MEASURING, MONITORING AND FORECASTING SYSTEM FOR CARBON MONOXIDE CONCENTRATIONS

by

USHARANI D/O RAJI

UNIVERSITI SAINS MALAYSIA

2005

MEASURING, MONITORING AND FORECASTING SYSTEM FOR CARBON MONOXIDE CONCENTRATIONS

by

USHARANI D/O RAJI

Thesis submitted in fulfillment of the

requirements for the degree

of Master of Science

May 2005

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor, Assoc. Prof. Dr. Mohd. Yusoff Mashor and my co-supervisor, Encik Ahmad Nazri Ali for their guidance, support, comments and discussions. With their help, this research work has improved tremendously.

My appreciation also goes to the lab technicians, Encik Khairul and Encik Latiff, for providing me with the necessary equipment to help in this research work.

A big thank you goes to my mother and sisters for their advice, love and prayers all the time. Special thanks also go to my friends, lab colleagues who have helped me one way or another to remove the obstacles and difficulties faced in this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

MEASURING, MONITORING AND FORECASTING SYSTEM FOR CARBON MONOXIDE

ABSTRACT

The impact of air pollution is broad towards the health of human beings, thus many studies have been focused on the forecasting of air pollutants in urban areas. In this research, a portable carbon monoxide forecaster has been developed in order to forecast the future level of carbon monoxide (CO) concentrations, since CO is primary pollutant which affects the air quality level. First, a PC based CO forecaster is developed in order to choose the best neural network and forecasting model. The chosen network together with its modeling will be later implemented in the portable system. From the comparison study, Hybrid Multilayered Perceptron (HMLP) network is chosen since it is found to perform better compared to Multilayered Perceptron (MLP), Recurrent and Radial Basis Function (RBF) networks, respectively. In this study, the HMLP network is trained with Modified Recursive Prediction Error (MRPE) algorithm to perform CO concentrations level forecasting. Besides that, only past CO concentration values are used as network input series, in order to perform forecasting. Meanwhile, on-line model is chosen since it is found to perform better compared to off-line model. This is due to its flexibility to update the network parameters for each data sample. The portable CO forecaster is developed by using dsPICDEM 1.1 development board, which uses digital signal controller dsPIC30F6014 device. The portable CO forecaster is able to perform CO concentrations measurement and forecasting up to 8 steps ahead. Besides that, the portable CO forecaster has the ability to store the measurement and forecasting values, in order to be transferred to PC through serial communication later on. HMLP network together with on-line model is

implemented into the architecture of portable system. The developed portable CO forecaster is evaluated by using 3 real data sets and 2 simulated environment data sets, respectively. The R^2 values achieved by these data sets for OSA test and 8 steps ahead forecasting are 0.9805 and 0.8501, 0.9677 and 0.7910, 0.8777 and 0.1524, 0.9966 and 0.8994, 0.9942 and 0.4097, respectively. The portable CO forecaster only uses 5 past CO concentration values, thus it can be concluded that the portable system gave excellent results over those data sets. Overall, it can be concluded that the developed portable CO forecaster can provide accurate and high accuracy of CO concentrations forecasting.

SISTEM PENGUKURAN, PENGAWASAN DAN PARAMALAN KEPEKATAN KARBON MONOKSIDA

ABSTRAK

Kesan pencemaran udara amat meluas terhadap kesihatan manusia, justeru itu banyak kajian telah diketengahkan untuk meramal pencemar udara di kawasan bandar. Dalam kajian ini, suatu alat mudah alih untuk meramal kepekatan gas karbon monoksida telah dibangunkan untuk meramal kepekatan karbon monoksida pada masa hadapan. Karbon monoksida telah dipilih kerana ia merupakan pencemar udara utama yang boleh menjejaskan kualiti udara. Pada mulanya, suatu peramal untuk kepekatan karbon monoksida telah dibangunkan berasaskan komputer, ia berfungsi untuk memilih rangkaian neural dan model peramalan yang terbaik. Rangkaian neural dan model yang dipilih, akan diimplementasikan pada sistem mudah alih. Daripada perbandingan prestasi, rangkaian Perseptron Berbilang Lapisan Hibrid (HMLP) telah dipilih kerana ia memberi prestasi yang lebih baik berbanding dengan rangkaian Perseptron Berbilang Lapisan (MLP), Fungsi Asas Jejarian (RBF) dan Perulangan (Recurrent). Rangkaian HMLP telah dilatih dengan menggunakan algoritma ralat ramalan jadi semula ubahsuai untuk meramal kepekatan karbon monoksida. Nilai kepekatan gas karbon monoksida yang lampau akan digunakam sebagai masukan kepada rangkaian neural tersebut. Manakala, model dalam talian telah dipilih kerana ia menunjukkan prestasi yang lebih baik berbanding dengan model luar talian. Ini kerana model dalam talian mengemaskinikan parameter rangkaian neural bagi setiap masukan data. Dalam kajian ini, sistem mudah alih untuk meramal kepekatan karbon monoksida dibina dengan menggunakan papan dsPICDEM 1.1, yang menggunakan pengawal isyarat digital, dsPIC30F6014. Sistem mudah alih tersebut boleh

mengukur dan meramal kepekatan karbon monoksida dari 1 hingga 8 langkah ke hadapan. Selain daripada itu, sistem tesebut juga dapat menyimpan nilai pengukuran dan peramalan, supaya ia boleh dihantar ke komputer melalui perhubungan data siri. Rangkaian HMLP dengan model dalam talian telah diimplementasikan dalam struktur sistem mudah alih tersebut. Sistem mudah alih yang dibina akan dinilaikan dengan menngunakan 3 data sebenar and 2 data daripada keadaan penyahlakuan. Nilai R^2 yang diperolehi bagi ujian OSA dan 8 langkah ke hadapan adalah 0.9805 dan 0.8501, 0.9677 dan 0.7910, 0.8777 dan 0.1524, 0.9966 dan 0.8994, 0.9941 dan 0.4097, masing-masing. Sistem mudah alih tersebut hanya menggunakan 5 nilai kepekatan karbon monoksida lampau, maka boleh disimpulkan bahawa sistem tersebut memberi keputusan yang baik bagi data-data tersebut. Secara keseluruhannya, boleh disimpulkan bahawa sistem peramal karbon monoksida mudah alih yang dibina mampu memberi ketepatan yang tinggi bagi meramal kepekatan karbon monoksida.

CHAPTER 1

INTRODUCTION

1.1 Introduction

The impact of air pollution is broad especially towards human beings (WHO, 1987) since it can cause irritation, odour annoyance, acute and long term toxic effects (Maffeis, 1999). In this preliminary study, Carbon Monoxide (CO) is chosen as an air pollution indicator since it is a primary pollutant in urban areas, due to the major emission from motor vehicles. The average atmospheric lifetime of CO is about two to four months, therefore human health will be directly affected in long term wise (City of Fort Collins, CO, 1996). According to the Journal of American Medical Association (JAMA), 1500 people die annually due to accidental CO poisoning and 10000 people seek medical attention. People suffering from asthma, heart or respiratory problem and small children can be affected with CO poisoning more quickly than others.

Effective strategies should be implemented in order to reduce the health risk of human beings, such as better car technologies, proper planning of building to reduce heating at certain area, traffic limitations and others. The successfulness of these strategies can only be achieved with widespread alert and warning towards people. Currently, traffic limitation can be considered in urban areas compared to the other approaches due to the time required to perform those strategies. Traffic limitation should be made before the CO concentrations level exceeds its hazardous value. In this case, a potential forecasting tool is required to forecast the future level of CO concentrations level in order to take preventive actions. Many researches have been carried out in order to perform CO and other pollutants gas forecasting with the main objective of reducing health risk of human beings and to give warning towards people (Viotti *et al*. (2002), Boznar *et al.* (1993), Gardner & Dorling (1999), Moseholm *et al.* (1996), Niska *et al.* (2004)).

In this research, a measuring, monitoring and forecasting system for CO concentrations is developed. CO forecasting are performed by two major components, referring to PC and dsPIC microcontroller embedded system, respectively. This research implements neural network to perform CO concentrations forecasting, since neural networks were reported to perform better compared to other statistical analysis approach. The warning system will be generated based on the forecasted value, enables preventive action to be taken before CO concentrations rise to dangerous level and give alert to the surrounding people.

1.2 Air Pollutants Forecasting Using Neural Networks

In the past few years, statistical analysis and neural networks were widely used to perform CO and other pollutants gas forecasting. Statistical analysis method widely known as time series model such as Box-Jenkins, regression and others were applied to perform forecasting. Recently, many studies have been reported on the implementation of neural networks in environmental studies, which were presented for air pollutants forecasting (Moseholm *et al.* (1996), Viotti *et al*. (2002), Gardner & Dorling (1999), Kolehmainen *et al.* (2001), Lu *et al.* (2004)). Neural networks have been mathematically proved to be capable of representing nonlinear mappings (Cybenko, 1989, Funahashi, 1989 and Hornik, 1991). Neural networks have the ability to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Other advantages include adaptive learning, self organisation and real time operation.

Several authors have made some attempts to perform comparative studies between time series analysis and neural networks implementation in application to atmospheric studies. The studies proved that neural networks performed better compared to statistical methods. Gardner & Dorling (1999) proved that neural network was found to give better results compared to multiple regression model and auto regressive model. In another study, Perez *et al.* (2000) presented performance comparison between persistence, linear regression and multilayer neural network. Multilayer neural network was found to perform better compared to the others. In another approach, Chelani *et al.* (2002) reported that neural network gave better predictions with those given by multivariate regression model.

In the present study, neural network will be implemented to perform CO concentrations forecasting. In literature, forecasting of CO and other pollutants gas using neural networks were always carried out using off-line modelling. The networks were only trained for a limited data set, and the final network parameters will be tested out using an independent testing set. Usually, the network is trained repeatedly until minimum prediction error is achieved. One of the disadvantage of this model, is it fails to produce good results if any drastic change occurs in testing phase. Besides that, heavy fluctuations of a data set affect the forecasting performance produced by off-line model. This problem could be resolved by using an on-line neural network model. On-line model will update the network parameters for each data sample received as input series. Thus, on-line model is capable of learning the changes occur in CO concentrations level, which leads to better forecasting performance.

Currently, only CO monitor is available in the market, this device measures CO concentrations only. Meanwhile, this research focuses on the development of portable CO forecaster unit, which has the capability to provide multiple steps ahead forecasting, other than measuring CO concentrations level. The portable unit implements Hybrid Multilayered Perceptron (HMLP) network together with on-line modelling to perform the task. The portable unit generates warning based on the forecasted value. Thus, preventive action can be taken earlier, where human's life can be saved from CO poisoning. The approach is totally different compared to normal CO monitor, where the device generates warning based on current CO concentrations level.

1.3 Objective and Scope of Research

The main objective of this research is to develop a measuring, monitoring and forecasting system for CO concentrations. The system will be developed with the implementation of neural network using an on-line model. The specific objectives of this research are mentioned as below:

- to develop a measuring system for CO concentrations level
- to develop a monitoring system for CO concentrations level
- to develop an off-line PC based intelligent forecasting tool for forecasting CO concentrations level based on neural network
- to develop an on-line PC based system for measuring, monitoring and forecasting CO concentrations level
- to develop a portable unit to measure and forecast CO concentrations level

The purpose of the first objective is to develop a measuring system for CO concentrations level. The measurement system is built based on 8051 microcontroller with 8 bit analog to digital converter (ADC). Measurement system is developed to capture the signal in form of output voltage from CO sensor and display the CO level in parts per million (ppm) unit on the Liquid Crystal Display (LCD) attach to the developed 8051 board.

Second objective of the project is to develop a monitoring system for CO concentrations level. The monitoring system is built based on Borland C++ Builder 6.0 software. Interfacing between measuring and monitoring system are carried out by using serial communication. The CO concentrations level will then be displayed on

Personal Computer (PC), and several data processing methods are presented. Warning will be generated when CO level rise to dangerous level.

The purpose of the third objective is to develop an intelligent forecasting tool to forecast the future level of CO concentrations. Performance comparisons between several neural networks are carried out by using off-line modelling, in order to choose the best neural network. The forecasting performance of these networks are evaluated using index of coefficient (R^2), mean squared error (MSE), one step ahead prediction (OSA) and multi steps ahead prediction (MSA). The best neural network will be implemented in software, and the forecasting will be carried out using off-line model.

The fourth objective of this research is to develop an on-line prototype system for measuring, monitoring and forecasting of CO concentrations level. The CO forecasting for the best network will be carried out using an on-line model as mentioned in the previous section. The forecasting performances of this model are evaluated using index of coefficient (R^2), mean squared error (MSE), one step ahead prediction (OSA) and multi steps ahead prediction (MSA) test. The forecasting performances produced by on-line and off-line model are compared and discussed.

The final objective of this research is to develop a portable unit to measure and forecast the future level of CO concentrations. This portable unit is built based on digital signal controller (dsPIC) technology. The chosen neural network is implemented on dsPIC device to perform on-line forecasting and the forecasting level will be displayed on LCD. The developed portable unit will also perform MSA test. The unit also works as a data logger, because it can transfer the collected data to PC for data processing purpose. The portable unit will generate warning based on the forecasted value.

The limitation of this research is meteorological inputs such as temperature and wind speeds are not included as input series to the network in order to perform CO concentrations forecasting. In this research, only past CO concentrations values are used as input series to perform the task. In Malaysia, temperature has always been in stable range, which is totally different compared to western countries. Thus, temperature is not included as input series, since it might not give significant improvements towards the forecasting performance. Sharma and Khare (2001) have proved that the inclusion of temperature series as input does not provide any significant improvement in the model performance. Besides that, the developed portable CO forecaster is targeted to be used in partially closed area or closed area, respectively. The portable unit might not perform well in open area, since wind speed might effects the forecasting performance of the portable system.

1.4 Thesis Guidelines

This thesis contains 5 chapters consisting of 1 chapter on introduction, 1 chapter on literature review, 2 chapters on results and discussion and 1 chapter on conclusions, respectively. The first chapter discusses about the risk of human health due to CO poisoning and the requirement of forecasting tool. Besides that, it also briefly discusses the current approach of performing CO concentrations level forecasting. This chapter also presents the objective and scope of the research together with thesis guidelines.

The second chapter discusses on literature review of CO concentrations forecasting and neural networks. First, the chapter discusses about health effects due to CO poisoning and the conventional methods used to perform CO concentrations forecasting. This chapter also discusses neural networks, referring to the architecture and learning algorithms available in the field. Then, the chapter reports the implementation of neural networks in CO and other pollutants gas forecasting. It also presents the implementation of neural networks in microcontroller embedded systems.

The third chapter discusses about PC based CO Forecaster. Firstly, the chapter reports the development of PC based CO Forecaster using the software and the data sets used to perform forecasting. Secondly, theoretical discussion of on-line model, off-line model and model validation test are reported. Comparison studies between Hybrid Multilayered Perceptron (HMLP), Multilayered Perceptron (MLP), Radial Basis Function (RBF) and Recurrent networks are also presented. Finally, the best network is chosen to perform on-line and off-line forecasting. The results obtained from these models are compared at the end of the chapter.

The fourth chapter consists of overall block diagram of the hardware development for portable CO forecaster. The hardware developments are divided into 4 sections consisting of description on CO sensor, dsPIC development board, program memory and functions available in portable unit. The description of interfacing between portable CO forecaster and PC will be also described. Finally, the results obtained from portable CO forecaster are discussed.

The final chapter presents the conclusion made from this research. In this chapter, a few suggestions are presented if the research is to be continued in the future. Appendixes are also attached at the end of thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents description of Carbon Monoxide (CO) consisting of the health effects due to CO poisoning and conventional methods for CO forecasting. Theoretical descriptions on biological and artificial neural networks, learning algorithms and architecture of neural networks are also described. Finally, literature survey on the implementations of neural network with application to gases forecasting and microcontroller embedded system are also presented.

2.2 Carbon Monoxide

Carbon monoxide (CO) is a colourless, odourless, toxic gas that gives no clear warning when a person exposed to dangerous level. Motor vehicles, heaters, appliances that use carbon based fuels, tobacco smoke and household fires are the main sources of CO poisoning. CO is a primary pollutant in urban areas due to the major emission from motor vehicles. According to the Journal of American Medical Association (JAMA), CO intoxication is the leading cause of death due to poisoning in United States. It has been reported that CO from motor vehicles exhaust is the most common cause of poisoning deaths in the United States (Centers for Disease Control and Prevention, 1996). Many researches have been carried out to forecast CO concentrations level, in order to control the gas level within certain standard value, thus implementing air quality management and public warning strategies.

2.2.1 Health Effects of Carbon Monoxide Poisoning

Normally, blood contains 0.5 to 0.7 percent of CO gas in healthy people. The level of CO gas for people with diseases such as anaemia may range normally from 1.0 to 1.5 percent. If CO level exceeds certain limit, people will experience many health related problems because CO prevents the blood's ability to carry oxygen to body tissues including vital organs such as heart and brain. CO combines preferentially with hemoglobin to produce carboxyhemoglobin (COHb), displacing oxygen and reducing systemic arterial oxygen (O_2) content. CO binds reversibly to hemoglobin with an affinity 200 to 230 times than oxygen (Rodkey *et al.*, 1974). According to a study conducted using 128-409 young volunteers, the mean half-life of COHb is 320 minutes (Varon and Marik, 1997). Exposures for certain time range in the pollution environment can result in toxic concentration towards human blood. Table 2.1 shows the health effects produced by COHb levels toward human beings. Possible mechanisms of toxicity include:

- Decrease in the oxygen carrying capacity of blood.
- Alteration of the dissociation characteristics of oxyhemoglobin, further decreasing oxygen delivery to the tissues.
- Decrease in cellular respiration.
- Cause myocardial and skeletal muscle dysfunction.

Table 2.1: COHb levels and symptoms (Varon and Marik, 1997)

COHb Level	Symptoms
10%	Asymptomatic or may have headaches
20%	Dizziness, nausea, and syncope
30%	Visual disturbances
40%	Confusion and syncope
50%	Seizures and coma
60+%	Cardiopulmonary dysfunction and death

In urban area, non smoking city residents have been found to have COHb levels in the range of 1 to 2% (Stewart *et al.*, 1971). Tobacco smoke is a significant source of CO, containing approximately 4% CO. Normally, smokers have been observed to have COHb levels typically in the 4% to 5% range and as high as 9%.

The health effects of breathing in CO depend on the concentration of CO in the air, the duration of exposure, and the health status of the exposed person. Table 2.2 shows the health effects produced by CO level for certain exposure of time toward normal people whom don't experience any serious illness. For most people, the first signs of exposure to low concentrations of CO include mild headache and breathlessness with moderate exercise. People with heart disease are more likely to be affected by CO, even at low concentrations. Continuous exposure can lead to flu like symptoms including more severe headaches, dizziness, tiredness, and nausea that may progress to confusion, irritability, and impaired judgment, memory and coordination. CO is called as "silent killer" because if the early signs are ignored, a person may lose consciousness and be unable to escape the danger. People with greater risk of CO poisoning are shown below:

- individuals with respiratory conditions (such as asthma and emphysema);
- individuals with cardiovascular disease;
- individuals with anaemia (such as sickle cell anaemia);
- individuals engaging in strenuous physical activity; and
- the elderly people, pregnant women, children and fetuses.

Primary standard rates for CO level are designed by many countries to protect human health and human welfare. According to United States Environment Protection Agency (EPA), standard rate for CO is 9 ppm (parts per million) averaged over an eight hour period and 35 ppm averaged over a one hour period (City of Fort Collins, CO, 1996). The EPA allows no more than one exceedance of the standard per year for any given location. The Malaysian Environmental Department also have designed standard air quality rate for CO which is 9 ppm averaged over an eight hour period, and 30 ppm over a one hour period (Department of Environmen t Malaysia, 2002).

Table 2.2: CO levels and exposure time versus symptoms (About Inc, 2004)

CO Level (PPM)	Time	Symptoms
35	8 hours	Maximum exposure allowed by OSHA in the workplace over an eight hour period.
200	2-3 hours	Mild headache, fatigue, nausea and dizziness.
400	1-2 hours	Serious headache and other symptoms intensify. Life threatening after 3 hours.
800	45 minutes	Dizziness, nausea and convulsions. Unconscious within 2 hours. Death within 2-3 hours.
1600	20 minutes	Headache, dizziness and nausea. Death within 1 hour.
3200	5-10 minutes	Headache, dizziness and nausea. Death within 1 hour.
6400	1-2 minutes	Headache, dizziness and nausea. Death within 25-30 minutes.
12,800	1-3 minutes	Death

2.2.2 Conventional Methods of Carbon Monoxide Forecasting

In this section, a brief overview of conventional methods used for CO forecasting will be presented. Zickus & Kvietkus (1998) implemented the application of gaussian and regression model for short time air pollution forecast. In this study, CO was chosen as an air pollution indicator. Forecasting for specific hours was needed to evaluate air pollution peak. The statistical data analysis was based upon 3 months of hourly measurement of CO concentrations and meteorological parameters in the Vilnius City, Lithuania. For short time forecast (1 to 2 hour), the regression model gave better results compared to Gaussian model. For regression model, R^2 obtained was 0.6, due to the usage of short time relationship within time series that allows the model to adjust with the changing emission and meteorological condition. The R^2 value achieved by Gaussian model was only 0.2, because the model is static. Gaussian model treats system in equilibrium therefore unable to treat short time non stationary situation such as air pollution peak. In order to take preventive measures, 9 hours period of forecast were also presented. From the results obtained, it showed that the prediction accuracy of regression model was lower and comparable with Gaussian model.

Polydoras *et al.* (1998) presented a comparison study between dispersion model and univariate Box-Jenkins stochastic model for air quality predictions. Dispersion model is a 3 dimensional prognostic non-hydrostatic numerical algorithm for atmospheric wind flow and dispersion of industry or traffic pollutant over terrain having complex topography. These approaches were implemented to predict the hourly mean value of CO concentrations at 3 different locations in Athens area. The best performance for dispersion model in terms of MAE (Mean Absolute Error) criterion is 67.89%. For 1 hour and 24 hours ahead prediction from stochastic model, the minimum value for MAE was 9.67% and 19.73% respectively. Stochastic model was found to perform better compared to dispersion model, due to its availability to predict on a real time basis.

In another study, Maffeis (1999) described a method called FOREPOLL (Forecasting Pollution), to forecast the exceedance probability of CO concentrations fixed threshold (15 mg m^{-3}). FOREPOLL was developed based on deterministic, stochastic and Bayesian models. The model was thought to be adjustable for different air quality stations and various pollutants. Three years of data (1992-1994) consisting of CO measurement and meteorological parameters for 35 monitoring stations in Lombardy were analysed. The best results were obtained for metropolitan area with strength of 60% and sensitivity of 75% respectively. The model seemed to work better in case of stronger and more diffusion pollution, so it gave worst results for smaller cities. It can be concluded that FOREPOLL was strictly dependent on the philosophy where the model was based.

In another approach, Comrie & Diem (1999) examined the relationship between meteorology, traffic patterns and CO at seasonal, weekly and diurnal time scales in Phoenix, Arizona. A suite of models to forecast hourly and 8 hour ambient CO concentrations were developed using stepwise multivariate regression. The model was developed using CO concentrations from 5 winter seasons (1990-1995), nocturnal low level temperature inversion (ΔT), wind speed and wind direction. The results obtained showed that two stages modelling perform better, which refers to initial prediction produced by first model, followed by second prediction to improve upon the first. The best combined models had R^2 value approaching 0.9, with mean error under 1 ppm.

Sharma & Khare (2000) made an attempt to determine the degree of prediction using a limited data set only, restricted to the past record of CO time series. Univariate linear stochastic models (ULSMs) based on the Box-Jenkins modelling technique, more commonly known as ARIMA (Autoregressive Integrated Moving Average) were developed to provide short–term, real time forecast of extreme CO concentrations for an air quality control region comprising a major traffic intersection in Delhi City. Two sets of standardised data for CO time series were developed

i) Maximum daily1-h average series,X_{1mn}-series
ii) Maximum daily "working hours" (8am to 8 pm) daily1-h average series, X_{1wn}-series

The index of agreement (*d*) for X_{1mn}- and X_{1wn}-series models were 0.789 and 0.815 respectively. The index of agreement (*d*) is calculated as shown in Equation (2.1).

$$d = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}\left(\left|P_i - \overline{O}\right| + \left|O_i - \overline{O}\right|\right)^2} \qquad 0 \le d \le 1 \qquad (2.1)$$

where O_i and P_i are the observed and predicted values. Meanwhile, $\overline{O}$ being the average observed values. The *d* value ranges from 0 to 1, higher d values will be achieved if the observed and predicted values are closer. The disadvantage of this model was the availability of sufficiently long historical data set for model formulation. Besides that, ULSM's also being site specific, separate model should be developed for different monitoring stations. The methodology was effective for temporarily autocorrelated data only. The forecasting accuracy of ULSM's decreases with time, so rapid availability of data sets are desirable to achieve sufficient accurate forecast.

Sharma & Khare (2001) presented a continuation study from the previous research published in Sharma & Khare (2000). In this study, Box-Jenkins transfer function noise (TFN) model was used to perform the same task mentioned above. The time series of the surface wind speed and ambient temperature were used as the exogenous variables in the TFN models. The results obtained showed a significant improvement compared to univariate ARIMA model used in Sharma & Khare (2000). The forecasting performance was found to improve with the inclusion of wind speed as input series, but did not provide any significant improvement towards the inclusion of temperature series. The index of agreement (*d*) for X_{1mn}- and X_{1wn}-series models with the inclusion of wind speed was 0.850 and 0.864 respectively.

2.3 Biological and Artificial Neural Networks

Artificial Intelligence (AI) appears to be recent development in daily human beings life. AI is being implemented in many applications to forecasting, pattern recognition or data classification, signal processing, biomedical system and others. One of the most common AI applied in current researches is Artificial Neural Network (ANN), which is inspired by the way biological nervous systems such as brain, process information.

2.3.1 Biological Neural Networks

Human nervous system can be viewed as a three stages system, shown as block diagram in Figure 2.1 (Arbib, 1987). Neural net is represented by brain, which continually receives and perceives information, to make the appropriate decision. The receptors convert stimuli from human body or external environment into electrical signals that transfer information to brain. The effectors convert the electrical impulses generated by the brain into their respective responses (Haykin, 1994).

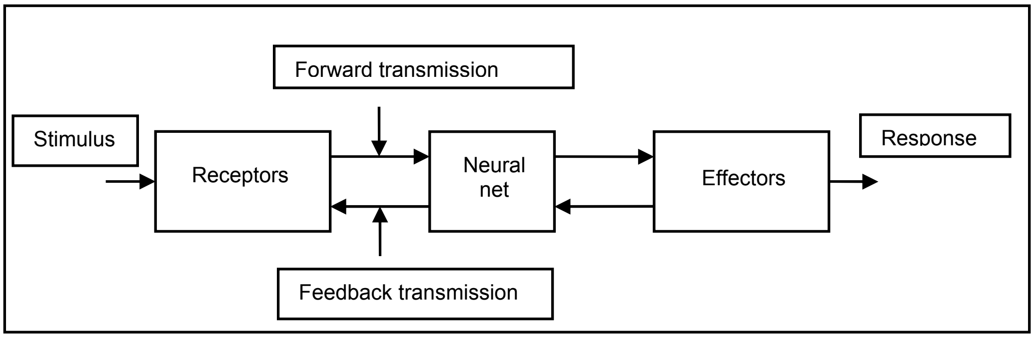

Figure 2.1: Block diagram of nervous system

Neurons are the fundamental element in human nervous systems found in human brain, the major components of a neuron are cell body, dendrites, axons and synapses as shown in Figure 2.2. Neurons process and collect information from others through a host of fine structures called dendrites, then transmit the information to other neurons or organ effectors. Dendrites collect information from other neurons through a long, thin stand known as an axon. Axons send out information from cells body from a neuron to other neurons.

Synapses are the elementary structural units that mediate interactions between neurons. Synapses convert the activity from axon into electrical effects that inhibit or excite activity from axon into electrical effects that inhibit or excite activity in the connected neurons. 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. While, small excitatory input will prevent the synapse to communicate between two neurons.

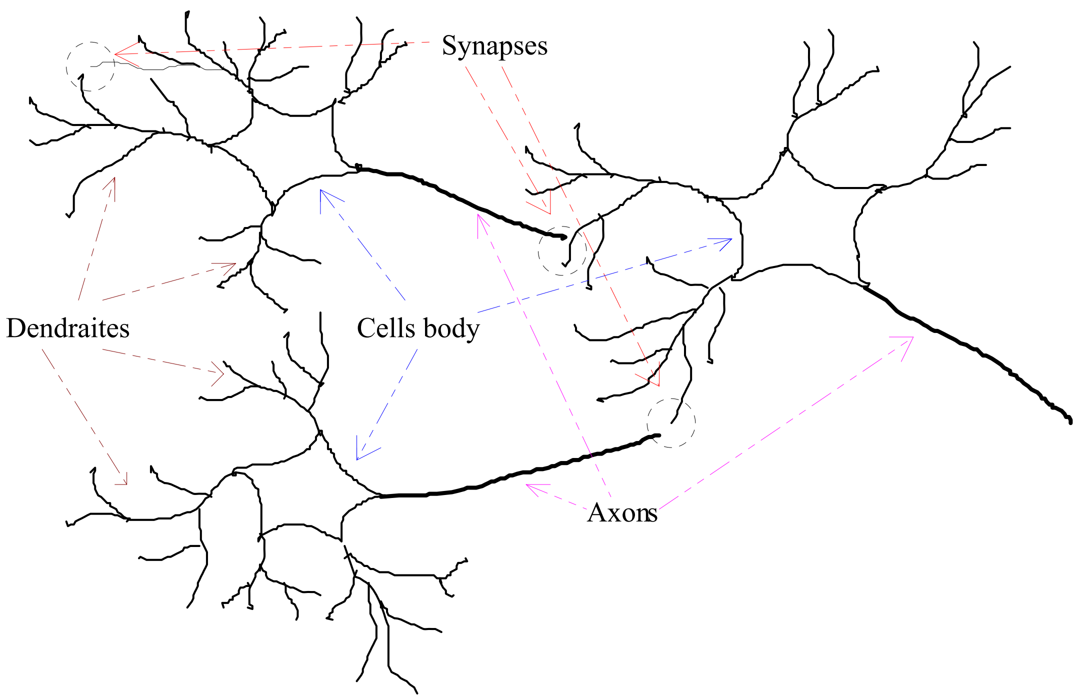

Figure 2.2: Three neurons and their respective components

2.3.2 Artificial Neural Networks

Artificial neural networks, commonly referred as "neural networks" have been motivated by human brain performing certain computations in an entirely different way from the conventional digital computer (Haykin, 1994). Neural networks are implemented by using a massive interconnection of simple computing cells called as "neurons" or "processing units" to perform its task. The artificial neurons that are used to build neural network are truly primitive compared to those found in brain, since the brain functionality are too complex to be implemented in neural network. Haykin (1994) describes neural network as a parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use in certain applications. It follows the brain functionality in two respects which is mentioned below:

1) Knowledge is absorbed by the network from its environment through learning process.
2) Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

Neuron is an information-processing unit, which is a fundamental element in order to develop a neural network. Figure 2.3 shows a block diagram of neuron, which consists of synaptic weights, adder and activation function.

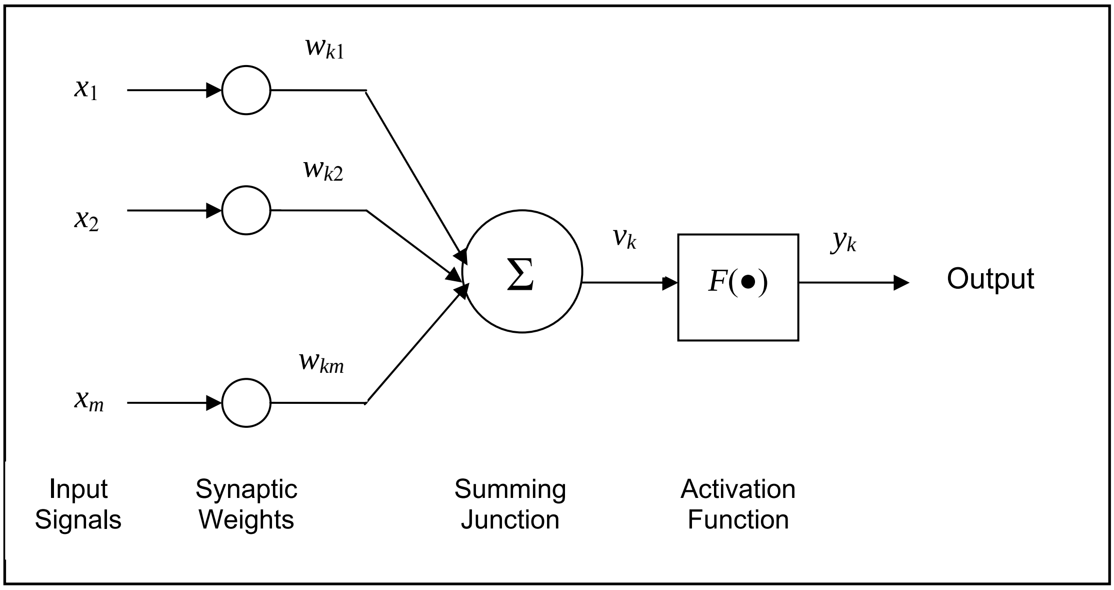

Figure 2.3: A simple model of neuron

Each synapse contains in a neuron will be represented by weight. Assume that *k*-th neuron has *m* synapses or input signals. Input signal x_j at the input of synapse *j* connected to neuron *k* is multiplied with synaptic weight w_{kj}. The synaptic weights of an artificial neuron can be in positive and negative values range. The weights will be adjusted according to their respective learning algorithm. The function of an adder is for summing the input signals, which is referring to the multiplication of signal x_j and their respective synaptic weight. The summing operation mentioned above constitutes a linear combiner. Finally, the activation function is used to limit the output of a neuron, allowing the amplitude range of the output signal within certain finite value. Typical choices of activation functions are hard limiter function, threshold function, sigmoidal function and sign function.

The mathematical equations to describe *k*-th neuron in a network are shown in Equation (2.2) and (2.3):

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \tag{2.2}$$

and

$$y_k = \varphi(u_k) \tag{2.3}$$

where x_j is the input signal; w_{kj} is the synaptic weight of neuron k; u_k is the linear combiner of the input signals; $\varphi(\bullet)$ is the activation function and y_k is the output signal of the neuron.

2.4 Architectures of Artificial Neural Networks

The architectures and learning algorithms used to train the network determine the capabilities of neural network achieving higher accuracy for a given application. The architectures of a network are identified by the manner in which the neurons are structured. The classification of neural network architectures can be divided into four components, which are single layer feedforward network, multilayer feedforward network, recurrent or feedback and hierarchical networks, respectively.

2.4.1 Single-Layer Feedforward Networks

Neurons in a layered neural network are organized in the form of layers. The simplest form of layered network consists of two layers which are input and output layers, respectively. Figure 2.4 is illustrated for the case of three nodes in both the input and output layers. The network is called as single–layer network, with the designation “single-layer” referring to output layer of computation nodes (neurons).

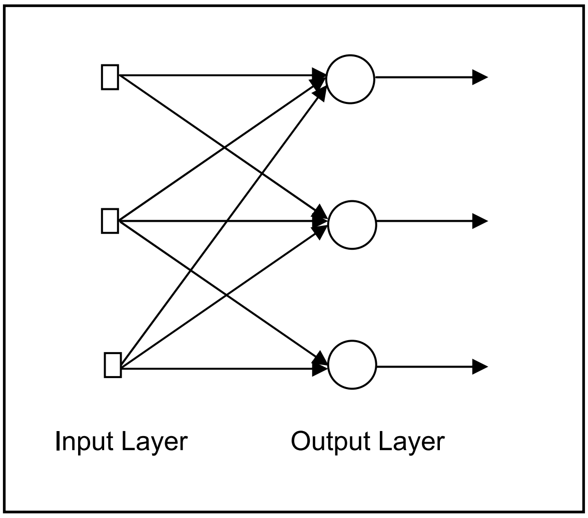

Figure 2.4: Feedforward network with single layer of neurons

2.4.2 Multilayer Feedforward Networks

The second class of feedforward neural network distinguishes itself by the presence of one or more hidden layers, to intervene between input and output layer, respectively. Some feedforward networks contain more than one hidden layer to enable them to extract higher order statistics (Churcland and Sejnowski, 1992). The ability of hidden neurons to extract higher order statistics is particularly valuable when the size of input layer is large. The input layer of the network supplies input signals for the hidden layer, and the output signals from hidden layer are used as input signals for the third layer and so on for the rest of network.

Feedforward network has two types of connection between the neurons, which are referring to fully connected and partially connected networks. The neural network is fully connected when every node in each layer of the network is connected to every other node in the adjacent forward layer. The network is partially connected if some communication links are missing. Figure 2.5 and Figure 2.6 show the architecture layout for fully connected and partially connected of multilayer feedforward neural network for the case of one hidden layer only.

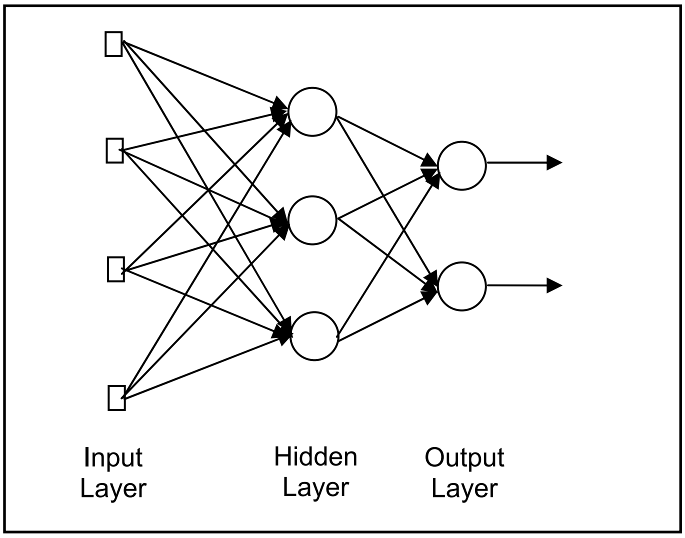

Figure 2.5: Fully connected multilayer feedforward network

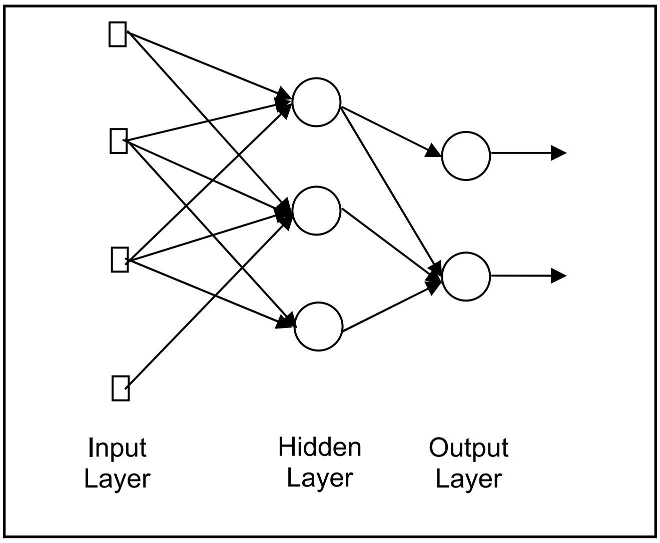

Figure 2.6: Partially connected multilayer feedforward network

2.4.3 Feedback/Recurrent network

A recurrent neural network distinguishes itself from a feedforward neural network and it has at least one feedback loop. The presence of feedback loops, result in a nonlinear dynamical behaviour, assuming that the network contains nonlinear units. Figure 2.7 shows a recurrent neural network, which contains 3 input neurons, 2 hidden neurons and 1 output neuron. The figure indicates that output signals from hidden and output neurons are feeding back to input neurons.

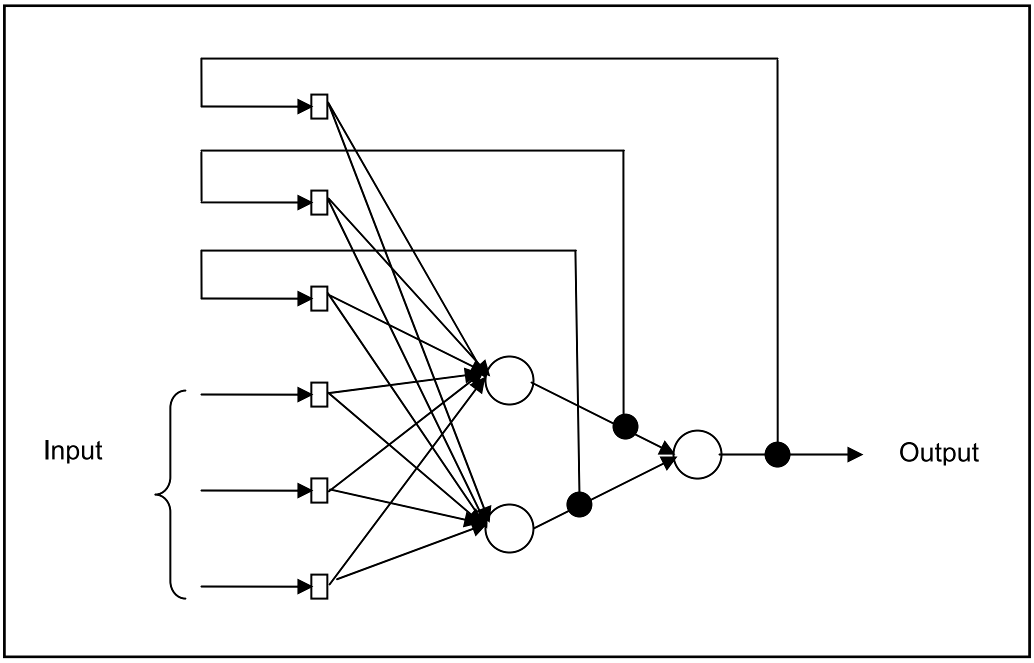

Figure 2.7: Recurrent network

2.4.4 Hierarchical Networks

Hierarchical network is built by using the same type of neural networks, architectures, and structures. The networks are cascaded together to built a hierarchically model. Usually, hierarchical networks are used to overcome classification problems, in which a network is required to perform two or more classifications for a given set of inputs. Mainly, these problems are found in application to pattern recognition. Each network in a hierarchical network has one output node only, the output node will classify their input into two classes only. Figure 2.8 shows a block diagram of a hierarchical network.

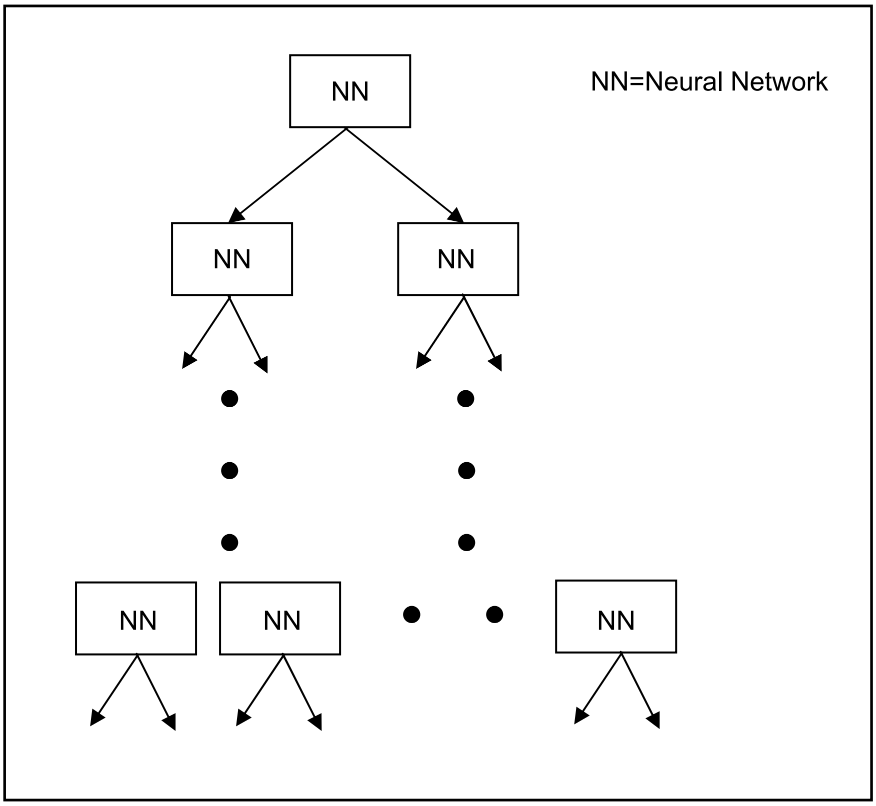

Figure 2.8: Block diagram of hierarchical network

2.5 Neural Network Learning Algorithms

Human brains learn from environment through experiences, while neural networks apply learning process such as learning algorithm to improve its performance. Learning algorithm is a prescribed set of rules which are applied to neural network for solving learning problems (Haykin, 1994). Neural network learns through an interactive process of adjustments applied to synaptic weights and bias levels. The network becomes more intelligent after each iteration of learning process.

Haykin (pp. 50, 1994) defines learning in the context of neural networks as: *Learning is a process by which the free parameters of a neural network are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning determined by the manner in which the parameter change take place.*

Five basic learning algorithms will be discussed below, consisting of error-correction learning, memory-based learning, Hebbian leaning, competitive learning, and Boltzman learning, respectively.

a) Error-correction learning

In this learning rule, network parameters are updated in order to minimize the error cost function of the output signal. An error signal denoted by $e_k(n)$ is calculated by using the following equation:

$$e_k(n) = d_k(n) - y_k(n) \tag{2.4}$$

Based on equation (2.4), $d_k(n)$ is the desired output and $y_k(n)$ is the real output of the given input signal. Some adjustments to synaptic weights are required in order to reduce the gap between the real and desired, output signal respectively.

The objective can be achieved by minimize the cost function or index of performance, $\xi(n)$, as shown below:

$$\xi(n) = \frac{1}{2} e_k^2(n) \tag{2.5}$$

where $\xi(n)$ is the instantaneous value of the error energy. The adjustments to the neuron *k* will be continued until the synaptic weights reach a stabil or steady state, in which the learning process are discontinued.

b) Memory-based learning

In memory-based learning, most of the past experiences are stored in large memories which are classified into two components, consisting of input and output of training data set. The desired response is restricted to be scalar (Haykin, 1994). When a testing data set is given, the network has to perform classification for a data set which is not seen before. At this moment, the algorithm responds to the network by retrieving and analyzing the training data in a “local neighborhood” towards the testing data samples. Local neighborhood is defined as training set which lies in the immediate distinctive characteristics of the given testing data set.

c) Hebbian Learning

Hebbian learning is the oldest and most famous learning rules compared to the others, it was introduced by Donald Hebb (Haykin, 1994). Donald Hebb quoted Hebbian learning in terms of neurobiological context in his book “The Organization of Behaviour” as mentioned below:

When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A’s efficiency as one of the cells firing B, is increased.

The statement mentioned above has been generated and modified in terms of neural network context into two rules (Stent, 1973; Changeux and Danchin, 1976):

1) If two neurons by the side of a synapse are activated at one time, the strength of the particular synapse will increase.
2) If two neurons on by the side of a synapse are activated unsynchronously, then the strength of the particular synapse will decrease.

d) Competitive learning

In competitive learning, only a single output neuron should be active at one time. Thus, the output neurons in a neural network have to compete among themselves to be active. Three basic elements in a competitive learning rule are shown below (Rumelhart and Zipser, 1985):

1) A set of neurons that are all the same except for some randomly distributed synaptic weights, and which therefore respond differently to a given set of input patterns.
2) Limit is applied on the strength of every neuron.