S. eds., GIS and Environmental Modeling: Progess and Research Issues. GIS World Inc., Fort Collins, CO, USA, pp. 419-422.

Lant, C.L., Kraft, S.E., Beaulieu, J., Bennett, D., Loftus, T., and Nicklow, J. 2005. Using GIS-based ecological-economic modeling to evaluate policies affecting agricultural watersheds. Ecol. Econ., **55**: 467-484.

Lee, D.S., Jeon, C.O., Park, J.M., and Chang, K.S. 2002. Hybrid neural network modeling of a full-scale industrial wastewater treatment process. Biotechnol. Bioeng., **78**: 670-682.

Lek, S., Guiresse, M., and Giraudel, J.-L. 1999. Predicting stream nitrogen concentration from watershed features using neural networks. Water Res., **33**: 3469-3478.

Li, X., Nour, M.H., Smith, D.W., and Prepas, E.E. 2008. Modelling Nitrogen Composition in Streams on the Boreal Plain using Genetic Adaptive General Regression Neural Networks. J. Environ. Eng. Sci. **7** (S1): 109 – 125

Livingstone, D., and Raper, J., 1994. Modelling environmental systems with GIS: theoretical barriers to progress. In: Worboys, M.F. ed., Innovations in GIS. Taylor & Francis, London, p. 229-240.

Loague, K., and Corwin, D.L., 2000. Regional-scale assessment of non-point source groundwater contamination. In: Gurnel, A.M., and Montgomery, D.R. eds., Hydrological Applications of GIS. John Wiley & Sons, pp. 137-145.

Luke, S.H., Luckai, N.J., Burke, J.M., and Prepas, E.E. 2007. Riparian areas in the Canadian boreal forest and linkages with water quality in streams. Environ. Rev., **15**: 79-97.

Maier, H.R., and Dandy, G.C. 1999. Empirical comparison of various methods for training feed-forward neural networks for salinity forecasting. Water Resour. Res., **35**: 2591-2596.

Maier, H.R., and Dandy, G.C. 2001. Neural network based modelling of environmental variables: A systematic approach. Math. Comput. Modell., **33**: 669-682.

Maier, H.R., Dandy, G.C., and Burch, M.D. 1998. Use of artificial neural networks for modelling cyanobacteria *Anabaena* spp. in the River Murray, South Australia. Ecol. Modell., **105**: 257-272.

Maier, H.R., Sayed, T., and Lence, B.J. 2000. Forecasting cyanobacterial concentrations using B-spline networks. J. Comput. Civil Eng., **14**: 183-189.

Maier, H.R., Sayed, T., and Lence, B.J. 2001. Forecasting cyanobacterium *Anabaena* spp. in the River Murray, South Australia, using B-spline neurofuzzy models. Ecol. Modell., **146**: 85-96.

Mandic, D.P., 2001. Recurrent neural networks for prediction: learning algorithms, architectures, and stability. New York : John Wiley, Chichester.

Mani, V., and Omkar, S.N. 2002. Understanding weld modelling processes using a combination of trained neural networks. Int. J. Prod. Res., **40**: 547-559.

Markus, M., Tsai, C.W.S., and Demissie, M. 2003. Uncertainty of weekly nitrate-nitrogen forecasts using artificial neural networks. J. Environ. Eng.-ASCE, **129**: 267-274.

Matejicek, L., Benesova, L., and Tonika, J. 2003. Ecological modelling of nitrate pollution in small river basins by spreadsheets and GIS. Ecol. Modell., **170**: 245-263.

Mattikalli, N.M., 2000. Integration of remotely sensed data into geographical information system. In: Schultz, G.A., and Engman, E.T. eds., Remote Sensing in Hydrology and Water Management. Springer.

Meireles, M.R.G., Almeida, P.E.M., and Simoes, M.G. 2003. A comprehensive review for industrial applicability of artificial neural networks. IEEE Trans. Ind. Electronic., **50**: 585-601.

Mikan, C.J., Schimel, J.P., and Doyle, A.P. 2002. Temperature controls of microbial respiration in arctic tundra soils above and below freezing. Soil Biol. Biochem., **34**: 1785-1795.

Mourad, D.S.J., Van der Perk, M., Gooch, G.D., Loigu, E., Piirimae, K., and Stalnacke, P. 2005. GIS-based quantification of future nutrient loads into Lake Pelpsi/Chudskoe using qualitative regional development scenarios. Water Sci. Technol., **51 (3-4)**: 355-363.

Mozejko, J., and Gniot, R. 2008. Application of neural networks for the prediction of total phosphorus concentrations in surface waters. Pol. J. Environ. Stud., **17**: 363-368.

Nagy, H.M., Watanabe, K., and Hirano, M. 2002. Prediction of sediment load concentration in rivers using artificial neural network model. J. Hydraul. Eng.-ASCE, **128**: 588-595.

Nayak, P.C., Sudheer, K.P., and Jain, S.K. 2007. Rainfall-runoff modeling through hybrid intelligent system. Water Resour. Res., **43 (7): W07415**.

Neelakantan, T.R., Brion, G.M., and Lingireddy, S. 2001. Neural network modelling of *Cryptosporidium* and *Giardia* concentrations in the Delaware River, USA. Water Sci. Technol., **43 (12)**: 125-132.

Neelakantan, T.R., Lingireddy, S., and Brion, G.M. 2002. Effectiveness of different artificial neural network training algorithms in predicting protozoa risks in surface waters. J. Environ. Eng.-ASCE, **128**: 533-542.

Ning, S.K., Chang, N.B., Jeng, K.Y., and Tseng, Y.H. 2006. Soil erosion and non-point source pollution impacts assessment with the aid of multi-temporal remote sensing images. J. Environ. Manage., **79**: 88-101.

Nour, M.H., Khan, A., Smith, D.W., and Gamal El-Din, M. 2005. On the potential of satellite derived vegetation phenology for watershed nutrient modelling: a neural network approach. Proc. WEFTEC 2005, Washington, DC.

Nour, M.H., Smith, D.W., El-Din, M.G., and Prepas, E.E. 2006a. Neural networks modelling of streamflow, phosphorus, and suspended solids: application to the Canadian Boreal forest. Water Sci. Technol., **53**: 91-99.

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

Olden, J.D., and Jackson, D.A. 2002. Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. Ecol. Modell., **154**: 135-150.

Olden, J.D., Joy, M.K., and Death, R.G. 2004. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecol. Modell., **178**: 389-397.

Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H.D., Koka, S., and Agrawal, A. 2006. ArcGIS-SWAT: A geodata model and GIS interface for SWAT. J. Am. Water Resour. Assoc., **42**: 295-309.

Panikova, N.S., Flanaganb, P.W., Oechelc, W.C., Mastepanovd, M.A., and Christensend, T.R. 2006. Microbial activity in soils frozen to below - 39 degrees C. Soil Biol. Biochem., **38**: 785-794.

Piatek, K.B., Mitchell, M.J., Silva, S.R., and Kendall, C. 2005. Sources of nitrate in snowmelt discharge: Evidence from water chemistry and stable isotopes of nitrate. Water Air Soil Poll., **165**: 13-35.

Qi, C., and Grunwald, S. 2005. GIS-based hydrologic modeling in the sandusky watershed using SWAT. Trans. ASAE, **48**: 169-180.

Raghuwanshi, N.S., Singh, R., and Reddy, L.S. 2006. Runoff and sediment yield modeling using artificial neural networks: upper Siwane River, India. J. Hydrol. Eng., **11**: 71-79.

Rai, R.K., and Mathur, B.S. 2008. Event-based sediment yield modeling using artificial neural network. Water Resour. Manage., **22**: 423-441.

Rankinen, K., Granlund, K., and Barlund, I. 2004. Modelling of seasonal effects of soil processes on N leaching in northern latitudes. Nordic Hydrol., **35**: 347-357.

Raper, J., and Livingstone, D. 1995. Development of a geomorphological spatial model using object-oriented design. Int. J. Geogr. Infor. Syst., **9**: 359-383.

Reed, B.C., Brown, J.F., and Loveland, T.R., 2002. Geographic data for environmental modelling and assessment. In: Skidmore, A. ed., Environmental Modelling with GIS and Remote Sensing. Taylor & Francis, London, pp. 52-69.

Ritchie, J.C., Zimba, P.V., and Everitt, J.H. 2003. Remote sensing techniques to assess water quality. Photogram. Eng. Remote Sens., **69**: 695-704.

Rivkina, E., Gilichinsky, D., Wagener, S., Tiedje, J., and McGrath, J. 1998. Biogeochemical activity of anaerobic microorganisms from buried permafrost sediments. Geomicrobiology J., **15**: 187-193.

Rivkina, E., Laurinavichius, K., McGrath, J., Tiedje, J., Shcherbakova, V., and Gilichinsky, D., 2004. Microbial life in permafrost. Space Life Sciences: Search for Signatures of Life, and Space Flight Environmental Effects on the Nervous System, pp. 1215-1221.

Robinson, J. 1995. EOS's Moderate Resolution Imaging Spectrometer (MODIS). IEEE Expert, [see also IEEE Intelligent Systems and Their Applications], **10**: 27.

Romanovsky, V.E., and Osterkamp, T.E. 2000. Effects of unfrozen water on heat and mass transport processes in the active layer and permafrost. Permaf. Periglac. Process., **11**: 219-239.

Ruess, R.W., Hendrick, R.L., Burton, A.J., Pregitzer, K.S., Sveinbjornsson, B., Allen, M.E., and Maurer, G.E. 2003. Coupling fine root dynamics with ecosystem carbon cycling in black spruce forests of interior Alaska. Ecol. Monogr., **73**: 643-662.

Rumelhart, D.E., McClelland, J.L., and Group., T.P.R., 1986. Parallel distributed processing : explorations in the microstructure of cognition Cambridge, Mass. : MIT Press.

Salehi, F., Prasher, S.O., Amin, S., Madani, A., Jebelli, S.J., Ramaswamy, H.S., Tan, C., and Drury, C.F. 2000. Prediction of annual nitrate-N losses in drain outflows with artificial neural networks. Trans. ASAE, **43**: 1137-1143.

Sallehuddin, R., Shamsuddin, S.M.H., Hashim, S.Z.M., and Abraham, A. 2007. Forecasting time series data using hybrid grey relational artificial neural network and auto regressive integrated moving average model. Neural Netw. World, **17**: 573-605.

Samanta, B., Bandopadhyay, S., Ganguli, R., and Dutta, S. 2005. A comparative study of the performance of single neural network vs. adaboost algorithm based combination of multiple neural networks for mineral resource estimation. J. South Afr. Inst. Min. Metall., **105**: 237-246.

Santhi, C., Muttiah, R.S., Arnold, J.G., and Srinivasan, R. 2005. A GIS-based regional planning tool for irrigation demand assessment and savings using SWAT. Trans. ASAE, **48**: 137-147.

Sarangi, A., and Bhattacharya, A.K. 2005. Comparison of artificial neural network and regression models for sediment loss prediction from Banha watershed in India. Agricul. Water Manage., **78**: 195-208.

Sarangi, A., Madramootoo, C.A., Enright, P., Prasher, S.O., and Patel, R.M. 2005. Performance evaluation of ANN and geomorphology-based models for runoff and sediment yield prediction for a Canadian watershed. Curr. Sci., **89**: 2022-2033.

Schimel, J.P., Bilbrough, C., and Welker, J.A. 2004. Increased snow depth affects microbial activity and nitrogen mineralization in two Arctic tundra communities. Soil Biol. Biochem., **36**: 217-227.

Schmidt, I.K., Jonasson, S., and Michelsen, A. 1999. Mineralization and microbial immobilization of N and P in arctic soils in relation to season, temperature and nutrient amendment. Appl. Soil Ecol., **11**: 147-160.

Schultz, G.A., and Engman, E.T., 2000. Introduction. In: Schultz, G.A., and Engman, E.T. eds., Remote Sensing in Hydrology and Water Management. Springer.

Shabin, M.A., Maier, H.R., and Jaksa, M.B. 2004. Data division for developing neural networks applied to geotechnical engineering. J. Comput. Civil. Eng., **18**: 105-114.

Sharma, V., Negi, S.C., Rudra, R.P., and Yang, S. 2003. Neural networks for predicting nitrate-nitrogen in drainage water. Agirc. Water Manage., **63**: 169-183.

Sinkevich, M.G., Walter, M.T., Lembo, A.J., Richards, B.K., Peranginangin, N., Aburime, S.A., and Steenhuis, T.S. 2005. A GIS-based ground water contamination risk assessment tool for pesticides. Ground Water Monit. Remedi., **25**: 82-91.

Skidmore, A.K. 2002. Taxonomy of environmental models in the spatial sciences. *In* Environmental Modelling with GIS and Remote Sensing. *Edited by* A.K. Skidmore. Taylor & Francis Inc., New York. pp. 8–23.

Skole, D.S., Justice, C.O., Janetos, A., and Townshend, J.R.G. 1997. A land cover change monitoring program: a strategy for international effort, Amsterdam, The Netherlands: Kluwer.

Specht, D.F. 1991. a general regression neural network. Ieee Trans. Neural Networks, **2**: 568-576.

Stottlemyer, R., and Toczydlowski, D. 1999. Seasonal relationships between precipitation, forest floor, and streamwater nitrogen, Isle Royale, Michigan. Soil Sci. Soc. Am. J., **63**: 389-398.

Sudheer, K.P. 2005. Knowledge extraction from trained neural network river flow models. J. Hydrol. Eng., **10**: 264-269.

Sudheer, K.P., and Jain, A. 2004. Explaining the internal behaviour of artificial neural network river flow models. Hydrol. Processes, **18**: 833-844.

Suen, J.P., and Eheart, J.W. 2003. Evaluation of neural networks for modeling nitrate concentrations in rivers. J. Water Resour. Plan. Manage.-ASCE, **129**: 505-510.

Sui, D.Z., and Maggio, R.C. 1999. Integrating GIS with hydrological modeling: practices, problems, and prospects. Comput. Environ. Urban Syst., **23**: 33-51.

Tayfur, G., and Guldal, V. 2006. Artificial neural networks for estimating daily total suspended sediment in natural streams. Nordic Hydrol., **37**: 69-79.

Tayfur, G., Ozdemir, S., and Singh, V.P. 2003. Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces. Adv. Water Resour., **26**: 1249-1256.

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

Terrado, M., Barcelo, D., and Tauler, R. 2006. Identification and distribution of contamination sources in the Ebro river basin by chemometrics modelling coupled to geographical information systems. Talanta, **70**: 691-704.

The ASCE Task Committee. 2000. Artificial neural networks in hydrology. I: preliminary concepts. J. Hydrol. Eng., **5**: 115-123.

Tsoukalas, L.H., and Uhrig, R.E., 1997. Fuzzy and Neural Approaches in Engineering. A Wiley-Interscience publication, New York : Wiley.

Tyler, A.N., Svab, E., Preston, T., Presing, M., and Kovacs, W.A. 2006. Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment. Int. J. Remote Sens., **27**: 1521-1537.

Valenzuelaa, O., Rojas, I., Rojas, F., Pomares, H., Herrera, L.J., Guillen, A., Marquez, L., and Pasadas, M. 2008. Hybridization of intelligent techniques and ARIMA models for time series prediction. Fuzzy Set Syst., **159**: 821-845.

Venterea, R.T., Groffman, P.M., Verchot, L.V., Magill, A.H., and Aber, J.D. 2004. Gross nitrogen process rates in temperate forest soils exhibiting symptoms of nitrogen saturation. Forest Ecol. Manage., **196**: 129-142.

Versace, M., Bhatt, R., Hinds, O., and Shiffer, M. 2004. Predicting the exchange traded fund DIA with a combination of genetic algorithms and neural networks. Expert Syst. Appl., **27**: 417-425.

Vignolo, A., Pochettino, A., and Cicerone, D. 2006. Water quality assessment using remote sensing techniques: Medrano Creek, Argentina. J. Environ. Manage., **81**: 429-433.

Vitousek, P.M., and Matson, P.A. 1984. Mechanisms of nitrogen-retention in forest ecosystems - a field experiment. Science, **225**: 51-52.

Vitousek, P.M., and Reiners, W.A. 1975. Ecosystem succession and nutrient retention - hypothesis. Bioscience, **25**: 376-381.

Vivoni, E.R., and Richards, K.T. 2005. Integrated use of GIS-based field sampling and modeling for hydrologic and water quality studies. J. Hydroinfor., **7**: 235-250.

Wang, J., and Yan, D.W., 2004. A high precision prediction method by using combination of ELMAN and SOM neural networks. Advances in Neural Networks - Isnn 2004, Pt 2, pp. 943-949.

Wang, W., Van Gelder, P., Vrijling, J.K., and Ma, J. 2006. Forecasting daily streamflow using hybrid ANN models. J. Hydrol., **324**: 383-399.

Wilby, R.L., Abrahart, R.J., and Dawson, C.W. 2003. Detection of conceptual model rainfall-runoff processes inside an artificial neural network. Hydrol. Sci. J.-J. Sci. Hydrol., **48**: 163-181.

Williams, M.W., Baron, J.S., Caine, N., Sommerfeld, R., and Sanford, R. 1996a. Nitrogen saturation in the Rocky Mountains. Environ. Sci. Technol., **30**: 640-646.

Williams, M.W., Brooks, P.D., Mosier, A., and Tonnessen, K.A. 1996b. Mineral nitrogen transformations in and under seasonal snow in a high-elevation catchment in the Rocky Mountains, United States. Water Resour. Res., **32**: 3161-3171.

Williams, M.W., Hood, E., and Caine, N. 2001. Role of organic nitrogen in the nitrogen cycle of a high-elevation catchment, Colorado Front Range. Water Resour. Res., **37**: 2569-2581.

Xu, F.L., Tao, S., Dawson, R.W., and Li, B.G. 2001. A GIS-based method of lake eutrophication assessment. Ecol. Modell., **144**: 231-244.

Xu, H., and Zeng, Y. 2003. Development of an environment detection information system and its application in the region of Longyangxia reservoir, upper reach of the Yellow River. International Conference of GIS and remote sensing in hydrology, water resources and environment, China, pp. 322-327.

Yacobi, Y.Z., and Schlichter, M. 2003. GIS application for mapping of phytoplankton using multi-channel fluorescence probe derived information. International Conference of GIS and Remote Sensing in Hydrology, Water

Resources and Environment (ICGRSHWE), Three Gorges Dam site, China, p. 301-307.

Yang, M.D., Merry, C.J., and Sykes, R.M. 1999. Integration of water quality modeling, remote sensing, and GIS. J. Am. Water Resour. Assoc., **35**: 253-263.

Yin, Q., Gong, C.L., Kuang, D.B., Zhou, N., Hu, Y., Zhang, F.L., Xu, W.D., and Ma, Y.Q. 2005. Method of satellite remote sensing of lake water quality and its applications. J. Infr. Millim. Waves, **24**: 198-202.

Yu, C., Northcott, W.J., and McIsaac, G.F. 2004. Development of an artificial neural network for hydrologic and water quality modeling of agricultural watersheds. Trans.ASAE, **47**: 285-290.

Zhang, Y., Pulliainen, J., Koponen, S., and Hallikainen, M. 2002. Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data. Remote Sens. Environ., **81**: 327-336.

Zhao, H., Hao, O.J., McAvoy, T.J., and Chang, C.-H. 1997. Modeling nutrient dynamics in sequencing batch reactor. J. Environ. Eng., **123**: 311-319.

Zhu, Y.M., Lu, X., and 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.

CHAPTER 3. MODELLING NITROGEN COMPOSITIONS IN STREAMS ON THE BOREAL PLAIN USING GENETIC ADAPTIVE GENERAL REGRESSION NEURAL NETWORKS[1]

3.1 Introduction

Nitrogen is one of the most significant water quality parameters that affect ecological health. Increased nitrogen loading into aquatic ecosystems after watershed disturbance (e.g., wildfire, forest harvest) has been associated with water quality problems like dissolved oxygen depletion, algal blooms, cyanobacterial toxin production and biodiversity disruption (Prepas et al. 2001). A number of comparative studies have linked watershed disturbance to deteriorated water quality, with consequently increased total organic nitrogen, nitrate (NO_3^-), total phosphorus and sediment concentrations in water (Carignan et al. 2000; Martin et al. 2000; McEachern et al. 2000; Ensign and Mallin 2001; Swank et al. 2001). At the same time, terrestrial loss of nitrogen may limit forest growth, because nitrogen is a significant soil nutrient controlling forest production (Vitousek and Howarth 1991). In upland boreal forest stands in Alberta, Canada, nitrification rates were high in soils in cut plots within an aspen/conifer-mixed stand during the growing season after disturbance, and remained high into the fall and winter (Carmosini 2000). The measurable rates of net mineralization and nitrification during winter on the cut plots indicated that the plots had the potential to lose nitrogen with snowmelt in the next spring, because nitrogen may leach from the plant-rooting zone before plant uptake becomes significant (Carmosini 2000).

Modelling nitrogen composition in streams is very complex because of the difficulty in mathematically representing factors like land use and land cover, soil and vegetation nitrogen dynamics, in-stream nitrification/denitrification and meteorological parameters. These factors are complex, nonlinearly related,

---

[1] A version of this chapter has been published. Li, X., Nour, M.H., Smith, D.W., and Prepas, E.E. 2008. Modelling Nitrogen Composition in Streams on the Boreal Plain using Genetic Adaptive General Regression Neural Networks. J. Environ. Eng. Sci. 7 (S1): 109 – 125

spatially distributed on a watershed scale and temporally variable. Also, hydrological, biological and chemical processes underlying nitrogen storage, transformation and release are not well understood. Hence, the application of mechanistic models is costly and impractical, because large amounts of data are required to establish parameters and verify model performance.

In contrast, artificial neural network (ANN) models can often capture data patterns without extensive knowledge of the particular site-related problems and can model complicated and non-linear processes with fewer input variables than mechanistic models. A number of studies have shown that ANN models are superior to regression models for some applications (Cannon and Whitfield 2002; Sarangi and Bhattacharya 2005; Chandramouli et al. 2007). Since they are capable of handling large-scale and complex problems, ANN models provide great advantages in a wide range of surface water quality applications, such as modelling salinity (Bowden et al. 2005) and color (Zhang and Stanley 1997), sediment concentrations (Agarwal et al. 2005; Nour et al. 2006*a*; Cigizoglu and Kisi 2006; Tayfur and Guldal 2006; Alp and Cigizoglu 2007), phosphorus concentrations (Nour et al. 2006a, b, c), turbidity, dissolved oxygen concentrations and pH (Sahoo et al. 2006), cyanobacterial blooms (Yabunaka et al. 1997; Maier et al. 2001, 2004a; Teles et al. 2006), *Cryptosporidium* spp. and *Giardia* spp. density in river waters (Brion et al. 2001; Neelakantan et al. 2001) and non-point sources of fecal contamination (Brion and Lingireddy 1999, 2003; Brion et al. 2002).

Application of ANNs for nitrogen modelling has mostly relied on multi-layer perception neural networks with back propagation (MLP-BP) and focused on NO_3^- in agricultural settings, such as annual NO_3^- loss into drainage (Salehi et al. 2000), simulating NO_3^- leaching in drainage effluent (Kaluli et al. 1998), forecasting NO_3^- loads on a watershed based on historical data (Yu et al. 2004) and predicting NO_3^- concentration in drainage water (Sharma et al. 2003) and groundwater (Khalil et al. 2005) after application of fertilizers and manure. Lek et al. (1999) applied MLP-BP to predict total and inorganic nitrogen concentrations in streams with correlation coefficients of 0.82 and 0.80, respectively, from watershed features. In another study, using weather station data, daily streamflow and the Julian day as model inputs (but not vegetation dynamics), Suen and Eheart (2003) developed a MLP-BP model with overall accuracy of 0.8 to predict if NO_3^- concentrations in a river were greater or less than 10 mg L^{-1}.

Although BP-MLP is the most popular ANN, BP-MLP takes a large number of iterations to converge on the desired solution (Specht 1991). In contrast, another type of ANN, general regression neural network (GRNN), is very fast to train because there is only one parameter to be optimized for model development. Developed by Specht (1991), the GRNN is a multilayer feedforward neural network that performs general regression analysis from sample data for the purpose of prediction. The GRNN is able to approximate continuous functions and model nonlinear relationships. Generally, GRNNs have the advantages of converging to the optimal regression surface and dealing with sparse data very efficiently in the real-time environment (Specht 1991). In particular, genetic adaptive GRNN used in this study applies a genetic algorithm to find appropriate individual smoothing factors for each input, as well as an overall smoothing factor during calibration. The Neuroshell 2 genetic adaptive calibration method (available from Ward Systems Group Inc., Frederick, MD) generally produces models that work much better on the test data set than the iterative calibration method.

The performance and generality of a model can be significantly affected by data division into data subsets (i.e., training, testing and validation data sets). The arbitrary division of data subsets, with inadequate knowledge of the patterns associated with them, can potentially result in randomness in the developed ANN model quality and performance (Bowden et al. 2006). Therefore another type of ANN, the self-organizing map (SOM), is useful because it facilitates the production of representative data subsets to improve model performance and generality (Bowden et al. 2002; Shabin et al. 2004). A SOM can identify the regularities and similarities in its inputs. Thus, a SOM is used to classify all of the available data to several clusters, then from each of the clusters training, testing and validation data are sampled. As a result, each of the data subsets is representative and evenly distributed for all of the parameters being modelled.

Previous studies applying ANN to nitrogen modelling in a watershed have considered only nitrogen load from external sources, nitrogen leaching and seasonal factors. With the rapid development of remote sensing (RS) technology and the reduced cost of acquiring RS data, it is now possible to take into consideration vegetation phenology, one of the most important factors affecting the nitrogen cycle. The Terra spacecraft launched by the United States National Aeronautics Space Administration in December 1999 has enhanced the Earth Observing System program capabilities significantly. A sensor called the Moderate Resolution Imaging Spectroradiometer (MODIS) on board Terra has greatly improved scientists' ability to measure plant growth on a global scale, with moderate spatial and temporal resolution. The MODIS Land Group provides not only satellite data, but also high-level data products that are specifically designed to support global to regional monitoring, modelling and assessment, such as vegetation indexes (VIs) (e.g., normalized difference vegetation index (NDVI), Enhanced Vegetation Index (EVI) and leaf area index) (Justice et al. 1998; National Aeronautics and Space Administration 2007). VIs have been successfully used to detect forest disturbance (Jin and Sader 2005), monitor vegetation dynamics (Beck et al. 2006), vegetation cover and condition (Fensholt 2004; Ben-Ze'Ev et al. 2006) and forage condition (Kawamura et al. 2005). The successful application of VIs to vegetation dynamics indicates that it has potential for constructing nitrogen models and improving prediction accuracy.

In this study, a GRNN was applied to model NO_3^-, ammonium (NH_4^+) and total dissolved nitrogen (TDN) concentrations in streams draining three watersheds on the Canadian Boreal Plain, within the study area of the Forest Watershed and Riparian Disturbance (FORWARD) project. The FORWARD project is a long-term study initiated to develop hydrological and water quality models that link water quantity and quality and biological indicators with watershed management on the Boreal Plain (Smith et al. 2003). Background information and objectives of the FORWARD project can be found in the Journal of Environmental Engineering and Science (2003) (Vol. 2, Suppl. 1). Nitrogen modelling can contribute to linking water quality and disturbance indicators to watershed management. The major objectives of this study are to: (1) develop a GRNN modelling tool that is able to predict daily NO_3^-, NH_4^+ and TDN concentrations in streams using easily accessible databases; and (2) test the usefulness of RS data in constructing predictive nitrogen models.

3.2 Theory

3.2.1 General Regression Neural Network

In the GRNN approach, a dependent scalar variable (y) can be estimated from independent random variables. A vector of **x** with *p* dimensions represents the p independent random variables and **X** is a particular measured value of the random variable **x.** The joint density function $f(\mathbf{X}, y)$ is assumed to be known, then the conditional mean of y given **X** (also call the regression of y on **X**) is calculated based on the following equation:

$$[1] \quad E[y|\mathbf{X}] = \frac{\int_{-\infty}^{\infty} y f(\mathbf{X}, y)dy}{\int_{-\infty}^{\infty} f(\mathbf{X}, y)dy}$$

However, the joint density function is not known. Therefore, a probability estimator $\hat{f}(\mathbf{X}, Y)$, based on a sample of observations of **x** and y is used (Specht 1991). $\mathbf{X}^i$ and $\mathbf{Y}^i$ are the sample values of the random variables **x** and y, whereas *n* is the number of sample observations, *p* is the dimension of the vector variable **x**, *T* is matrix transpose, and σ is the smoothing factor:

[2]

$$\hat{f}(\mathbf{X},Y)=\frac{1}{(2\pi)^{(p+1)/2\cdot\sigma^{(p+1)}}}\cdot\frac{1}{n}\sum_{i=1}^{n}\exp\left[-\frac{(\mathbf{X}-\mathbf{X}^i)^T(\mathbf{X}-\mathbf{X}^i)}{2\sigma^2}\right]\cdot\exp\left[-\frac{(Y-Y^i)^2}{2\sigma^2}\right]$$

In essence, the probability estimator $\hat{f}(\mathbf{X},Y)$ assigns sample probability of width σ for each sample $\mathbf{X}^i$ and $\mathbf{Y}^i$, and the sum of those sample probabilities is the probability estimate. Substituting equation [2] into equation [1] and performing integration, Specht (1991) has shown the estimator of the conditional mean, $\hat{Y}(\mathbf{X})$, is:

[3] $$\hat{Y}(\mathbf{X})=\frac{\sum_{i=1}^{n}Y^i\exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n}\exp\left(-\frac{D_i^2}{2\sigma^2}\right)}$$

where $D_i^2=(\mathbf{X}-\mathbf{X}^i)^T(\mathbf{X}-\mathbf{X}^i)$. This regression in equation [3] can be implemented by a GRNN with four layers (Figure 3-1). The pattern layer consists of one neuron for each training pattern. The weights of the A summation layer are set to the actual output values, i.e., $A^i=Y^i$; i=1, …, n and weights of the B summation layer are unity i.e., $B^i=1$; i = 1, …, n. The output merely divides the results of the A summation by that of the B summation to provide the desired estimate of Y. To develop a GRNN with good performance, the optimal smoothing factor σ must be found.

The GRNN does not produce wild estimates outside of the range of observation, because the estimate is bounded by the minimum and maximum values of the observations (Specht 1991). On the other hand, it means that GRNN will not give reliable predictions beyond the range of the training data sets. In turn, the training data set must encompass the full range of expected output values to achieve accurate results.

3.2.2 Kohonen's Self-organizing Map Networks

SOM networks can learn to detect regularities and correlations in their inputs and order the inputs by similarity, as well as adapt their future responses to the inputs accordingly (Kohonen 1982). The SOM network generally has an input layer and a Kohonen layer. As shown in Figure 3-2, the input layer, fully connected to the Kohonen layer, provides the inputs to be ordered to the Kohonen layer. In the Kohonen layer, there is one neuron for each output category. The neurons in the Kohonen layer measure the distance of their weights to the input pattern. Based on the distance, the winning neuron is determined as follows, where *X* represents the input data with *N* inputs:

[4] $X = \{x_i; i = 1,..., N\}$

If there are *M* neurons in the Kohonen layer, each of the *M* neurons in the Kohonen layer will also have *N* weight values:

[5] $W_{ji} = \{w_{ji}; j = 1,..., M; i = 1,..., N\}$

First, it is determined how much the weights of each neuron match the corresponding input pattern. For each of the *M* Kohonen neurons, the distance, such as the Euclidean distance, is calculated as:

[6] $S_j = \|X - W_j\| = \left[\sum_{i=1}^{N} (x_i - w_{ji})^2\right]^{\frac{1}{2}}$, **j = 1,…,M.**

The winner neuron is the one having the smallest value of S_j. During the learning process, only the winner and the neurons close to the winner have their weights updated. The extent to which a neuron is close to the winner is defined by a neighborhood set. The radius of a neighborhood set can decrease as the training proceeds. After having the weights updated, the network will take the next input data pattern and the learning process continues until the stopping criteria are reached. It is the weights that provide topological information existing within the input data sets.

3.3 Study Area and Database

GRNN was applied to model nitrogen composition in streams on three forested watersheds located in the Swan Hills, northwest of Edmonton, Alberta, Canada, namely Burnt Pine (8 km^2), Willow (16 km^2) and Two Creek (129 km^2) (Figure 3-3). The Burnt Pine watershed was 100% burned in a severe wildfire in June 1998, whereas Willow and Two Creek are relatively undisturbed watersheds. The dominant soils in the study area are Luvisols, Organics, Brunisols and Gleysols, but Regosols also exist (Ecological Stratification Working Group 1996). The Boreal Plain supports mixed-wood forests, characterized by black spruce (*Picea mariana* (Mill.) BSP) and tamarack (*Larix laricina* (Du Roi) K. Koch) in poorly drained sites, and trembling aspen (*Populus tremuloides* Michx.), balsam poplar (*P. balsamifera* L.), white spruce (*Picea glauca* (Moench) Voss), lodgepole pine (*Pinus contorta* Douglas ex Louden var. *latifolia* Engelm.) and jack pine (*P. banksiana* Lamb.) in well-drained sites (Ecological Stratification Working Group 1996).

The required input data were obtained from the Whitecourt A meteorological station (Environment Canada 2007), FORWARD meteorological stations (Nour et al. 2006a,b; Prepas et al. 2008), and MODIS VIs data sets (National Aeronautics and Space Administration 2007) from the years 2002 through 2005. The data layer of MODIS EVI was exported from the MODIS VI images using the software *Geomatica* V9.1 (PCI Geomatics, Richmond Hill, ON). The exported

MODIS EVI layer was then loaded into ArcGIS 9.2 (ESRI, Redlands, CA), overlaid by the watershed shapefiles, and the corresponding EVI data for each watershed were extracted and averaged using the spatial analyst tools. Stream water quality parameters, including NO_3^-, NH_4^+ and TDN concentrations from May to October for the years 2002 through 2005, were obtained from the FORWARD data repository (Prepas et al. 2008). TDN was measured on a sample that had been filtered through GF/C Whatman glass fibre filters and within 48 hours photocombusted in a UV digester. The products of photocombustion (ammonia, nitrate, nitrite, some nitrous oxide and some nitric oxide) were passed through a zinc reduction column to be reduced to ammonia and then analyzed.The uncertainty of these measurements is shown in Table 3-1.

3.4 Model Development

To ensure the models' performance, all models were developed following the standard procedure for ANNs in a systematic way, which includes input determination, data division, determination of model internal parameters, training and stopping and model evaluation (e.g., Maier and Dandy 2000; the ASCE Task Committee 2000; Dawson et al. 2002).

3.4.1 Input Determination

In order to develop a robust ANN model, it is critical to carefully select a representative and significant set of input variables. Typically, the variables that describe the system being modelled are not equally prominent. The discrepancies in their importance are mainly due to their differences in causing changes in the modelled system. However, some input variables may be correlated, noisy or have no significant relationship with the output variable, increasing the computational complexity of the developed model (Bowden et al. 2006). Developing parsimonious models is always the objective of modelling efforts (El-Din and Smith 2002).

This study focused on data-driven means of water quality modelling. Mechanistic studies specially targeting on nitrogen dynamics in streams can be found in other literature (e.g., Chen et al. 2002; Fukuzawa et al. 2006; Oczkowski et al. 2006; Rusjan et al. 2008). Thus, model inputs in terms of cause/effect factors, time-lagged inputs and inputs reflecting seasonal cyclic nature were determined based on a combination of *a priori* knowledge of the system being modelled, cross-correlation analysis and trial-and-error screening by GRNN. On a watershed scale, the nitrogen concentrations in streams are influenced by atmospheric deposition, soil leaching and nitrogen transport along streams. These processes are very complex and are in turn influenced by land use and cover, soil and vegetation nitrogen dynamics, soil characteristics, in-stream nitrogen transformations and meteorological factors. Each of these factors is also a consequence of other more fundamental interactions. For example, nitrification and denitrification by bacteria are controlled by dissolved oxygen, alkalinity, pH, temperature and carbon source. The equilibrium between NH_4^+ and ammonia (NH_3) in water is governed by pH. Below pH 9.5, which is the case for the study streams, NH_4^+ is predominant.

Nitrogen concentrations in streams in the study area peak during early spring snowmelt (Pelster et al. 2008). The two major sources of nitrogen loading into streams after spring break-up are the snowpack and soils; this nitrogen primarily originates from atmospheric nitrogen deposition and nitrification, respectively. Total nitrogen deposition rates of approximately 420 to 2200 mg m^{-2} yr^{-1} have been estimated for western Canada, with higher rates close to urban development (Shaw et al. 1989; Kochy and Wilson 2001). A number of *in-situ* and laboratory studies have shown that nitrogen transforming processes occur under snow cover, and demonstrated that these processes should be considered when modelling nutrient leaching and concentrations in surface waters (Brooks and Williams 1999; Brooks et al. 1999). Snowpack insulates soil from the very low atmospheric temperatures in winter and thus enables microbial processes (e.g., nitrification) to occur. Consequently, NO_3^- concentrations can increase during the winter in soil under certain snow conditions. The relative contribution of atmospheric and microbial sources is determined primarily by the combined effects of air pollution, weather and watershed characteristics. The fact that nitrogen concentration peaks occur during early spring snowmelt makes snowmelt a significant indicator of this phenomenon. 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.

Based on an understanding of the processes involved in nitrogen modelling, the most significant cause/effect factors of concern are rainfall, temperature, a snowmelt indicator and a vegetation growth indicator. The cumulative degree-days (dd_t) 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.

$$[7] \quad dd_t = \sum_{i=0}^{i=N-1} \left(T_{avg(i)} - T_{b(i)}\right) \cdot \left(t_{i+1} - t_i\right)$$

Here, 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 $T_{avg} \geq T_b$, dd_t is the total degree days at time t in °C day, and $(t_{i+1} - t_i)$ is typically taken as 1 day. Vegetation growth was taken into consideration through the RS VIs, particularly EVI derived from MODIS images, which has been successful in monitoring vegetation phenology (Khan 2005). EVI is a relatively new data product developed from the MODIS Science Team to improve upon the quality of the NDVI for forested ecosystems. The atmospheric resistance was calculated by adding information from the blue wavelength and two constants, C_1 and C_2. The canopy adjustment to minimize the effect of the changes of optical properties of soil background was calculated by introducing a constant, L. EVI is formulated as represented by equation 8 (Huete and Liu 1994):

$$[8] \quad EVI = G \times \frac{\rho_{NIR} - \rho_{\mathrm{Re}d}}{\rho_{NIR} + C_1 \rho_{\mathrm{Re}d} - C_2 \rho_{Blue} + L}$$

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 to 0.52 μm), red (0.6 to 0.7 μm), and near-infrared (0.7 to 1.1 μm) wavelengths, respectively. Compared to NDVI, EVI corrects for some distortions in the reflected light caused by particles in the air, as well as the ground surface below the vegetation. In addition, it is more near-infrared reflectance sensitive than the NDVI and responds to canopy structural variations such as canopy type and architecture (Pettorelli et al. 2005).

Rainfall, snowmelt and vegetation growth probably have time-lagged effects on nitrogen concentrations in this study, because other studies indicate that environmental variables have time-lagged effects on water quality parameters (Niu et al. 1998; Nour et al. 2006a). In addition, examination of a time series plot of daily nitrogen concentrations indicated seasonal periodicity. Thus, not only cause/effect inputs but also inputs reflecting time correlation and seasonal periodicity were considered when determining the inputs for the nitrogen models. The cross correlation function (CCF) was estimated between the potential time-lagged inputs (e.g., rainfall, snowmelt, EVI) and the model outputs, and the possible time-lagged inputs were identified based on the selection criterion of 95% confidence intervals. Although the accuracy of the CCF is limited by the fact that the time series in this study are not weakly stationary, it can still provide information to determine the possible time lags of input variables. Then, the selected possible inputs using CCF were further run through GRNN. At each time, only one input was removed from GRNN inputs and the GRNN was trained and evaluated again. It was a trial-and-error method at this stage and the final parsimonious inputs to GRNN model were determined based on the GRNN's performance on the validation data set. The seasonal periodicity was accounted for by assigning Julian day of the year to each daily record. Assigning a time index to each data record has been successful in helping the ANN to identify the periodicity of data series (Gregory et al. 1991; Zhang and Stanley 1997; Sharma et al. 2003).

In summary, for all the developed models in this study, the cause/effect inputs were rainfall, daily mean air temperature, cumulative degree-days and EVI (Table 3-2). As determined through cross-correlation analysis and the trial-and-error screening by GRNN, the input having the time-lagged effect on nitrogen in the studied watersheds was cumulative degree-days and the time lags were three days before the current day. The periodicity of the model output was considered in the model development by introducing an additional input representing the Julian day of the year. The inputs were scaled linearly into open intervals between -1 and 1.

3.4.2 Data Division

Data division is also an important step in ANN model development because an ANN's performance can be significantly affected by the representativeness of subsets. Each of the subsets should represent all the patterns contained in the available data. A SOM was implemented using NeuroShell 2 (Ward Systems Group Inc., Frederick, MD) to divide the available data into training, testing and validation data sets, which were able to statistically represent the same population. To cluster the data, the inputs to the SOM network were the patterns of variables to be predicted, namely, NO_3^-, NH_4^+ and TDN concentration. The learning, neighborhood size and number of epochs were set by using default parameters and the number of clusters was five. After the clusters were formed, data sets for training, testing, and validation were randomly sampled from each of the clusters at a ratio of 3:1:1. After the three data sets were formed, further statistical analysis was performed on each of them to ensure they represented the same population. This data division method using SOMs has been tested to generate representative data sets and improve model performance and applied in the literature (Bowden et al. 2002; Maier et al. 2004b; Shabin et al. 2004). Since nitrogen concentrations in the study streams were not normally distributed, Kolmogorov-Smirnov tests were performed using MATLAB R2007a (The MathWorks Inc., Natick, MA) to verify that the three data sets represented the same population.

3.4.3 Determination of Model Architecture

The parameters to be determined are the number of neurons in the input, pattern and output layers and the smoothing factor. In fact, the number of neurons in the input layer and the output layer correspond to the number of input variables and output variables, respectively. The number of neurons in the pattern layer is usually the number of patterns in the training data set, because this layer consists of one neuron for each pattern in the training data set. The number can be set larger if more patterns are added later, but it should not be smaller than the number of training patterns. Hence, the only parameter that needs to be optimized is the smoothing factor. The smoothing factor must be greater than 0 and usually ranges from 0.01 to 1 to provide good results (required by NeuroShell 2). For all nine models developed in this study, the smoothing factor was 0.3 and not sensitive. Given the same data, this work can be easily reproduced using NeuroShell 2 by setting the only internal parameter of GRNN to 0.3.

3.4.4 Training Criteria and Stopping Criteria

During calibration, a genetic algorithm was applied to find an appropriate individual smoothing factor for each input, as well as an overall smoothing factor. Genetic algorithms basically work through selectively breeding an initial population of individuals, each of which is a potential solution to the problem, based on an objective function (Tsoukalas and Uhrig 1997). At each step, the genetic algorithm selects individuals randomly from the current population to be parents and uses them to produce the children for the next generation. Eventually, the population "evolves" toward an optimal solution through successive generations. Training using a genetic algorithm proceeded in two steps, with an objective function to minimize the mean squared error of the test data set (Ward Systems Group Inc., Frederick, MD). At first, the network was trained with the training data set. Then with the network created, a whole range of smoothing factors was tested to try to find a combination that worked best on the test data set. The learning process was terminated when an individual that improved the mean squared error by at least 1% was not produced within 20 successive generations of the whole population.

3.4.5 Model Evaluation

The developed models were evaluated based on the following criteria:

1. The coefficient of multiple determination (R^2) and root mean squared error (RMSE):

[9] $$R^2 = 1 - \frac{\sum_{i=1}^{n} (y - \hat{y})^2}{\sum_{i=1}^{n} (y - \bar{y})^2}$$

[10] $$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y} - y)^2}{n}}$$

where y is the actual value, $\hat{y}$ is the predicted value, $\bar{y}$ is the mean of the y values, and n is the number of data observations. R^2 is a statistical indicator for multivariate regression analysis. The better the model fit, the closer the R^2 is to 1.

2. Model robustness based on swapping the testing and validation data sets, followed by retraining the genetic adaptive GRNN (GA-GRNN) model and assessment of the performance of the new model. A robust model should still perform well when the testing and validation data sets are swapped.
3. Graphical examination of predicted and measured NO_3^-, NH_4^+ and TDN concentration profiles over time.

3.5 Results and Discussion

3.5.1 Case Study 1: Willow watershed

GA-GRNN models for NO_3^-, NH_4^+ and TDN concentrations were developed using inputs from the databases mentioned earlier following the model development procedure outlined above. The cumulative degree-days had a time-lagged impact on NO_3^-, NH_4^+ and TDN concentrations (Table 3-2). Table 3-3 presents the minimum, maximum, mean, period and number of data patterns for all three data subsets. The statistical measures of the GA-GRNN models' performance based on R^2 and RMSE are summarized in Table 3-4. The training data sets were well predicted (indicated by R^2 and RMSE performance data) because of the nature of the GRNN architecture (i.e., the pattern layer consists of one neuron for each training sample). Also, the GA-GRNN models demonstrated high accuracy for the testing data set, because it was used to optimize the smoothing factors. However, for the validation data sets that the models had never seen before, the R^2 values of GA-GRNN models were 0.95 for all modelled parameters (daily NO_3^-, NH_4^+ and TDN concentration) within the Willow watershed (Table 3-4). The average measurement errors for the modelled parameters were calculated by taking the average of daily measured values multiplied by their corresponding expanded uncertainty through the modelling period. The RMSEs for all of the validation data sets were 4.47, 4.97 and 26.9 µg L^{-1} and the average measurement error was 0.89, 1.55 and 37.09 µg L^{-1} for NO_3^-, NH_4^+ and TDN, respectively. The RMSEs were close to the measurement errors. The RMSEs were very low compared to the high concentrations of NO_3^-, NH_4^+ and for TDN and they were not low for the low concentrations of nitrogen. This is typical for stream water quality parameters because there is a large difference between water quality parameter values corresponding to baseflow conditions and those corresponding to snowmelt and storm events. Furthermore, these GA-GRNN models proved to be stable and consistent in predicting average daily NO_3^-, NH_4^+ and TDN concentration, because model performance was consistent even after swapping the testing and validation data sets (Table 3-4). The developed

GA-GRNN models successfully simulated most of the peak concentrations for NO_3^-, NH_4^+ and TDN (Figure 3-4a-c).

3.5.2 Case Study 2: Two Creek watershed

The Two Creek watershed is more than eight times larger in area than Willow watershed. Larger catchments typically have lower nutrient exports because they have more storage capacity for sediments and associated nutrients (Skidmore 2002). The developed models were successful in predicting average daily NO_3^-, NH_4^+ and TDN concentrations in the stream draining this larger watershed. For the validation data sets, the R^2 values of the NO_3^-, NH_4^+ and TDN models exceeded 0.90 (Table 3-4). The model performance was preserved when the testing and validation data sets were swapped, which further proved the model's power. Visualization of Figure 3-5b showed that several peak concentrations of NH_4^+ were not well modelled. Therefore, the model's prediction accuracy about these peaks was further investigated and the results shown in Table 3-5. Except for 30-Apr-03 and 28-Arp-04, all the peaks that appeared to be missed by the models for 4 years (from 2002 to 2005) were actually well modelled because the modelled results fell within or were very close to (within 2 μg L^{-1} of) the confidence intervals. The network modelled results for 30-Apr-03 and 28-Arp-04 were within about 8 μg L^{-1} and 3 μg L^{-1} of the confidence intervals, respectively. Based on these evaluations, the developed models demonstrated good simulation for peak concentrations of NO_3^-, NH_4^+ and TDN (Figure 3-5a-c). The effect of EVI on nitrogen was examined by comparing the models' performance both with and without EVI as an input. By including EVI, there was marginal enhancement in the validation data R^2 for the NO_3^- and TDN models (from 0.97 to 0.98 and from 0.91 to 0.94, respectively), while the R^2 for the NH_4^+ model was significantly enhanced (from 0.62 to 0.95).

3.5.3 Case Study 3: Burnt Pine watershed

Changes to the vegetation canopy after wildfire in the Burnt Pine watershed are captured in the EVI, because it shows the density of plant growth. Higher EVI values indicate a more dense vegetation canopy. Compared to the MODIS EVI of the Willow and Two Creek watersheds (approximately 1000 to 5000, with no annual trend from 2001 to 2005), the MODIS EVI of the Burnt Pine watershed was lower and demonstrated an increasing trend from 2001 through 2005, due to the recovery of vegetation with time since disturbance (Figure 3-6). Also, the EVI of these three watersheds showed common features in terms of seasonal vegetation growth pattern within a given year.

Average daily NO_3^-, NH_4^+ and TDN concentrations were well predicted for Burnt Pine using GA-GRNN models of very similar model architecture (Figure 3-7*a-c*). The number of hidden neurons was approximately the number of training data patterns and the initial smoothing factor was 0.3. In fact, model calibration using a genetic algorithm showed that the NO_3^-, NH_4^+ and TDN models were not sensitive to the initial smoothing factors, which is the only parameter to be determined for GRNN model architecture. Moreover, these three nitrogen models for Burnt Pine watershed had similar inputs as models for Willow and Two Creek watersheds (Table 3-2). The optimal inputs were derived from the five major parameters: rainfall, mean air temperature, cumulative degree-days, EVI and Julian day of that year, as well as related time lags (Table 3-2). Although several peak concentrations seemed to be poorly modelled (Figure 3-7b), in fact, with the consideration of measurement errors, the modelled results were close to the confidence intervals (within 5 μg L^{-1}) (Table 3-5). For all of the data (May to October from 2002 to 2005), the only peak poorly modelled was on 17-Aug-04 (Table 3-5) from the validation data set. This was probably because the training data set did not contain a pattern similar to 17-Aug-04. Data patterns that are not included by the training data set cannot be well modelled due to the nature of ANNs. However, in this study, the data division already verified that the training and validation data sets generally represented the same population. Thus, overall, the Burnt Pine nitrogen composition models predicted the base and peak concentrations of NO_3^-, NH_4^+ and TDN moderately well, with all R^2 values exceeding 0.83. The RMSEs for validation data sets were 3.40, 4.92 and 29.67 μg L^{-1} for NO_3^-, NH_4^+ and TDN, respectively, and were comparable, with corresponding measurement errors that were 1.11, 1.22 and 17.35 μg L^{-1}.

Including EVI as model input, the validation data R^2 values for the NO_3^- and TDN models were marginally enhanced (from 0.97 to 0.98 and from 0.93 to 0.96, respectively), while the R^2 for NH_4^+ was significantly enhanced (from 0.58 to 0.87). This indicates that inclusion of RS information, which provides additional information on nitrogen processes, contributed to yielding a consistent predictive relationship.

3.6 Conclusions

The factors of concern in terms of nitrogen modelling are complicated and non-linearly related, which makes ANN a suitable application for this problem. This study provided an efficient and cost-effective approach to simulate nitrogen concentrations in streams using easily accessible data as model inputs. GA-GRNN models were developed following strict procedures and applied to simulate daily mean NO_3^-, NH_4^+ and TDN concentration in streams at three watersheds in the Swan Hills of Alberta, Canada. The optimal inputs were derived from five major variables: rainfall, daily mean air temperature, cumulative degree-days, EVI and Julian day of the year. All such variables are easily accessible for the Boreal Plain of Canada. The consistent performance of GA-GRNN models for two relatively undisturbed watersheds, as well as a burned watershed, was obtained with the inclusion of the RS-derived EVI as one of the model inputs. This index was capable of describing vegetation canopy differences among watersheds, as well as vegetation phenology. In terms of model architecture, the developed models were not sensitive to the initial smoothing factor and training with a genetic algorithm improved model performance on testing data sets. The developed models successfully simulated NO_3^-, NH_4^+ and TDN concentrations for three streams, with R^2 values exceeding 0.83 for all data sets. The power of these models was demonstrated by the consistent magnitude of the performance measures achieved (R^2 and RMSE) when swapping the testing and validation data sets.

This study distinguished itself from other nitrogen modelling studies (Kaluli et al. 1998; Lek et al. 1999; Sharma et al. 2003; Yu et al. 2004; Khalil et al. 2005;

Almasri and Kaluarachchi 2005) in that it: (1) explored the water quality modelling capability of GRNNs trained with a genetic algorithm; (2) took into consideration the dynamics of vegetation phenology on nitrogen modelling by using RS data and; (3) developed GA-GRNN models that successfully predicted not only NO_3^-, but also NH_4^+ and TDN concentrations. Based on a series of model evaluations, the successful application of GA-GRNN models to predict three nitrogen constituents during dry and wet weather conditions by using five major input parameters and relevant time-lagged inputs, demonstrated the models' generality in the studied watersheds. More importantly, it implies the high potential of applying GA-GRNN models for predicting other surface water quality parameters on other similar or different watersheds. To strengthen the results even more, further investigations are needed to determine the consistency of the presented modelling approach in different geomorphological and spatial settings, and with other water quality parameters.

Table 3-1. Uncertainties of measurements for NO_3^-, NH_4^+ and TDN (Biogeochemical Analytical Laboratory 2008).

Analyte	Detection Limit ($\mu g\ L^{-1}$)	Concentration Range ($\mu g\ L^{-1}$)	Expanded Uncertainty (95% confidence interval, %)
NH_4^+	2	20 to 200	9.0
NH_4^+	2	200 to 2000	5.3
NO_3^-	1	10 to 200	7.1
NO_3^-	1	200 to 2000	5.6
TDN	7	70 to 500	9.1
TDN	7	500 to 2000	7.7
TDN	7	2000 to 6000	5.7

Note: Expanded uncertainty defines an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. The fraction is a level of confidence of the interval. To associate a specific level of confidence with the interval defined by the expanded uncertainty explicit or implicit assumptions regarding the probability distribution characterized by the measurement result and its combined standard uncertainty are required.

Table 3-2. Input parameters to nine genetic adaptive general regression neural network models.

Model	Inputs							
	R_t	T_{mean}	dd_t	$dd_{t\text{-}1}$	$dd_{t\text{-}2}$	$dd_{t\text{-}3}$	EVI_t	Julian day
Willow								
NO_3^-	Y	Y	Y	Y	Y	---	Y	Y
NH_4^+	Y	Y	Y	Y	Y	Y	Y	Y
TDN	Y	Y	Y	Y	Y	---	Y	Y
Two Creek								
NO_3^-	Y	Y	Y	Y	Y	---	Y	Y
NH_4^+	Y	Y	Y	---	---	---	Y	Y
TDN	Y	Y	Y	Y	Y	---	Y	Y
Burnt Pine								
NO_3^-	Y	Y	Y	Y	Y	---	Y	Y
NH_4^+	Y	Y	Y	Y	Y	---	Y	Y
TDN	Y	Y	Y	Y	Y	---	Y	Y

Note: R_t is rainfall in mm; T_{mean} is mean daily air temperature in degrees C; dd_t, $dd_{t\text{-}1}$, $dd_{t\text{-}2}$ and $dd_{t\text{-}3}$ are cumulative degree days at lags of 0, 1, 2, and 3 days, respectively; EVI_t is the enhanced vegetation index. Y indicates the parameter was used, '---' indicates it was not used.

Table 3-3. The minimum, maximum, mean and number of data points (*n*) for input data to the nitrogen genetic adaptive general regression neural network model.

		Rainfall (mm)*			Mean daily air temperature (°C)*			Cumulative degree days*			Enhanced Vegetation Index†		
		T‡	S_1‡	S_2‡	T	S_1	S_2	T	S_1	S_2	T	S_1	S_2
Willow	Minimum	0.00	0.00	0.00	-1.2	-0.2	-0.3	193.7	194.6	194.7	1928	1934	1936
	Maximim	42.42	14.73	19.56	23.5	23.4	21.5	2328.1	2333.6	2337.7	4966	4966	4963
	Mean	1.39	1.27	1.45	12.49	12.36	11.91	1255.0	1271.8	1283.8	3678	3703	3687
	n	383	127	127	383	127	127	383	127	127	383	127	127
Two Creek	Minimum	0.00	0.00	0.00	-1.20	-0.30	0.10	193.7	194.6	194.7	2101	2117	2122
	Maximim	22.07	33.56	26.07	23.40	22.20	23.50	2401.1	2380.9	2388.9	4493	4479	4464
	Mean	1.41	2.06	1.99	12.33	11.97	12.20	1287.3	1272.5	1294.4	3473	3451	3494
	n	391	130	129	391	130	129	391	130	129	391	130	129
Burnt Pine	Minimum	0.00	0.00	0.00	-2.90	-3.30	-4.30	82.0	97.8	104.1	145	104	58
	Maximim	29.21	16.76	22.86	23.40	21.80	23.50	2339.1	2333.6	2337.7	4846	4813	4820
	Mean	2.14	1.26	1.82	11.85	11.79	12.08	1196.5	1213.9	1205.7	2952	2994	2967
	n	413	138	138	413	138	138	413	138	138	413	138	138

* Input data on a daily basis, generally from May to October for each year of 2002, 2003, 2004 and 2005.
† Input data on a bi-weekly basis.
‡ T is the training data set; S_1 is the testing data set; S_2 is the validation data set. The data patterns are evenly distributed from May to October for each year of 2002, 2003, 2004 and 2005.

Table 3-4. Statistical measures of performance of the genetic adaptive general regression neural network models with EVI as model inputs.

Model	Measure	S_1 as testing data set			S_2 as testing data set		
		Training (T)	Testing (S_1)	Validation (S_2)	Training (T)	Testing (S_2)	Validation (S_1)
Willow							
NO_3^-	R^{2*}	0.99	0.98	0.95	0.99	0.98	0.97
	RMSE†	0.29	2.31	4.47	1.17	3.19	3.32
NH_4^+	R^2	0.98	0.97	0.95	0.98	0.96	0.95
	RMSE	2.97	3.94	4.97	2.96	4.49	5.12
TDN	R^2	0.98	0.97	0.95	0.97	0.96	0.96
	RMSE	18.34	19.61	26.90	20.18	24.74	23.14
Two Creek							
NO_3^-	R^2	0.99	0.98	0.98	0.99	0.98	0.98
	RMSE	1.89	2.54	3.17	1.62	2.51	2.78
NH_4^+	R^2	0.98	0.96	0.95	0.99	0.96	0.95
	RMSE	2.05	3.20	3.78	1.43	3.48	3.61
TDN	R^2	0.99	0.96	0.95	0.99	0.97	0.94
	RMSE	8.78	25.02	28.23	5.98	22.79	31.32
Burnt Pine							
NO_3^-	R^2	0.99	0.98	0.98	0.99	0.98	0.98
	RMSE	2.03	3.77	3.40	2.06	3.17	3.75
NH_4^+	R^2	0.98	0.93	0.83	0.97	0.92	0.87
	RMSE	1.81	3.14	4.92	1.95	3.43	4.31
TDN	R^2	0.99	0.97	0.96	0.99	0.97	0.96
	RMSE	13.56	26.13	29.67	13.61	26.49	29.15

* R^2 is the coefficient of determination.
† RMSE is root mean square error in μg L^{-1}.

Table 3-5. Investigation of the ammonium model prediction accuracy about peak concentrations for the Two Creek and Burnt Pine watersheds.

Date	Measured ($\mu g\ L^{-1}$)	Modelled ($\mu g\ L^{-1}$)	Measurement Error ($\mu g\ L^{-1}$)*	Confidence Interval ($\mu g\ L^{-1}$) †	Data Set‡
Two Creek					
29-Apr-03	56.28	50.69	5.07	51.22 to 61.35	T
30-Apr-03	67.07	53.54	6.04	61.04 to 73.11	S_2
01-May-03	59.30	53.38	5.34	53.96 to 64.64	S_1
22-Jul-03	118.44	109.74	10.66	107.78 to 129.11	T
28-Apr-04	55.05	47.17	4.95	50.10 to 60.00	S_2
13-Sep-04	103.52	94.97	9.32	94.20 to 112.84	S_2
14-Sep-04	99.28	94.19	8.93	90.34 to 108.21	S_1
Burnt Pine					
18-Jun-02	72.96	67.83	6.57	66.39 to 79.52	S_1
4-Jul-02	29.45	24.35	2.65	26.80 to 32.10	S_2
10-Jun-03	21.89	18.38	1.97	19.92 to 23.85	S_1
3-Jul-03	44.08	40.70	3.97	40.12 to 48.05	S_1
17-Aug-04	64.16	29.69	5.77	58.39 to 69.93	S_1
18-May-05	54.64	47.08	4.92	49.72 to 59.56	S_2
22-Jun-05	53.63	44.86	4.83	48.80 to 58.46	S_2

* Measurement Error = Measured × Expanded Uncertainty

† Confidence Interval = (Measured – Measurement error) to (Measured + Measurement error)

‡ T is the training data set; S_1 is the testing data set; S_2 is the validation data set.

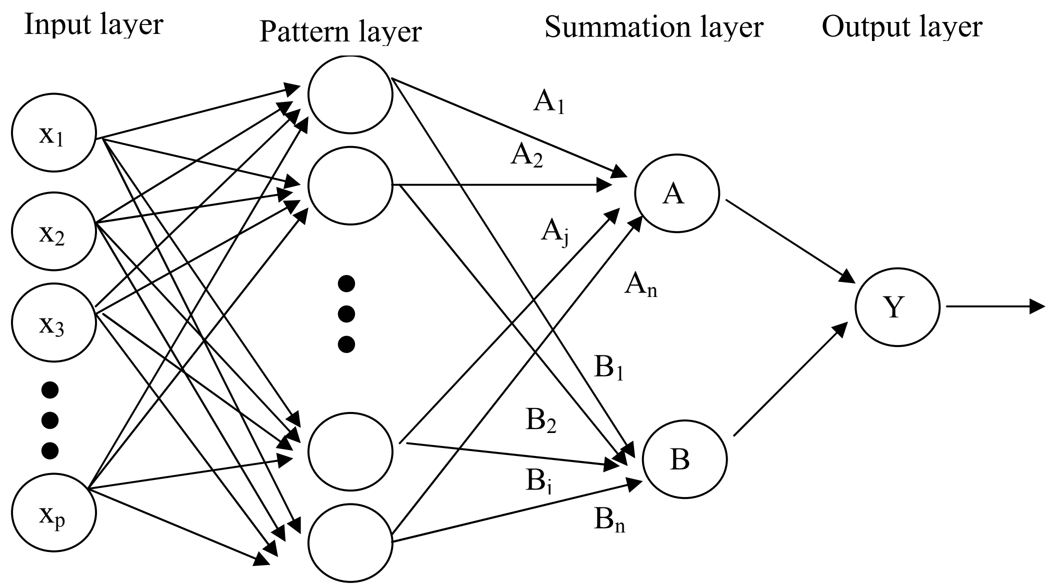


Figure 3-1. The architecture of general regression neural network models.

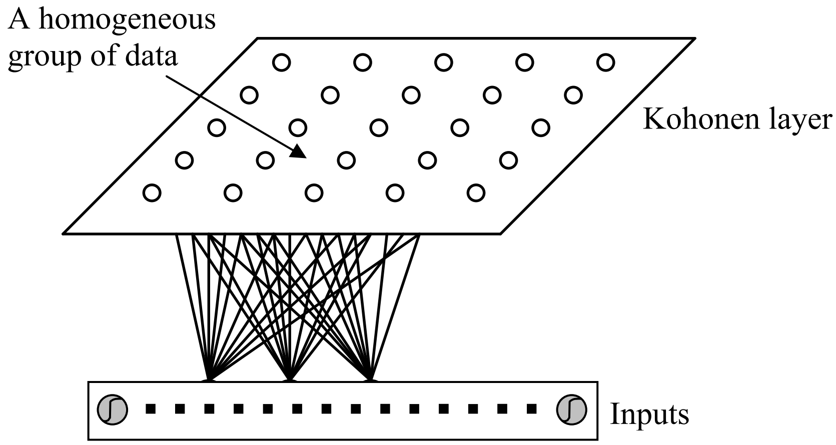


Figure 3-2. The self-organizing map, consisting of *n* inputs and a 5 by 5 Kohonen layer.

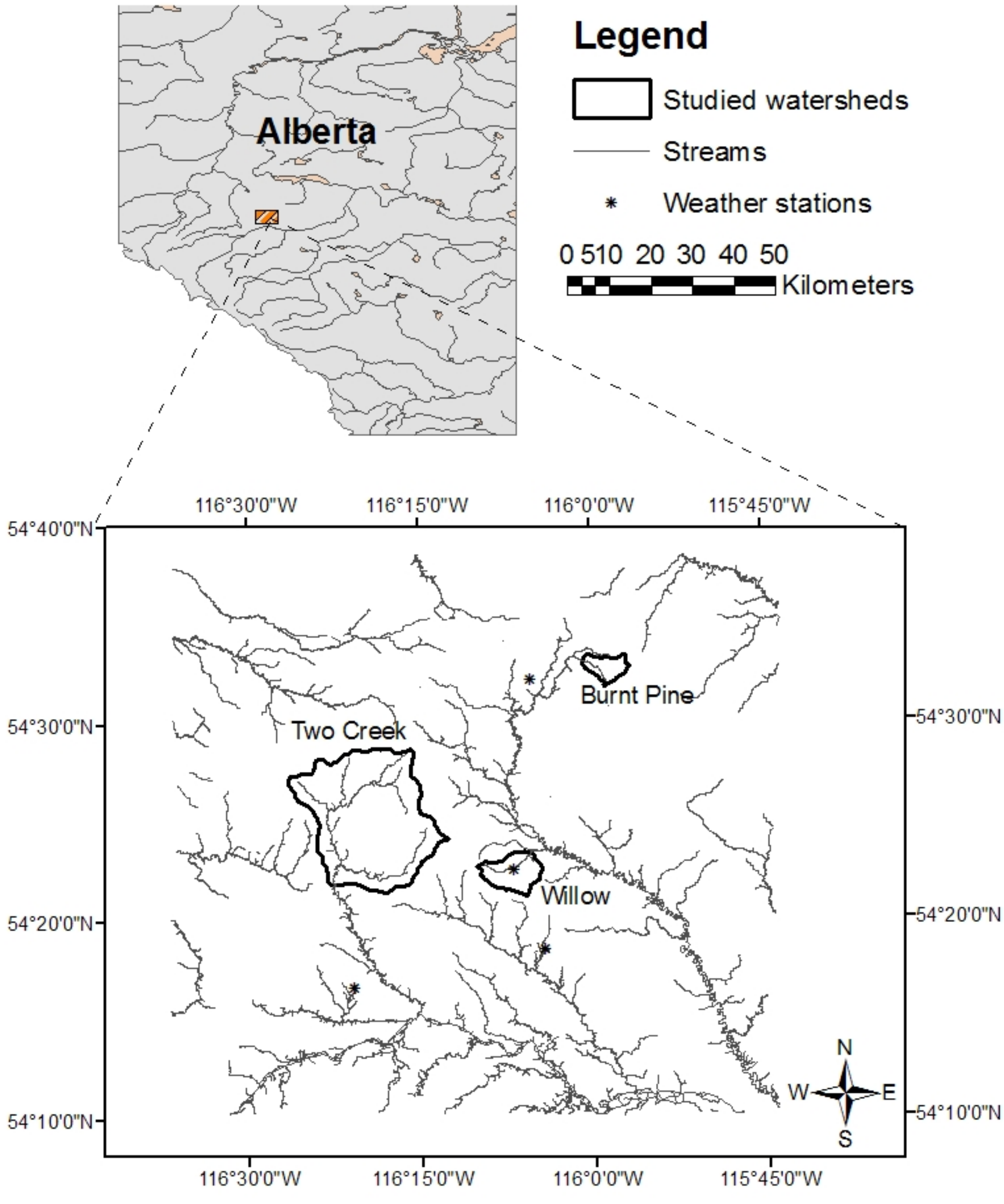


Figure 3-3. The three watersheds under study in the Swan Hills, Alberta, Canada.

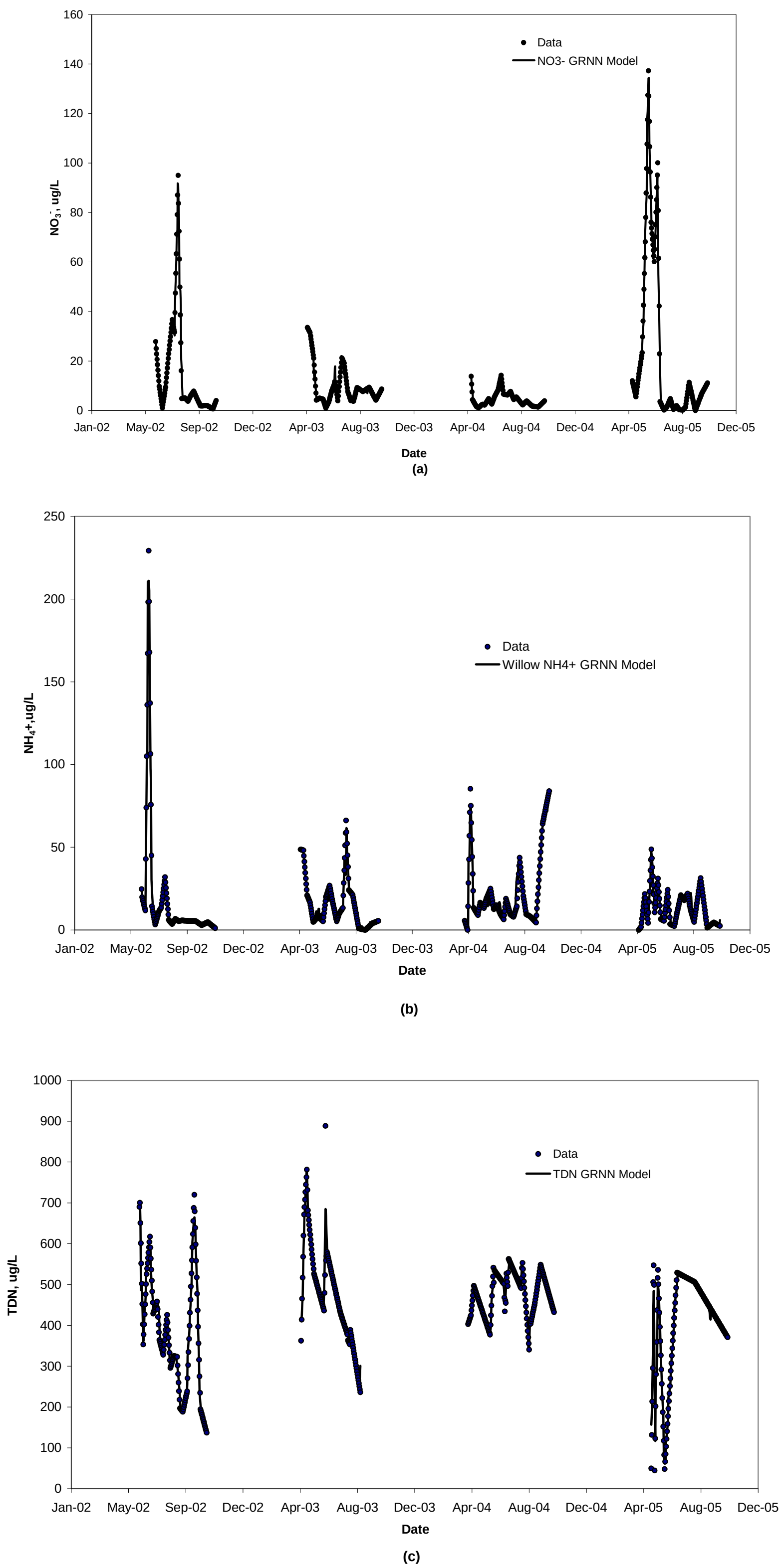
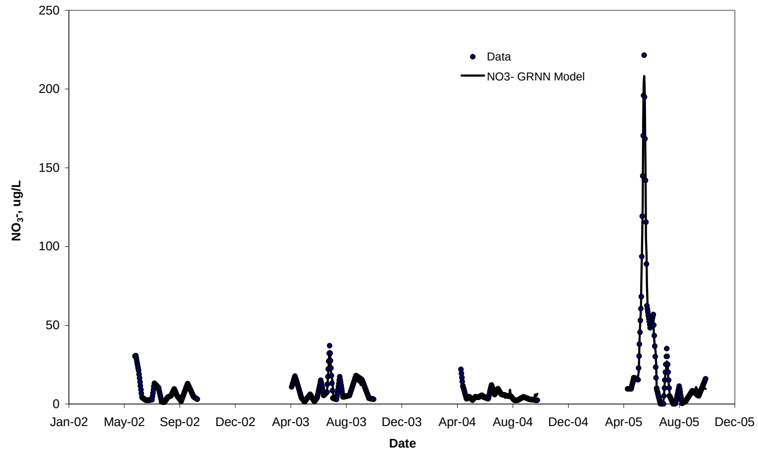
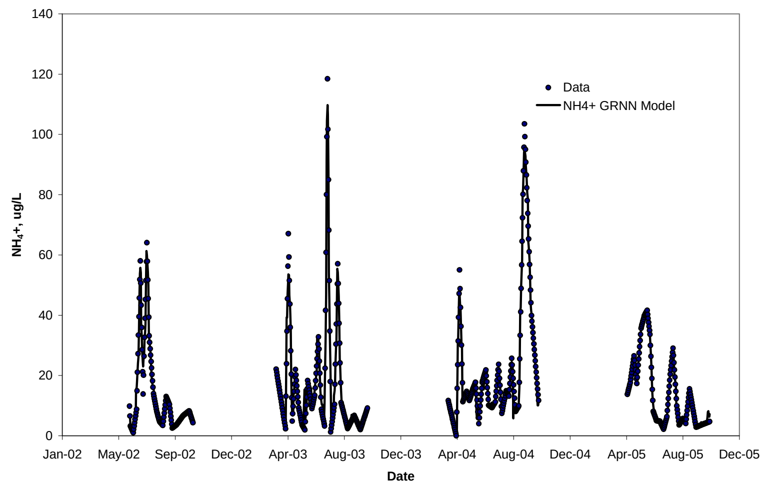


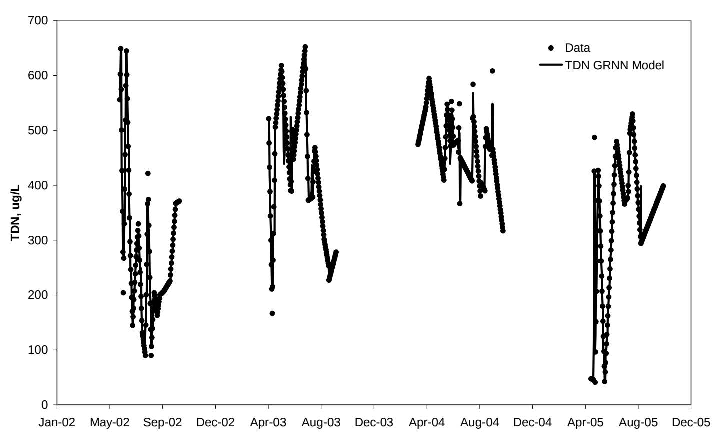
Figure 3-4. Time series plot of measured and GA-GRNN predicted concentrations of (a) NO_3^-, (b) NH_4^+, and (c) TDN in the stream draining the Willow watershed.



(a)



(b)



(c)

Figure 3-5. Time series plot of measured and GA-GRNN predicted concentrations of (a) NO_3^-, (b) NH_4^+, and (c) TDN in the stream draining the Two Creek watershed.

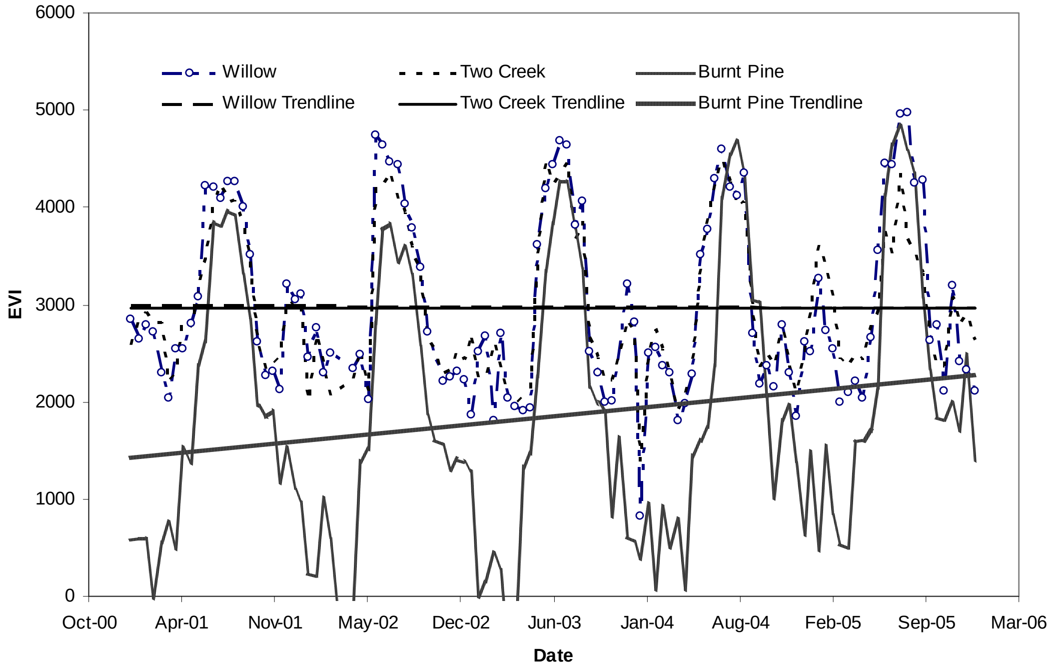
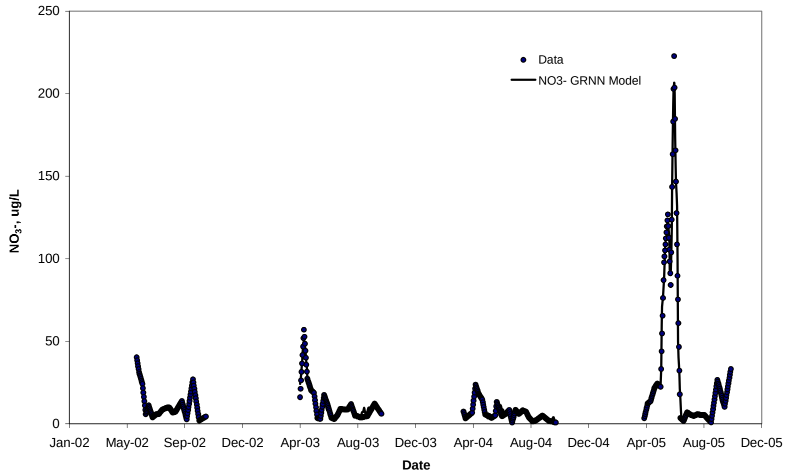
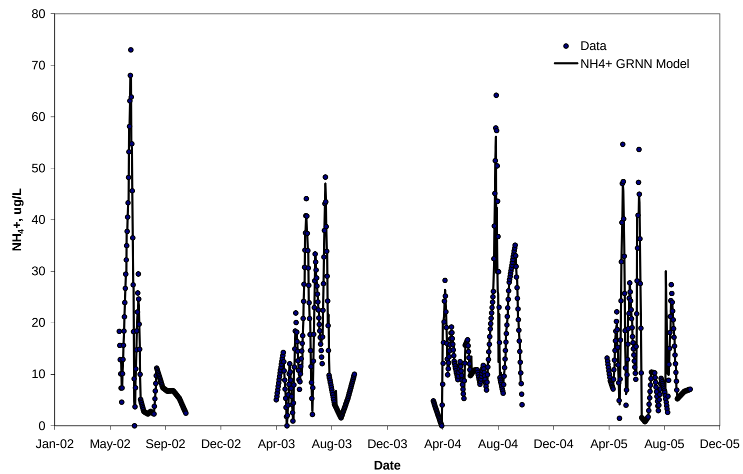


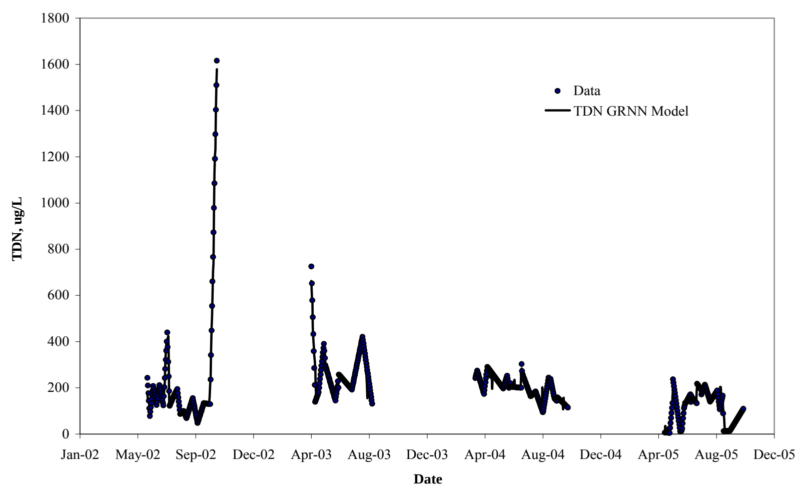
Figure 3-6. EVI of the reference watershed (Willow and Two Creek) and burned watershed (Burnt Pine). Note: The valid range of EVI is from 0 to 1. However, the data product of MODIS EVI is scaled up by a factor of 10,000 with a fill value of -3000 (Huete et al. 1999).



(a)



(b)



(c)

Figure 3-7. Time series plot of measured and GA-GRNN predicted concentrations of (a) NO_3^-, (b) NH_4^+, and (c) TDN in the stream draining the Burnt Pine watershed.

3.7 References

Agarwal, A., Singh, R.D., Mishra, S.K., and Bhunya, P.K. 2005. ANN-based sediment yield river basin models for Vamsadhara (India). Water SA **31**: 95–100.

Almasri, M.N., and Kaluarachchi, J.J. 2005. Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. Environ. Model. Soft. **20**: 851–871.

Alp, M., and Cigizoglu, H.K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ. Model. Soft. **22**: 2–13.

Beck, P.S.A., Atzberger, C., Hogda, K.A., Johansen, B., and Skidmore, A.K. 2006. Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. Remote Sens. Environ. **100**: 321–334.

Ben-Ze'ev, E., Karnieli, A., Agam, N., Kaufman, Y., and Holben, B. 2006. Assessing vegetation condition in the presence of biomass burning smoke by applying the Aerosol-free Vegetation Index (AFRI) on MODIS images. Int. J. Remote Sens. **27**: 3203–3221.

Biogeochemical Analytical Laboratory. 2008. Department of Biological Sciences, University of Alberta. Edmonton, Alberta.

Bowden, G.J. Maier, H.R., and Dandy, G.C. 2002. Optimal division of data for neural network models in water resources applications. Water Resour. Res. **38**: 1–10.

Bowden, G.J., Maier, H.R., and Dandy, G.C. 2005. Input determination for neural network models in water resources applications. Part 2. Case study: forecasting salinity in a river. J. Hydrol. **301**: 93–107.

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

Brion, G.M., and Lingireddy, S. 1999. A neural network approach to identifying non-point sources of microbial contamination. Water Res. **33**: 3099–3106.

Brion, G.M., and Lingireddy, S. 2003. Artificial neural network modelling: a summary of successful applications relative to microbial water quality. Water Sci. Technol. **47 (3)**: 235–240.

Brion, G.M., Neelakantan, T.R., and Lingireddy, S. 2001. Using neural networks to predict peak *Cryptosporidium* concentrations. J. Am. Water Works Assoc. **93**(1): 99–105.

Brion, G.M., Neelakantan, T.R., and Lingireddy, S. 2002. A neural-network-based classification scheme for sorting sources and ages of fecal contamination in water. Water Res. **36**: 3765–3774.

Brooks, P.D., and Williams, M.W. 1999. Snowpack controls on nitrogen cycling and export in seasonally snow-covered catchments. Hydrol. Process. **13**: 2177–2190.

Brooks, P.D., Campbell, D.H., Tonnessen, K.A., and Heuer, K. 1999. Natural variability in N export from headwater catchments: snow cover controls on ecosystem N retention. Hydrol. Process. **13**: 2191–2201.

Cannon, A.J., and Whitfield, P.H. 2002. Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. J. Hydrol. **259**: 136–151.

Carignan, R., D'Arcy, P., and Lamontagne, S. 2000. Comparative impacts of fire and forest harvesting on water quality on Boreal Shield Lakes. Can. J. Fish. Aquat. Sci. **57**(S2): 105–117.

Carmosini, N. 2000. Net and gross nitrogen mineralization and nitrification in upland stands of the mixedwood boreal forest following harvesting. M. Sc. Thesis. University of Alberta, Edmonton, Alta.

Chandramouli, V., Brion, G., Neelakantan, T.R., and Lingireddy, S. 2007. Backfilling missing microbial concentrations in a riverine database using artificial neural networks. Water Res. **41**: 217–227.

Chen, L. D., Fu, B. J., Zhang, S. R., Qiu, J., Guo, X. D. and Yang, F. L. 2002. A comparative study on nitrogen-concentration dynamics in surface water in a heterogeneous landscape. Environ. Geol. **42**: 424–432.

Cigizoglu, H.K., and Kisi, O. 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol. **317**: 221–238.

Dawson, C.W., Harpham, C., Wilby, R.L. and Chen, Y. 2002. Evaluation of artificial neural network techniques for flow forecasting in the river Yangtze, China. Hydrol. Earth. Sys. Sci. **6**: 619–626

Ecological Stratification Working Group. 1996. A national ecological framework for Canada. Centre for Land and Biological Resources Research, Research Branch, Agriculture and Agri-food Canada, Ottawa, Ont. 125 p.

El-Din, A.G., and Smith, D.W. 2002. A neural network model to predict the wastewater inflow incorporating rainfall events. Water Res. **36**: 1115–1126.

Ensign, S.H., and Mallin, M.A. 2001. Stream water quality changes following timber harvest in a coastal plain swamp forest. Water Res. **35**: 3381–3390.

Environment Canada. 2007. Canadian Climate Data [online]. Available from http://www.climate.weatheroffice.ec.gc.ca/climateData/canada_e.html [cited 15 September 2007]

Fensholt, R. 2004. Earth observation of vegetation status in the Sahelian and Sudanian West Africa: comparison of terra MODIS and NOAA AVHRR satellite data. Int. J. Remote Sens. **25**: 1641–1659.

Fukuzawa, K., Shibata, H., Takagi, K., Nomura, M., Kurima, T., Fukazawa, T., Satoh, H., and Sasa, K. 2006. Effects of clear-cutting on nitrogen leaching and fine root dynamics in a cool-temperate forested watershed in northern Japan. For. Ecol. Manage. **225**: 257–261

Gregory, S.V., Swanson, F.J., McKee, W.A., and Cummins, K.W. 1991. An ecosystem perspective of riparian zones. BioScience, **41**: 540–551.

Huete, A.R., and Liu, H.Q. 1994. An error and sensitivity analysis of the atmospheric- and soil-correcting variants of the NDVI for the MODIS-EOS. IEEE Trans. Geosci. Remote Sensing, **32**: 897–905.

Huete, A.R., Justice C., and Leeuwen, W.V. 1999. MODIS Vegetation Index (MOD 13). Algorithm Theoretical Basis Document, Version 3. Available from Available from http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf. [Accessed on 26 May 2008].

Jin, S.M., and Sader, S.A. 2005. MODIS time-series imagery for forest disturbance detection and quantification of patch size effects. Remote Sens. Environ. **99**: 462–470.

Justice, C.O., Vermote, E., Townshend, J.R.G., Defries, R., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.L., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Wan, Z.M., Huete, A.R., van Leeuwen, W., Wolfe, R.E., Giglio, L., Muller, J.P., Lewis, P., and Barnsley, M.J. 1998. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. IEEE Trans. Geosci. Remote Sensing, **36**: 1228–1249.

Kaluli, J.W., Madramootoo, C.A., and Djebbar, Y. 1998. Modeling nitrate leaching using neural networks. Water Sci. Technol. **38**(7): 127–134.

Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, G., and Wang, S. 2005. Monitoring of forage conditions with MODIS imagery in the Xilingol steppe, Inner Mongolia. Int. J. Remote Sens. **26**: 1423–1436.

Khalil, A., Almasri, M.N., McKee, M., and Kaluarachchi, J.J. 2005. Applicability of statistical learning algorithms in groundwater quality modeling. Water Resour. Res. 41: Art. No. W05010, 16 p.

Khan, A. 2005. Satellite characterization of vegetation dynamics in the western Canadian Boreal Plain. M. Sc. Thesis. University of Alberta, Edmonton, Alta.

Kochy, M., and Wilson, S.D. 2001. Nitrogen deposition and forest expansion in the northern Great Plains. J. Ecol. **89**: 807–817.

Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. **43**: 59–69.

Lek, S., Guiresse, M., and Giraudel, J.L. 1999. Predicting stream nitrogen concentration from watershed features using neural networks. Water Res. **33**: 3469–3478.

Maier, H.R., Burch, M.D., and Bormans, M. 2001. Flow management strategies to control blooms of the cyanobacterium, *Anabaena circinalis*, in the River

Murray at Morgan, South Australia. Regul. Rivers Res. Manage. **17**: 637–650.

Maier, H.R. and Dandy, G.C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model. Soft. **15**: 10–124.

Maier, H.R., Kingston, G.B., Clark, T., Frazer, A., and Sanderson, A. 2004a. Risk-based approach for assessing the effectiveness of flow management in controlling cyanobacterial blooms in rivers. River Res. Appl. **20**: 459–471.

Maier, H.R., Morgan, N., and Chow, C.W.K. 2004b. Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. Environ. Modell. Softw. **19**: 485–494.

Martin, C.W., Hornbeck, J.W., Likens, G.E., and Buso, D.C. 2000. Impacts of intensive harvesting on hydrology and nutrient dynamics of northern hardwood forests. Can. J. Fish. Aquat. Sci. **57**(S2): 19–29.

McEachern, P., Prepas, E.E., Gibson, J.J., and Dinsmore, W.P. 2000. Forest fire induced impacts on phosphorus, nitrogen, and chlorophyll *a* concentrations in boreal subarctic lakes of northern Alberta. Can. J. Fish. Aquat. Sci. **57**(S2): 73–81.

National Aeronautics and Space Administration. 2007. Data Gateway Interface [online]. Available from: http://modis.gsfc.nasa.gov/ [cited 15 June 2007]

Neelakantan, T.R., Brion, G.M., and Lingireddy, S. 2001. Neural network modelling of *Cryptosporidium* and *Giardia* concentrations in the Delaware River, USA. Water Sci. Technol. **43**(12): 125–132.

Niu, X.F., Edmiston, H.L., and Bailey, G.O. 1998. Time series models for salinity and other environmental factors in the Apalachicola estuarine system. Estuar. Coast. Shelf Sci. **46**: 549–563.

Nour, M.H., Smith, D.W, Gamal El-Din, M., and Prepas, E.E. 2006a. 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**: 19–32.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006b. Neural networks modelling of streamflow, phosphorus, and suspended solids: application to the Canadian Boreal forest. Water Sci. Technol. **53**(10): 91–99.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006c. Artificial neural networks and time series modelling of TP concentration in boreal streams: A comparative approach. J. Environ. Eng. Sci. **5**(S1): 39–52.

Oczkowski, A.J., Pellerin, B.A., Hunt, C.W., Wollheim, W.M., Vorosmarty, C.J., and Loder, T.C. 2006. The role of snowmelt and spring rainfall in inorganic nutrient fluxes from a large temperate watershed, the Androscoggin River basin (Maine and New Hampshire). Biogeochemistry. **80**: 191–203

Pelster, D.E., Burke, J.M., and Prepas, E.E. 2008. Runoff and inorganic nitrogen export from Boreal Plain watersheds six years after wildfire and one year after harvest. J. Environ. Eng. Sci. **7**: S51-S61.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J., and Stenseth, N.C. 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol. Evol. **20**: 503–510.

Prepas, E.E., Pinel-Alloul, B., Planas, D., Méthot, G., Paquet, S., and Reedyk, S. 2001. Forest harvest impacts on water quality and aquatic biota on the Boreal Plain: Introduction to the TROLS lake program. Can. J. Fish. Aquat. Sci. **58**: 421–436.

Prepas, E.E., Smith, D.W., Putz, G. and Germida, J.J. 2008. FORWARD database, The Forestry Corp., Edmonton, Alta.

Rusjan, S., Brilly, M., and Mikos, M. 2008. Flushing of nitrate from a forested watershed: An insight into hydrological nitrate mobilization mechanisms through seasonal high-frequency stream nitrate dynamics. J. Hydrol. **354**: 187–202

Sahoo, G.B., Ray, C., and De Carlo, E.H. 2006. Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii. J. Hydrol. **327**: 525–538.

Salehi, F., Prasher, S.O., Amin, S., Madani, A., Jebelli, S.J., Ramaswamy, H.S., Tan, C., and Drury, C.F. 2000. Prediction of annual nitrate-N losses in drain outflows with artificial neural networks. Trans. ASAE, **43**: 1137–1143.

Sarangi, A., and 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.

Shabin, M.A., Maier, H.R., and Jaksa, M.B. 2004. Data division for developing neural networks applied to geotechnical engineering. J. Comput. Civil. Eng. 18: 105–114.

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

Shaw, R.D., Trimbee, A.M., Minty, A., Fricker, H., and Prepas, E.E. 1989. Atmospheric deposition of phosphorus and nitrogen in central Alberta with emphasis on Narrow Lake. Water Air Soil Poll. **43**: 119–134.

Skidmore, A.K. 2002. Taxonomy of environmental models in the spatial sciences. *In* Environmental Modelling with GIS and Remote Sensing. *Edited by* A.K. Skidmore. Taylor & Francis Inc., New York. pp. 8–23.

Smith, D.W., Prepas, E.E., Putz, G., Burke, J.M., Meyer, W.L., and Whitson, I.R. 2003. 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** (S1): 1–13.

Specht, D.F. 1991. A general regression neural network. IEEE Trans. Neural Netw. **2**: 568–576.

Suen, J.P., and Eheart, J.W. 2003. Evaluation of neural networks for modeling nitrate concentrations in rivers. J. Water Resour. Plan. Manage. ASCE, **129**: 505–510.

Swank, W.T., Vose, J.M., and Elliott, K.J. 2001. Long-term hydrologic and water quality responses following commercial clearcutting of mixed hardwoods on a southern Appalachian catchment. For. Ecol. Manage. **143**: 163–178.

Tayfur, G., and Guldal, V. 2006. Artificial neural networks for estimating daily total suspended sediment in natural streams. Nord. Hydrol. **37**: 69–79.

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

The ASCE Task Committee. 2000. Artificial neural networks in hydrology I: preliminary concepts. J. Hydrol. Eng. **5**: 115–123.

Tsoukalas, L.H., and Uhrig, R.E. 1997. Fuzzy and neural approaches in engineering. Wiley-Interscience, New York.

Vitousek, P.M., and Howarth, R.W. 1991. Nitrogen limitation on land and in the sea: how can it occur. Biogeochemistry, **13**: 87–115.

Yabunaka, K. Hosomi, M. and Murakami, A. 1997. Novel application of a back-propagation artificial neural network model formulated to predict algal bloom. Water Sci. Technol. **36**: 89–97.

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

Zhang, Q., and Stanley, S.J. 1997. Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modeling. Water Res. **31**: 2340–2350.

CHAPTER 4. NEURAL NETWORKS MODELLING OF NITROGEN EXPORT: MODEL DEVELOPMENT AND APPLICATION TO UNMONITORED BOREAL FOREST WATERSHEDS[1]

4.1 Introduction

Water quality modeling of 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 modeling complicated and non-linear processes; therefore they have gained popularity in water quality modeling applications over the past decade. Additional features of ANN models that contribute to their utility for surface water quality modeling are (ASCE Task Committee 2000): (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 numerical models, especially when only sparse, gapped data are available for model training (Abrahart and See 2000; Srivastava et al. 2006). Conventional statistical modelling approaches (e.g., multivariate linear regression and time series modelling) make assumptions about the system under study and develop equations to describe the problem to achieve statistical optimality. In contrast, ANN practitioners do not need to make such assumptions (Amari et al. 1994). ANN models have also proven to be superior to traditional statistical time series models by a number of case studies (e.g., Chang et al. 2004; Kisi 2004; Nour et al.

[1] A version of this chapter was accepted with revisions by *Environmental Technology*.

2006b; Bowden et al. 2006), because they can handle non-linear problems and non-stationary data sets (Chang et al. 2004; Nour et al. 2006b) that are often the cases for environmental variables. A number of case studies have documented the superiority of the ANN modelling approach over other statistical techniques (e.g., Chang et al. 2004; Kisi 2004; Nour et al. 2006b; Bowden et al. 2006). Compared to numerical models, ANN models can provide comparable modelling accuracy but are more applicable in practice when professional expertise and data are limited. A number of case studies selected ANN models because they can satisfy the modelling objectives by using only routine monitoring data (e.g., Brion et al. 2001; Nour et al. 2006a; Li et al. 2008).

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 (Aber et al. 2003). 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 nonlinearly 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 trained with the error back-propagation algorithm (MLP-BP) is the most widely used neural network for water quality modelling (Zealand et al. 1999; Maier and Dandy 2000). MLP has been used to model a variety of water quality parameters such as sediment loads (Cigizoglu 2004; Cigizoglu and Kisi 2006; Tayfur and Guldal 2006), phosphorus concentrations (Lek et al. 1996; Nour et al. 2006a, 2006b), microbial contamination (Brion et al. 2001; Neelakantan et al. 2001; Brion and Lingireddly 2003), and phytoplankton (including cyanobacteria) communities (Recknagel 1997; Maier et al. 1998, 2001; Hou et al. 2006; Recknagel et al. 2006; Teles et al. 2006). In terms of N modelling, MLP-BP has been applied to simulate nitrate leaching in agricultural drainage effluent (Kaluli et al. 1998; Sharma et al. 2003), forecast nitrate loads on an agricultural watershed based on historical data (Yu et al. 2004), predict total and inorganic N concentrations in 927 streams in the United States from features in mixed-use (forested, agricultural and urban) watersheds (Lek et al. 1999) 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 (Clair and Ehrman 1996).

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. 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 (Smith et al. 2003; Prepas et al. 2008). 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 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.

4.2 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 to 129.4 km^2

(Table 4-1, Figure 4-1). The peatland and riparian cover data summarized in Table 4-1 were previously documented (Prepas et al. 2006). Rainfall and temperature data were acquired from seven public-domain weather stations and fire towers close to the study area (Figure 4-1). 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 (PCI Geomatics, Richmond Hill, ON). The exported MODIS images were then loaded into ArcGIS 9.2 (ESRI, Redlands, CA), overlaid by the watershed shape files, and the corresponding 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 (Prepas et al. unpubl. data) 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. (2006). The areal TDN export (TDN_E) ($g/km^2/d$) for each watershed was calculated as indicated by Equation 1:

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

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.

4.3 MLP Model Development

To ensure adequate model performance, all models were systematically developed following the standard procedures for ANN model building, including input determination, data pre-processing, data division, determination of model internal parameters, selection of the model training algorithm and stopping criterion and model evaluation (ASCE Task Committee 2000; Maier and Dandy 2000; Nour at al. 2006b).

4.3.1 Input Determination

To develop a robust ANN model, it is critical to carefully select a representative and significant set of input variables. 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 (e.g., Creed and Band 1998; de Wit and Behrendt 1999). Generally speaking, TDN export is highly correlated with time and has seasonal fluctuation. Also, the influence of factors like 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 (Hatano et al. 2005; Inamdar et al. 2006; Ide et al. 2007). 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 Equation 2, 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.

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

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 (Clair and Ehrman 1996), particularly in the study area (Prepas et al. 2006).

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 (Hatano et al. 2005). 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 (Pelster et al. 2008) and the uptake of N moving with snowmelt runoff by forest vegetation is minimal (Lauren et al. 2005). 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 (250 m x 250 m pixel size) and temporal resolution. The RS EVI provided by the MODIS Land Group has shown a high correlation with vegetation conditions (Ferreira and Huete 2004; Kawamura et al. 2005; Cheng et al. 2006; Silveira et al. 2008). 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 Equation 3:

[3] $$EVI = G \times \frac{\lambda_{NIR} - \lambda_{\mathrm{Re}d}}{\lambda_{NIR} + C_1 \lambda_{\mathrm{Re}d} - C_2 \lambda_{Blue} + L}$$

where $G = 2.5$, $C_1 = 6$, $C_2 = 7.5$, and $L = 1$. The terms λ_{Blue}, λ_{Red}, and λ_{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 (Li et al. 2008). 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 cyclisity, 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 (Gregory et al. 1991; Zhang and Stanley 1997; Sharma et al. 2003).

4.3.2 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 Equations 4 to 6:

$$[4] \quad r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

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

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

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.

4.3.3 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. Complete data records were divided into independent data sets for training, testing and validation at a ratio of 3:1:1. After data division, Kolmogorov-Smirnov tests were performed using MATLAB R2007a (The MathWorks Inc., Natick, MA) to verify that the three data subsets represented the same population because 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 model had never seen, was then used to test model generality.

4.3.4 Model Architecture

At this step, it is important to determine the number of hidden layers, the number of nodes in the hidden layers and the type of activation function(s). A typical MLP ANN with a single hidden layer can model most applications. However, the single hidden layer with only one activation function did not produce acceptable N export results during preliminary model runs. Nitrogen export is highly correlated to streamflow, which is influenced by snowmelt in spring and by rainfall in summer and autumn. One hidden layer with more than one activation function has been successful in capturing the different driving forces for streamflow in the study area (Nour et al. 2006b). Thus, the same architecture was used in this application as well. The optimum number of nodes in the hidden layer and the activation functions were determined using the systematic approach presented in Maier and Dandy (2000) and Nour et al. (2008).

4.3.5 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 (Ward Systems Group 1996). 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.

4.3.6 Model Evaluation

There is not a 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 (Legates and McCabe 1999). 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.

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

[8] $$E = 1 - \frac{\sum_{i=1}^{N} (O_i - P_i)^2}{\sum_{i=1}^{N} (O_i - \overline{O})^2}$$ <-∞, 1.0>

[9] $$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| O_i - P_i \right|$$

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

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 (Legates and McCabe 1999). 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.

4.4 Model Performance

The daily TDN export of five watersheds was modelled using the above noted ANN modelling algorithm. Table 4-2 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 4-3. 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 4-4) and by examining the time series plot of the measured and modelled TDN export profiles. The 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 4-2). The MAE values for all data subsets were small compared to the peaks of TDN export. The peaks of TDN export occurred in spring during snowmelt and in summer during heavy rainfall events (Figure 4-2). 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 training data set did not contain enough data patterns similar to the ones to be predicted (2002 was the driest of the 5 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.

4.5 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.

4.5.1 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). The more similar is each pair of watersheds, 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 (Gan and Burges 2006).

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 (Everitt et al. 2002; Shanmugam et al. 2006). 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 (Zhu and Mazumder 2008) and peatland cover was positively associated with ammonium (a fraction of TDN) export in the study watersheds (Prepas et al. 2006). 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 (Luke et al. 2007). In general, riparian areas have the potential to reduce excess N export into surface water (Willems et al. 1997; Cey et al. 1999; Baker et al. 2001). 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 (Chen et al. 2005). The NDWI was calculated using the near- (λ_{NIR}) and mid-infrared (λ_{MIR}) frequency ranges downloaded from MODIS (Equation 11).

$$[11] \quad NDWI = \frac{\lambda_{NIR} - \lambda_{MIR}}{\lambda_{NIR} + \lambda_{MIR}}$$

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 analyzed, 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:

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

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

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

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

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

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 4-5.

4.5.2 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 data. The statistical measures of model performance when the calibrated models were transferred to different watersheds were calculated (Table 4-6). 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 4-3 (a)). The peaks were predicted fairly well for five years except for 2002 and 2005 when the Willow had lower peak concentrations than the Thistle. As a result of the nature of data-driven model, the Willow model, which were trained using based on monitoring data from Willow, produced underestimations when applied to Thistle. 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 4-1). They are also very similar in terms of rainfall and vegetation dynamics (Table 4-5).

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 4-3 (b)), probably because of its high percentage of peatland coverage (Table 4-1), 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 because 5 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 4-4. Among the individual watershed indices, rainfall_index explained the most variation in E (r^2

= 0.71, P ≤ 0.05), followed by NDWI_Index (r^2 = 0.69, P ≤ 0.05). Peatland_index and EVI were related to *E*, with r^2 values of 0.45 (P ≤ 0.05) and 0.44 (P = 0.06), respectively. The riparian index did not provide an important measure of watershed similarity in this study. 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 ≤ 0.05) (Figure 4-4 (f)). If the three most important factors were considered, their contribution to the variation in *E* was marginally improved to r^2 = 0.74 (P ≤ 0.05) (Figure 4-4 (g)).

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*: $Peatland\_index_{i\text{-}j} + Rain\_index_{i\text{-}j} + NDWI\_Index_{i\text{-}j}$, (3) select the most similar monitored watershed *k* to the unmonitored watershed, which have 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.

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.

4.6 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 5 to 130 km^2 in area demonstrated the success of the ANN modelling approach in predicting daily TDN using public-domain, readily available data with reasonable accuracy, 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 to 0.62 and 0.44 to 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 towards expanding and refining hydrologic similarity indices relies on free-of-cost RS information.

Table 4-1. 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

Table 4-2. 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}, Tmean, 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}, Tmean, 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}, Tmean, 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},Tmean, 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}, Tmean, 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, and 3 days, respectively; *Tmean* is mean daily air temperature in degree C; *ddt*, *ddt*-1, *ddt*-2 and *ddt*-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 4-3. 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 weights	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, which means the weights are updates after training proceeds through the entire patterns in the training data set.

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

Measures	Willow Model			Cassidy Model			Two Creek Model			Thistle Model			1A		
	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.

Table 4-5. 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 4-6. 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

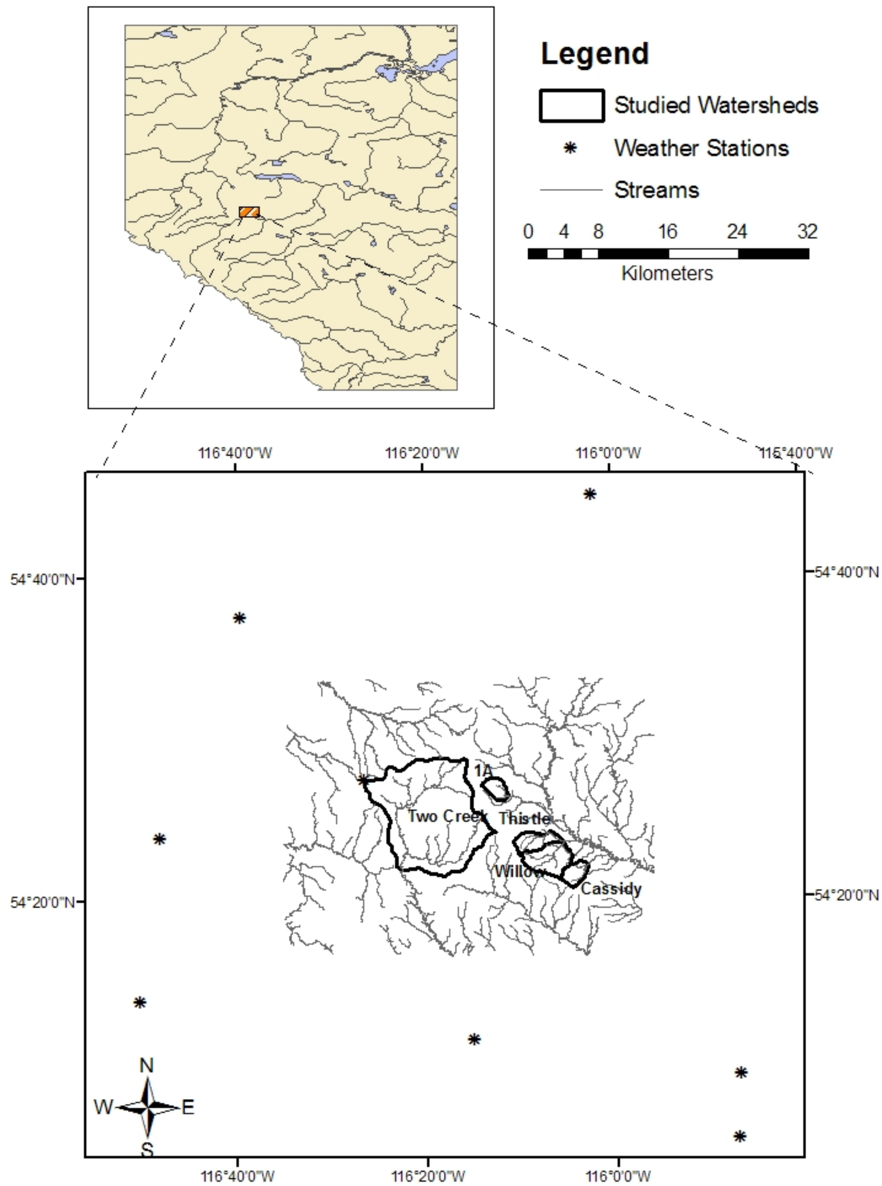
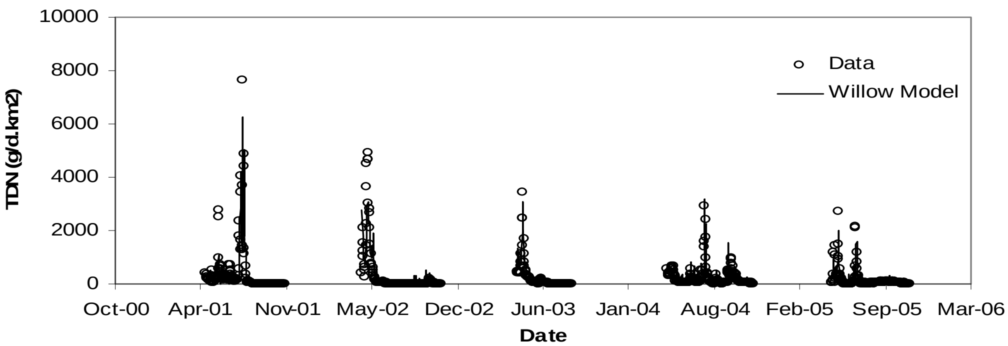
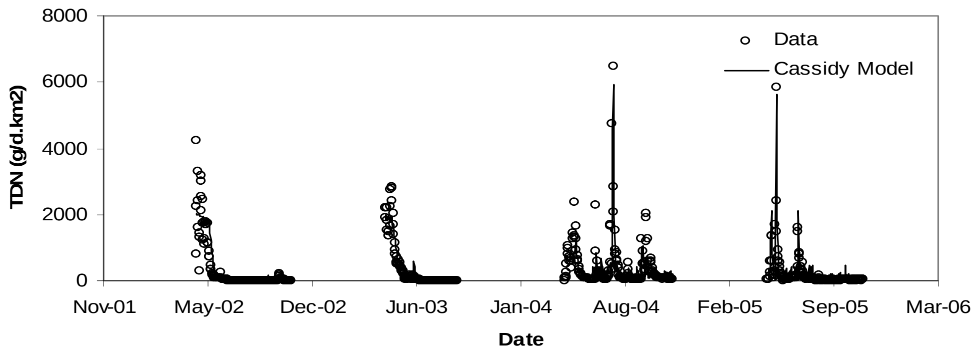


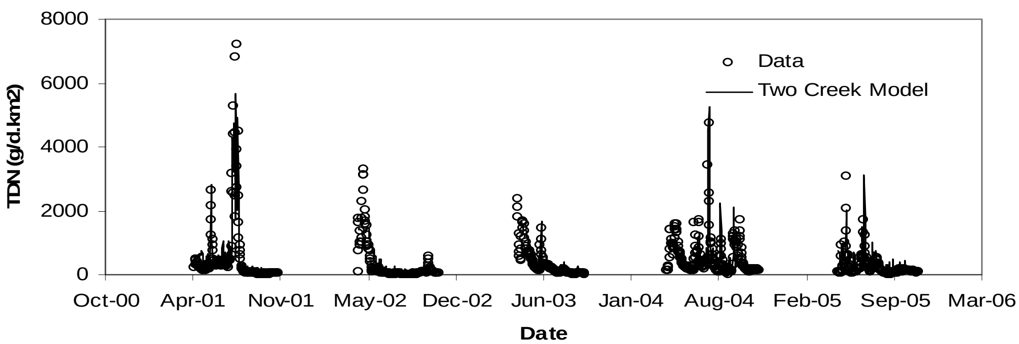
Figure 4-1. Study area: the watersheds under study and the weather stations.



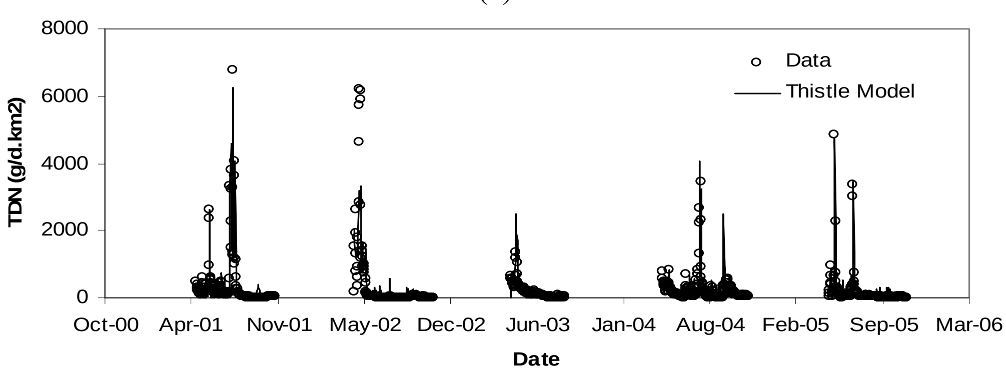
(a)



(b)

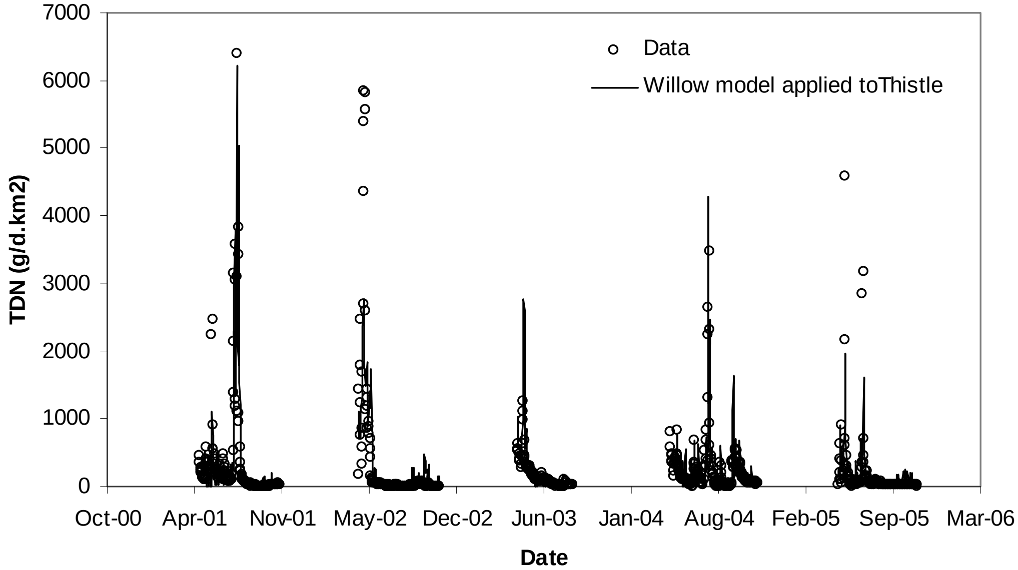


(c)

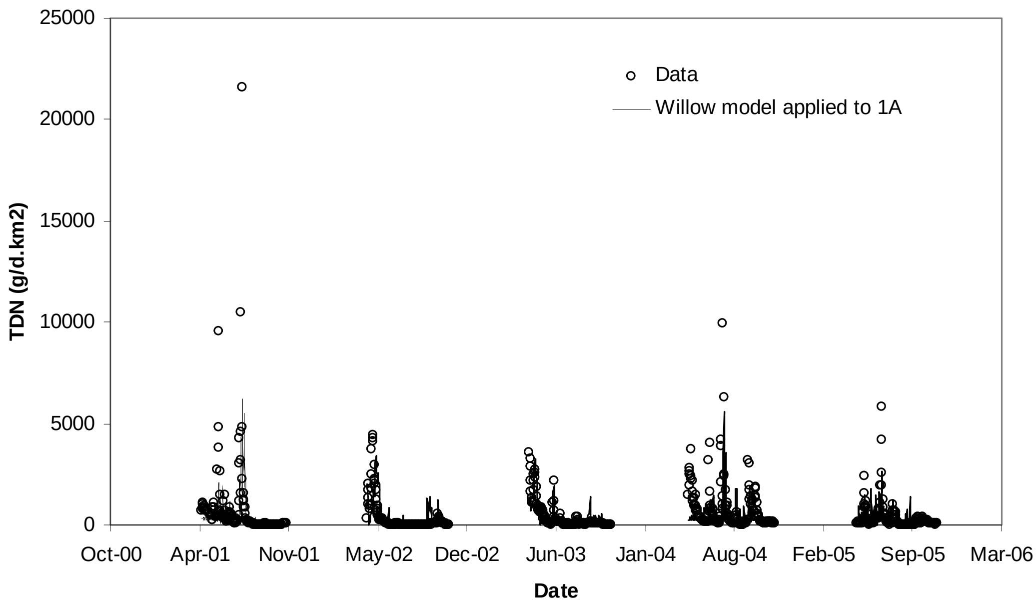


(d)

Figure 4-2. Time series plot of measured versus modelled daily TDN export for (a) Willow, (b) Cassidy, (c) Two Creek, and (d) Thistle watershed.



(a)



(b)

Figure 4-3. Time series plot of measured versus modelled daily TDN export (a) when Willow model applied to Thistle, and (b) Willow model applied to 1A.

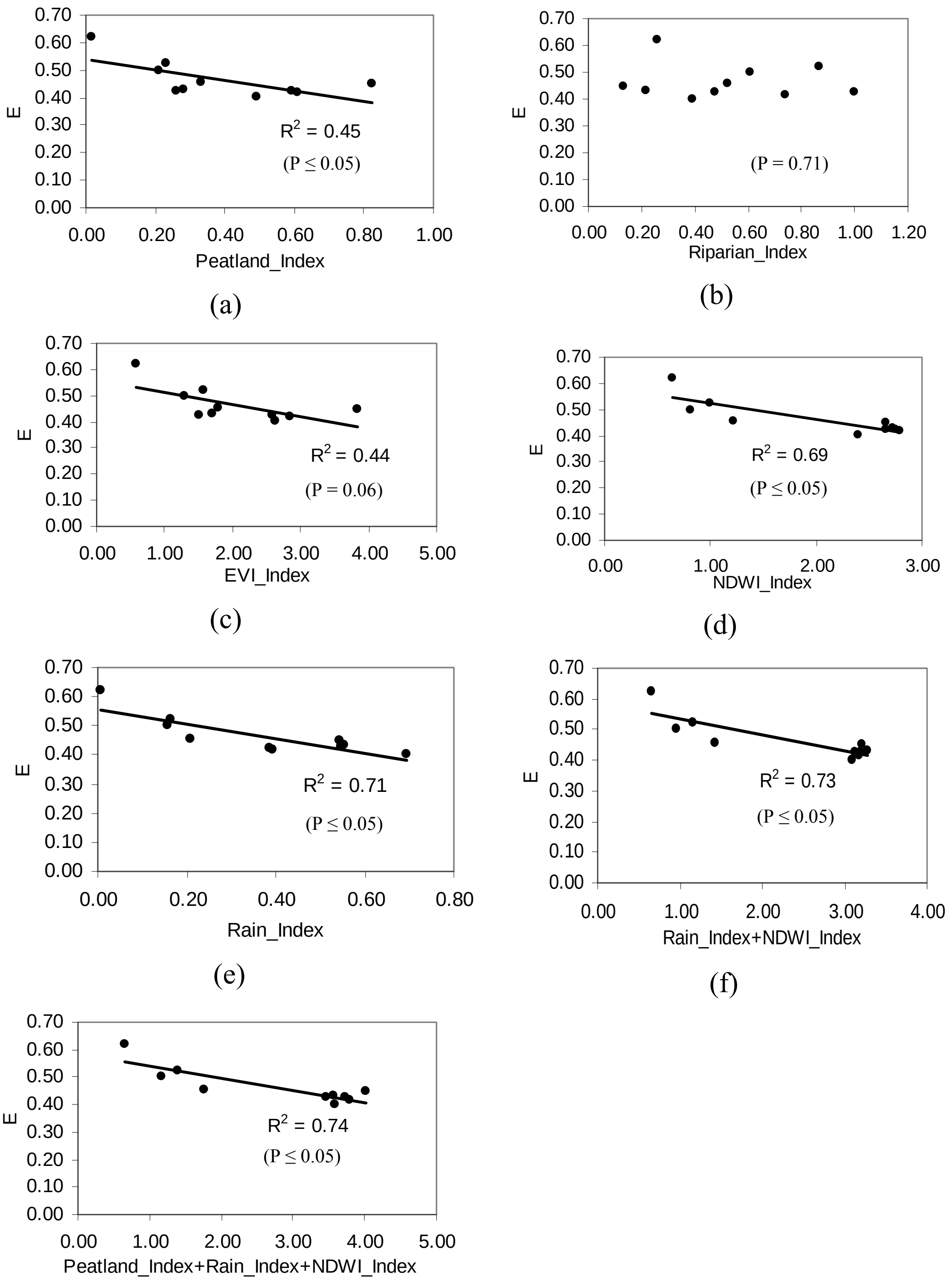


Figure 4-4. Plot of model transferability performance measure *E* versus watershed similarity indices.

4.7 References

Aber, J.D., Goodale, C.L., Ollinger, S.V., Smith, M.L., Magill, A.H., Martin, M.E., Hallett, R.A., and Stoddard, J.L. 2003. Is nitrogen deposition altering the nitrogen status of northeastern forests? BioScience. **53**: 375-389.

Abrahart, R.J. and See, L. 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrol. Process. **14**: 2157-2172.

Amari, S., Barron, A.R., Bienenstock, E., Geman, S., Breiman, L., McClelland, J.L., Ripley, B.D., Tibshirani, R., Cheng, B., and Titterington, D.M. 1994. Neural networks-a review from statistical perspective-comments and rejoinders. Stat. Sci. **9**: 31-54.

Baker, M.E., Wiley, M.J., and Seelbach, P.W. 2001. GIS-based hydrologic modeling of riparian areas: Implications for stream water quality. J. Am. Water Resour. Assoc. **37:** 1615-1628.

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

Brion, G.M., and Lingireddy, S. 2003. Artificial neural network modelling: a summary of successful applications relative to microbial water quality. Water Sci. Technol. **47 (3):** 235-240.

Brion, G.M., Neelakantan, T.R., and Lingireddy, S. 2001. Using neural networks to predict peak Cryptosporidium concentrations. J. Am. Water Works Assoc. **93 (1)**: 99-105.

Cey, E.E., Rudolph, D.L., Aravena, R., and Parkin, G. 1999. 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**: 45-67.

Chang, L.C., Chang, F.J., and Chiang, Y.M. 2004. A two-step-ahead recurrent neural network for stream-flow forecasting. Hydrol. Process. **18**: 81-92

Chen, D.Y., Huang, J.F., and Jackson, T.J. 2005. Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. Remote Sens. Environ. **98**: 225-236.

Cheng, Y.B., Zarco-Tejada, P.J., Riano, D., Rueda, C.A., and Ustin, S.L. 2006. Estimating vegetation water content with hyperspectral data for different canopy scenarios: Relationships between AVIRIS and MODIS indexes. Remote Sens. Environ. **105:** 354-366.

Cigizoglu, H.K. 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. Adv. Water Resour. **27:** 185-195.

Cigizoglu, H.K., and Kisi, O. 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol. **317:** 221-238.

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

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

de Wit, M., and Behrendt, H. 1999. Nitrogen and phosphorus emissions from soil to surface water in the Rhine and Elbe basins. Water Sci. Technol. **39 (12):** 109-116.

Everitt, J.H., Yang, C., Escobar, D.E., Lonard, R.I., and Davis, M.R. 2002. Reflectance characteristics and remote sensing of a riparian zone in south Texas. Southw. Natural. **47**: 433-439.

Ferreira, L.G., and Huete, A.R. 2004. Assessing the seasonal dynamics of the Brazilian Cerrado vegetation through the use of spectral vegetation indices. Int. J. Remote Sens. **25:** 1837-1860.

Gan, T.Y., and Burges, S.J. 2006. Assessment of soil-based and calibrated parameters of the Sacramento model and parameter transferability. J. Hydrol. **320:** 117-131.

Gregory, S.V., Swanson, F.J., McKee, W.A., and Cummins, K.W. 1991. An ecosystem perspective of riparian zones. BioScience. **41:** 540-551.

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

Haykin S. 1994. Neural networks: a comprehensive foundation. Macmillan College Publishing, NY.

Hou, G.X. Li, H.B., Recknagel, F., and Song, L.R. 2006. Modeling phytoplankton dynamics in River Darling (Australia) using the radial basis function neural network. J. Freshw. Ecol. **21:** 639-647.

Ide, J., Nagafuchi, O., Chiwa, M., Kume, A., Otsuki, K., and Ogawa, S. 2007. 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:** 45-56.

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

Kaluli, J.W., Madramootoo, C.A., and Djebbar, Y. 1998. Modeling nitrate leaching using neural networks. Water Sci. Technol. **38 (7):** 127-134.

Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, G., and Wang, S. 2005. Monitoring of forage conditions with MODIS imagery in the Xilingol steppe, Inner Mongolia. Int. J. Remote Sens. **26**: 1423-1436.

Khan, A., 2005. Satellite Characterization of Vegetation Dynamics in the Western Canadian Boreal Plain. Department of Civil and Environmental Engineering. University of Alberta, Edmonton, pp. 162.

Kisi, O. 2004. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. Hydrol. Sci. J. **49:** 1025-1040.

Lauren, A., Finer, L., Koivusalo, H., Kokkonen, T., Karvonen, T., Kellomadki, S., Mannerkoski, H., and Ahtiainen, M. 2005. 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:** 657-674.

Legates, D.R., and McCabe, G.J. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. **35:** 233-241.

Lek, S., Dimopoulos, I. and Fabre, A. 1996. Predicting phosphorus concentration and phosphorus load from from watershed characteristics using backpropagation neural networks. Acta Oecol.-Int. J. Ecol. **17:** 43-53.

Lek, S., Guiresse, M., and Giraudel, J.-L. 1999. Predicting stream nitrogen concentration from watershed features using neural networks. Water Res. **33:** 3469-3478.

Li, X., Nour, M.H., Smith, D.W., and Prepas, E.E. 2008. Modelling nitrogen composition in streams on the Boreal Plain using genetic adaptive general regression neural networks. J. Environ. Eng. Sci. **7** (Suppl. 1): 109-125.

Luke, S.H., Luckai, N.J., Burke, J.M., and Prepas, E.E. 2007. Riparian areas in the Canadian boreal forest and linkages with water quality in streams. Environ. Rev. **15:** 79-97.

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

Maier, H.R., Dandy, G.C., and Burch, M.D. 1998. Use of artificial neural networks for modelling cyanobacteria *Anabaena* spp. in the River Murray, South Australia. Ecol. Model. **105:** 257-272.

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

Neelakantan, T.R., Brion, G.M., and Lingireddy, S. 2001. Neural network modelling of *Cryptosporidium* and *Giardia* concentrations in the Delaware River, USA. Water Sci. Technol. **43**: 125-132.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006a. 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:** 19-32.

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

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2008. Towards a generic neural network model for the prediction of daily streamflow in ungauged boreal plain watersheds. J. Environ. Eng. Sci. **7**, (Suppl. 1): 79-93.

Pelster, D., Burke, J.M., and Prepas, E.E. 2008. Runoff and inorganic nitrogen export from Boreal Plain watersheds six years after wildfire and one year after harvest. J. Environ. Eng. Sci. **7** (Suppl. 1): 51-61.

Prepas, E.E., Smith, D.W., Putz, G., and Germida, J.J. Unpublished data. FORWARD database, The Forestry Corp., Edmonton, AB.

Prepas, E.E., Burke, J.M., Whitson, I.R., Putz, G., and Smith, D.W. 2006. 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** (Suppl. 1): 27-37.

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

Recknagel, F. 1997. ANNA-Artificial neural network model for predicting species abundance and succession of blue-green algae. Hydrobiologia. **349:** 47-57.

Recknagel, F., Cao, H., Kim, B., Takamura, N. and Welk, A. 2006. Unravelling and forecasting algal population dynamics in two lakes different in morphometry and eutrophication by neural and evolutionary computation. Ecol. Inform. **1:** 133-151.

Shanmugam, P., Ahn, Y.H., and Sanjeevi, S. 2006. 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**: 379-394.

Sharma, V., Negi, S.C., Rudra, R.P., and Yang, S. 2003. Neural networks for predicting nitrate-nitrogen in drainage water. Agric. Water Manage. **63**: 169-183.

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

Smith, D.W., Prepas, E.E., Putz, G., Burke, J.M., Meyer, W.L., and Whitson, I. 2003. 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** (Suppl. 1): 1-13.

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

Tayfur, G., and Guldal, V. 2006. Artificial neural networks for estimating daily total suspended sediment in natural streams. Nord. Hydrol. **37**: 69-79.

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

The ASCE Task Committee. 2000. Artificial Neural Networks in Hydrology. I: Preliminary Concepts. J. Hydrol. Eng. **5**: 115-123.

Willems, H.P.L., Rotelli, M.D., Berry, D.F., Smith, E.P., Reneau, J., Raymond B., and Mostaghimi, S. 1997. Nitrate removal in riparian wetland soils: Effects of flow rate, temperature, nitrate concentration and soil depth. Water Res. **31**: 841-849.

Yu, C., Northcott, W.J., and McIsaac, G.F. 2004. Development of an artificial neural network for hydrologic and water quality modeling of agricultural watersheds. Trans. ASAE. **47**: 285-290.

Zealand, C.M., Burn, D.H., and Simonovic, S.P. 1999. Short streamflow forecasting using artificial neural networks. J. Hydrol. **214:** 32-48.

Zhang, Q., and Stanley, S.J. 1997. Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modeling. Water Res. **31:** 2340-2350.

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

CHAPTER 5. INCORPORATING WATER QUANTITY AND QUALITY MODELING INTO FOREST MANAGEMENT[1]

5.1 Introduction

In Alberta, the forest industry is an important land-based resource activity, following only the oil and gas and agricultural sectors in economic importance. Millar Western Forest Products Ltd. (MWFP) is an Alberta-based company for pulp and lumber production. The company's harvesting activities are governed by a Forest Management Agreement (FMA) with the Province of Alberta. The obligations of an FMA permit holder are to harvest no more than the amount of timber stated in the FMA, and to promptly regenerate and maintain the harvested areas in a forested condition. In addition, the FMA holder should also plan the harvest strategy to accommodate other interests in its FMA area, as they are impacted by the forest management operations. To accommodate possible conflicting interests within an FMA area, the agreement holder prepares a Detailed Forest Management Plan (DFMP) that sets the strategic planning process. This process assures that each FMA holder is following the guidelines set out by the provincial government.

A set of harvesting control policies are currently used as best practices, attempting to minimize the adverse impacts of forest harvesting on biodiversity, ecological sustainability, water quantity and quality and forest sustainability. For example, rules like the 50% maximum allowable cut with a 40 ha area maximum clearcut size, and harvesting that mimics natural disturbance are being used to preserve our forests. However, there is no research basis to substantiate these approaches. In an era of increased land use and resource development, forest management planning processes will require increasingly more sophisticated

[1] A version of this chapter has been published. Li, X., Nour, M.H., Smith, D.W., Prepas, E.E., Putz, G., and Watson, B. 2008. Incorporating Water Quantity and Quality Modelling into Forest Management. Forest. Chron. **84:** 338 – 348.

modelling tools to identify and avoid significant impacts on the environment (Smith *et al.* 2003a).

As a first step toward developing the required modelling tools, MWFP initiated the Forest Watershed and Riparian Disturbance (FORWARD) project in collaboration with researchers from Lakehead University, the University of Alberta, and the University of Saskatchewan to develop a better understanding of the impact of harvesting activities on soils, hydrology and water quality (Smith *et al.* 2003b). Over the past 6 years, FORWARD research has provided a detailed database of soil properties, streamflow and water quality within the MWFP FMA area and a better understanding of the link between land-based activities and water resource impacts. Initial streamflow and water quality simulation models have been formulated and tested on small pilot-scale forested watersheds near Whitecourt, Alberta. To apply these models on a scale comparable to an entire industrial forest management area, there is an urgent need to adapt them to be less reliant on data intensive inputs and to provide the means to include these models in operational forest management and planning. This paper summarizes the available water quantity and quality models, presents the FORWARD stream flow and quality modelling approach and initial modelling results, and proposes a framework toward incorporating these modelling efforts in the DFMP process.

5.2 Water Quantity and Quality Modelling

Simulation models are very useful tools to analyze watershed processes, and to develop and assess watershed management scenarios. A multitude of applications (e.g., streamflow and water quality parameters forecasting, the evaluation of the impact of different forest management and agricultural activities on water quantity and quality and the evaluation of watershed responses to different climate change scenarios) have contributed to the development of a vast number of watershed models, starting in the early 1960s (Wagener 2005). These models are usually a mixture of linear and non-linear functions, combined to represent those processes occurring in a specific watershed and important for the study objectives at hand.

5.2.1 Review of Available Watershed Models

Watershed models can be classified based on the degree of spatial resolution into: (1) lumped models that use average values of input variables over the entire watershed area, and thus have minimal data requirements; (2) semi-distributed models that divide the watershed into sub-watersheds, in which each sub-watershed carries a distinct set of input variables; and (3) distributed models that are pixel-based in terms of input representations and parameter routing, and therefore having huge data requirements. Although using distributed models is conceptually appealing, the superiority of the more complex semi-distributed and distributed models over the simpler lumped models is still an issue of debate (Wilcox *et al.* 1990, Michaud and Sorooshian 1994, Hauhs *et al.* 1996, Donnelly-Makowecki and Moore 1999).

Watershed models can also be classified according to physical conceptualization into: (1) empirical (also called data-driven); (2) physically-based (also called mechanistic); and (3) conceptual (also called parametric) models. Empirical models use available time-series of input and output variables (nutrient concentrations, precipitation, streamflow, temperature, etc.) to derive both the model structure and the corresponding parameter values. They therefore do not need a complete prior knowledge about the physical, chemical, hydro-morphological and biological processes controlling flow processes and contaminant transport mechanisms. Physically-based models mathematically describe a process using a set of principles, based on the conservation of mass, momentum and energy. They are distributed models and have intensive data requirements. Conceptual models include both simplified physically-based components and empirical components. The modeller, based on a conceptualization of the watershed, specifies the structure of these models in advance and uses observations of the watershed response to find appropriate values for the model parameters through empirical relations. Conceptual models form the large majority of models used in practice.

Conceptual watershed-scale water quantity and quality models include; but are not limited to: the Soil and Water Assessment Tool (SWAT) developed by Arnold *et al.* (1998); aerial non-point source watershed environment simulation-2000 (ANSWERS-2000) (Beasley *et al.* 1980, Bouraoui and Dillaha 1996); Hydrologic Simulation Program Fortran (HSPF) (Johanson *et al.* 1984); erosion productivity impact calculator (Sharpley and Williams 1990); annualized agriculture non-point source pollutant loading model (AnnAGNPS) (Bingneer *et al.* 2001); and the Guelph model for evaluating the effects of agricultural management systems on erosion and sedimentation (GAMES) by Cook *et al.* (1985). The use of this class of models presents the challenge of estimating or calibrating a large number of model parameters from information with limited availability. Obtaining the information necessary for model calibration is time consuming and expensive.

Data-driven models have been successful in capturing patterns in data with less knowledge of the behavior of the system in terms of interactions among the biological, geological, chemical and physical processes affecting the modelled system. Consequently, they are attractive alternatives to traditional conceptual models. Among those techniques, artificial neural network (ANN) models hold promise for water quantity and quality modelling. ANN models can often capture data patterns without extensive knowledge of the particular site-related problems and can model complicated and non-linear processes with fewer input variables than mechanistic models. Since they are capable of handling large-scale and complex problems, ANN models provide great advantages in a wide range of water quality applications, such as modelling sediment concentrations (Cigizoglu and Alp 2006, Cigizoglu and Kisi 2006, Nour *et al.* 2006b, Tayfur and Guldal 2006, Alp and Cigizoglu 2007), phosphorus concentrations (Nour *et al.* 2006b, c, d), and cyanobacteria blooms (Maier *et al.* 2004, Teles *et al.* 2006) in surface waters.

5.2.2 Data Requirements

It is always a concern to obtain the data required to calibrate watershed-scale models (e.g., topography, vegetation cover, soil characteristics, stream channel characteristics and subsurface infiltration), thus hindering their applications in practice. The current resurgence in earth-observing satellite and airborne platforms, along with the advancements in computer and software technology, has made it possible to evaluate and quantify large numbers of watershed physical characteristics and state variables via remote sensing (RS). It is a cost-efficient way to improve the spatial and temporal coverage of surface water and watershed monitoring (Koponen *et al.* 2004). RS techniques have expanded widely, to the point that they now include most of the electromagnetic spectrum. Different sensors can provide unique information about properties of the surface or shallow layers of the Earth.

The application of RS information to watershed modelling, as well as management, can be divided into three main categories: (1) to delineate surface features, such as snow-covered areas, surface water extent or sediment plumes; (2) to retrieve information such as land cover, geological features or other hydrologic parameters through interpretation and computer classification of remotely sensed data; and (3) to directly use RS digital data to estimate hydrological state parameters. The third application is the most important to watershed modelling and is normally achieved through electro-optical or statistical modelling of known hydrometric data with satellite data. Although there has been some success in the application of RS data in hydrology, the incorporation of RS information in watershed water quality modelling still requires more effort.

The Moderate-resolution Imaging Spectroradiometer (MODIS) sensor on board Terra (United States National Aeronautics Space Administration (NASA)) allows for measurement of plant growth on a global scale at moderate spatial and temporal resolution. The data provided by the MODIS Land Group (e.g., vegetation indexes (VIs) like the normalized difference vegetation index (NDVI), Enhanced Vegetation Index (EVI) and leaf area index (LAI)) support global to regional monitoring, modelling and assessment (Justice *et al.* 1998, National Aeronautics and Space Administration 2007, Li et al. 2008). Furthermore, MODIS data are freely available, thus providing a means of acquiring time series representations of vegetation dynamics at an affordable cost. For instance, a successful nutrient model requires information regarding soil and vegetation nutrient status. RS VIs can represent vegetation health and stress in terms of the vegetation chlorophyll content and the leaf water content, which can be linked to soil/vegetation nutrient interactions and thus can aid in formulating relatively accurate and usable nutrient watershed models (Cheng *et al.* 2006). Such information can potentially act as a surrogate for soil/vegetation nutrient transport and therefore can potentially represent vegetation dynamics in nutrient model formulation.

5.2.3 Integration of RS and GIS with Simulation Models for Watershed Management

The complexity of decision-making, as well as data requirements, has created a need to integrate RS and Geographical Information Systems (GIS) technology with simulation models for watershed management. GIS technology is an essential tool in a variety of fields where spatial information processing is involved, such as forest management, urban planning and agriculture. RS is commonly used in conjunction with GIS to provide spatial data in GIS databases. GIS and RS have been combined with simulation models for many applications, such as vegetation mapping and monitoring, biodiversity mapping and modelling, hydrological modelling, land use planning and environmental impact assessment (Skidmore 2002). For example, a GIS-based ANN model was developed to simulate spatial distribution of nitrate (NO_3^-) concentrations in groundwater with land use information and site-specific hydrogeological properties in an agricultural region (Wang *et al.* 2006). GIS tools were used to prepare and process input–output vectors data for the ANN, which efficiently simulated groundwater NO_3^- concentrations and captured the general trend of groundwater NO_3^- pollution patterns (Wang *et al.* 2006). The use of multi-temporal RS images in support of environmental modelling analysis in a GIS environment has contributed to identify a variety of long-term interactions between resources and land use (Ning *et al.* 2006). The spatial and temporal variation in NO_3^- in the basin of a small river was simulated by combining NO_3^- dynamic modelling and GIS with the use of RS NDVI (Matejicek *et al.* 2003). The NDVI was implemented in the dynamic model to estimate the level of denitrification. In summary, the integration of simulation models, GIS techniques and RS information is necessary to improve watershed management and decision-making process when ground-based data are limited.

5.3 FORWARD Modelling Approach

The FORWARD project study area includes 16 watersheds (3 to 250 km^2) located in the Virginia Hills, Alberta (locations shown in Prepas *et al.* 2008). Seven of the watersheds are relatively undisturbed systems, four were up to 100% burned during the Virginia Hill fire in 1998, four were harvested in 2004, and one was harvested in 2000. Data collection on soil, vegetation, meteorology, water quantity and water quality began in 1998. Based on 11 selection criteria for modelling streamflow and nutrient concentrations in Boreal Plain streams, two modelling approaches incorporating the SWAT model and ANNs have been used in the FORWARD Project (Putz *et al.* 2003). A detailed description of the FORWARD Project can be found in the J. Environ. Eng. Sci. 2 (Suppl. 1).

5.3.1 SWAT Modelling

The SWAT model is a physically-based distributed watershed-scale model that operates on a daily time step. It was developed to predict the impact of watershed management on water, sediment, nutrient and agricultural yields in large basins for a long simulation period (Arnold *et al.* 1998). Spatially, the SWAT model simulates a basin by subdividing it into sub-basins based on topographic information. The components of sub-basins include eight major divisions: hydrology, weather, sedimentation, soil temperature, crop growth, nutrients, pesticides and agricultural management (Arnold *et al.* 1998). The sub-basins can be further divided into smaller modelling units called Hydrologic Response Units (HRUs) depending on the heterogeneity of land uses and soil types within the sub-basin (Muleta and Nicklow 2005). Watershed characteristics and management features are considered homogeneous within an HRU.

Hydrologic processes are simulated within the SWAT model, including surface runoff, percolation, lateral subsurface flow, groundwater flow, snowmelt and water storage (Arnold *et al.* 1998). The SWAT model also simulates a variety of other watershed processes (e.g., crop growth and nutrient cycling) using watershed information like weather, soil, topography, vegetation and land management practices. The integration of GIS and the SWAT model has proven to be an effective and efficient means for input data preprocessing and output data visualization (Arnold *et al.* 1999). It has been successfully used for many case studies (Di Luzio *et al.* 2005, Qi and Grunwald 2005, Santhi *et al.* 2005, Grunwald and Qi 2006, Olivera *et al.* 2006). The SWAT model is a comprehensive hydrologic model that can simulate streamflow, as well as water quality parameters. It is a public-domain model and is conceptually sound. However, it was initially developed to simulate agricultural watersheds. A number of modifications are required for the model to be able to simulate a forested ecosystem.

5.3.2 ANN Modelling

An ANN model is a data-driven modelling alternative that was developed in an attempt to mimic the learning of human brains. ANNs consist of a large number of simple, highly interconnected processing elements (neurons) in an architecture inspired by the structure of the cerebral cortex of the brain (Tsoukalas and Uhrig 1997). The successful application of ANNs in water quantity and quality modelling indicates that ANN models are useful in supporting environmental decision making (Maier *et al.* 2001, Rudra *et al.* 2005, Dakou *et al.* 2006, Diamantopoulou *et al.* 2007, Elhatip and Komur 2008).

There are seven major components to an ANN architecture: (1) processing neurons; (2) a state of activation; (3) an output function for each neuron; (4) a pattern of connectivity or weights between neurons; (5) a propagation rule; (6) an activation function to combine the inputs impinging on a neuron with the current state of that neuron to produce a new level of activation for that unit; and (7) a learning rule whereby weights are adjusted for model calibration (Rumelhart and McClelland 1986). The processing neurons are generally organized in layers. The multi-layer perceptron neural network trained with the error back-propagation training algorithm (MLP-BP) is by far the most popular of all neural networks (Maier and Dandy 2000, Dawson and Wilby 2001). Considering the popularity of the algorithm and because the authors have applied this algorithm successfully in different applications, it was used in this study.

The development of ANN models generally includes the following steps: (1) input determination; (2) data division into training, testing and validation datasets; (3) determination of model architecture (e.g., number of hidden layers, number of neurons in each layer, activation function, and learning rate); (4) model calibration; and (5) model evaluation. The determination of model inputs is one of the most important steps in developing a successful ANN model. Training data quality is of paramount importance to a data-driven modelling approach like ANN. Inclusion of noisy and correlated input variables can increase computational complexity and deteriorate model performance. Thus, input variables should be carefully selected to closely describe the physical system being modelled. Ideally, the available input/output data pairs should be divided into three data sets for training (calibrating), testing and validating the model. The training data set is used for model training and for the optimization of the model connection weights. The testing set is used to decide when to stop training to avoid model overfitting. The cross-validation data set is used to evaluate the model against a totally independent data set. Careful data division is important to assure good model generalization ability. Model architecture determination generally relies upon modellers' experience, however guidelines have been proposed in recent studies (e.g., Maier and Dandy 2000, Nour *et al.* 2006d). For an ANN to generate output predictions that are as close as possible to the objective values, model calibration is an essential step to find optimal weights, minimizing a predetermined error function. After model calibration, the power of the models in terms of prediction accuracy, robustness and generalization ability should be verified before the models are put in use. The models usually are tested through several criteria including: the coefficient of multiple determination (R^2), correlation coefficient, root mean squared error, the multivariate corrected Akaike's information criterion (AICc), the Bayesian information criterion (BIC), swapping the testing and validation data sets, plotting measured and predicted values over time and graphing measured and predicted values for training, testing and validation data sets (Nour *et al.* 2008a).

5.3.3 Link between SWAT and ANN

The SWAT model is widely used for hydrologic processes simulation and to support decision-making within watershed management. It has been upgraded over time and integrated with a GIS database component. However, as a physically-based distributed basin scale model, the SWAT model is very data intensive. In some situations, the power of SWAT is compromised due to data limitation. On the other hand, ANNs are able to identify the relationship between inputs and outputs, without fully understanding the mechanistic principles behind them. ANN models have proved to be superior to mechanistic models when data are limited and numerous assumptions have to be made to solve the physically-based equations. For instance, in-stream nitrogen concentrations, which are affected by land use/land classification, vegetation dynamics and in-stream nitrogen transformations, were simulated reasonably well with the ANN approach (Lek *et al.* 1999; Khalil *et al.* 2005).

A study on comparing the performance of SWAT and ANN models in simulating hydrologic processes in an agricultural watershed found that the ANN monthly predictions were closer to the observed flows than the monthly predictions from the SWAT model (Srivastava *et al.* 2006). This and other studies suggest that ANN is an attractive modelling alternative for hydrologic and water quality modelling. However, current ANNs efforts cannot take into account spatial variation within a watershed. More efforts are needed to explore the applicability of developing semi-distributed watershed models using ANN.

There have been very limited attempts in the literature to develop a hybrid SWAT/ANN approach. For example a hybrid rainfall-runoff model integrating ANNs with a conceptual model was introduced by Chen and Adams (2006). In their hybrid model, the spatial variation of rainfall, the heterogeneity of watershed characteristics and their impacts on runoff were investigated by the semi-distributed conceptual rainfall-runoff model, while the nonlinear transformations of the runoff generated from the individual subcatchments into the total runoff at the watershed outlet were performed by the ANNs. This hybrid model took into account both the spatial variation presented by the semi-distributed conceptual model and the nonlinear mapping ability of ANNs, highlighting the possibility of integrating the two kinds of models.

FORWARD researchers use one physically-based data-intensive approach to understand the biological, geological, chemical and physical behaviors of the system, and one data-driven approach that is flexible in terms of data requirements to develop a modelling tool that is less-data-intensive. In addition, attempts to link both the SWAT and the ANN approaches are ongoing to capitalize on the strength of each technique.

5.4 Initial Modelling Results

To date, FORWARD researchers have made significant modifications to the SWAT model to better model watersheds located in forested ecosystems. More modifications are underway. They developed a step-by-step framework for modelling time-correlated variables using ANN and a protocol for utilizing RS information in water quality modelling, and they have applied the developed models successfully in experimental small watersheds in the Boreal Plain. Both these modelling efforts require significant additional advancements before they can to be applied to a larger landbase (like an FMA area) due to the lack of detailed distributed land base data. The experience gained from implementing ANN and SWAT models in small experimental watersheds will pave the road for further development towards a robust technique for utilizing RS data at the wider FMA scale. The following sections summarize the conducted modifications for the SWAT model, and the initial results of the ANN modelling approach.

5.4.1 Modification of SWAT

Because the SWAT model was originally developed for agricultural watershed management, a series of modifications have been incorporated to make it effective for boreal forest management. The modified model, called SWAT-Boreal Forest (SWAT-BF), incorporates the following major changes:

1. A litter layer component adopted from Wattenbach *et al.* (2005) has been incorporated. The litter layer contains many nutrients and is an important component of forest ecosystem. It is able to store water and reduce peak flows during rain events.
2. An algorithm was incorporated to account for the effects of slope and aspect on incoming solar radiation.
3. A new wetlands model was incorporated because the original wetlands model was not deemed suitable for the wetlands found in boreal forests. The new wetlands model is based on a bucket model approach. It has an upper organic layer and a lower organic layer. A non-linear function is used to determine the amount of lateral flow from each layer. Unlike the original wetlands model in SWAT, the new model accounts for water uptake by vegetation, surface runoff, percolation and base flow.
4. A new feature has been incorporated into SWAT that enables lateral flow and base flow from upland HRUs to be routed through lowland wetlands that are found in valley bottoms. Previously with SWAT, this was not possible. Instead, all HRUs simply contributed to the stream and there was no consideration of the position of the HRUs in the landscape. Lowland wetlands tend to retain water and dampen peak flows, so this new arrangement in SWAT attempts to reproduce that process.

5.4.2 ANN Modelling of Streamflow and Water Quality Parameters

The FORWARD researchers have been successful in using ANN modelling to predict streamflow, and water-phase total suspended solids (TSS), total phosphorus (TP), and nitrogen components (NO_3^- and total dissolved nitrogen (TDN)) in undisturbed FORWARD watersheds (e.g., Willow, Two Creek, and Cassidy), disturbed watersheds (e.g., Burnt Pine) and watersheds with a relatively large percentage of wetlands (e.g., 1A) (Table 5-1). The main advantage of the developed ANN models is that they use as inputs meteorological data, which are easily accessible in the boreal forest via Environment Canada weather stations and provincial fire towers, and public-domain, free-of-cost RS MODIS-derived VIs. The model inputs were carefully selected to reflect causality, time correlation and Q/TSS/nutrient hystereses loops. The inputs are composed of cause/effect factors (e.g., rainfall (R), snowfall (S), temperature (T), degree-days (dd), and EVI), time-lagged inputs (e.g., R_t, R_{t-1}, …$R_{t-\mathrm{N}}$) and inputs reflecting seasonal cyclisity and Q/TSS/nutrient hystereses behavior (e.g., $\sin(2\pi\nu t)$, $\cos(2\pi\nu t)$, Julian day), which are determined based on a combination of prior knowledge of the system being modelled, as well as statistical analysis of the data (Table 5-1).

The developed models accurately predicted the data that the models had never seen during calibration (R^2 of validation > 0.76), which highlights the good generalization ability of these models. Also, they captured the seasonal cycle of the modelled parameters and accurately predicted both base and peak concentrations (see Fig. 5-1 as an example). To demonstrate the applicability of the developed models to unmonitored watersheds, the calibrated models were used to predict N export in other different watersheds (unmonitored watersheds in this perspective) without further calibration. The results of transferring the calibrated models to other unmonitored watersheds were promising in some applications with Nash Sutcliffe coefficient E, the square of the Pearson's product-moment correlation r^2 and Index of degree d values being as high as: 0.62, 0.63, and 0.88, respectively. Furthermore, the usefulness of the following indices (individual or combined) was examined in simulating watershed hydrologic similarity, and thus reflecting the success of model transferability from one watershed to the other: peatland index, riparian index, rainfall index, EVI, and remotely-sensed normalized difference water index (NDWI). The best watershed similarity index was found to be the combined (Rainfall_Index+ Peatland_Index+ NDWI_Index) with $r^2 = 0.74$ when the similarity index values were regressed to the E values of model predictions (Figure 5-2). The results can be used to predict the expected success of model transferability from gauged to unmonitored Boreal forest watersheds. Although the proposed indices are not mature enough to have meaningful thresholds above which the models should not be transferred, it is an initiative towards expanding and refining hydrologic similarity indices that rely on free-of-cost remote sensing information.

To quantify the impact of land use activities on water quality, the model must divide the watershed into subwatersheds, to be able to recognize the locations of disturbances and hence simulate the corresponding impacts. FORWARD researchers conducted a leading study on the impact of watershed subdivision on the prediction accuracy of TP concentration in Boreal Plain streams (Nour *et al.* 2008b). Although the statistical model evaluation favored the finest spatial resolution, all model performance indicators were satisfactory for the four models devised for different watershed subdivisions for the Willow watershed. The differences in performance indicators were not significant for any practical application. Therefore, it was concluded that the choice of the optimum watershed subdivision should depend upon the modelling objective (Nour *et al.* 2008b). Lumped parameter models are easy to construct and rely on affordable land base information, but cannot address questions related to the impact of different land use scenarios on water quality. Therefore, if the objective is to forecast real-time water quality (likely used for post-harvesting assessment), lumped parameter modelling can be used without jeopardizing prediction accuracy. On the other hand, if the objective is to quantify the impact of different land use activities, then the watershed must be divided into subwatersheds to make the model able to recognize the locations of disturbances and thus, able to simulate the corresponding impacts on water quality. Based on our results, we conclude that only in this case is the added time, cost and effort of preparing and processing distributed landbase information justifiable.

Successful ANN modelling of water quantity and quality parameters relying only on meteorological and RS data suggests that the proposed models can potentially be applied to modelling watersheds in an FMA-scale land base. The use of MODIS-derived VIs has played an important role in representing phosphorus and nitrogen dynamics within the soil/vegetation phases in the developed water quality models. For example, changes to the vegetation canopy after wildfire in the Burnt Pine watershed (burned in 1998) are captured by the MODIS-derived EVI (Figure 5-3). Higher EVI values indicate higher density of vegetation canopy. Compared to the EVI of the Willow watershed, the EVI of the Burnt Pine watershed was lower and demonstrated an increasing trend from 2001 through 2005, which was resulted from the recovery of vegetation with time since disturbance (Figure 5-3). The following section proposes a plan to extrapolate the application of the currently developed models to the entire MWFP FMA area.

5.5 A Framework to Include Modelling in the DFMP Process

The goal of the FORWARD Project is to develop an improved decision support tool (essentially integrated streamflow and water quality models), which is capable of predicting changes in streamflow and water quality parameters as a result of proposed spatial and temporal patterns of forest harvesting. The improved decision support tool will be incorporated into MWFP's next DFMP to provide the capability to limit and control disturbance effects on water resources in the MWFP FMA area. This task is challenging, due to our immature understanding of the hydrological, biological and chemical mechanisms that control contaminant transport at the large watershed-scale, as well as the lack of pertinent data for model calibration. Providing the human and infrastructure resources to gauge streamflow and measure water quality in all watersheds of interest in an FMA area is impractical, thus a class of models that could simulate the response of unmonitored watersheds by relying on easily accessible information (like meteorological and RS inputs) is important to forest management planning. The previously described ANN models use readily available, easily accessed meteorological information and public-domain free-of-cost information as model inputs. Such models can have a good chance of application in unmonitored watersheds due to the relative ease of obtaining these inputs for an entire FMA area. However, initial results from the FORWARD modelling efforts 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 from one watershed to the other is hydrologic similarity. Thus, the following steps summarize the proposed FORWARD approach toward incorporating water quantity and quality modelling in the MWFP DFMP process:

1. Delineate the digital elevation model (DEM) of the MWFP FMA area into 1st order watersheds (~ 5 km^2) using the eight-direction pour-point algorithm and a reasonable threshold for flow accumulation. This task is already completed and details are presented in Prepas *et al.* (2008).
2. Use rainfall interpolation techniques like kriging and inverse distance weighted interpolation techniques to estimate daily rainfall intensity in the centroid of each watershed using data from surrounding weather stations (fire towers, Environment Canada, and FORWARD project stations). FORWARD researchers have established the interpolation algorithm (Nour *et al.* 2006a). The interpolation weights can be assumed fixed with time, and thus can be used for future scenario simulations.
3. Formulate streamflow (Q), TSS and nutrient models for several of the 16 FORWARD experimental watersheds and validate with the remaining watersheds. A series of combinations of calibration and validation watersheds will be examined to assure model stability and parsimony. This process is still ongoing. For each model to be constructed, the following must be done: (1) MODIS images must be downloaded from NASA; (2) data quality has to be assessed; (3) a yearly stack of images will be constructed for each spectral band of importance (e.g., red, near infrared, mid-infrared, and blue bands) using Erdas Imagine image processing software (from Leica Geosystems);

(4) map algebra will be used to calculate relevant VIs that will later serve as inputs to the water quality models; (5) the DEM-derived watersheds will be overlapped on the satellite images to extract relevant VIs for each delineated watershed; (6) cross correlation and spectral analyses will be used to identify time-lagged inputs and feed the model with Q/TSS/nutrient hystereses loops, respectively; and (7) different ANN models will be formulated for predicting Q, TSS and nutrient concentrations for the FORWARD experimental watersheds.

4. The previously delineated watersheds, including the FORWARD study watersheds, will be grouped into different categories according to hydrologic homogeneity in term of VIs, average slope, % wetland composition, yearly precipitation, and basin area (Figure 5-4). It is proposed to use an advanced classification technique (Kohonen neural networks; Kohonen 1982) to classify the MWFP FMA area delineated watersheds into groups of hydrologically similar watersheds. Data must be collected for a member of each group where none of the FORWARD study watersheds exist. Models are to be calibrated for these additional watersheds as described in No. 3 above. This results in a calibrated model for each group of watersheds.
5. Each calibrated model will then be run for all the watersheds falling into its group of similar watersheds. Upon successful implementation of models to the whole FMA area, scenario-based analysis that forces harvesting disturbance on the landbase, will be fed to the models to identify the impact of different land use activities on water quality and quantity. To design these scenarios, a relation has to be established between VIs (currently used in modelling) and typically used vegetation metrics (e.g., timber volume, average age, height, and diameter at breast height). This relation can be used to translate vegetation cut into values of VIs that can be fed to the models to predict changes in streamflow, water-phase solids, and nutrients in response to harvesting scenarios.
6. The previous demonstration will finally be repeated with a hybrid ANN/SWAT modelling approach, which would likely create boundaries and

reduce the number of parameters for the ANN modelling (a data-driven approach) based on the conceptual representation of the SWAT model (a conceptually based approach) and thus make it easier for ANN parameter estimation.

5.6 Conclusions

Over the past 6 years, the FORWARD research project has developed a detailed database of soil properties, streamflow and water quality within the MWFP FMA area and a better understanding of the impact of land-based activities on water resources. Simulation models capable of modelling initial streamflow, TSS and nutrient concentrations have been developed and tested on pilot-scale forested watersheds near Whitecourt, Alberta. To apply these models on a full-scale industrial FMA area, it is necessary to adapt them to be less data-intensive and to provide the means to incorporate these models in operational forest management and planning. Therefore, FORWARD attempted to rely on one physically-based approach to understand the biological, geological, chemical, and physical behaviors of the system, and one data-driven approach, which is flexible in terms of data requirements, to develop a modelling tool that is less data-intensive. In addition, attempts to link both approaches are ongoing in order to capitalize on the strengths of each technique.

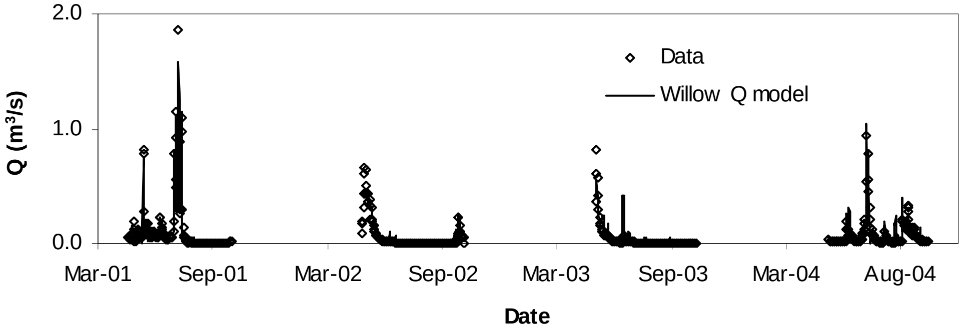
To date, FORWARD researchers have made significant modifications to the SWAT model to better model watersheds located in forested ecosystems. More modifications are underway. They developed a step-by-step framework for modelling time-correlated variables using ANN, a protocol for utilizing RS information in water quality modelling, and they have applied the developed models successfully in experimental watersheds on the Boreal Plain.

The experience gained from implementing ANN and SWAT models in experimental watersheds will pave the road for further development towards a robust technique for utilizing RS data at the broader FMA scale. The results of the model application within the MWFP FMA area will be documented and included in the company's next DFMP. These efforts will provide a leading example for similar forest industry companies on how to predict and mitigate disturbance on the landscape wisely, so that water quality is not impaired. This work will also likely provide guidelines for including such modelling tools in operational forest management and planning for possible use by other forest industry companies.

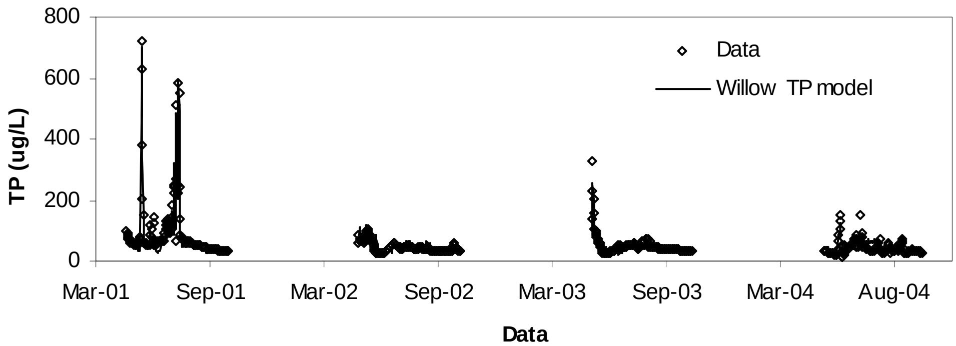
Table 5-1. Initial ANN modelling results for streamflow and water quality parameters.

Modelled parameter	FORWARD watershed[a]	Input parameters[b]	Model performance (R^2)		References
			Calibration	Validation	
Q	Cassidy, Willow, 1A, Two Creek	R, dd, S, T, seasonal cycle indicators ($\sin(2\pi vt)$, $\cos(2\pi vt)$)	>0.90	>0.84	Nour *et al*. 2006b,c, 2008a
TSS	Two Creek	R, dd, S, seasonal cycle indicators ($\sin(2\pi vt)$, $\cos(2\pi vt)$)	>0.90	>0.90	Nour *et al*. 2006b
TP	Willow, 1A, Two Creek	R, dd, S, T, EVI, and seasonal cycle indicators ($\sin(2\pi vt)$, $\cos(2\pi vt)$)	>0.86	>0.76	Nour *et al*. 2005, Nour *et al*. 2006b, 2008b
TDN	Willow, Two Creek, Burnt Pine	R, dd, T, EVI, seasonal cycle indicator (Julian day)	>0.90	>0.90	Li *et al*. 2008
NO_3^-	Willow, Two Creek, Burnt Pine	R, dd, T, EVI, seasonal cycle indicator (Julian day)	>0.90	>0.90	Li *et al*. 2008
NH_4^+	Willow, Two Creek, Burnt Pine	R, dd, T, EVI, seasonal cycle indicator (Julian day)	>0.90	>0.83	Li et al. 2008

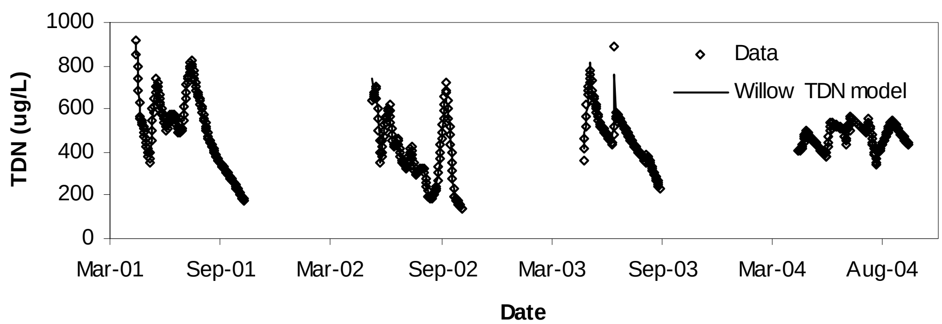
[a]Cassidy, Willow and Two Creek are undisturbed watersheds, 1A is undisturbed and has a large percentage of wetlands, Burnt Pine watershed was 100% burned in 1998.
[b]R: rainfall, dd: degree-days, S: snowfall, T: air temperature, EVI: enhanced vegetation index.



(a)



(b)



(c)

Figure 5-1. Time series plot of measured and ANN predicted (a) streamflow (Q), (b) TP, and (c) TDN for the Willow watershed.

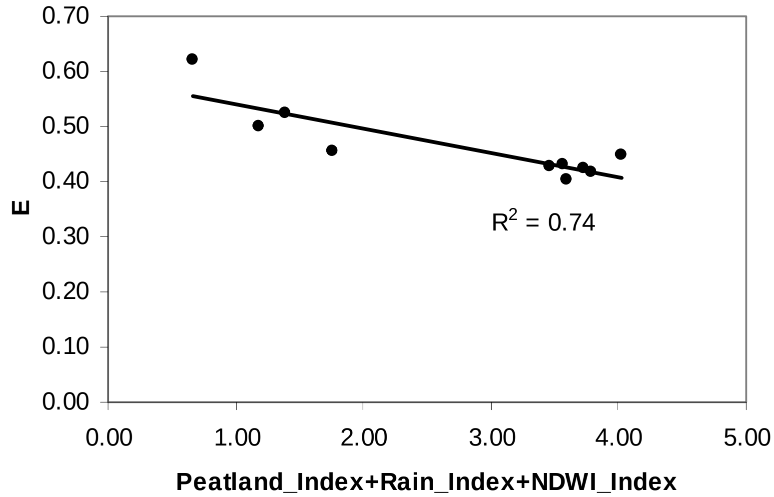


Figure 5-2. Plot of model transferability performance measure *E* versus watershed similarity indices.

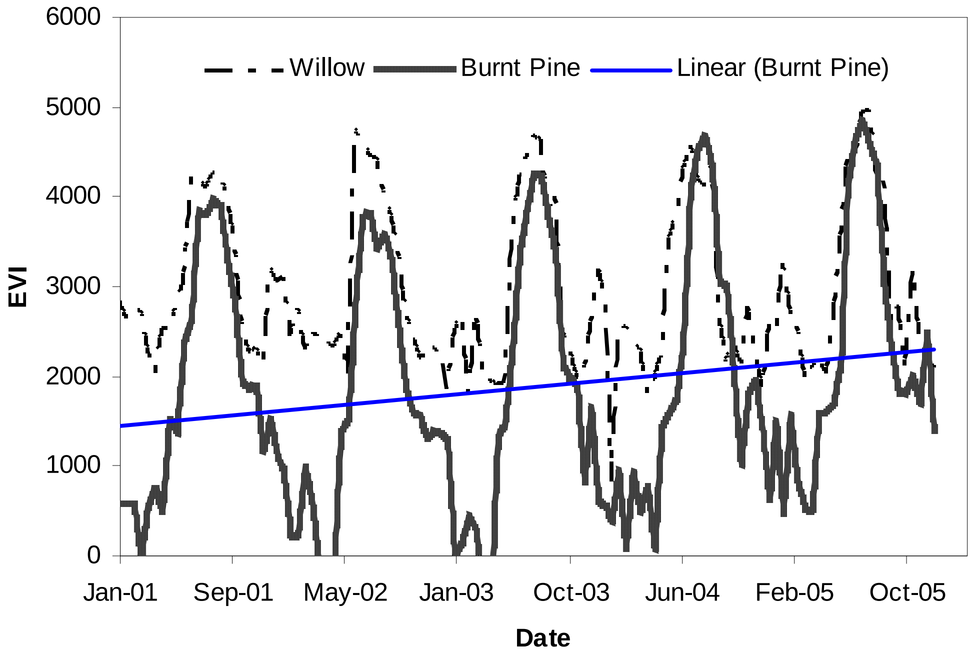


Figure 5-3. Comparison of EVI for the Willow and Burnt Pine watersheds for 2001 to 2005.

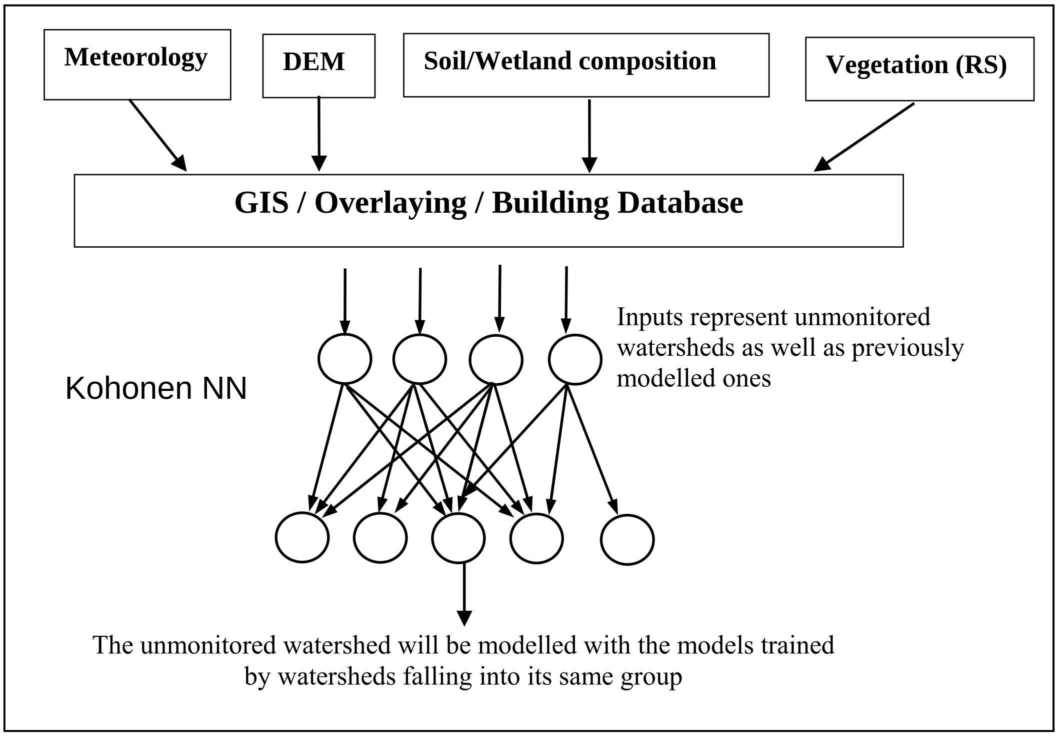


Figure 5-4. A framework to model unmonitored watersheds using ANN models. DEM: digital elevation model, RS: remote sensing, GIS: geographical information systems, NN: neural network.

5.7 References

Alp, M. and Cigizoglu, H.K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. Environ. Model. Soft. **22**: 2-13.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., and Williams, J.R. 1998. Large area hydrologic modeling and assessment - Part 1: Model development. J. Am. Water Resour. Assoc. **34**: 73-89.

Arnold, J.G., Srinivasan, R., Ramanarayanan, T.S., and Luzio, M. Di. 1999. Water resources of the Texas Gulf Basin. Water Sci. Technol. **39 (3)**: 121-133.

Beasley, D.B., Huggins, L.F., and Monke, E.J. 1980. ANSWERS - A model for watershed planning. Trans. ASAE **23**: 938-944.

Bingneer, R.L., Theurer, F.D., Cronshey, R.G., and Darden, R.W. 2001. AGNPS 2001. Available from http://msa.ars.usda.gov/ms/oxford/nsl/AGNPS.html.

Bouraoui, F. and Dillaha, T.A. 1996. ANSWERS-2000: Runoff and sediment transport model. J. Environ. Eng.-ASCE **122**: 493-502.

Chen, J. and Adams, B.J. 2006. Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. J. Hydrol. **318**: 232-249.

Cheng, Y.B., Zarco-Tejada, P.J., Riano, D., Rueda, C.A., and Ustin, S.L. 2006. Estimating vegetation water content with hyperspectral data for different canopy scenarios: Relationships between AVIRIS and MODIS indexes. Remote Sens. Environ. **105**: 354-366.

Cigizoglu, H.K. and Alp, M. 2006. Generalized regression neural network in modelling river sediment yield. Adv. Eng. Soft. **37**: 63-68.

Cigizoglu, H.K. and Kisi, O. 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol. **317**: 221-238.

Cook, D.J., Dickinson, W.J., and Rudra, R.P. 1985. GAMES-the Guelph Model for Evaluating the Effects of Agricultural Management Systems in Erosion and Sedimentation. User's Manual. Guelph, Ont.

Dakou, E., Goethals, P.L.M., D'Heygere, T., Dedecker, A.P., Gabriels, W., Pauw, N. De, and Lazaridou-Dimitriadou, M. 2006. Development of artificial neural network models predicting macroinvertebrate taxa in the river Axios (Northern Greece). Annales De Limnologie-Int. J. Limnol. **42**: 241-250.

Dawson, C.W. and Wilby, R.L. 2001. Hydrological modelling using artificial neural networks. Progr. Phys. Geogr. **25**: 80-108.

Di Luzio, M., Arnold, J.G., and Srinivasan, R. 2005. Effect of GIS data quality on small watershed stream flow and sediment simulations. Hydrol. Proc. **19**: 629-650.

Diamantopoulou, M.J., Georgiou, P.E., and Papatnichail, D.M. 2007. Performance of neural network models with Kalman learning rule for flow routing in a river system. Fresenius Environ. Bull. **16**: 1474-1484.

Donnelly-Makowecki, L.M. and Moore, R.D. 1999. Hierarchical testing of three rainfall-runoff models in small forested catchments. J. Hydrol. **219**: 136-152.

Elhatip, H. and Komur, M.A. 2008. Evaluation of water quality parameters for the Mamasin dam in Aksaray City in the central Anatolian part of Turkey by means of artificial neural networks. Environ. Geol. **53**: 1157-1164.

Grunwald, S. and Qi, C. 2006. GIS-based water quality modeling in the Sandusky Watershed, Ohio, USA. J. Am. Water Resour. Assoc. **42**: 957-973.

Hauhs, M., Neal, C., Hooper, R., and Christophersen, N. 1996. Summary of a workshop on ecosystem modeling: The end of an era? Sci. Total Environ. **183**: 1-5.

Johanson, R.C., Imhoff, J.C., Kittle, J.L., and Donigian, A.S. 1984. Hydrological Simulation Program-FORTRAN (HSPF) User's Manual. U.S. Environmental Protection Agency, Athens, GA.

Justice, C.O., Vermote, E., Townshend, J.R.G., DeFries, R., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.L., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Wan, Z.M., Huete, A.R., Leeuwen, W. van, Wolfe, R.E., Giglio, L., Muller, J.P., Lewis, P., and Barnsley, M.J. 1998. The Moderate Resolution Imaging

Spectroradiometer (MODIS): Land remote sensing for global change research. IEEE Trans. Geosci. Remote Sens. **36**: 1228-1249.

Khalil, A., Almasri, M.N., McKee, M., and Kaluarachchi, J.J. 2005. Applicability of statistical learning algorithms in groundwater quality modeling. Water Resour. Res. **41**: Art. No. W05010.

Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. **43**: 59-69.

Koponen, S., Kallio, K., Pulliainen, J., Vepsalainen, J., Pyhalahti, T., and Hallikainen, M. 2004. Water quality classification of lakes using 250-m MODIS data. IEEE Geosci. Remote Sens. Lett. **1**: 287-291.

Lek, S., Guiresse, M., and Giraudel, J.-L. 1999. Predicting stream nitrogen concentration from watershed features using neural networks. Water Res. **33**: 3469-3478.

Li, X., Nour, M.H., Smith, D.W., and Prepas, E.E. 2008. Modelling nitrogen composition in streams on the Boreal Plain using genetic adaptive general regression neural networks. J. Environ. Eng. Sci., **7** (S1): 109-125.

Maier, H.R. and Dandy, G.C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model. Soft. **15**: 101-124.

Maier, H.R., Burch, M.D., and Bormans, M. 2001. Flow management strategies to control blooms of the cyanobacterium, *Anabaena circinalis*, in the River Murray at Morgan, South Australia. Regul. Rivers Res. Manage. **17**: 637-650.

Maier, H.R., Kingston, G.B., Clark, T., Frazer, A., and Sanderson, A. 2004. Risk-based approach for assessing the effectiveness of flow management in controlling cyanobacterial blooms in rivers. River Res. Appl. **20**: 459-471.

Matejicek, L., Benesova, L., and Tonika, J. 2003. Ecological modelling of nitrate pollution in small river basins by spreadsheets and GIS. Ecol. Model. **170**: 245-263.

Michaud, J. and Sorooshian, S. 1994. Comparison of simple versus complex disturbed runoff models on a midsized semiarid watershed. Water Resour. Res. **30**: 593-605.

Muleta, M.K. and Nicklow, J.W. 2005. Decision support for watershed management using evolutionary algorithms. J. Water Resour. Plan. Manage. - ASCE **131**: 35-44.

National Aeronautics and Space Administration. 2007. Data Gateway Interface [online]. Available from: http://modis.gsfc.nasa.gov/ [cited 15 June 2007]

Ning, S.K., Chang, N.B., Jeng, K.Y., and Tseng, Y.H. 2006. Soil erosion and non-point source pollution impacts assessment with the aid of multi-temporal remote sensing images. J. Environ. Manage. **79**: 88-101.

Nour, M.H., Khan, A., Smith, D.W., and Gamal El-Din, M. 2005. On the potential of satellite derived vegetation phenology for watershed nutrient modelling: a neural network approach. Proceedings of the Water Environment Federation, WEFTEC®2005. Washington, D.C. 23 pp.

Nour, M.H., Smith, D.W., and Gamal El-Din, M. 2006a. Geostatistical mapping of precipitation: Implications for rain gauge network design. Water Sci. Technol. **53 (10)**: 101-110.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006b. Neural networks modelling of streamflow, phosphorus, and suspended solids: application to the Canadian Boreal forest. Water Sci. Technol. **53 (10)**: 91-99.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006c. 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**: 19-32.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2006d. Artificial neural networks and time series modelling of TP concentration in boreal streams: a comparative approach. J. Environ. Eng. Sci. **5** (S1): 39-52.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2008a. Towards a generic neural network model for the prediction of daily streamflow in unmonitored Boreal Plain watersheds. J. Environ. Eng. Sci. **7** (S1): 79-93.

Nour, M.H., Smith, D.W., Gamal El-Din, M., and Prepas, E.E. 2008b. Effect of watershed subdivision on water-phase phosphorus modelling: an artificial neural network modelling application. J. Environ. Eng. Sci. **7** (S1): 95-108.

Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H.D., Koka, S., and Agrawal, A. 2006. ArcGIS-SWAT: A geodata model and GIS interface for SWAT. J. Am. Water Resour. Assoc. **42**: 295-309.

Prepas, E.E., Putz, G., Smith, D.W., Burke, J.M., and MacDonald, J.D. 2008. The FORWARD Project: Objectives, framework and initial integration into the Detailed Forest Management Plan process in Alberta. For. Chron. **84**: 330-337.

Putz, G., Burke, J.M., Smith, D.W., Chanasyk, D.S., Prepas, E.E., and Mapfumo, E. 2003. Modelling the effects of boreal forest landscape management upon streamflow and water quality: Basic concepts and considerations. J. Environ. Eng. Sci. **2** (S1): 87-101.

Qi, C. and Grunwald, S. 2005. GIS-based hydrologic modeling in the Sandusky watershed using SWAT. Trans. ASAE **48**: 169-180.

Rudra, R.P., Negi, S.C., and Gupta, N. 2005. Modelling approaches for subsurface drainage water quality management. Water Qual. Res. J. Can. **40**: 71-81.

Rumelhart, D.E. and McClelland, J.L. 1986. Parallel distributed processing: explorations in the microstructure of cognition. Vol. 2. MIT Press, Cambridge, Mass.

Santhi, C., R.S. Muttiah, J.G. Arnold and R. Srinivasan. 2005. A GIS-based regional planning tool for irrigation demand assessment and savings using SWAT. Trans. ASAE **48**: 137-147.

Sharpley, A.N., and Williams, J.R. 1990. EPIC-Erosion/Productivity Impact Calculator: I. Model Documentation, Tech. Bull.

Skidmore, A.K. 2002. Introduction. In: Skidmore, A. ed., Environmental Modelling with GIS and Remote Sensing. Taylor & Francis Inc., New York.

Smith, D.W., Russell, J.S., Burke, J.M., and Prepas, E.E. 2003a. Expanding the forest management framework in the province of Alberta to include landscape-based research. J. Environ. Eng. Sci. **2** (S1): 15-22.

Smith, D.W., Prepas, E.E., Putz, G., Burke, J.M., Meyer, W.L., and Whitson, I. 2003b. 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** (S1): 1-13.

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

Tayfur, G. and Guldal, V. 2006. Artificial neural networks for estimating daily total suspended sediment in natural streams. Nordic Hydrol. **37**: 69-79.

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

Tsoukalas, L.H. and Uhrig, R.E. 1997. Fuzzy and Neural Approaches in Engineering. Wiley-Interscience, New York.

Wagener, T. 2005. Watershed Modelling. Water Encyclopedia. John Wiley & Sons, Inc.

Wang, M.X., Liu, G.D., Wu, W.L., Bao, Y.H., and Liu, W.N. 2006. Prediction of agriculture derived groundwater nitrate distribution in North China Plain with GIS-based BPNN. Environ. Geol. **50**: 637-644.

Wattenbach, M., Hattermann, F., Weng, R., Wchsung, F., Krysanova, V., and Badeck, F. 2005. A simplified approach to implement forest eco-hydrological properties in regional hydrological modelling. Ecol. Model. **187**: 40-59.

Wilcox, B.P., Rawls, W.J., Brakensiek, D.L., and Wight, J.R. 1990. Predicting runoff from rangeland catchments: A comparison of two models. Water Resour. Res. **26**: 2401-2410.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

6.1 General Overview

Prediction of N using a modelling tool is a significant component of watershed planning and management. Watershed models can be classified into data-driven models, physically-based models and conceptual models according to physical conceptualization. The factors that change the availability and cycling of N in forest ecosystems include nitrogen deposition, climate gradients and variations, species effects, hydrologic pathways and disturbance. And the effects of the factors on nitrogen in streams are not completely understood. Also, it is very difficult to mathematically represent these factors that are nonlinearly interacted and spatially distributed. In addition, it is cost prohibitive and/or physically unavailable to monitor all the watersheds of interest and in addition some of them are very remote. Therefore, data-driven models (artificial neural networks (ANNs)) that require less data than the physically-based models and conceptual models were adapted to this study to simulate N forms in streams.

The overall objective of this thesis was to develop nitrogen models with reasonable predictive capabilities that rely only on meteorological data and satellite information. The models will be further used to predict N in unmonitored watersheds where on-site measurements are not needed. Ultimately, a framework on incorporating these models into forestry planning and management is proposed. To achieve these targets, it is crucial to get understanding of ANN models, N dynamics and the way to couple models to support decision making. Understanding the standard model development procedure, the advantages and limitation of ANN models and the methods to reduce their limitations in practice is needed to develop reliable ANNs models and apply them properly (Chapter 2). It is very important to select a set of representative input parameters for the system being modelled to develop ANN model with good generality. To select the appropriate inputs, the studied system needs to be understood as much as possible, which is provided in Chapter 2. Chapter 2 also provides a review on the combination of water quality models with GIS and RS to support decision making in watershed management. Given the solid understanding of the above perspectives, the nitrogen compositions in streams on Canadian Boreal Plain were modelled using ANNs and the results were presented in Chapter 3. The successful modelling N using only meteorological data and RS data highly indicated the proposed ANN approach is promising to be used to model N in unmonitored watersheds. Therefore, transferring calibrated models to other different watersheds without further calibration was investigated and a watershed similarity index correlated to the possibility of model transferability was proposed (Chapter 4). Based on solid understanding of water quality models and the promising results of N modelling, Chapter 5 proposed a framework to incorporate these models into forestry planning and management.

6.2 Conclusions

Based on the literature study and the development of nitrogen models, the major conclusions can be drawn as follows:

1. GA-GRNN models were developed following strict procedures and applied to simulate daily mean NO_3^-, NH_4^+ and TDN concentrations in streams at three watersheds, Willow (reference, 15.6 km^2), Two Creek (reference, 129.4 km^2), and Burnt Pine (burnt, 7.7 km^2) in the Swan Hills of Alberta, Canada. The optimal inputs were derived from five major variables: rainfall, daily mean air temperature, cumulative degree-days, EVI and Julian day of the year. In terms of model architecture, the developed models were not sensitive to the initial smoothing factor and training with a genetic algorithm improved model performance on testing data sets.

2. The consistent performance of GA-GRNN models for two relatively undisturbed watersheds, as well as a burned watershed, was obtained with the inclusion of the RS-derived EVI as one of the model inputs. Without EVI, some models could not steadily perform well on validation data sets. This index was capable of describing vegetation canopy differences among watersheds, as well as vegetation phenology. The developed models

3. A MLP-BP algorithm using low-cost, readily available meteorological data and satellite data was proposed to model N export in forested watersheds. The performance of the models was evaluated using correlation-based measures, absolute error measures and time series plots of measured and modelled values. Although the modelled parameter varied in a big range (i.e., the peak values were over thousands of the low values), it was simulated fairly well. The best MLP-BP architecture for all the models had a single hidden layer with three activation functions. Modelling nitrogen export only using readily available data with reasonable accuracy indicates its potential application to unmonitored watershed.

4. The calibrated models were used to predict N export in other different watersheds without further calibration to demonstrate the applicability of the developed models to unmonitored watersheds. The Nash Sutcliffe coefficient E is greater than 0, which means that the models produced better estimates than the mean of the observed values. The correlation coefficient r^2 and index of agreement d were in the range of 0.44 to 0.63 and 0.73 to 0.88, respectively. The transferred models could catch the seasonal and annual periodicity of nitrogen export even though some peak values were not well predicted.

5. A watershed similarity index was proposed to measure watershed similarity and model transferability. The usefulness of the following indices (individual or combined) was examined in simulating watershed hydrologic similarity, and thus reflecting the success of model transferability from one watershed to the other: peatland index, riparian index, rainfall index, EVI, and remotely-sensed normalized difference water index (NDWI). For each single index, Rainfall_Index had the highest correlation with model transferability (r^2 = 0.71), and the next one was NDWI_Index (r^2 = 0.69). The best watershed similarity index was found to be the combined

6. A framework to incorporate the developed ANNs models into forestry management was proposed based on the success of N modelling. This framework basically includes watershed delineation, model development and watershed similarity measurement and can be proceeded in the steps: (1) Delineate the digital elevation model of the study area into 1st watershed; (2) Prepare rainfall data from weather stations and extract RS VIs for the studied watersheds; (3) The watersheds are classified into groups of hydrological similarity using Kohonen neural network or the proposed watershed similarity index in Chapter 4; (4) An ANN model is developed for a member of each group and validated for the other watersheds in the same group; (5) A relation need to be established between VIs (currently used in modelling) and typically used vegetation metrics (e.g., timber volume, average age, height, and diameter at breast height) for successful implementation of models to direct forestry planning based on scenario analysis. This relation can be used to translate vegetation cut into values of VIs that can be fed to the models to predict changes in streamflow, water-phase solids, and nutrients in response to harvesting scenarios.

In addition, although GRNNs can model continuous functions and are fast to train, MLP-BP is more preferable over GRNN to simulate time series variables that vary in a large range of magnitude. This is because that GRNN models generate blank prediction results for input data that are not similar to its training data patterns. The nature of GRNNs determines that their predictions are bounded by the minimum and maximum of the training data.

6.3 Recommendations for Future Studies

This study developed N models only relying on meteorological data and satellite data, which can be applied to unmonitored watersheds, and proposed the application of these models to forestry planning and management. Based on the current results, future studies can be recommended as follows:

1. The model transferability was studied only to relatively undisturbed reference watersheds having different features and scales (from ~ 5 km^2 to ~ 129 km^2). To apply these models to forecast the effect of watershed disturbances (e.g., fire, harvesting) on water quality, several representative disturbed watersheds with several years of monitoring history should be selected. The change of seasonal and annual variation and the magnitude of the water quality parameters will be analyzed and compared to pre-disturbed status. Then the calibrated models will be modified based on the analysis of pre-disturbed and post-disturbed status. Finally, the modified calibrated models will be used to predict the post-disturbed water quality.
2. The proposed watershed similarity index can measure the transferability of calibrated nitrogen models from one watershed to other different watersheds. Further investigation should be conducted to use this similarity index guide transferring other nutrient like phosphorus models to different watersheds.
3. The watershed similarity index was formulated from watersheds on Canadian Boreal Plain and the soil chemical properties were assumed uniform within the watersheds. These watersheds have relatively high peatland cover. To apply this watershed similarity index to other watersheds with completely different features than the Boreal Plain watersheds, further investigation is needed.
4. RS VIs (e.g., EVI) is a significant input of the N models. To enable the models to provide scientific basis for harvesting planning not to impair water quality, the relationship between the change of EVI post-harvesting and the harvesting activity related parameters such as harvested volume and replanting should be established. In turn, the post-harvest EVI can be predicted and included in model inputs.
5. This study demonstrated the significance of using RS EVI as model inputs to represent vegetation dynamics and the correlation of RS EVI and RS NDWI to model transferability. Improved estimation of vegetation dynamics and coverage, precipitation, soil moisture and snow is promising with the rapid

6. The hybrid models coupling ANNs and physically based models can improve the extrapolation ability of ANNs by incorporating mechanistic principles about the modelled system. Also, the hybrid models do not have such intensive data requirement as physically based models as ANNs can be used to predict some of the model parameters. It is plausible to investigate the hybrid ANNs-physical models for water quality prediction.