DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR RIVER WATER QUALITY CLASSIFICATION BASED ON ALGAE COMPOSITION

By

Fong Wai Mei

Dissertation submitted to
UNIVERSITI SAINS MALAYSIA

In partial fulfillment of the requirements
for degree with honors

BACHELOR OF SCIENCE (MECHATRONIC ENGINEERING)

School of Electronic and
Electrical Engineering
Universiti Sains Malaysia

Mac 2006

ABSTRACT

Throughout the years, many researches have been conducted on the potential applications of Artificial Intelligence (AI) in the biological monitoring of river quality. This project will provide an overview regarding the feasibility of the application of neural networks for direct classification of river water quality based on algae composition. A brief introduction to neural networks and the suitability of neural network for use in river water quality determination will be investigated. In this project, several neural networks will be developed and their performance are compared to yield the most suitable network that will be used to model the classification system for determination of river water quality based on algae composition. Among the types of neural network that will be developed are Multilayer Perceptron network (MLP), Radial Basis Function (RBF) network and Hybrid Multilayer Perceptron (HMLP) network. This study proves that the HMLP network trained using the MRPE algorithm achieves the best performance as compared to the MLP and RBF network. The HMLP network produces 90% accuracy. In this study, an intelligent system is developed for the classification of river water quality using the HMLP network. The proposed system provides several advantages in terms of its applicability, high accuracy, user-friendliness and as well as yields faster results compared to conventional system.

ABSTRAK

TAJUK: PEMBANGUNAN SISTEM PINTAR UNTUK PENGKELASAN KUALITI AIR SUNGAI BERDASARKAN KOMPOSISI ALGA

Banyak penyelidikan telah dijalankan untuk mengkaji potensi penggunaan rangkaian neural buatan dalam pengawasan biologi kualiti air. Projek ini akan mengkaji kesesuaian penggunaan rangkaian neural buatan untuk membangunkan sistem penunjuk kualiti air sungai berdasarkan komposisi alga. Pengenalan tentang asas rangkaian neural dan seni bina beberapa rangkaian neural akan dibincangkan. Di dalam projek ini, beberapa rangkaian neural akan dibangunkan. Diantaranya ialah Perceptron Berbilang Lapisan (MLP), Perceptron Berbilang Lapisan Hibrid (HMLP) dan Fungsi Asas Jejarian (RBF). Rangkaian-rangkaian neural yang telah dibangunkan akan dibandingkan dari segi keringkasan seni bina dan kejituan dalam pengkelasan kualiti air sungai. Projek ini telah membuktikan bahawa rangkaian HMLP yang dilatih dengan algoritma pembelajaran MRPE menghasilkan kejituan yang paling tinggi jika dibandingkan dengan rangkaian MLP dan RBF. Kejituan rangkaian HMLP dalam pengkelasan kualiti air sungai adalah setinggi 90%. Suatu sistem pintar yang menggunakan rangkaian HMLP telah dibangunkan untuk tujuan pengkelasan kualiti air sungai berdasarkan komposisi alga. Sistem pintar yang dicadangkan ini mempunyai beberapa kelebihan seperti mesra pengguna, mempunyai kejituan yang tinggi dan menjimatkan masa.

ACKNOWLEDGEMENT

I would like to express my gratitude to the many people who have made my undergraduate studies possible and who have made it such a rewarding experience. First and foremost, my greatest honour and appreciation go to Dr. Nor Ashidi Mat Isa, my supervisor cum lecturer for his tireless dedication in guiding his students. His thoughts, encouragement and suggestions have contributed to the success of this project.

Special thanks to Mr. Khoo Wei Qi, my fellow coursemate, for his time and guidance which inspire me in working my project. His opinion, idea sharing and guidance in C++ Builder programming language proved to be very important during my project execution period. Thanks to Mr. Fakroul, postgraduate student for his ideas throughout the implementation of this project. Many thanks to Ms. Lim Chia Li who had assisted me unconditionally.

Next, I would like to thank my family and friends who have supported me throughout the past 4 years of my studies in USM. Thanks to my parents, who have always encouraged me to do my best. A special acknowledgment goes to my brother for his thoughtfulness of buying me a new notebook for the convenience of doing my final year project.

My acknowledgments would not be complete without expressing my personal belief in and gratitude towards God, our creator and redeemer. I feel very fortunate to have had the opportunity to study in the field of engineering in such a prestige university in Malaysia. I believed that it is truly God’s wisdom and blessings that had brought me thus far.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Abbreviation	Description
AI	Artificial Intelligence
ANNs	Artificial Neural Networks
ASPT	Averaqe Score Per Taxon
BBNs	Bayesian Belief Networks
BMWP	British Monitoring Working Party
BP	Back-Propagation
BR	Bayesian Regularization
GLS	Givens Least Square
GUI	Graphical User Interface
HMLP	Hybrid Multilayered Perceptron
LM	Levenberg-Marquardt
MLP	Multilayered Perceptron
MOPED	MOdelling Patterns in Environmental Data
MRPE	Modified Recursive Prediction Error
NIWA	National Institute of Water and Atmospheric Research
RBF	Radial Basis Function
RPE	Recursive Prediction Error
SOM	Self-Organizing Map

CHAPTER 1

INTRODUCTION

1.1 Background of River Water Quality Classification

"Water quality" is a term used here to express the suitability of water to sustain various uses or processes. Any particular use will have certain requirements for the physical, chemical or biological characteristics of water; for example limits on the concentrations of toxic substances for drinking water use, or restrictions on temperature and pH ranges for water supporting invertebrate communities. Consequently, water quality can be defined by a range of variables which limit water use. There is increasing recognition that natural ecosystems have a legitimate place in the consideration of options for water quality management. This is both for their intrinsic value and because they are sensitive indicators of changes or deterioration in overall water quality, providing a useful addition to physical, chemical and other information (Meybeck et. al., 1996).

Biological monitoring of river quality has grown in importance over the past few decades due to the recognition of important advantages over chemical monitoring. Bio-monitoring requires the development of tools with the capacity to interpret biological and environmental variables in terms of chemistry and vice versa. The response of benthic macroinvertebrate and macrophyte communities in rivers to environmental stresses of various types is acknowledged, and scientifically utilized as a means of assessing water quality. Bio-monitoring of water, water bodies and effluents is based on five main approaches:

1. Ecological methods
    - analysis of the biological communities (biocenoses) of the water body,
    - analysis of the biocenoses on artificial substrates placed in a water body, and
    - presence or absence of specific species.
2. Physiological and biochemical methods
    - oxygen production and consumption, stimulation or inhibition,
    - respiration and growth of organisms suspended in the water, and
    - studies of the effects on enzymes.
3. The use of organisms in controlled environments
    - assessment of the toxic (or even beneficial) effects of samples on organisms under defined laboratory conditions (toxicity tests or bioassays), and
    - assessing the effects on defined organisms (e.g. behavioral effects) of waters and effluents in situ, or on-site, under controlled situations (continuous, field or “dynamic” tests).
4. Biological accumulation
    - studies of the bioaccumulation of substances by organisms living in the environment (passive monitoring), and
    - studies of the bioaccumulation of substances by organisms deliberately exposed in the environment (active monitoring).
5. Histological and morphological methods
    - observation of histological and morphological changes, and
    - embryological development or early life-stage tests.

The following table summarizes the basic characteristics of three of the commonly used European methods of biological monitoring. The methods considered are the Saprobic System, the Trent Biotic Index, the BMWP score and ASPT (Walley, 1994).

Table 1.1: Summary of the characteristics of biological monitoring methods.

No.	Methods	Description	Remarks
1	The Saprobic System	▪ All taxa are identified to species level. ▪ Each species is allocated an indicator value. ▪ Abundance levels are included. ▪ 'Absent' evidence is excluded. ▪ No allowance is made for seasonal behavior.	This system implicitly assumes independence. It is based upon "frequency" distributions but the uncertainties they imply are not preserved in the method of combination. It utilize an ad hoc weighted average procedure.
2	The Trent Biotic Index	▪ Some key indicator 'groups' are used. ▪ Abundance of individuals is not used but abundance of species within 'groups' and of 'groups' are used. ▪ The absence of key threshold 'groups' can have a major impact on the classification. ▪ No allowance is made for seasonal behavior.	Independence is not required in this system. It is a kind of pattern recognition system which uses key threshold 'groups' and the total number of 'groups' as the main features on which to base its classification. It disregards the inherent uncertainties in the presence/absence of the key 'groups' and thus it is somewhat 'brittle'. It is an ad hoc look up procedure based upon experience. The 'groups' are defined at various taxonomic levels.

3	BMWP score and Average Score per Taxon (ASPT)	▪ Identification is to family level only. ▪ No distinction is made between families in terms of their indicator value. ▪ Abundance levels are not included. ▪ ASPT specifically excludes 'absent' evidence but BMWP includes it, giving it a zero score. ▪ No allowance is made for seasonal behavior.	This system implicitly assumes independence. The method allocates a score (1-10) to each family on the basis of the pollution sensitivity of its least sensitive species. The BMWP Score is the total of the scores of each family found in the sample and the ASPT is the average of these scores. This method is ad hoc and disregards uncertainty.

1.2 River Water Classification based on Algae Composition

Periphyton are benthic algae that grow attached to surfaces such as rocks or larger plants. Due to the sedentary nature of periphyton, the community composition, structure, and biomass are sensitive to changes in water quality and are often used as indicators of ambient conditions. Because periphyton are attached to the substrate, this assemblage integrates physical and chemical disturbances to the stream reach. The periphyton assemblage serves as a good biological indicator due to:

- its rapid reproduction rates and very short life cycles - indicate short-term impacts
- its naturally high number of species
- its rapid response time to both exposure and recovery
- its direct connection with physical and chemical factors
- its identification to a species level by experienced biologists
- the ease of sampling which require only a few people

- the tolerance or sensitivity to specific changes in environmental condition are known for many species

Diatoms in particular are useful indicators of biological condition because they are ubiquitous and found in all lotic systems. By using algal data in association with macroinvertebrate and fish data, the strength of biological assessments is optimized. The objectives of a rapid bioassessment protocol for periphyton could include, but would not be limited to, assessment of biomass (*chlorophyl a* or ash-free dry mass), species, composition and biological condition of periphyton assemblages. These information are cited from internet sources URL *http://www.epa.gov/bioindicators/html/periphyton.html*, 2005.

Studies of algal, combined with macroinvertebrate communities, provide a valuable assessment of the overall health of aquatic systems. The objectives of a rapid bioassessment protocol for periphyton could include assessment of biomass (chlorophyl a or ash-free dry mass), species, composition and biological condition of periphyton assemblages. The following information are cited from internet sources URL *http://www.dep.state.fl.us/labs/biology/algae.htm*, 2005. There are several methods for the measurements of algal community health. These are:

1. Taxa richness which indicates the number of different types of organisms present in a system.
2. Shannon-Weaver diversity – an index that measures the distribution of organisms present. Low diversities represent conditions where only a few organisms are abundant, to the exclusion of other taxa.

3. Numbers of pollution sensitive taxa for example certain taxa are labeled as sensitive or tolerant to pollution.
4. Community structure which is the measurements of shifts in proportions of major groups of organisms, compared to reference conditions.
5. Algal biomass which indicates the amount of algal growth that a water body can support, measured as algal density or chlorophyll *a*.
6. Habitat Assessment which refers to the quality of the local environment with respect to the needs of the organisms investigated.

A review of published reports in the United States surface water quality management programs identified six states that use information derived from communities of algae as a water quality management tool (Kroeger et. al., 1999). Many respondents acknowledged the merits of algae as a biological indicator, but some states have chosen not to use algae because of the lack of qualified staff or the costs associated in developing a new program that would only complement an existing one (benthic macroinvertebrates or fish).

The six states that utilize information from periphyton communities as a water quality management tool are Florida, Idaho, Kentucky, Massachusetts, Montana and Wyoming. In general these states use information derived from a quantitative enumeration of species observed at a site. Information may include species diversity, a summary of pollution tolerant (or intolerant) species and a comparison of species composition with a control (unimpacted) site, if one was available.

1.3 Objective and Scope of Project

The main objective of this project is to develop an intelligent system using neural networks to classify river water quality based on algae composition. The performance of various neural networks using different training algorithm will be compared to identify the best network with the highest accuracy. The feasibility of the application of neural networks in the classification of river water quality will be examined. The software system and the final user interface will be developed using Borland C++ Builder.

The scope of this project includes the development of Multilayered Perceptron network (MLP) and Radial Basis Function (RBF) network using the Neural Network toolbox in MATLAB as well as the Hybrid Multilayered Perceptron (HMLP) network using Borland C++ Builder version 6.0. The RBF network is trained using the moving k-means and givens least squares (GLS) algorithm and the MLP network is trained using the Back-propagation (BP), Levenberg-Marquardt (LM) as well as the Bayesian Regularization (BR) algorithm. The Hybrid Multilayered Perceptron (HMLP) network is trained using the modified recursive prediction error algorithm or better known as MRPE algorithm.

The data for this project were obtained from the 'Pinang' river and were divided into two parts which is training data and testing data. There were altogether 200 data; 120 data for training and 80 data for testing. The suitability of these data will also be investigated based on the network performance.

This project however, focuses mainly on the development of HMLP network which is predicted to yield the best results as compared to MLP and RBF networks. Finally, the user's interface with basic features such as processing the input data by user and displaying the results in the form of a graph is developed.

1.4 Report Layout

Chapter 1 is an introduction intended to provide a short review regarding the background of river water quality classification. A brief discussion about the various river water classification schemes is presented. Objectives and scope of the project is specified in this chapter as to clarify the purpose of this thesis.

Chapter 2 provides a short insight into the fundamental of artificial neural networks. It is a short introduction about the various types of neural network and the commonly used training algorithm. A literature review regarding the application of artificial intelligence and neural networks in river water quality classification is enclosed.

Chapter 3 outlines the methodology for the implementation of this project. It contained a thorough explanation regarding the basic architecture of each neural network and its training algorithm. A step-by-step procedure for data preparation, the development and training of neural networks and the complete AI system with user's interface will be presented in this chapter.

Chapter 4 contained the results obtained from the simulation of the various neural networks developed. The performance of all the neural networks developed will be compared in terms of optimum structure and accuracy. Relevant discussions regarding the performance of the final AI system are enclosed.

Chapter 5 draws the final conclusion of this project and contains future suggestions for the improvement and continuation of this project.

CHAPTER 2

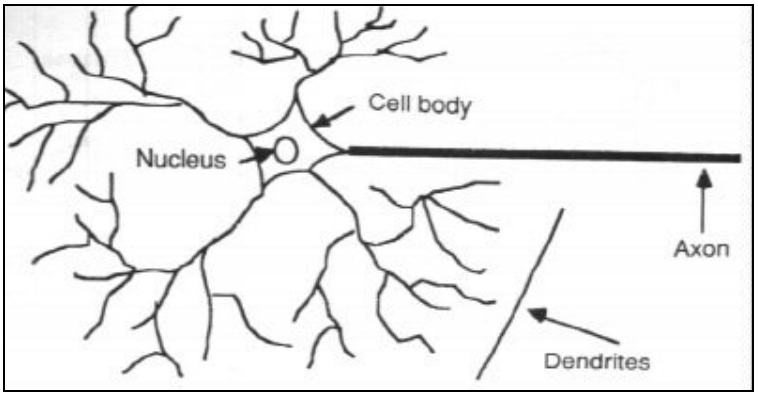
LITERATURE REVIEW

2.1 Introduction

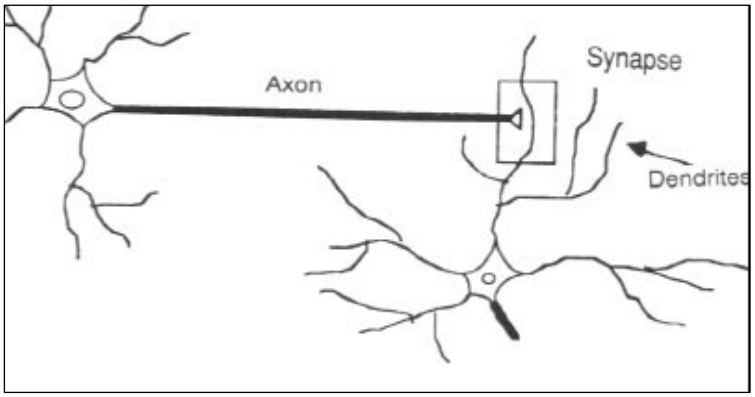
This chapter provides an overview about biological neurons and artificial neural networks. A few types of neural networks will be discussed briefly together with the commonly used training algorithm. Literature reviews regarding the application of neural networks in river water classification will be presented at the end of this chapter.

2.2 The Biological Neuron

The nervous system of living organisms as shown in Figure 2.1(a) is a structure consisting of many elements working in parallel and in connection with one another. In the human brain, a typical neuron collects signals from others through a host of fine structures called dendrites. The neuron is a many-inputs / one-output unit. The neuron sends out spikes of electrical activity through an axon, which splits into thousands of branches. The information transmission happens at the synapses. The spikes traveling along the axon of the pre-synaptic neuron trigger the release of neurotransmitter substances at the synapse (refer Figure 2.1(b)). The neurotransmitters cause excitation or inhibition in the dendrite of the post-synaptic neuron. When a neuron receives excitatory input that is sufficiently large compared with its inhibitory input, it sends a spike of electrical activity down its axon. The contribution of the signals depends on the strength of the synaptic connection (Stergiou et. al., 2005)



(a)



(b)

Figure 2.1: (a) Biological neuron and (b) The synapses

2.3 Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) are computational systems whose architecture and operation are inspired from biological neural cells (neurons) in the brain. ANNs can be described either as mathematical and computational models for non-linear function approximation, data classification, clustering and non-parametric regression or as simulations of the behavior of collections of model biological neurons. These are not simulations of real neurons in the sense that they do not model the biology, chemistry, or physics of a real neuron. They do, however, model several aspects of the information combining and pattern recognition behavior of real neurons

in a simple yet meaningful way. Neural modeling has shown incredible capability for emulation, analysis, prediction, and association. ANNs can be used in a variety of powerful ways for example to memorize characteristics and features of given data and to match or make associations from new data to the old data.

ANNs are simple models of the structure and function of the brain. ANNs are capable of solving difficult problems in a way that resembles human intelligence. What is unique about neural networks is their ability to learn by example. Traditional artificial intelligence (AI) solutions rely on symbolic processing of the data, an approach which requires a priori human knowledge about the problem. Neural networks also have an advantage over statistical methods of data classification because they are distribution-free and require no knowledge about the statistical distributions of the classes in the data sources in order to classify them. Unlike these two approaches, ANNs are capable of solving problems without any assumptions.

The first model of a neuron was proposed in 1943 by McCulloch and Pitts when they described a logical calculus of neural networks. The McCulloch-Pitts neuron models connected up in a simple fashion (forming a single layer), were given the name "perceptrons" by Frank Rosenblatt in 1962. In his book "Principles of Neurodynamics" he described the properties of these neurons, but more importantly he presented a method by which the perceptrons could be trained in order to perform simple patterns recognition tasks. He also provided a theorem called the perceptron convergence theory which guarantees that if the learning task is linearly separable (that is, if the data classes can be separated by a straight line in input space) then the perceptron will yield a solution in a finite number of steps (Ampazis, 2005).

A neuron model consists of three main elements; synapses or connecting links, adder and activation function (as shown in Figure 2.2). The neuron has K input lines and a single output. Each input signal is weighted. It is multiplied with the weight value of the corresponding input line (by analogy to the synaptic strength of the connections of biological neurons). The neuron will combine these weighted inputs by forming their sum and, with reference to a threshold value and activation function, it will determine its output. The model of neuron j can be represented by:

$$net_j = \sum_{k=1}^{K} w_{jk} x_k - \theta_j \tag{2.1}$$

$$y_j = g\left(net_j\right) \tag{2.2}$$

where

$x_1, x_2, \ldots, x_K$ are the input signals

$w_{j1}, w_{j2}, \ldots, w_{jK}$ are the synaptic weights converging to neuron j

net_j is the cumulative effect of all the neuron connected to neuron j

θ_j is the threshold of neuron j

$g(\bullet)$ is the activation function

y_j is the output signal of the neuron

The most common form of activation function used is the sigmoid function.

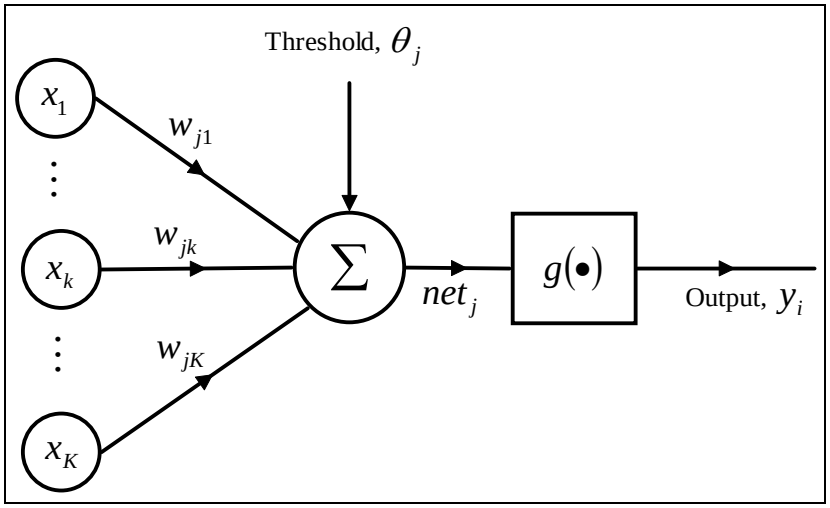


Figure 2.2: A neuron model

There are several types of ANN, however only two types are discussed here. They are feed-forward ANN (Figure 2.3) and recurrent ANN (Figure 2.4). Feed-forward neural networks allow signals to propagate in one direction; from input to output. There is no feedback (loop) i.e. the output of any layer does not affect that same layer or the layer before. Feed-forward neural networks tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition.

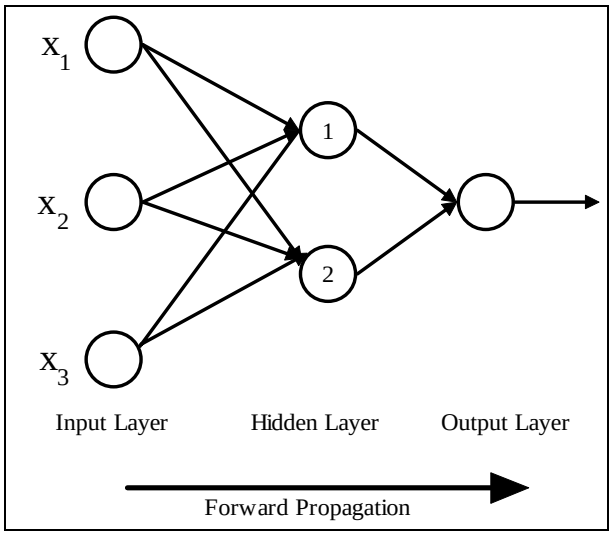


Figure 2.3: Feed-forward ANN

On contrary, feedback networks can have signals traveling in both directions by introducing loops in the network. It also allows self feedback represented by dotted loop as shown in Figure 2.4. Feedback networks are dynamic; their 'state' is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feedback architectures are also referred to as interactive or recurrent, although the latter term is often used to denote feedback connections in single-layer organizations. From training examples recurrent netwroks can learn to map input sequences to output sequences (Haykin, 1994).

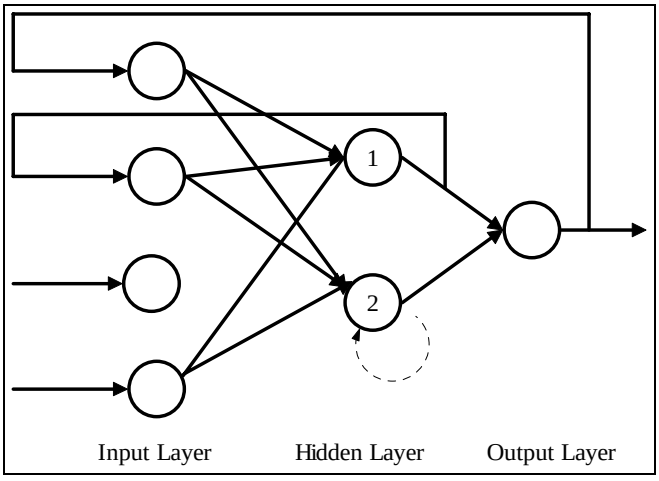


Figure 2.4: Recurrent ANN

2.3.1 Learning in Artificial Neural Networks

There are two basic types of learning paradigms: supervised learning and unsupervised (self-organized) learning. Supervised learning is performed under the supervision of an external teacher. The teacher provides the network with a desired or target response for any input vector. The actual response of the network to each input vector is then compared by the teacher with the desired response for that vector, and the network parameters are adjusted in accordance with an error signal which is defined as

the difference between the desired response and the actual response. The adjustment is carried out iteratively in a step-by-step fashion with the aim of eventually making the error signal for all input vectors as small as possible. In contrast to supervised learning, unsupervised or self-organized learning does not require an external teacher. During the training session, the neural network receives a number of different input patterns and learns how to classify input data into appropriate categories (Ampazis, 2005).

2.3.2 Multilayered Perceptron (MLP) networks

A typical Multilayered feed-forward ANN is shown in Figure 2.5. This type of network is also known as a Multilayer Perceptron (MLP) network. The units (or nodes) of the network are nonlinear threshold units. The units are arranged in layers and each unit in a layer is connected to the units of a preceding layer but it does not have any connections to units of the same layer to which it belongs. The layers are arrayed one succeeding the other so that there is an input layer, multiple intermediate layers and finally an output layer. Intermediate layers are called hidden layers. Figure 2.5 shows a MLP with only one hidden layer. Back-propagation neural networks such as MLP are usually fully connected i.e every node in each layer is connected to every other node in the adjacent forward layer. Generally, the input layer is not considered as neurons as they do not perform any computation. Therefore, the hidden layer is referred as the first layer of the network (Ampazis, 2005).

The most common learning algorithm for MLP networks is the backpropagation (BP) algorithm. The following discussion regarding the BP algorithm is based on the book by Haykin (1994). Back-propagation is a gradient descent procedure that attempts to reduce the errors between the output of the network and the desired result. The back-

propagation training consists of three phases of computation: forward propagation, backward propagation and weight adaptation phase. In the forward propagation an input pattern vector is applied to the sensory nodes of the network that is, to the units in the input layer. The signals from the input layer propagate to the units in the first layer and each unit produces an output. The outputs of these units are propagated to units in subsequent layers (in this case the hidden layers) and this process continues until the signals reach the output layer where the actual response of the network to the input vector is obtained. During the forward propagation the synaptic weights of the network are fixed. The error terms of hidden and output nodes are computed in the backward propagation and propagated backward through the network against the direction of synaptic connections. Finally the weight for each connection is adapted accordingly. BP will stop once the convergence criterion is met.

Other learning algorithms for MLP networks are Bayesian Regularization (BR) algorithm and the Levenberg-Marquardt (LM) algorithm. Levenberg-Marquardt is a popular alternative to the Gauss-Newton method of finding the minimum of a function F(x) that is a sum of squares of nonlinear functions. It outperforms simple gradient descent and other conjugate gradient methods in a wide variety of problems. Bayesian Regularization is a powerful add-on to the training of neural networks which enables an automatic optimization of the weight decay parameter. Thus, both over-fitting and over-smoothing are prevented efficiently. A combination of the Levenberg-Marquardt and Bayesian Regularization implies very little additional computational costs, since it exploits the approximation to the Hessian. Therefore, the weight decay can be updated easily after each training cycle (Haykin, 1994).

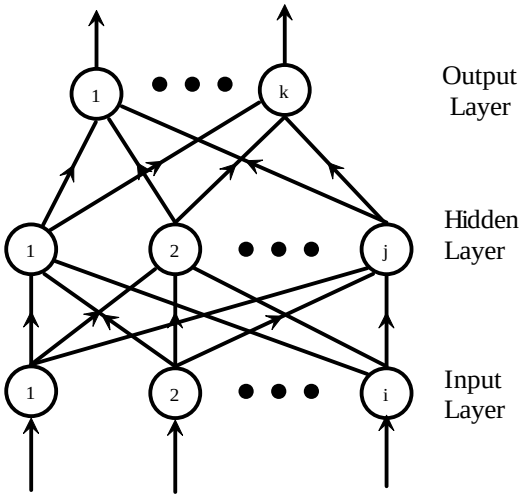


Figure 2.5 Multilayered Perceptron (MLP) network

2.3.3 Hybrid Multilayered Perceptron (HMLP) networks

According to the paper by Mashor (2000), HMLP network is the combination of the conventional MLP network with additional linear input connections. HMLP allows direct connection from network inputs to output nodes via some weighted connections forming a linear model in parallel with the nonlinear original MLP model. A HMLP network with one hidden layer is shown in figure 2.6. The HMLP network offers a faster learning rate and gives better performance than the standard MLP network with more hidden nodes. The additional linear input connections in the HMLP network do not significantly increase the complexity of the MLP network since the connections are linear. In fact, by using the linear input connections, the required number of hidden nodes can be reduced, which will also reduce the computational load (Mashor, 2000).

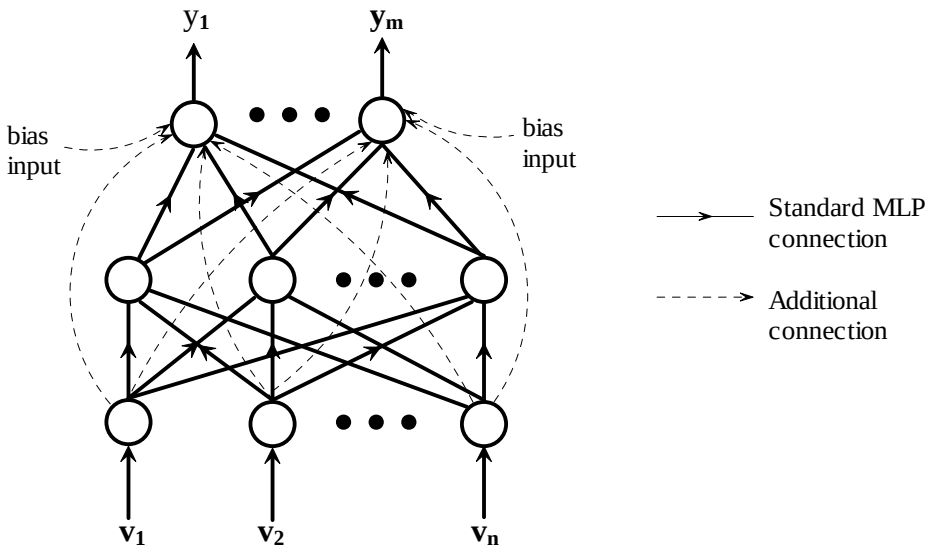


Figure 2.6 Hybrid Multilayered (HMLP) network

The learning algorithm that will be used to train the HMLP network is known as the modified recursive prediction error (MRPE) algorithm. MRPE algorithm is actually a modified version of the RPE algorithm originally derived by Ljung and Soderstrom (1983) to train MLP networks. The RPE algorithm which is a Gauss Newton type algorithm generally yield a better performance compared to the steepest descend type algorithm such as the back-propagation algorithm. This is due to the ability of RPE algorithm to provide a faster convergence rate and better final convergence values of weights and thresholds for the MLP network. The architecture of HMLP network and the MRPE algorithm will be discussed further in section 3.2.2 in chapter 3.

2.3.4 Radial Basis Function (RBF) networks

Radial basis function (RBF) networks shown in Figure 2.7 are supervised feed-forward networks. They are typically configured with a single hidden layer of units which activation function is selected from a class of functions called basis functions. While similar to MLP networks with back propagation algorithm in many respects,

radial basis function networks have several advantages. They are usually trained much faster than back propagation networks (Haykin, 1994). They are less susceptible to problems with non-stationary inputs because of the behavior of the radial basis function hidden units. Radial basis function networks are similar to the probabilistic neural networks in many respects (Wasserrnan 1993). Popularized by Moody and Darken (1989), radial basis function networks have been proven to be an useful neural network architecture. The major difference between radial basis function networks and back propagation networks is the behavior of the single hidden layer. Rather than using the sigmoidal or S-shaped activation function as in back propagation, the hidden units in RBF networks use a Gaussian or some other basis kernel function. Each hidden unit acts as a locally tuned processor that computes a score for the match between the input vector and its connection weights or centers. In effect, the basis units are highly specialized pattern detectors. The weights connecting the basis units to the outputs are used to take linear combinations of the hidden units to produce the final classification or output.

In radial basis function networks, the weights into the hidden layer basis units are usually set before the second layer of weights is adjusted. As the input moves away from the connection weights, the activation value falls off. This behavior leads to the use of the term “center” for the first-layer weights. These center weights can be computed using statistical methods such as K-Means clustering. In any case, they are then used to set the areas of sensitivity for the RBF hidden units, which then remain fixed. Once the hidden layer weights are set, a second phase of training is used to adjust the output weights. This process typically uses the standard back propagation training rule (Haykin, 1994).

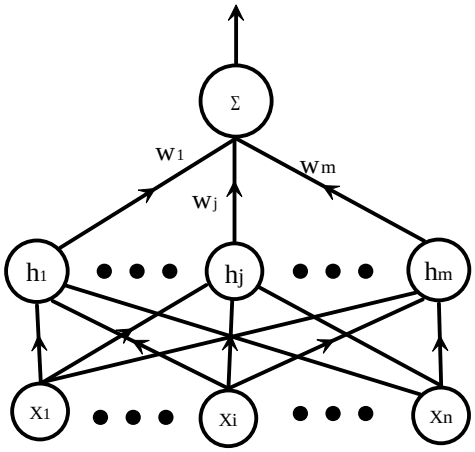


Figure 2.7: Radial Basis Function Network

2.4 Application of Artificial Intelligence in River Water Classification

Artificial Intelligence is defined as a discipline concerned with the building of computer programs that perform tasks requiring intelligence when done by humans. In this project it refers to the task of interpreting biological data into river water quality. When this is done by human experts it involves two complimentary mental processes: scientific reasoning and pattern recognition. The Artificial Intelligence (AI) field offers two powerful techniques for the interpretation of biological data namely, Bayesian belief networks and neural networks. Bayesian belief networks provide a method of probabilistic reasoning that is capable of emulating the essential features of human reasoning. They make good use of existing scientific knowledge and understanding, but generally little use of field data.

Professor Bill Walley, the head of the Centre for Intelligent Environmental Systems at Staffordshire University, UK, is a world leader in the application of artificial intelligence (AI) techniques to environmental science. Walley et. al. (2000) through Environment Agency R&D Technical Reports and Toxicology and

Ecotoxicology News has approached the issue of bio-monitoring in rivers by trying to model how expert stream ecologists assess stream health. In general, experts use two complementary mental processes when diagnosing or predicting problems in their domain of expertise. These are:

(a) plausible ("probabilistic") reasoning based upon their scientific knowledge;

(b) pattern recognition based upon their experience of previous cases.

The robust and holistic nature of Bayesian belief networks (BBNs) makes them well-suited to modelling complex systems, like river ecology, where the interaction between the elements is probabilistic in nature.

Walley (November–December 2000) provided insights into a range of AI techniques that seem ideally suited to both the interpretation of complex ecological patterns. Two of the areas of particular interest were the use of Bayesian Belief Networks (BBNs), also known as Plausible Reasoning Networks, for relating stream health to environmental factors and the use of unsupervised neural networks (such as Self-Organising Maps, or SOMs) for mapping patterns in environmental data. In year 2000, NIWA (National Institute of Water and Atmospheric Research) in USA has developed a software for the construction of models that can predict biological assemblages from habitat data (for example, predict what species should be present in a certain stream). This software, called MOPED (MOdelling Patterns in Environmental Data), uses the SOM approach in conjunction with discriminant functions analysis

2.5 Application of Neural Networks in River Water Classification

Artificial neural networks have become a popular modeling tool for highly complicated phenomena. The main function of an ANN is to map between input and output data sets. They have been successfully applied to hydrological processes (Hsu et. al. 1995), modeling of plankton dynamics and algal blooms (Recknagel et. al. 1997) and water quality modeling (Jian and Eheart 2003).

ANNs were first applied to the task of learning to predict algal blooms from water quality databases by French and Recknagel (1994). In this application, a feedforward ANN was trained to make predictions of abundance of species of phytoplankton in the Saidenbach reservoir, Germany. The architecture of the ANN was such that the database measurements fed to the input layer represented forcing functions controlling in-lake processes such as underwater light conditions, temperature, nutrient dynamics and zooplankton. Phytoplankton productivity or algal blooms was represented in the output layer. The input and output layer were interconnected via a hidden layer consisting an arbitrary number of neurons. The backpropagation training algorithm was applied to the ANN structure to map the relationship between the input and output layers given data collected over several years from the reservoir. The model was validated by accessing its predictive performance on a subsample of 2 years of data held out from training.

Since then, applications of ANN modeling for predicting algal biomass in freshwater ecosystems based on this methodology have been developed for Lake Kasumigaura (Japan), Lake Biwa (Japan), Lake Tuusulunjarvi (Finland) and the Darling River (Australia) (Recknagel et. al 1997). The potential for elucidation of interactions in freshwater ecosystems from ANN models by means of sensitivity and scenario analyses has been demonstrated in Recknagel et. al. (1997). Recknagel et. al.

(1998) further enhanced the model for predicting biomass of algal species in Lake Kasumigaura through the development of and ensemble ANN where separate ANNs were trained for different years in the time series and the best model year was selected.

Further improvement towards dynamic ecosystem modeling was seen in Kwang et. al (2001) whereby the recurrent artificial neural network was used for time series modeling of phytoplankton dynamics in the hypertrophic Nakdong river system in Korea. The recurrent algorithm was adopted for ANN training based on 4 years (1995-1998) of meteorological, physico-chemical and biological data of the river and ANN validation by means of data for an independent year (1994). The results from this study indicate that time series modeling of the Nakdong River by ANN proved to be suitable and useful for both prediction and elucidation of algal dynamics.

In conclusion, all the previous researches conducted by various individuals discussed above have proven the feasibility of ANN in modeling and prediction of very complex and nonlinear ecological phenomena such as algal blooms.

2.6 Summary

Artificial intelligence (AI) especially artificial neural networks (ANNs) have been used widely in various applications such as pattern recognition, vision, speech recognition, classification, and control systems. Their ability to learn by example makes them very flexible and powerful. Furthermore there is no need to devise an algorithm in order to perform a specific task; i.e. there is no need to understand the internal mechanisms of that task. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex. Therefore, ANNs have been successfully applied in the modeling and prediction of highly complicated ecological systems. The study of various types of neural networks and its characteristics is crucial to enhance the neural network to its maximum potential.