Water Treatment Plant Clarifier Control:

An Artificial Intelligence Approach

By

Setta Sasananan

B.Eng (Water Resources Engineering), Kasetsart University
B.Eng (Civil Engineering), Kasetsart University
M.Eng (Sanitary Engineering), Chulalongkorn University
M.Sc (Structural Engineering), University of Liverpool

Submitted in fulfilment of the requirement for
Degree of Doctor of Philosophy

School of Engineering, University of Tasmania
November 2009

Statement of Originality & Authority of Access

This thesis contain no material which has been accepted for a degree or diploma by the university of Tasmania or any other institution, except by way of background information and has been duly acknowledged in this thesis, and to the best of the author's knowledge and belief no material has previously been published or written by another person except where due acknowledgement is made in the text of this thesis.

Signed:..... Setta Sasananan .....

(Setta Sasananan)

Date: 15th November, 2009

This thesis is not to be made available or copying for two years following the date this statement was signed. Following that time the thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968*

Signed:..... Setta Sasananan .....

(Setta Sasananan)

Date: 15th November, 2009

ABSTRACT

The water treatment industry is currently facing growing pressure to provide efficient treatment at economical cost. The clarifier is one of the most important units in a water treatment plant. The impurity removal efficiency of the clarifier relies heavily on the skills of the operators and the fluctuation of raw water quality. The behaviour of the clarifier is complex and not able to be explained by simple mathematic equations.

The goal of this thesis was to develop and evaluate intelligent control methods for clarifier control. The intelligent control system proposed conceptually controls the clarifier by mimicking human operator control. The control objectives here were not only to minimise the clarified water impurity but also minimise the operational cost. The intelligent control system developed in this thesis was an integrated system employing an Artificial Neural Network (ANN) and Genetic Algorithms (GAs), which were linked by Model Predictive Control (MPC) architecture. The intelligent control system was composed of three components: (i) an ANN clarifier process model, (ii) an Multiobjective genetic algorithm (MOGA) optimiser and (iii) a decision maker. The system was designed to alleviate the operational problems of clarifier control and help the operators to choose suitable control actions. The clarified water qualities and operational cost were mathematically taken into consideration for control action optimisation via the MOGA optimiser. Two case studies of solid contact clarifier operation were investigated in this thesis, at Bryn Estyn water treatment plant (BEWTP), Hobart, Australia and Bang Khen water treatment plant (BKWTP), Bangkok, Thailand.

The ANN clarifier process models were developed using past operational data from both case studies. Past operational data were divided into training, validation and testing sets. To ensure the statistical equivalence among these three sets, clustering methods based on Self-Organising Map (SOM) networks were employed. The optimal temporal spans and architectures of the models were found by trial and error using the associated testing set. For the BEWTP case study, excellent predictive performance of clarified water colour (one time step ahead) is demonstrated with a coefficient of correlation (r^2) of 0.88 and Mean Absolute Error (MAE) of 0.50 HU. There is also excellent prediction of clarified water turbidity, evidenced by coefficient of correlation (r^2) of 0.89 and Mean Absolute Error (MAE) of 0.11 NTU. For the BKWTP case study, the ANN model showed good

performance prediction (one time step ahead) with a coefficient of correlation (r^2) of 0.71 and Mean Absolute Error (MAE) of 0.65 NTU. These ANN models were reliable since their prediction errors were of the same magnitude as measurement errors. However, prediction further than one step ahead was not recommended since this resulted in prediction errors that were larger than input measurement errors.

The intelligent control system was simulated using the testing set in order to assess its performance. Control actions were optimised to minimise clarified water impurities and operational cost by the MOGA optimiser and help the decision maker to choose the best set of control actions using the shortest normalised distance from the utopia point. However, defining of MOGA optimiser parameters was a prerequisite. Some parameters were chosen using guidelines from historical research, while the population size and number of generations were defined using a trial and error process. It was found that the optimal number of generations and population size was 200 and 80 respectively for the BEWTP case study and 400 generations and 100 populations for the BKWTP case study. According to the simulation results, all clarified water qualities complied with their operational targets and the operational cost was reduced by 8.3 percent and 3.4 percent for BEWTP and BKWTP respectively.

A full-scale pilot plant test was conducted at BKWTP in Bangkok. Unfortunately, it was not possible to conduct a pilot test at BEWTP due to health and safety regulations. The full-scale pilot plant test was for verifying the performance of the proposed control system in real life situations and in comparison with the performance of human operators. One BKWTP clarifier was chosen to be controlled by the intelligent system and the other was selected for normal control by the human operators. For a period of about one month from August to September 2007, the full-scale pilot plant was run continuously with both intelligent and human clarifiers running simultaneously. In terms of operational cost, the intelligent control system ran the clarifier with a 2.4 percent saving over that of the human controlled clarifier. If the intelligent control system was installed for all clarifiers at BKWTP and they ran with full production capacity of 3.5 million cubic metres per day, the operational cost saving would have been about $AUS 153,000 per *annum*. The reduction in operator salary costs was another saving which was not included. In terms of water impurity removal, clarified water with mean turbidity of 6.4±1.5 NTU discharged from the intelligent clarifier which was found to comply with the operational target of

seven NTU. However the human operator performed slightly better with mean clarified water turbidity of 6.3±1.5 NTU.

The novel aspects of this work are in the establishment and implementation of intelligent control systems which are a combination of ANN modelling and a MOGA optimiser. Its performance is also assessed in real practice. This is the first time that an intelligent approach fully mimics how human operators control the solid contact clarifier. All of the control actions (i.e. both chemical and physical control actions) are used by the intelligent control system to minimise the clarified water turbidity and operational cost. The contributions based on the result of this research work include the following items.

- ANN process model method and the unique use of the SOM are successfully developed based on the particular operational data of each of these two case studies.

- To achieve the best prediction of the clarifier process model, there is a need for using both present and past (temporal) data. Although this was expected, the study shows exactly how long a period of data is needed to optimise the control actions. The optimal model architecture of each process model is obtained based on the operational data, and their time lags are 8 hours for BEWTP and 12 hours for BKWTP.

- The prediction of the clarifier process model deteriorates when the prediction is taken outside the training domain. One time step ahead (4 hour) prediction is reliable when compared with the measurement error. However, multiple steps (long range) prediction is not in favour of the ANN process model because errors accumulate over each time step.

- When the intelligent control system works outside its training domain, it has been shown that its performance is less satisfactory than when working inside its training domain since the performance of the intelligent control system depended upon the predictive performance of the clarifier process model.

- In this thesis, the "ill posed problem" is avoided by the division of operational range according to raw water turbidity. Therefore, the set of control actions were optimised in the specifically defined ranges of raw water turbidity. This method has been shown to be effective in real life application.

In conclusion, with the successful results of the pilot plant test, both the operational targets of high water quality and cheaper operational cost are achieved using intelligent control. The intelligent control methods are proved able to work in real practice. This will overcome the limit of the manual skill of human operators and frees the clarifier operation from human error.

ACKNOWLEDGMENTS

The complementation of this thesis was accomplished because of kindness and support by several people. I would like to give the credit and share my indebtedness to everyone who was involved and who helped me to complete this thesis.

The first person whom I will always think of when I see this thesis is Dr. Tim Gale, My supervisor. I would like to sincerely express my thoughtful gratitude for his advise, support and inspiration throughout all my study. My appreciation also goes to my co-supervisor, Prof. Christ Letchford for his support.

All of my successes in my life are devoted to my parents and Montalee, my lovely wife. I would like to heartily express my profound gratitude and love for their immeasurable kindness, support, encouragement; love and care without any hesitation.

I would like to really give thanks to Mr. Russell Fox and Mr. Brett Pritchard from Bryn Estyn water treatment plant for providing operational guidance and for his assistance in data collection. On the other side of the world, at Bang khen water treatment plant, special credit must be mentioned to Mr. Wirachat Olapiriyakul and Mr Samai Anuwatkaseam, for their kind corporation that provided all facility and support for the full-scale pilot plant test and accommodation. Many thanks are also given to Mr. Surachet Panishkarn, Mr. Termsak Chotwanwirach, Mrs. Sopin Paethong, and all operators for providing data collection, the training of clarifier operation, practical suggestions, and warm hospitality during periods of full-scale pilot plant test. I also like to thank Dr Steve Carter for his inputs.

I would also like to express my deep gratitude and appreciation to Dr. Neil McGlasham and Jill Saunders, Iona Vialle and Diane Smith for their great support, good relationship, sincerity and happiness. Thanks for always being there during my trouble time and for making my life much more pleasurable.

Many thanks are given to all staff and friends at the School of Engineering and all library staff, for the good relationship, sincerity and happiness that I have had during my postgraduate life at UTAS.

PUBLICATION AND SUPPORTING WORKS

The thesis is supported by the following publications:

Sadsananan, S., Carter, S. and Gale, T. Prediction of Water Clarifier Performance: An Artificial Intelligence Approach. IPWEA Conference, Institute of public works Engineering Australia (IPWEA), Adelaide, Australia, August 2005. This paper presented some of the initial ideas of modelling the clarifier as presented in Chapter 5

Sadsananan, S., Gale, T. Artificial Neural Network Modelling of Water Clarifiers: Bang Khen Water treatment Plant Case study. NEC8-2009 Conference, Environmental Engineering Association of Thailand (EEAT), Nakhonratchasima, Thailand, March 2009. This paper drew on the work presented in Chapter 5 of this thesis.

Planed Publications (currently being drafted):

One of the most essential parts of this research is a full-scale pilot plant test. It has been dogged by concern about the fluctuation of raw turbidity in the Chao Phraya River. The test should be ideally taken during the early rain season, but it was only in mid-2007 that it finally became clear that the research methodology had been successful. Therefore, a full description of this full-scale pilot plant results and intelligent control system is planned to be published at a later date.

Supporting work:

This research spanned the five-plus year period from mid 2003 to late 2008. Although the focus is on the intelligent control and modelling of the complex system of the clarifier, the author had gained hands-on experience of clarifier control by undergoing operator training at BKWTP before the full-scale pilot plant was to be tested.

Poster Publication:

Sadsananan, S., Carter, S. and Gale, T. Improving water treatment plant clarifier control: An artificial intelligence. Poster presented in the School Research Evening, University of Tasmania, Hobart, Australia, 1 June 2005.

Presentation:

Sadsananan, S., Carter, S. and Gale, T. Improving WTP clarifier control: An artificial intelligence. Research presentation given to Hobart Water, Negara Crescent, Hobart, Australia, 17 October 2005.

CONTENTS

Abbreviations and Glossary

Abbreviations

ANN	Artificial Neural Network
BKWTP	Bang Khen Water Treatment Plant
BEWTP	Bryn Estyn Water Treatment Plant
CMD	Cubic Metres per Day
r	Coefficients of correlation
r^2	Coefficient of determination
GA	Genetic Algorithm
HU	Hazen Unit
IMC	Internal Model Control
MCMD	Million Cubic Meters per Day
MLD	Million Litres per Day
MOGA	Multiobjective Genetic Algorithm
MPC	Model Predictive Control
NTU	Nephelometric Turbidity Unit
PAC	Poly Aluminium Chloride
PFQI	Pareto Front Quality Index
SD	Standard Deviation
SCD	Streaming Current Detector
SCADA	Supervisory Control and Data Acquisition system
RPM	Round per Minute
WTP	Water Treatment Plant

Glossary

Upset number

The number of runs which clarified water qualities do not comply with the operational targets.

Upset percentage

The percentage of runs which clarified water qualities do not comply with the operational targets.

Hydraulic retention time

The time that the water is retained in the clarifier

1.0 INTRODUCTION

1.1 Background

The water treatment industry is currently facing growing pressure to provide efficient treatment in an increasingly stringent regulatory environment. The operation of a Water Treatment Plant (WTP) differs significant from other industrial plant operations in the aspect of raw material. The resource of the plant is raw water which is occasionally subject to unpredictable changes of quality. Adherence to a standard quality of effluent water must be strictly maintained otherwise community health will suffer.

Water clarifiers lie at the heart of modern water treatment. The clarifier is governed by complex non-linear relationships between numerous physical, chemical, and operational parameters. It is well known that about two thirds of the operational costs are spent on the clarifier and about eighty percent of impurities are expected to be trapped in this unit. An early incident which pointed to the importance of successful clarifier operation is the *Giardia* and *Cryptosporidium* pathogen crisis in the 1990s. These pathogens are widespread in treated water in a number of WTPs. Consequently the *multiple barrier strategy* was first launched by the United States Environmental Protection Agency (USEPA) which emphasised the importance of successful clarifier operation rather than, as previously, relying on filtration performance. This strategy gained momentum by improving clarifier operation, and a number of plant managers responded by raising the standard of clarified water quality (USEPA, 1999).

In order to improve clarification processes, plant operators need tools that will allow them to select which control actions are appropriate to achieve a desired quality of effluent. Historically, attempts have been made to obtain optimal control by fitting bench-scale data to mathematical formulae (e.g. a lookup table for coagulation dosage) or using the Jar test results (Kawamura, 1991; Kerri, 1996). Such attempts have generally been unable to account for simultaneous change in more than one or two of the key process parameters and often fail when applied to full-scale systems. As a result, although a number of online sensors are typically installed in clarifiers, current clarifier control is not automatic, but rather relies upon expert knowledge of plant operators (Mirsepassi et al., 1995).

This research has been motivated by the belief that there can be better control of clarifiers by doing more than simply relying on operator skill. The author's appreciation of the need for this research derives from work experience in water supply engineering and lecturing in this area. In the author's own experience, a number of clarifiers are at present operating far beyond their designed conditions because of a high demand for treated water. For example, in one of the case studies presented in this thesis, Bang Khen water treatment plant in Bangkok, the clarifiers are always operated at 50 percent above the design capacity.

Naturally the standard of the performance of manually controlled clarifiers relies entirely on the judgement of the human operators. Associated operational problems have historically hindered the development of clarifier control including

- Control actions based on operators' experience frequently worsens the clarified water quality and thus results in uneconomical operation.
- Traditional laboratory test results (i.e. Jar test) have been used to determine appropriate control actions such as coagulation dosages, however, this takes time and, results in improper control actions during the period of delay.

Several attempts have been made to improve clarifier control with new techniques, such as fitting sensor to set the chemical dosage. For example, the Streaming Current Detector (SCD) is used to measure streaming current potential which directly related to optimal chemical dosages (Liu & Wu, 1997). However, expert knowledge is required to set the rest of the control actions such as turbine speed. Improvement of clarifier control cannot be accomplished as long as the operator skill plays a significant role. This is especially so if the intention is to change manual clarifier control to intelligent control. With recent developments in desktop technology Artificial Intelligence (AI) techniques have been used and have shown promise in a large number of industrial control applications. There is considerable potential for AI techniques to overcome some of the drawbacks of traditional methods when applied to clarifier control.

1.2 Research Goal and Approach

The principle research goal of this study is to develop and evaluate the effectiveness of methods for an artificial intelligence based control and optimisation of clarifier

operations. To address clarifier operation problems, the author proposes that intelligent process modelling with Artificial Neural Networks (ANNs) may alleviate the problem of delay within the traditional laboratory tests. Also, an intelligent control system incorporating an ANN and genetic algorithm optimiser so called intelligent control system will be used to address suboptimal control actions due to human limitation. Rather than using only the "single" objective of improving clarified water quality, it is necessary to consider operational cost in the optimisation of the control actions. Therefore, the intelligent control system will be designed to incorporate multiple objective optimisations.

Two aspects to improving clarifier performance can be addressed based on the need to understand the clarification process and to provide better control actions. The first is modelling the clarifier based on pattern recognition using an Artificial Neural Network (ANN). ANN techniques should be able to describe the complex behaviour of a clarifier by utilising the ANN's learning and recognition process. The details of the ANN modelling method are given in Chapter 4. The ANN process models are particularly established for both case studies and their predictive performances are evaluated in Chapter 5.

The second is "an intelligent operator" to search for an optimised solution. It is well known that the Genetic Algorithm (GA) is one of the best tools for optimisation. The GA is a searching technique adapted from the Dravidian theory of survival. It can be utilised as an optimisation tool for a set of control actions for the problem at hand. Linking an ANN model and a GA based intelligent operator is proposed as an intelligent control system for a clarifier. Relevant details of GA with multiple objective optimisations and control simulation results are described in Chapter 6.

To prove the reliability and to assess the performance of the designed control system in a real life situation, a full scale pilot plant test will be conducted at Bang Khen water treatment plant (one of case studies) and compared with that of the human operator. The relevant details are presented in Chapter7. The approaches taken with this research are listed as follows.

- Review of background clarifier control systems, their operation problems and

performance evaluation.

- Collect data from two case studies in order to carry out statistical analysis and determine the need for improvement.
- Develop and validate clarifier process models in order to predict water quality using an ANN approach.
- Develop intelligent control systems utilising an integrated system of ANN based modelling and Genetic Algorithm (GA) optimisation in order to minimise the water impurity and operational cost using historical data to explore the performance of the intelligent control system compared to control by human operators.
- Use a full scale pilot plant test to assess one of the developed intelligent control systems and compare with manual operator control of a parallel system.

1.3 Research Partners

Research partner with the University of Tasmania in this research were Hobart Water in Tasmania, and Bangkok Water Metropolitan in Thailand. Two case studies were used to demonstrate the performance of the intelligent control system. One is in Tasmania, Australia, at the Bryn Estyn Water Treatment Plant (BEWTP) and the other one is in Bangkok, Thailand, at Bang Khen Water Treatment Plant (BKWTP). Both plants are extremely different in terms of sizing, raw water quality and operational problems. For this reason, they provide excellent cases for evaluating the performance of control systems under different conditions.

1.4 Thesis Organization

The structure of this thesis is as follows

Chapter1: Introduction

Chapter2: Water Treatment Plant: Background and Performance Evaluation:
This chapter provides background understanding of the clarifier process and control, including clarifier performance evaluation and identification of operation problems.

Chapter3: Water Treatment Plant Data Collection, Manipulation and Analysis:
This chapter provides understanding of the characteristics of the operational data of BEWTP and BKWTP, including statistical

analysis.

Chapter4: Artificial Neural Network Modelling: Review of Methodology and Application to Current Research:

This chapter presents a background to artificial neural networks and clarifier modelling development using artificial neural networks.

Chapter5: Clarifier Artificial Neural Network Model Development and Optimisation:

This chapter presents model development and model performance validation using numerical simulation and validates this with real data for both case studies.

Chapter6: Clarifier Intelligent Control: Genetic Algorithm Approach:

This chapter presents the development of an integrated control system combining ANN and GAs for optimising control actions, and system simulation using real data for both case studies.

Chapter7: Clarifier Intelligent Control Performance: Full-scale Pilot Plant Test:

This chapter examines the use of intelligent control systems by performing a pilot plant test at BKWTP with the aim of optimising clarified water quality and operational cost.

Chapter 8: Conclusion and Recommendations

2.0 WATER TREATMENT PLANTS: BACKGROUND AND PERFORMANCE EVALUATION

2.1 Introduction

The overall performance of any Water treatment plant (WTP) is mainly dependent on the efficiency of each major unit such as its clarifier (USEPA, 1998). The efficiency of the clarifier dominantly affects the overall WTP as well as the sequential unit process performance since a number of critical chemical and physical actions take place there to achieve coagulation, flocculation, and sedimentation (Kawamura, 1991). These complex interactions between physical and chemical reactions in a clarifier are difficult to understand and control. Although there are several on-line sensors currently available in the water supply industry, an automated control system for a clarifier is not yet fully employed and manual operations based on operator experience are still utilised in most WTPs.

This chapter starts with a brief background of water treatment processes. Clarifier control and operational problem are then discussed to guide the design needed for a new control system. Both case studies (BEWTP and BKWTP) are later introduced, followed by evaluations of their clarifier performances in order to examine a need for control action improvement.

2.2 General Background of Water Treatment Process

Water treatment processes for domestic consumption involves raw water treatment at several stages to remove impurities such as bacteria and suspended solids before feeding into the distribution network. Figure 2.1 illustrates a typical WTP unit process which includes a pre-treatment and pH adjustment unit, clarification unit, filtration unit and disinfection unit (AWWA, 1991).

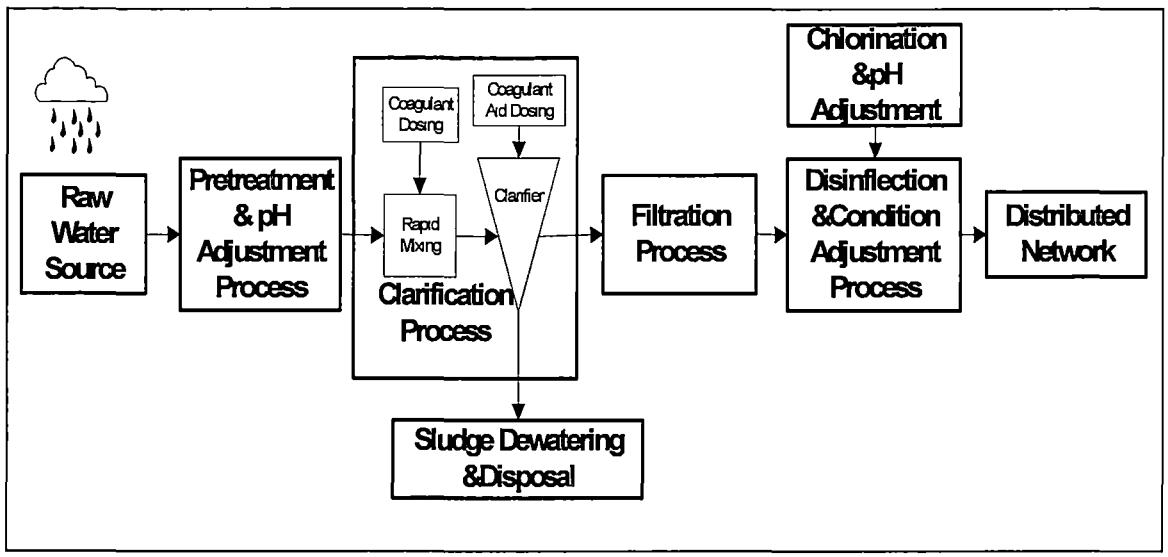


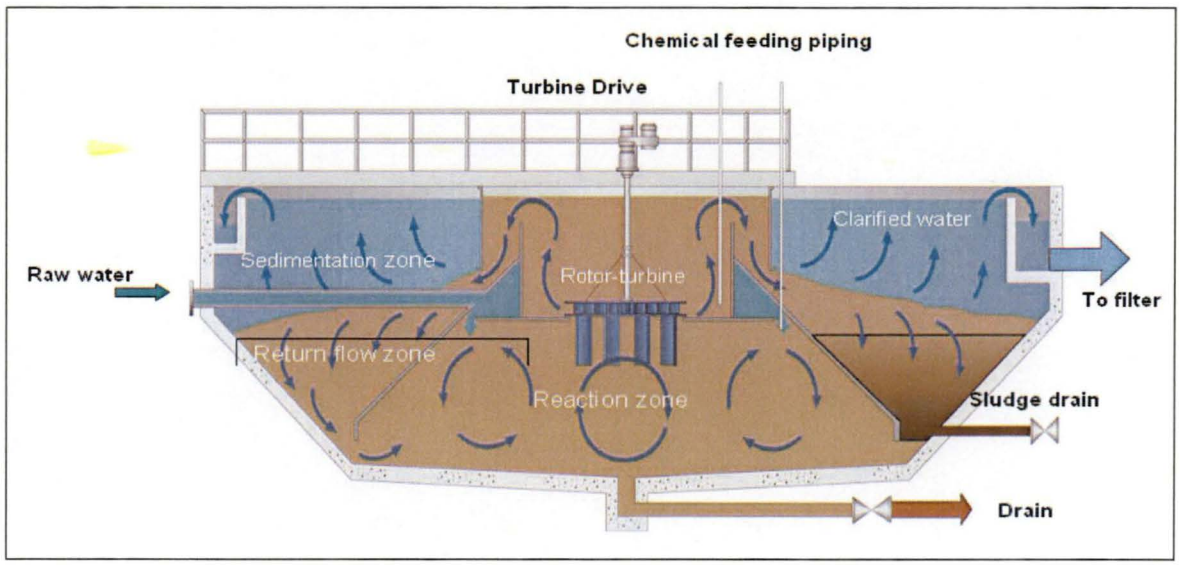
Figure 2.1 Water Treatment Process

From the raw water source, the raw water is passed through a pre-treatment unit, where the acidity or alkalinity for pH adjustment is introduced. In a clarifier, coagulant and coagulant aid are typically introduced in a rapid mixing stage. The coagulant serves to promote the coagulation and destabilization of the colloidal particles in the raw water, and the coagulant aid encourages the flocculation and growth of these destabilized particles. Alum is one of the most used coagulants. The dosages required depend on the source water characteristics, such as turbidity, pH, temperature, and conductivity. The large particles or flocs formed after flocculation are separated in the clarifier via sedimentation. Subsequently, most of the remaining solids that escape the clarifier are entrapped in the filtration process. Finally, to meet with hygienic standards, disinfection substances are applied (e.g. chlorine). The general detail of the clarification process is outlined below. More details of the other unit processes can be found in a number of textbooks (ASCE. & AWWA., 1990).

2.2.1 Clarification Process

A clarifier contains a three-stage process consisting of coagulation, flocculation and sedimentation. In the coagulation process, coagulation agent such as Alum is initially fed and rapidly distributed by fast agitation in order to destabilise suspended solid. This coagulation is normally completed in a line mixer ahead of the actual clarifier. The degree of destabilisation is further increased by aggregation of newly coagulated solids as well as reused particles in the flocculation process. Flocculation is promoted by gentle agitation resulting in larger size particles named flocs. Specifically, due to reuse of particle, this process can be called a solid contact process (Degrémont S. A., 1991). In

some situations, dosing of coagulant aid (e.g. polyelectrolyte) enhances the flocculation process. This is followed by the sedimentation process, which is a solid-liquid gravitational separation. Most of the larger flocs are trapped in a sedimentation process. A typical cross section of an *Accelator sludge recirculation* clarifier as equipped at BEWTP is shown in Figure 2.2



Source: www.degremont-technologies.com

Figure 2.2 Typical cross section of an Accelator sludge recirculation clarifier

After raw water is mixed with a coagulant (e.g. Alum), raw water flows through the gentle mixing in the reaction zone to promote floc amalgamation and then flows to the sedimentation zone. The clear clarified water flows upward to filter while flocs settle to bottom of the clarifier. Within a certain period of time, larger and heavier flocs settle to the return flow zone. The clarified water flows to the filter unit. Some of the flocs are reused and returns to reaction zone. In order to keep mass balance in a clarifier, excess sludge is routinely drained out by a drain system at the bottom of the clarifier.

Generally, water is retained in the clarifier for 1.5-4.0 hours. The solid-contact clarifier is categorised into two types, the low production rate of sludge blanket units and the high production rate of sludge recirculation units. For the first one, sludge blanket is formed in the sedimentation zone and help turbidity entrapping. For the second one, the sludge is separated form the clarified water in sedimentation zone, and it then returns to a reaction zone equipped with mechanical agitation to promote flocculation. Degremont (1991) provides a good review of solid-contact clarifier. For both of the case studies in this thesis, the clarifiers are the sludge recirculation type.

2.2.2 Clarifier Operation Concept

The prime objective of the clarifier unit is to promote settlability of flocs and entrap the turbidity. To achieve this, larger floc size is preferable (Tambo & Watanabe, 1979). Successful floc amalgamation involves two issues: the chemical, and the physical. For the chemical issues, correct chemical agents and amounts need to be used for linkage of small flocs to enhance their size to become larger ones. The alum is used as the main coagulation agent and polyelectrolyte is applied merely to enhance it effect. Using polyelectrolyte effectively reduces the alum dosage (Shanks, 1978). For the physical issues, smooth agitation from the turbine increases floc amalgamation and proper sludge drainage rate aims for mass balance control. In addition, a decrease in sludge draining and reduction of the plant flow rate result in more sludge remaining in the clarifier. This too benefits floc amalgamation. In particular, with a sludge recirculation clarifier, mixing energy by turbine speed helps to recycle sludge in the reaction zone. In practice, the operators use the Jar test results to guide both alum and polyelectrolyte dosages. On the other hand, turbine speed and sludge drainage can only be judged by the operator's experience.

Conceptually, if raw water qualities are within treatable ranges and all the control actions can be suitably provided, some degree of interaction between these control actions can be expected. If increases in the chemical dosages are within the optimal range, this increase will promote floc amalgamation and lead to improvement in turbidity removal. In practice alum is generally used as the main chemical agent to achieve turbidity precipitation. Polyelectrolyte only enhances floc amalgamation. With no polyelectrolyte and using alum only, it is possible to precipitate the turbidity. However this is always uneconomic especially in high turbidity (Kawamura, 1991).

Increasing turbine speed also generally promotes floc amalgamation but only when used within in a suitable range. Aggressive agitation will cause floc to break up and result in turbidity leakage. The sludge drainage rate aims to control the mass balance in the clarifier and also controls the reuse rate of any sludge. Since the clarifiers at BEWTP and BKWTP are of sludge recirculation type, a longer retention of sludge will benefit floc amalgamation. However, if the sludge is retained beyond an optimal time, for instance if a sludge drainage rate was assigned, then too much floc will be kept in the clarifier and again lead to the turbidity leakage. An increase of plant flow rate does not promote the

performance of turbidity removal since it lessens the time for floc amalgamation. Figure 2.3 graphically illustrates the interactions among these control actions (Δ and ∇ signs mean "increasing' and "decreasing"). Again, these relations are based on the assumptions that raw water qualities fall in the treatable range and that all the control actions vary in appropriate ranges.

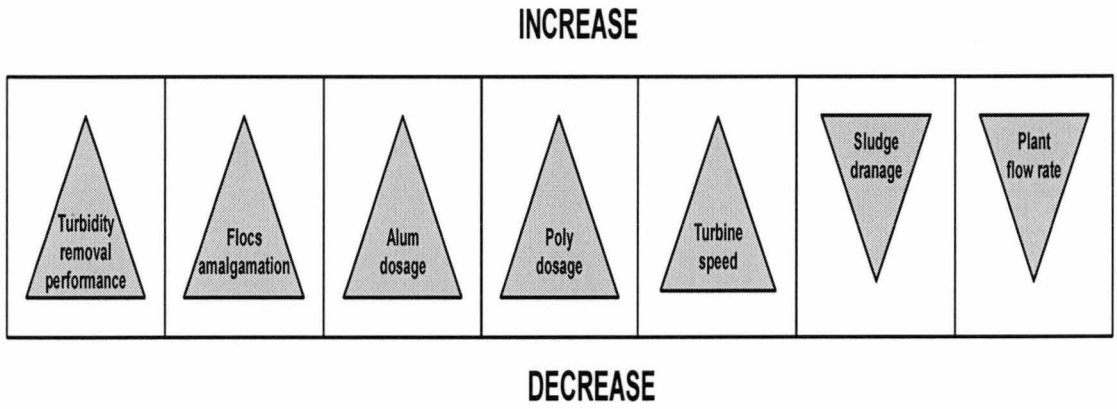


Figure 2.3 The interaction of clarifier control actions: (Δ and ∇ signs mean "increasing' and "decreasing")

2.3 Clarifier Control and Operational Problems

One of the main concerns of this section is to review conventional clarifier control. It is important to understand typical clarifier controls and identify their disadvantages, to recognise the operational problems and to see how the operators deal with the problems conventionally. In both case studies, they adopted almost the same procedure in common which was originally based on the guideline from (ASCE. & AWWA., 1990).

2.3.1 Current Clarifier Control

Generally, clarifier control involves five control actions. These are feeding coagulant and coagulant aid dosages, changing turbine speed, plant flow rate control, and sludge drain control. Feeding chemicals is used for flocs forming, and optimal turbine speed helps to enlarge the floc size and provides flow circulation in the clarifier tank. Sludge drain control is set to avoid sludge carry over. Besides meeting the clarified water standard, the difficulty encountered here is economically to provide suitable control actions in a timely manner, in correspondence with each to change of raw water quality (Kerri, 1996). Detailed description of each control action is well documented in a number of water treatment plant design textbooks and quoted in Table 2.1(ASCE. & AWWA., 1990; Degrémont S. A., 1991).

Control Actions	Description
Feeding coagulant (e.g. Alum)	• Regularly used for floc agglomeration • Costly operating costs
Feeding coagulant aid (e.g. polyelectrolyte)	• Flocs enlargement and settleability enhancement • More costly operating costs
Plant Flow Control	• Relating to hydraulic retention time • Under constrain of water demand • Cheapest control action
Mixing Intensity Control (Turbine speed)	• Used for floc agglomeration and circulation • Moderate Operating Costs
Sludge drainage	• Keep mass balance in clarifier • Cheapest control action

Table 2.1 Clarifier control actions

All of these control actions were optimised by the plant operator's experience using the data from online sensors and laboratory results (e.g. Jar test, described below). Typically, these online data are associated with raw water quality such as turbidity, alkalinity, pH, temperature and conductivity. They are known to be significant to clarifier performance (ASCE. & AWWA., 1990; Degrémont S. A., 1991). Their short descriptions are listed in Table 2.2. Normally, sampling frequency of online data is set to 15 to 30 minutes. On the other hand, laboratory test results (e.g. Jar test, colour test) were conducted every four hours. Therefore, every four hours, both online and laboratory data were ready for the operators to execute the optimal set of control actions.

In most water treatment plants, the Jar test is set as a standard laboratory test in order to give a guideline for coagulant and coagulant aid dosages, which are the first line control actions (Kawamura, 1991). The Jar test is done by taking samples of raw water and adding a different proportion of coagulant to each sample. After rapid and slow mixing then a certain period of settling time, each sample is assessed for water quality. The dosage that produces the optimal result is found and used as the recommended coagulant and coagulant aid dosage. The details of the Jar test process are well documented which can be found in (Clesceri et al., 1989). However, the Jar test is a time-consuming test, and

it cannot reveal the results without time delay. In Figure 2.4, the operator was using a Jar test.

Parameter	**Importance of Parameter** (Degrémont S. A., 1991)	**Application in WTP Control** (Pollack et al., 1999)
Turbidity	Measuring total suspended solids which is an indicator of degree of impurity of water and also the potential of microbial contamination	• Overall performance of unit process monitoring in terms of solid removal • Utilised for coagulant approximation and early warning of abnormal operational condition
Conductivity	Measuring of water's ability to conduct electrical current which is an indicator of dissolved solids	• Utilised for coagulant dosage approximation and early warning of abnormal operational condition
pH	Measuring of hydrogen ions, which is important for process optimization, corrosion control and aesthetic objective	• Utilised for coagulant dosage approximation and early warning of abnormal operation condition
Water temperature	Directly relates to chemical reaction, flocculation time viscosity of water and settleability of particles.	• Utilised for coagulant dosage approximation • Measuring of degree of density current and early warning of abnormal operational condition

Table 2.2 Parameters relating to clarifier control

Figure 2.4 Jar test demonstration (BKWTP)

In general, during the normal operational period, primary tests of raw water quality and clarified water quality changed in a narrow range and the operator had time to slowly optimise control actions. The operators were able to justify a set of control actions for the next four hours on the test results presently at hand and online data. The coagulant dosage would be fed based on the guideline of the Jar Test. A plant constant flow rate was preferable since it was easier to maintain suitable *hydraulic retention time* (i.e the period of detention of water in clarifier) whist the demand for the produced water was under control. Mixing intensity and sludge drain rate could be maintained within the optimum range by setting a proper turbine speed and correct sludge drain time. The sludge drain time and turbine speed was set based on the operator's experience. However, in some plants, there are lookup tables available that can be used as general guidelines.

Occasionally, during times of irregular conditions, primary tests may fluctuate highly resulting in poor quality of clarified water. Operational stability may be disturbed due to a number of factors such as abrupt changes in raw water quality, plant flow rate and environment. To recover the stability, a request for a new Jar test is recommended (City Water Technology, 2001; MWA, 2000). Increasing coagulant dosage is normally used as a first line control action. However, if the clarified water remains poor, increasing coagulant aid dosage is considered. Implementing other control actions such as reducing plant flow rate, reducing turbine speed are recommended. Nevertheless, these are based on operators' experience. A schematic diagram of clarifier control with flow of information is shown in Figure 2.5.

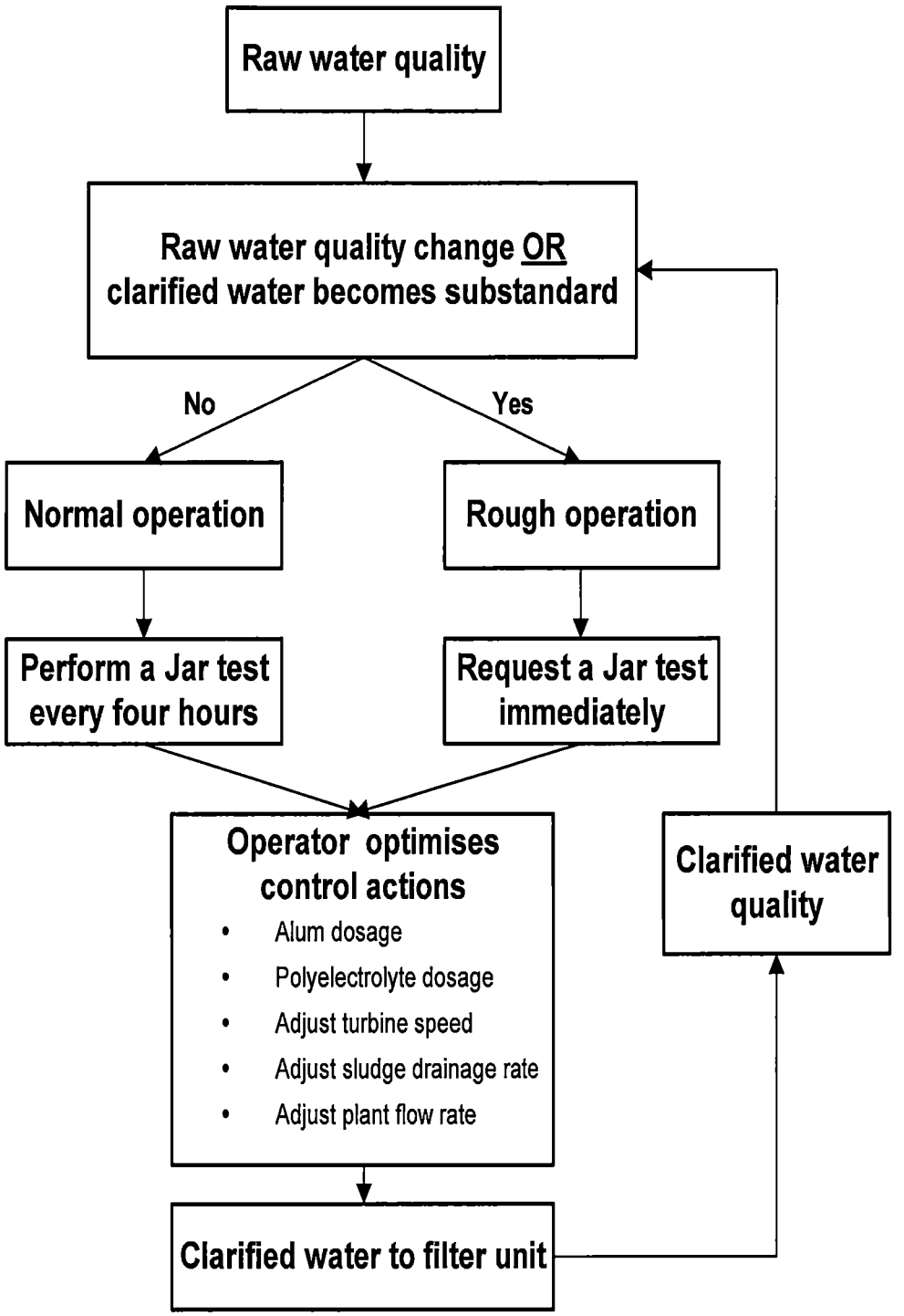


Figure 2.5 Schematic diagram of clarifier conventional control

In general, upset conditions often occur because of abrupt changes in raw water colour and turbidity. Such conditions are predictable and most seasonally occur. On top of that, due to the large size of clarifier, thermal gradient induced density currents in the clarifier tank are a frequent cause of poor clarified water.

2.3.2 Operational Problems

The most serious limitation of clarifier operation is found in instability during an abrupt change of flow rate, influent turbidity and temperature(Degrémont S. A., 1991; Kerri, 1996). To deal with such limitations, a number of approaches are recommended such as adding coagulant aid and adjusting turbine speed. These operational problems are mainly due to an inability to provide proper control actions (i.e. coagulation dosage, plat flow, etc) in a timely manner since control actions are guided by delay in receiving Jar test results. Additionally, all control actions depend upon using the operator's experience. However, reliance on standard operating procedures may occasionally provide consequential problems due to human error, especially during an upset condition. Specifically, clarifier control problems can be outlined as follows.

- The Jar test results are not immediately available while the data from online sensors are readily at hand. Therefore, operators have to wait for the results of a Jar test to know what coagulant dosages are justified. This limits the overall performance of the control system in spite of the results of other faster on-line sensors (e.g. turbidimeter, pH meter, conductivity meter) being known (James M. Montgomery Consulting Engineers Inc., 1985). In order to cope with a change of raw water quality, one attempt has been made which is to use a lookup table. It can roughly estimate the coagulant dosages needed in some WTPs. Nevertheless, only one variable typically is used (e.g. raw water turbidity), resulting in inaccurate dosages. This results in less than ideal control, especially during upset conditions.

- Apart from coagulant dosage controls, which are guided by Jar test results, the other control actions are determined purely by the operator's experience. Occasionally, human error may result in suboptimal control during upset conditions.

- The operators can use only the data presently at hand (e.g. Jar test results and online data) to select a set of control actions. These may be insufficient. More data covering earlier condition (i.e. time lag data) should give better results. This is because data with a larger time lag covering sludge age in the clarifier would provide information of sludge conditions. Sludge concentration is known to be a very important parameter for clarifier performance.

- Generally, to obtain good clarified water quality is set as a high priority objective. It would be better if costs consideration were also considered as a constraint in optimising control actions. For example, feeding proper coagulant dosage results in good clarified water, it is costly. Alternative such as suitable combination of turbine speed and sludge drainage rate can effectively reduce the chemical dosage during a period of control low water turbidity.

How these operational problems affect clarified water quality cannot be given in quantified terms. However, according to interviews with plant operators, they agreed that any seasonal changes of raw water quality were likely to be predictable. Early warning was possible and most diurnal variations were not of severe degree. They might often harm the operation during the first one or two days after a shock load approaches, and operational stability could afterwards be recovered. Although this occurred only for a short period, each error resulted in poor clarified water quality. On the other hand another major cause of operational instability was environmental changes, such as density current and wind induced surface waves. Although each incidence caused clarified water quality to be slightly below standard, they happened a few times a day. This was particularly found in the case of the large clarifier.

2.4 Artificial Intelligence in Water Supply Engineering

The acceptance of artificial intelligence (AI) application in the water supply industry is a relatively new construct compared with other industries, such as chemical processing. Although several applications have been developed in the field in last several years, it is not yet a mature research area especially in the area of clarifier control (Maier et al., 2004; Riyaz et al., 2004).

2.4.1 Past Research work

There have been a number of AI approaches developed and applied in process control of water treatment. Most of them include the expert systems and ANNs.

Works related to Expert system and Fuzzy logic

Expert system and Fuzzy log was the first AI approach introduced to the water supply engineering research area. Most of the work was related to chemical dosage suggestion.

With the same approach, expert rule (or fuzzy rule) was established by using operational data. Nonetheless, it became less popular after the ANN's appearance in this area.
Zhu and Simpson (1991) applied an expert system to the conventional Anstey Hill water treatment plant, in Adelaide to advise plant operators on alum and activated silica dosing. Additionally, the expert system was used for fault diagnosis and it was used to train new operators. It showed potential to improve operation. However, Anstey Hill water treatment plant is a conventional plant operated without the complication of a clarifier unit. The rapid mixing unit (coagulation process), slow mixing (flocculation process) and sedimentation are separately located.

Liu and Wu (1997) applied a fuzzy logic controller incorporating a Streaming Current Detector (SCD) for automatic control of coagulation. This system was tested with a bench-scale water treatment plant with synthetic raw water and the results showed that this combination functioned satisfactorily for coagulation. However, the main problems with using the expert system and fuzzy logic are that general rules may be too simplistic to describe the complex processes of the clarifier which require understanding of process. The difficulty is most obvious when fuzzy logic deals with a large dimension model (e.g. large number of inputs and outputs)(Zhang, G. et al., 2001).

Work related to Artificial Neural Network

Many researchers focus on ANNs, since they have the potential to cope with nonlinear problems. Most of the associated studies focus on process modelling or chemical dosage modelling by using an ANN. The ANN applications in water supply industries fall into three categories: (i) prediction, (ii) process control and (iii) others related to water supply. The following paragraphs discuss these studies based on these categories of the ANN application.

- **Predictive model**

Most studies used ANN models mainly for prediction applications, which can be categorised into two types depending on the model architect used. The first one is called a forward model or process model. It replicates how the process works. The input and the output of the model is equivalent to the real process. For example, in a clarifier process model, model inputs are raw water qualities and outputs are predictive clarified water turbidity. The second category is known as a process inverse model involving mapping the inverse of the process, and the outputs are predicted control actions (e.g. alum dosages) that will be required to meet the target value of the process.

Mirsepassi et. al. (1995) used five years of daily average operational data of Wyong WTP in New South Wales, Australia to develop an ANN inverse process model to predict the alum and polymer dosages. The inputs were a time series of daily average raw water quality variables and did not include the clarified water quality. Guiding by sensitivity analysis, the number of model input variables was reduced. They obtained a very good result, with coefficient of determination (r^2) of 0.94.

Without using real operational data, Maier et al (2004) developed model processes and inverse processes by using a set of 202 data from Jar test results. This water samples were collected from several water sources in southern Australia. Process models were used to predict effluent water qualities (e.g. turbidity, colour, pH). The inverse model was used to predict Alum dose. The performance of both process and inverse process showed excellent promise with coefficients of determination (r^2)of 0.90 and 0.98, respectively.

- **Process control**

In process control applications, a few AI control systems have been developed and applied to clarification processes. Most of the current studies are related only to chemical dosage control.

Zhang and Stanley (1999) established ANN process and inverse process models using the daily average operational data from a conventional cross flow clarifier at Rossdale water treatment plant. They applied internal model control (IMC) architecture for which the clarified turbidity process model served as reference model and linked with an inverse process model to predict the alum dose. The process model was not only trained with good cases but it was also trained with the poor cases (if the clarified water turbidity complied with operational target, it was considered "good case" otherwise it was poor case). This was to ensure the process model was trained by variety pattern of operations and it could predict the clarified water turbidity for overall operating ranges. The inverse model was trained by the good cases only since it was expected to predict the best possible alum dose. The concept of an IMC system for alum dosage suggestion as proposed by Zhang and Stanley is shown in Figure 2.6.

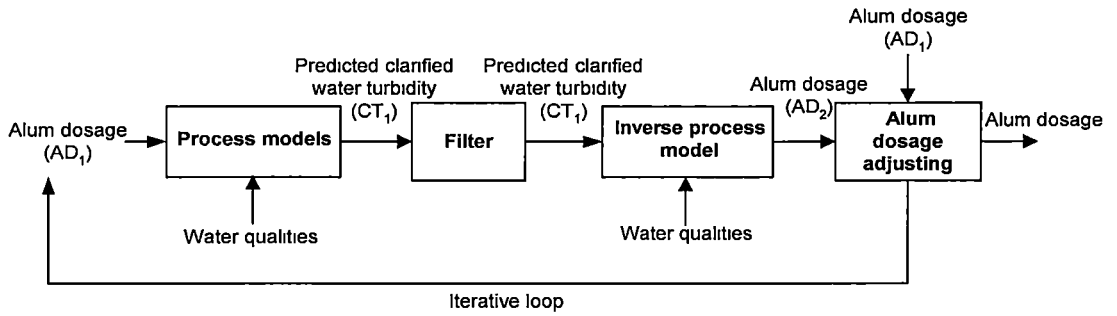


Figure 2.6. An IMC system for alum dosage suggestion, presenting the process model and inverse process model linked in IMC architecture and flow of information

At the beginning, raw water quality variables were iteratively input to the process model together with the generated values of alum dose (AD_1) until the predicted value of clarified water turbidity (CT_1) complied with the operational target. Next, the filter would allow CT_1 to be input as the inputs of inverse process model. Actually, the input and output variables for the inverse model were of the same group as that of the process model, except that the predicted value of clarified water turbidity became an input variable and alum dosage ,(AD_2) became the sole output of the inverse model. Theoretically, if the inverse model was a perfect invert of the process model, AD_1 and AD_2 would be identical and the usage of AD_1 would be justified. However, if it was *not* the case (since process model and inverse process model were deliberately built from different data), then AD_1 and AD_2 would not be identical, and then the alum dosage would be adjusted in the vicinity of AD_1 using an iterative process to minimise these differences. In this work, the results from the simulation showed r^2 values of 0.86 and 0.96 for the process reference model and process inverse model respectively.

Baxter et al. (2001) developed a softening process control system, using the same approach as Zhang and Stanley (1999) with their internal model control strategy. Daily average operational data of the conventional cross flow clarifier at Rossdale WTP was used to develop the models. The raw water qualities were input to the process model to predict the total hardness of clarified water, which became the input of the inverse process model, and lime dose became the sole output. The specific goal of control system was to estimate the total hardness in the softening clarified water and the softening lime dose required. The target was to maintain the clarified water hardness of 135 mg/L (as

$CaCO_3$). The results showed that the r^2 values of 0.84 and 0.95 for the process reference model and plant inverse model, respectively.

- **Others related to water supply**

With regard to raw water quality forecasting, ANN models have been developed to predict *Cryptosporidium* concentration (Brion et al., 2001), source water salinity (Dandy G. C. et al, 1991; DeSilets et al., 1992) and raw water colour (Zhang, Q & Stanley, 1997). The goal of these forecasts was to provide the operators with an early warning system for raw water quality changes during water quality fluctuations. An ANN model has also been developed for the prediction of residual chlorine in the distribution system in Quebec City (Rodriguez MJ et al., 1997).

2.4.2 Strength and Weakness of Previous Approaches

Among past research works, the study of Zhang and Stanley (1999) used the best approach (section 2.4.1). They developed a very interesting approach to control the cross flow clarifier. They used two ANN models joined together under IMC architecture. The first part was a clarifier process model (for prediction of water quality), and the second part was an inverse process model (for suggesting a dosage). The inverse model was set up by selectively using only good operational data. It used the predicted clarified water quality (as given by the process model) to suggest a dosage. Despite its ingenious approach, the inverse model only mimics how a good operator works with regard to selecting chemical dosage. The Zhang method might not provide an optimised solution although it might equal the only solution from a good operator. That is, it might only be close to the optimal point. In addition, only chemical dosages are output, as those are the main control actions for a conventional cross flow clarifier. This would not be suitable for a solid contact clarifier such as is the case for the studies in this thesis. The values of optimal turbine speed and sludge drainage rate must also be included.

Most of these models were developed using daily average data to predict the clarified water quality. This is reasonable for use with an early warning system. However for predictive models or process control it is questionable if these process models need to predict (with different frequencies) either a faster or slower sampling frequency than that for which they were developed. In real practice human operators usually select the set of control actions at least every four hours. This period is approximately equal to the time

that the water spends in the WTP (hydraulic retention time) and ensures that any faults can be detected and timely corrective actions taken (Kerri, 1996). If these models are to work in real practice, they need to be accurate enough to predict at this sampling frequency (Riyaz et al., 2004)

In addition in all these past research works, most focused on the clarified water qualities. The operational cost was not included in optimising the control actions. This is major drawback of their approach used, since operational cost is a key factor in WTP. The disadvantages of the previous approaches can be outlined as follows

- Most of these ANN process models utilised low-resolution daily average operational data which might not be accurate enough to use in different time intervals.
- Only the works of Zhang and Stanley (1999) and Baxter et al. (2001) were concerned with clarifier control. In addition to their process model they also used an inverse model to predict a single set of control actions with no means of verifying whether these control actions are optimal. It was just one of good set of control actions chosen from the "good case".
- None of these previous approaches found the optimal control action for achieving, the operational targets by including operational cost. Actually, in the previous work, the control actions were optimised the control actions by considering only clarified water qualities not including the operational cost.

All these limitations are used as a platform to develop the proposed clarifier intelligent control system. A clarifier intelligent control system is proposed that overcomes these limitations and its detail is shown in the next section.

2.5 Clarifier Intelligent Control System

Although in both case studies, the mean clarified water quality generally conformed to their targets, operational stability could not always be maintained due to operational problems as discussed in the previous section. In order to achieve operational stability, control actions need to be provided to cope with any changes in a prompt and timely manner. With regard to the clarifier operational problems, an intelligent control system is proposed to enable the control with automatic mode. This control system is a control system using computers to control the process, replacing human operators. It provides the

means by which to avoid errors resulting from human sensory perception and mental ability. It also brings the control actions to cope with environmental changes.

In general, a control system with automatic mode consists of three major integrated components: a process controller for supplying control logic, on-line sensors for collecting information, and a Supervisor Control and Data Acquisition (SCADA) system for communicating, executing control actions and evaluating process performance (Degrémont S. A., 1991; Pollack et al., 1999). The schematic diagram of a basic process control system is shown is Figure 2.7.

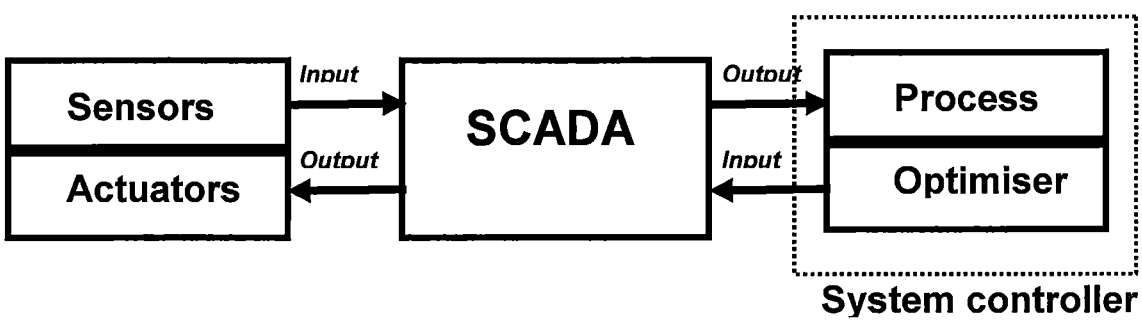


Figure 2.7 Basic structure of automated control system

Automated control begins with data collection from the installed sensors on a continuous basis or from manual input. Sensed signals are relayed to the SCADA system, which is responsible for all communication between sensors, actuators and the process model, process performance evaluation, and execution of control logic. After process performance evaluation, the data is sent to a system controller which consists of process model and optimiser. After an optimising algorithm, the controller provides control logic as controller outputs. A SCADA system transfers the control logic to control actions sent to actuators. SCADA also hosts the interfacing between the control system and plant operators, and has the ability to store operational data of the actions it has initiated.

Conceptually, the clarifier automation controller consists of two key components, which are designed to cope with the operational problems recognised above. First, the clarifier process model can be used automatically to predict clarifier water quality. This provides an analogy of how a Jar test works, but it can be used regardless or minimise of time delay. The second component is that rather than using the operators' experience, an intelligent optimiser is utilised automatically to optimise the set of control actions. It

should be able to optimise more than one constraint such as clarified water quality and operational cost. The concept of clarified automated control system is presented in Table 2.3.

Clarifier operational problems	Potential solutions to problems
Time delay of laboratory (e.g. Jar test, colour test)	• Clarifier process model
Suboptimal control actions	• Automatic control optimiser

Table 2.3 Clarifier operation problems and solutions

2.6 Artificial Intelligence Approach for Clarifier Automated Control

In recent years, there have been improvements in computer hardware and excellent artificial intelligence software packages are now available. This has resulted in these methods being widely applied in process modelling. Due to the complex behaviour of the clarifier, clarifier modelling and control optimisation are the challenging tasks. **It is proposed that artificial intelligence methods such as ANNs and GAs have potential to move clarifier operation towards automated control by using them as process modelling tool and intelligent optimisers, respectively. The focus of this thesis is to implement this application and evaluate these methods.**

Artificial Neural Network

An ANN is a type of artificial intelligence technique that attempts to imitate the way a human brain works. An ANN is a data information processing paradigm that is inspired by biological nervous systems such as brain processes (Negnevitsky, 2005). An ANN, with its remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and recognise trends that are too complex to be noticed either by humans or by other computer techniques. One key significant advantage of using ANN modelling is that ANN can map the set of inputs and outputs without any previous knowledge of process relationships concerned (Masters, 1993).

By utilising its advantages, ANN provides a potential process-modelling tool for the complex relations of clarifiers. By analogy to the jar test, the virtual ANN version of the

clarifier should be able to work automatically in a real-time mode in line with online sensors. This should well match the need of the clarifier ANN modelling and performance evaluation.

Genetic Algorithm

Genetic Algorithm (GA) is another artificial intelligence approach, which is adapted from evolutionary processes in nature. GA is an exploratory search and optimisation procedure. It is based on a Darwinian-type survival of the fittest strategy with sexual reproduction, where stronger individuals in the population have a higher chance of creating offspring. GA generally includes the three fundamental genetic operations of selection, crossover and mutation. These operations are used to modify the chosen solution and select the most appropriate offspring to pass on to succeeding generations (Negnevitsky, 2005). GA provides a rapid convergence to a near optimum solution in many types of problems. Unlike other optimisation techniques, GA does not required mathematical descriptions of the optimisation problem but instead it relies on a objective function in order to assess the fitness of a particular solution to the problem in question (Jain & Martin., 1999).

The proposed approach to improve clarifier control in this thesis is to use a GA-based intelligent optimiser to optimise the set of control actions. The set of control actions should be able to optimise with respect to multiple objectives such as water quality and operating cost. In particular, the intelligent optimiser requires the GA, which is able to handle multi-objective optimisation. Temporal history is taken into account for instead of using only current information as is more conventional. The details of using GA for the intelligent optimiser with an intelligent control are in Chapter 6.

2.7 Case Studies

Two selected case studies are utilised in this study. They are extremely different in terms of production capacity, environmental conditions, and raw water quality. The first case study is Bryn Estyn Water Treatment Plant (BEWTP) located near New Norfolk, Tasmania. The second case study is Bang Khen Water Treatment Plant (BKWTP) in Bangkok, Thailand. The details of both case studies are as follows.

2.7.1 Bryn Estyn Water Treatment Plant Case Study

BEWTP was first operated in 1960. It is presently under Hobart Water and located 45 km from Hobart city. The full capacity during the summer is about 160,000 cubic metres per day (CMD), and can drop to 20,000 CMD in winter according to community needs. The water produced serves the Hobart metropolitan area and several districts near New Norfolk.

Raw water is drawn from the Derwent River and it is directly fed to the treatment process without any buffering reservoir. This way results in quality fluctuation. However, the raw water is of good quality and is typically low in turbidity. Nonetheless, its colour is characterised as moderate *true colour*, which does not contribute to water turbidity. The water is also quite cold for much of the year, as compared to other Australian plants. The raw water pH is usually within a reasonable range for treatment. Typical alkalinity and calcium hardness levels in the raw water are relatively low. Low alkalinity, high colour waters can sometimes be difficult to treat (City Water Technology, 2001). The associated statistical analyses are presented in Chapter 3.

BEWTP utilises two kinds of treatment systems. The first one is a direct filtration process with a capacity of 60,000 CMD and the second is a clarification-filtration process with a capacity of 100,000 CMD. During a period of good raw water quality, BEWTP operates in the direct filtration mode by which coagulant and polymer are mixed with raw water and directly fed to the filter units. However, most of the time BEWTP employs only a clarification-filtration process.

In the clarification process, there are two five-metre depth INFILCO ACCELATOR clarifiers with diameters of 26 and 33 metres, with capacities of 40,000 CMD and 60,000 CMD respectively. Both of them are linked to the filter units. The larger clarifier is used for this case study since it has been continuously operated through the years to cater for basic community demand. If the water demand increases, then the smaller clarifier and direct filtration unit are added to the system. Both clarifiers are fully operated under a Supervisor Control and Data Acquisition, SCADA system which is used only for data display and logging. The figure of the larger clarifier is shown in Figure 2.8.

Source:(City Water Technology, 2001)

Figure 2.8 INFILCO ACCELATOR, Bryn Estyn water treatment plant

2.7.2 Bang Khen Water Treatment Plant Case Study

BKWTP is located 10 km north of Bangkok Metropolis, Laksi district. It started operating in 1979. The maximum capacity is 3.2 million cubic metres per day (MCMD), and it normally serves the community demand during summer periods and drops to a minimum of about 30 percent of capacity during the rainy season. It is responsible for 75 percent of Bangkok's water demand. Raw water is taken from the Chao Phraya River in Pathum Thani 18 km north of BEWTP. Before feeding to the plant, there is a buffering reservoir located at the beginning of the treatment process which regulates raw water quality. This raw water is very high in suspended solids resulting in high turbidity and colour. However, the colour, which is called *apparent colour*, as it is mainly caused by suspended solids, can normally be treated at the same time as turbidity. Typically, turbidity and colour peaks are found during high flows in the rainy season. Alkalinity and pH are in the normal range. Occasionally, peak turbidity occurs from river sand mining about 30 km from the intake (MWA, 2000). In addition, due to the hot climate, temperature changes during the day are high thus the vertical thermal gradient along the depth of the tank conducts a current as called a density current. This results in turbidity overflow.

BKWTP has been continually operating in clarification-filtration mode. BKWTP contains eighteen flat conical tanks of clarifiers, and solid contact units of slurry

recirculation type, which are 58 metres in diameter and 6 metres deep as shown in Figure 2.9. Each tank's clarification rate is about 200,000 CMD. All of these units are linked to 44 units of filters. These clarifiers are operated manually even though a number of sensors are installed. According to the BKWTP development plan, the SCADA system will be fully employed by the year 2008. However, in the first state, SCADA will be used for data display and logging only.

Figure 2.9 Solid contact clarifier, Bang Khen water treatment plant

2.8 Clarifier Performance Evaluation

In order to examine the need for operational improvement, clarifier performance of both case studies are outlined. Generally, the treatment performance of each clarifier can be evaluated by turbidity removal efficiency although other measurement methods can potentially be utilised as benchmark parameters (e.g. water colour and particle counting). This is because the degree of impurity directly relates to the level of turbidity and turbidimeters are extensively employed in a large number of WTPs to monitor turbidity at various locations throughout the process (Pizzi, 2005; USEPA, 1998).

However, every WTP specifically sets its own operational target according to their raw water quality and environment. For Bryn Estyn Water Treatment plant (BEWTP), the

performance criteria are set so that the clarified turbidity and colour should be less than 1 NTU and 10 HU, respectively. The clarified water colour is used due to moderate *true colour* raw water thus it is difficult to treat concurrently with turbidity (City Water Technology, 2001).

For BKWTP, a target of clarified water turbidity of 7 NTU is set, which is higher than that of BEWTP's since its intake of raw water is of extremely poor quality. A clarified water colour target is not used because *apparent colour* of raw water can be precipitated with the turbidity coincidently (MWA, 2000). The site-specific operational criteria of both case studies are tabulated in Table 2.4.

Item	Performance Criteria
Bryn Esytrn Water treatment plant	• The Clarified water turbidity is less than 1 NTU. • The Clarified water colour is less than 10 HU (City Water Technology, 2001).
Bang Khen Water treatment plant	• The Clarified water turbidity is less than 7 NTU (MWA, 2000).

Table 2.4 Operational target of clarifiers used in case studies in this thesis

In order to examine the clarifiers' performance, trend and statistical analyses were used. The trend analysis was primarily concerned with considering the characteristics of the raw and clarified water. It was not an exact statistical method, but provided an indication of the stability of the clarifier operation. The maximum daily turbidity values were used for trend analysis as guided by USEPA. The maximum values were used since the goal was to assess the integrity of the clarifier operation when it was most vulnerable (USEPA, 1998, 1999). To achieve operational stability, a clarifier must demonstrate that it could take a raw water source of various quality and consistently produced high quality clarified water. On the other hand, the statistical analysis could be employed to determine the mean values of clarified water quality and the percentage of upset conditions when clarified water s less than a certain turbidity and colour.

A minimum of 12 months of maximum daily values of water turbidity and colour is recommended to assess clarifier performance since it provides a good indicator for long-term performance and covers seasonal effects (USEPA 1998). However, for BEWTP, clarifier performance was evaluated from 10 months of data from August 2002 to May

2003, as a longer period of data collection was unavailable. For the case of BKWTP, 18 months of data from February 2003 to October 2004 were used. The discussions of both case studies are given separately as follows.

2.8.1 Bryn Estyn Water Treatment Plant Performance Evaluation

For Bryn Estyn Water Treatment Plant (BEWTP), the treatment performance of the clarifier unit was evaluated by utilising a set of data that describes raw water, and clarified water quality was evaluated in terms of colour and turbidity removal using trend and statistical analysis. The associated statistic values are shown in Table 2.5 and described as follows:

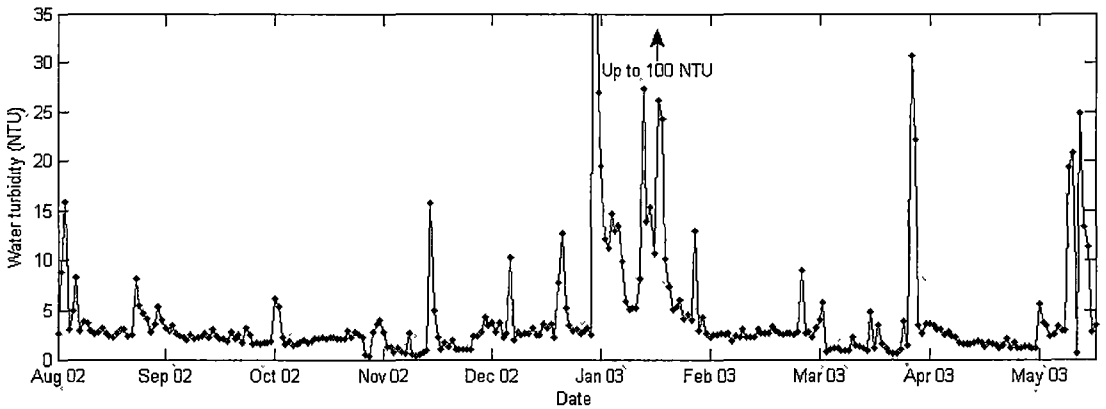
Performance parameters	Value
Raw water turbidity (Mean±1SD)	2.92±3.6 NTU
Clarified water turbidity (Mean±1SD)	0.69±0.50 NTU
Raw water & clarified water turbidity correlation (r)	0.48
Raw water colour (Mean±1SD)	44.41±28.53 HU
Clarified water colour (Mean±1SD)	6.69±2.54 HU
Raw water & clarified water colour correlation (r)	0.51

Table 2.5 BEWTP performance analysis

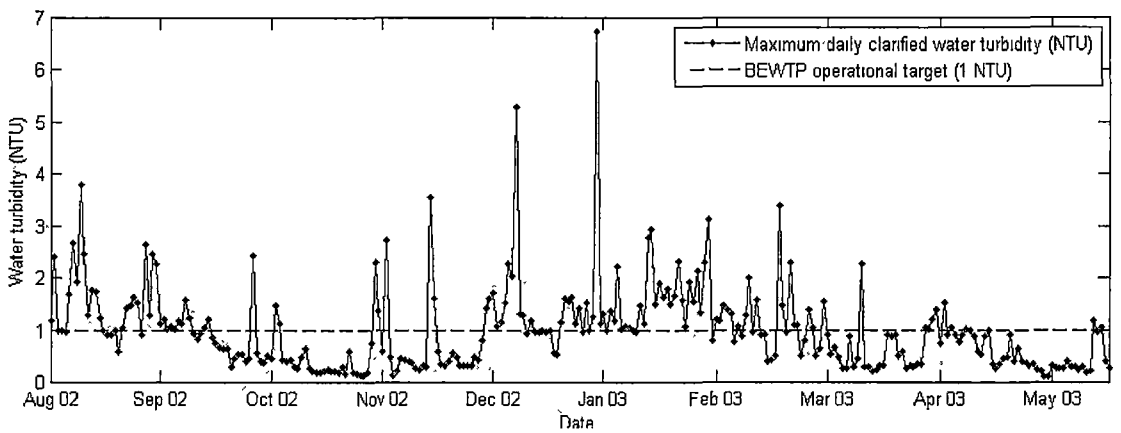
Turbidity Removal Analysis

According to the BEWTP operation guideline, 1 NTU of clarified water turbidity and 10 HU of clarified water colour are set as an acceptable threshold (MWA, 2000).The turbidity trend analysis was conducted by considering any changes in maximum daily value of raw and clarified water turbidity as shown in Figure 2.10a and Figure 2.10b. During the period of August 02 to November 02, a few spikes in the clarified water plot were found when raw water turbidity varied to a mild degree. The operators did fairly well in maintaining operational stability. However, in the next period of December 2002 to March 2003, the operational stability was interfered with when the raw water turbidity showed several peaks and variability to a severe degree. The variability was also evident

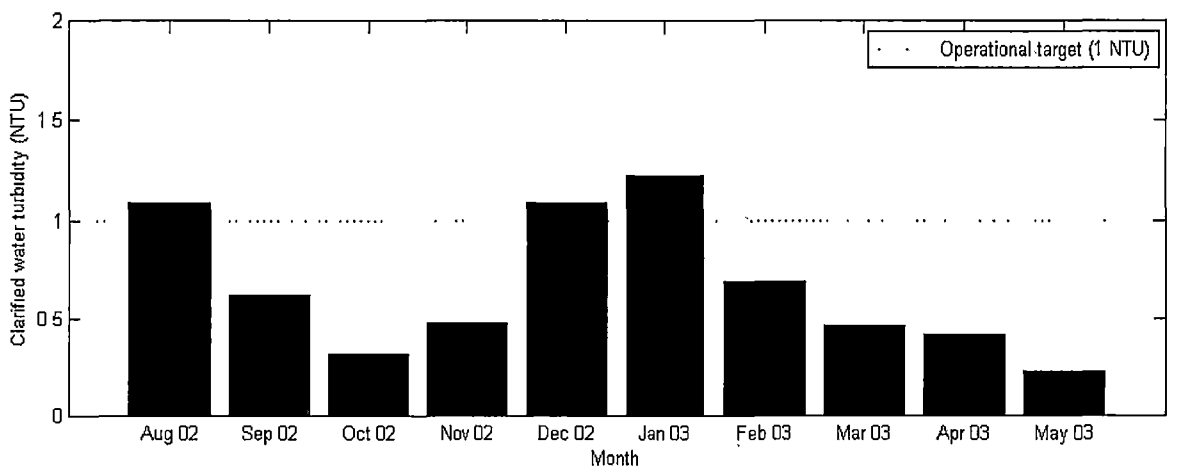
in the clarified water turbidity. These pass-through variations and spikes indicate that the performance of the clarifier was below optimum. After March 2003, the operational stability recovered when the clarified water turbidity became consistent, complying with an operational target of 1 NTU in spite of significant variations in raw water turbidity.



(a) Maximum daily raw water turbidity



(b) Maximum daily clarified water turbidity



(c) Monthly clarified water turbidity

Figure 2.10 Clarifier Performance evaluations (Turbidity approach, BEWTP)

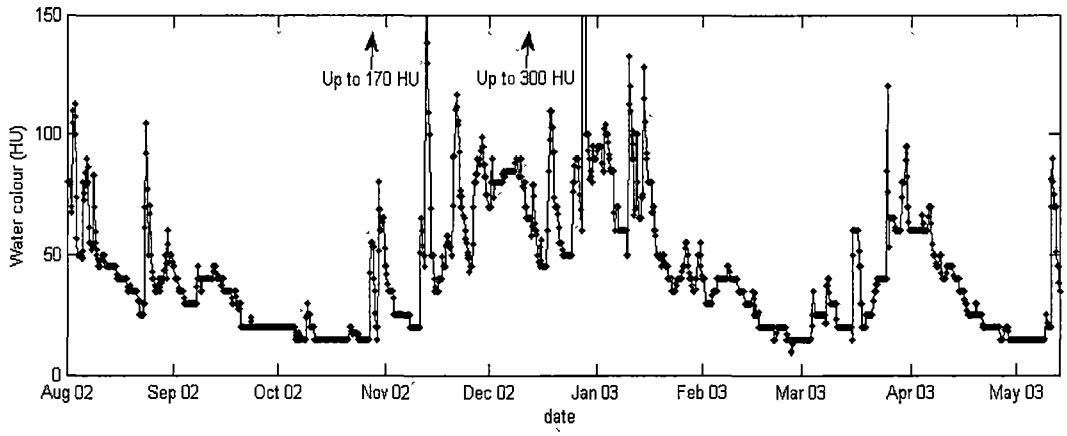
In order to consider seasonal changes, a plot of monthly mean of clarified water turbidity is shown in Figure 2.10c. Since all monthly means of clarified water turbidity were less than 1 NTU except those which marginally failed during the two periods of August 2002 and December to January 2002, the clarifier performance met target turbidity adequately. Poor performance in August during a period of the first flood in early spring was expected when peaks of raw water turbidity occurred. A density current due to high temperature in the summer period was likely to be the main cause of poor performance in December and January. The overall mean of clarified turbidity was 0.69 NTU with standard deviation of 0.50, which well met the operational target of 1 NTU. However, according to the data at hand, it was found that 339 out of 1700 runs (about 20 percent of data) were above the threshold limit of 1 NTU. Most of the turbidity upsets that occurred resulted in the raw water turbidity appearing in outlier range (i.e. outside of the mean ± one standard deviation). The plant operators could recognise raw water turbidity changes as evidenced by r of 0.48 to clarified water turbidity.

Colour Removal Analysis

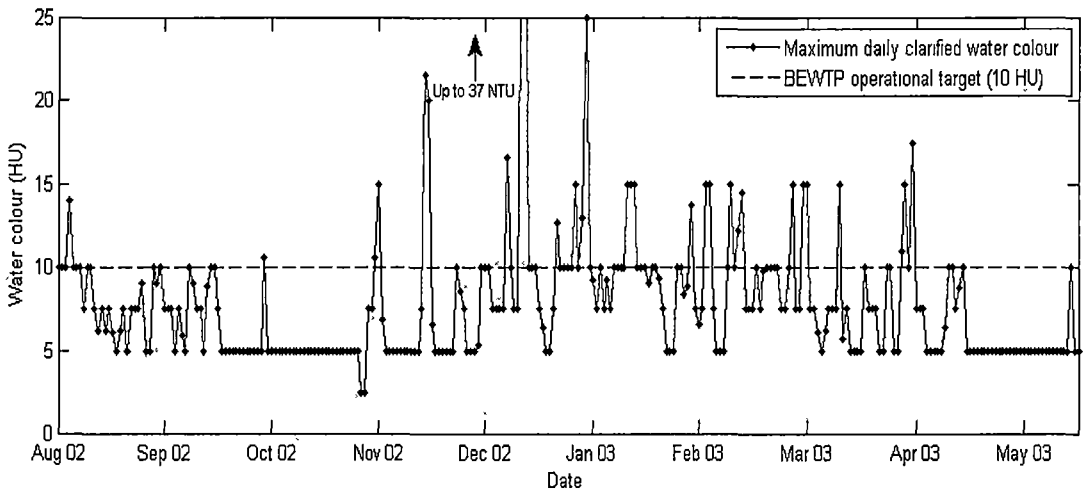
The plots of daily maximum raw water and clarified water colour are shown in Figure 2.11a and 2.11b, respectively. In the same manner as previously discussed, it was found that during the period of August 2002 to November 2002, the operators performed very well to keep the clarified water colour satisfying the operational target of 10 NTU. There were some evidence of fluctuations, but to a mild degree. It is a good example of a stable operation. During the period of December 2002 to April 2003, the operational stability was severely interfered by several peaks of raw water colour as evidenced by a number of spikes in the clarified water colour plot. After April 2003, there were a few peaks of raw water passing to the plant but the operational stability could be maintained until May 2003.

In Figure 2.11c, the monthly means of clarified water colour completely complied with BEWTP operational criteria regardless of seasonal changes. All monthly means of clarified water colour were less than target of 10 HU. By using data of every month, the overall mean of clarified water was 6.69 HU with standard deviation of 2.54. Therefore, it met the operational target of 10 HU. The number of upset runs was 63 runs from 1700 runs, which was 3.7 percent. This was less than that of turbidity since the operators intentionally focussed on colour removal due to raw water having very low turbidity. The operators were able to handle a change of water colour better than that of turbidity as

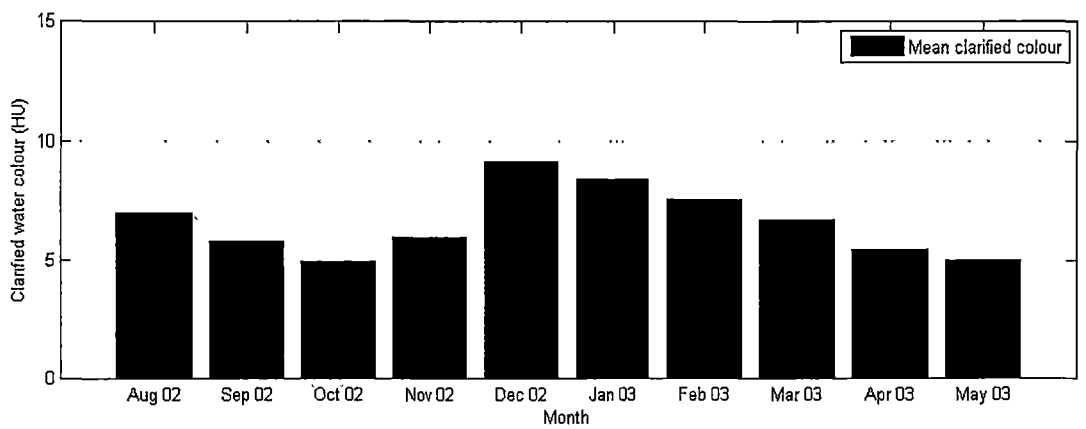
evidenced in a higher correlation (r) of 0.51 to clarified water colour. The associated statistical values are shown in Table 2.5.



(a) Maximum daily raw water colour



(b) Maximum daily clarified water colour



(c) Monthly clarified water colour

Figure 2.11 Clarifier Performance Evaluations (Colour approach, BEWTP)

In conclusion, according to the BEWTP operational criteria of 1 NTU clarified water turbidity and 10 HU clarified water colour, the overall clarifier performance was very good. It almost perfectly complied with colour operational criteria although there was evidence of some unacceptable levels. The mean clarified water turbidity of 0.69 NTU was well within to the requirement of 1 NTU, and most of the turbidity peaks occurred when the raw water was most turbid. Fortunately, the raw water quality was generally excellent and the succeeding unit of slow sand filters were designed to take care of clarified water with turbidity up to 5 to 7 NTU. Thus the clarifier operation became less critical.

2.8.2 Bang Khen Water Treatment Plant Performance Evaluation

For BKWTP, the treatment performance of the clarifier unit is evaluated by utilising a set of data spanning from February 2003 to October 2004 that contains turbidity of raw water and clarified water at four-hour sampling intervals. In the same manner as the BEWTP case study, the maximum turbidity values each day are used in the trend analysis of performance. The associated statistical parameters are shown in Table 2.6. Only a clarified water turbidity of seven NTU is set as acceptable criteria for clarified water quality since raw water colour is *apparent colour* (MWA, 2000).

Statistic parameters	values
Raw water turbidity (Mean±1SD)	93.27±34.37 NTU
Clarified water turbidity (Mean±1SD)	6.34±1.75 NTU
Raw water & clarified water turbidity correlation (r)	0.29

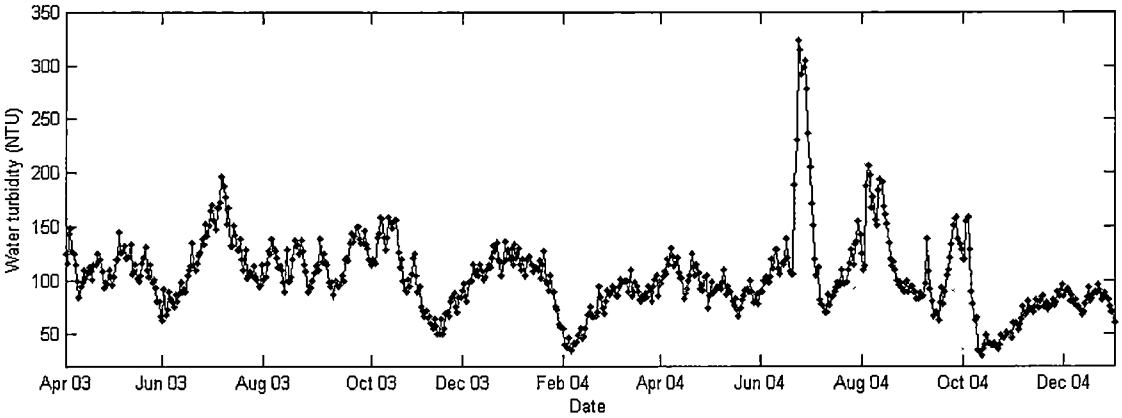
Table 2.6 BKWTP performance analysis

The clarifier performance trends are shown in Figure 2.12a and 2.12b, which are the plots of the maximum daily raw and clarified water turbidity respectively. By comparing these two trend plots, it can be noted that the raw water turbidity showed a large number of spikes and considerable variability. Variability is obviously evident in the clarified water turbidity plot. There are also a larger number of spikes, which indicates operational instability, and most of them rose above the operational criteria of 7 NTU.

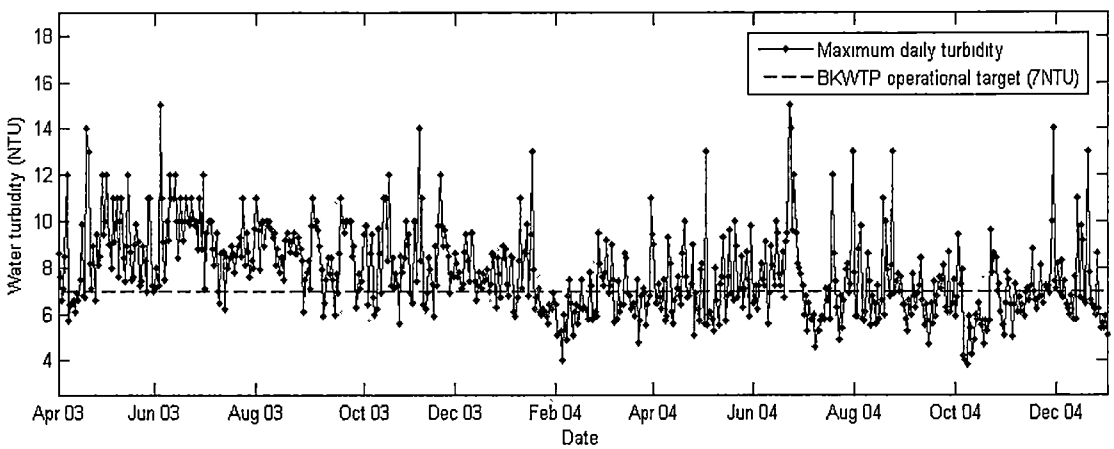
The variation and peaks in raw water turbidity were expected and apparently showed some links with the seasonal change. Instead of occurring in line with the peaks of raw water turbidity, the spikes in clarified water turbidity plots occurred in complete disorder as evidenced by a small r of 0.29. The reason is that variation in raw water turbidity may not be the only source contributing to operational instability, and the other sources can be environmental changes (e.g. density current). This can be confirmed by the chief operators' interviews that raw water quality did not change much but there were large fluctuations of clarified water quality during the day. This was because large changes in water temperature led to density current, especially in the afternoons of summer months. This situation may repeat several times each day.

The effects of seasonal change can be determined in the plot of monthly mean of clarified water turbidity as shown in Figure 2.12c. The monthly means of clarified water usually conformed to the operational target of 7 NTU. An exception was found in those of the rainy season of 2003 (i.e. May 2003 to August 2003) since a number of raw water turbidity peaks reached the plant during a flooding period of rain. This is also evidenced in the slight pass of June in the early rain of 2004.The overall mean of clarified water turbidity was about 6.34 NTU with standard deviation of 1.75, which marginally complied with the operational target of 7 NTU. However, for about 30 percent of the time the clarified water turbidity exceeded 7 NTU. Additionally, the raw and clarified water turbidity correlation (r) of 0.29 to clarified water turbidity is small, meaning that the operators were not able to effectively recognise changes of quality in raw water.

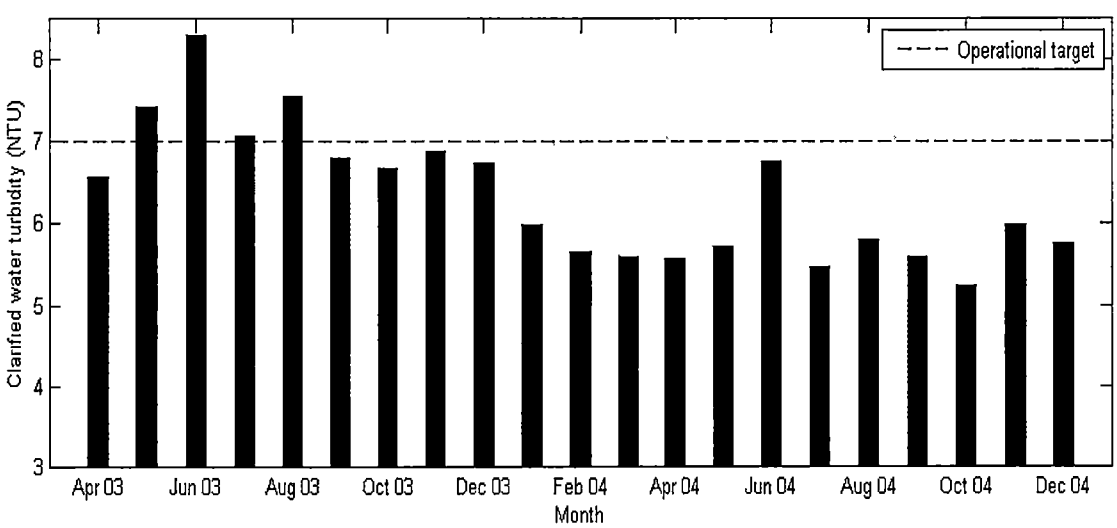
According to BKWTP performance analysis, it can be concluded that there is a prime need for BKWTP to improve operational performance. This is due to the fact that the mean clarified water turbidity only marginally met the operational target. Also there were always operational instabilities as evidenced by the fact that for about 30 percent of the time, clarified water turbidity exceeded 7 NTU. This placed a heavy burden on the succeeding filter units since they were designed for maximum load of 5 to 7 NTU. This overload resulted in expensive filter backwash. However, this is an unavoidable situation in BKWTP due to the extensive water demand. On top of that, to keep operational stability during environmental changes, such as variations in density current, which occur repeatedly during the day, control actions need to be provided promptly, in order to keep the clarified water turbidity within the operational criteria.



(a) Maximum daily raw water turbidity trend



(b) Maximum daily clarified water turbidity trend



(c) Monthly clarified water turbidity

Figure 2.12 Clarifier performance evaluations (Turbidity approach, BKWTP)

2.9 Chapter Conclusion

In this chapter, the background of water treatment plants and their clarifiers were given in order to identify problems usually found in clarifier control. The application of artificial intelligence in clarifier control is not fully mature yet. Most past work using ANN was related to prediction and clarifier control. Attempts to mimic how the human operators control the clarifier were not entirely complete since most focus only on the chemical dosages. The physical control actions (i.e. turbine speed and sludge drainage rate) were not included in past studies. All early works also focused only on improving the qualities of clarified water but not upon the operational cost. Mathematically, all past research work was accounted as single objective optimisation problem rather than a multiple objectives optimisation problem.

Past operational data from two case studies were used to conduct performance evaluations in order to characterise the past operations and to give evidence of the need for improvement. The operators in BEWTP performed very well since the mean quality values of clarified water routinely satisfied their operational criteria. In addition, this performance was superior to that of BKWTP where its operation was often under stress conditions. Although the mean BKWTP clarified turbidity generally meets its target (marginally) by about 30 percent of the time, BKWTP operators fail to meet their own criteria. Improving the BKWTP operation is set therefore as a priority task.

In common to both case studies, their control procedures and operational problems were reviewed to set out a conceptual model for improving performance. Delay in receiving laboratory test results and human error in selecting control actions were recognised as the main sources of operational problems. The approach proposed is to employ artificial intelligence techniques to address these perceived problems. The first approach is to use ANN modelling to model the clarifier process. This is analogous to the Jar test and will minimise the time delay. The second is to employ a GA to optimise the set of control actions to minimise the operational cost and improve water quality, which would avoid human errors.

3.0 WATER TREATMENT PLANT DATA: COLLECTION, MANIPULATION AND ANALYSIS

3.1 Introduction

This chapter presents a detailed description of data collection, manipulation and statistical data analysis for two cases studies, Bryn Estyn Water Treatment Plant (BEWTP) and Bang Khen Water Treatment Plant (BKWTP). The primary objectives of these data analyses are to gain familiarity with the study domain and to prepare available data for the clarifier process model development in Chapter 5. In terms of data preparation, the irrational data will be eliminated to ensure data quality before modelling. Additionally, the data which are collected with the different measurement frequency will be discretised to the same time interval.

It is essential to ensure that the data are fully representative of normal and upset operation conditions and cover all the effects of seasonal changes. The sampling period should and cover at least a full operational year since the data are site specific and are influenced by seasonal fluctuations in raw water quality and the frequency of process upset conditions (USEPA, 1998).

3.2 Data Manipulation and Analysis

The data routinely collected at BEWTP and BKWTP were grouped into four categories: raw water quality, control variables, process variables and clarified water quality. The raw water quality data consisted of the various water quality variables (e.g. water turbidity and colour), which were available from online sensors and laboratory results. These parameters were known to affect treatment processes (ASCE. & AWWA., 1990). The control data was a set of data concerning those variables that can be changed at will by operators, such as chemical dosages. The process variables were observed in order to track any change inside the clarifiers. The performance of a clarifier was measured via clarified water qualities, which were water turbidity and/or colour.

Data manipulation was concerned with two issues, the first one was data filtration to eliminate irrational data and the second one was a manipulation of measurement frequency. To ensure quality of data, it was necessary to filtrate irrational data. Irrational data normally occur from a typing error in entering data by hand. Sometimes, irrational

data are included during the sensor calibrations. This is always the case in online sensors. The filter criteria were set for each case study according to whatever ranges were possible in each variable. After all of the irrational data were eliminated, the new data would be inserted. The new data values could be obtained by the interpolation of the nearby data. This method was recommended in a number of works (Baxter et al., 2001; Maier & Dandy, 1998; Zhang & Stanley, 1997) .

For both case studies, the data were taken at several measurement frequencies as shown in Table 3.1 and 3.5. However, in this thesis, the data would be discretised by interpolation to be at the same measurement frequency of every four hours which is how the operators control the clarifier. In common with the BEWTP and BKWTP operational manuals, the control actions need to be selected twice a shift or every four hours at least. This is because the water is retained in water treatment plant for about four hours, and for most of that time it is actually retained in the clarifier where it varies from 1.5 to 3.5 hours depending on type of clarifier (Kerri, 1996). Therefore, if anything goes wrong, the operator should be able to detect it and take action. Additionally, in most past research, the clarifiers were modelled on the basis of daily average data. This is questionable procedure for predicting in the lower or higher frequencies (Riyaz et al., 2004).Therefore, if the model is based on the data discretised to four hours time steps, then the model should predict and work in line with the operators. In this thesis, all data from both case studies will be discretised to four hours time steps.

3.3 Bryn Estyn Water Treatment Plant Data

Operational data were collected from August 2002 to May 2003. These included the data from online sensors and the laboratory test results. These variables are known to affect the clarifier performance significantly (ASCE. & AWWA., 1990). Both data from online sensors and laboratory test results were logged to a Supervisory Control and Data Acquisition (SCADA) system. From the online sensors, the data were inputted into SCADA every 30 minutes. The set of online sensors are shown in Figure 3.1. The laboratories' results were manually entered into the SCADA system every four hours. The measurement schedules are shown in Table 3.1.

Figure 3.1 Set of online sensors at BEWTP

Variables	Online /Lab	Time					
		0:00	4:00	8:00	12:00	16:00	20:00
Sludge concentration	Test kit	X	X	X	X	X	X
Turbine speed	Visual	X	X	X	X	X	X
Raw water Colour	Lab	X	X	X	X	X	X
Clarified water colour	Lab	X	X	X	X	X	X
Raw water turbidity	Online	Every 30 minutes					
Clarified water turbidity	Online	Every 30 minutes					
Raw water Ph	Online	Every 30 minutes					
Raw water temperature	Online	Every 30 minutes					
Alum dosage	Online	Every 30 minutes					
Polyelectrolyte dosage	Online	Every 30 minutes					
Plant flow rate	Online	Every 30 minutes					

Table3.1 BEWTP measurement Schedule

Calibration tests of associated sensors, dosage pumps and chemical weighing machines were routinely conducted every few months. However, no records of calibration data had been kept. Although online sensors (e.g. turbidimeter) benefit to the clarifier operation, it should be noted that without the support of quality assurance and control, the accuracy of an online turbidity meter is questionable in the low turbidity range (Burlingame et al., 1998; USEPA, 1999). In the operator's experience, the measurement errors of online

turbidity were sometimes up to 30 percent. All the associated measurement errors were approximated by the chief operator as tabulated in Table 3.2.

Variables	Measurement Error
Raw water turbidity (NTU)	10%
Raw water pH (pH units)	5%
Raw water temperature (C°)	5%
Raw water colour (HU)	1 HU
Plant flow rate (MLD)*	15%
Alum dosage (mg/L)	10%
Polyelectrolyte dosage (mg/L)	10%
Sludge concentration (% by volume)	5%
Turbine speed (RPM)	10%
Clarified water colour (HU)	1 HU
Clarified water turbidity (NTU)	10-15%

* Million litres per day.

Table 3.2 Estimated measurement errors (BEWTP)

3.3.1 Bryn Estyn Water Treatment Plant Data Preparation

In order to ensure that correct data are used, the data were filtered to remove occasional outlier values that the plant chief operator advised are likely to be associated with sensor calibration and maintenance. The other source of irrational data is human error. These are especially found in hand entered data. Those irrational data were filtrated out. Filter criteria are site-specific. The Filter thresholds were set according to normal range of each variable shown in the plant operation manual (City Water Technology, 2001) and combined with the chief operators' suggestions as shown in Table 3.3. These values have all been replaced by interpolated data.

Variable	Filter
Raw water turbidity (NTU)	> 0.2
Raw water pH (pH units)	6.0-8.8
Raw water temperature (°C)	2-25
Sludge concentration (% by volume)	15-30
Plant flow rate (MLD)	15-110
Turbine speed (rpm)	5-15
Alum dosage (mg/L)	15-60
Polyelectrolyte dosage (mg/L)	0.10-0.30
Target clarified water quality (NTU)	< 3.0

Table3.3 BEWTP data filter

3.3.2 Bryn Estyn Water Treatment Plant Data Analysis

After the irrational data were filtered out, the statistical data analysis was conducted. As will be discussed, statistical analysis of raw water qualities, control variables, clarified water qualities and process variables are shown. Their associated correlations (r) with operational targets (i.e clarified turbidity and colour) are used to illustrate their relations. Their coefficient of correlations (r) will be applied for primary model input selection guideline in Chapter 5. The associated statistical parameters are summarised in Table 3.4. In this section, graphical presentations are provided for raw clarified water turbidity and the all of variables are shown in Appendix A.

Variable	Nature	Range (Min-Max)	Mean	SD	Correlation (r) to	
					Clarified turbidity	Clarified colour
Raw water turbidity (NTU)	Raw water	0.20 -99.9	2.92	3.60	0.48	0.26
Raw water pH	Raw water	6.63-8.16	7.26	0.27	-0.37	-0.25
Raw water temp. (°C)	Raw water	5.30-21.33	13.62	4.65	0.43	0.37
Raw water colour (HU)	Raw water	10-300	44.41	28.53	0.50	0.51
Sludge concentration. (%)	Process	16-46	25.07	5.47	0.07	0.02
Plant flow rate (ML/day)	Control	18.83-102.19	58.20	21.75	-0.07	-0.09
Alum dosage (mg/L)	Control	15.15-52.06	26.07	6.87	0.42	0.45
Polyelectrolyte dosage (mg/L)	Control	0.10-0.30	0.15	0.039	0.06	0.11
Turbine speed (RPM)	Control	7-15	10.4	1.3	0.09	-0.23
Clarified water turbidity (NTU)	Clarified water	0.09-6.72	0.69	0.50	1.00	0.47
Clarified water colour (HU)	Clarified water	2.5-36.68	6.69	2.54	0.47	1.00

Table 3.4 BEWTP operation data analysis

Raw Water Quality Data

For BEWTP, raw water quality data contained raw water turbidity, pH, temperature and colour data. Only raw water colour was obtained from laboratory test results. The other variables were available via online sensors. The sampling rates of associated online sensors were every 30 minutes, and the colour test was routinely taken every four hours in the laboratory. The sampling points were located at the raw water pumping station. All the raw water qualities were of the raw water from the river without pre-coagulation pH adjustment. This unit was not available during period of data collection (City Water Technology, 2001).

Most of the raw water parameters varied due to seasonality in the Derwent River's flow characteristics. The Derwent River carried a small amount of sediment because of low erosive energy, resulting in low raw water turbidity. However, there was a large amount of humus and peat in the river, which gave it a characteristic brown true colour (City Water Technology, 2001). The plots of raw water turbidity and colour are shown in Figure 3.2 and 3.4, respectively. The raw water turbidity varied in the range from 0.2 NTU to 100 NTU. The raw water colour ranged from 10 to 300 HU. The water turbidity and colour values were reported with mean and standard deviation of 2.92 ± 3.60 NTU and 44.41 ± 28.53 HU, respectively (Table 3.4). There was a slight link between raw water colour and turbidity as evidenced by a small correlation (r) of 0.54. Since raw water was true colour type, it was a slight contribution for turbidity. Raw water turbidity varied extensively, because raw water was directly pumped from the Derwent River. Thus there was no reservoir to buffer turbidity variation. The peaks of raw water turbidity (99.91 NTU) and colour (300 HU) were found concurrently during the autumn period, associated with snowmelt following the first thawing.

Raw water pH varied between 6.63 and 8.16, its mean and standard deviation was 7.26 ± 0.27 (Table 3.4), which was in a normal range and suitable for treatment. The correlation (r) to clarified water turbidity and colour were -0.37 and -0.25, respectively. In addition, City Water Technology (2001) provided information that hardness and alkalinity levels were in the range of 10-30 and 7-23 mg/L as CaCo3 was relatively low. Occasionally raw water with high colour and low alkalinity was difficult to treat.

Control Variable Data

The plant operators could manipulate control variables in order to attain the criteria for clarified water quality. In practice, there are five control actions related to clarifier control. These include adjusting alum and polyelectrolyte dosages, changing the mixing intensity (i.e. the turbine speed), changing the plant flow rate and the sludge drainage rate. The sludge drain rate is defined as the time (in seconds) that the valve at the bottom of the clarifier is opened to drain the excess sludge in one hour. Therefore, the sludge drain rate is assigned in terms of seconds per hour. In BEWTP, the sludge drain rate was set to 10 to 20 seconds for every one hour. However, no records were available.

Operators could change the alum, polyelectrolyte dosages, turbine speed and, to a lesser extent, the plant flow rate. Plant flow rate could not be changed without considering the water demand. The plant flow rate data were logged into the SCADA system every 30 minutes by an online flow meter. By knowledge of the plant flow rate, alum and polyelectrolyte dosages could be calculated and logged into SCADA at the same frequency. The turbine speed data was manually recorded every four hours.

Plant flow rate ranges from 18.83 to 102.19 million litres per day (MLD). The reported mean and standard deviation was 58.20 ± 21.75 MLD (Table 3.4). The correlation (r) to clarified water turbidity was -0.07 and to colour was -0.09. This inverse correlation was negligible. The presence of treated water storage at the end of the process resulted in a regulated plant flow rate. Thus the operator could keep the plant flow nearly constant during a shift as long as it met community water demand.

Alum and polyelectrolyte were used as chemical bases for coagulation to control turbidity and colour. Alum dosage ranged from a low of 15 mg/L to a high of 53 mg/L. Polyelectrolyte was an anionic polyacrylamide type with a commercial name of Magnafloc LT22. Polyelectrolyte dosage was in the range of 0.10 to 0.30 mg/L. The means and standard deviations of alum was 26.07 ± 6.87 mg/L and that of polyelectrolyte dosages was 0.15 ± 0.039 mg/L (Table 3.4). The correlation (r) of alum to clarified water turbidity was 0.42 and to clarified water colour was 0.45. On the other hand, the correlation (r) of polyelectrolyte to clarified water turbidity was 0.06 and to clarified water colour was 0.11. Consequently, adding alum showed a larger effect on turbidity and colour removal than adding polyelectrolyte. This was evidenced by a larger correlation (r) as above. These amounts of chemical dosages in the treatment process

complied with drinking water standard guided by United States Environmental Protection, (USEPA) and Department of Environment, UK (e.g. for the polyelectrolyte, it was allowed to be used up to 0.5 mg/L and exceptional dosage of 1.0 mg/L) (Masschelein, 1992).

Turbine speed varied from 7 to 15 RPM. Its mean and standard deviation were reported as 10.4 and 1.3 RPM respectively (Table 3.4). The turbine speed weakly correlated to clarified water turbidity with a correlation of 0.09 but it inversely correlated to clarified water colour with correlation of -0.23.

Clarified water quality

The clarified water turbidity and colour were used as a measure of the clarifier performance. The clarified turbidity and colour were measured by online sensors and were subsequently tested in the laboratory. In addition to analysis concerning clarified water quality using trend and statistic analysis in Chapter 2, more details are presented here.

- **Clarified Water Turbidity**

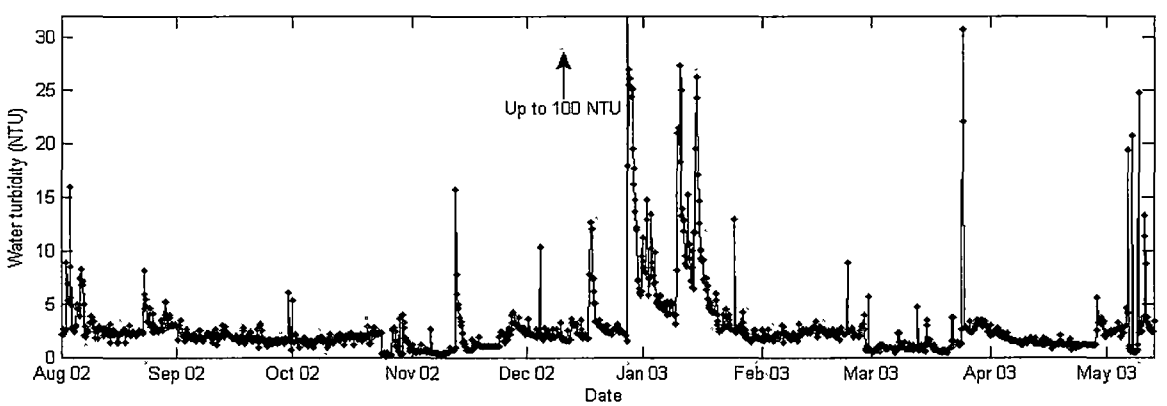


Figure 3.2 Raw water turbidity (BEWTP)

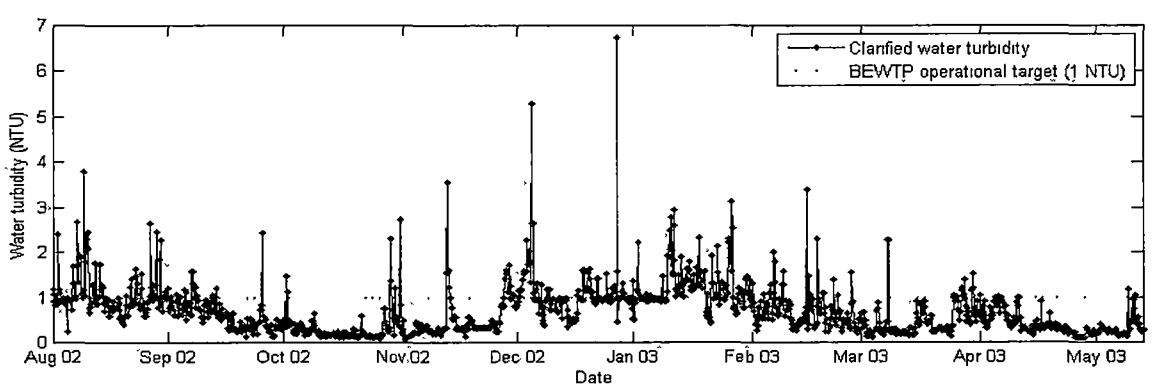


Figure 3.3 Clarified water turbidity (BEWTP)

The plot of clarified water turbidity (Figure 3.3) varied in line with raw water turbidity. The clarified water turbidity ranged from 0.09 to 6.72 NTU. The mean and standard deviation was 0.69 ± 0.50 NTU. The best clarified water turbidity was found in spring, and the worst were in early autumn in periods of snow melt. The clarified water turbidity slightly correlated with raw water turbidity as evidenced by a correlation of 0.48, implying that the operators managed to catch up with raw water turbidity to a moderate degree.

- **Clarified Water Colour**

The plot of clarified water colour is presented in Figure 3.5. The clarified water colour ranged from 2.5 to 36.7 HU. In common with clarified water turbidity, the best clarified water turbidity and colour were concurrently found in spring, and the worse were in early autumn in periods of snow melt. Referring to the discussion in Chapter 2, the mean and standard deviation was 6.69 ± 2.54 HU, complying with target water colour of 10 HU. The upset percentage to 10 HU was about 3.68 percent reflecting that operators performed very well. The clarified water colour showed comparative relation with raw water colour as evidenced by correlation of 0.51, which was almost equal to that of clarified water turbidity (r =0.48). This raised the question why operators performed much better in colour removal compared with turbidity removal. This was because the raw water was *true colour* which could not precipitate concurrently with turbidity. In addition, the operators focused on colour removal rather than turbidity removal since the raw water turbidity was always low.

In general all variables were in normal ranges and suitable for treatment, and there were some physical links between each other as previously explained. One of the key difficulties for the operation was the raw water being *true colour*. Thus it was difficult to precipitate at the same time with turbidity. The other difficulty was that raw water turbidity fluctuated since the water was directly pumped from the river without a buffering reservoir. However, the BEWTP clarifier generally operated at or near optimum.

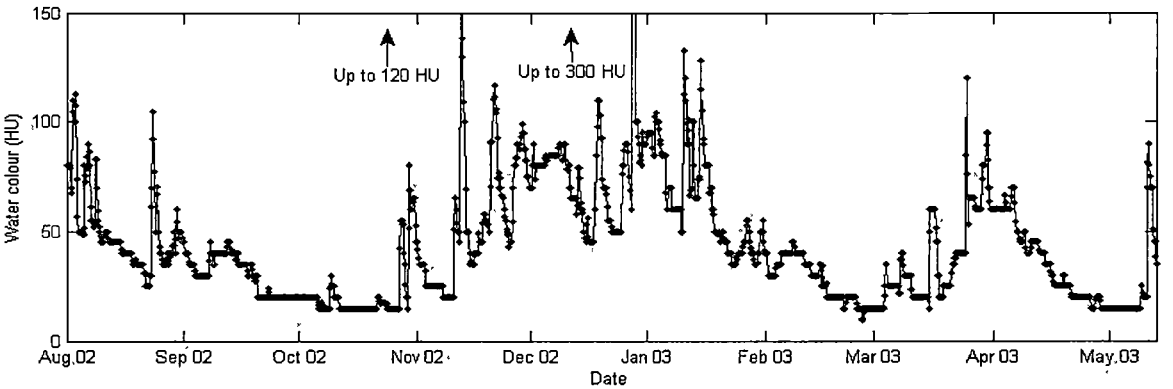


Figure 3.4 Raw water colour (BEWTP)

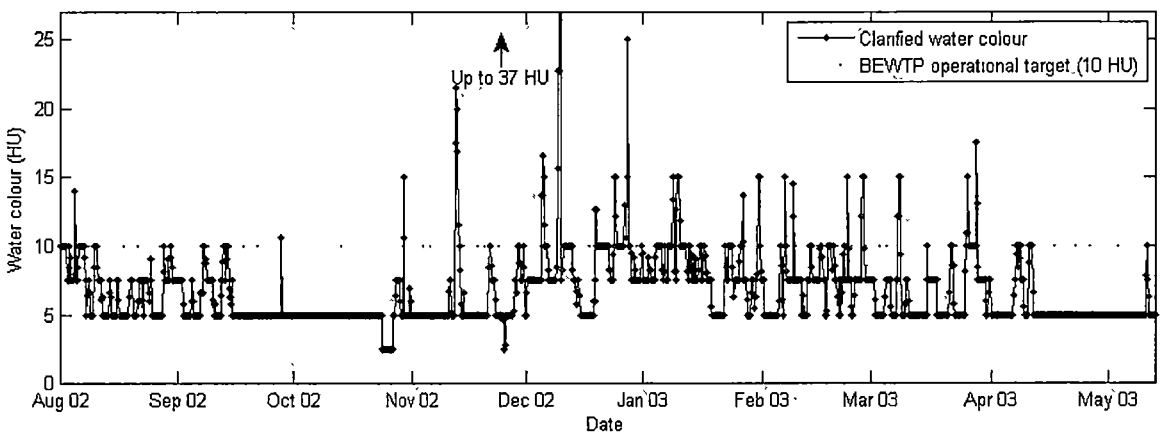


Figure 3.5 Clarified water colour (BEWTP)

3.4 Bang Khen Water Treatment Plant Data

The operational data used for the clarifier process model of BKWTP was from February 2003 to December 2004. The list of variables and associated measurement schedule is shown in Table 3.5. However, the sludge concentration has had to be ignored since its data was not available. At BKWTP, plant operators collected data from online sensors and laboratory results. Although a number of online sensors were present in most clarifier units, the full linkage with the SCADA system was not yet available. It was planned to be fully implemented by the end of 2008. Therefore, online data was read from the meters and manually recorded in hard copy format every four hours. Simultaneously, lab test results were reported and recorded in the same format.

The chemical dosages were set according to a combination of the Jar test results and the operators' experience. Jar tests were conducted twice a day at 8:00 and 16:00. The chemical dosages were actually calculated after the plant flow rate was reported. The

chemical weighting machine and chemical flow meter errors might contribute to chemical dosage errors. The operators also adjusted the turbine speed and sludge drain rate. According to the operational manual, all these actions should be set at the beginning of the first hour and fifth of each 8 hour shift. It could be more often during period of bad water quality.

Variables	Online /Lab	Time					
		0:00	4:00	8:00	12:00	16:00	20:00
Raw water turbidity	Online	X	X	X	X	X	X
Raw water pH	Lab	X	X	X	X	X	X
Alkalinity	Lab	X	X	X	X	X	X
Slurry Concentration	Test kit			X		X	
Alum dosage	Calculated	X	X	X	X	X	X
Polyelectrolyte dosage	Calculated	X	X	X	X	X	X
Plant flow rate	Online	X	X	X	X	X	X
Turbine speed	Visual	X	X	X	X	X	X
Auto Drainage	Visual	X	X	X	X	X	X
Clarified water turbidity	Lab	X	X	X	X	X	X

Table 3.5 BKWTP measurement schedule

Although sensor calibrations and adjustment were occasionally conducted, these were not recorded. Therefore the associated measurement errors were all approximated according to consultations with the chief operator (Table 3.6).

Variables	Measurement Error
Raw water turbidity (NTU)	10%
Raw water pH (pH unit)	5%
Raw water Alkalinity (mg/L of $CaCo_3$)	10%
Raw water conductivity (μS/cm)	10%
Alum dosage (mg/L)	10%
Polyelectrolyte dosage (mg/L)	10%
Plant flow rate (MCMD)	15%
Turbine speed (RPM)	10%
Auto Drainage (Sec/hour)	5%
Clarified water turbidity (NTU)	10-15%

Table 3.6 Estimated BKWTP measurement errors

For the BKWTP case, the raw water colour was largely attributed by suspended solids (i.e. turbidity). This was considered *apparent colour* type, which could be removed with turbidity. For this reason, target clarified water quality criteria was set only to a clarified turbidity of 7 NTU and did not include colour (MWA, 2000).

3.4.1 Bang Khen Water Treatment Plant Data Preparation

There were a number of incomplete and missing data since the BKWTP operational data were not stored in an electronic version. All of them were recorded manually in hard copy format. In the same manner as the BEWTP case, to avoid irrational data, the data filter criteria were set at the suggestion of the chief operator. Missing data were filled by using an interpolation method. The associated filter thresholds are shown in Table 3.7. The range of each variable was widely extended compared to BEWTP since here the raw water varied greatly in quality.

Variable	Filer
Raw water turbidity (NTU)	>15
Raw water (pH unit)	6.8-8.5
Raw water Alkalinity (mg/L as $CaCo_3$)	35-125
Raw water conductivity (μS/cm)	>90
Plant flow rate (MCMD*)	15-30
Alum dosage (mg/L)	10-75
Polyelectrolyte dosage (mg/L)	0.0-.08
Turbine speed (rpm)	0.8-3
Auto Drainage (min/hour)	>1
Clarified water turbidity (NTU)	<20

MCMD is million cubic meter per day

Table3.7 Bang Khen water treatment plant data filter

3.4.2 Bang Khen Water Treatment Plant Data Analysis

After any irrational data had been removed, statistical analyses were conducted. In case of BKWTP, the operational data were categorised into three types: raw water quality, control variables and clarified water quality. No process variable was used in this case. In the same manner, the significant variables, raw and clarified water turbidity plots were presented in this section. All of variables were graphically shown in Appendix B. The associated statistical analyses are outlined as follows and summarised in Table 3.8.

Variable	Nature	Range (Min-Max)	Mean	SD	Correlation (r) to clarified turbidity
Raw water turbidity (NTU)	Raw water	23-324	93.27	34.37	0.29
Raw water pH	Raw water	7.05-8.13	7.57	0.15	-0.07
Raw water alkalinity ($CaCo_3$)	Raw water	48-114	86.35	12.96	-0.015
Raw water conductivity (μS/cm)	Raw water	103-473	249.01	46.21	-0.05
Plant flow rate (MCMD)*	Control	16-26	23.21	1.26	0.004
Alum dosage (mg/L)	Control	15-70	31.38	10.23	0.11
Polyelectrolyte dosage (mg/L)	Control	0.01-0.07	0.023	0.019	0.19
Turbine speed (rpm)	Control	0.97-2.14	1.45	0.25	-0.07
Sludge drain rate (sec/hour)	Control	20-120	40.38	18.52	0.10
Clarified water turbidity (NTU)	Clarified water	2.2-15	6.34	1.75	1.00

MCMD* = million cubic metres per day

Table 3.8 BKWTP operation data analysis

Raw Water Quality

Raw water quality included raw water turbidity, pH, alkalinity and conductivity. Each of these had a significant effect on clarifier water quality. Raw water turbidity, pH and conductivity data were manually recorded from online sensors every four hours. On the other hand, water alkalinity data were received from laboratory tests every four hours. All the raw water qualities were sampled after pre-coagulation pH adjustment before the clarifier.

The source of raw water and its quality varied considerably due to substantial seasonal variation in the Chao Phraya River flow rate as well as ambient air temperature. The raw water turbidity was extremely high since the Chao Phraya River was located upon the Chao Phraya alluvium bed with high erosion during especially the rainy season. This introduced a great deal of clay and suspended solids to the water source. Raw water turbidity ranged from 23 NTU during the winter to 324 NTU during the first flood period of an early rainy season. The raw turbidity was reported to have mean and standard deviation of 93.27 ± 34.37 NTU (Table 3.8). Apart from first flood effects, there was another source of raw water turbidity peaks. The turbidity peaks occasionally occurred at times out of the rainy season because of sand mining and washing operations on the Chao Phraya River located 30 kilometres upstream from the plant site. However, since the year 2006, sand mining was illegal and banned. The plot of raw water turbidity variation was shown in Figure 3.6. Raw water turbidity correlated only insignificantly to clarified water turbidity as evidenced in a small correlation of 0.29 (Table 3.8). This confirmed that the operators were unable to control for a change of raw water turbidity and /or there were other sources of clarified water turbidity upset, possibly the density current.

Raw water conductivity, pH and alkalinity were sampled after lime dosing. In BKWTP pre-coagulation pH adjustment was conducted by using lime (MWA, 2000). Raw water conductivity was reported with a mean and standard deviation of 249.02 ± 46.21 μS/cm (Table 3.8). The pH and alkalinity of raw water varied with a similar small variation with mean and standard deviation of 7.56 ± 0.15 and 86.35 ± 12.96 mg/L ($CaCo_3$) respectively (Table 3.8). In summary, raw water quality for BKWTP had very high turbidity. The pH and alkalinity fell into a normal range, which was suitable for turbidity precipitation with alum (MWA, 2000).

Control Variables

Similarly to BEWTP, control actions included adjusting plant flow rate, changing alum dosages and polyelectrolyte dosages, changing turbine speed and sludge drainage. The statistical analyses of the control action data are shown as follows.

The plant flow rate depended on water supply demand. It ranged from a minimum of 16 million cubic metres per day (MCMD) to a maximum of 26 MCMD. The plant flow rate was reported with mean and standard deviation of 23.21 ± 1.26. MCMD (Table 3.8). The

small value of standard deviation indicated that there was not much variation in the plant flow rate since there was a large treated water reservoir located at the end of the treatment process to regulate the plant flow rate. The plant flow rate changes also weakly affected the clarified water turbidity (r of 0.004).

Alum dosage was used as a main chemical to remove turbidity. Besides the use during high turbidity, polyelectrolyte was frequently applied to assist turbidity removal. The operators only used alum to precipitate turbidity when raw water turbidity was about 20 NTU or less. Alum dosages varied concurrently with raw water turbidity and ranged from 15 mg/L to 70 mg/L. It had a mean and standard deviation of 31.38 ± 10.23 mg/L (Table 3.8). During period of April, June and October, the raw water turbidity was low with a level less than 50 NTU, and sometimes they went down to about 20 NTU. Therefore a small amount of alum was used. However, the alum dosage was never below 15 mg/L. Polyelectrolyte was an anionic polyacrylamide type with a commercial name of Super Floc A100.The polyelectrolyte dosages also varied substantially with mean and standard deviation of 0.023 ± 0.019 mg/L (Table 3.8). Alum and polyelectrolyte dosages correlated to clarified water turbidity with low correlation of 0.11 and 0.19 respectively. All chemical dosages used in BKWTP were harmless to health. Especially for the polyelectrolyte, it was allowed to be used up to 0.5 mg/L and exceptional dosage of 1.0 mg/L as guided by United States Environmental Protection, (USEPA) and Department of Environment, UK (Masschelein, 1992).

The turbine speed varied in the range of 0.97 to 2.14 RPM through the whole year. It had a mean and standard deviation of 1.45 ± 0.25 RPM. Its correlation to clarified water turbidity was vary small (r = -0.07) (Table 3.8).

Normally, it was necessary to drain sludge in order to maintain the sludge concentration in a clarifier otherwise the turbidity would be carried over. The automatic drainage system was set by setting the drainage time in seconds per one hour. The sludge drainage rate varied with mean and standard deviation of 40.38 ± 18.52 second per hour, and ranges from 20 to 120 second per hour (Table 3.8). The sludge drain rate related slightly to clarified turbidity with a small inverse correlation (r of 0.1).

Clarified Water Quality

In the case of BKWTP, only the clarified water turbidity was used to measure the clarifier performance. Clarified water samples were sampled from the outlet of the clarifier, and then a turbidity test was conducted using an online turbidimeter every four hours and recorded by hand. Clarified water turbidity was plotted in Figure 3.7. It ranged from 2.2 to 15 NTU with mean and standard deviation of 6.34 ± 1.75 NTU (Table 3.8). According to an interview with the chief operator, poor raw water quality during the rainy season always induced an operation difficult to correct but it was foreseen and able to be handled. However, the unpredicted raw water turbidity outside the rainy season from sand mining gave far more severe results. In addition, there were some turbidity leaks caused by a density current. This always occurred on a day of high temperatures.

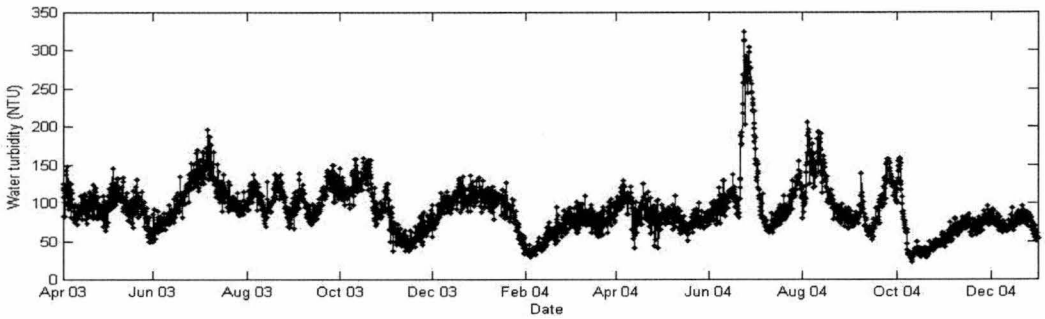


Figure 3.6 Raw water turbidity (BKWTP)

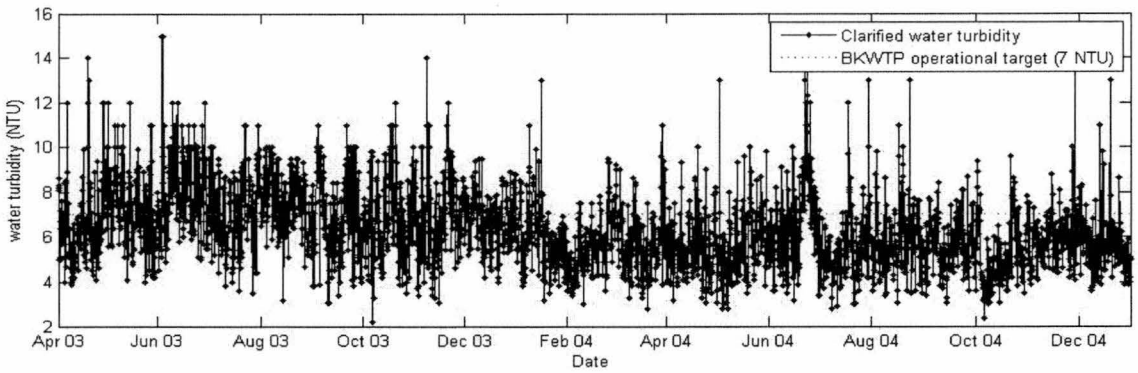


Figure 3.7 Clarified water turbidity (BKWTP)

Most of the water qualities were shown to be in the normal range and suitable for treatment. Their relations to the operational target (i.e. clarified water turbidity) were small implying complicated relations in the clarifier. These variations at BKWTP were not large in comparison with those at BEWTP. This implied that the presence of a reservoir at the beginning of the process helped to buffer any changes. However, one of the key difficulties for the operation was a density current in the clarifier. Existence of this density current may be confirmed by a small correlation of clarified to raw water turbidity. This implied that there were other interferences in addition to the ever changing raw water turbidity. The density current phenomenon was documented as being strongly

linked with water temperature, and it normally occurred in large uncover clarifiers (Hudson, 1981), such as the ones being used at BKWTP. Unfortunately, no thermal sensors had ever been sited in the clarifier to provide further information and proof of this inference.

3.5 Chapter Conclusion

This chapter has presented the statistics of the operational data of both case studies in order to identify the manner of variation of each variable. Measurement errors were also included. The linkages of raw water qualities to operational targets were examined by using the coefficient of correlation (r) to guide the selection of process model inputs in Chapter 5. All the data which had different measurement frequencies were discretised to a four hour time step to use the same time interval which would benefit prediction and concur with how the operator controls the clarifier in real practice

In the BEWTP case study, the operational data analyses revealed that the raw water qualities were excellent and well within the treatable range, with an exception that the raw water colour was moderate and *true colour*. There was some inconsistency between the clarified water colour and turbidity due to its being *true colour*. Most changes in raw water quality could be foreseen as varying seasonally. A significant finding was the large fluctuations in raw water turbidity which was a typical characteristic of water treatment plants without a buffering reservoir.

From the analysis of BKWTP operational data, the variables proved to be widely different to BEWTP's case study because of seasonal changes. Raw water was of poor quality with routinely high turbidity. Nonetheless, the raw water qualities were within treatable ranges and they fluctuated more moderately due to the existence of a buffering reservoir. The coefficient of correlation of clarified to raw water turbidity, being only small, implied that other and extraneous sources might occur in the clarifier, possibly arising from a density current.

In both case studies, all the associated raw water quality, process variables and control variables were shown to correlate weakly to the operational targets (i.e. water turbidity and colour). This reflected the complexity of the clarifier processes, and it confirmed that the artificial neural network should be a suitable modelling tool for the clarifier process.

4.0 ARTIFICIAL NEURAL NETWORK MODELLING: REVIEW OF METHODOLOGY AND APPLICATION TO THE CURRENT RESEARCH

4.1 Introduction

Artificial Neural Networks (ANNs) are information-processing systems that attempt to mimic the operation of human brain and nervous system. They can be employed in a wide variety of applications such as pattern recognition, prediction and process control and clustering (Demuth & Beale, 2000). ANNs possess a remarkable capacity to learn and update themselves from examples of input-output patterns, which are used to determine the rules that direct the relationship between the variables. ANNs are non-linear interpolator of multi-dimensional data are able to *generalise* complex relations among inputs and outputs. Accordingly, ANNs are well suited for modelling complex problems where there is unknown relationships between the variables and when nonlinearity is suspected (Haykin, 1999; Norgaard et al., 2000).

ANNs were first introduced in 1943 (McCulloch & Pitts, 1943) and were put into wide usage in 1986 after the introduction of the backpropagation training algorithm for feed forward ANNs (Rumelhart et al., 1986). The history of neural networks is well documented in detail in several textbooks such as (Negnevitsky, 2005).

Among the data-driven modelling methods, ANNs have shown promise and superiority to other classical modelling methods such as multiple regressions. Its superiority has been proved in a number of comparative studies related to environmental system modelling and water supply and wastewater engineering (Chen et al., 2008; Chowdhury et al., 2009; Dellana & West, 2009; May & Sivakumar, 2009; Perendeci et al., 2008; Sun et al., 2009).

The objective of this chapter is to provide background of the Artificial Neural Network (ANN) modelling methodology and to select suitable modelling methods to use in this thesis. Two types of neural network are used here: (i) Feed forward backpropagation network (backpropagation network in short) and (ii) Self-Organising Map (SOM) network. These will be used for process modelling and data clustering, respectively. Initially, the backgrounds of two networks are introduced. Then, ANN modelling will be

described by focusing on determination of inputs, data division, and data pre-processing and model performance evaluation. At the end of this chapter, all the selected modelling methods will be summarised and used in the next chapter with real operational data from both case studies for clarifier process modelling.

4.2 Classification of Artificial Neural Network

The classification of artificial neural networks can be addressed on the basis of two main criteria: (i) the learning rule used and (ii) the connection between neurons. Based on learning rules, ANNs can be categorised into supervised and unsupervised learning. In a supervised network, the weights are adjusted in the direction that minimises the error between its output and the targets to be learned. Several algorithms concerning supervised learning (e.g. backpropagation) are frequently reported in the modelling of nonlinear systems. Unsupervised training differs from supervised training in that it requires only inputs to train the network. During unsupervised training, the weights are adjusted in such a way that similar inputs produce similar outputs. For instance, unsupervised learning is used in Self-Organising map (SOM) networks to discover clusters of structure in data (Baughman & Liu, 1995; Pham & Liu, 1995; Tsoukalas & Uhrig, 1997)

Based on the connection between neurons, ANNs can be divided into feedforward or feedback network. A feedforward network has only a forward direction in the connections between the neurons. This type of network contains no closed loops and is suitable for pattern recognition or modelling application. In a feed back network, the outputs are connected to its input, allowing the network to recognise temporal behaviour (Demuth & Beale, 2000; Kosko, 1992; Pham & Liu, 1995).

The feedforward backpropagation and SOM networks are used in the clarifier process model and will be introduced later in this chapter.

4.3 Structure and Operation of Feedforward backpropagation Artificial Neural Networks

In this section, feedforward backpropagation ANNs (Backpropagation ANNs in short) are introduced with a discussion of neural network structure and links to how a neural

network learns and generalises the information from sets of data. It is necessary to understand the classic weight updating method and the learning rule called gradient descent algorithm widely used in backpropagation networks.

4.3.1 Feed Forward Artificial Neural Network Structure

ANNs are parallel systems consisting of several neurons. Each neuron in a specific layer is fully or partially connected to other neurons via connection weights. The weight in each connection link corresponds to the strength of the connections between interconnected neurons. A zero weight refers to no connection between two neurons. An ANN's architecture is defined by the number of layers, the number of neurons in each layer, and the type of activation function used by neuron in each layer. All the neurons in a given layer are assumed to use the same transfer function.

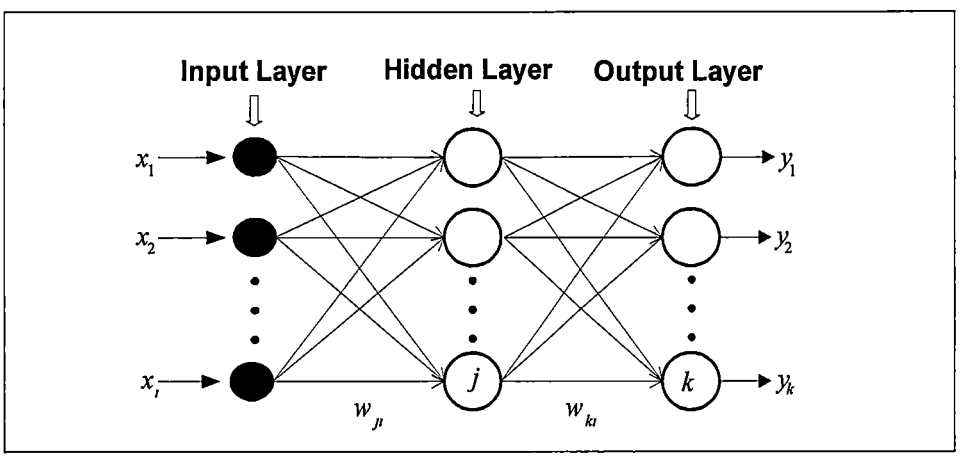


Figure 4.1 A feedforward artificial neural network structure

A feedforward artificial neural network structure is illustrated in Figure 4.1. The n input values $\{x_1 ... x_i\}$ are entered into the neurons in the first layer of the network (the input layer). Thus the number of input neurons, i, is equal to the number of input values. Between the input layer and output layer are layers of neurons called hidden layers, and there is no restriction on the number of layers and neurons in these layers. There are j neurons, $\{h_1 ... h_j\}$, in the single hidden layer, and these neurons take as their input the output from the i neurons of the input layer, plus bias b_i. The network's output layer has k neurons, $\{y_1 ... y_k\}$, where k is the number of required output values. These neurons receive as their input the output from the k neurons of the hidden layer, plus bias b_j. The word "feedforward" shows information flow direction from the input to the output. When

the network has more than one hidden layers, it can be entitled as Multiple-layer feedforward network.

Depending on the objective of the study, the output layer can have one or more neurons that produce the output. Multiple outputs are feasible but they lead to increased network complexity. Thus it is recommended that one separate ANN should be developed for each output to achieve better results (Haykin, 1999).

4.3.2 Artificial Neuron

Fundamental to ANNs are artificial neurons ("neurons" in short) as they make up the processing units for such systems. Based on the basic concepts of biological neurons, artificial neurons accept inputs from other similar neurons process the inputs and send a single output to other neurons. The basic structure of an artificial neuron is shown in Figure 4.2.

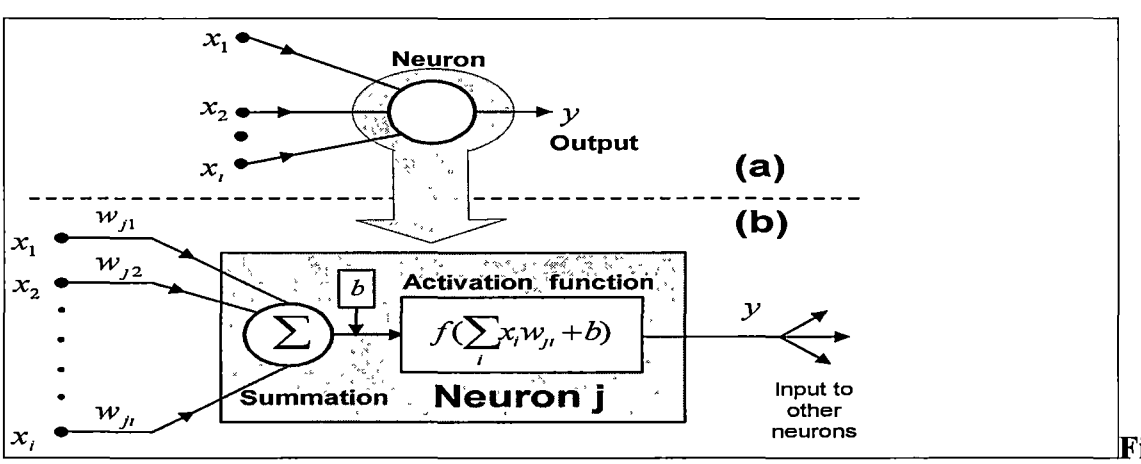


Figure 4.2 an artificial neuron

(a) A single neuron with I input (x_1 to x_i), produces a scalar output y

(b) The neuron *Transfer function* : the summation of input is passed to an activation function including b , the value assigned to any possible bias, which produces the output)

The neuron's processing capability is attributed to three main functions: (i) the input function, (ii) the activation function, and (iii) the output function. These three functions are combined and called the *transfer function.* To provide a better understanding of the transfer function, these three functions should be considered separately. In Figure 4.2, the input function performs an algebraic summation of multiplication of inputs $\{x_1,.....,x_i\}$

with the corresponding connection weights, $\{w_{j1},.....,w_{ji}\}$. The net input function to the j^{th} neuron unit, net_j, can be written as:

$$net_j = \sum_i x_i w_{ji} + b_i \qquad (4.1)$$

where, i index denotes the number of input neurons and b_i is a bias (Freeman & Skapura, 1991).

The output of the input function, net_j then is the necessary input for the activation function. The activation function is a differentiable function that determines the output of those neurons based on the net input function (Masters, 1993). Although the activation function can be any function, it is most often monotonic and a bounded function (Haykin, 1999; Norgaard et al., 2000). The output values for the activation function are generally bounded by a range 0 to 1 or -1 to 1 (depending on the activation function used).

In early neural models, a simple threshold function, or step function was used as the activation function (Figure 4.3(a)). This particular type of activation function allowed a value of 1 to generate from the neuron if the weighted sum of the inputs exceeds some selected threshold, otherwise the output is 0.

More recently, the threshold function has been substituted by a more general nonlinear sigmoid function producing an S-shape curve (Figure 4.3(b)). Generally a Sigmoid function is of real value and differentiable. It actually refers to the special case of the logistic function and has often been reported as a popular alternative (Rumelhart et al., 1986). Besides the logistic function, the hyperbolic tangent function is another popular activated function (Norgaard et al., 2000) (Figure 4.3(c)). Both the activation function and their derivatives can be outlined as follows:

Logistic function and its derivative,

$$f(net) = \frac{1}{1+\exp(-net)} \quad \text{and,} \quad f'(net) = f(net)(1-f(net)) \qquad (4.2)$$

Hyperbolic tangent and its derivative,

$$f(net) = \frac{\exp(net)-\exp(-net)}{\exp(net)+\exp(-net)} \quad \text{and} \quad f'(net) = 1-f(net) \qquad (4.3)$$

These particular functions have an advantage in that their derivatives, which will later be shown, provide significant aspects for neural computation (Section 4.2.5). Since their derivatives are low order polynomials and are easily computed. Choice of activation function is discussed in section 4.3.4.

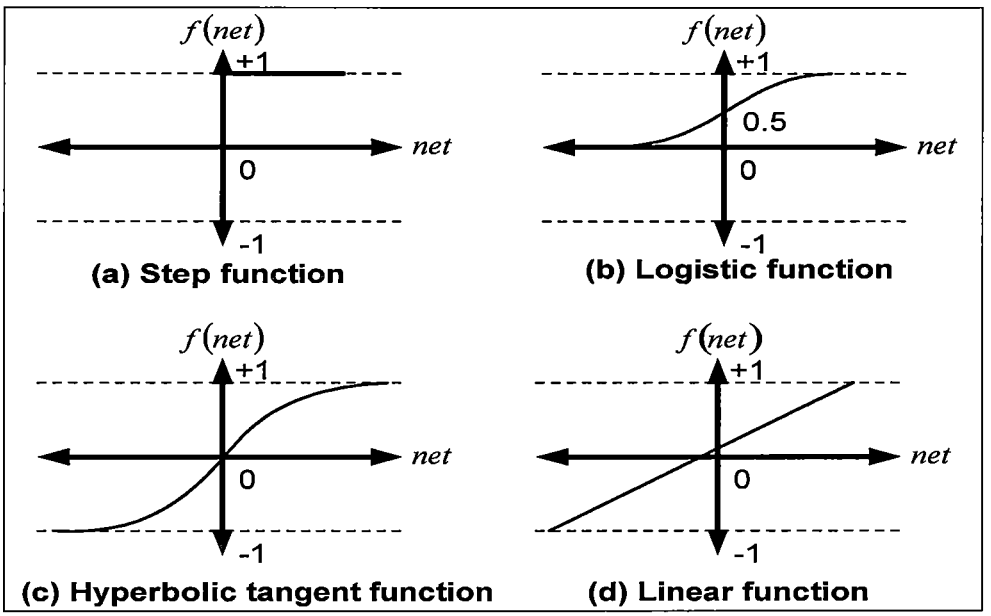


Figure 4.3 Typical Activation functions

The third component of the transfer function is the output function. It is usually chosen to be equivalent to the output of the activation function (i.e. the output of neuron will be the same as that of the activation function). Hence, in this Figure 4.3 the abscissa corresponds to the *net* while the ordinate represents the neuron output. The output function distributes the neural outputs to the other neurons in the next layer (Caudill & Butler, 1990) .

4.3.3 Perceptron and Linear Network

The perceptron is among the simple forms of a neural network. It was developed Frank Rosenblatt in the early 1960s. It is a feedforward network that utilises a step function as an activation function. Each neuron computes the weighted sum of inputs and compares the resulting net weighted input to a threshold value, T . If the net input is larger than or equal to the threshold, the neuron yields an output value of 1.0, otherwise it gives an output value of -1.0.

The transfer function of the perceptron is written as:

$$y = \begin{cases} +1, if\ \_net \geq T \\ -1, if\ \_net \leq T \end{cases} \qquad (4.4)$$

where y = output of the neurons,

net = net weighted input to the neuron and is calculated as follows:

$net = \sum_j w_{ji} x_i$

where w_{ji} = component of weight vector associated with the input i to neuron j and

x_i = component of input vector

The perceptron training rule adjusts neuron weights using the formula:

$$\Delta w_{ji} = w_{ji}^{new} - w_{ji}^{old} = \varepsilon_i x_j \qquad (4.5)$$

where $\varepsilon_j = t_j - y_j$ is the error of actual output t_j and desired output y_j .A perceptron is useful as a classifier. A perceptron network is only capable of training a single layer of neurons, it can only solve linearly separable problem but not for nonlinear separable problem (Demuth & Beale, 2000; Negnevitsky, 2005). One solution to this difficulty is to use a pre-processing method that results in nonlinear separable vectors or multiple perceptron in multiple layers. However this adds more complexity to the networks.

A linear network is another form of neural network that is similar to the perceptron. However its activation function is a linear function (Figure 4.3(d)) rather than a step function (Figure 4.3(a)), allowing its result to take on any value, while the perceptron output is limited to either 0 or 1. It has a similar limitation to the perceptron in that it can only solve linearly separable problem (Demuth & Beale, 2000). Consider a training data set containing n records, with each record consisting of an input vector x_i (the predictor variables), and the correct target response, t_i. The network's response is y_i and the Sum-Square Error (SEE) is:

$$SSE = \sum_{j=1}^{n} \left(y_j - d_j \right)^2 \qquad (4.6)$$

The derivative of the mean SSE with respect to weight w_{ji} is (Zurada, 1992):

$$\frac{\partial SSE}{\partial w_{ji}} = -2\varepsilon_j x_i \qquad (4.7)$$

In training of the network, all such derivatives are calculated, and the neuron weights are adjusted to decrease (and eventually minimise) the mean SSE, the basic adjustment being:

$$\Delta w_{ji} = \eta \varepsilon_j x_i \qquad (4.8)$$

Although this gradient descent approach can be based on different error functions, the most common function is SEE (Bishop, 1995). The rule is often referred to as the *Generalised delta rule*, or the *Widrow-Hoff rule* (Widrow & Hoff, 1960). Its training rule is similar to the perceptron, except for the use of a learning rate, η which is a computation acceleration parameter. The technique is used in the numerical method in this thesis.

4.3.4 Feedforward Backpropagation Network

A Feedforward backpropagation network (backpropagation network in short) is a feedforward network that adjusts the connection weights by using a *backpropagation* algorithm. The word backpropagation is an abbreviation for "backwards propagation of error"(Baughman & Liu, 1995). It is a network that employs the generalised delta rule to govern the learning process. The backpropagation algorithm involves two steps. First is a forward pass, where the effect of the input is passed forward through the different layers of neurons in the network reaching the output layer. Once the error is calculated by comparing network output and desired output, a second step starts backward through the network. The errors at the output layer are propagated back towards the input layer in order to adjust the weight. Standard backpropagation uses a gradient descent algorithm, as does the Generalised delta rule, or the Widrow-Hoff rule (Demuth & Beale, 2000; Hagan et al., 1996). The hidden layer neurons commonly use logistic or hyperbolic tangent activation functions while the output layer neurons usually adopt a linear transfer function (Carter, 1999)

Essentially backpropagation is a gradient descent technique that minimises the network error. Each input pattern passes through the network from the input layer to output layer. After passing the final layer, the actual output of the network is compared with the desired output. The error of each is computed based on (Dayhoff, 1990):

$$E = \frac{1}{2}\sum_{J}(y_J - d_I)^2 \qquad (4.9)$$

where

E = the global error function;

y_J = the prediction output by the network; and

d_J = the desired output.

The global error function is minimised by adjusting the weight using the gradient descent rules:

$$\Delta w_{JI}(n) = -\eta \frac{\partial E}{\partial w_{JI}} \qquad (4.10)$$

where:

Δw_{JI} = weight increment from neuron j to neuron i; and

η = The learning rate, by which size of the step taken along the error surface is determined.

Equation (4.10) can be further defined by the generalised delta rule as follows:

$$\Delta w_{JI} = \eta \delta_J x_I \qquad (4.11)$$

where:

x_I = input from neuron $i, i = 0,1,2...,n$;

δ_J = error value between the predicted and desired output for node j.

If the neuron j is in the output layer, it can be computed by applying the generalised delta rule as follows:

$$\delta_J = (y_J - d_J) f'(net_J), \qquad (4.12)$$

where $f'(net_j)$ is the derivative of the activation function with respect to the weighted sum of inputs of neuron j.

If node j is in the hidden layer, the generalised delta rule can be used as illustrated in the equation (4.13) and Figure 4.4 (Dayhoff, 1990).

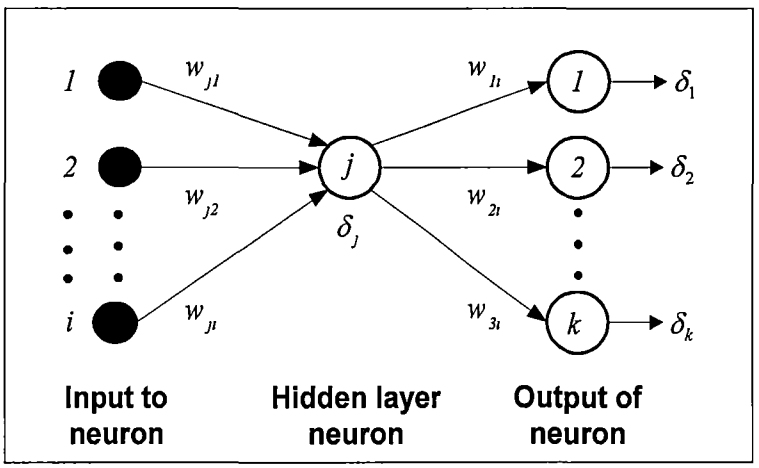


Figure 4.4 Hidden layer neuron used in backpropagation neural network

$$\delta_j = \left[\sum_k \delta_k w_{kj}\right] f'(net_j) \qquad (4.13)$$

The weights are then updated by adding the delta weight, to the corresponding previous weight as follows:

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} \qquad (4.14)$$

where:

$w_{ji}(n)$ = the value of a weight from node i to node j at step n (before adjustment); and

$w_{ji}(n+1)$ = the value of the weight at step $(n+1)$ (after adjustment).

The backpropagation algorithm is sensitive to the initial condition, i.e. the initial value of the weights due to its gradient descent nature. For instance, training may begin with a set of initial weights that are positioned in the flat region of the error surface when convergence becomes especially slow (Hassoun, 1995). Moreover, training may start from an unfavourable position in weight space where the network may be trapped in a local minimum and cannot escape (Maier, H. R. & Dandy, 1998).

The choice of the learning rate is crucial and an optimum learning rate is commonly found by trial-and error. If the learning rate is chosen to be small, convergence will eventually be achieved. However it will be extremely slow. In addition convergence will be subject to the local minimum in the error surface that is closest to the random starting position. On the other hand, convergence may sometimes not be attained using a large learning rate (Haykin, 1999).

To solve this problem without leading to oscillation, Rumelhart et al.(1986) described a process that adds a momentum term, μ to weight adjustment that is proportional to the amount of the previous weight change. After an adjustment is performed, it is saved and used to change all subsequent weight adjustments. This implies that the weight change of the current step should carry some momentum of weight change from the previous step. The modified adjustment equation is as follows:

$$\Delta w_{ji}(n+1) = -\eta \frac{\partial E}{\partial w_{ji}} + \mu \Delta w_{ji}(n); \qquad (4.15)$$

and

$$w_{ji}(n+1) = w_{ij}(n) + \Delta w_{ji}(n+1) \qquad (4.16)$$

Several other algorithms are available for training the networks as described by Hertz et. al.(1991). A majority of these algorithms are founded upon the assumption that the learning rate is constant from one epoch to the next and from one weight to another. Their learning rates and momentums are found by a trial-error approach. This is time consuming and has no guarantee for trapping in local minimum. Both the magnitude of the learning rate and momentum are in the range of 0 to 1. If the learning rate is small, the weight will be modified by smaller increments with less oscillation. When the learning rate is large the weight will be drastically changed and often causes oscillation. In general, a small learning rate is well adapted for use with noisy data. (Garret et al., 1992). Besides use in alleviating of local minimum trapping, momentum is a smooth factor that permits faster learning without oscillation by making the weight change by a function of the final step weight change (Garret et al., 1992; Hagan et al., 1996)..

However some researchers (Bishop, 1995; Demuth & Beale, 2000; Kosko, 1992; Ripley, 1996) have challenged the above assumptions by proposing a learning rule that uses a varying learning rate and momentum. Demuth and Beale (2000) recommended an adaptive learning rate back propagation with momentum that can converge from ten to a hundred times faster than a standard gradient descent with momentum back propagation. In addition, no time consuming trial-error processes are necessary for finding the optimal learning rate and momentum, if this training algorithm is used. Therefore variable learning rate backpropagation will be used throughout this research.

4.4 Self-Organising map Network

Self-Organising Maps (SOM) belong to the group of unsupervised neural networks and were proposed and developed by Kohonen (1982). Unsupervised neural networks are normally used for data clustering to optimise and identify similarities associated with raw data. SOMs will be used in the process model development to cluster the data variables in order to divide data into training, testing and validation sets. The typical structure of SOMs consists of two layers: an input layer and a Kohonen layer. In the Kohonen layer, a number of competitive neurons are arranged in a one or two dimensional array as shown in Figure 4.5 (a).

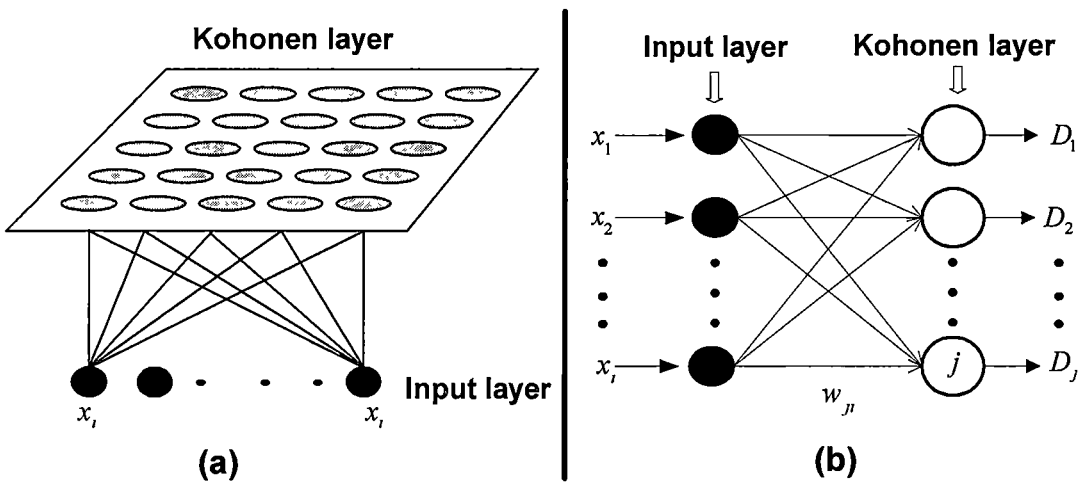


Figure 4.5 Self-Organising Map network

(a) The topology of Self-Organising map network

(b) Unsupervised learning Self-Organising map network

In Figure 4.5(b), the input from each neuron in the input layer (x_i for $i = 1,2,....n$) is fully connected to the Kohonen layer through connection weights (w_{ij} for $j = 1,2,....m$). These weights are randomly initialised at the start of the self-organising process, t. At each

neuron in the Kohonen layer, the input (x_i) is present without proving the desired output, and a matching value is computed. This value is typically the *Euclidean* distance. For node j in the Kohonen layer, the Euclidean distance (D_j) between the weights and the corresponding input values, is expressed by:

$$D_j = \sum_{i=1}^{n} (w_{ji} - x_i)^2, j = 1,2,....., m \qquad (4.17)$$

The neuron that has the minimum Euclidean value is declared the "winner". In other words, the winner is the neuron whose weight is most similar to the input value. Subsequently the weight of the winner neuron and its neighbouring neurons, in terms of topology, are updated to get a closer match with the input values. The incremental weight update for neuron j is as follows:

$$\Delta w_{ji} = \eta(x_i - w_{ji}) \qquad (4.18)$$

where:

η = learning rate.

At step n of training, node j can be updated as in equation (4.19):

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} \qquad (4.19)$$

The process is repeated by successively inputting new data records to the model and modifying the connection weights until they stay unchanged. This results in a topological map where similar data records are clustering together. For a detailed description of the operation of SOMs see (Kohonen, 1997).

4.5 Artificial Neural Network modelling

To improve performance, ANN models need to be developed in a systematic manner. In such an approach it is important to address major factors such as the determination of adequate model inputs, data division and pre-processing. The choice of a suitable network to control the optimisation method, the stopping criteria and model performance evaluation are also essential.

In the process of ANN modelling, a general practice is to classify the available data into three sets: training, validation and testing. The training set is used for the purpose of adjusting the connection weights. The validation set is used to check the performance of the model at various training stages, and to decide when to stop training to avoid overfitting. Finally the test set is used to estimate the performance of the trained network in the environment in which it is to be deployed (Demuth & Beale, 2000).

4.5.1 Determination of Model Inputs

In developing ANN models an important step is to select the input variables with the most significant impact on model performance. Especially in the case of modelling of temporal system, the number of the selective inputs is perhaps the most important parameter since it corresponds to the number of lagged observations used to discover the underlying pattern in temporal data (Zhang, G. et al., 2001).

A good subset of input variables will substantially improve model performance. If ANN models are presented with a number of input variables, the network size is enlarged and results in increased model complexity, computational time and reduced efficiency (Lachtermacher & Fuller, 1994). Even though ANNs do not themselves consider the physical properties of their inputs, a clear understanding of the system to be modelled is an important prerequisite for successful application. In addition, to minimise the training samples and training time, only the factors that have strong influence on the specific problem should be considered for the input neurons(Masters, 1993; Zhang, G. et al., 1998).

A common approach to input determination is to employ an *expert knowledge* of the system to select a set of candidate inputs for the model. Although an expert identification is widely adopted in several ANN applications, dependence upon an expert's knowledge is both subjective and problematic. It becomes very important to have a good level of understanding of the system being modelled in order to select appropriate model inputs. If important inputs were to be neglected, some system information might be lost. On the other hand, if spurious inputs are included, the training process may be confused and delayed.

Generally the potential number of input variables to an ANN model tends to be large, and complete expert knowledge is not available to suggest which variables should be

included. Intuitively the preferred approach to this problem involves a combination of available expert knowledge and a statistical analytical approach such as *correlation* and *autocorrelation function (ACF)* (Maier, H.R. & Dandy, 2000; Qi & Zhang, 2001).

For example, in system modelling using multivariate time series it is often possible to define the appropriate kind of variables for model inputs by using expert knowledge. However, one must question how far the selected data should span into the past (i.e. what time lag of time series data is appropriate). Therefore an analytical method such as autocorrelation analysis may become a more suitable approach for this task.

The autocorrelation function (ACF) is the popular statistic tool for measuring correlation over a time series data. An ACF physically measures if present values in the time series have some relation to earlier values. Initially autocorrelation is used to define a significant time lag for system with the assumption of the system being linear (Box & Jenkins, 1976). For nonlinear system modelling, although there are a number of arguments, a number of researchers (Filho et al., 2006; Haykin, 1999; Ikonen & Najim, 2002; Saad et al., 1999; Venema et al., 1996) use ACF as a guideline for the time lag.

Let a time series data $x_1,.....x_n$. The autocorrelation coefficient at lag k is given as:

$$C_k = \frac{1}{N-k}\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x}) \qquad (4.20)$$

where k is lag, $k = 0,1...,N$. If is set as zero that means the present value and $\bar{x}$ is the overall mean of the series, defined as:

$$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t . \qquad (4.21)$$

The autocorrelation function at lag k, $ACF(k)$ is given by:

$$ACF(k) = \frac{C_k}{C_0}, \qquad (4.22)$$

where C_0 is the autocorrelation coefficient when the time lag is set as zero, (present value). The suitable time lag is defined as the smallest k that makes $ACF(k) < K$ where is usually taken as 0.5 or $\frac{1}{e}$, e=2.71828 (Filho et al., 2006; Venema et al., 1996).

Defining the suitable time lag for a nonlinear system modelling is problematic. A trial-error approach by varying the time lag and considering model performance is recommended. The time lag suggested by autocorrelation is set as the upper boundary for the trial-error procedure, since time steps beyond this threshold are less significant to the present information.

In this thesis, in addition to expert knowledge, statistical approaches (i.e. correlation and autocorrelation function) are used with the multivariable time series data in two case studies to guide the input selection both of selection of the significant inputs for the model and of suitable time lag.

4.5.2 Data Division

Like conventional statistical models, ANNs' model parameters (e.g. connection weights) are modified in the training model so as to minimise the error between model output and target in the training set. ANNs exhibit best performance when they do not extrapolate beyond the range of data used for training (Flood & Kartam, 1994; Tokar & Johnson, 1999). In order to develop the best ANN model for the available data, all the patterns contained in the data set need to be included in the training set (Masters, 1993).

For instance, if available data contain extreme data points that were excluded from the training set, the model is unlikely to perform well as the test data will test the model's extrapolation ability, and not its interpolation ability. If all of the patterns that appeared in the available data are included in the training set, the most rigorous evaluation of the generalisation ability of the model occurs if all the patterns (and not just a subset) are contained in the test set.

In order to avoid model overfitting, Stone (1974) suggested a approach which he called *cross-validation.* Here the data are divided into three sets: training, validation, and testing. The training and testing set are used for model training and evaluation,

respectively. The validation set guards the model from overfitting by stop training. However, no guidelines exist in the literature for optimising which portion of the data should be used for training, validation and testing sets in the attempt to determine optimal proportions of division of the data

Cross-validation may be employed as the stopping criterion. The results obtained from the validation set have to be representative of those obtained from the training set as the validation set is used to decide when to stop training. Consequently, it is important that the statistical properties (i.e. mean and standard deviation) of the various data subsets (i.e. training, validation and testing) are similar to ensure that each subset represents the same statistical population. If this is not the case, it may be difficult to judge the validity of the ANN model (Maier, H. R. & Dandy, 1998; Masters, 1993).

However, it was not until recently that systematic approaches for data division have been proposed in the literature. Before dividing the data, two main questions need to be answered; first, *"what portion"* of the data should be used for each of the training, validation and testing set. Second, *"which data"* should be included in each of the sets. Again the statistical equivalence among these three subsets must not be compromised (Bowden et al., 2002).

Bowden et al.(2002) adopted a genetic algorithm to minimise the objective function which was set as the difference between the mean and standard deviations of the data in the training, validation and testing sets. While this approach ensures that the statistical properties of the various data subsets are similar, there is still a need to choose what proportion of data are suitable for training, validation and testing. In this approach, the question of *"which data"* was answered but no answer for *"what portion"*

Kocjancic and Zupan (2000) and Bowden et al (2002) used a Self-Organising map network (SOM) to cluster high-dimension input and output data in two dimensional space and divided the available data so that values from each cluster are represented in the various data subsets. This ensures that data in different subsets are representative of each other. It also has the additional advantage that there is no need to decide what percentage of the data should be used for training, validation and testing.

- **Data division using Self-Organising map network**

SOM is used to cluster the input and output data into training, validation and testing sets that have similar statistical properties. It can be implemented using software package MATLAB (version 6) with the Neural Network Toolbox (version 4), both from Mathworks INC, Natick, MA, U.S.A. To cluster the data, the inputs and their corresponding output are presented to the network as SOM's inputs. The output of the SOM is obtained using a selective grid topology (e.g. rectangular grid), which displays a representation of the neurons that are winning each pattern. For simplicity, default parameters are used such as *Kohonen rule* for the learning rule, Euclidean distance for neighbourhood distance and the *Rectangular grid* for grid topology. The associated details are well documented (Demuth & Beale, 2000).

Each individual cell in the grid represents a neuron in the Kohonen layer (Figure 4.6). At present there is no theoretical principle for determining the optimum size of the Kohonen layer. The Kohonen layer was kept large enough to ensure that the maximum numbers of clusters were formed from the training data. It has been suggested that the optimal number of epochs and optimal size of the Kohonen layer (number of neurons) for obtaining the maximum number of clusters are data specific and found by trial and error process. In particular, the size of the Kohonen layer is initially set, and then the number of epochs is increased until the number of clusters becomes unchanged. (Of course, the selected size of the Kohonen layer is larger than the mature number of clusters).

According to the original work of Bowden et al (2002), when the maximum number of clusters is formed, and if any cluster contains three data or more, the data should be randomly chosen, and assigned one for each of the training, validation and testing sets. If two data are in a cluster, then one datum is randomly selected for the training set and the other is placed in the validation set. In the instance that a cluster only contains one data, then this data is placed in the training set.

However, Bowden' approach is modified to assure that the training set covers as many varied patterns as possible from the data in this research. When the cluster contains more than three data, two of them are randomly selected and assigned for each of the validation and the test sets, the rest are included in the training set. This modified Bowden method will be used throughout this research.

In Figure 4.6, each square grid denotes a neuron in the Kohonen layer and the black dots represent the data contained in each cluster. In this example, for the cluster A containing only two data, all of them are assigned to the training set. For the cluster B, two records are randomly sampled from the cluster, one datum for the validation set and the one for the testing set. The rest of the data in this cluster are added to the training set. This process is repeated for each cluster in the Kohonen layer.

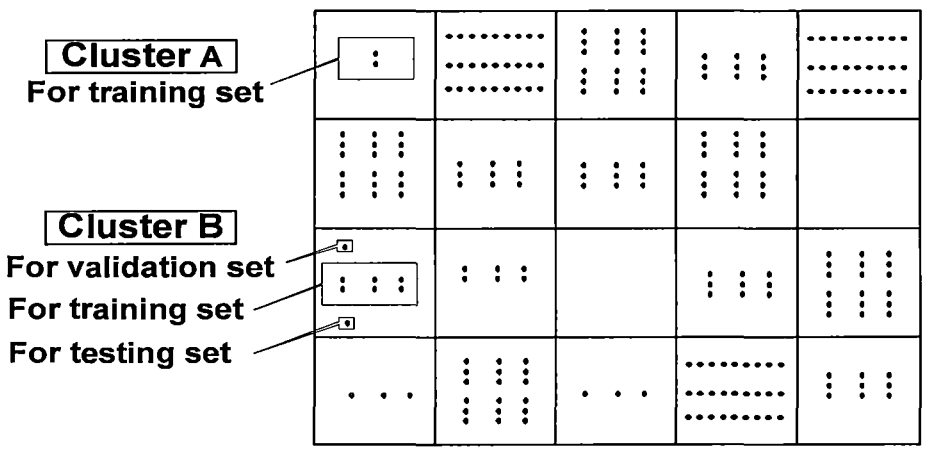


Figure 4.6 illustrates the sampling process used in the SOM data division method.

4.5.3 Data Pre-processing

After the available data have been divided into their subsets (i.e. training set, validation set and testing set), it is important to pre-process the data in order to have the same magnitude before they are fed to the ANN. One of most popular pre-processing techniques is normalisation. ANNs require that all inputs and outputs are normalised to have the same order of magnitude in order to ensure that all variables receive equal attention during the training process. If the input and output variables are not of the same order of magnitude, some variables may appear to have more significance than they actually do. The training algorithm has to compensate for order of magnitude differences by adjusting the connection weights. This is a very ineffective way of training algorithms (Baughman & Liu, 1995).

Additionally, inputs need to be commensurate and fall within the limits of the activation function. This is between -1.0 to 1.0, if using a hyperbolic tangent function and 0.0 to 1.0 for the sigmoid function (Masters, 1993).

Baughman and Liu (1995) recommended normalizing all input and output variables and introduced a comparative review of three normalisation methods as follows.

Method 1: normalise each variable, x_i, in the data set to between 0 and 1 by dividing its value by a selected upper limit of that variable, $x_{i,upper}$. The normalised variable, $x_{i,norm}$ is calculated from:

$$x_{i,norm} = \frac{x_i}{x_{i,upper}}. \qquad (4.23)$$

For example, variable x_i has a range of 500 and 900 and the normalisation factor is chosen as $x_{i,\max} = 1000$. Using the equation (4.23), the normalised variable values will be 0.5 and 0.9. One limitation of this method is that it does not utilise the entire range of the transfer function. Figure 4.7 (a) shows that only a small portion of the transfer function corresponds to x_i values of 0.5 to 0.9. The connection weights can be broadened and their range can be shifted to include a larger region of the transfer function. However, as the number of variables and connection weights increase, these adjustments become more difficult for training algorithms.

Method 2: Expand the normalisation range so that the minimum value of the normalised variable, $x_{i,norm}$ is set to 0 and the maximum value, $x_{i,norm}$ is set at one. We define the normalised variable $x_{i,norm}$ by using the minimum and maximum values of the original variable, $x_{i,\min}$ and $x_{i,\max}$ respectively.

$$x_{i,norm} = \frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}} \qquad (4.24)$$

This method offers a significant improvement on the Method 1 by using the entire range of the transfer function, as Figure 4.7 (b) illustrates. Moreover, with this method every input variable in the data set has a similar distribution range, a feature which improves the efficiency of training.

Method 3: Normalises data set between limits of -1 and +1 with the mean value set to zero. This technique is named *zero-mean normalisation.* The normalisation variable, $x_{i,norm}$ is obtained by:

$$x_{i,norm} = \frac{x_i - x_{i,mean}}{R_{i,\max}}, \tag{4.25}$$

and

$$R_{i,\max} = Maximum \quad [(x_i - x_{i,mean}), (x_{i,mean} - x_i)] \tag{4.26}$$

where x_i is an input or output variable, $x_{i,mean}$ is the mean value of the variable over the data set. $R_{i,\max}$ is the maximum range between the average value and either the minimum or the maximum value whereby $x_{i,\max}$ and $x_{i,\min}$ are the maximum and minimum value respectively.

As in the Method 2, the zero-mean normalisation method utilises the entire range of transfer function, and every input variable in the data set has a similar distribution range (Figure 4.7(c)). This allows the weight connections to follow a more standard distribution, without requiring a shift and broadening of the input variables to match their respective output variables. This method gives meaning to the value of the normalisation variable; 0 represents the normal state (or average) of the variable;-1 represents the lowest possible level of the variable, and +1 represents the highest possible level of variable. The normalised range of -1 to +1 is well matched with a hyperbolic tangent function which is itself superior to a sigmoid function (section 4.3.2).

In addition, by setting all the normal states of variables to zero, the network will have a standard structure that makes the training both efficient and consistent from one problem to the next. That is, all networks should normally predict output responses of approximately 0 (normal value) whilst the input variables are set at their normal values. Therefore, the network is essentially only training deviations in the output variables which are due to various deviations in the input variables. The zero-mean normalisation method is therefore used throughout this research.

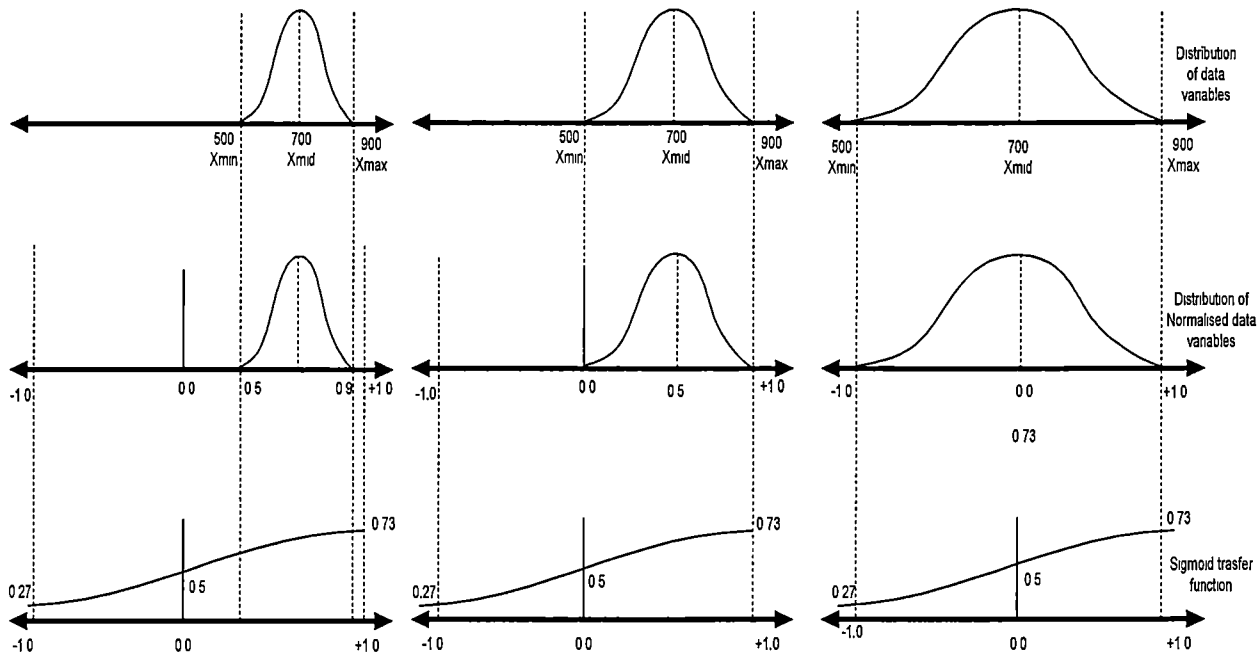


Figure 4.7 Input normalisation methods

4.5.4 Selection of Proper Activation Function

There are some heuristic rules for the selection of the activation function. For example, Kilmasauskas (1988) suggested use of a logistic activation function for classification problems which involve learning about average behaviour, and use of hyperbolic tangent function if the problem involves learning about deviations from the average such as in a forecasting problem.

Based on the function of the network being used, a transfer function also needs to be chosen. The hyperbolic tangent and sigmoid function are appropriate for most types of networks, especially for predictive problems. A Gaussian function is only recommended for classification networks. Baughman and Liu (1995) demonstrated that the hyperbolic tangent function performs better than the sigmoid function for the following reasons:

1. With the output varying from -1 to +1 for the hyperbolic tangent and only 0 to 1 for the sigmoid function means that the hyperbolic tangent function has both a negative input value and a positive response for a positive input value, while the sigmoid function is limited by always having a positive response.

2. The slope of the hyperbolic tangent is much greater than the slope of the sigmoid function. This has the effect that the hyperbolic tangent function is more sensitive to small changes of input.

Consequently, the hyperbolic tangent will be used as an activation function for every neuron in the hidden layers and the linear activation function will be employed for every neuron in the output layer in this research.

4.5.5 Determination of Model Architecture

One of the most important and difficult tasks in ANN model development is determining the network architecture. The use of a large number of hidden layers and neurons may be counterproductive because an excessive number of parameters will encourage overfitting of the network solution to the training data, and so reduce the generalisation capabilities of the final product. The selection of a suitable architecture is problematic. Therefore, both empirical methods or trial and error process must be used to determine the optimal network architecture (Maier, H.R. & Dandy, 2000; Masters, 1993).

The number of neurons in the input can be chosen using expert knowledge and statistic approaches (section 4.3.1) and the number of neurons in the output layer is obviously fixed according by the number of outputs. Selection of the optimal number of hidden layers and neurons is, however, a different matter. In practice, the optimum number of hidden layers and neurons in each layer can be empirically achieved by fixing the number of layers and optimising the number of neurons in each layer by trial and error.

One hidden layer only is recommended by several researchers (Hecht-Nielsen, 1989; Masters, 1993) because the use of more than one will increase the number of parameters to be estimated, which may slow down the training process without substantially improving the efficiency of the network. Theoretical results have shown that a one hidden layer feedforward network with an appropriate activation function is capable of approximating any measurable function to a reasonable degree of accuracy even if some errors occur. Errors will be due to either inadequate learning or too few hidden neurons. Also the data may not reveal sufficiently deterministic or causal relationships. However,(Hagan et al., 1996; Maren et al., 1990; Tarassenko, 1998; Weigend et al., 1990) all suggest that using a two hidden layers network is sufficient to model continuous

functions. Therefore, in this research, a two hidden layer network is considered the largest network practicable for model optimisation.

The number of neurons in the input and output layers are fixed by the number of inputs and outputs respectively. It is a common practice to firstly fix the number of hidden layers in the network and then secondly to choose the number of neurons in each of these layers. The usual method of selecting the number of hidden neurons is a rule of thumb approach. There are few relevant studies but, unfortunately these guidelines are for one hidden layer network. Hecht-Nielsen (1989) proposed the upper limit for the number of hidden layer neurons:

$$N^H \leq 2N^I + 1 \tag{4.27}$$

where N^H is the number of hidden layer neurons and N^I is the number of inputs. Fletcher and Goss (1993) proposed a number of hidden layer nodes as follows:

$$\left[2(N^I)^{0.5} + N^0\right] \leq N^H \leq \left(2N^I + 1\right) \tag{4.28}$$

where N^O is the number of outputs. However, Rogers and Dowla (1994) proposed the following relationship to avoid overfitting the training data by adding the term number of training samples defined as N^{TR} :

$$N^H \leq \frac{N^{TR}}{N^I + 1}. \tag{4.29}$$

For a network with two hidden layers, a 3:1 ratio of neurons between the two layers is suggested (Maier and Dandy 1996). In this thesis, the maximum number of hidden layers is set to two in order to avoid model complexity. The numbers of neurons in each hidden layer are empirically determined (i.e. trial and error process) by choosing those with the best predictive performance for network. However, it is necessary to realise that the number of neurons in the hidden layers (during trial and error process) increases the amounts of connections and weights to be fitted. This number cannot be increased without limit because one may reach a situation where the number of weights to be fitted is larger than the number of data sets available for training. Though the neural network

can still be trained, the case is mathematically undetermined. Mathematically it is not possible to determine more fitting weights than the available training data sets (Sha, 2007). For example, two data sets are required as a minimum for linear regression, three data sets for second order polynomial (parabolic) regression and so on. In practice, for reliable regression, much more data than the minimum amounts are used to increase statistical significance. For example, if only two data sets are used to determine a slope through linear regression, the standard error of the slope calculated will be infinite (infinitely large). A slope determined through two data sets has no statistical significance (Harnett & Murphy, 1975; Mendenhall & Beaver, 1994).

4.5.6 Choice of Stopping Criteria

To decide when to stop the training process, stopping criteria are used. If a lengthy training period is chosen, there is the possibility that the network will try to fit noise in the data. However assuming that the noise is small relative to the main features, it is only in the later stage that the network will attempt to fit noise. Stopping the training before the noise is learned will improve the generalisation (Weigend et al., 1990)

Stopping criteria determine whether the model has been optimally or poorly trained. Several approaches can be used to determine when to stop training. Training can be stopped after the presentation of a finite number of epochs; a maximum number of 2000 epochs has been suggested (DeSilets et al., 1992) despite the fact that this number should be experimentally determined. Training could be stopped when the training error reaches a sufficiently small value; or when no or slight changes in the training error occur. However, in these examples, the stopping criteria may lead to overfitting.

The *cross-validation* technique (Stone, 1974) is an approach that can be used to overcome such a problem. It is considered to be the most valuable tool to ensure that overfitting does not occur (Smith, 1993). The Cross-validation technique is concerned with avoiding overfitting by the use of the validation set. The validation set measures the ability of the model to generalise, and the performance of the model using the particular set is checked at many stages of the training process. Training is stopped when the error of the validation set starts to increase. The Cross-validation technique was adapted and used in Neural Network Toolbox (version 4) which referred to it as the *Early Stopping* technique. This validation set has to be monitored during the training process. The

validation error will normally decrease at the beginning phase of the training. The validation error will then increase as the network begins to overfit the data. When the validation error increases over a pre-set number of epochs, the training process will be stopped. The weight and the bias of the minimum validation error are returned (Demuth & Beale, 2000). The Early Stopping technique is employed for model development throughout this research.

4.5.7 Neural Network Model Performance Evaluation

In order to evaluate the predictive performance of a neural network model, a variety of qualitative and quantitative means can be employed. Selective testing sets are utilised as benchmarks for performance comparison. A visual inspection of a plot of actual outputs versus predicted outputs can serve as a qualitative measure of the predictive performance of a model (Flood & Kartam, 1994). Quantitatively, the predictive performance of a model can be assessed using statistical indicators such as r^2, the coefficient of determination, as well as the mean absolute error (MAE) (Demuth & Beale, 2000). These two statistical indicators will be used for model performance evaluation throughout this research.

The r^2 calculation compares the accuracy of the ANN model to the accuracy of a trivial benchmark model in which the prediction is just the mean of all the samples. A perfect fit would result in a r^2 value of 1. A very good fit would be near 1, and a very poor fit near 0. The equation of r^2 is given as follows:

$$r^2 = 1 - \frac{SS_r}{SS_T}, \qquad (4.30)$$

where, SS_r is residual sum square and SS_T is the total sum square, which are denoted as:

$$SS_r = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (4.31)$$

and,

$$SS_T = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2 \qquad (4.32)$$

where, y_i is the observed value, $\hat{y}_i$ is the predicted vale for y_i and $\bar{y}$ is the mean of the observation. The r^2 calculation is applied to the entire test set data and serves as a bulk indicator to measure candidate model performance.

The second statistical indicator, the Mean Absolute Error (MAE), is used to highlight inconsistencies in ANN model prediction and can also be used to determine whether the model prediction are adequate for process control. The MAE has dimension and units. Thus it is easy to compare with any particular thresholds such as the operational targets of the clarified turbidity and colour in a water treatment plant model. The calculation of MAE is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i), \qquad (4.33)$$

where , y_i is the observed value, $\hat{y}_i$ is the predicted vale for y_i.

4.6 Chapter Conclusion

This chapter has reviewed the artificial neural network as an advanced modelled technique. It has demonstrated that ANNs are a form of artificial intelligence, which, by means of their architecture, attempt to mimic the biological structure of the human brain and nervous system. It is evident from this chapter that in using ANN modelling for prediction, the purpose of the model is to capture the patterns between a set of model inputs and the corresponding outputs. This involves pattern recognition.

To achieve this, ANNs rely on the existing data alone to determine model structure and parameters. Any development of ANN models must address several factors which include the determination of adequate model inputs, data division and pre-processing, the choice of suitable network architecture, careful selection of model performance evaluation, the stopping criteria and model validation. The ANN modelling methods selected for this thesis are listed in Table 4.1 and they will be used to model the clarifier process model in Chapter 5.

Model element	Selected ANN methods
Type of network	Feedforward network with Backpropagation learning
Learning rule	Adaptive learning rate with momentum
Inputs selection method: (i) Type of variables (ii) Suitable temporal span	 (i) Expert knowledge & Coefficient of correlation (r) (ii) Autocorrelation function (ACF)
Data division method	Data clustering based SOM
Data pre-processing	zero-mean normalisation
Training stopping criteria & Model validation	Early stopping
Determination of model architecture	Trial-error process
Model performance evaluation	Coefficient of determination (r^2) and Mean absolute error (MAE)

Table 4.1 the selected ANN methods for the clarifier model development.

5.0 CLARIFIER ARTIFICIAL NEURAL NETWORK MODEL :DEVELOPMENT AND EVALUATION

5.1 Introduction

Artificial Neural Network (ANN) model falls into two different types: *process models* and *inverse process models*. In the former the model predicts the value of one or more process outputs given the values of the process input variables. For the clarifier modelling, this type of the model predicts clarified water qualities using the influent water quality and control action variables as inputs for the ANN model. In the latter, the ANN model is often used to predict the value of the control actions required to give a desired output.

This chapter describes the development and application of the ANN process models to prediction problems. Neural networks were developed for process models to predict clarified water qualities. Two case studies were considered, BEWTP and BKWTP. Past operational data taken from both case studies were used as a source of data for the predictive water quality models. For the BEWTP case study, the operational targets were the clarified water colour and turbidity. Model were developed to predict each target. Targets were set individually as outputs of the associated process models. These models were separately developed and evaluated. On the other hand, for BKWTP, only clarified water turbidity process models were developed since the sole operational target was clarified turbidity. In common for both case studies, the clarifier process models were temporal systems which were designed to cope with the changes of the raw water qualities varying with time. The past operational data were discretised and input to the models.

This chapter begins with description of the process model development processes used in common for both case studies. This includes systematic modelling methods concerning data manipulation, selection of the inputs, data clustering and ANN model architecture optimisation to ensure the prediction performance of the model. The process model performances and reliability will be separately evaluated at the end of the chapter.

5.2 Clarifier Process Model Development

This section describes the general stage in the development of the ANN clarifier process models that predict the clarified water qualities. All the ANN process models described herein were developed using the commercial software package MATLAB (version 6) with the Neural Network Toolbox (version 4), both from Mathworks INC, Natick, MA, U.S.A. Backpropagation networks using adaptive learning rate were employed to predict the clarified water qualities and self-organised map (SOM) neural networks were used to cluster the data during at the pre-process stage. The development of the models consisted of three stages: the pre-processing stage of data collection and manipulation, the ANN process modelling stage and the post-process stage relating to model performance evaluation. The systematic modelling procedure is graphically shown in Figure 5.1

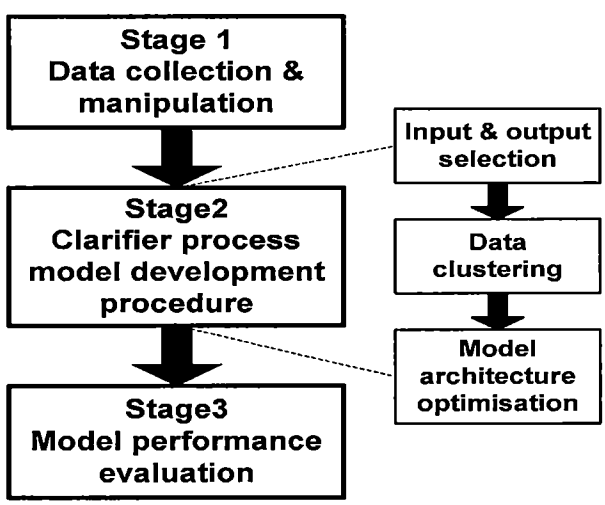


Figure 5.1 The modelling development process, illustrating that modelling development consists of three stages: (i) data collection and manipulation, (ii) clarifier process model development procedure and (iii) model performance evolution.

Stage 1 Data collection and manipulation

The first stage was concerned with data collection and manipulation. Past operational data were collected from BEWTP and BKWTP. To ensure quality of data, the irrational data were filtered out and substituted by interpolated data. The past operational data with different measurement frequency were discretised to four hours time step.

In real practice human operators have been recommended to select the control action set at least every four hours (City Water Technology, 2001; MWA, 2000). This period is approximately equal to the time that the water spend in the WTP (hydraulic retention time). This frequency of control actions ensures that faults can be detected and corrective actions taken place in time (Kerri, 1996). The process models were to work in line with the operator therefore time step data was set equal to how the operators control their clarifier. It is not recommend to use a model which is developed from data with one measurement frequency but predicts at a different frequency (Riyaz et al., 2004). For example, if the ANN process model was developed from daily average data, it should not be used in prediction for a frequency of every four hours. After this state, all the erroneous data were replaced by interpolated ones and formatted as temporal data as if with four hours' time step. All of these works were completely done in Chapter 3

Stage 2 Clarifier model development procedures.

The second stage was the most important as it concerned how to systematically develop the clarifier process model using the existing temporal operation data. The main objective of this stage was to optimise the architecture of the process model to achieve the best predictive ability of clarified water quality in one time step (four hours) into the future. This stage was composed of three steps as follows.

- **Input and output selection**

The initial step in model development was to select input and output variables. The operational targets (i.e. clarified water qualities) were individually set as the output of the clarifier process model. Even though ANNs are capable of multiple output modelling, better performance is obtained by using a single output as this results in reduced model complexity (Haykin, 1999). Therefore a single model for each output was used in this thesis.

The inputs selected were those variables which had or were thought to have an effect on clarified water qualities. Two main subjects were of concern: (i) what type of data was selected and (ii) temporal span. The selection type of the input variables (e.g. raw water pH, alum dosage) can be obtained by using expert knowledge and making use of the correlation coefficient (r). By temporal span we meant how long to backdate data from the present (i.e. the suitable temporal span of each selected input variable). The

autocorrelation function (ACF) was utilised to suggest the optimal temporal data (section 4.3.1).

- **Data clustering**

After the inputs were selected (types and temporal span were defined), the self–organised map (SOM) approach was used to cluster the data (section 4.3.2) to ensure that the training set, validation and testing set were statistically equivalent. To verify this, the T-test examined the null hypothesis of no difference in the means of two data sets and the F-test examined the null hypothesis of no difference in the standard deviations of two sets. Kreyszig et al.(2006) gave a good description of these tests and recommended the selective level of significance of 0.05. This meant that the training, validation and testing sets achieved statistical consistency at the confidence level of 95 %.

The ANN models worked well for interpolation (Masters, 1993). The process model should be tested and validated in the interpolation mode which meant that the range of the training set should cover the range of testing and validation set for each variable. It was necessary to check whether any variables of the test and validation sets were out of range of that of the training set. If the maximum value of any test or validation set variable was greater than the maximum training set value or the minimum test or validation set value was less than the equivalent training set value, then the null hypothesis was rejected.

- **Model architecture optimisation**

To optimise model architecture, the optimum number of hidden layers and hidden neurons was investigated. The best architecture networks were selected as those with the best ability to predict the clarified water qualities one time step of four hours ahead of present time. These networks gave the lowest mean absolute error (MAE) and the highest value of coefficient of determination (r^2) in predicted versus actual output. Optimising the model architecture was conducted by systematic trial and error by varying the number of hidden layers and hidden neurons. However, it was unnecessary to have a network with more than 2 hidden layers (section 4.3.5). All of these trial and error processes were conducted on the test set because it was used as the benchmark and was statistically representative of the training set.

Other model architectural elements have already been discussed (Table 4.1). The conclusions were that: the tangential hyperbolic and linear activated functions were

selected for hidden and output neurons respectively. The *zero mean* method was used to normalise these data. *Early stopping* was set as the stopping criterion for training in order to avoid overfitting. Training was stopped after any continuous five epochs for which the MAEs of the validation set were more than those of training set. This process obtained the clarifier process model with optimised architecture to best predict one step ahead performance.

Stage 3 Model performance evaluations

The performance evaluation of the optimised architecture process model was conducted in two schemes: one step prediction performance (four hours ahead) and multiple step prediction performance evaluation (up to 24 hours ahead).

- **One step prediction**

One step prediction performance was evaluated by using the MAE and r^2. In addition, the reliability of model was evaluated by comparing it with measurement error of the output of the model. If the prediction error was less than the measurement error, the models were considered reliable. Both prediction performance and reliability of the process models were assessed via the associated test set.

- **Multiple step prediction**

Once the best models for one step prediction were established, their capacity was challenged by long range prediction. The main objective of long range prediction was to investigate how far the clarifier process model could predict the clarified water qualities with the present information. To predict the clarified water qualities at t+N steps in the future, the predicted values of clarified water qualities from t+N-1 to the present time were input into the model. For example, to predict the clarified turbidity water at 3 steps ahead from now (i.e. 12 hours ahead), the predicted values of clarified water turbidity (and colour) at the 4^{th} hour were input into the process model to predict clarified water qualities at the 8^{th} hour. Consecutively, both predicted values were again input into the clarifier process model to predict clarified water qualities at the12^{th} hour. The multiple step prediction concept was employed in case studies. There was a link between the clarified water turbidity and colour models for the BEWTP case. Both models used the clarified water turbidity and colour as inputs to the model. On the other hand, the clarified water turbidity model worked individually in the case of the BKWTP.

The model development process was applied for both case studies and the details of each case study will be separately discussed in the following sections.

5.3 Bryn Estyn Clarifier Process Modelling

This section focuses on the specific modelling process for the BEWTP clarifier process model for each of the three development stages (Figure 5.1). Two ANN clarifier process models were developed using site specific operational data to predict the clarified water turbidity and colour for the BEWTP case study. In short, one ANN process model was used to predict clarified water turbidity and another one for clarified colour prediction. Operational data from August 2002 to May 2003 and their associated statistical analyses were described in Chapter 3. After erroneous operational data had been filtrated out, the sampling rates were formatted to four hours.

5.3.1 Determination of the Significant Inputs

The type of inputs and their temporal spans were important to model performance. Suitable inputs for the clarified water colour and turbidity model were determined as described in this section. The list of available variables is shown in Table 5.1. The type of input variables and suitable temporal spans for the clarifier models are discussed separately as follows.

Type of input variables

The coefficients of correlation (r) of each input variable to the model outputs (i.e. the clarified water turbidity and colour) were low (in the absolute value range of 0.02 to 0.51, Table 3.4) due to the complex behaviour of the clarifier. Unfortunately, this meant that no firm thresholds of r could be set to ignore some variables. Some variables such as the turbine speed, plant flow rate and polyelectrolyte dosage showed insignificant effect on clarifier performance since their coefficients of correlation (r) were very low. To add these inputs to the model might only add noise to the model. However it was a very indecisive choice to ignore them since all these input variables were recommended as being important to clarifier operation (ASCE. & AWWA., 1990; Kawamura, 1991). In addition, according to the interview with the head of operators and the BEWTP operational manual (City Water Technology, 2001), the operators at BEWTP have in practice been using values of these variables as the primary information to operate their

clarifiers. Therefore all WTP inputs were used as the inputs for the two process models. These are tabulated in Table 5.1.

Suitable temporal span

The clarifier process model is a temporal system and it is necessary to determine the optimal temporal span for each of the input data. It was assumed that the optimal temporal span for an input variable could be found by maximising r^2 and minimising the MAE in a rigorous series of trials where different temporal spans of inputs were evaluated. The suitable temporal span can be guided by considering the Autocorrelation function (ACF) (section 4.5.1) with a threshold of 0.5. Any time steps beyond this threshold are less significant to the present step and they can be negligible (Filho et al., 2006; Venema et al., 1996). The ACF of the clarified water turbidity was critical and as shown in Figure 5.2, was a relatively low value, particularly at temporal spans longer than 16 hours (four time step lags). Considering this, a five time step lag (20 hours time lag) was assigned as the maximum temporal span considered for the investigations of suitable temporal span.

One should know that the clarifier at BEWTP is an *Accelator*, which is solid contact type with sludge recirculation. Some sludge was kept and reused for flocculation enhancing. It was necessary to make sure that the maximum temporal span covered the residential time of the sludge in the clarifier since it was the evidence of how well the operator did and how serious the clarifier was annoyed by ambient changes. The sludge concentration at BEWTP was normally in the range of 20 to 30 percent (by volume) (City Water Technology, 2001). If, for example, the sludge concentration was below 10 percent at 12 hours previously, then it would indicate that there was abnormal situation happened that time and if this sludge still remained in the clarifier, its effects should be taken in to consideration. Actually there was no report on the exact residential time of sludge in the clarifier. The best guess was 20 hours temporal span since it was much longer than the time that water was retained in the clarifier (hydraulic retention time) of about 2 to 3.5 hours. Although it might or might not cover the residential time of sludge, it should be long enough to get more information for the process model.

The trial and error was performed by varying the temporal span by step of four hours up to the upper limit of 20 hours. Therefore the choices of temporal spans were present time,

4 hours, 8 hours, 12 hours, 16 hours and 20 hours. The list of inputs for the clarified water turbidity process models and colour are tabulated in Table 5.1

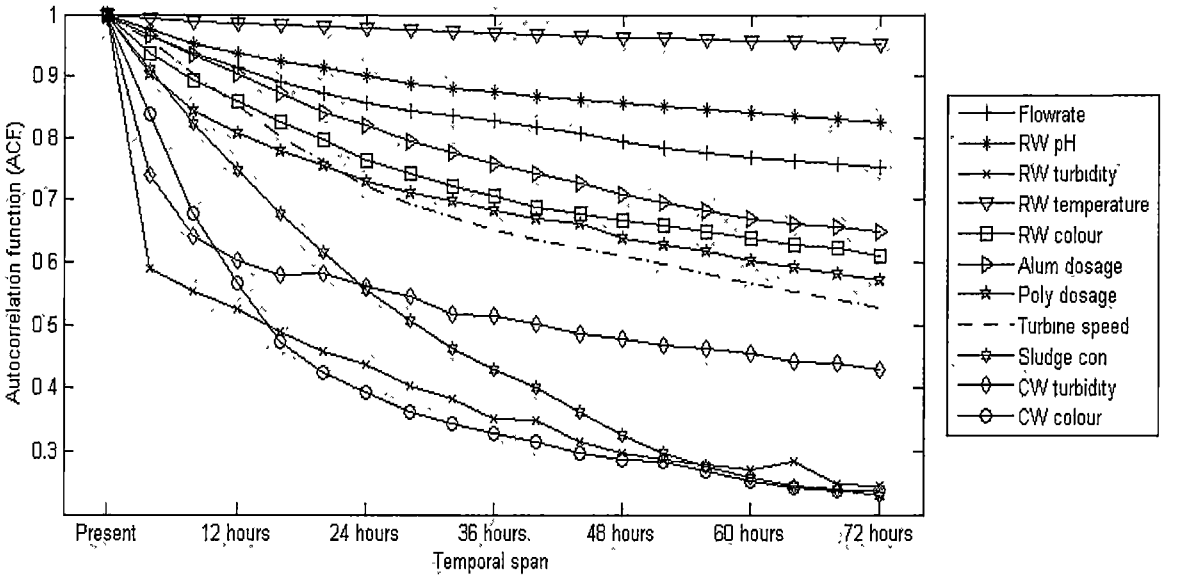


Figure 5.2 Autocorrelation function (ACF) of input variables (BEWTP), showing the ACF values of the inputs of the clarifier process model in the increase of temporal span

Variable	Temporal span	Remark
Raw water turbidity (NTU)	t, t-1, t-2, t-3, t-4, t-5	Model input
Raw water colour (HU)	t, t-1, t-2, t-3, t-4, t-5	Model input
Raw water pH (pH units)	t, t-1, t-2, t-3, t-4, t-5	Model input
Raw water temperature ($^{\circ}C$)	t, t-1, t-2, t-3, t-4, t-5	Model input
Sludge concentration (% by volume)	t, t-1, t-2, t-3, t-4, t-5	Model input
Plant flow rate (MLD)	t, t-1, t-2, t-3, t-4, t-5	Model input
Turbine speed (rpm)	t, t-1, t-2, t-3, t-4, t-5	Model input
Alum dosage (mg/L)	t, t-1, t-2, t-3, t-4, t-5	Model input
Polyelectrolyte dosage (mg/L)	t, t-1, t-2, t-3, t-4, t-5	Model input
Clarified water turbidity (NTU)	t, t-1, t-2, t-3, t-4, t-5	Model input
Clarified water colour (HU)	t, t-1, t-2, t-3, t-4, t-5	Model input
Clarified water turbidity (NTU)	**t+1**	**Model output**
Clarified water colour (HU)	**t+1**	**Model output**

Table 5.1 The selected inputs with temporal spans and outputs for clarified water turbidity and colour models for BEWTP

5.3.2 Data Clustering

After selecting suitable variable types and temporal spans, the input and associated output data was divided into three subsets: training, validation and testing sets comprising a total of 1700 data sets each with 20 hours temporal span inputs and 2 outputs (1 data set contains 68 variables). These subsets were divided in such a way that they were statistically consistent and thus represented an identical statistical population.

The SOM data clustering technique was used for this task. SOM was used with a 25 x 25 Kohonen layer with a rectangular grid topology and the Kohonen learning rule. Choosing the maximum number of clusters was obtained by an empirical approach where the number of epochs was increased until the number of clusters remained unchanged. Figure 5.3 shown that any iteration further than 5000 epochs did not increase the number of clusters. Therefore it was assumed that 355 clusters was the maximum number of cluster of the data at hand. Additionally, rather than being a monotonic trend, fluctuations in the number of clusters with number of the epochs were assumed to be the effect of random searches in the SOM algorithm.

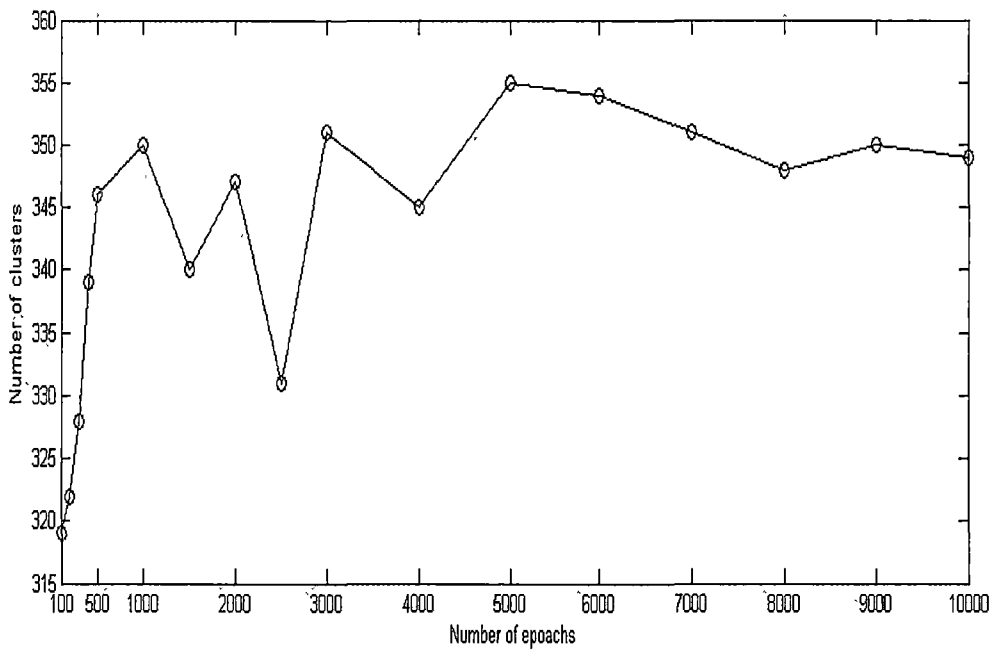


Figure 5.3 The number of clusters in BEWTP operational data, illustrating the number of clusters at each number of epochs obtained by using clustering method based SOM

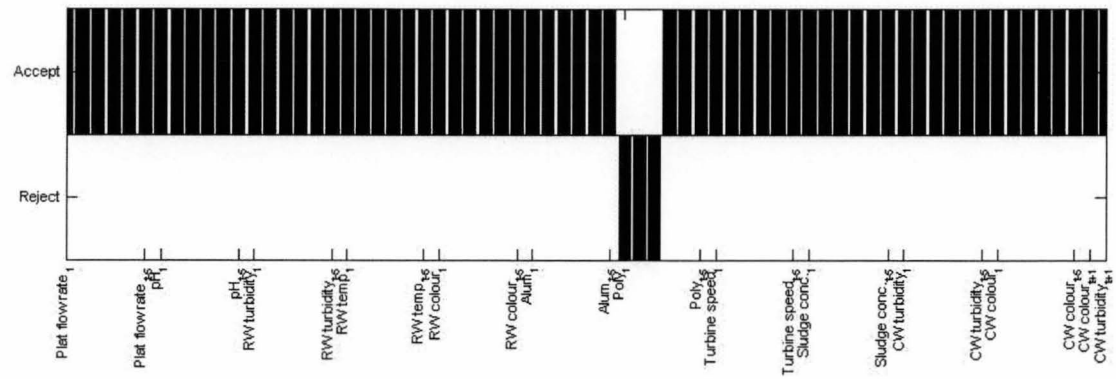
The 1700 data sets were clustered into 355 clusters. There were 148 clusters containing less than three data sets these were all assigned to join the training set. There were 207 clusters containing at least three data sets. From these clusters, two were randomly selected for the validating and testing sets and the rest were assigned to the training set. Therefore the training set contained 1286 data sets and the validation and testing set each contained 207 data sets.

The result of comparing the statistical equality between both the testing and the validation sets to the training set, for each of 68 variables, was that some of the T-test and F-test null hypotheses were rejected. At most, only 3 variables of the testing and validation sets (from 68 variables) were not statistically equal to those of the training set. Therefore the testing and the validation sets could be assumed to be statistically similar to the training set. The numbers of rejected null hypotheses as well as the associated variables were shown in Table 5.2 and Figure 5.4.

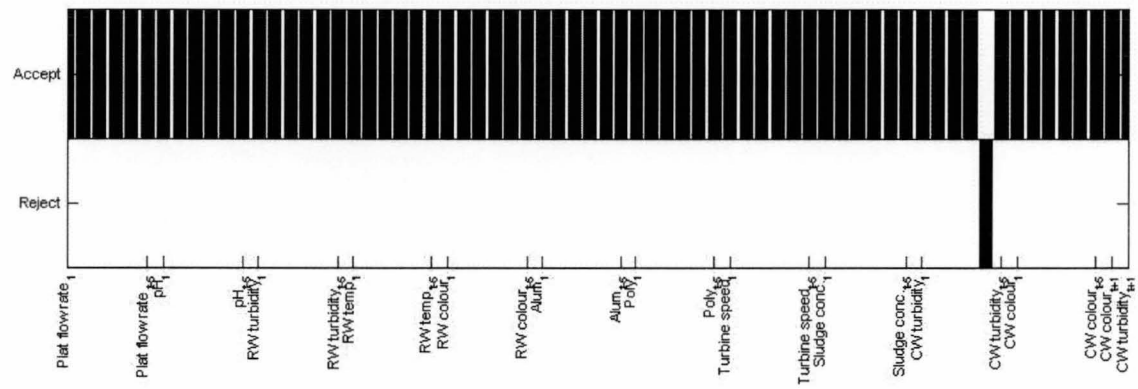
The range of the training set should cover the range of testing and validation sets for each variable. It was necessary to check whether any variables belonging to the test and validation set were beyond the range. By using the range null hypothesis test, it was found that eight and seven variables from the testing and validation sets, respectively were out of range of the training set respectively. The number of rejections was about double that of the T-test and F-test because of the random selection. The results of hypothesis testing were given in Table 5.2 and shown graphically in Figure 5.4

	Number of rejections	
	Testing & training set	Validation & training set
T-test null hypothesis	3	1
F-test null hypothesis	3	3
Range test null hypothesis	8	7

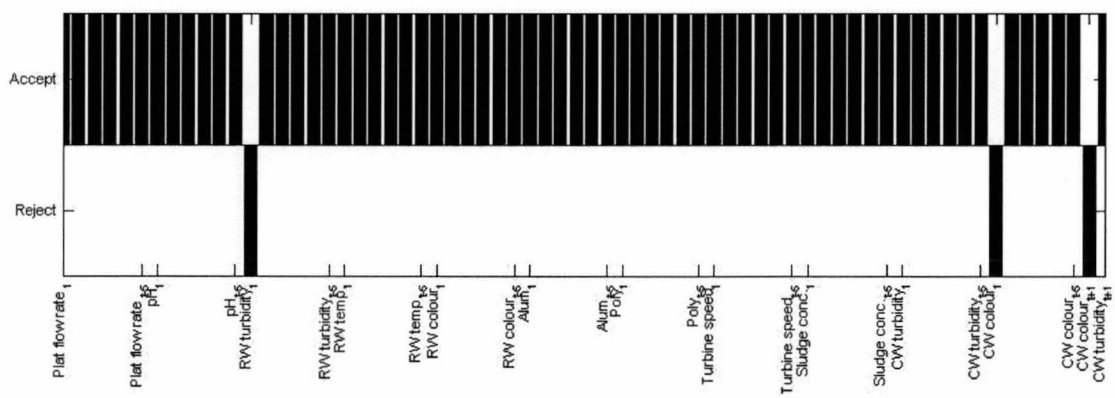
Table 5.2 The numbers of rejected null hypotheses (BEWTP), showing number of rejected null hypotheses between testing, validation and training sets



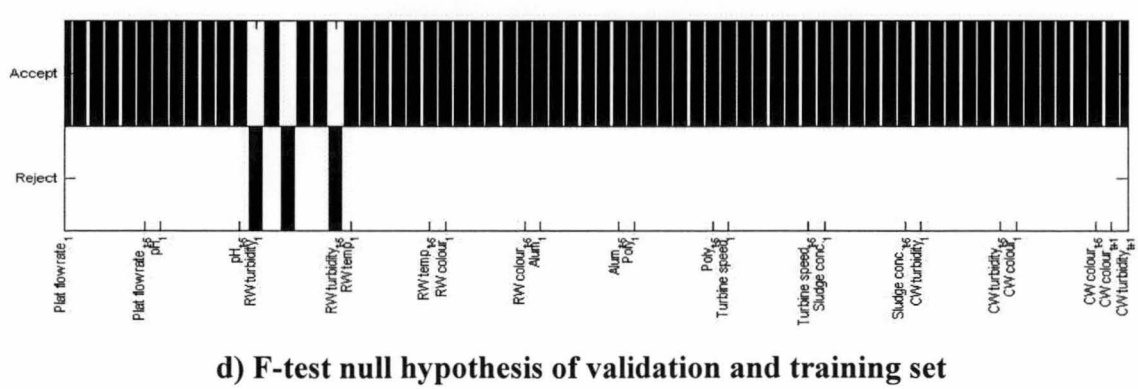
a) T-test null hypothesis of testing and training set



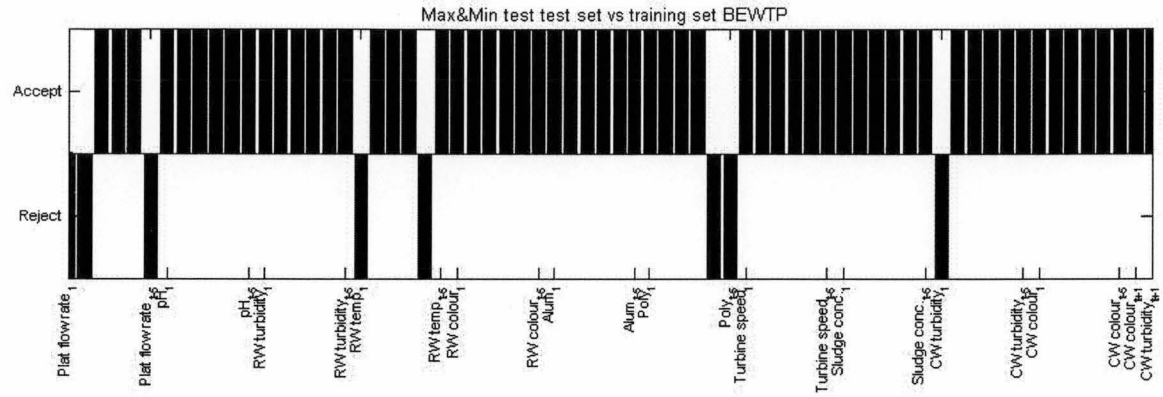
b) T-test null hypothesis of validation and training set



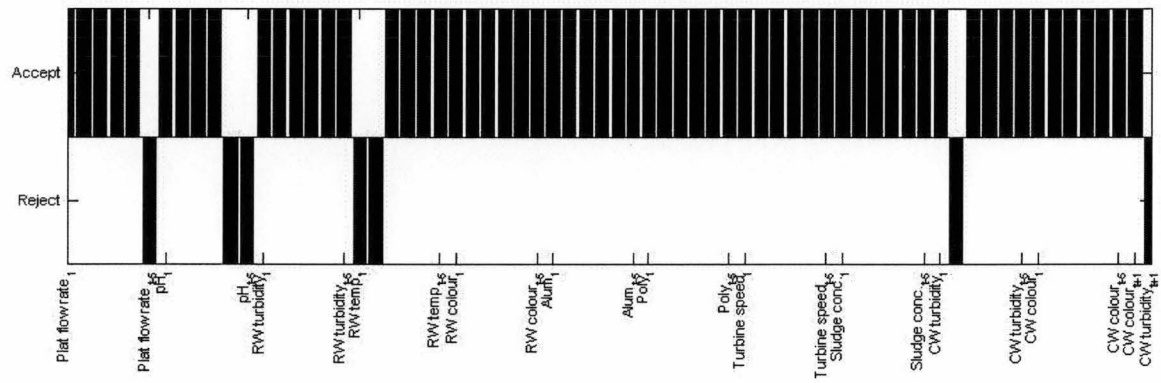
c) F-test null hypothesis of testing and training set



d) F-test null hypothesis of validation and training set



e) Range-test null hypothesis of test and training set



f) Range-test null hypothesis of validation and training set

Figure 5.4 The results of the T-test, F-test and range null hypotheses for BEWTP case study; the variables with time lag is listed on the X-axis, and results of hypothesise test are shown on the Y-axis:

(a) T-test null hypothesis of testing and training set

(b) T-test null hypothesis of validation and training set

(c) F-test null hypothesis of testing and training set

(d) F-test null hypothesis of validation and training set

(d) Range-test null hypothesis of testing and training set

(e) Range-test null hypothesis of validation and training set

5.3.3 Determination of Optimal Model Architectures

Optimisation of the model architecture was done to find the number of hidden layers and their associated neurons which gave the best predictive performance in the four hour period ahead of attaining clarified water turbidity and colour. The predictive performances of various models were evaluated using the test sets. Other model elements and associated frameworks were predetermined and shown in Table 5.3.

The optimal number of hidden layers and hidden neurons were obtained by systematic trial and error. The best model was the one that gave the largest r^2 and smallest MAE between predicted versus actual results in test set when the models were used to predict the clarified water turbidity and colour four hours ahead. According to the discussion in section 4.3.5 the maximum number of layers was limited to two. In each hidden layer, the number of hidden neurons was varied. Nonetheless, the guidelines available by which to find the suitable number of hidden neurons only quote for a network with one hidden layer. As it was reviewed in Chapter 4, among these authorities, only the work of Roger and Dowla (1994) accounted for overfitting by considering the number of samples, N^{TR}, and the number of inputs, N^I, in the training set. Roger and Dowla's formula was $N^H \leq \frac{N^{TR}}{N^I + 1}$ (Equation 4.29), where N^H was the recommended number of hidden neurons. Specific to BEWTP case study N^{TR} was 1700 samples and N^I for both turbidity and colour models was 66. This means that the recommended number of hidden neurons was 26.

However, to ensure that the optimal architecture networks were obtained, the number of neurons in the first and the second hidden layers were varied in increments of 5 over the range of 5 to 100 neurons. The maximum number of hidden neurons became 200 neurons (100 neurons for each of the first and second hidden layers) which was much less than the number of training set (1286 data sets) resulting in mathematically determined system which was in favor of ANN computation (Sha, 2007). Finally, systematic trial and error processes were conducted using the framework tabulated in Table 5.3. The resulting structure of the ANN process models for clarified water turbidity and colour are shown in Figure 5.5

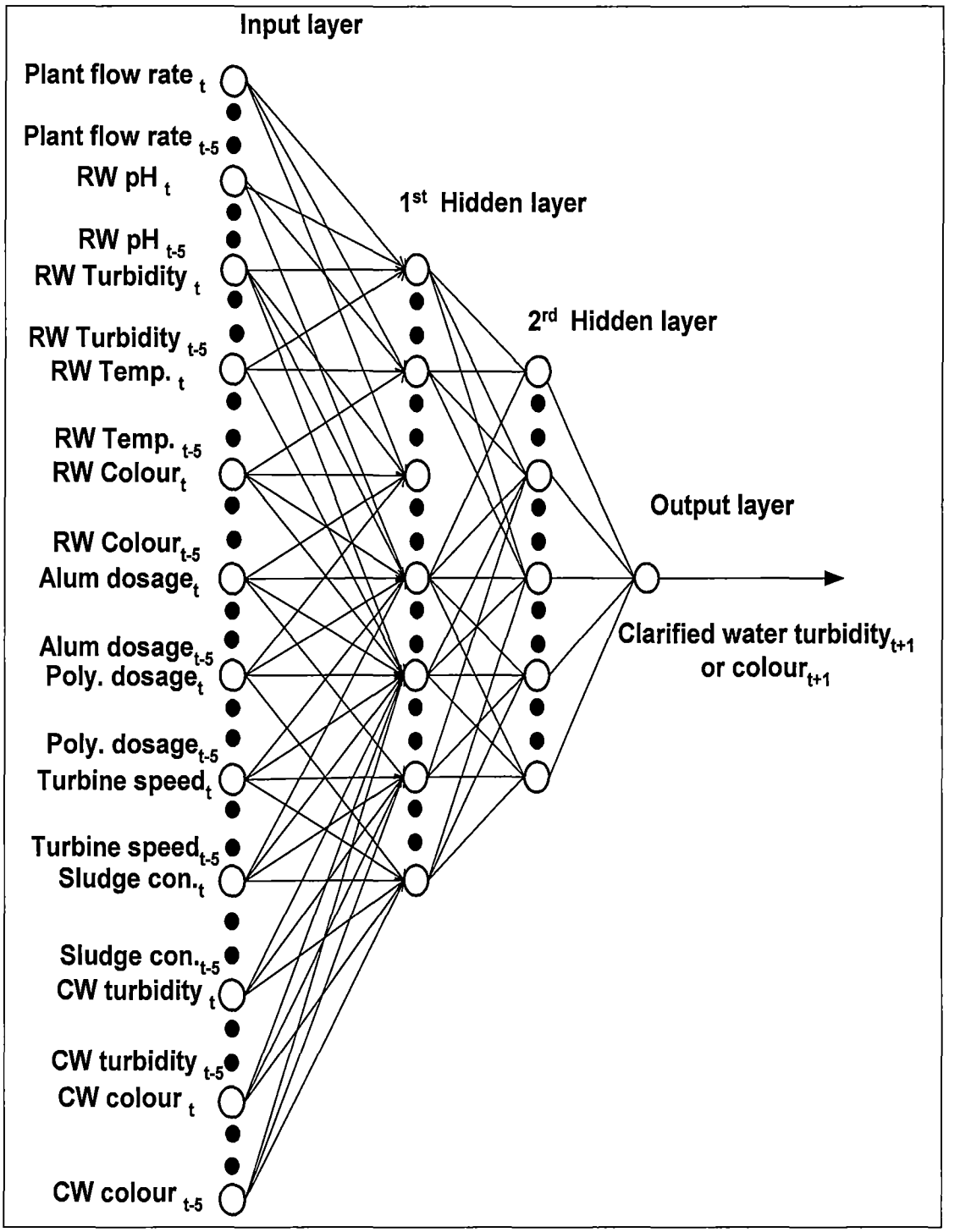


Figure 5.5 Structure of clarifier ANN models for prediction of clarified water turbidity and colour for BEWTP: showing the inputs with temporal span and the outputs of the process models

Model elements	Network architecture (varied)
Temporal span	t, t-1, t-2, t-3, t-4, t-5
Maximum number of hidden layer	1 or 2
Number of neuron in 1st hidden layer	5 to 100 with 5 increments
Number of neuron in 2rd hidden layer	5 to 100 with 5 increments
Model elements	**Network architecture (fixed)**
Neural network type	Back propagation
Activation functions	• Hyperbolic tangent function for (1st & 2rd hidden layers) • Linear function for output layer
Data normalisation	Zero mean method
Learning algorithm	Adaptive learning rate with momentum
Model validation criteria	Early stopping
Model performance evaluation	• Coefficient of determination (r^2) • Mean absolute error (MAE) • Mean percentage error (MPE)
Model output	Predictive clarified water turbidity & colour at one time step ahead,(CW turbidity $_{t+1}$) & (CW colour $_{t+1}$)
Prediction horizon	t+1, (one time step or four hours ahead)

Table 5.3 Model elements and architecture for BEWTP clarifier model development, showing which elements were varied and which were fixed in the trial and error development process

Optimal model architecture for the clarified water turbidity process model

A bar chart of the best performance ANN clarifier process models for clarified water turbidity prediction of one and two layer architectures at each time lag is shown in Figure 5.6 and tabulated in Table 5.4. The model with the best performing architecture (the "optimal model") had two hidden layers, a temporal span of 2 time lags (8 hours time lag data), 35 neurons in the first hidden layer and 5 neurons in the second layer. This model gave the result of 0.89 for r^2 and smallest MAE of 0.11 NTU. The associated mean percentage error was 16.73 percent.

Temporal span (Hours)	No of neuron	r^2	Mean absolute error (NTU)	Mean percentage error (%)
Present	80	0.87	0.13	21.60
4 Hours	85	0.87	0.13	21.46
8 Hours	10	0.88	0.12	20.27
12 Hours	70	0.87	0.13	21.77
16 Hours	35	0.87	0.12	21.12
20 Hours	25	0.87	0.13	22.37

a)One hidden layer network

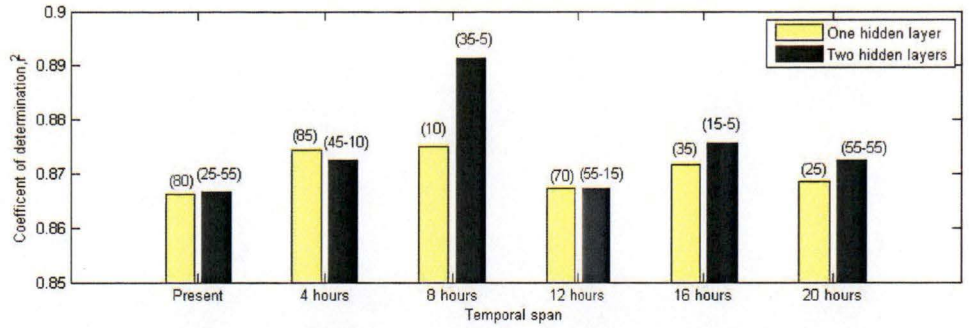
Temporal span (Hours)	No. of neuron 1st hidden layer	No. of Neuron 2nd hidden layer	r^2	Mean absolute error (NTU)	Mean percentage error (%)
Present	25	55	0.87	0.13	20.68
4 Hours	45	10	0.87	0.12	19.08
8 Hours	**35**	**5**	**0.89**	**0.11**	**16.74**
12 Hours	55	5	0.87	0.12	19.74
16 Hours	15	5	0.87	0.12	20.08
20 Hours	55	55	0.87	0.12	20.95

b) Two hidden layer network

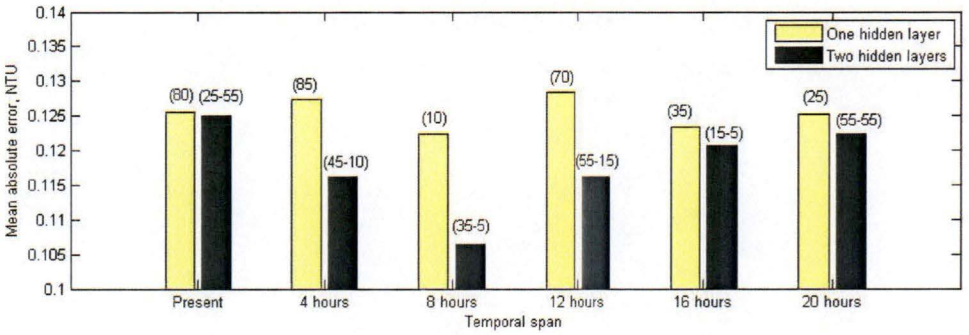
Table 5.4. The performance in prediction of clarified water turbidity for model with one and two hidden layers (BEWTP), showing the optimal model architectures and their predictive performances at each temporal span.

(a) One hidden layer network , (b) Two hidden layer network set

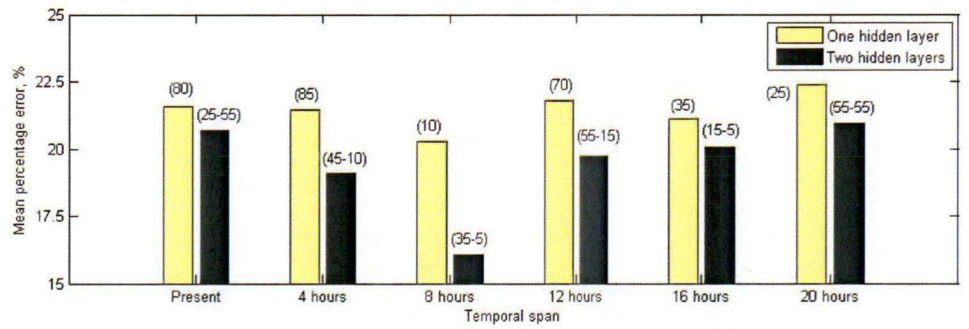
In Figure 5.6a and b, the r^2 and MAE vary only over a small range which means that their performance did not show much difference when the temporal spans were varied when the single hidden layer model was used. However, it was noted that the performance of the two hidden layer networks were superior to the one hidden layer network for every temporal span. Therefore, the affects of the time lag used could not be recognised by simple single hidden layer architecture. When the two hidden layer architecture models were employed, these models were more sensitive to a change of time lags and were able to recognise the patterns of time lag changes. Inputting of temporal spans of more than 8 hours (i.e.12, 16 and 20 hours) definitely contains the information of 8 hours temporal span. However, according to empirical results at hand, the best performance was obtained from the ANN model which temporal span of 8 hours. Temporal span of more than 8 hours might introduce noise to the process model resulting in deterioration of predictive performance.



a) Coefficient of determination (r^2)



b) Mean absolute error



c) Mean percentage error

Figure 5.6 Performance in prediction of clarified water turbidity for one and two hidden layer models with temporal span from present to 20 hours for BEWTP case study,

(a) Coefficient of determination (r^2) between predicted versus actual clarified turbidity used test set data

(b) Mean absolute error between predicted versus actual clarified turbidity used test set data

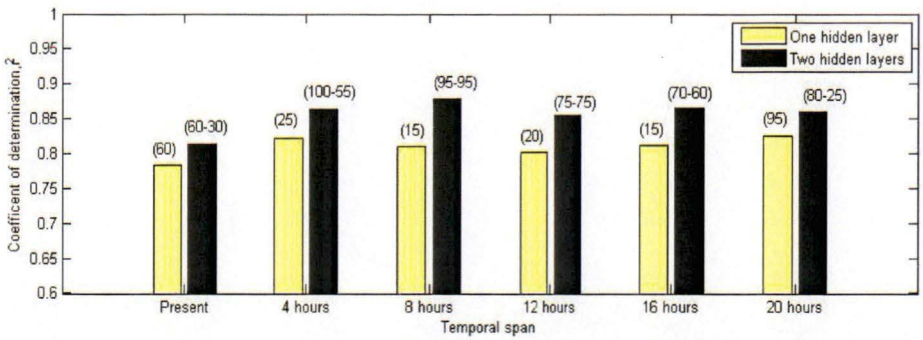
(c) Mean percentage error between predicted versus actual clarified turbidity used test set data

From Figure 5.6c, the reliability of the "optimal model" can be evaluated by comparing the mean percentage error with the turbidimeter measurement error which was in the range of 10 to 15 percent (Table 3.2). The other models were less reliable since their mean percentage errors were about 20 percent, a value well above the measurement error of turbidimeter. The "optimal model" gave the lowest mean percentage error of 16.73 percent, which was just above the measurement error of turbidimeter. This model's reliability could be marginally accepted.

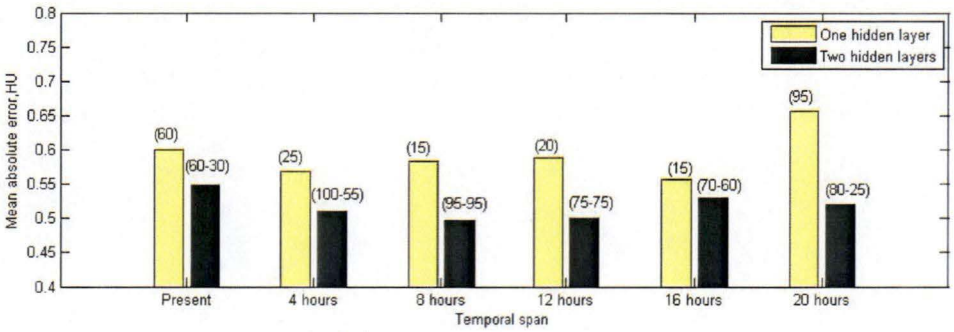
Optimal model architecture for the clarified water colour model

For clarified water colour prediction, a model with two hidden layers and a temporal span of two time lags (8 hours) showed the best performance according to empirical results at hand. Longer temporal spans (12, 16 and 20 hours) was unfavourable to predictive performance even though they contained the information of 8 hours temporal span. Temporal span of more than 8 hours might introduce noise to the process model resulting in deterioration of predictive performance.

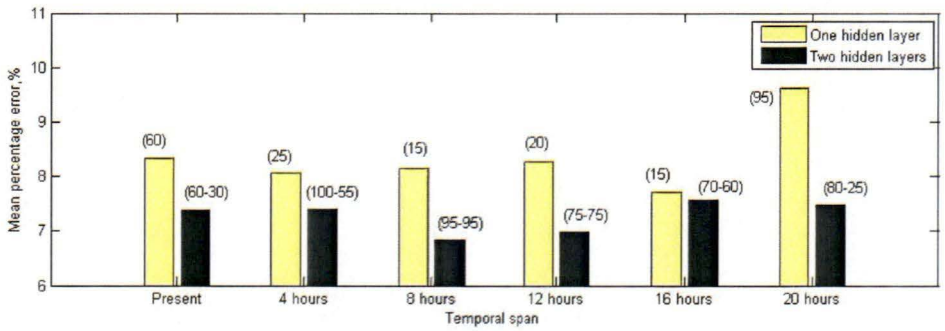
This model had the largest r^2 of 0.88 and the smallest MAE of 0.50 HU. Its mean percentage error was 6.83 percent. The optimal numbers of hidden neurons were 95 for both the first and second hidden layers. The r^2 and MAE of the best candidate model in each time lag are graphically shown in Figure 5.7 and shown in Table 5.5. The r^2 changed only over the range of 0.78 to 0.88 and MAE from 0.50 to 0.66 HU, with temporal spans from the present up to 20 hours. The two hidden layer networks always showed superior performance to the one hidden layer network at the same temporal span. Every model regardless of temporal span showed a reasonable performance and all were reliable since their MAEs were well below the water colour measurement error of 1 HU (Table 3.2).



a) Coefficient of determination (r^2)



b) Mean absolute error



c) Mean percentage error

Figure 5.7 Performance in prediction of clarified water colour for one and two hidden layer models with temporal span from present to 20 hours for BEWTP case study,

(a) Coefficient of determination (r^2) between predicted versus actual clarified colour used test set data

(b) Mean absolute error between predicted versus actual clarified colour used test set data

(c) Mean percentage error between predicted versus actual clarified colour used test set data

Temporal span (Hours)	No of neuron	r^2	Mean absolute error (HU)	Mean percentage error (%)
Present	60	0.78	0.60	8.34
4 Hours	25	0.82	0.57	8.07
8 Hours	15	0.81	0.58	8.14
12 Hours	20	0.80	0.59	8.28
16 Hours	15	0.81	0.56	7.72
20 Hours	95	0.83	0.66	9.62

a) One hidden layer network

Temporal span (Hours)	No. of neuron 1st hidden layer	No. of Neuron 2nd hidden layer	r^2	Mean absolute error (HU)	Mean percentage Error (%)
Present	60	30	0.81	0.55	7.39
4 Hours	100	55	0.86	0.51	7.40
8 Hours	**95**	**95**	**0.88**	**0.50**	**6.83**
12 Hours	75	75	0.86	0.50	6.99
16 Hours	70	60	0.87	0.53	7.56
20 Hours	80	25	0.86	0.52	7.49

b) Two hidden layer network

Table 5.5. The performance in prediction of clarified water colour for model with one and two hidden layers (BEWTP), showing the optimal model architectures and their predictive performances at each temporal span.

(a) One hidden layer network

(b) Two hidden layer network set

5.3.4 Performance of Clarifier Process Models

The performance of clarified water colour and turbidity process models was measured by the use of testing sets. The performance could be measured under two schemes: one step prediction and multiple step prediction.

One time step ahead prediction is demonstrated in Figure 5.8 showing the flow of the data for the real clarifier and the associated predictive models. Since the clarified water turbidity and colour were set as operation targets, two process models were separately used to predict the clarified water qualities. When the switch was turned to the "A" position, the clarified water turbidity and colour worked individually. Without linkage between the two process models, they used the previous operational data to predict water qualities only one step ahead (four hours).

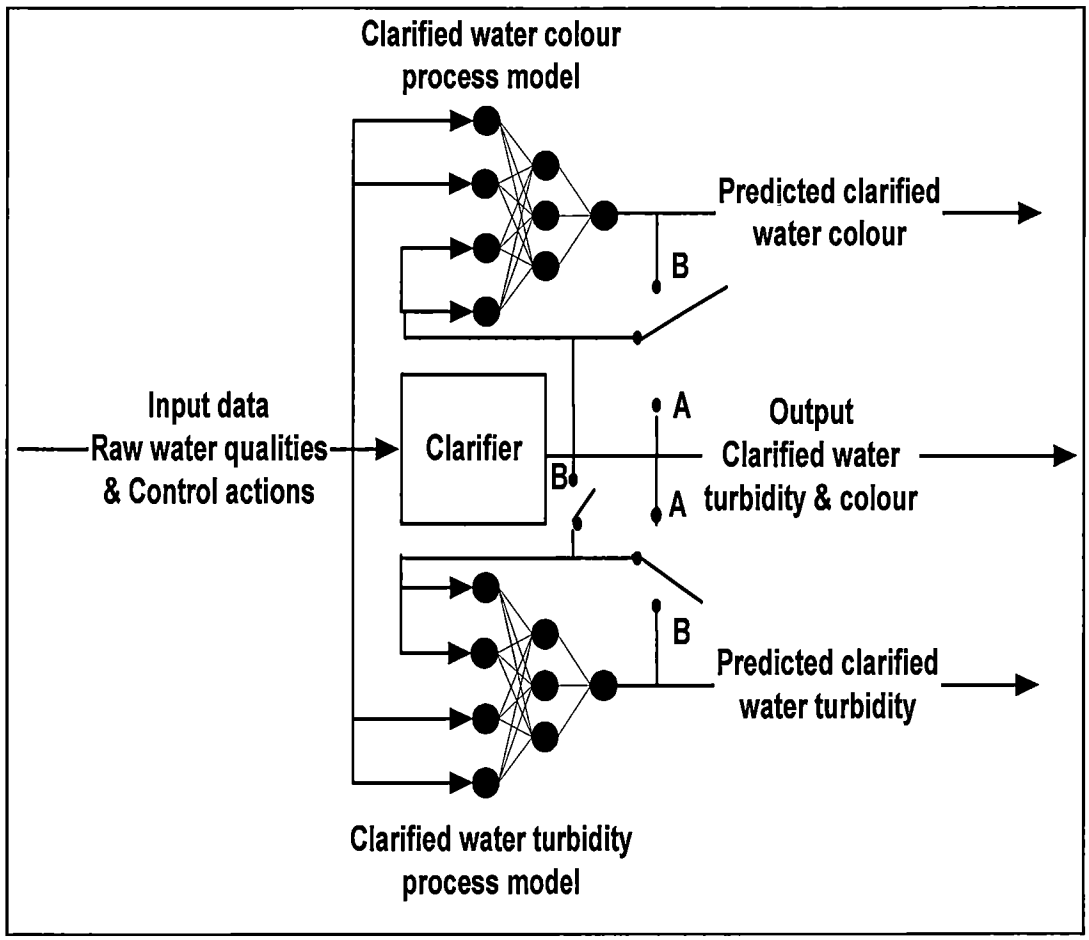


Figure 5.8 One and multiple steps predictions by two clarifier process models (BEWTP), when the switch was turned to "A" position, it was for one step prediction, and "B" position, it was for multiple steps prediction

Multiple step predictions were performed by inputting the predicted water qualities to the individual process models as shown in Figure 5.8. When the switch was turned to the "B" position, the clarified water turbidity and colour models were linked together in order to share information. Both predicted clarified water turbidity and colour (from the previous time step) were simultaneously inputted to both turbidity and colour process models to predict the following time steps. This iteration kept going on until six time steps ahead (24 hours). The predictive ability of clarified water turbidity and colour process models were separately evaluated as follows:

Clarified water turbidity model performance evaluation

- **One step predictive performance**

With the inputs from the optimal temporal span of two steps (eight hours), the best performing ANN clarified turbidity process model had two hidden layers and 35 neurons in the first and 5 neurons in the second hidden layer. The predicted and actual clarified water turbidity is shown in Figure 5.9. The clarified turbidity process model could track the trend very well with the r^2 of 0.89, and good predictive accuracy with a small MAE of 0.11 NTU. In terms of model reliability, the overall mean percentage error of 16.08 percent was slightly greater than the measurement error of the online turbidimeter (10 to15 percent). This was marginally acceptable. The associated statistic parameters are shown in Table 5.6

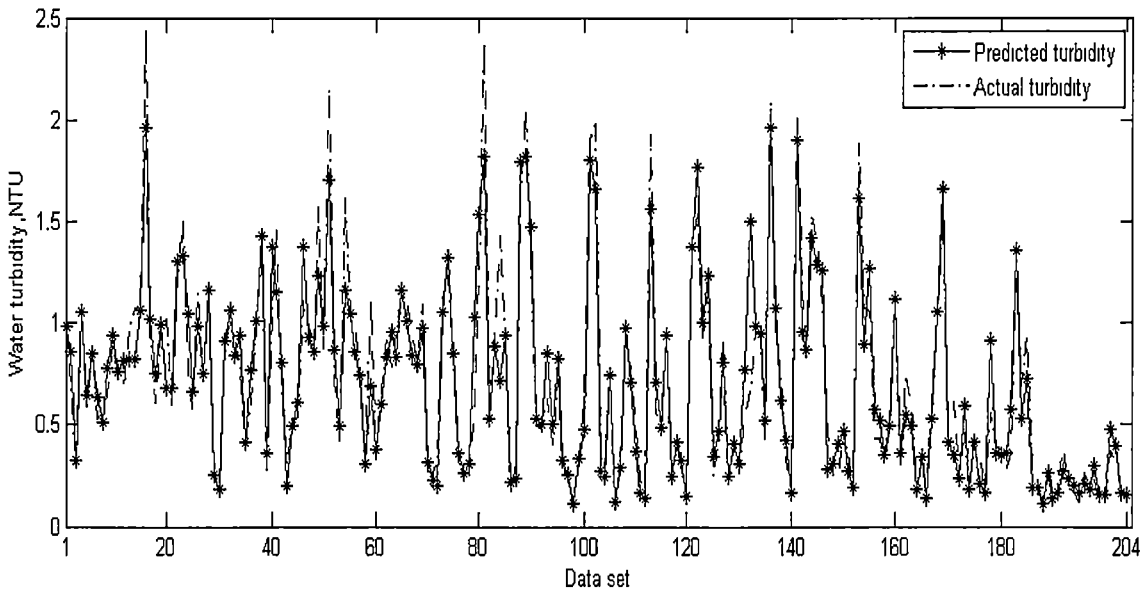


Figure 5.9 Predicted clarified turbidity, showing the predicted and actual clarified water turbidity of the test set data

Statistical parameters	Values
Coefficient of determination (r^2)	0.89
Mean absolute error	0.11
Mean percentage error	16.08

Table 5.6 Performance of clarified water turbidity prediction (BEWTP), showing performance parameters when the process model predicted the clarified water turbidity in four hours ahead using test set data

The raw water turbidity was very similar in quality to the clarified water turbidity as denoted by the large correlation (r) of 0.51 comparing to the other variables (Table 3.4) and therefore it was useful to evaluate the model's performance through the domain of raw water turbidity input. The plot of the predictive percentage error against the increase of raw water turbidity is shown in Figure 5.10. The predictive percentage error slowly dropped from 17.1 to 16.2 while the raw water turbidity was extended to 27.5 NTU. The model reliability is marginally acceptable since it levelled to the maximum measurement error of turbidimeter (10-15 NTU).

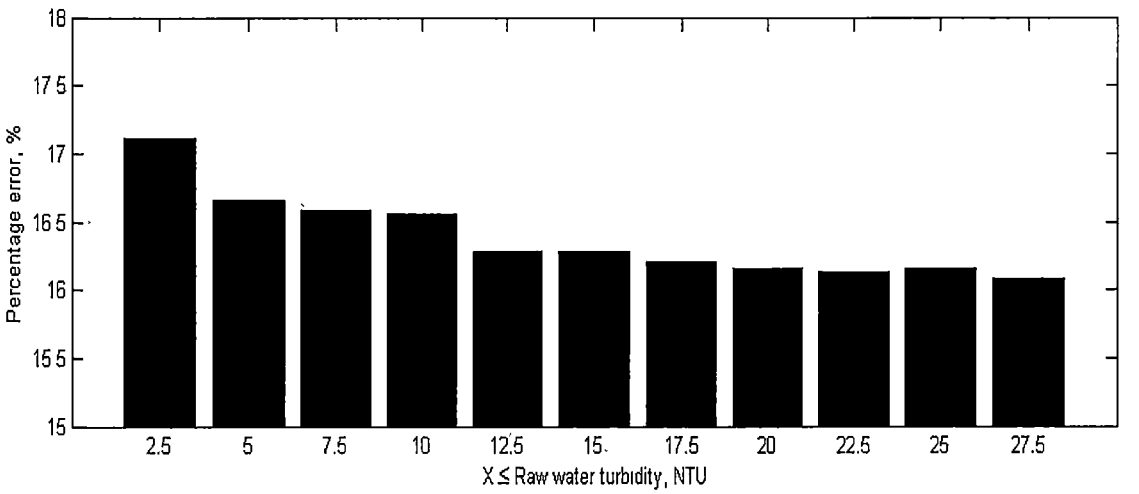
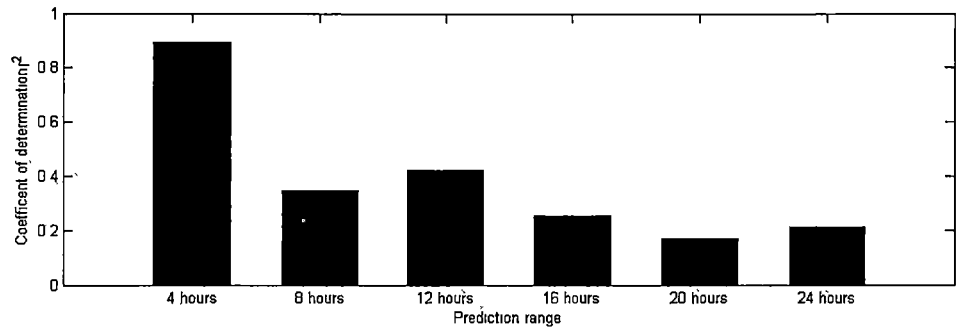


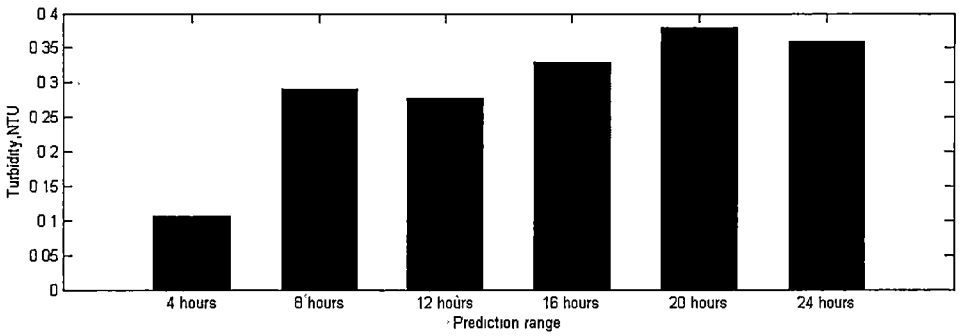
Figure 5.10 Reliability of clarified water turbidity model (BEWTP): showing predictive percentage error value when the raw water turbidity increased

- **Multiple step prediction performance**

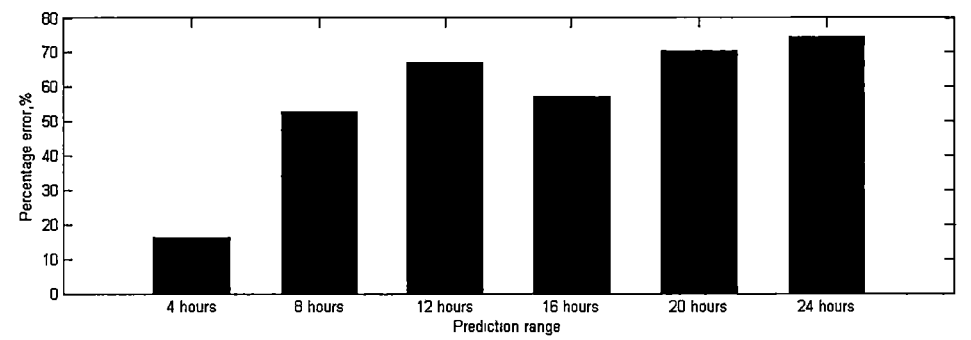
The multiple step prediction of clarified turbidity predicted up to 6 time steps (24 hours) ahead. The bar plots of r^2 and MAE of each time step are shown in Figure 5.11a and 5.11b. The reduction of r^2 and the increase of MAE implied the existence of a deterioration of predictive performance with the prediction range. After the first lag of four hours, the predictive performance dropped dramatically, due to the accumulation of prediction errors when the ANN process model was operated recursively. In terms of reliability, the mean percentage error rose above the measurement error of the turbidimeter after the first lag. Therefore, prediction beyond that first lag of four hours is not recommended.



a) Coefficient of determination (r^2)



b) Mean absolute error



c) Mean percentage error

Figure 5.11. Multiple steps prediction performance of clarified water turbidity (BEWTP) from four hours to 24 hours ahead using test set data

(a) Comparison of coefficient of determination (r^2)

(b) Comparison of mean absolute error

(c) Comparison of mean percentage error

Clarified colour process model performance

- **One step predictive performance**

Using the inputs of the optimal time lag of two steps (eight hours), the clarified water colour model with two hidden layers of 95 neurons in both the first and second hidden layers showed the best performance. This model could recognise the inputs and predicted the clarified water colour vary well with a high r^2 of 0.88 and small MAE of 0.50. The plot of predicted and actual clarified water colour is shown in Figure 5.12 and the associated mean percentage error was 6.83 percent. This process model was reliable since the mean absolute error (MAE) was 0.50 HU which was far less than the water colour measurement error of 1HU (Table3.2). All statistic parameters are shown in Table 5.7

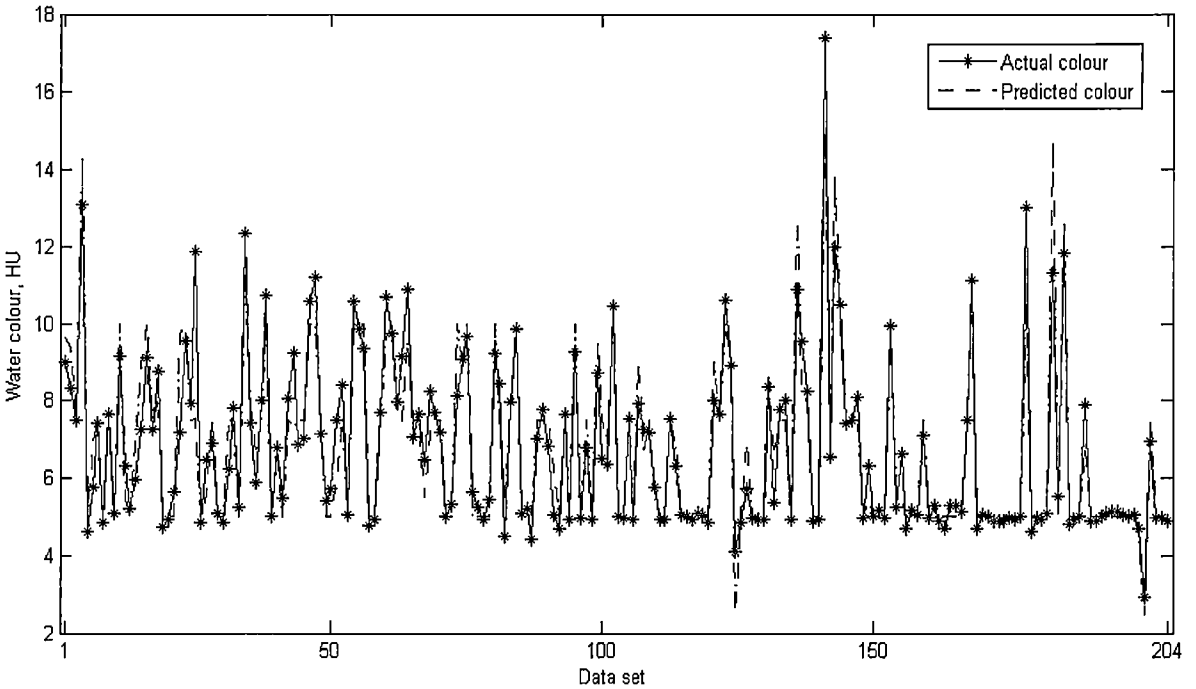


Figure 5.12 Predicted clarified colour, showing the predicted and actual clarified water turbidity of the test set data

Statistic parameters	Values
Coefficient of determination (r^2)	0.88
Mean absolute error (HU)	0.50
Mean percentage error (%)	6.83

Table 5.7 Performance of clarified water colour prediction (BEWTP), showing performance parameters when the process model predicted the clarified water colour in four hours ahead using test set data

Since raw water colour was closely related to the clarified water colour, the predictive performance of the model should be evaluated through raw water colour. The plot of the increasing of raw water colour against the predictive MAE of clarified water colour is shown in Figure 5.13. The predictive error, MAE increased from 0.18 to 0.50 HU while the raw water colour extended to 65 HU. Beyond this point, the MAE became nearly constant through the further range of raw water colour. If water colour measurement error of 1 HU was set as the acceptable threshold, this process model was reliable.

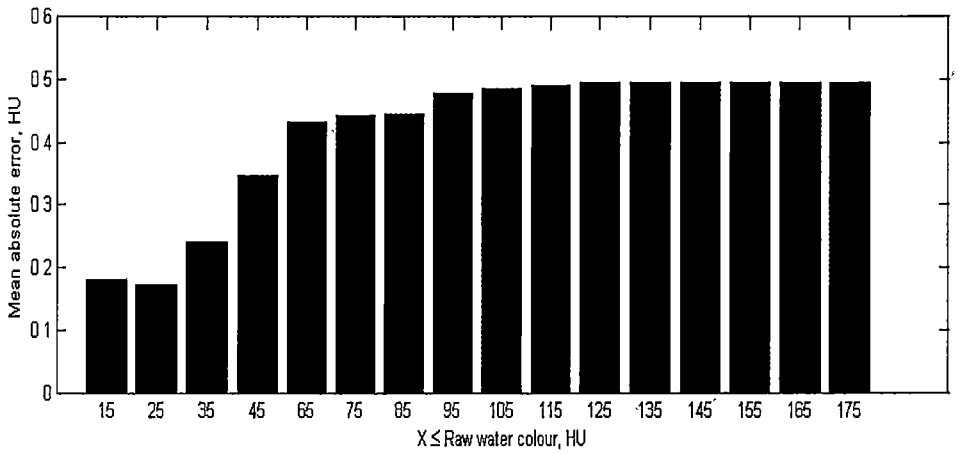
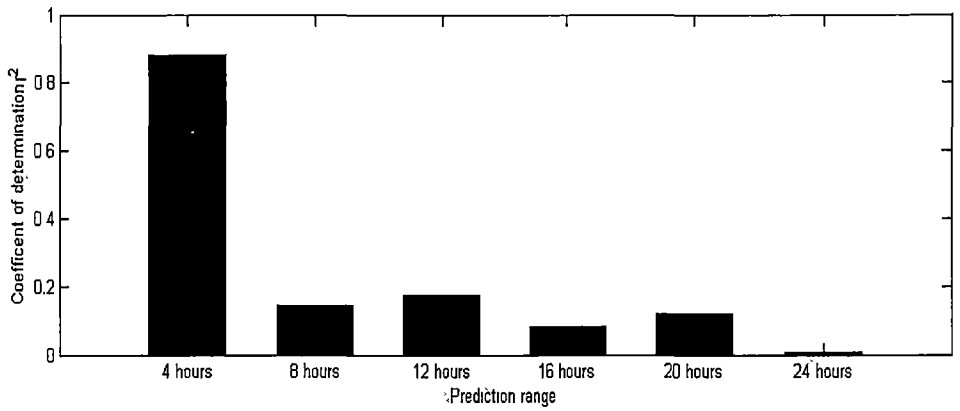


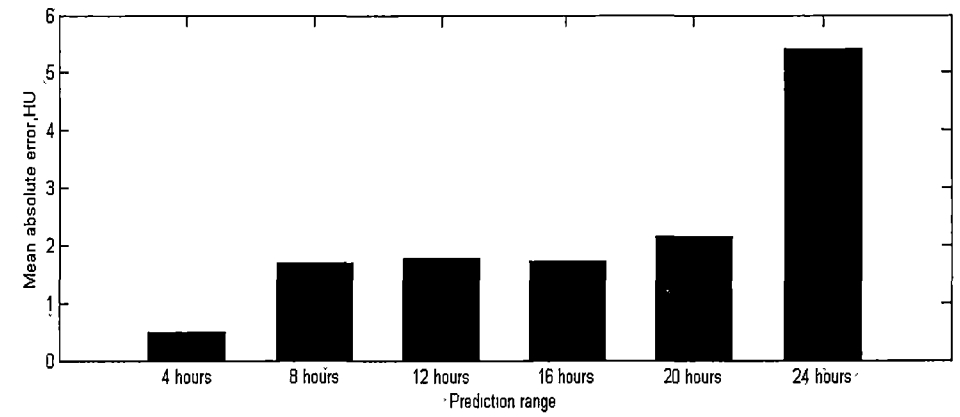
Figure 5.13 Reliability of clarified water colour model (BEWTP): showing predictive percentage error value when the raw water colour increased

Multiple steps prediction performance

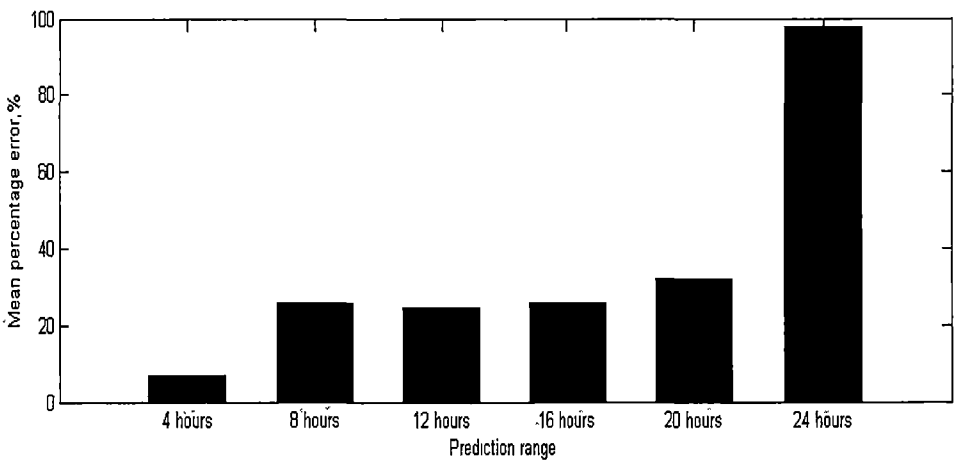
In the same manner as the clarified water turbidity process model, the multiple step prediction performance was evaluated as far as 6 time steps (24 hours ahead). Plots of r^2 and MAE for each time step were shown in Figure 5.14a and 5.14b respectively. The associated mean percentage error of each time step is graphically shown in Figure 5.14c. The longer the prediction range was made, the more the predictive performance deteriorated due to the accumulation of errors. The predictive performance deteriorations clearly found after the first time step (four hours) were also evidenced in a dramatic reduction of r^2, increasing both MAE and mean percentage error. The model reliability suffered severely since the MAE drastically rose higher than the water colour measurement error after the first time step prediction (evidenced in the plot of mean percentage error and prediction horizon as shown in Figure 5.14b). Thus, using the clarified water colour process model to predict further than one time step was not recommended.



a) Coefficient of determination (r^2)



b) Mean absolute error



c) Mean percentage error

Figure 5.14. Multiple steps prediction performance of clarified water colour (BEWTP) from four hours to 24 hours ahead using test set data

(d) Comparison of coefficient of determination (r^2)

(e) Comparison of mean absolute error

(f) Comparison of mean percentage error

5.4 Bang Khen Clarifier Process Modelling

This section focuses on the specific modelling process for the BKWTP clarifier process model for each of the three development stages (Figure 5.1).Since only the clarified water turbidity is an operational target at BKWTP, development only of the ANN clarifier process model was needed to predict clarified water turbidity. Development of the model was based on a total of 3788 data sets taken from past operational data from February 2003 to December 2004. Statistical analyses of the data were shown in Chapter 3. The blank and erroneous operational data were removed and replaced by interpolated data in the same manner as in the BEWTP case study. Data whose measurement frequency differed from four hours were discretised to a time interval of four hours.

5.4.1 Determination of the Significant Inputs

The number and type of the inputs and associated time lags is directly related to the size of the model input. The quantity and quality of inputs will have significant effect on the model's ability to predict. The type of inputs and a suitable temporal span for the clarified water turbidity model were separately determined. The list of associated variables is shown in Table 5.8

Type of input variables

The clarifiers at BKWTP behaved in a complex manner. All the coefficients of correlation (r) of input data to the modelling output of the clarified water turbidity were very small, varying in the range of 0.015 to 0.29 (Table 3.8) reflecting how complicate the relations were. Even though some inputs with low coefficients of correlation (r) to the model's output (i.e. clarified water turbidity) should not be omitted by setting up thresholds, every input variable was known to be significant to the clarifier's operation (ASCE. & AWWA., 1990). Following the BKWTP operational manual (MWA, 2000), the clarifiers had routinely been operated by using all these variables as operation guidelines. For example, before feeding raw water to the clarifier, the recommended water pH was in the range of 6.8 to 7.8. Therefore, this "expert knowledge" of the significant inputs implied that all types of inputs should be used as the inputs for this ANN process model.

Suitable temporal span

The plot of the Autocorrelation function (ACF) with time lag was shown in Figure 5.15 in the same manner as in the BEWTP case study. Any results in lag data that produced an ACF value less than 0.5 had insufficient affect and can therefore be ignored (Filho et al., 2006; Venema et al., 1996). From Figure 5.15, the ACF of the clarified water turbidity was at its lowest value and its ACF value beyond the first lag (four hours from present time) was less than 0.5.

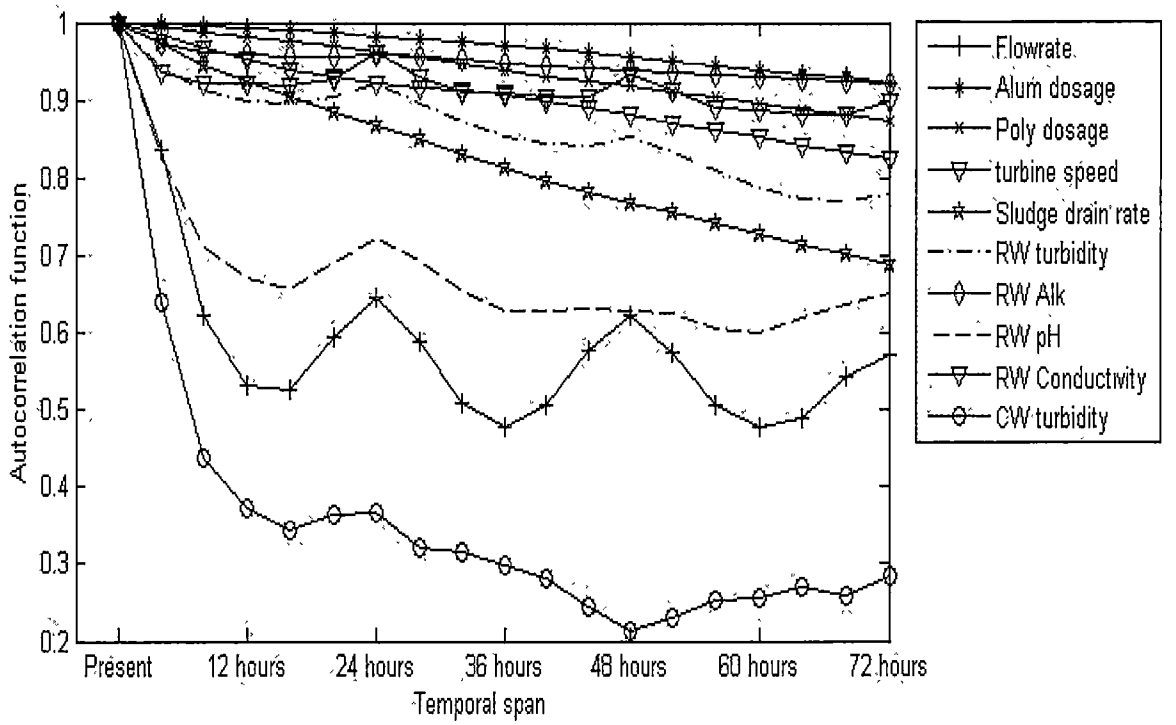


Figure 5.15 Autocorrelation function (ACF) of input variables (BKWTP), showing the ACF values of input of the clarifier process model values when the increasing of temporal span

In a manner similar to the BEWTP case, the temporal span of five lags (up to 20 hours) was assigned as the maximum temporal span to ensure that there was enough information for processing the model. Consequently, the set of maximum input data contained ten types of inputs with temporal span of five lags (60 inputs in total). The size of the input was refined by systematic trial and error, by varying the temporal span by any number of steps of four hours up from the present time to the upper limit of 20 hours. Therefore, the choices of the temporal spans were present time, 4 hours, 8 hours, 12 hours, 16 hours and 20. The list of the inputs for the clarified water turbidity process models are given in Table 5.8

Variable	Temporal span	Remark
Raw water turbidity (NTU)	t, t-1, t-2, t-3, t-4, t-5	Model input
Raw water pH (pH unit)	t, t-1, t-2, t-3, t-4, t-5	Model input
Raw water Alkalinity (mg/L as $CaCo_3$)	t, t-1, t-2, t-3, t-4, t-5	Model input
Raw water conductivity (μS/cm)	t, t-1, t-2, t-3, t-4, t-5	Model input
Plant flow rate (MCMD*)	t, t-1, t-2, t-3, t-4, t-5	Model input
Alum dosage (mg/L)	t, t-1, t-2, t-3, t-4, t-5	Model input
Polyelectrolyte dosage (mg/L)	t, t-1, t-2, t-3, t-4, t-5	Model input
Turbine speed (rpm)	t, t-1, t-2, t-3, t-4, t-5	Model input
Sludge drainage rate (min/hour)	t, t-1, t-2, t-3, t-4, t-5	Model input
Clarified water turbidity (NTU)	t, t-1, t-2, t-3, t-4, t-5	Model input
Clarified water turbidity (NTU)	**t+1**	**Model output**

Table 5.8 The selected inputs with temporal spans and output for clarified water turbidity models for BKWTP

5.4.2 Data Clustering

After input selection, the 3688 operational data set, each containing the 20 hours time lagged input variables including 60 variables and one output of clarified water turbidity in the next four hours (Table 5.8), were divided into three subsets: training, validation and testing sets using the SOM clustering technique to ensure they were statically equivalent.

The maximum number of clusters was empirically obtained using SOM with a 25 x 25 Kohonen layer with the rectangular grid topology and the Kohonen rule for weight updating. In Figure 5.16, the number of clusters became unchanged after the number of epochs reaches 7500. It varied from 421 to 439 clusters. The maximum number of clusters was assumed to be 439 which occurred when the number of epochs was 8500. Some fluctuation of the number of clusters while the number of epochs increases was due to the random nature during of searching by SOM.

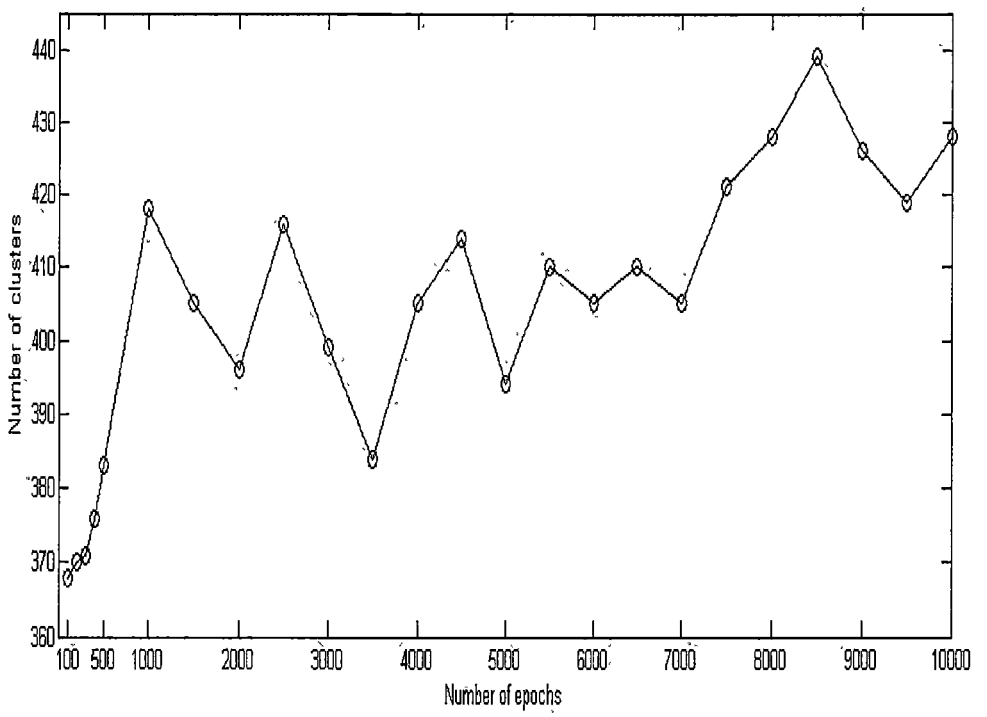


Figure 5.16 Autocorrelation function (ACF) of input variables (BKWTP), showing the ACF values of the inputs of the clarifier process model when the increasing of temporal span

The 3788 data sets were clustered into 439 clusters. Of these, 167 clusters contained less than three data in a cluster, and these were all assigned to join the training set. The remaining 272 clusters contained at least three data. From these, one data in each cluster were randomly selected for each of the validating and testing sets. The rest became members of the training set. Therefore the training set contained 3516 data and the validation and testing set each contained 272 data.

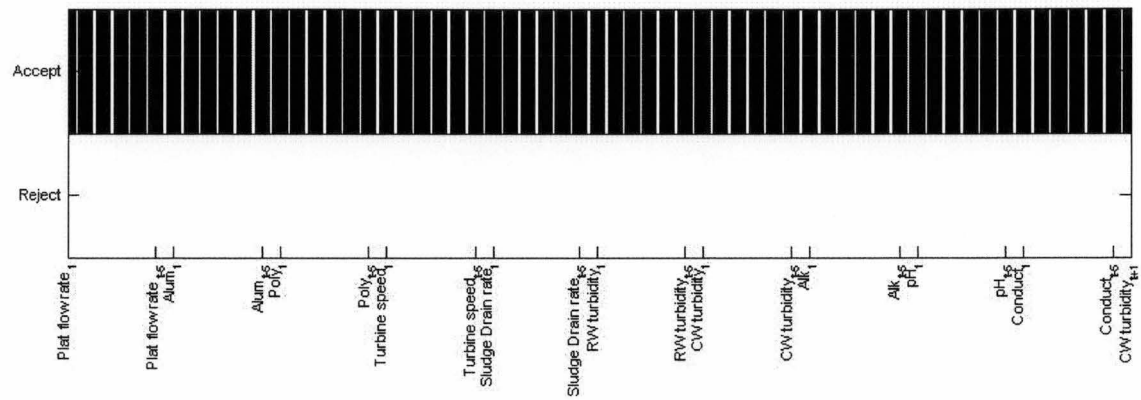
The testing and validation sets were tested to see if they had the same mean and standard deviation as that of the training set using the T-test and F-test of the null hypothesis with significant level of 0.05. In order to ensure that the networks were tested and validated in the interpolation rather than extrapolation domain, a range test was conducted to check whether any of the variables used in the testing and validation sets were beyond the range of the training set. These statistical tests were conducted upon all 61 variables (i.e. 60 input variables and one output variable). The numbers of rejections are shown in Table 5.9. For more detail, the associated statistic test results are graphically shown in Figure 5.17.

For the statistical test of the testing set and the training set, all variables were accepted for the T-test null hypothesis. Four variables failed the F-test null hypothesis and eight variables were rejected by the range test. These results were graphically presented in Figure 5.17a, Figure 5.17c and Figure 5.17e respectively. For the statistical test of the validation set and the training set, only three variables failed the T-test hypothesis, one variable failed to satisfy the F-test null hypothesis and one variable was rejected for the range test hypothesis. These results were graphically presented in Figure 5.17b, Figure 5.17d and Figure 5.17f respectively.

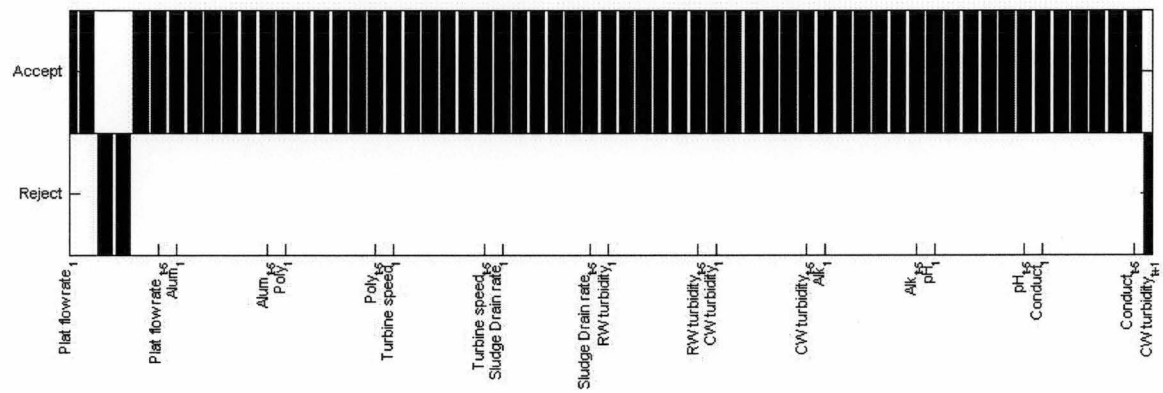
The test and validation sets were assumed to be statistically equivalent to the training set since only a very few variables failed the T-test and F-tests. This meant that means and standard deviations of most variables in the test and validation set were similar to that of the training set. However, eight variables of testing set failed the range-test due to the fact that the nature of the random selection within the SOM clustering approach could not perfectly meet all constraints.

	Number of rejections (percentage)	
	Testing & training set	Validation & training set
T-test null hypothesis	0	3
F-test null hypothesis	4	1
Range test null hypothesis	8	1

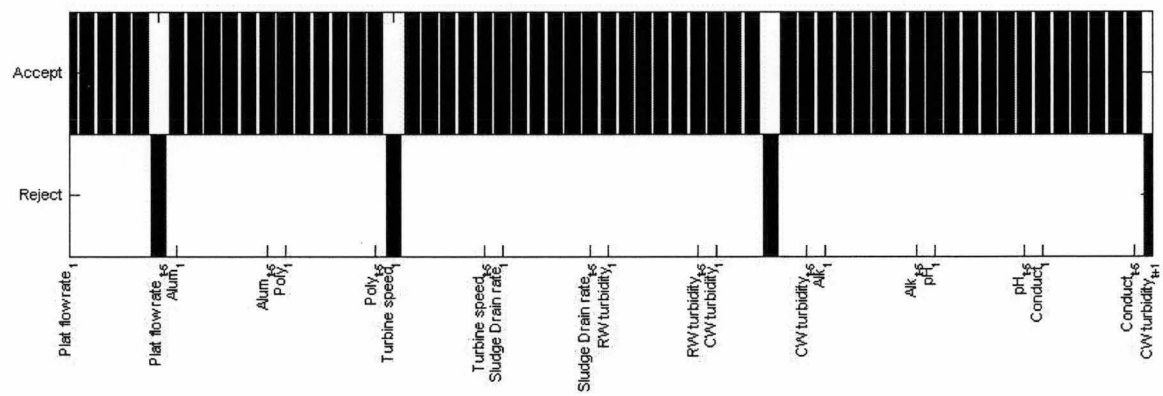
Table 5.9 The numbers of rejected null hypotheses (BKWTP), showing number of rejected null hypotheses between testing, validation and training sets



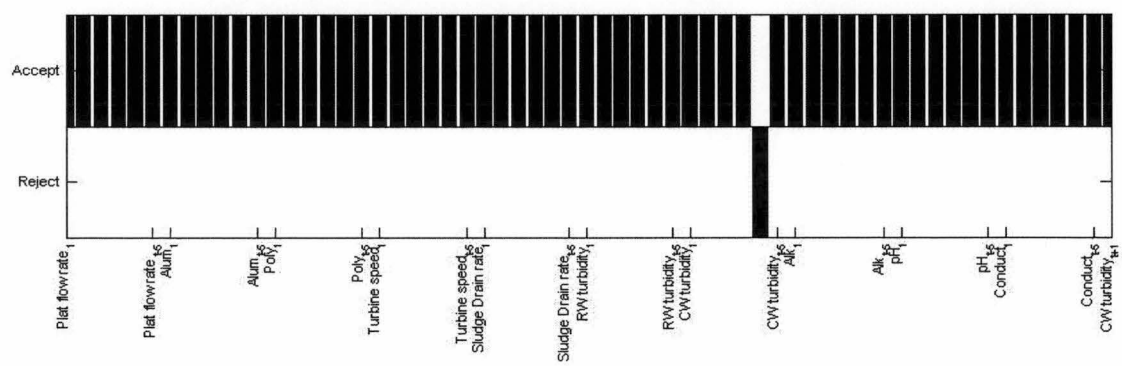
a) T-test null hypothesis of testing and training set



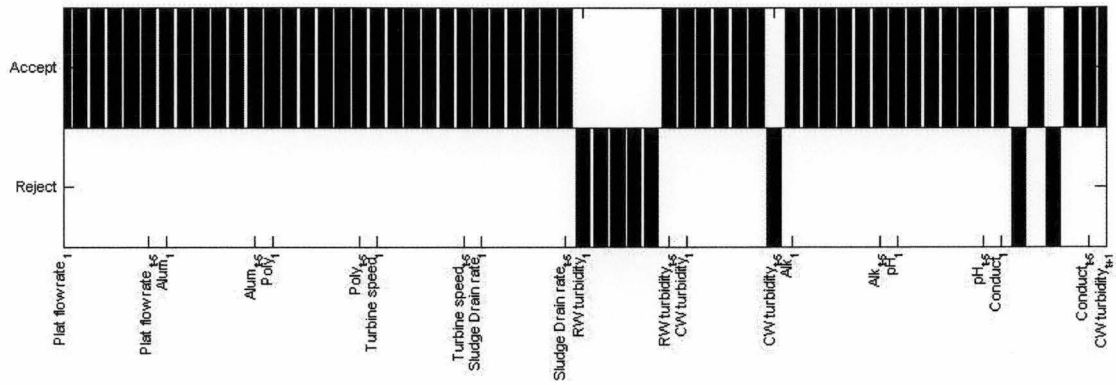
b) T-test null hypothesis of validation and training set



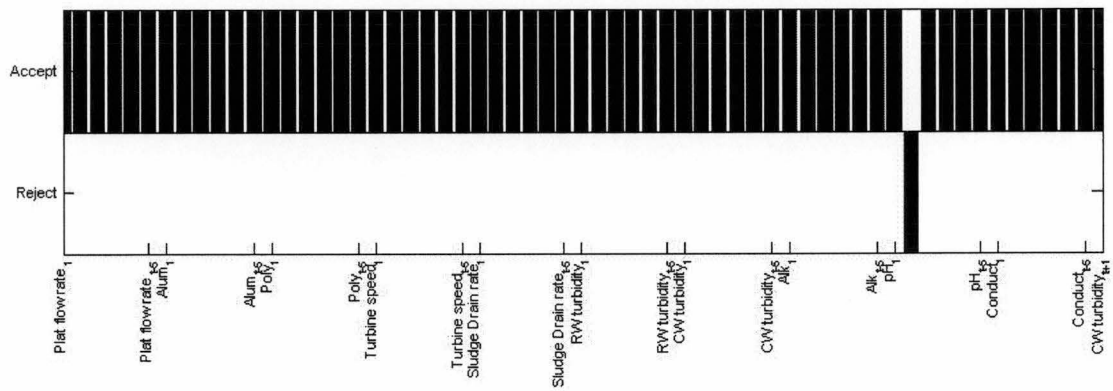
c) F-test null hypothesis of test and training set



d) F-test null hypothesis of validation and training set



e) Range test null hypothesis of test and training set



f) Range test null hypothesis of validation and training set

Figure 5.17 The results of the T-test, F-test and range null hypotheses for BKWTP case study; the variables with time lag is listed on the X-axis, and results of the hypothesise test are shown on the Y-axis:

(a) T-test null hypothesis of testing and training set

(b) T-test null hypothesis of validation and training set

(c) F-test null hypothesis of testing and training set

(d) F-test null hypothesis of validation and training set

(d) Range-test null hypothesis of testing and training set

(e) Range-test null hypothesis of validation and training set

5.4.3 Determination of Optimal Model Architectures

In the same manner as in the BEWTP case study, the optimising of model architecture aimed to find the number of hidden layers and their associated hidden neurons which gave the best predictive performance in the four hours of attaining clarified water turbidity. The predictive performance of the candidate models was measured using the test set. The associated modelling elements were shown in Table 5.10.

The predictive performance of each model was measured by its r^2 and MAE. On the other hand, model reliability was checked by comparison of percentage prediction accuracy with the measurement error of turbidimeter (10-15 NTU, Table 3.6).

In the same manner as in the BEWTP case study, the maximum number of layers was limited to 2 layers as discussed in section 4.3.5. Although it was recommended for use with a single hidden layer network only, the number of hidden neurons was guided by Equation 4.29 proposed by Roger and Dowla (1994): $N^H \leq \frac{N^{TR}}{N^I + 1}$. Specific to the BKWTP case study, the number of samples, N^{TR} was 3788 and number of inputs, N^I was 60. Therefore the recommended number of hidden neurons was 62 neurons. Systematic trial and error was performed over this recommended number of neurons by varying of number of hidden neurons over the range 5 to 100 neurons with increments of 5 neurons. It was important to ensure that the maximum number of hidden neurons was less than the number of training data set to avoid mathematically undetermined system (Sha, 2007). In this case, the maximum number of 200 neurons (100 neurons from each of the first and second hidden layers) was much less than the number of training set (3516 data sets) and it would keep ANN model away from being mathematically undetermined system. Finally, the systematic trial and error processes were conducted using the framework tabulated in Table 5.10. The resulting structure of the ANN process model for clarified water turbidity is shown in Figure 5.18

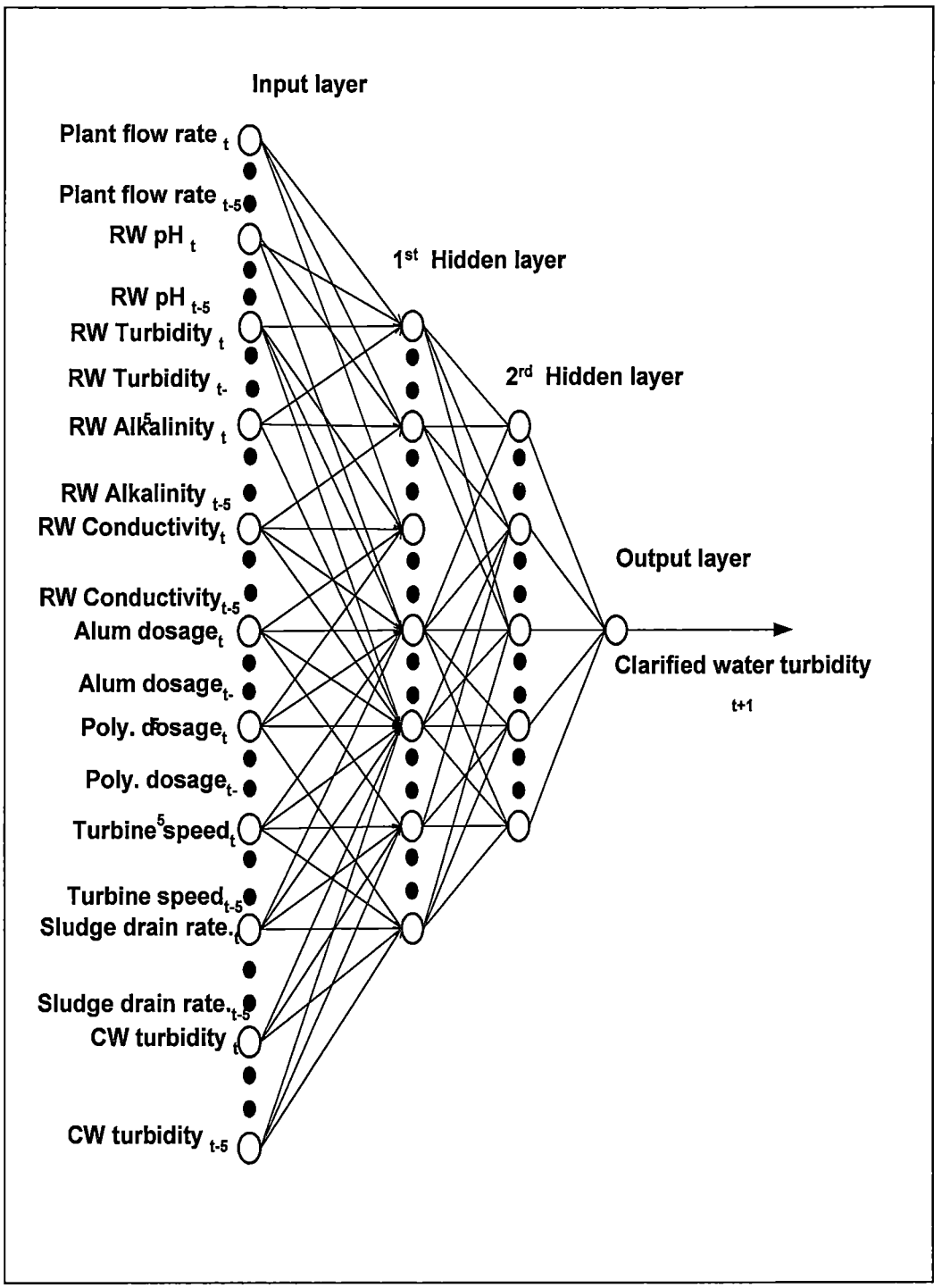


Figure 5.18 Structure of clarifier ANN models for prediction of clarified water turbidity for BKWTP: showing the inputs with temporal span and the output of the process model

Model elements	Network architecture (varied)
Temporal span	t, t-1, t-2, t-3, t-4, t-5
Maximum number of hidden layer	1 and 2
Number of neuron in 1st hidden layer	5 to 100 with 5 increment
Number of neuron in 2rd hidden layer	5 to 100 with 5 increment
Model elements	**Network architecture (fixed)**
Neural network type	Back propagation
Activation functions	• Hyperbolic tangent function for (1st & 2rd hidden layers) • Linear function for output layer
Learning algorithm	Adaptive learning rate with momentum
Stopping criteria and model validation	Early stopping
Model performance evaluation	Coefficient of determination ,(r^2) Mean absolute error (MAE) Mean percentage error (MPE)
Model output	Predictive clarified water turbidity at one time step ahead,(CW turbidity $_{t+1}$)
Prediction horizon	t+1, (one time step or four hour ahead)

Table 5.10 Model elements and architecture for BKWTP clarifier model development, showing which elements were varied and which were fixed in the trial and error development process

Optimal model architecture for the clarified water turbidity process model

The best performances of process model of each time lag from present time to 20 hours time lag for one and two hidden layers architecture are tabulated in Table 5.11, and graphically presented in Figure 5.19.

From Figure 5.19a and 5.19b, although the predictive performances of one and two layers architecture did not greatly differ, the optimal architecture was 20 neurons in the first hidden layer and 5 neurons in the second layer. The best performance model was the two hidden layers ANN model with temporal span of 12 hours according to empirical results at hand. Longer temporal spans (16 and 20 hours) were unfavourable to predictive performance even though they contained the information of 12 hours temporal span.

Temporal span of more than 12 hours might introduce noise to the process model resulting in deterioration of predictive performance. This model gave the best result of 0.71 for r^2 and smallest MAE of 0.65 NTU. In addition, it was noted that the performance of two hidden layer networks was superior to the one hidden layer network for every temporal span. This implied that single hidden layer network was not be able to recognise complicated patterns of the BKWTP operational data, it required more computational neuron to recognise the operational patterns.

In terms of model reliability, each of the models' reliability was acceptable since they all gave the mean percentage error similar to the lower bound of turbidimeter measurement error of 10 percent (Table 3.6) which could be seen in Figure 5.19c and the best model suggested gave a low percentage error of 10.73 percent. The predicting ability of this ANN process model will be discussed in the next section.

Temporal span (Hours)	No of neuron	R^2	Mean absolute error (NTU)	Mean percentage error (%)
Present	30	0.67	0.68	10.98
4 Hours	25	0.68	0.68	11.09
8 Hours	10	0.68	0.68	11.15
12 Hours	5	0.69	0.68	10.80
16 Hours	5	0.69	0.68	10.87
20 Hours	10	0.68	0.68	11.08

a) One hidden layer network

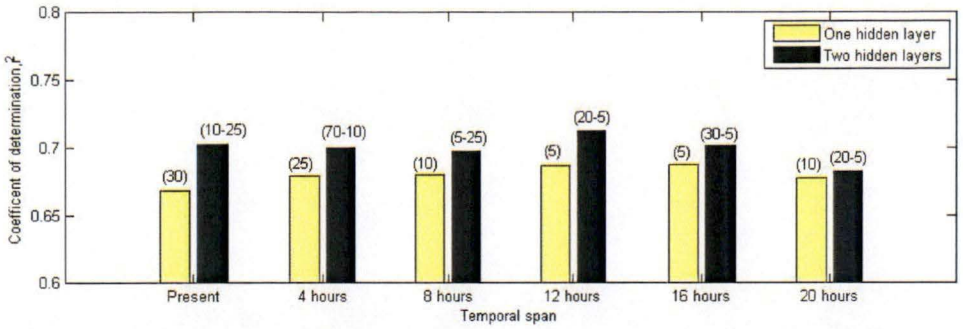
Temporal span (Hours)	No. of neuron 2nd Layer	No. of Neuron 3rd Layer	R^2	Mean absolute error (NTU)	Mean percentage error (%)
Present	10	25	0.70	0.67	10.67
4 Hours	70	10	0.70	0.66	10.90
8 Hours	5	25	0.70	0.67	10.84
12 Hours	**20**	**5**	**0.71**	**0.65**	**10.73**
16 Hours	30	5	0.70	0.66	10.63
20 Hours	20	5	0.68	0.65	10.61

b) Two hidden layer network

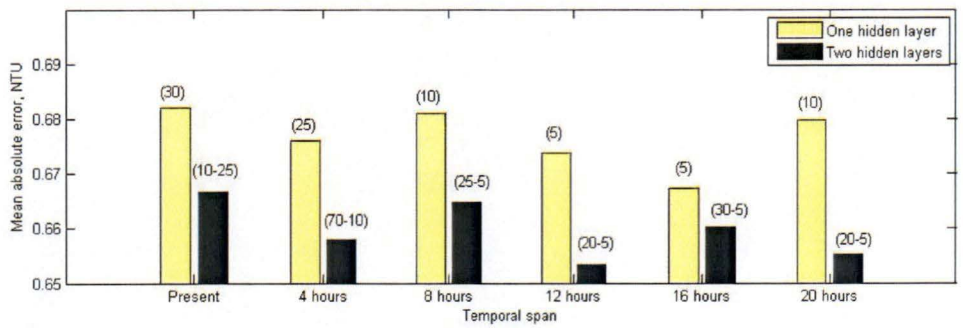
Table 5.11. The performance in prediction of clarified water turbidity for model with one and two hidden layers (BKWTP), showing the optimal model architectures and their predictive performances at each temporal span.

(a) One hidden layer network

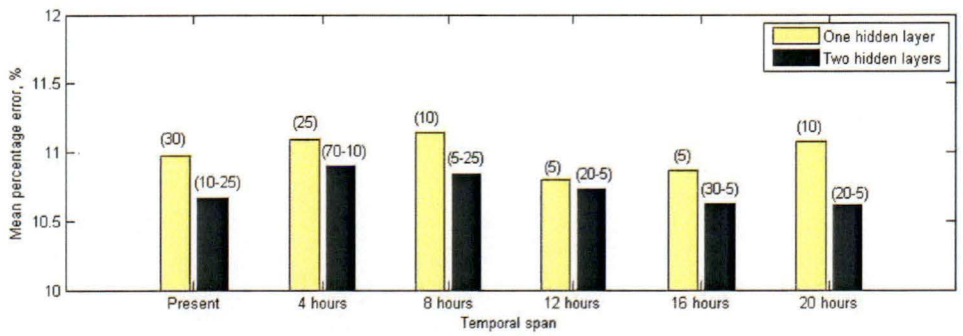
(b) Two hidden layer network set



a) coefficient of determination



b) Mean absolute error



c) Mean percentage error

Figure 5.19 Performance in prediction of clarified water turbidity for one and two hidden layer models with temporal span from present to 20 hours for BKWTP case study,

(a) Coefficient of determination (r^2) between predicted versus actual clarified turbidity used test set data

(b) Mean absolute error between predicted versus actual clarified turbidity used test set data

(c) Mean percentage error between predicted versus actual clarified turbidity used test set data

5.4.4 Performance of the Clarifier Process Model

The performance of the clarifier process model was measured via the test set. Similar to the BEWTP model, the predictive performance could be measured for four hours ahead of prediction (one time step ahead) and multiple step prediction. The flow of data for one time step and multiple time step prediction are shown in Figure 5.20. When the switch was turned to the "A" position, the process model worked in one time step prediction. None of predicted value of clarified water turbidity was inputted to the model. Therefore the model used only previous data to predict the clarified water turbidity.

On the other hand, if the switch was turned to the "B" position, the predicted values of clarified water turbidity were repeatedly inputted to the model. Now the model would use the predicted value and previous data to predict the clarified water turbidity in the following time step. The multiple time step predictions were limited to six time steps ahead (24 hours)

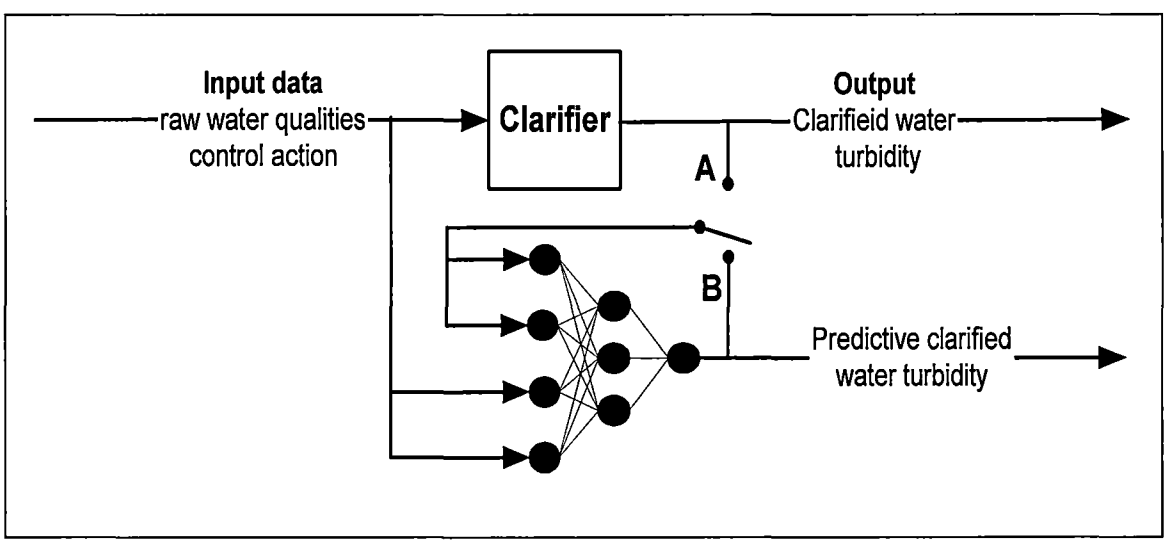


Figure 5.20 One and multiple steps predictions of clarified water turbidity (BKWTP), when the switch was turned to "A" position, it was for one step prediction, and "B" position, it was for multiple steps prediction

Clarified turbidity process model performance evaluation

- **One step predictive performance**

With input at the optimal temporal span of 12 hours, the process model with 20 neurons in the first and 5 neurons in the second hidden layers showed the best performance. The predicted and actual clarified water turbidity is shown in Figure 5.21. The overall ability

to predict was good with the r^2 of 0.71 and a small MAE of 0.65 NTU. The overall mean percentage error of 10.73 percent was similar to the level of the measurement error of the online turbidimeter (10 to15 percent, Table 3.6). Thus the process model was reliable.

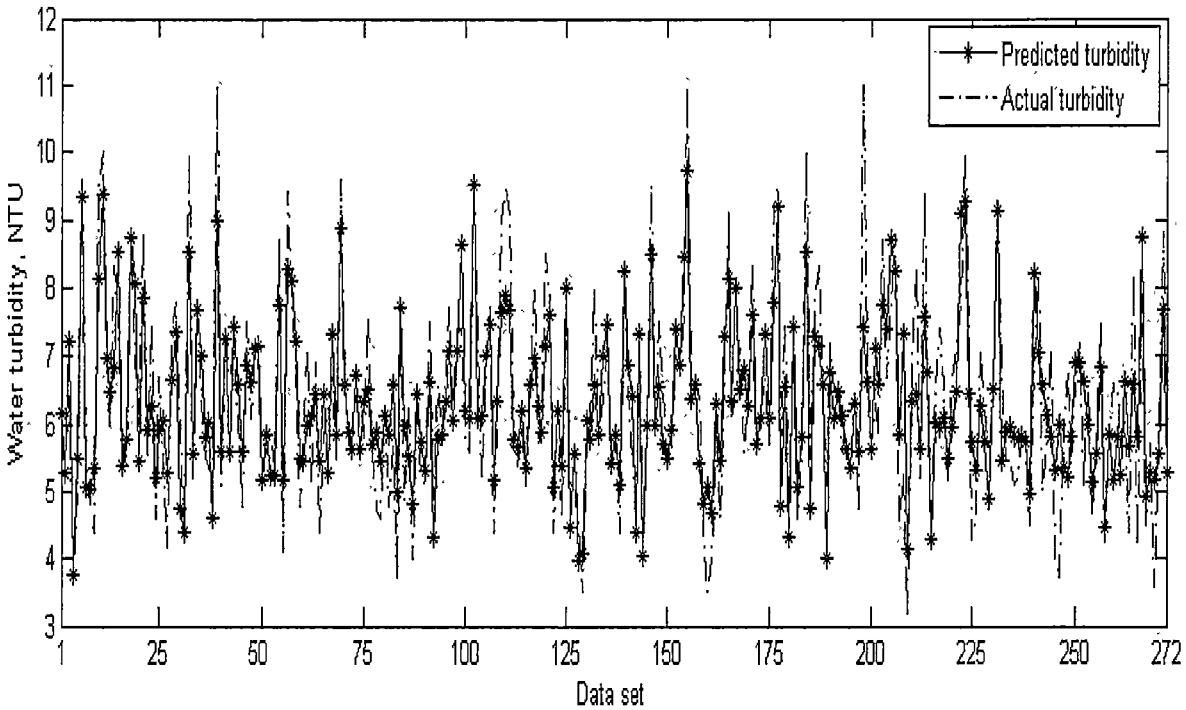


Figure 5.21 Predicted clarified turbidity, showing the predicted and actual clarified water turbidity of the test set data

Statistic parameters	Values
Coefficient of determination (r^2)	0.71
Mean absolute error	0.11
Mean percentage error	10.73

Table 5.12 Performance of clarified water turbidity prediction (BKWTP), showing performance parameters when the process model predicted the clarified water turbidity in four hours ahead using test set data

To assess model reliability from the input domain due to the raw water turbidity that was closely related to the clarified water turbidity, the bar plot of raw water turbidity and the predictive percentage error is shown in Figure 5.22. This figure was drawn in order to investigate any effects on percentage error when the raw water turbidity domain was enlarged and to investigate the consistency of the percentage error. The predictive percentage error was slightly reduced from 12.47 to about 11 percent when the raw water

turbidity increases to about 100 NTU. The percentage error became constant at about 10.75 percent for any further increasing of raw water turbidity upwards from 100 NTU. Thus this clarified water turbidity process model was reliable since its percentage error was closely similar to turbidimeter measurement error (i.e. 10 to 15 NTU, Table 3.6). .

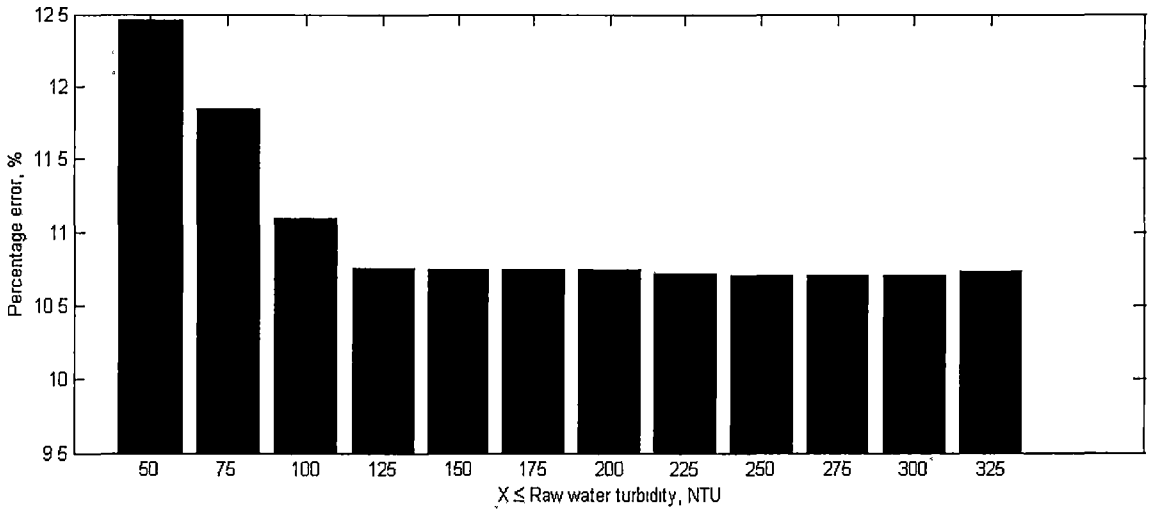
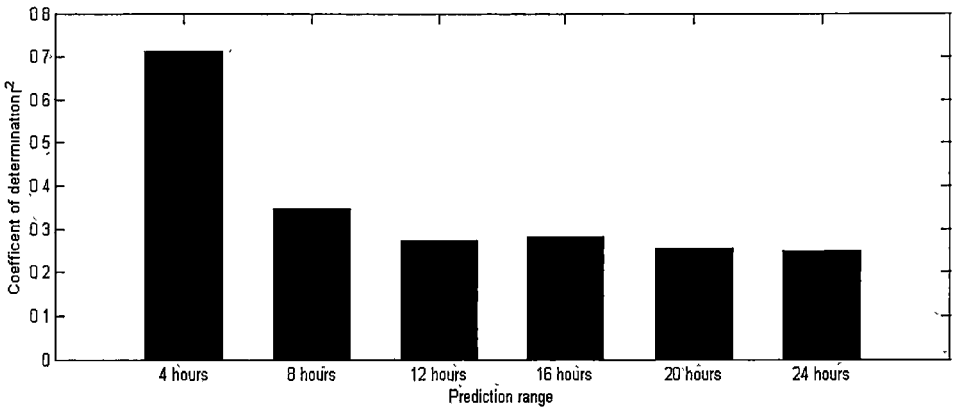


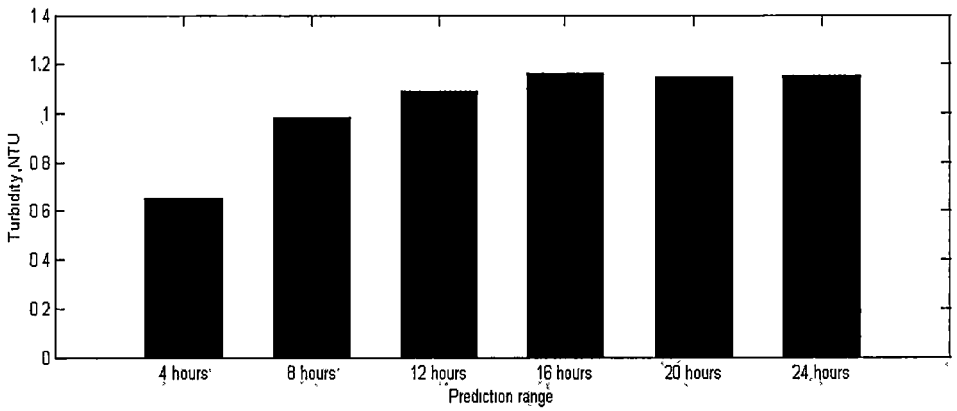
Figure 5.22 Reliability of clarified water turbidity model (BKWTP): showing predictive percentage error value when the raw water turbidity increased

- **Multiple steps prediction performance**

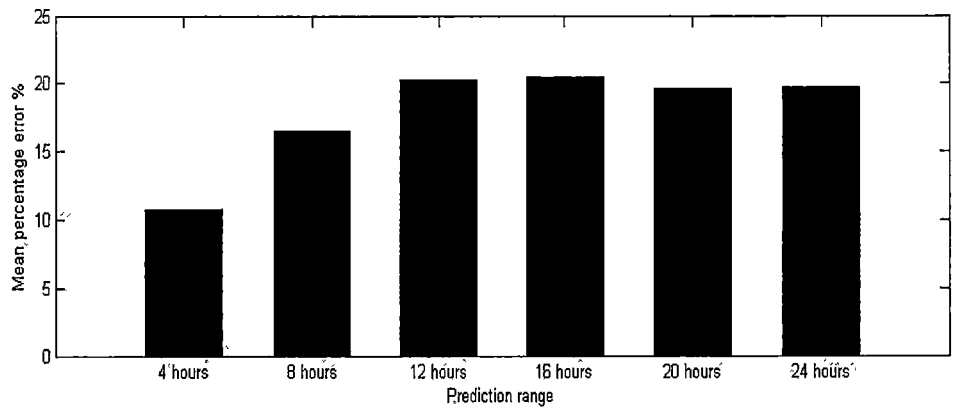
The multiple steps clarified turbidity was predicted by up to 6 time steps (24 hours). The bar plots of r^2 and MAE of each time step were shown in Figure 5.23a and Figure 5.23b respectively. The reduction of r^2 and the increase of MAE by about 50 percent after first step prediction implied the existence of the deterioration of predictive ability. Reliability also suffered from the range of prediction being extended. The evidence for this lay in the increase of percentage error of 10.73 to about 20 percent. After the first time step (four hours), the predictive performance fell dramatically since the accumulation of predictive errors occurred when operating the ANN process model recursively.



a) Coefficient of determination (r^2)



b) Mean absolute error



c) Mean percentage error

Figure 5.23. Multiple steps prediction performance of clarified water turbidity (BKWTP) from four hours to 24 hours ahead using test set data

(a) Comparison of coefficient of determination (r^2)

(b) Comparison of mean absolute error

(c) Comparison of mean percentage error

5.5 Chapter Conclusion

In the case of the BEWTP, the optimal architectural network for clarified water turbidity prediction was the network with 35 for the first layer and 5 neurons for the second hidden layer. On the other hand, the best performance network for water colour prediction was a network with 95 neurons in both the first and second hidden layers. The optimal temporal span in both cases was 2 time lags or 8 hours from the present. Both the clarified water turbidity and colour models could track the trend in the data very well resulting in large r^2 of 0.89 and 0.88 respectively. The associated small MAEs of (0.11 NTU and 0.50 HU) indicated good accuracy in both process models.

The clarified water colour model reliably predicted clarified water colour since the mean predictive error (MAE) of 0.50 HU was less than the associated water colour measurement error of 1 HU. On the other hand, the clarified water turbidity model reliability was marginally acceptable since its percentage error lied at in the same level as the associated measurement error of 10-15 percent.

Prediction by longer than one time step (four hours) for both clarified water turbidity and colour were not recommended since the errors increased by more than any of the related measurement errors.

In the case of the BKWTP, the optimal architectural network was the network with 20 and 5 neurons for first and second hidden layer respectively. The optimal temporal span was 3 time steps (12 hours). The model performance was inferior to the BEWTP. The data trend was recognised in a moderate degree with r^2 of 0.71 and accuracy with MAE of 0.65 NTU. This process model showed a promising degree of reliability since its mean predictive percentage error of 10.73 percent equalled to measurement error of the turbidimeter (10 to 15 percent).

No additional step beyond one single step of four hours was recommended since, after one step, the percentage error would increase above the turbidimeter measurement error and therefore the process model would become unreliable.

The performance of the BKWTP clarifier compared unfavourably with that of BEWTP clarifier model since the quality of the data obtained from the BEWTP was more reliable.

This was because the data received from sensors (from a SCADA system) at BEWTP were electronically sensed and the data were recorded every 30 minutes, rather than manually recorded only every four hours as in case of the BKWTP. Therefore, the data inputted to the BEWTP clarifier model were likely to contain less error and omissions and be more representative of how the clarifier behaved. In addition, the operational data of BKWTP varied in greater range than that of BEWTP as it already discussed in Chapter3. Therefore with low data quality and greater variation range, this was the cause of poor prediction performance of clarifier process model of BKWTP.

The BKWTP clarifier required a longer suitable temporal span (12 hours) than the BEWTP clarifier (8 hours) to reach its best performance. Any difficulties or how well the operators operate were also recorded in the period of temporal span. Although a longer temporal span may be able to provide more information, the BKWTP clarifier model will never perform better than the BEWTP one. This was because sludge concentrations and water temperature data were not available for use as inputs for the BKWTP clarifier model. These variables have been known to be important indicators to density current phenomena, which were known to effect clarifier performance especially in the case of high ambient temperature of BKWTP.

For both of these case studies, the optimal temporal spans and architectures of process models and their performance evaluation were the main concerns of this chapter. The results of this chapter show that these models may be useful in modelling clarifiers. The real test for performance of these models will be when used in a case study involving intelligent control of the full scale pilot plant test, and the results are compared to control by human operators. This is the subject of the next two chapters.

6.0 CLARIFIER INTELLIGENT CONTROL: GENETIC ALGORITHM APPROACH

6.1 Introduction

ANN clarifier process models were developed in the previous chapter. This chapter describes the development of the method for use of these models in control of a simulated clarifier. Several model based control strategies are reviewed and compared. ANN Model Predictive Control (MPC) architecture is selected due to its superiority to all others. This combines with an evolutionary optimiser that enables the control system both to handle multi-objective optimisation and gives the flexibility to add operational constraints. The optimiser's evolutionary search facility adds an intelligent feature to this control scheme. For clarity, the background of evolutionary search is described before the control simulations are performed on selected test sets derived from BEWTP and BKWTP case studies (i.e. the same test sets as used in Chapter 5). The performance of the intelligent control system is evaluated and then compared with that of the human operator's. The outcome of these simulations will be used for the full scale pilot plant test in Chapter 7

The idea of using an intelligent MPC for the ANN process model with an evolutionary optimiser to control the clarifier (as presented in this chapter) is novel. It is an original idea to use these methods and combine the methods in this way. Novelty is also involved in the development of an evolutionary optimiser to handle multi-objective optimisation of the clarifier control. It is also the first time that control actions are numerically optimised with respect to water quality and operational cost.

6.2 Review of ANN Control Strategies

There are a number of ANN control strategies which are based on various types of process models. Most of these control designs employ a forward and/or an inverse model. These nonlinear model control strategies are well established and often benefit from the ability to incorporate improved robustness more directly in the controller design. Narendra and Parthasarathy (1990) suggested that many of these model based control strategies employ neural network models and thus benefit from the nonlinear approximation properties of ANNs. The following section reviews the main types of

ANN model based control structures and begins with the problem of identifying inverse process models that utilise ANN model based control strategies.

6.2.1 Inverse Modelling

Inverse modelling plays a crucial role in a range of control structures. The ANN models considered so far are called *forward models* since the direction of information flow through the model is from input to process output. Conversely, the direction of information flow through an inverse process is opposite to that of the forward model. Consequently, an *inverse model* predicts the control action. The objective is to use the inverse model to formulate a controller (Hunt & Sbarbaro, 1995).

Conceptually, the simplest approach is direct inverse process modelling as shown in Figure 6.1. The process input, including a manipulated variable, is applied to the process. The process output is then used as input to the ANN controller (inverse model). The ANN controller output (the predicted input) is compared with the process input and used as the error, ε to train the ANN model. This will clearly force the ANN model to represent the inverse of the process.

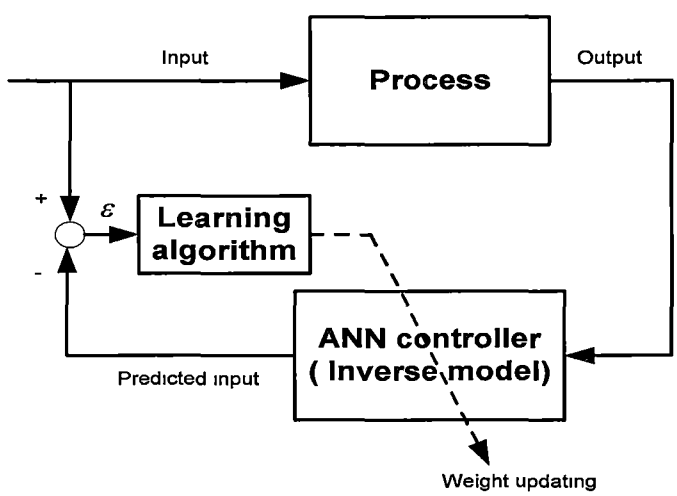


Figure 6.1 Direct inverse modelling

However, this approach is disadvantaged by the need to specify the manipulated variable input signal since it may be difficult to define the operational process input as a priori. Thus, the direct inverse modelling structure is not "goal directed" since the process output cannot be explicitly forced to cover the desired operating range (Irwin et al., 1995; Moâsciânski & Ogonowski, 1995).

The root of the problem with direct inverse modelling is that the training of the inverse model attempts to minimise error in the predicted manipulated input variables and this does not correspond to the control objective which is to minimise the process output error. An alternative method, specialised inverse modelling, (Figure 6.2), overcomes this by using the process output error, or alternatively a forward process model output error (when the system is noisy), to generate the error signal. The error, ε between the setpoint and the process model output is passed back through the forward model to give the manipulated variable error which is used to train the inverse model. Hence specialised inverse modelling is goal directed since it is based on the error between desired system output and actual output (Psaltis et al., 1988).

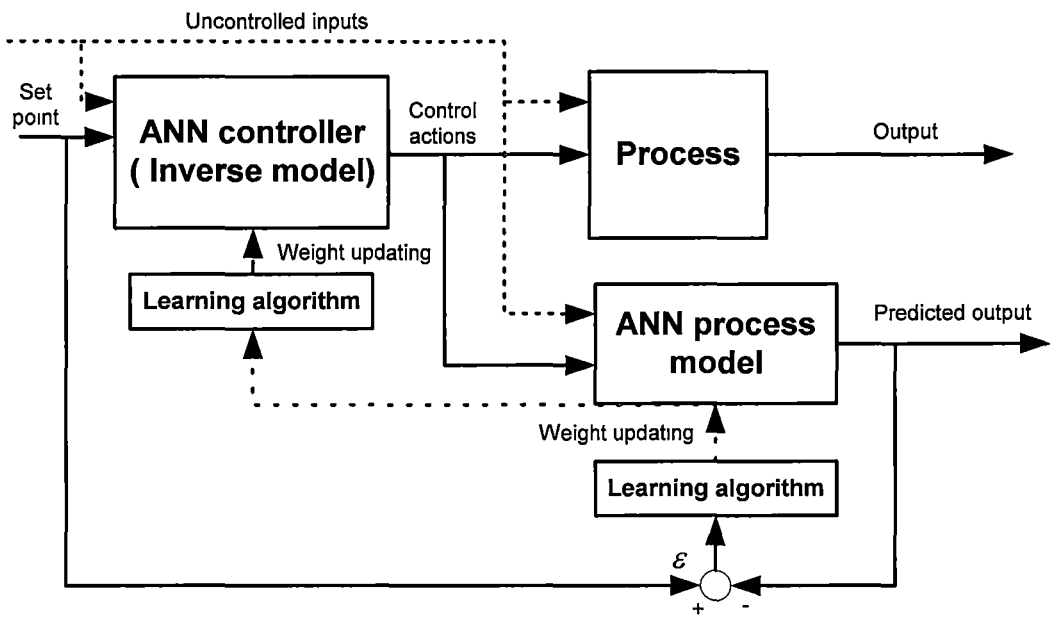


Figure 6.2 Specialised inverse modelling

There is a potential problem associated with developing an inverse model of a nonlinear process. For instance, some classes of nonlinear system cannot be inverted (Economou et al., 1986) or the nonlinear model performs a many to one type mapping from input to output (so there is no unique inverse solution). In some textbooks, it is referred as an ill posed problem (Swingler, 1996). The nonlinear system must have a unique (i.e. one to one) forward mapping if the inverse is to be unique. If this is not the case then the incorrect inverse model may result and the associated model based control system will be unstable. Therefore his situation should be avoided (Nahas et al., 1992).

The simplest use of an inverse process model for control is to use it as a controller by placing it at the front of the process so that the composite system results in an identity mapping from set point to process output (Figure 6.3). This approach is called *Direct inverse control.* Although it is applied in a number of applications, the absence of feedback results in a lack of robustness for the practical case of an imperfect inverse model (Irwin et al., 1995).

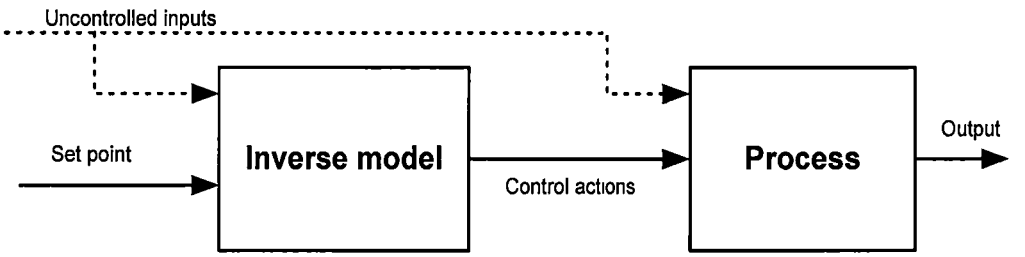


Figure 6.3 Direct inverse control

There are several well established ANN model based control structures incorporating inverse models which have robustness superior to that of direct inverse control. Some of these are discussed in the following section.

6.2.1 Internal Model Control

The use of neural networks in the internal model control structure has been proposed by several researchers (Bhat & McAvoy, 1990; Hunt & Sbarbaro, 1991; Hunt et al., 1992). This has been implemented for the control of simulated processes under assumptions of an open loop stable system as well as for the perfect forward and inverse models (Hunt et al., 1992; Nahas et al., 1992).

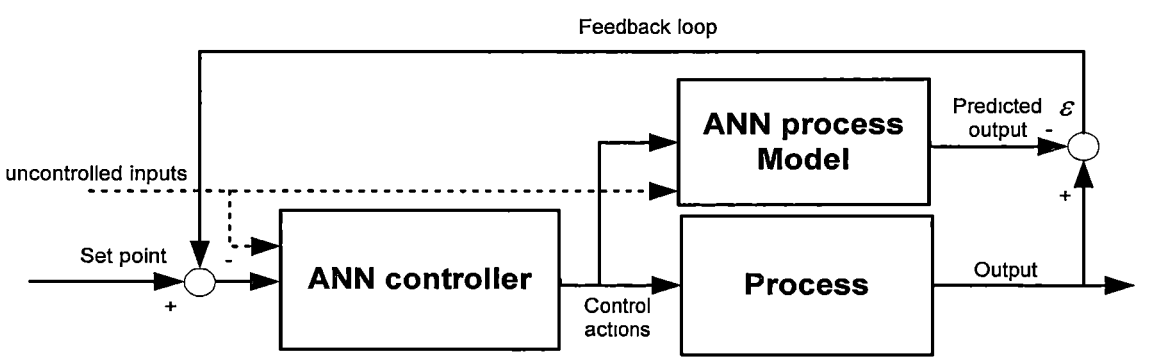


Figure 6.4 Internal model control

In Figure 6.4, the control structure uses both a forward and an inverse ANN process model. The forward model is placed in parallel with the process. The difference between process output and model output (predicted output) is used as the feedback signal, and this is passed to the ANN controller (i.e. ANN inverse model). The internal model controller is generally designed to be an inverse process model (when it exists). Many theoretical stability results relating to internal model control loops are available (Hunt & Sbarbaro, 1991) although they generally make assumptions on the open-loop stability of the system, exact modelling and /or inverse modelling. In spite of these assumptions, it is claimed that this approach readily extends to nonlinear systems and yields to robustness and stability analysis (Harris, 1994; Hunt et al., 1992)

6.2.2 Model Reference Control

Model reference control has been widely used in linear adaptive control application. This control system is composed of two models, one is for process model and the other one is the controller. The process is initially modelled by the process model and the controller is trained so that the process output follows the reference model output. Narendra and Parthasarthy (1990) have suggested that it could be used for nonlinear control by employing artificial neural networks. The model reference control structure for this use is shown in Figure 6.5. It incorporates a plant reference model (i.e. forward model) and an ANN controller. The control objective is to force the process output to asymptotically track the output of the reference model. The error between the reference model and the process output is used to train the inverse process model acting as the controller.

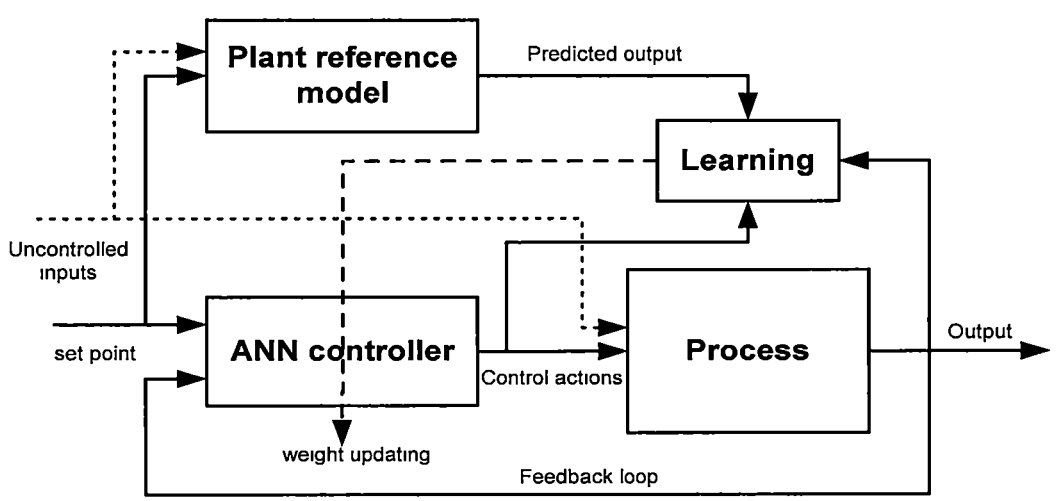


Figure 6.5 Model reference control

The performance of model reference control is highly dependent on the choice of a suitable reference model and the derivation of an appropriate learning rule. Harris (1994) reported that simple gradient-based learning rules were sometimes insufficient and there is no reason why this should not also be the case for more general nonlinear process models and controllers. For the case of a reference model which based on the identity mapping, the controller learns the inverse process model. However, in practice, the reference model is chosen to have be dynamically stable and these result in the ANN controller being a "detuned" inverse model (Hunt et al., 1992). Applications of ANN model reference control are not common although some simulation studies have been demonstrated (Hassibi & Loparo, 1991; Narendra & Parthasarathy, 1990).

6.2.3 Model Predictive Control

Model Predictive Control (MPC) is one of the most important control technologies for the process industries in general. Traditionally, MPC is used in linear control. A MPC control system is composed of the process model and optimiser. The process model replicates the process used to predict the outputs of the process. Initially, this control system uses the process model to predict future plant performance. The optimiser (i.e. the controller) then calculates the control actions that will optimise the process performance over specified future time horizon. It has been postulated that MPC can be implemented on nonlinear control schemes as surveyed by (Henson, 1998; Rawlings et al., 1994).

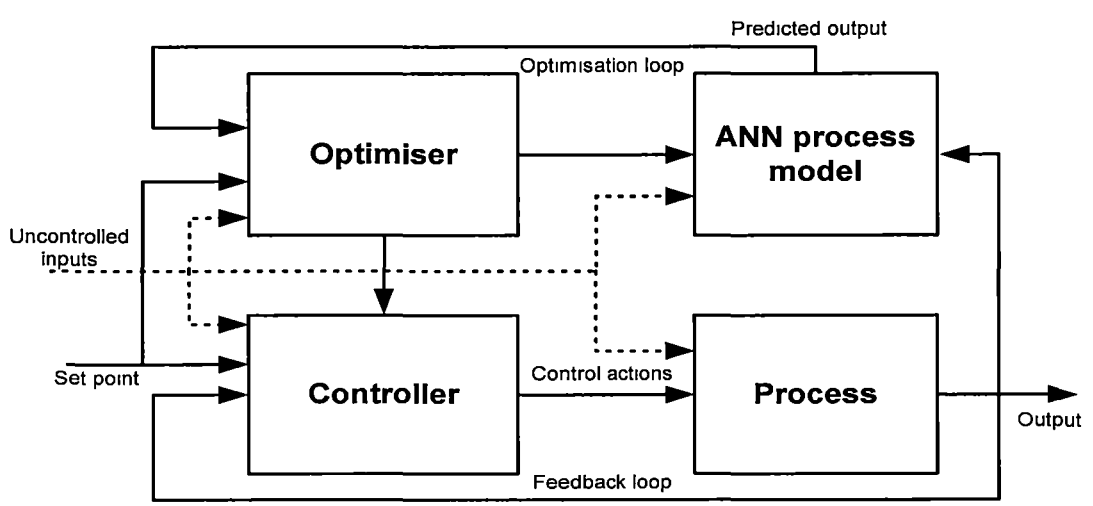


Figure 6.6 Model predictive control

Figure 6.6 shows an artificial neural network MPC scheme which, in contrast to the other control strategies reviewed so far, does not incorporate an inverse process model. A prime characteristic that distinguishes MPC from other control systems is the idea of a receding horizon. At each time step, control actions are determined to achieve the desired behaviour in the following time steps. This idea also has appeal because it relates to many of the control tasks that simulate a human being at work on a daily basis (Norgaard et al., 2000). In this approach an artificial neural network model provides a prediction of the process's future over a specified horizon of time. The prediction supplied by the ANN model is passed to a nonlinear optimisation routine which attempts to minimise a specified performance criterion. Generally, the control action is chosen to minimise the quadratic performance criterion which compromises between the tracking error (i.e. the error between the set point and the predictive output) and the control objective subject to the constraints. For a general objective function this can be shown as follows (Baughman & Liu, 1995; Mills et al., 1996; Moâsciânski & Ogonowski, 1995):

$$J = \sum_{i=N_1}^{N_2} (\hat{y}(t+1) - r(t+i))^2 + \lambda \sum_{i=1}^{N_3} (u(t+i-1) - u(t+i-2))^2 \qquad (6.1)$$

subject to $u_{\min} \le u(t) \le u_{\max}$,

Where J is objective function, r and $\hat{y}$ are the set point and the predictive output respectively; u is the set of manipulated variables (i.e. control actions) bounded between their maximum value, $u_{\max}$ and minimum value, $u_{\min}$; The predictive horizon, N_1 and N_2 are defined as minimum and maximum the number of future time steps for which the process model will give the prediction results; the term N_3, control horizon is defined as the number of future time steps for which the MPC system will provide the set of control action. λ denotes a weight factor which will penalise any changes in the control.

One useful feature of MPC is that the control actions are optimised. The optimiser can also take account of process constraints which is an important practical feature of MPC since all real processes are subject to some constraints. Typically, the optimisation is solved subject to hard constraints (boundaries) on the manipulated variables and some soft constraints on process outputs which may have to be violated in order to form a feasible optimisation problem.

If the process model can predict the process outputs properly and the performance function and the optimiser are appropriately chosen, this control scheme can provide excellent closed loop control. In practise, the multi-step ahead optimisation may result in some error amplification since in every single step of the multiple step prediction, since the error is enlarged by accumulation of predictive error. (Harris, 1994).

In the light of these points, the Model Predictive Control (MPC) architecture has been chosen for the clarifier control since MPC architecture is superior to all other architectures, and a MPC system is goal directed and is flexible to the addition of operational constraints.

6.3 Overview of Genetic Algorithms

Genetic algorithms (GAs) are stochastic search methods use for optimisation. They are based on analogies to the mechanics of natural selection and genetics (Darwin's theory of survival), and combine "survival of the fittest" among "generation" of chromosome structures with structured yet randomised information exchange incorporated into a search algorithm (Goldberg, 1989). In each generation, a new set of chromosomes is created by selection, crossover and mutation processes involving their fitness. GAs are not simple random walks but they efficiently exploit historical information in finding new search points, to find solution that give progressively for each generation during optimisation (Goldberg, 1989; Zalzala & Fleming, 1997).

In this section, classical GAs and their main operations are introduced for single objective optimisation. The main consideration of this section is how to apply classical GAs to multi-objective optimisation problems. The basic operations of classical GAs will then be merged to the multi-objective optimisation problem via one of the most commonly used algorithms, Multiobjective Genetic Algorithm (MOGA).

6.3.1 Operating Principle of Genetic Algorithms

Genetic Algorithms are one of the best tools for quickly searching for an approximate global maximum or minimum (Haupt et al., 2004). GAs require the parameter set of the optimisation problems, which are encoded as chromosomes, over some finite alphabets. Parameters can be encoded by binary, ternary, integer, real-valued and other (Zalzala & Fleming, 1997). After coding of parameters, the parent pools are generated and assessed

for their fitness values through the associated objective function. The chromosomes that have their values in the top ranks of objective value are chosen to perform GA operations for the next generation. Successively generated populations (offspring pool), which have been expected to improve over time, can be yielded by a set of simple operations illustrated by three operators involving nothing more complex than copying and switching chromosome: (i) selection, (ii) Crossover and (iii) Mutation (Goldberg, 1989). These iteration processes are consecutively executed until a predefined stopping criterion is satisfied or the convergence condition is reached (Zalzala & Fleming, 1997). The stopping criteria can be activated as soon as one of the following conditions are satisfied: (i) the maximum number of generation is reached; (ii) the best acceptable solution is obtained; or (iii) the best solution cannot be improved by increasing the number of generations (Haupt et al., 2004). The simple genetic algorithm procedure is illustrated in Figure 6.7.

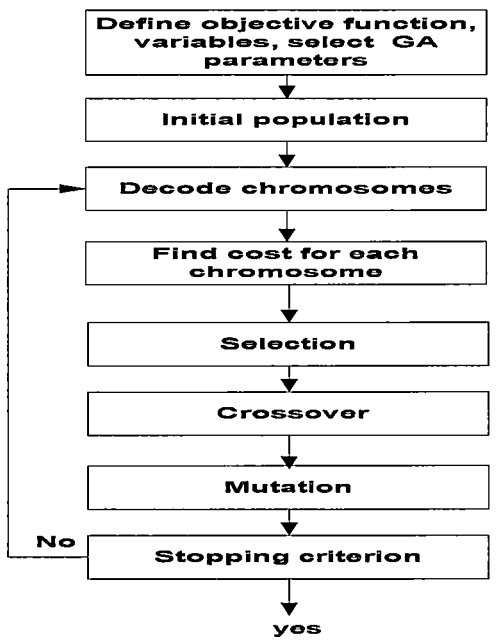


Figure 6.7 Genetic algorithm process, illustrating a schematic diagram of a classical genetic algorithm and its operations

Selection

In some textbooks, selection is referred to as reproduction (Haupt et al., 2004). The selection process begins with identifying "good" (usually above-average) solutions in a population, and then making copies of chromosomes based on their fitness (i.e. objective function values). This process means that the chromosomes with better fitness values have a high probability of creating offspring in the next generation (Goldberg, 1989).

There are three commonly used operators for this process which are the roulette wheel, stochastic universal sampling and tournament selection (Goldberg, 1989; Haupt et al., 2004).

- *Roulette wheel selection* is one of the most used chromosome selection techniques. This selection implements the concept of the roulette wheel. The individual will have a roulette wheel slot size proportional to their fitness. The chromosomes are selected as the results of spinning the wheel N times (N is number of selections required), meaning that the higher probability of chromosomes to go to the next generation depends on the better fitness values.
- *Stochastic universal sampling* is a single-phase sampling algorithm which has a zero bias and short time performance. Instead of spinning the roulette wheel N times, this method spins only once. Stochastic universal sampling uses N equally spaced pointers to select the individual whose fitness spans the position of pointers.
- *Tournament selection*, the tournaments are played among a few individuals which are randomly selected from the population and the best individual form this group is selected as parent for the crossover process. In practice, the number of tournaments is predefined, the best solutions are those that win the entire tournament and the worse solution will loss all the tournaments and will be eliminated from the population.

The first two selection methods can be categorised as *proportionate selection methods*. The assigned fitness is proportional to the objective value (Goldberg & Deb, 1991). However, it has a scaling problem. The outcome of these selection methods depends on the true value of the fitness, instead of the relative fitness values of the population members. For instance, if all populations have more or less the same fitness values so that every solution has a similar probability of selection, it leads to a single assignment of each solution. This phenomenon is equivalent to not performing the selection operation at all. If the selective pressure is defined as the ratio of the probability that the most fit chromosome is selected as a parent to the probability that average chromosome is selected, then proportionate selection method cannot ensure selective pressure. On the other hand, if in the population, one solution with large fitness has raw fitness compared with the other, it will be selected for crossover since its possibility of being selected would be close to one (based on proportionate selection method). It would lead to

dominating solution and to loss of diversity. However this scale problem does not exist when the tournament selection method is applied (Deb, 2001).

In order to circumvent both of these difficulties, Goldberg (1989) suggested that raw fitness should be ranked by comparing with each other solutions and sorted according to their relative fitness, from the worst (rank 1) to the best (rank N), (N is number of population). Thereafter, the proportionate selection operators will be applied with the ranked fitness value but not raw fitness. This method is called rank selection method and it will pay significant rule in the multi-objective optimisation.

Crossover

Crossover (occasionally called recombination) when it is used with other encoding methods from binary encoding (Haupt et al., 2004) is the operator that produces new chromosomes, which is performed after selection process. It pairs up two parents and creates two new offspring that inherit both parents. The genes for the offspring will be inherited just like those of the parents, except that some genes that came from the father in the first offspring will come from the mother and some of those inherited from the mother will come from the father. However, to reserve some good chromosomes from the selection process, some of chromosomes in the population are not used in a crossover. If a crossover probability of p_C is used, then $100\,p_C$ percent of chromosomes in the population are used in the crossover process and $100(1-p_C)$ percent of the population are simply copied to the new population. The process of crossover can be described as follows (Goldberg, 1989):

- An integer position k, is randomly selected in the two-parent (father and mother) chromosome length, l the value is between 1 and $l-1$.
- Two offspring are generated from parent chromosome by switching all characters from position $k+1$ to l of both parents.

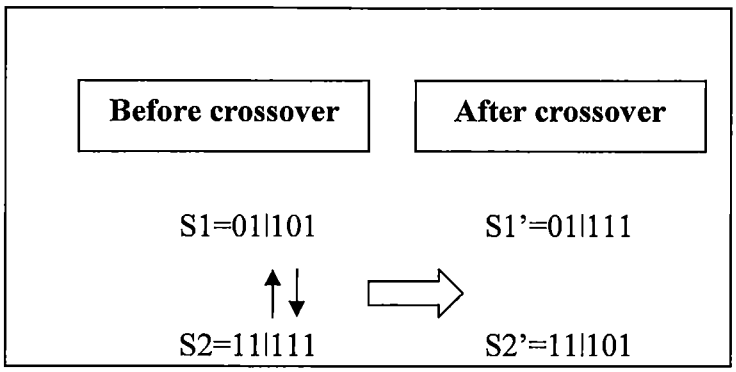


Figure 6.8 Simple crossover operation (single-point crossover)

Various forms of crossover can be presented as follows:

- *Simple crossover operation* (single-point crossover) is illustrated in Figure 6.8 which supposed that the random position along chromosome length, k is 2 (represented by a vertical line). Two offspring, $S1'$ and $S2'$, are created from two parents (father, $S1$,and mother, $S2$).
- *Double-point crossover* operates upon the same idea as single-point crossover, but two random positions are generated and chromosome will be switched between the points.
- *Multi-point crossover*: the randomly selected point can be varied from 1 to l-1, where l is the length of the chromosome. They are chosen at random with no duplication and sorted into ascending order. The genes are then exchanged between the selected positions while the rest will not be exchanged.
- *Shuffle crossover* is modified from single-point crossover. Both parents are randomly shuffled before the chromosomes are exchanged. The offsprings are unshuffled after recombination.

Mutation

Mutation is an operator representing the natural evolution. It is a random process with a low probability that one gene of chromosome is replaced by another to generate a new genetic structure (Zalzala & Fleming, 1997). Its role is to provide a guarantee that the search algorithm is not trapped on a local minimum (Haupt et al., 2004). The sequence of selection and crossover operations may stagnate at any homogeneous set of solutions. The solution might appear to become optimal, or rather locally optimal, only because the search algorithm is not able to proceed any further. Mutation aids the searching to avoid

loss of diversity (Negnevitsky, 2005). Figure 6.9 illustrates the mutation operation when the position of the fourth gene of the parent ($S3$) is randomly selected.

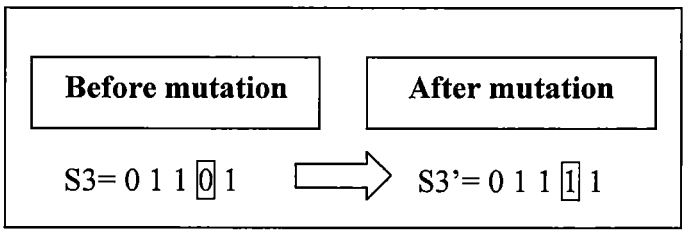


Figure 6.9 Mutation operation

6.3.2 Multiple Objective and Genetic Algorithms

Most optimisation models involve single objective problems where the best solution can easily be obtained by comparing the objective function values. A major difference between single and multiple objectives is the definition of the optimum concept. Multi-objective optimisation considers several objective functions simultaneously and does not have merely a single justification of the optimal solution. Generally, one objective cannot be sustained without being at expenses of the other objectives (Haupt et al., 2004).

In the past, most multi-objective optimisation problems traditionally dealt with scalar objective functions, which apply the weights into the variable of the function and sum them up into one variable to be optimised. Physically, this approach transforms the multiple objectives to a single objective. However finally, the decision makers have to decide in order to obtain a single unique solution. Coello Coello et. al (2002) classified the multi-optimisation problem by the way of solution searching and decision making:

A priori Preference Articulation (decide⇒ search): The decision maker selects the weighting before conducting the optimisation algorithm. In practice, it is the method that combines the objective functions into a scalar cost function (linear or nonlinear combination). This effectively converts a multi-objective problem into a single objective one.

A Posteriori Preference Articulation (search⇒ decide): there are no specified weightings before or during the optimisation process. The optimisation gives a set of efficient candidate solutions from which the decision maker may select.

Progressive Preference Articulation (decide⇔search): The decision maker interacts with the optimisation program during the optimisation process by using his/her judgement upon whether or not to change the weight of objective function, while the system provides an update of the solution.

The first and second approaches are very popular. There are two major techniques for the definition of multi-objective optimisation problem which can be described as follows (Coello Coello et al., 2002):

- *Weight sum of objective function*: The multi-objective problem is converted into a single objective one (scalar application) by a using weighted sum of the objective functions as a single representative objective function. Then the problem will be solved as a single objective one. It implies for a *priori* preference articulation.

- *Pareto optimisation*: The multi-objective problem is solved by applying the Pareto optimisation approach. The optimal solution is selected from the resulting Pareto-optimal set (non-dominated solutions) which can be considered as vector applications. It is representative for a *posteriori* preference articulation.

The best weight to apply is very difficult to determine and multi-objective problems have no unique answer. The most efficient way to solve the multi-objective optimisation problem is by using Pareto optimisation. Therefore, a set of non-dominated solutions is needed to be determined from all feasible solutions (Deb, 2001; Goldberg, 1989).

A non-dominated solution is the point at which the value of one of the objective functions cannot be improved without degrading others (Haupt et al., 2004). The non-dominated solutions, called the Pareto-optimal set, define the Pareto-optimal front (Deb, 2001; Fwa et al., 2000). The Pareto-optimal front can be easily described as the set X which is Pareto optimal, if there is no other set Y dominating the set X with respect to a set of fitness values. In addition, set X dominates set Y if X is better than Y at least once and is not worse with respect to all other objective functions. An example of a Pareto-optimal front of a minimisation problem is shown in Figure 6.10.

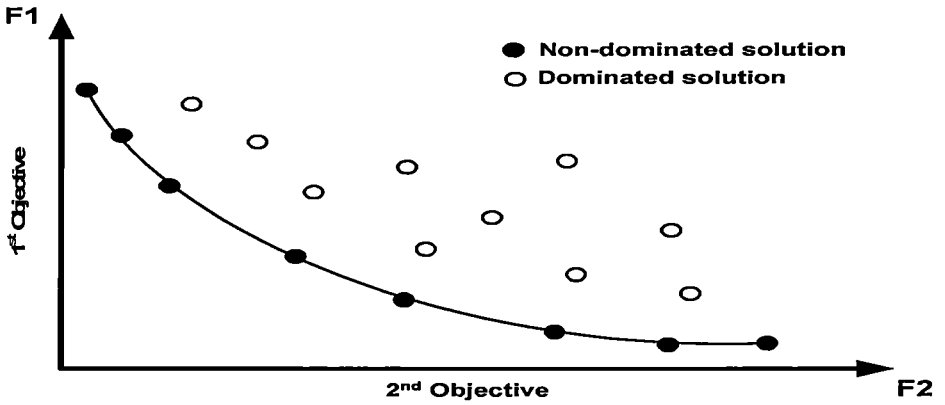


Figure 6.10 The Pareto-optimal front, showing dominated solutions and non-dominated solutions which were on the Pareto front of a minimisation problem

The Pareto-optimal set can rarely be obtained with conventional optimisation because (i) it is incapable of finding out whether the obtained solution can really be optimal, and (ii) the solution tends to be trapped by the first local minimum or maximum. Genetic Algorithms (GAs) are good at overcoming the problems resulting from conventional optimisations. GAs locate the global optimal solutions because they have the ability to preserve the populations and look for Pareto-optimal solutions. GAs utilise constraints for fitness consideration and selection of nondominated solution of the multi-objective optimisation problem (Cheng & Li, 1997; Haupt et al., 2004).

In a multi-objective optimisation problem, after Pareto-optimal solutions are provided, the decision maker will select the best solution. Different decision makers with different preferences may select different solutions. Therefore, it is important to find as many Pareto-optimal solutions as possible so that the decision maker can choose the preferred one. Therefore, there are two goals in Pareto-optimal solutions (Coello Coello & Lamont, 2004; Deb, 2001):

- To find a set of solutions as close as possible to the true Pareto-optimal front.
- To find a set of solutions as diverse as possible.

6.3.3 Multiobjective Genetic Algorithm (MOGA)

The consideration of multiple objectives in evolutionary based search algorithms has received much interest. A number of algorithms have been proposed including the Multiobjective Genetic Algorithm (MOGA) (Fonseca & Fleming, 1993), Niched Pareto Genetic Algorithm (NPGA) (Horn et al., 1994), Nondominated Sorting Genetic

Algorithm (NSGA) (Srinivas & Deb, 1994) and Strength Pareto Evolutionary Algorithm (SPEA) (Zitzler & Thiele, 1999). These algorithms and variations thereof comprise the basis of most popular multi-objective evolutionary algorithms currently in use. To compare performance between algorithms, Coello Coello and Lamont (2004) reviewed a number of different applications. Most of these algorithms were tested against selected functions and it was found that their performances were problematic.

One of the most commonly used multi-objective methods based on Pareto optimisation is the Multiobjective Genetic Algorithm (MOGA). It was introduced by Fonseca and Fleming (1993). They were the first to suggest a multi objective GA which explicitly emphasises the non-dominated solution and simultaneously maintains its diversity. Rather than using raw fitness values, the MOGA differs from a standard GA in the way fitness is assigned relatively to each solution in the population. This is called Pareto-optimal rank strategy and uses techniques such as fitness sharing which aid in spreading the solution along the Pareto-optimal front. The rest of the algorithm (selection, crossover and mutation) is employed in the same way as that in classical GAs (Deb, 2001). The schematic diagram of The MOGA is shown in Figure 6.11

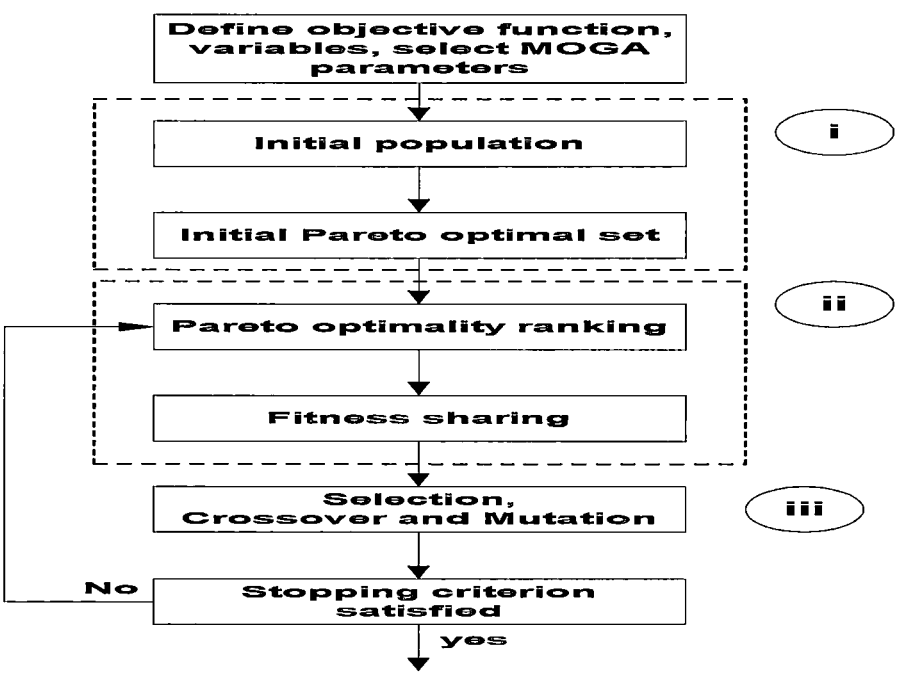


Figure 6.11 Multiobjective Genetic Algorithm (MOGA), showing that the three steps of the MOGA algorithm: Step (i) Initial population and initial Pareto-optimal set, Step (ii) Pareto optimality ranking and Fitness Sharing and Step (iii) Selection, Crossover and mutation

The three main modules in the process are numbered in Figure 6.11. These are: (i) production of the initial generation and establishment of an initial Pareto-optimal set, (ii) application of the MOGA technique of finding the Pareto-optimal set, and (iii) evolution by selection, crossover and mutation operators and revision of the Pareto-optimal set. For each generation, MOGA firstly determines the fitness function of individuals of the previous generation and then generate chromosome by selecting two parents on the basis of their fitness and reproducing them by crossover and mutation until the whole population is recreated. Finally, MOGA decodes and evaluates the chromosomes of this new generation and revises the Pareto-optimal set of the previous generation. This procedure is repeated many times until either: (i) the maximum generation number is reached or (ii) the convergence and diversity index are satisfied or not further improved. The first termination criterion is necessary to prevent a run of excessively long time. The second termination criterion is important to check the convergence and diversity of the optimisation procedure.

Module 1 Initial population and initial Pareto-optimal set

In the same manner as in classic GAs, the initial population can be established by randomly choosing a population in a predefined range. After the associated objective functions are determined, a search space can be formed.

Module2 Pareto optimality ranking and Fitness Sharing

The nearly identical convergence and bunching of solutions at optimum regions is called genetic drift phenomenon. It misses the inheritance of parent's best solution, and can be misleading to the global optimum goals (Cheng & Li, 1997). In order to retain the diversity of the solution and ensure the global optimum, Pareto optimality ranking and fitness sharing should be employed to solve these weaknesses (Cheng & Li, 1997; Goldberg, 1989). Rank-based fitness assignment is a prerequisite for applying fitness sharing. Fonseca and Fleming (1993) explained an excellent theoretical approach and a numerical example was clearly illustrated by Deb (2001). Two main steps concerned with Pareto-optimal ranking and fitness sharing are described as follows:

- **Pareto optimality rank based fittest assignment**

Rather than using a raw fitness value, the fitness is assigned to the solution in proportion to their raw fitness value. By MOGA, each population in the search space is ranked according to the following expression (Fonseca & Fleming, 1993):

$$r_i = 1 + n_i, \qquad (6.2)$$

where r_i presents the rank of the i^{th} solution. n_i is the number of the solutions, which dominate the solution i. Obviously, the minimum rank is 1 (for non-dominated solutions with respect to population) and the maximum rank is no more than the size of the population, N. It is also clear that the ranking procedure may not assign all possible ranks (between 1 to N). Additionally, the smaller r_i is the better the solution will be (Deb, 2001).

After ranking, the population is sorted according to its ranks and the associated raw fitnesses (i.e. objective function values) are assigned to a solution by using a linear mapping function (which is normally linear but not necessarily so) from the best solution to the worst solution. Usually the mapping function is chosen so as to assign fitness between N (for the best rank solution) and 1 (for the worst-rank solution). Thereafter, the raw fitnesses of the solutions within the same rank are averaged. This average fitness is now called the assigned fitness to each solution of the rank (Fonseca & Fleming, 1993).

Rank based fitness assignment overcomes the scale problem of the proportionate fitness assignment. When the selective pressure is too small (i.e. all the raw fitnesses are more or less the same value) where selection has caused the search to narrow down too quickly this will result in no solution being generated and an excessive number of offspring. The average fitness by ranging base fitness assignment introduces a uniform scaling across the population and provides a simple and effective way of controlling the selective pressure. For the special case of linear mapping, the assigned fitness of the i^{th} solution, f_i. is given as follows (Deb, 2001):

$$f_i = N - \sum_{k=1}^{r_i - 1} \mu(k) - 0.5(\mu(k) - 1). \qquad (6.3)$$

where $\mu(k)$ is the number of solutions in rank r_i and the assigned fitness is now a function of number of the dominated solution, not the objective functions.

- **Fitness Sharing**

In order to maintain population diversity and avoid genetic drift, a fitness sharing technique (Goldberg & Richardson, 1987) is usually applied in multi-objective evolution

algorithms. The basic idea is to reduce the reproduction ability of a solution crowded by many solutions through degrading its assigned fitness value f_i using the sharing function concept. Typically the following function is defined as the sharing function $Sh(d_{ij})$ of two solutions i and j :

$$\text{If } d_{ij} < \sigma_{share}, Sh(d_{ij}) = 1 - \left(\frac{d_{ij}}{\sigma_{share}}\right)^{\alpha} \qquad (6.4)$$

Otherwise, $Sh(d_{ij}) = 0$

where, α is a constant and if α =1 is used, the effect linearly reduces from one to zero. It has been found that α does not have too significant an effect on the performance of sharing function and is commonly set to one (Fleming & Pashkevich, 2007). σ_{share} is the niche radius, which represents the minimal distance between two solutions desired by the user. The parameter d_{ij} is the distance between any two solutions i and j which is defined as:

$$d_{ij} = \sqrt{\sum_{k=1}^{M}\left(\frac{f_k^{(i)} - f_k^{(j)}}{f_k^{\max} - f_k^{\min}}\right)^2} \qquad (6.5)$$

where $f_k^{\max}$ and $f_k^{\min}$ are defined as maximum and minimum objective function values of k^{th} objective.

Therefore, from equation (6.4), if $d_{ij} = 0$, then sharing function, $Sh(d_{ij}) = 1$. This means that a solution has full sharing on itself. If $d_{ij} \geq \sigma_{share}$, then $Sh(d_{ij}) = 0$. This means two solutions which are at least σ_{share} distance away from each other do not have any sharing effect on each other. For any other distance, d_{ij} two solutions have a partial effect. A niche count, nc_i which gives an estimation of the extent of the crowding near the i^{th} solution, is then calculated as follows:

$$nc_i = \sum_{j=1}^{N} Sh(d_{ij}). \qquad (6.6)$$

The sum includes the i^{th} solution itself, thus nc_i is always greater than or equal to one since at least $Sh(d_{ij})=1$. Finally, the shared fitness value $f_i^{'}$ for i^{th} solution is calculated as follows:

$$f_i^{'}=\frac{f_i}{nc_i}. \qquad (6.7)$$

Thus, if the i^{th} solution does not have any sharing effect on any other solution in the population, namely $nc_i=1$, its fitness value will not be degraded. Otherwise, the sharing function will degrade fitness according to the extent of crowding near the i^{th} solution.

The choice of σ_{share} has a significant impact on the performance of MOGA (Fonseca & Fleming, 1993). It is very difficult to determine a suitable fixing value of σ_{share} since it is problematic. It is recommended to obtain it empirically by a trial and error process (Obayashi et al., 1998). However, Fonseca et al (1993) suggested a dynamic updating strategy for σ_{share}. The value of σ_{share} gives a simple estimation of σ_{share} in the objective function space with an adaptability feature as:

$$N\sigma_{share}^{q-1}-\frac{\prod_{i=1}^{q}(M_i-m_i+\sigma_{share})-\prod_{i=1}^{q}(M_i-m_i)}{\sigma_{share}}=0. \qquad (6.8)$$

Where N is a population size, q is number of objectives, M_i and m_i are maximum and minimum values of each objective, respectively.

The suitable σ_{share} is updated in every generation and it will be adaptive to the population during the evolutionary process. Clear numerical presentation of this strategy is demonstrated by Deb (2001).

Module3 Selection, Crossover and mutation

After the shared fitness $f_i^{'}$ is calculated for each population, the typical GA operations are applied. Regardless of the selection method, any population with higher shared fitness has a higher chance to select for crossover and mutation operations. In this work, the stochastic universal selection (with shared fitness), the single-point crossover and the bit-

wise mutation operators are applied to create the new offspring. After, these three operations and if the stopping criteria are satisfied, then the MOGA will stop evolution of solutions and the last set of solutions (at the Pareto-optimal front) are assumed to be optimal.

6.3.4 The Assignment of MOGA Parameters

The success of GAs depends on the specification of several parameters such as the population size, the number of generations, together with the probabilities of crossover and mutation. Goldberg (1989) has suggested that good GA performance requires a choice of high crossover and low mutation probabilities and a moderate population size. However, it is difficult to predict how the various parameters interact with each other. Indeed there is no general theory for selecting optimal GA parameters and this is always a problem. Therefore trial and error processes with the problem at hand should be used as guidelines to find the suitable parameters (Grefenstette, 1986).

Population size and number of generations

The population size and number of generations affect both ultimate performance and the efficiency of GAs. It is a very sensitive parameter for the GA process (Halhal et al., 1997). De Jong (1975) suggested that with small population size, the initial performance can be improved. However a large population size improves long-term performance and the optimal number of populations should be in the range of 50 to 100 individuals. GAs generally perform poorly with a very small population because the population provides an insufficient sample size and will converge rapidly to a local optimum. A large population, on the other hand, requires more evaluations per generation and results in unnecessary computational cost and unacceptably slow convergence. McKinney and Min-Der (1994) reported that a population of 50 to 100 individuals taken through 10 to 20 generations have a high probability of finding an optimal or near optimal solution. However the best value is problem dependent and requires trial and error to verify it.

Crossover rate

The crossover rate controls the frequency with which the crossover operator is applied. The higher the crossover rate, the more quickly new chromosomes are introduced into the population. If the crossover rate is too high, high performance structures are discarded faster than selection can produce improvement. If the crossover rate is too low, the search

may stagnate due to the lower exploration rate (Eshelman & Schaffer, 1991). Negnevitsky (2005) suggested that a value of 0.7 for crossover probability generally gives good results.

Mutation rate

The need for mutation is to keep diversity in the population. When using selection and crossover, some potentially useful genetic material might be lost. Its role is to guarantee that the search algorithm is not trapped at a local optimum (Holland, 1992). The mutation operator protect against such irrecoverable premature loss. A mutation probability is a very small number typically in the range of 0.001 to 0.01 (Negnevitsky, 2005). De Jong (1975) originally suggested using a mutation probability inversely proportional to the population size.

6.3.5 Pareto Front Quality

It is well recognised that there are two goals when solving a multi-objective optimisation problem: (i) the convergence to the Pareto-optimal set, and (ii) the diversity of solutions in the Pareto-optimal set, since these two goals are distinct. The first goal requires search *towards* the Pareto-optimal solution, while the second goal requires a search *along* the Pareto-optimal front (Deb, 2001).

Deb and Goldberg (1989) use a chi-square-like deviation measure to assess diversity of solutions. Zitzler and Thiele (1999) used the maximum *spread* for diversity assessment. The maximum spread is defined as the Euclidean distance between the extreme solutions of each objective. The larger the maximum spread, the more diversity there is. However convergence measuring is not possible since the true Pareto-optimal front is unknown in this method. Therefore, only a relative assessment of convergence can be made as opposed to an absolute comparison. For example, Reed and Goldberg (2003) used the percentage of change in the number of non-dominated solutions as the termination criterion.

The Pareto Front Quality Index (PFQI) was introduced in (2003) by Kazancioglu et al.. It is a relative global assessment approach. It can assess the convergence and diversity of the solution simultaneously, regardless of the value of real Pareto-optimal front being unknown. Before assessments are taken, all the objective values are normalised by their

associated maximum and minimum values. Thus PFQI assesses the quality of the Pareto-optimal front in a normalised scale of range zero to one.

PFQI qualifies solutions in the Pareto-optimal front in terms of (i) their closeness to the utopia point, d_C which is defined by the average distance from every solution on the Pareto front (in a normalised scale) , (ii) the range of the Pareto front, d_R defined as the summation of the distance between the extreme point and the closest solution on the Pareto front and (iii) the evenness of the spread, d_E which can be measured as the maximum distance to between two neighbouring solutions in the Pareto front. From this formulation, the lesser the value of PFQI, the better quality of Pareto fronts is. Consequently, The Pareto Front Quality Index (PFQI) is defined as:

$$PFQI = d_C + d_R = d_E \tag{6.9}$$

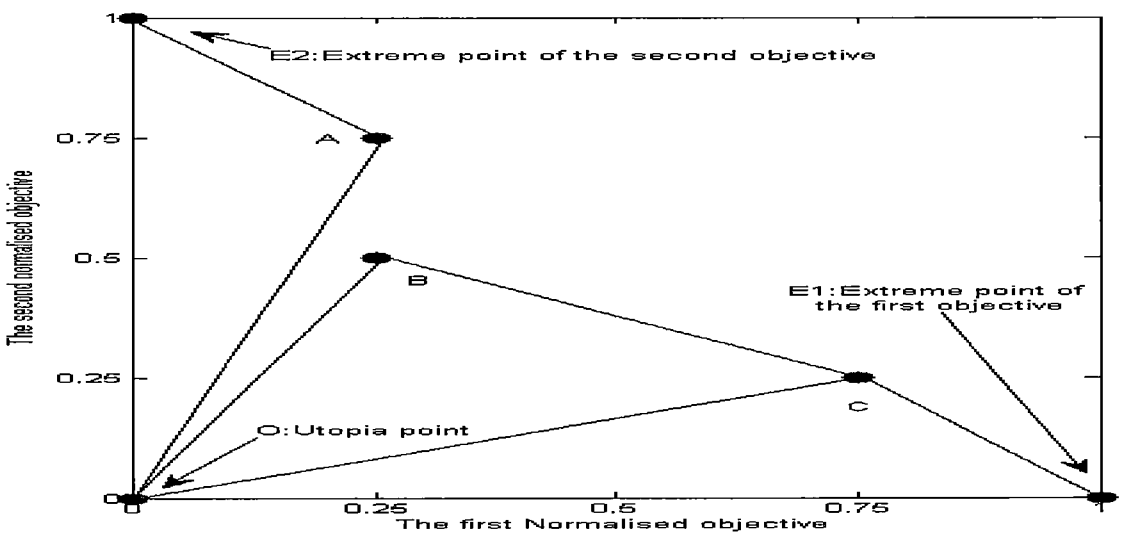


Figure 6.12 Pareto front quality index, demonstrating Pareto front within normalised space of two objective functions, A , B and C are the solutions on the Pareto front.

For clarity, as an example of minimisation problem of a two objective function, consider that the solutions A, B and C are in a two-dimensional normalisation objective function space. The coordinates of point A, B and C are (0.25, 0.75), (0.25, 0.50) and (0.75, 0.25) respectively. The origin O is the utopia point and two unit vectors (1, 0) and (0, 1) correspond to the extreme points in the first (denoted as E1) and second objective (denoted as E2), respectively. The PFQI can be obtained as follows (Kazancioglu et al., 2003):

d_C: to measure how close the Pareto-optimal front lies to the utopia point (i.e. the convergence of the solutions), the average distance of points A, B and C (on the Pareto-optimal front) from the utopia point is:

$$d_C = \frac{OA + OB + OC}{3} \approx 0.71 \tag{6.10}$$

d_R: to measure the diversity of the solution and the range of the solution, Point A and C are the closest to the extreme points E_1 and E_2. The average closest distance to both extreme points, d_R is given by:

$$d_R = \frac{E_1C + E_2A}{2} \approx 0.35 \tag{6.11}$$

d_E, the uniformity of the solutions in the Pareto-optimal front is how well the Pareto-optimal front spreads, and can be measured by using d_E, the closest distance of each single point on the Pareto-optimal front. In Figure 6.12, for point A, point B is the closest. For point B, certainly point A is closest. For point C, point B is the closest one.

$$d_E = \max(AB, BA, BC) = BC \approx 0.56 \tag{6.12}$$

Therefore, the PFQI value for the Pareto-optimal front in Figure 6.12 is given as:

$$PFQI = d_C + d_R + d_E = 0.71 + 0.35 + 0.56 = 1.62 \tag{6.13}$$

The general form of PFDI with a mathematical description was also provided in the work of Kazancioglu et al.(2003), it could be used when more than two objectives are considered.

6.3.6 Goal Programming and Constraint Handling

Goal programming methods have been used in innumerable applications in engineering (Romero, 1991; Steuer, 1986). Since goal programming attempts to find one or more solutions which satisfy a number of goals to the extent possible, therefore it is different in concept from nonlinear programming or optimisation techniques. Instead of finding a solution which absolutely minimises or maximises objective functions, goal

programming finds a solution that, if possible, satisfies a set of goals, or otherwise violates the goals minimally. This makes the approach more appealing to practitioner compared to conventional optimisation techniques. In goal programming, each goal is converted into an objective function of the minimising the different between the goal and its target. Generally, the conversion process depends on the type of goals used. They are presented in the Table 6.1 (Deb, 2001).

Type	Goal	Objective function
$\leq$	$f_j(x) \leq t_j$	Minimize $\langle f_j(x) - t_j \rangle$
$\geq$	$f_j(x) \geq t_j$	Minimize $\langle t_j - f_j(x) \rangle$
$=$	$f_j(x) = t_j$	Minimize $\lvert f_j(x) - t_j \rvert$
Range	$f_j(x) \in [t_j^l, t_j^u]$	Minimize $\max\left(\langle t_j^l - f_j(x) \rangle, \langle f_j(x) - t_j^u \rangle\right)$

Table 6.1 Goal Programming using multiple objective optimisation

For the bracket operator, if the operand is positive, it returns the value of the operand. Otherwise it will return zero. For an absolute operator, it returns an absolute value of the operand. In this way, goal programming is converted to a multiobjective optimisation problem. The advantages are: (i) no additional constraint for each goal is needed, and (ii) objective functions can be simply used since GAs do not require objective functions to be differentiable.

6.4 The Decision Maker: Shortest Normalised Distance Selection

It is possible to have more than one "best" solution from a Pareto-optimal set. In order to select only one optimal solution, the shortest normalised distance method is employed. It is the smallest Euclidean distance from the theoretical best optimal solution found in the normalised objective space (Fwa et al., 2000). In order to clarify this selection procedure, an example of two-objective minimisations is illustrated in Figure 6.13(a). The optimal solutions A and B are determined where A is the minimum solution of the first objective, and B is the lowest solution of the second objective respectively.

In Figure 6.13(b), The origin point in the normalised space, M is a utopian point defined as the theoretical best optimal solution (Fwa et al., 2000). The selected best solution is

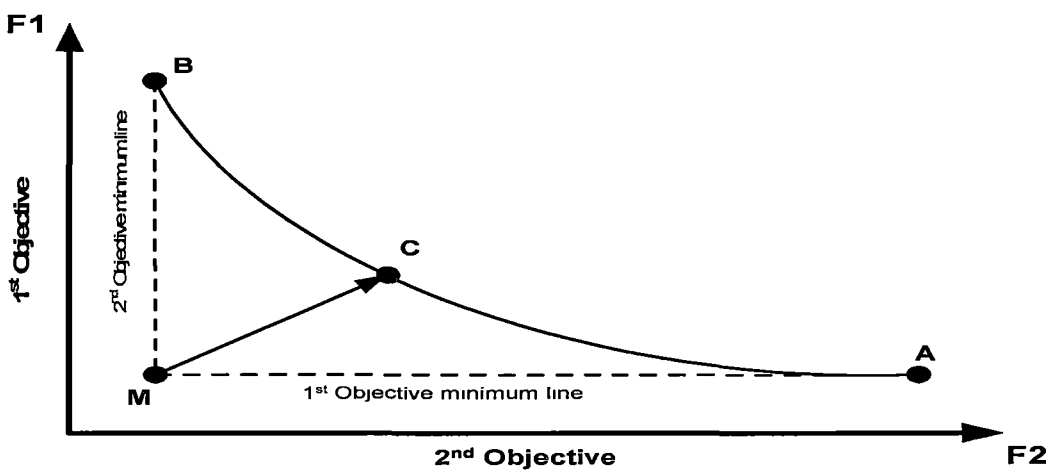
the solution that has the nearest point with respect to point M in terms of normalisation distance. The normalised objective function value is determined over a scale of 0 to 1 for each objective based on the following expression (Fwa et al., 2000):

$$\tilde{o}(i,x) = \left(\frac{o(i,x) - o(i,\min)}{o(i,\max) - o(i,\min)} \right) \tag{6.14}$$

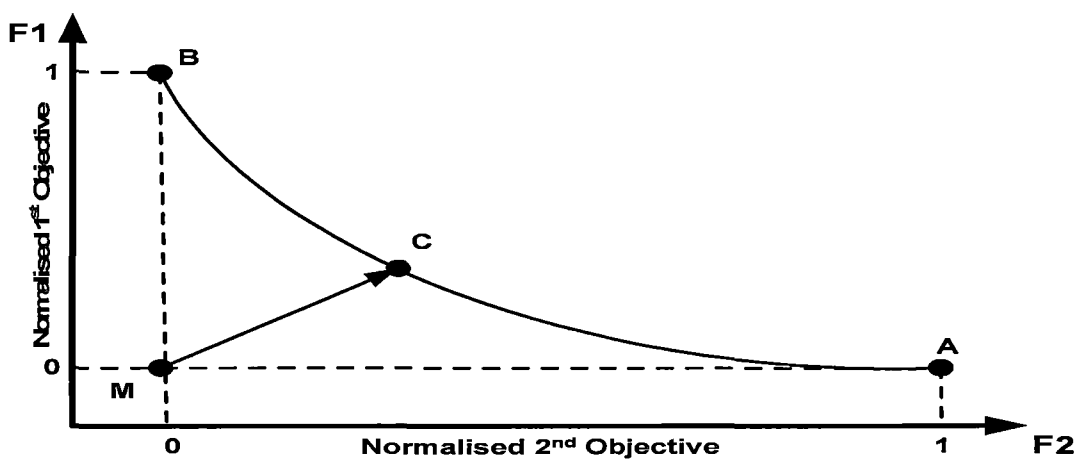
$\tilde{o}(i,x)$ = normalised parameter value of solution x of objective i

$o(i,\min)$ = minimum parameter value of objective i

$o(i,\max)$ = maximum parameter value of objective i



(a) Establishment of reference points



(b) The shortest distance from utopia point M,

Figure 6.13 Illustration of the shortest normalisation distance in the case of minimisation: (a) Establishment of reference points; (b) The shortest distance from utopia point M, MC in normalised scale

Two objective function values are transformed to the normalisation scale shown in Figure 6.13 (b). The best solution is performed by detecting the shortest normalised distance from point M and d_m is defined as the normalised distance to the population in the Pareto-optimal front presented as follows:

$$d_m = \sqrt{\sum_{i=1}^{n} \left(\tilde{o}(i,x) - \tilde{o}(i,\min)\right)^2}, \qquad (6.15)$$

Where; n=number of objectives under consideration.

Specifically, the short distance in Figure 6.13(b) is shown as MC which is selected on the basis of the minimisation of two objectives simultaneously.

6.5 Clarifier Intelligent Control Simulations

The clarifier control simulations were conducted using the proposed "*intelligent control system*". The intelligent control system consisted of three components: (i) ANN clarifier process model, (ii) intelligent optimiser based MOGA and (iii) decision maker as shown in Figure 6.14.

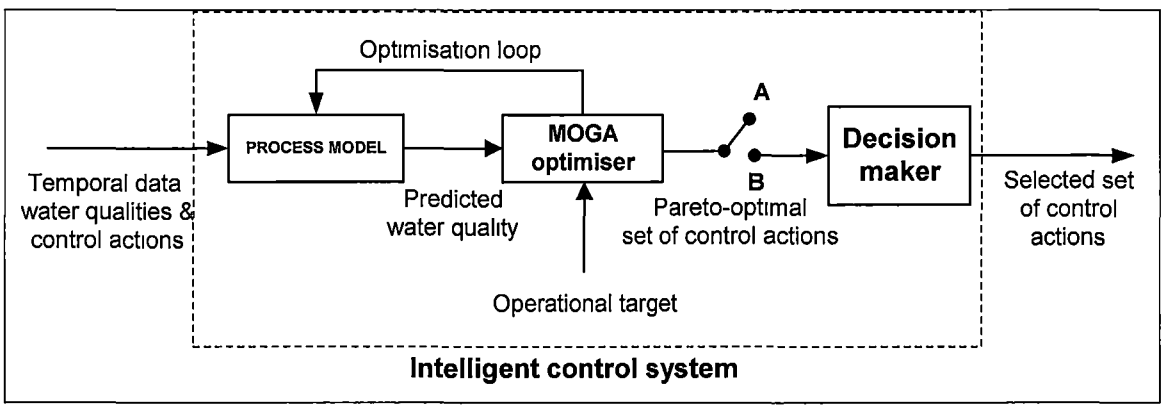


Figure 6.14, intelligent control system, was composed of three components: (i) process model, (ii) intelligent optimiser and (iii) decision maker.

At the beginning, the switch was turned to position "A". A clarifier process model was linked with the MOGA optimiser under MPC architecture. After the temporal operation data was input to the process model, it would predict clarified water quality and then give to the MOGA optimiser. The intelligent optimiser would optimise the set of control actions in order to satisfy the operational target (e.g. clarified water turbidity is less than 7 NTU for case of BKWTP) and minimise the operational cost. These two components would work repeatedly until predefined number of iterations was reach then the switch

would be turned to position "B". The last set of solutions, which was assumed to be the optimal one (i.e. Pareto front), was sent to the decision maker. Here, the best solution was chosen from the set of solutions on the Pareto front using shortest normalised distance selection.

The objectives of the clarifier control simulation were concerned with two issues: (i) the primary object was to find the optimal number of generations and the population size for the MOGA optimiser. Although the other MOGA parameters (e.g. GA operators and associated coefficients) could be predefined (section 6.3.4), the population size and number of generations were problematic and should be obtained empirically by trial and error. (ii) The other objective was to assess the performance of intelligent control system. Next the optimal number of generations and the population size were to be defined in line with the other predefined parameters. With past operation data, the clarifier controls were simulated by using an intelligent control system and their performances could be evaluated and compared with the human operators' performance. These performance assessments were conducted using the test sets of BEWTP and BKWTP case studies. To obtain these objectives, there were a few prerequisite tasks such that the objective functions and the agreement of simulations need to be defined.

6.5.1 Objective Function Formulation

Successful clarifier operation not only kept the clarified water quality satisfactory in terms of the operational targets, but to achieve economical operation presented a further challenge. Operation of clarifiers under both operational targets could be mathematically transformed to a multi-objective optimisation problem. MOGA was employed as an intelligent operator to search for the optimal set of control actions.

In the model predictive control scheme, the optimiser worked in line with the ANN process model. The ANN process model would predict the clarified water qualities in predefined future time steps (the so called predictive horizon) and the optimiser evaluated the set of control actions (e.g. chemical dosages, sludge drainage) and determined the optimal control actions by satisfying the clarified water quality (i.e. clarified water turbidity and colour) together with minimising the operational cost and the control effort. The objective function was minimised for every time step in the future, (so called control horizon). The control effort was also minimised to avoid aggressive control. To apply

MPC with the clarifier control, the control system objective was to improve the clarifier operation so that:

- The clarified water quality should not be higher than the operational targets.
- Operational cost was minimised.
- The clarifier was operated without aggressive control.

Referring to Equation (6.1), particularly to clarifier control, the objective function was given by:

$$\text{Minimise } J = \sum_{i=N_1}^{N_2} \langle \text{CWQ} - \text{Operational\_Target} \rangle + \sum_{i=N_1}^{N_2} OC + \lambda \sum_{i=1}^{N_3} (CA)^2 , \tag{6.16}$$

Subject to CWQ $\leq$ operational targets ,

Control actions bounds $u_{\min} \leq u(t) \leq u_{\max}$,

where λ was the penalising weight factor for control effort and N_1, N_2 were defined for Predictive horizon and N_3 was control horizon.

Predictive and control horizon

$\mathbf{N_1}, \mathbf{N_2}$ **and** $\mathbf{N_3}$ terms, were used to defined the predictive and control horizon. The performance of the MPC system was heavily reliant on the accuracy of the predictive model. Using MPC with long range prediction (multiple steps prediction) was not recommended (Norgaard et al., 2000). If this were done, predictive error would be added in every single future time step. In the specific case study of BEWTP and BKWTP, the results of prediction simulations (in chapter 5) confirmed that predictive errors would be greatly increased after one step prediction (section 5.3.4 and section 5.4.4). Therefore, the control and predictive horizons were set to a single step or the control system simulation was conducted for only one time step ahead which meant that $\mathbf{N_1}, \mathbf{N_2}$ and $\mathbf{N_3}$ were equal to one.

Clarified water quality

$\langle \mathbf{CWQ - Operational\_Target} \rangle$ stood for the difference between the predicted clarified water qualities (CWQ) and the operation targets. The bracket operator, $\langle\ \rangle$, from goal programming was applied here (Table 6.1, section 6.3.6) to ensure that the clarified water qualities were less than the operational target. If the "CWQ – Operational_Target" term

was positive (i.e. the clarified water quality upset the operational target), it returned to the value of the operand. Otherwise it would return to zero. For BEWTP, the set point was for clarified water colour to be less than ten HU and clarified water turbidity less than one NTU. On the other hand, for the BKWTP, the only set point was that the clarified water turbidity be less than seven NTU.

Control effort

$(CA)^2$ stood for the mean of the sum square of the normalised control action changes. In practice, any abrupt changes of control actions or aggressive control actions were not desirable and would result in operational instability (Mills et al., 1996). If changes of control actions between the present and the previous step were minimised while the other objectives (i.e. the clarified water qualities and operational cost) were satisfied, then the controls could be sustained smoothly. Since the magnitude of each control action was different (e.g. polyelectrolyte was in the range of 0.01 to 0.07 mg/L but that of alum dosage is 20 to 70 mg/L), it was necessary to alleviate the bias due to these different magnitudes. In the light of this, the control changes were normalised to the range of 0 to 1 by using associated boundary conditions. After normalisation, each square of the changes in control action was summed and averaged by being divided by the number of control actions which were used.

Operational cost, OC

OC took into account only for the chemical cost, and was calculated based on the price of per one cubic metre of clarified water. The electricity costs for turbine speed and sludge drain system were excluded since data were not available. Labour costs were also excluded. In a similar manner, the operational cost was also scaled into the range of 0 to 1 by associated maximum and minimum values which were functions of the control actions. On the other hand, the bias among different magnitudes of clarified water qualities could not be avoided since normalising of the clarified water qualities was not possible since their ranges were unknown.

Weight coefficient penalising control effort, λ

Weight coefficient penalising control effort, λ was used to tune the effect of the control effort. Norgarrd (2000) suggested that parameter λ could be obtained by trial and error in a control simulation. Typically, a suitable value of λ minimises the mean square error of setpoint tracking and the control effort simultaneously. This approach was well suited for

single objective optimisation but was not suitable for multi-objective optimisation of clarifier control. This was because the set point tracking error was defined as minimising the clarified water qualities rather than minimising the difference between the set points and clarified water qualities. Additionally, if clarified water qualities were considered only suitable for λ, this would lead to a bias in the operational cost minimisation. Therefore, λ was set to a value of one and used for the whole simulation.

Control action boundary $u_{\min}, u_{\max}$

One seriously weak point of the ANN model based control system was that it could not map the input to the output on a *one to one* mapping. Occasionally the process model performed a *many to one* mapping in which a number of inputs were mapped to one output. This was an ill posed problem (Swingler, 1996). For example, with the same raw water quality, more than one set of control actions gave the same clarified water quality. If this was the case, the non-sensible control actions were suggested from the process model that were not logical. For example, with very low quality of raw water, the process model suggested using very low dosage of chemical, which was clearly not sensible and was therefore excluded.

On occasions it might be advantageous to partition the input domains into multiple subsets or operational domains. This strategy was originally inherent in local modelling techniques (Murray-Smith & Johansen, 1997). It would here be adapted for the global process model of clarifier, using a single model to predict the output for the whole operational range. Thus, the input domain partition would be limited by having a bounded control action domain. Non-sensible control actions would not be available in this bounded search space. Selection of a proper boundary limited for the control actions required expert knowledge. The recommended control action ranges appropriate to raw water quality value were documented in the operating guidelines of both case studies (City Water Technology, 2001; MWA, 2000). Thus they could be utilised for partition of the operational domain. Not only alleviating the ill posed problem and providing the input domain partitioning, it also helped the MOGA to optimise the solutions under selection boundaries. This avoided looking for the solution from the whole search space and certainly reduced computational time.

6.5.2 Agreement of Clarifier Control Simulation

In this simulation, the test sets of each case study were to be recalled and used as benchmarks for performance evaluation since they were statistically equivalent to the operational data involved (section 5.3.2 and 5.4.2). The control system performance was evaluated in terms of clarified water quality and the operational cost in comparison with that of human operators.

Predefined MOGA parameters, population size and number of generations

Most of the MOGA parameters including those with associated GA operators (e.g. selection operator) were set according to the guidelines in section 6.3.4 and shown in Table 6.2. Only two values, the population size and the number of generations, were obtained empirically by trial and error processes since they were reported to be important to the convergence of the Pareto-optimal front (Deb, 2001). The MOGA was implemented using the genetic algorithm toolbox of GEATbx (Pohlheim, 2005).

MOGA parameters	**Parameter (varied)**
Number of generations	Trial and error process. (20 and 200 populations with the increments of 20 populations)
Population size	Trial and error process, (50, 100, 200, ...600 generations)
MOGA parameters	**Parameter (fixed)**
Population encoding	Real-valued
Selective function	Stochastic universal sampling
Selective pressure	2.0
Crossover function	Intermediate recombination
Crossover probability	0.01
Mutation function	Bit-wise mutation operators
Mutation probability	0.01

Table 6.2 Multiobjective Genetic Algorithm (MOGA) parameters showing which parameters were varied and which were fixed for the MOGA optimiser development process

The population size and generation were important to the convergence and diversity of the Pareto-optimal front (Deb, 2001). The Pareto Front Quality Index (PFQI) was used to

qualify the Pareto-optimal front (section 6.3.5), since the convergence and diversity of the Pareto-optimal front were assessed at the same time. The optimal numbers of generation and population size were systematically varied with a constant increment. For both case studies, the population sizes were in the range between 20 and 200 population with the increment of 20 population, and the number of generations is in the range of 50 to 600 generations (50, 100, 200, ...600 generations). For each case study, this trial and error process would be conducted with only one run which was randomly selected from the associated test set. The number of generations and associated population size which identified the smallest PFQI value were then assumed to be optimal. Additionally, the optimal population size and number of generations would be used for the whole control simulation. The associated number of generations would be used as the stopping criterion for the MOGA optimiser.

Decision maker

When the number of generations reached the optimal number, the MOGA would stop evaluating the populations. The populations (i.e. sets of control action) on the last Pareto-optimal front were assumed to be optimal. Consequently, the decision maker would choose the best set of control actions from the population in the last Pareto-optimal front using the shortest normalised distance (section 6.4).

6.6 Clarifier Intelligent Control System Simulation Results

This section concerns the result of the simulations. Initially it is necessary to lay out the specific background of each case study since each operates in a different environment and the operational targets of each case study are different. Each case study is now to be separately discussed.

6.6.1 Bryn Estyn Clarifier Control System

For the BEWTP case study, a clarified turbidity of 1 NTU and colour of 10 HU were set as the operational targets (City Water Technology, 2001). With predictive and control horizon of only one step ahead and the objective functions were shown as follows:

$$\text{Minimise } J = \sum_{t=1}^{1} \langle CWC - 10 \rangle + \sum_{t=1}^{1} \langle CWT - 1 \rangle + \sum_{t=1}^{1} OC + \sum_{1}^{1} (CA)^2 \qquad (6.17)$$

Subject to CWT ≤ 10 and CWC ≤ 1,

Control actions bounds $u_{\min} \leq u(t) \leq u_{\max}$,

where CWC and CWT were predicted clarified water turbidity and colour values at (t+1) or four hours ahead. $u(t)$ were the set of three control actions at the present time (t) (i.e. turbine speed, alum and polyelectrolyte dosages). OC was the operational cost ($AUS per cubic metre), which was a function of the alum and polyelectrolyte dosage at the present time (t).

Before the control simulation, it was necessary that the optimal population size and number of generations were identified. Systematic trial and error processes were conducted with a single run which was randomly selected from the test set. The population size and number of generations that gave the smallest PFQI were assumed to be optimal.

The control simulation was carried out on the test set which contains 204 runs. For the MOGA optimiser, the evolution would be terminated after the optimal number of generations was reached, and the decision maker would select the best solution based on the shortest normalised distance. The schematic diagram of BEWTP clarifier intelligent control is shown in Figure 6.15. Specific to this case, the two process models for predictive clarified water turbidity and colour are linked together. They gave the predictive water turbidity and colour at one future time step, (t+1) in conjunction with the MOGA optimiser that would evaluate the optimal solution set. After the predefined maximum number of generations was reached, the last Pareto-optimal front was assumed to be the optimal one, the switch would be turned to "B" position, then the set of control actions had been finally selected by the decision maker from the last Pareto-optimal front.

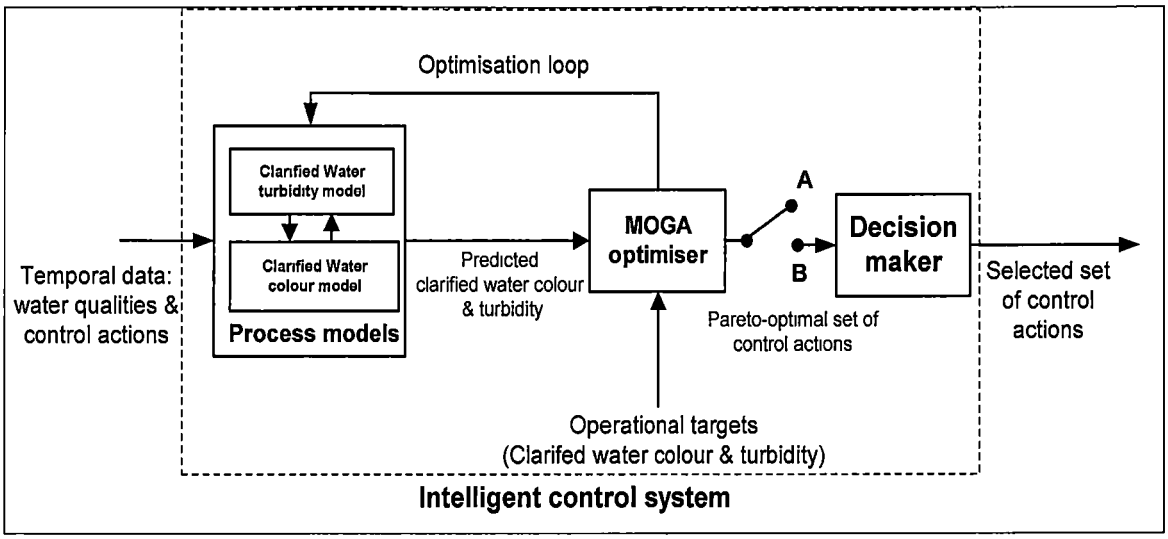


Figure 6.15 Schematic diagram of the intelligent control system (BEWTP), showing the intelligent control system for the clarifier at BEWTP; the process models and MOGA optimiser were linked under MPC architecture. After the optimisation process finished (switch would be turned to "B" position), the decision maker selected the set of control action from the last set of Pareto-optimal solutions.

Temporal span of inputs and predictive horizon

The two process models were used to predict clarified water turbidity and colour in the next four hours (one time step) only. The operation data at t, t-4, t-8 and t-12 were input to the model since that was empirically proved to be the optimal set (section 5.3.4).

Control action boundaries

In the BEWTP case study, only three control actions were involved (i.e. turbine speed, alum and polyelectrolyte dosages). The ranges of the control actions were set according to the operational guidelines shown in Table 6.3 (City Water Technology, 2001). Colour removal was the main issue at BEWTP. Thus, during the optimisation process, the raw water colour value at the present time was used to interpolate with the raw water colour in Table 6.3 for recommended control actions. After the control actions (as recommended) had been tried out, the chief operator suggested that another ± 10 percent should be added in practice to the upper and lower boundaries due the uncertainty of the raw water quality and its ambient conditions.

Raw water colour (HU)	Alum (mg/L)	Polymer (mg/L) (Summer time)	Polymer (mg/L) (Winter time)	Turbine speed (RPM)
20	17	0.12	0.13	12
25	18	0.12	0.13	12
30	20	0.12	0.13	12
35	22	0.12	0.13	12
40	24	0.13	0.14	12
45	25	0.13	0.14	12
50	26	0.13	0.14	12
55	27	0.135	0.14	12
60	28	0.135	0.15	12
65	29	0.14	0.15	12
70	30	0.14	0.15	12
75	31	0.145	0.15	12
80	32	0.145	0.16	12
85	33	0.145	0.16	11
90	34	0.145	0.16	11
95	35	0.15	0.17	11
100	36	0.15	0.17	11
110	37	0.15	0.18	11
120	38	0.15	0.18	11
130	39	0.15	0.19	11
140	40	0.15	0.19	11
150	41	0.15	0.20	9
160	42	0.15	0.20	9
170	43	0.15	0.20	8
250	45	0.18	0.22	8
300	55	0.22	0.26	8
350	65	0.26	0.30	8

Table 6.3 Recommended Control actions (BEWTP), showing the suitable ranges of control actions according to raw water colour values (City Water Technology, 2001).

Operational cost

The operational cost was based on the chemical cost only of alum and polyelectrolyte. Their unit cost per cubic metre (at a concentration of 1 mg/L of alum and polyelectrolyte) were show in Table 6.4. The electricity cost for turbine agitation and labour cost could not be included since the electricity and labour cost was recorded for the whole plant use and not specific for the clarifier use. These data were recorded monthly.

Chemical	Unit cost ($AUS per cubic metre)
Alum (1 mg/L)	0.000428
Polyelectrolyte (1 mg/L)	0.00785

Table 6.4 Unit cost of chemical (BEWTP)

6.6.2 Bryn Estyn Clarifier Control System Simulation Results

Firstly optimal population size and number of generations must be found next, the performance of the intelligent control system could be assessed and compared with the human operators' performance.

The optimal number of generations and population size

The optimal population size and number of generations were shown in Figure 6.16. By considering this contour plot, the PFQI value varied in the range of 1.45 to 2.65. The minimum value of PFQI was 1.45 and it is found where the population size and number of generations were 80 populations and 200 generations, respectively. Larger values of PFQI (low quality Pareto-optimal front) were located where the number of generations was small and population size was large (lower left hand side in Figure 6.16). This physically implied that, at such a large size of populations, the small number of generations was not enough to obtain a good quality Pareto-optimal front. However larger numbers of generations could improve the quality of the Pareto-optimal front. This was evidenced by a lower value of PFQI zone found at the top half side in Figure 6.16.

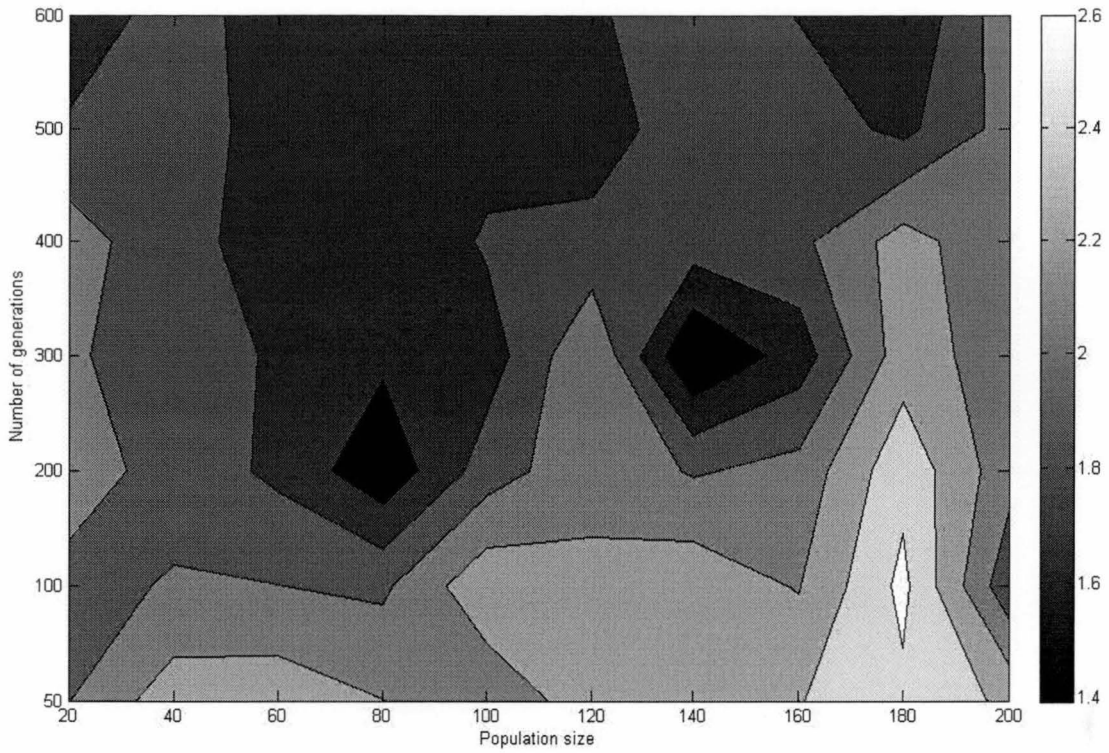


Figure 6.16 Pareto front quality index for the BEWTP case study, showing values of Pareto front quality index (PFQI) when the population size and number of generations were varied. The greyscale of the PFQI is shown on the right side of the figure, the lower the PFQI value the higher the quality of Pareto-optimal front.

Intelligent control system performance

Rather than measuring the performance of the intelligent control system by itself, it was preferable to compare the performance of the intelligent control system with that of human operators. All the associated statistical parameters and the upset number were shown in Table 6.5. The upset number was defined as number of runs which would cause an upset in the operational targets (more than 1 NTU of clarified water turbidity or 10 HU for clarified water colour).

Turbidity removal performance

Figure 6.17 shows the clarified water turbidity resulting from the human operator and from using the intelligent control system. Both were similar with coefficients of determination (r^2) of 0.88 and their means and standard deviations were also nearly identical, i.e. 0.72±0.49 NTU and 0.72±0.45 NTU respectively. They also both performed well since neither of their means was larger than the operational target of 1 NTU. However, the intelligent control system performed marginally better than human

operators since its upset number (47 runs) was slightly smaller than that of human operators (49 runs).

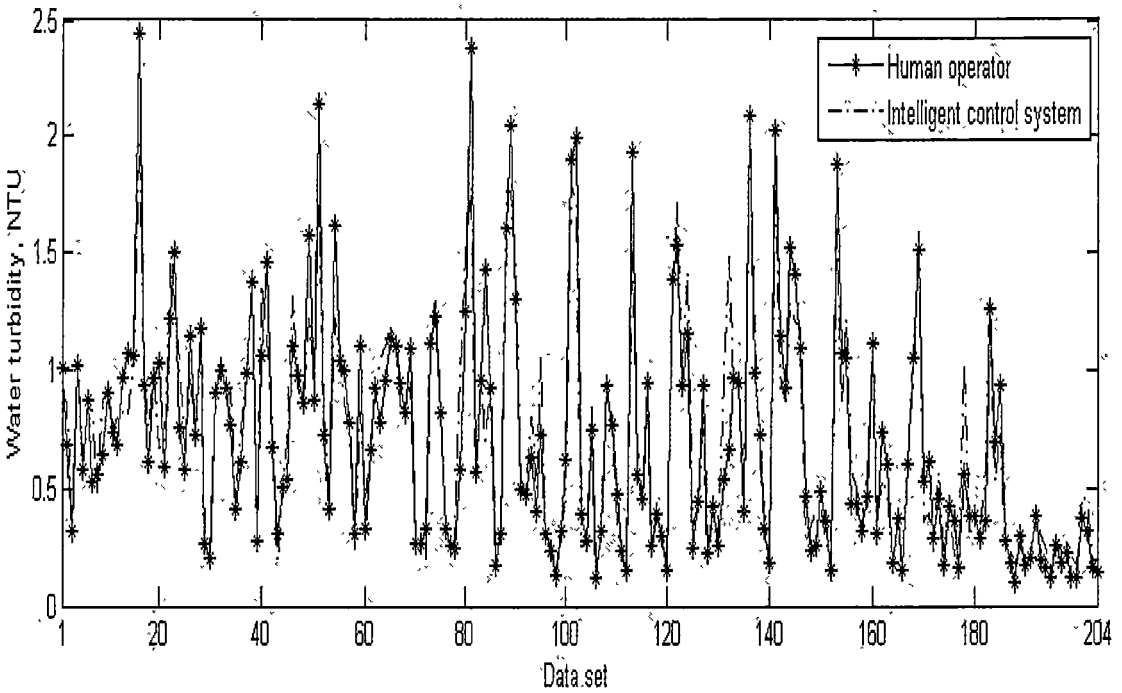


Figure 6.17 Turbidity removal performances (BEWTP case study), showing the simulation results of actual values of clarified water turbidity versus that of intelligent control system using the test set

Colour removal performance

The colour removal performance of both the intelligent control system and the human operators were shown in Figure 6.18 with means and standard deviation of 5.94±1.86 HU and 6.76±2.23 HU respectively. Therefore, both of them worked well and satisfied the operational target of 10 HU. Both of them varied with a low correlation as evidenced by the smaller coefficient of determination (r^2) of 0.63 (compared with clarified water turbidity). However, the intelligent control system works better than human operators since the upset number (7 runs) and the means of clarified water colour (5.94±1.86 HU) of the intelligent control system were smaller than that of human operators (9 upset runs, 6.76±2.23HU).

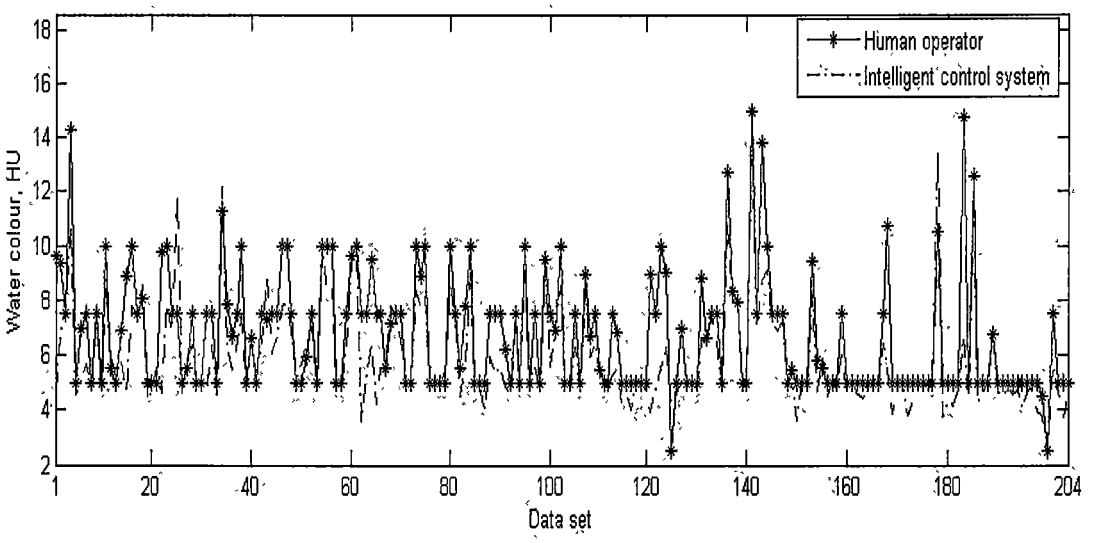


Figure 6.18 Colour removal performances (BEWTP case study), showing the simulation results of actual values of clarified water colour versus that of intelligent control system using the test set

Parameters	Human operator	Intelligent control
Clarified water turbidity (NTU) (Mean±std)	0.72±0.49 NTU	0.72±0.45 NTU
Clarified water turbidity (upset number*, runs)	49 runs	47 runs
Clarified water turbidity (upset percentage**,%)	24%	23%
Clarified water colour (HU) (Mean±std)	6.76±2.23 HU	5.94±1.86 HU
Clarified water colour (upset number)	9 runs	7 runs
Clarified water colour (upset percentage,%)	4.3%	3.4%
Alum dosage (mg/L)	26.07±6.64 mg/L	22.53±5.21 mg/L
Polyelectrolyte dosage (mg/L)	0.0141±0.0334 mg/L	0.0136±0.0154 mg/L
Turbine speed (RPM)	10.46 ± 1.21 RPM	11.85 ± 0.92 RPM
Operational cost (AUS $ per thousand cubic metres)	$AUS 12.00 ±3.0	$AUS 11.00 ±2.3

* Upset number is defined as number of runs which upsets the operational target

** Upset percentage = upset number/204*100

Table 6.5 Intelligent control system performance (BEWTP), showing the summary of the clarified water qualities and control actions of actual values of the human operator versus that of the intelligent control system.

Control actions and operational cost

In table 6.5, the human operators used a larger amount of alum and polyelectrolyte than the intelligent control system. This was evidenced by the larger means (26.07±6.64 mg\L of alum and 0.0141±0.0334 mg\L of polyelectrolyte) compared with those of the human operators (22.53±5.21 mg\L of alum and 0.0136±0.0154 mg\L of polyelectrolyte). Therefore, the intelligent control system reduced the alum and polyelectrolyte dosages by about 13.6 and 3.5 percent, respectively.

One should realise that the intelligent control system was only provided with a set of three control actions (i.e. turbine speed, alum and polyelectrolyte dosage). The sludge drainage was excluded since it was not initially included in the set of the process model input (due to data unavailability). However to keep the mass balance in the clarifier with no sludge drainage, it was logical that the intelligent control would reduce chemical dosages.

With the small amounts of chemical used, the intelligent control system ran the clarifier with a mean operational cost of 11.0 ±2.3 dollars per thousand cubic metres. On the other hand, the human operators spent 12.0 ±3.0 dollars per thousand cubic metres. Therefore the intelligent control system had the potential to save up to 8.33 percent of the operational cost. However, the intelligent control system required the turbine to agitate more rapidly than its human operators since the means and standard deviations of turbine speed were 11.85 ± 0.92 and 10.46 ± 1.21 RPM respectively.

From this simulation with the test set, it was obvious that the intelligent control system had a performance superior to human operators in terms of colour and turbidity removal efficacy as well as minimising the operational cost. Although it used smaller amounts of chemical dosage, the intelligent control system had proved its ability to select the proper turbine speed (for which the cost is negligible) and it could drive the control scheme in such a way as to achieve the operational targets.

Any decision to reduce the chemical dosage and to increase the turbine speed certainly reduced the operational cost. In addition, it was theoretically logical that increasing the turbine speed successfully improved the floc agglomeration rate (section 2.2.2). Thus combination of control actions represented a good decision when chemical dosages were lowered.

Another interesting point was that the intelligent control chose to lessen the alum dose rather than polyelectrolyte dose. Certainly reducing the alum dosage drove the control scheme efficiently to minimise the operational cost. Although the unit cost of alum dosage was cheaper than that of polyelectrolyte dosage (Table 6.4), the alum was used in much larger amounts than the polyelectrolyte.

However, for polyelectrolyte dosage, the intelligent control system could only save up to 3.5 percent compared with the human operators. This was much less than the case of the alum dosage where the intelligent control could save up to 13.6 percent. This might imply that it was the uncompromised need of polyelectrolyte which was necessary to keep stability of the operation during a period of external interference, such as possibly a change of density current. Further reducing polyelectrolyte dosages might upset the operational targets. In the light of these simulation results, this intelligent control system was able to recognise operation patterns and optimised the operation.

6.6.3 Bang Khen Clarifier Control System

For the BKWTP case study, clarified turbidity was the only operational target. This sole operational target was to achieve not more than 7 NTU of clarified water turbidity(MWA, 2000). The objective functions were shown as follows:

$$\text{Minimise } J = \sum_{t=1}^{1} \langle CWT - 7 \rangle + \sum_{t=1}^{1} OC + \sum_{1}^{1} (CA)^2 , \qquad (6.17)$$

subject to CWT $\leq$7 NTU,

control action bounds $u_{\min} \leq u(t) \leq u_{\max}$,

where CWT was the predicted clarified water turbidity value at one time step ahead, (t+1) (four hours ahead). An operational constraint of CWT $\leq$7 NTU was merged with the objective function by using goal programming. If "CWT-7" term was positive (i.e. the clarified water quality was more than seven NTU), it returns to the value of the operand. Otherwise it would return to zero. $u(t)$ was the set of four control actions at the present time (t) (i.e. sludge drain rate, turbine speed, and alum and polyelectrolyte dosages). The ranges of control actions are shown in Table 6.6 which were set in accordance with the operational manual and the suggestions of the chief operators (MWA, 2000). OC was

defined as operational cost per cubic metre, but only alum and polyelectrolyte dosages were accounted as included in the operational cost. The labour and electricity costs were not included since these data were not available and they were nearly constant and did not contribute to any change in the operational cost. Aggressive control action was avoided by minimising the terms CA^2 (which was the mean of sum square of all normalised control action changes).

In the same manner as in the BEWTP case study, the optimal number of generations and population size were initially identified. Systematic trial and error processes were conducted with a single run which was randomly selected from the test set. The Pareto Front Quality Index (PFQI) was used to decide whether the number of generations and population size were optimal for convergence and diversity. The number of generations and population size giving the lowest PFQI were assumed to be optimal, and were used for the whole simulation.

The control simulation was carried out on the test set which contained 271 runs. The process model gave only one future time step (t+1) of prediction values of the clarified water turbidity to the MOGA optimiser. Both of them worked in a manner similar to the case of BEWTP except that they were concerned only with predictive clarified water turbidity, set as the operational target (MWA, 2000). The schematic diagram of the BKWTP clarifier intelligent control is shown in Figure 6.19. The temporal data were input to the process model and the predicted values of clarified water turbidity were given to the MOGA optimiser. The sets of control actions were optimised until a predefined optimal number of generations was reached. The last Pareto-optimal front was assumed to be the optimal one. The switch was turned to "B" position. Thereafter, only one solution on the last Pareto-optimal front with the shortest normalised distance to the utopian point was selected by the decision maker.

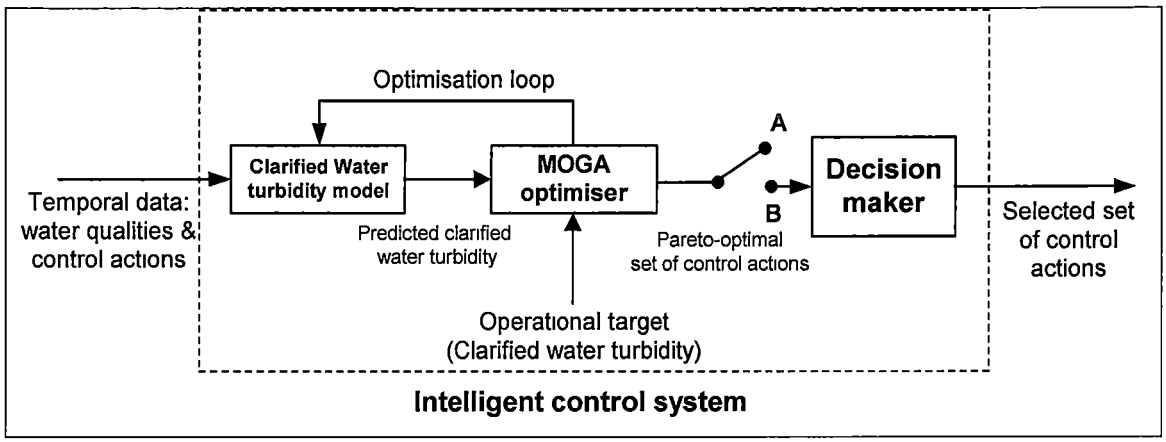


Figure 6.19 schematic diagram of the intelligent control system (BKWTP), showing the intelligent control system for the clarifier at BKWTP; the process models and MOGA optimiser were linked under MPC architecture. After the optimisation process finished (switch was turned to "B"), the decision maker selected the set of control action from the last set of Pareto-optimal solutions.

Temporal span of inputs and the predictive horizon

The ANN clarifier process model was used to predict clarified water turbidity in the next four hours (one time step) only. Operation data at t, t-4, t-8, t-12, and t-16 were input to the model since it was proved to the optimal set (section 5.4.3).

Control action boundaries

In the BKWTP case study, four control actions were utilised (i.e. sludge drain range, turbine speed, and alum and polyelectrolyte dosages). The ranges of the control actions were set according to the operational guidelines and the chief operator's suggested values are shown in Table 6.6 (MWA, 2000). All these ranges of control actions were set in line with the raw water turbidity since that was the main concern of BKWTP.

In practice, the chief operator decided daily the optimal dosages of chemical usage(based on the Jar test results) but not from the look-up table (Table 6.6). The existence of guidelines for control actions was just for emergency use and they were just approximated value (MWA, 2000). However, when using them to define the boundaries of control actions, the chief operator suggested that a further ± 10 percent should be added to the upper and lower limits of the alum dosage. Actually, in the BKWTP, alum rather than the polyelectrolyte was the main dosage vehicle for turbidity removal. It is therefore logical to use raw water turbidity as the value to guide the alum dosage. On the

other hand, the polyelectrolyte dosage was used to enhance operational stabilisation especially during external interference, for instance for the density current. Therefore to use only the value of the raw water turbidity as a guideline for the polyelectrolyte dosage might not be the best choice to maintain enough search space for the MOGA optimiser. The upper and lower limits of polyelectrolyte dosage are set equal to the upper and lower limits tabulated in Table 6.6. For example, if the raw turbidity at the present lag was 57 NTU (which fell into the range of 50 to 75 NTU) then without interpolation, the upper limit and lower limit of polyelectrolyte dosages were set to 0.015 and 0.020 mg/L respectively. This approach was also applied to identify turbine speed and sludge drain rate boundaries

Raw water turbidity (NTU)	**Alum (mg/L)**	**Polymer (mg/L)**	**Turbine speed (RPM)**	**Sludge drain rate (sec/hours)**
20	20	0.01	2	20
30	22			
40	23			
50	24	0.015		
60	26			
70	28		1.54	30
75		0.020		
80	30			
90	31			
100	33	0.030		
110	36		1.30	
120	39			
125		0.040		
130	41			
140	43			
150	47	0.050	1.11	60
160	51			
170	55			
175		0.060		
180	59			
190	64			
200	69	0.070	0.83	120

Table 6.6 Recommended Control actions (BKWTP), showing the suitable ranges of control actions according to raw water colour values (MWA, 2000).

Operational cost

The operational cost was based on the chemical cost of alum and polyelectrolyte only. For their unit cost per cubic metre (at a concentration of 1 mg/L of both alum and polyelectrolyte) see Table 6.7. The electricity and labour costs were excluded since the data were not available. The electricity cost was only recorded for the whole plant use and it was monthly cost.

Chemical	Unit cost ($ AUS per cubic metre)
Alum (1 mg/L)	0.000132*
Polyelectrolyte (1 mg/L)	0.00385

* based on 1 $ AUS = 28.00 Baht

Table 6.7 Unit cost of chemical (BKWTP)

6.6.4 Bang Khen Clarifier Control System Simulation Results

In a similar manner to the BEWTP case study, the optimal population size and number of generations were initially defined. Then the performance of the intelligent control system was assessed and compared with the human operators.

The optimal number of generations and population size

The optimal population size and number of generations were found in Figure 6.20. Considering this contour plot, the PFQI value varied in the range of 1.24 to 1.65. The minimum value of PFQI (1.24) was found where the population size and generation were 100 populations and 400 generations respectively. The larger value of the PFQI zone (with a low quality Pareto-optimal front) was located where the numbers of generations were less than 200 generations, regardless of the number of initial population. This physically implied that for any population size, a small number of generations was not enough to achieve good quality of Pareto-optimal front. This might be due to the complexity of solution space. The better quality Pareto-optimal front (lower PFQI value) was found when the number of generations increases above 200 generations.

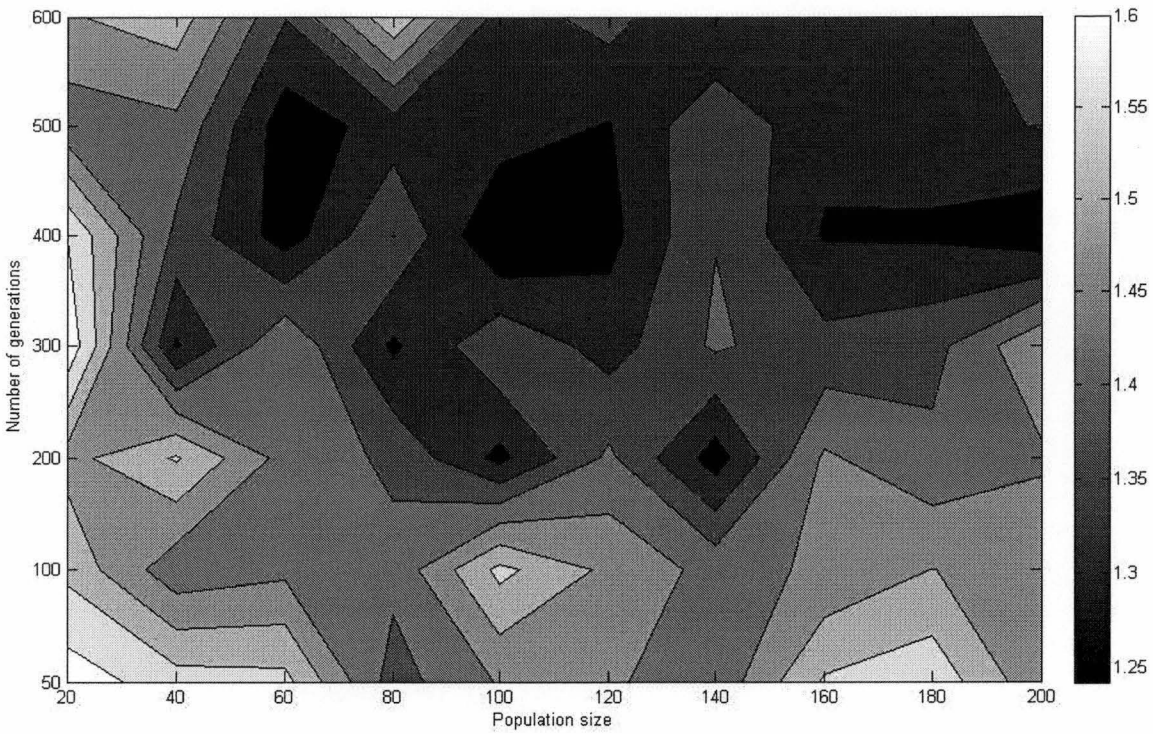


Figure 6.20 Pareto front quality index for BKWTP case study, showing value of Pareto front quality index (PFQI) when the population size and number of generations were varied. The greyscale of the PFQI was shown on the right side of the figure. The lower The PFQI value the higher the quality of Pareto-optimal front obtained

Intelligent control system performance

In a similar manner to the BEWTP, the performance of the intelligent control was compared with that of its human operators. All the associated statistical parameters and the upset number are shown in Table 6.8.

Clarified water turbidity

Figure 6.21 shows the clarified water turbidity from the human operator and intelligent control. Both of them varied with coefficients of determination (r^2) of 0.77 and their means and standard deviations were 6.35±1.52 NTU and 6.20±2.44 NTU respectively. Their performances were marginally acceptable with the target at 7 NTU. Although the intelligent control system with the smaller mean of clarified water turbidity worked better than human operators, it had a higher standard deviation. This implied the existence of operational instability which was also evidenced by the larger number of runs upsetting

the operation target (88 runs upset the target of 7 NTU) compared with that of the human operators (76 runs upset the target of 7 NTU).

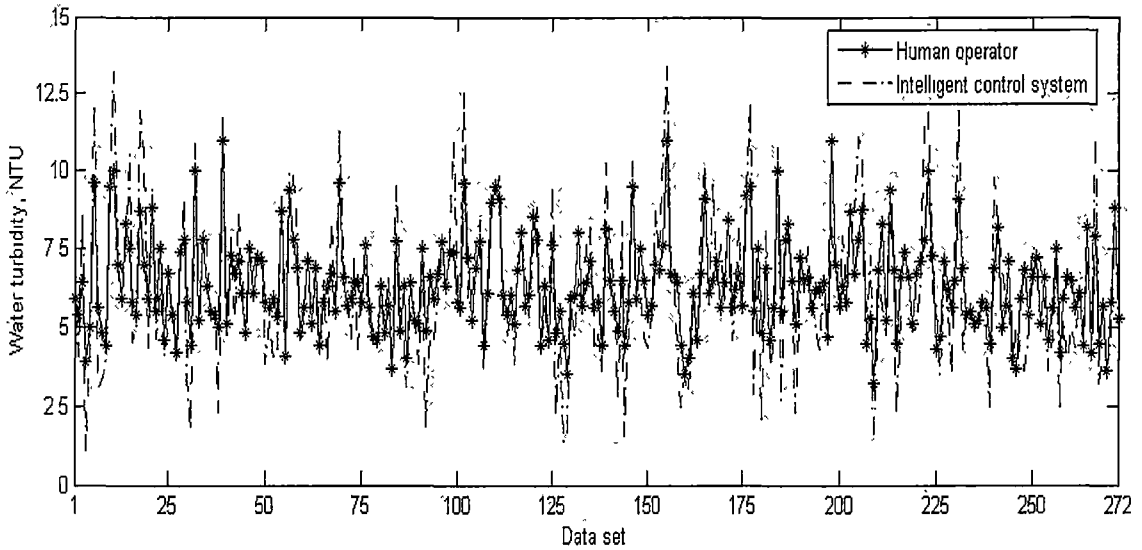


Figure 6.21 Turbidity removal performances (BKWTP case study), showing the simulation results of actual values of clarified water turbidity versus that of the intelligent control system using test set

Parameters	Human operator	Intelligent control
Clarifier water turbidity (Mean±std),(NTU)	6.35±1.52 NTU	6.20±2.44 NTU
Clarifier water turbidity (upset number*, runs)	76 runs	88 runs
Clarifier water turbidity (upset percentage**, %)	28.0 %	32.5%
Alum dosage (mg/L)	30.79±9.96 mg/L	29.48±6.61 mg/L
Polymer dosage (mg/L)	0.020±0.010 mg/L	0.025±0.010 mg/L
Turbine speed (Sec/Rev.)	1.43 ± 0.22 RPM	1.49 ± 0.22 RPM
Sludge drainage rate (Hours)	39.98±17.79 Sec/Hours	34.57±11.01 Sec/Hours
Operational cost (AUS $ per thousand cubic metres)	$AUS 4.14 ±1.07	$AUS 4.00 ±0.89

* Upset number is defined as number of runs which upsets the operational target

** Upset percentage = upset number/272*100

Table 6.8 Intelligent control system performance (BKWTP), showing the summary of the clarified water qualities and control actions of actual values of the human operator versus that of the intelligent control system.

Control actions and operational cost

In Table 6.8, operating the clarifier by human operators has been shown to be slightly more costly by about 3.45 percent more than by using the intelligent control system. The associated operational costs were $AUS 4.14 ±1.07 and $AUS 4.00 ±0.89 per thousand cubic metres for human operators and for the intelligent control system, respectively.

The human operators fed a slightly higher amount of alum (30.79±9.96 mg\L) than did the intelligent control system (29.48±6.61 mg\L). On the other hand, much less polyelectrolyte was fed by human operators. That was 0.020±0.0102 mg\L of polyelectrolyte fed by human operators and 0.025±0.010 mg\L of polyelectrolyte given by intelligent control. Consequently, the intelligent control system reduced the alum dosage by 4.25 percent but increased polyelectrolyte dosages by about 25 percent.

For the turbine speed and sludge drainage rate, the human operators agitated the water to more gently (1.43 ± 0.22 RPM) in comparison with the intelligent control (1.49 ± 0.22 RPM). On the other hand, human operators tended to drain more sludge (39.98±17.79 Sec/Hours) than the intelligent control system did (34.57±11.01 Sec/Hours).

From this simulation using the test set, the intelligent operator performed slightly better with a lower mean than the human operators in terms of turbidity removal. However, with a higher standard deviation of clarified water turbidity, this resulted in the intelligent control being unable to stabilise the operation. This was also evidenced by a larger upset number. External interference by a density current might also be the cause. If this was the case and raw water temperature were to be the input to the process model, then the clarifier model would recognise the density current phenomenon. This would considerably help to improve the operational stability. However, considering operational costs, the intelligent control system ran the clarifier slightly more cheaply.

One interesting point found in this simulation was that intelligent control drives the control scheme towards a cheaper solution by selecting the cheaper control action of polyelectrolyte dosage the same as it did in the case of BEWTP. However, 25 percent more polyelectrolyte dosage seems a large figure but operational stability (possibly due to current density) could still not be achieved. This was evidenced by a larger upset percentage and standard deviation. The intelligent control system might possibly not receive enough information. One should realise that the intelligent control had no

information about water temperature and this was directly concerned with the current density phenomenon.

In this simulation, the intelligent control agitated more gently (resulting in a lower amount of floc being produced) but at the same time had to retain larger amounts of flocs in order to maintain mass balance in the clarifier by lessening sludge drainage (section 2.2.2). This signified that the intelligent control system well recognised the control pattern from past operational data.

6.7 Chapter Conclusion

The intelligent control system was composed of an ANN clarifier process model and MOGA optimiser under MPC architecture. By simulations using selected test sets (the same as used in Chapter 5), the performance of the intelligent control systems were compared with human operators in both case studies. Most of the MOGA parameters were selected according to guidelines from previous studies. Only population size and number of generations were optimised by trial and error processes. The intelligent control systems were designed specifically for each case study because each had different operational targets.

In recognition of the result from Chapter 5, the optimal temporal span of 8 hours for BEWTP and of 12 hours for BKWTP were used for each of the intelligent control systems. Furthermore, one time step ahead of four hours was set as the control and predictive horizon for both systems. The shortest normalised distance approach was employed as the decision maker and control action boundaries were limited in order to alleviate ill posed problems in the process models.

From the simulation result of these two case studies, the intelligent control system operated clarifiers at a lower operational cost compared with human operators. In terms of impurity removal performance, the intelligent control system showed a performance superior to human operators at the BEWTP but satisfied the operational target only marginally for the case of the BKWTP.

The control actions given by the intelligent control systems were theoretically logical. For instance, in the case of the BEWTP, with a lower mean of chemical dosages, the

intelligent control system increased the turbine speed to generate more flocs for mass balancing in the clarifier. In case of BKWTP, with a lower dosage of alum used (compared with that of human operators), the intelligent control decided to raise the polyelectrolyte dose which resulted in lessening the overall operational cost.

From the simulation results of the BKWTP case study and comparing between the intelligent control system and human operators, their turbidity removal performance was closely similar. However, the intelligent control system gave a cheaper operational cost. Although the mean of clarified water turbidity marginally satisfied the operational target, but in 88 runs from 272 runs (about 32.5 percent) the clarified turbidity target was not met. This was rather high if it was compared with that of the BEWTP case study, in which 47 runs from 204 runs (about 23.0 percent) of target were not met for clarified turbidity target and 7 runs from 204 runs, (3.4 percent) were unmet for clarified water colour.

In the BEWTP case study the intelligent control system performed better than human operators in terms of both impurity removal and operational cost control. All the operation targets were well satisfied. To confirm the result of the simulations, a full scale pilot plant test should ideally be implemented. Unfortunately a pilot plant test could not be conducted in BEWTP because there is only a single clarifier which is continuously operated at BEWTP. A second clarifier exists but it is used only for batch operation. It was, of course, not possible to implement the intelligent control system while the human operators were running the clarifier. Safety and community health issues were also of concern. If any errors occurred in trialling the pilot plant, major worries would arise in the local community. Additionally, one of the control actions, sludge drainage, was excluded from the input set since the data were not available. Therefore, it was not reasonable to try to conduct a pilot plant test. The human operators also have greater freedom of choices of the control actions to be taken whilst choices would be limited if the intelligent control were to be employed.

However, it could be claimed that in terms of the simulation results, the performance of the clarifier with an intelligent control system at BEWTP was superior to that of BKWTP. Therefore, an implementation of a pilot plant test at BKWTP should be set as a priority. Other additional factors supporting this decision are that (i) all of the control actions that the operators used were all utilised for the intelligent control system. (This

was not in case of BEWTP) and (ii) many (18) clarifiers operate at BKWTP, and it would be possible to compare the performance of the intelligent control system and human operators in identical conditions. (iii) A single clarifier could be used for a full scale pilot plant test while the other 17 clarifiers were operated as normal, and so there would be little potential impact on safety or on community health.

The novel idea of using an intelligent MPC with the ANN process model with an evolutionary optimiser MOGA to control the clarifier was confirmed by the simulation results for both case studies. They showed potential for use in real life situations. The other configurations of the intelligent control system, such as the decision maker based shortest normalised distance and using the control action bounded for partition of the operational domain were contributed to the originality of this mode of the clarifier control. The next chapter describes the full-scale pilot plant test at the BKWTP.

7.0 CLARIFIER INTELLIGENT CONTROL PERFORMANCE: FULL-SCALE PILOT PLANT TEST

7.1 Introduction

This chapter describes the pilot plant test of the BKWTP intelligent control system that was developed in Chapter 6 and is based on ANN Model Predictive Control (MPC) (composed of the BKWTP ANN clarifier process model and the MOGA optimiser). The full-scale pilot plant test was conducted at Bang Khen Water Treatment Plant (BKWTP), Bangkok, Thailand, during the period of August and September 2008 which was the early rainy season. This was the first flood period of the year with high flow rate in Chao Phraya River causing high degree of erosion and raw water turbidity increase. Therefore, it was a suitable period for full-scale pilot plant test since the intelligent control system would be tested in the wider range of raw water turbidity. The performance of the intelligent control system was compared with that of human operators in real situations using the criteria of turbidity removal and operational cost. Some practical aspects concerning BKWTP will be introduced as background before the pilot plant test results are discussed. In sequence, this chapter presents: (i) the intelligent control system integration which describes how the intelligent control system is merged into the existing control system, (ii) the agreement of the full-scale pilot plant test and some practical issues of BKWTP, and (iii) the evolution of intelligent control system performance and comparison with that of human operators.

7.2 Intelligent Control System Integration

In general, operators control the clarifier by using four control actions: alum dosages, polyelectrolyte dosages, turbine speed and the sludge drainage rate. Selection of suitable control actions are taken using information concerning raw water qualities and guided by Jar test results. At BKWTP, the clarifiers are routinely operated manually. Although a number of online sensors are presented in the system, these are not fully electronically linked together (because a *Supervisory Control And Data Acquisition*, (SCADA) system being installed is not yet available). Manual meter readings are taken from each sensor and an internal telephone is used to communicate information. Data is manually entered to spreadsheet software together with hard copy used solely for data storage. Data is

routinely collected for every four hours in each working day. This includes raw water qualities from the laboratory and meter readings. However the Jar test is conducted twice a day. From these the suitable alum and polyelectrolyte dosages are recommended. The remaining control actions (i.e. the turbine speed and sludge drainage) are optimised by using the operator's experience. The flow of information is illustrated in Figure 7.1.

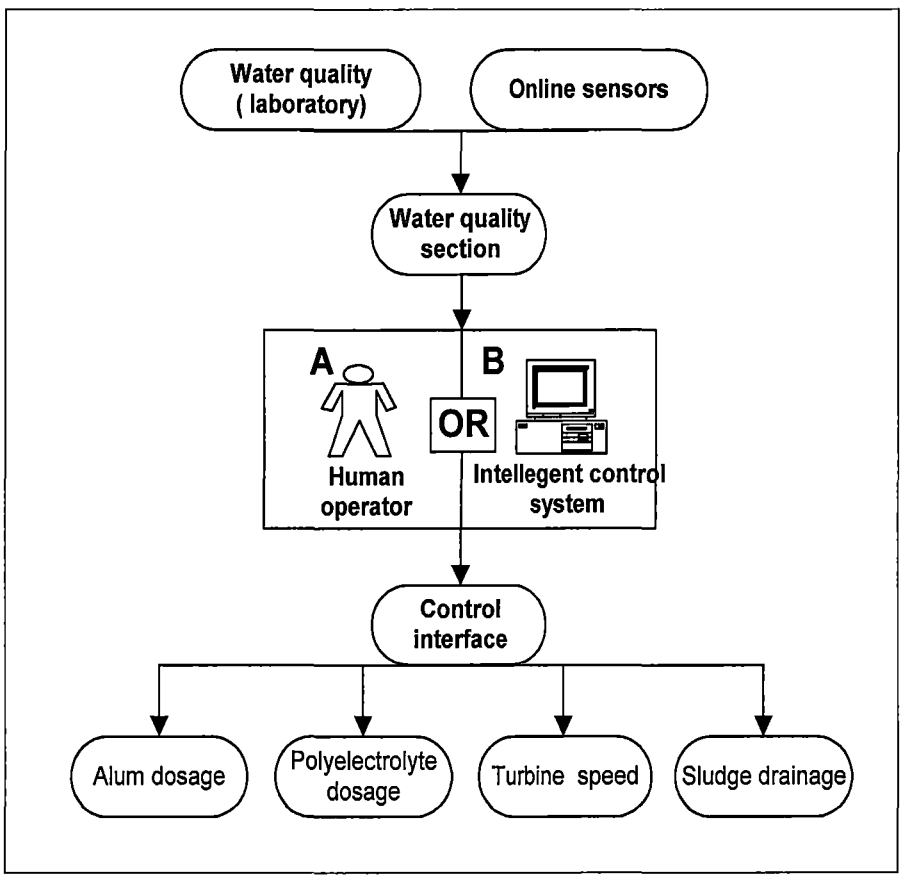


Figure 7.1 The flow of information: A) the human operators receive data from laboratories and online sensors, then they select the control actions, B) the intelligent control system uses the same data during the pilot plant test.

Two clarifiers were controlled, one by human operators and one by the intelligent control system. The intelligent control system was substituted for the human operator as shown in Figure 7.1. During the full-scale pilot plant test, both clarifiers were updated with new data every four hours (i.e. the same frequency as that of human operators). Both the intelligent control system and the human operators independently selected a set of appropriate control actions every four hours.

7.3 Practical Aspect: BKWTP Clarifier Control

In addition to details initially given in Chapter 2, this section presents more practical information concerning the clarifiers at BKWTP. The means by which the operators control the clarifiers are discussed. Then the layout of BKWTP and the design and capacity of the clarifiers are introduced, and finally the clarifier operation in practice is addressed.

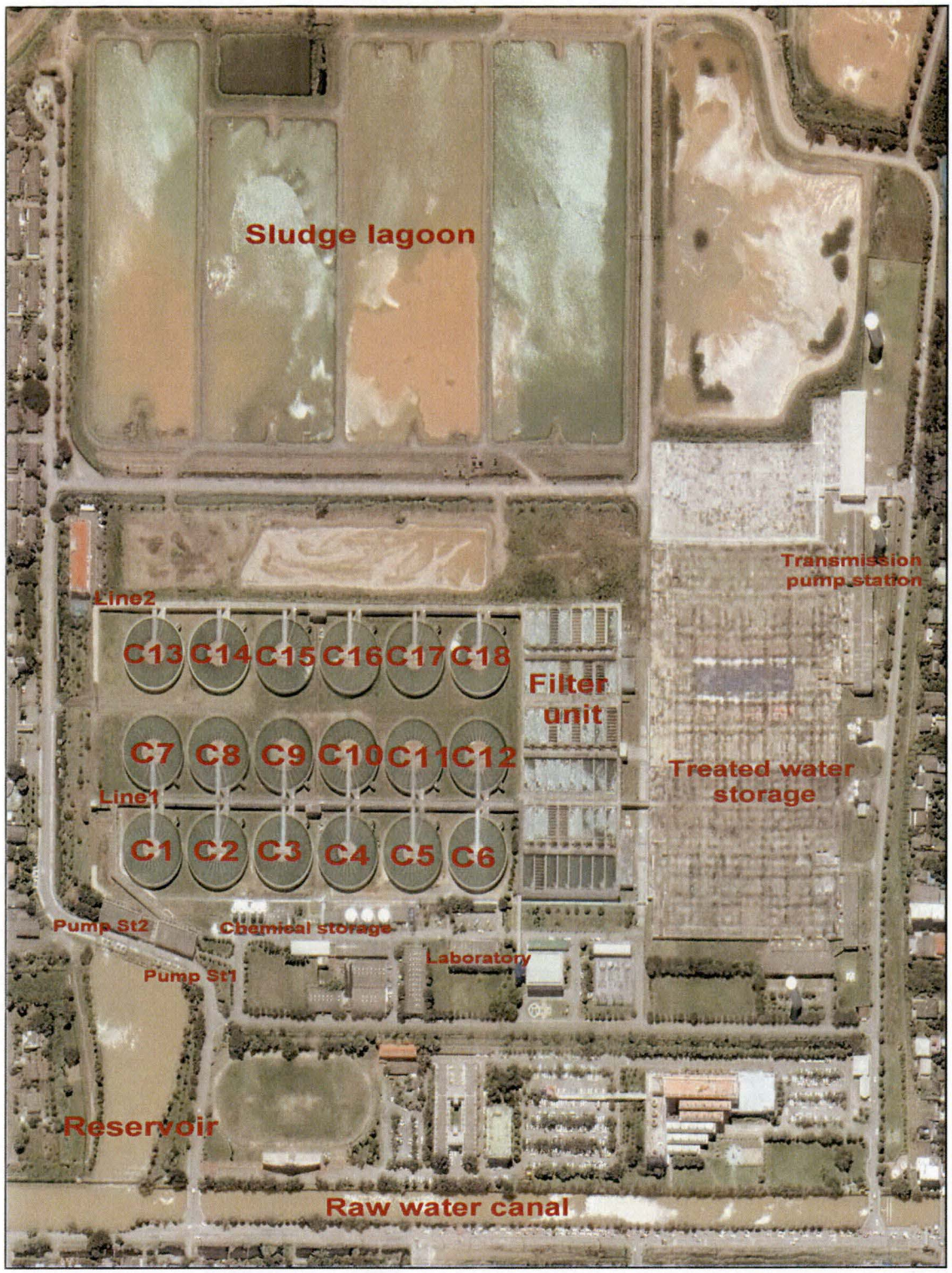
7.3.1 Layout, Design and Capacity

In BKWTP, there are 18 clarifiers of the solid contact sludge recirculation type allocated in two production lines. The first production line contains twelve clarifiers and six are in the second production line. Two raw water pumps separately drive raw water up to their head tanks then gravity feed the raw waters to the associated production lines. The solid contact sludge recirculation type is designed for high production rates (surface overflow rate) in the range of 2.0 to 2.4 m/hour. The recycling of the sludge is introduced to enhancing floc amalgamation and to minimise chemical use. The clarifiers at BKWTP are designed to operate in a range of not more than 150 NTU of raw water turbidity (MWA, 2000).

All clarifiers are of the same size, 58 metres in diameter, each with a capacity of about 200,000 Cubic Metres per Day (CMD). With different sizes of the reaction well, they are categorised into three groups as shown in Table 7.1. These clarifiers are numbered and located in two production lines as shown in Figure 7.2. The alum and polyelectrolyte are dosed to the first production line containing 12 clarifiers. The other six clarifiers are located in the second production line and dosed by Poly Aluminium Chloride (PAC) and polyelectrolyte.

Reaction zone size	Clarifier number	Designer
Small reaction zone	1-6	Eimco
Medium reaction zone	7-10 & 13-18	Fuler & Smith
Large reaction zone	11&12	Ebara

Table 7.1 Clarifier types in BKWTP



Source: www.earth.geogle.com

Figure 7.2 Layout of BKWTP, showing the satellite image of BKWTP with 18 clarifiers in two production lines. Water is sourced from the raw water canal and retained in the reservoir before pumped to the treatment process. After the treatment process, the treated water is stored in the treated water storage and fed to the distribution network. The drained sludge is discharged to the sludge lagoon.

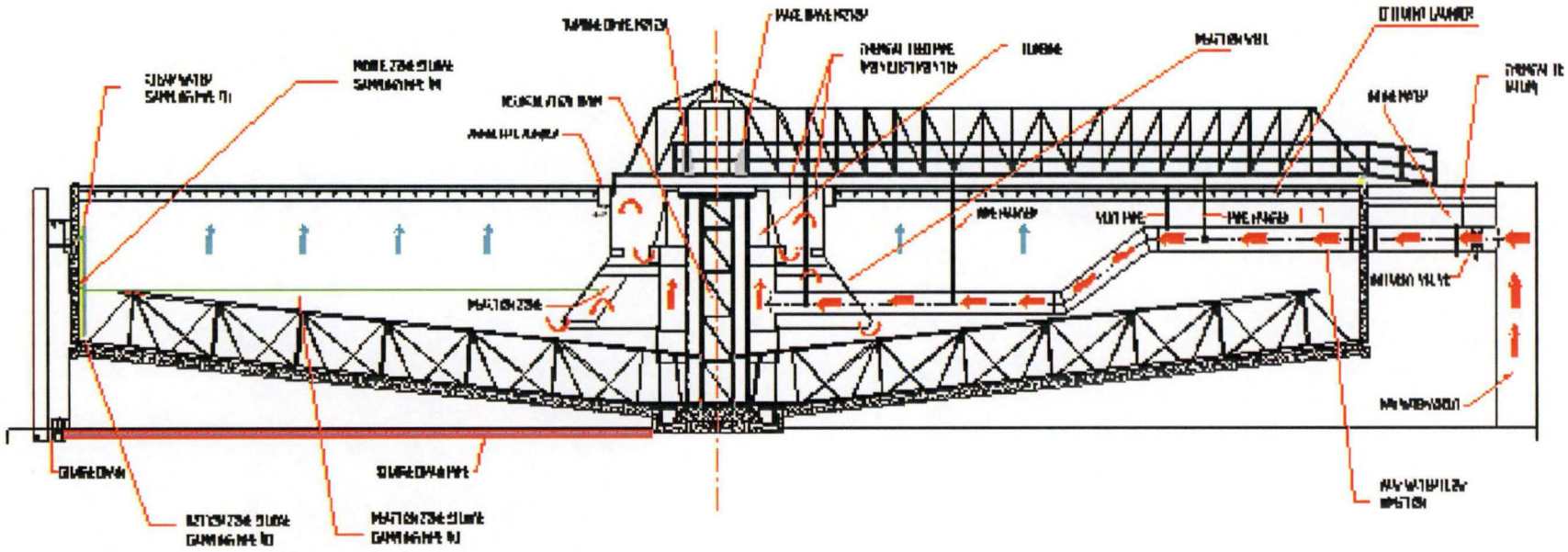


Figure 7.3 Cross section view of the solid contact clarifier at BKWTP, showing the direction of flow of the raw water and the equipment in the clarifier. The red arrows represent raw water and blue arrows represent clarified water.

7.3.2 Bang Khen Clarifier Operation

Figure 7.3 presents the cross section of the clarifier and illustrates the detail of each unit in the clarifier and shows the direction of the raw water flowing through the clarifier. The pH adjusted raw water is initially pumped to the clarifier through the inline mixer installed at the inlet of the clarifier. Alum is also dosed at the inline mixer and assumed to be completely mixed with raw water. In the clarifier, raw water flows to the reaction well via the raw water conduit. A recirculation drum and turbine spin in the direction that forces the raw water to flow up to the top of the reaction zone. Here, the polyelectrolyte is fed to help floc amalgamation. Suspended solid, alum and polyelectrolyte initially form larger particles (called flocs) by gentle agitation of the turbine. Collision between flocs (acting under soft agitation) further promotes floc amalgamation in the reaction zone. The large flocs continuously settle to the bottom zone. While the scalper is sweeping around the floor to collect the settle flocs, some small portion of settle flocs are mixed with new coming water to promote floc amalgamation (this explains the term "slurry recirculation" in the name of the type of clarifier). The excess settled flocs are drained through the sludge drainage system for mass balancing in the clarifier. The clarified water flow upward to effluent launder and discharge to filter unit. The relationships between these control actions (i.e. turbine speed, sludge drainage rate, alum and polyelectrolyte dosage) are already described in section 2.2.

7.3.3 Monitoring and Control

Figure 7.4 presents a diagram of clarifier sampling points and control actuators. In terms of process monitoring, the operators routinely measure the water qualities every four hours. The measurements are taken six times daily at 00:00, 04:00, 08:00, 12:00, 16:00 and 20:00 hours. Raw water qualities are sampled after the water has been conditioned by lime (at the inlet of the clarifier) and clarified water turbidity is monitored at the outlet of the clarifier. The sludge concentration is also measured every four hours. As this is the only information available from inside the clarifier, sludge concentrations are sampled at the reaction zone (at R in Figure 7.4) and at three other points along the depths of the clarifier. One sample point is at the bottom (at B in Figure 7.4) and the other two are at the middle and top of the tank (at M and T in Figure 7.4). For clarity, the sludge sampling pipes and the sampling points (B, M, R and T) also shown in Figure 7.3.

For process control, operators normally use four control actions (i.e. alum dosage, polyelectrolyte dosage, turbine speed, and sludge drainage), and they will take action at least every four hours. The chemical dosages are guided by Jar tests which are conducted twice a day (at 08:00 and 16:00 hours). Proactive control actions are also possible and this occasionally happens during abrupt changes in raw water qualities. In this cases the operators will request an extra jar test and make a decision for a new set of control actions.

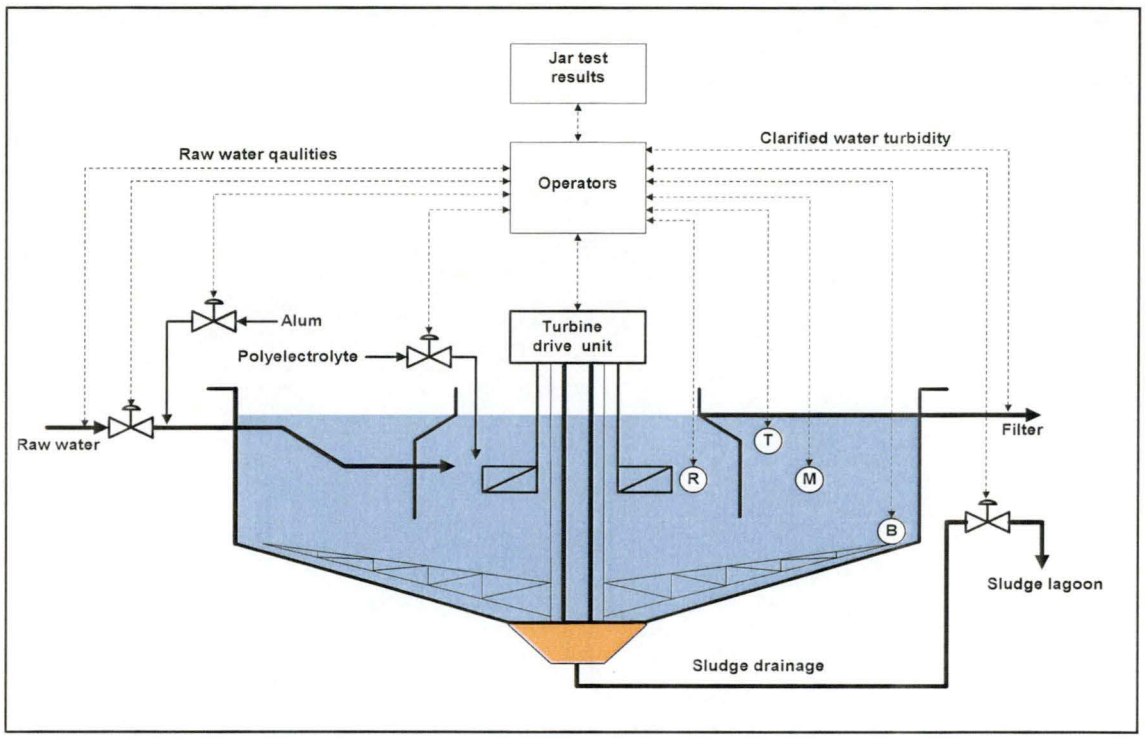


Figure 7.4 Clarifier sampling points and control actuators; presenting the flow of information from the sampling points to the control actuators for clarifier control at BKWTP

Besides clarified water turbidity, sludge concentration is used to monitor the completeness of the operation. Normally, regardless of chemical dosage and raw water quality, from two to three percentage of sludge concentration should be maintained in the reaction zone (position 'R') and less than one percent in the middle zone (M). In practice, there should be strictly no sludge in the top zone (T), and not more than five percent in the bottom zone (B) (MWA, 2000). The sludge concentration sampled from inside the reaction well is an indicator of floc amalgamation. Samples from inside and outside of the reaction well provide an alert for an upset condition. In Figure 7.5 the operator is collecting sludge from the sludge sampling pipes (from B, M, T and R positions) located beside the clarifier tank.

Figure 7.5, Sludge samplings showing the operator sampling the sludge from the sampling points. The labels B, M, T and R indicated that the sludge was being sampled from the Bottom, Middle, Top and Reaction zone of the clarifier.

7.3.4 Chemical Dosages

Two chemical agents, alum and polyelectrolyte are directly concerned with BKWTP clarifier control. The Jar test results are used to guide these chemical dosages. In BKWTP, the chemical section staff feed alum and polyelectrolyte in amounts appropriate to the whole production line. It is the responsibility of the operators to adjust the chemical flow rates into each clarifier. Theoretically, the chemical agent should be fed equally to each clarifier but this may not be the case in practice due to the different sizes of the reaction well and the hydraulic condition within each clarifier. This adjustment relies purely on the operator's experience. In Figure 7.6, the author is trained to adjust the valves for optimal alum dosage.

Figure 7.6 Adjusting chemical dosages. (The author in training for chemical dosage control with the senior operator)

Turbidity precipitation is the main target at BKWTP, because the turbid raw water from its source is highly variable in turbidity on a seasonal basis. The density current is another of the main concerns because the clarifier is large in size, and there is severe difference between day and night temperatures of Bangkok. A recommendation has been made to use a density current baffle system to alleviate this problem. However, this solution is not possible in the short term because it would be costly and operations would have to stop during the installation. It may be less efficient to confront this problem by using slightly raised overdose of alum and polyelectrolyte to enhance floc amalgamation. In the opinion of the chief operator, operational stability can occasionally be achieved by slightly overdosing with the chemical agents especially if the day is hot.

Until the year 2006, high corrosion in the river caused of peaks of raw water turbidity. This was especially common in the early periods of the rainy season (July to August). River sand mining is also often the cause of raw water peaks outside the rainy period. However, since 2006, river sand mining is legally prohibited and this has resulted in the water being clearer. Unfortunately, clear water promotes photosynthesis and results in algae blooms. When raw water containing algae is used (even though it brings a lower turbidity), it results in shorter filter run times. Usually, filters requires back washing twice a day (when clarified water turbidity is not more than 7 NTU). During algae crises even with the same turbidity value, the filters need to be backwashed much more often (MWA, 2006). Sometimes washing is needed more than four times a day. This problem was first discovered in the summer of year 2006, when water flooded over the filter units. To alleviate this crisis, some actions have been taken: (i) intermediate chlorination, where the chlorine can be fed at the intake of the clarifiers, or (ii) Jar tests are taken frequently for algae removal (rather than only for turbidity) (MWA, 2006). The chief operator believes that increasing the alum and (especially) polyelectrolyte dosages helps in algae removal.

7.3.5 Operational Target Upset and Alleviation

In some situations, the clarified water turbidity may be considerably above the operational target of 7 NTU. According to the operation manual (MWA, 2000), the following actions are suggested. If the turbidity leakage is not severe (say between 7 and 15 NTU), the operators should reduce the turbine speed and increase the sludge drainage. However, if the clarified water turbidity reaches more than 15 NTU even after the rapid

slowing down the turbine speed, the emergency drain is suggested as a first line control. Then the operators should increase the chemical dosages and a request for an extra Jar test is permissible. However, after using emergency sludge drain, the operator should keep the sludge concentration in the reaction zone (R) and in the bottom zone (B) up to about 15 to 20 percent higher than the normal level (equivalent to the starting up state) and later slowly reduce them again.

7.4 Agreement of Full-scale Pilot Plant Test

The first production line is the main production line of the BKWTP and has a productivity of twice the second production line. Clarifier number nine, which is located in the middle of this production line, was chosen for installing the intelligent control system. This avoided any bias resulting from varying distance from the head tank leading to different hydraulic conditions (Figure 7.2). In addition, it alleviated the bias due to the reaction well sizes since Clarifier number nine's reaction well is of middle size (as shown in Table 7.1)

It was necessary to select one clarifier from the other eleven clarifiers in this production line to represent the human operator. The clarified water turbidity needs to show similarity with that of Clarifier number nine. This task was actually done by comparing the coefficients of determination (r^2) between clarified water turbidity at Clarifier number nine to the others. These correlation tests were conducted on past operational data from January to April 2008 (a period of four months before the pilot plant test took place). After that one clarifier was chosen to represent a human operator, and its performance was compared with that of the intelligent control system.

Not all the control actions could actually be executed and implemented punctually (for every four hours) due to delay in obtaining manual and laboratory results. However, it has been assumed that all the control actions were applied promptly.

The control actions were optimised by the intelligent control system in order to (i) minimise the clarified water turbidity and (ii) minimise the operational cost. The intelligent control system optimises its control actions within predefined boundaries (as shown in Table 6.6) to alleviate any ill posed problems. Other components of the MOGA optimiser were also predefined and shown in Table 6.2. The control and predictive

horizons were four hours ahead. As it had been proved to be the optimal temporal span (section 5.3.3), the inputs at present time and three further lags were input to the process models. (i.e. t, t-4, t-8 and t-12 hours). The list of input variables is shown in Table 5.8. Operational cost was calculated per one cubic metre of raw water. Only the chemical costs of alum and polyelectrolyte were accounted for in the optimisation process. Electricity and labour costs were not included.

7.5 Selection of a Representative for the Human Operated Clarifier

A clarifier to represent the human operated clarifiers needed to be selected. The performance of this clarifier was to be compared to the performance of the intelligent control system operated clarifier. Selection was based on similarity in performance measured by a correlation test for performance using past operational data of 718 runs from January to April 2008 (which was the four months' period before the full-scale pilot plant test). This was at the end of the winter close to the summer period when the raw water turbidity is low (in the range of 23 to 78 NTU and with mean±standard deviation of 37.00 ± 5.35 NTU). The performance of Clarifier number eight was not accounted for since its operational data was incomplete. The r^2 between the clarified water turbidity of Clarifier number nine and the other clarifiers are shown in Table 7.2 including the upset numbers (i.e. the upset number is the number of runs whose clarified water turbidity is more than the operational target of seven NTU). The associated means and standard deviation are also given. The associated r^2 and the upset numbers of these clarifiers are graphically plotted in Figure 7.7 and Figure 7.8 respectively.

All these clarifiers operated very well since the means of clarified water turbidity all satisfy the operational target of seven NTU. However it was found that clarified water turbidity of Clarifier number nine was the closest to that of Clarifier number ten. Although their r^2 was small (0.47), it was the largest in comparison with all the others. Some similarities might be expected and detected since they were of the same design as a medium size reaction well by Fuler & Smith, and each one was located next to each other. Therefore, as they are at the same distance from their respective head tanks this results in having hydraulic similarity.

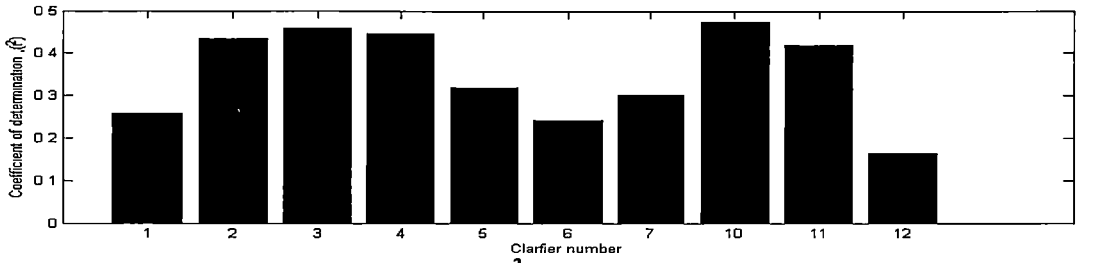


Figure 7.7 Coefficients of determination, (r^2) between the clarified water turbidity of Clarifier number nine to the others

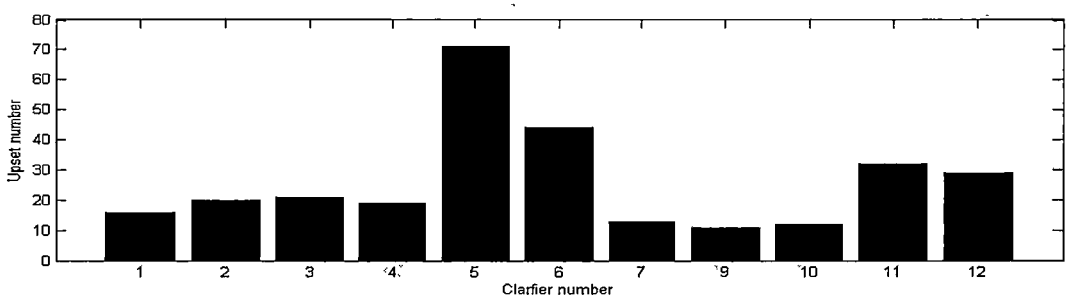


Figure 7.8 Upset numbers, presenting the number of runs of the clarifiers of the first production line which upsets the operational target of seven NTU. (Data from clarifier number eight was not available)

Clarifier Number	Coefficient of determination to Clarifier nine,(r^2)	Upset number	Clarified water turbidity ,NTU (Mean ± Standard deviation)
1	0.26	16	4.17±1.14
2	0.43	20	4.47±1.18
3	0.46	21	4.31±1.15
4	0.45	19	4.33±1.30
5	0.32	71	4.72±1.68
6	0.24	44	4.40±1.73
7	0.30	13	4.66±1.06
8	Data unavailable	-	-
9	1.00	11	4.45±0.92
10	0.47	12	4.54±0.97
11	0.42	32	4.87±1.19
12	0.17	29	4.33±1.61

Table 7.2 Coefficients of determination, (r^2) and performance of each clarifier in the first production line, showing r^2 between the clarified water turbidity of clarifier number nine and the others, together with their clarified water turbidity

Difficulties were recognised when the operators operated Clarifier number 5, 6, 11 and 12 as evidenced by the large upset numbers (as presented in Figure 7.8). Since their locations are near the dead end of the conduit (Figure 7.2), blocking the flow at the end of conduit always results in eddy flows which introduce the high turbidity to those clarifiers. In addition, the clarifiers with smaller reaction zone size (Clarifier numbers 5 and 6) are more sensitive to this effect than the larger ones (Clarifier numbers 11 and 12). This is because the larger size of their reaction zones produces more flocs which are able to buffer these situations.

7.6 Bang Khen Clarifier Full-Scale Pilot Plant Test: Method and Results

This section describes full-scale pilot plant method and compares the performance between human operators and the intelligent control system in terms of turbidity removal, operational cost and control actions in predefined ranges of raw water turbidity. The predictive performance of intelligent control system is also examined.

7.6.1 Pilot Plant Test Method

Clarifier number nine is operated by the intelligent control system. On the other hand, Clarifier number ten is selected as representing the human operator. Consequently, Clarifier number 9 is called an "Intelligent clarifier" and Clarifier number 10 is called a "Human clarifier".

For every four hours, the plant status (e.g. raw water qualities) in the format of temporal data of present, t-4, t-8 and t-12 hours was input into the intelligent control system. The intelligent control system then solved values for the optimum set of control actions (alum dosage, polyelectrolyte dosage, turbine speed and sludge drainage rate). These control actions were then applied to the intelligent clarifier plant. All the temporal data were informed via the telephone line. The control actions had to be applied manually to the plant, since the SCADA system was not available in BKWTP. From receiving the data to implementing the control actions, it took about 20 to 30 minute. Most of this delay was contributed by laboratory process, communication and manual implementation of the control actions. The data input and computational time was less than 5 minutes.

Raw water turbidity is recognized as one of the most significant factors in the performance of any clarifier (ASCE. & AWWA., 1990). Figure 7.9 presents raw water turbidity during the pilot plant test from September to August 2008. At the beginning of the period, raw water was at about 70 NTU and rose to its first peak of 145 NTU. After a while, the raw water again rose to the second peak of 192 NTU and it dropped to about 70 NTU at the end of the pilot plant test. Statistically, the maximum and minimum were 192 and 67 NTU respectively and their mean ± standard deviation was 118.94 ± 30.14 NTU.

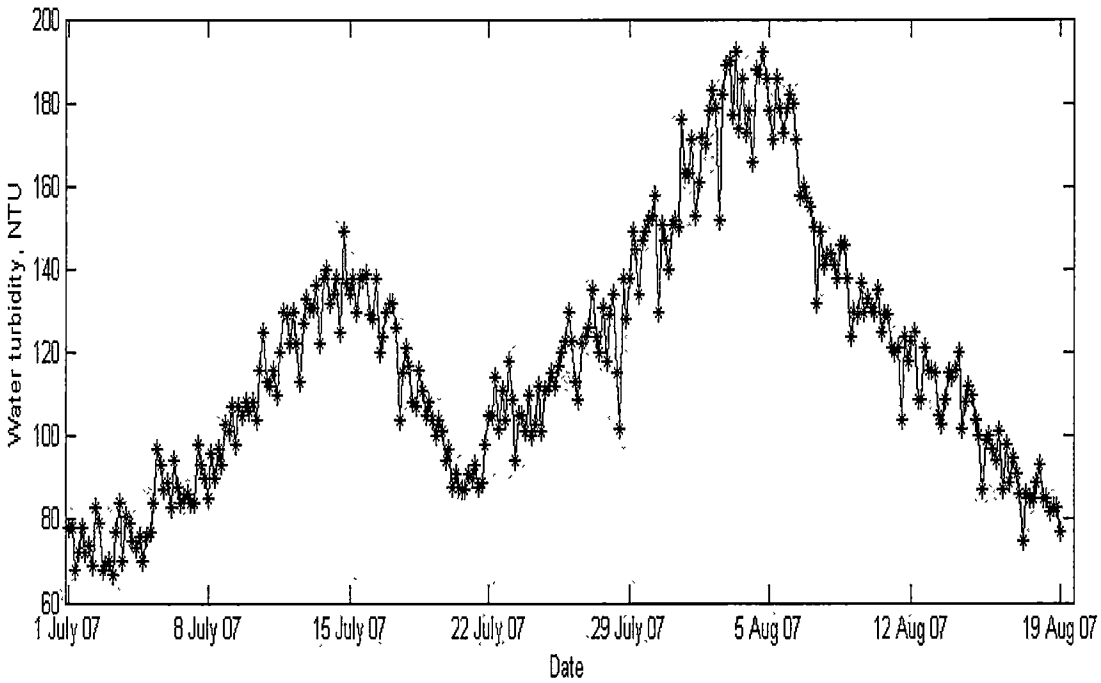


Figure 7.9 Raw water turbidity during the period of the full-scale pilot plant test

The performance in prediction and turbidity removal of the intelligent control system were compared with those of human operators in the two schemes according to raw water turbidity values: Firstly the comparisons were taken for the overall range of raw water turbidity and, secondly, in the separate ranges bounded by the mean ± standard deviation of the raw water turbidity of the historical operational data (93.58±35.10 NTU, Table 3.8) used to develop the intelligent controller. These comparisons and assessments were taken in two ranges. The first range was when the raw water turbidity was in the range of mean ± standard deviation (93.58±35.10 NTU) of the past operational data, (i.e. 58.90 NTU≤ raw water turbidity ≤ 127.64 NTU). For the second range, the assessments were done when the raw turbidity was distant from the means or when the turbidity was beyond the range of mean ±standard deviation as recorded from past operational data. However,

during the period of the full-scale pilot plant test, the raw water turbidity was never lower than 58.90 NTU. Therefore, the upper raw water turbidity value of 127.64 NTU was used as the only threshold.

7.6.2 Pilot Plant Predictive Performance

For the overall range of raw water turbidity, the predicted and actual clarified water turbidity of Clarifier number nine (controlled by the intelligent control system) are graphically shown in Figure 7.10. Generally, the predicted values of clarified water turbidity were lower than the actual values. This was evidenced by the smaller means of 5.94 NTU while the actual clarified water turbidity value was 6.36 NTU. Their associated standard deviations were similar, with the values of 1.59 and 1.54 NTU respectively.

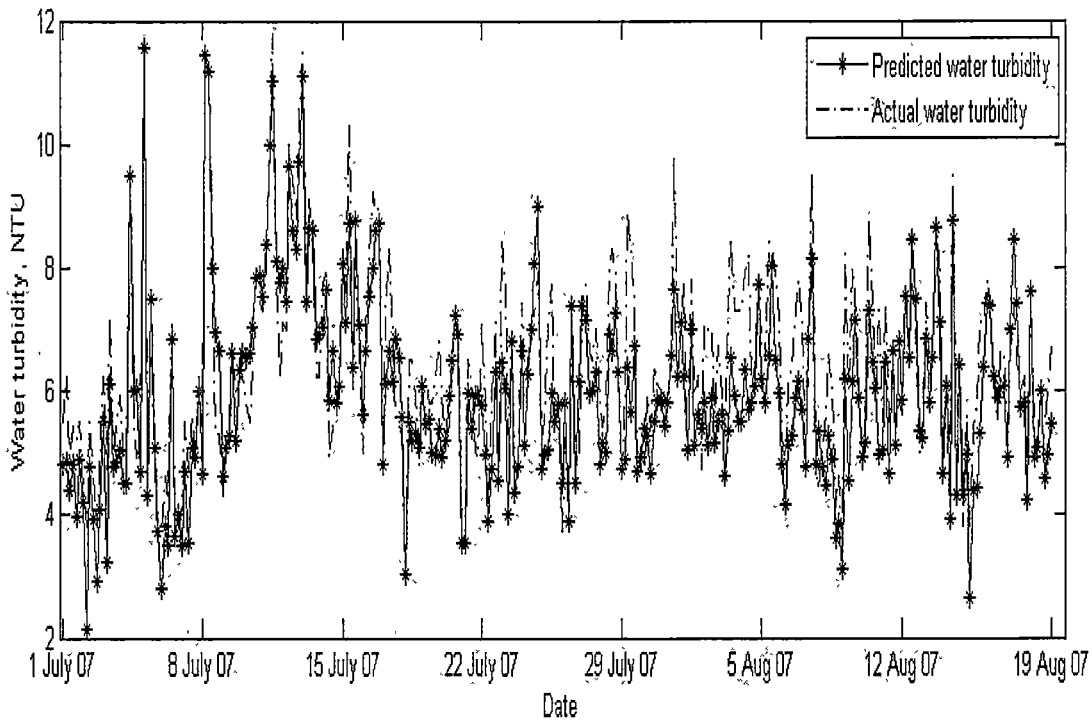


Figure 7.10 Predicted clarified water turbidity, showing the predicted versus actual turbidity of clarified water turbidity during full-scale pilot plant test

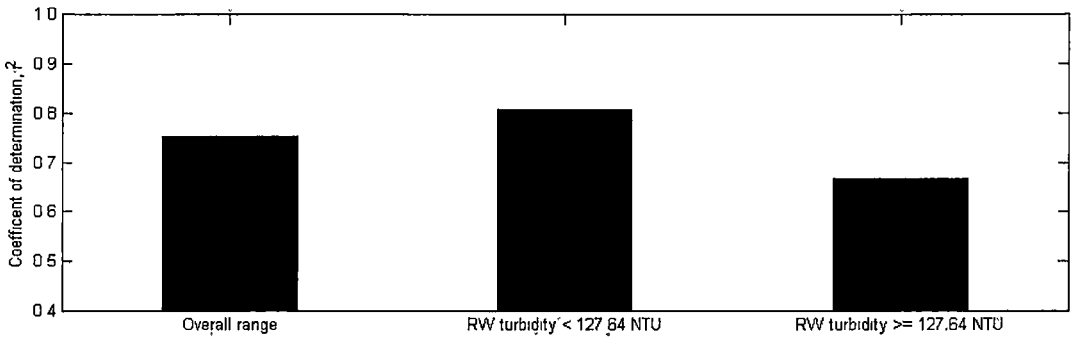
The r^2 is 0.75, and its associated MAE was less than one NTU (0.72 NTU). Although the predictive error was slightly high despite a moderate r^2, the ANN process model was able to recognise the occurrence of changes in actual clarified water turbidity. Figure 7.11 shows that the predictive reliability of the ANN model was marginally acceptable for the overall range of raw water turbidity since the percentage error of 11.31 percent was only just above the lower measurement error of the turbidity meter of ten percent (Table 3.6). Their associated statistical parameters are tabulated in Table 7.3.

Raw water turbidity range		Actual clarified water turbidity	Predicted clarified water turbidity
Overall range	Mean±Std (NTU)	6.36±1.54	5.94±1.54
	r^{2*}	0.77	
	MAE, (NTU)	0.75	
	Percentage error**	11.31	
Raw water turbidity less than 127.6 NTU	Mean±Std (NTU)	6.14±1.58	5.85±1.68
	r^{2*}	0.81	
	MAE, (NTU)	0.65	
	Percentage error**	10.60	
Raw water turbidity more than 127.6 NTU	Mean±Std (NTU)	6.73±1.46	6.09±1.27
	r^{2*}	0.67	
	MAE, (NTU)	0.84	
	Percentage error**	12.43	

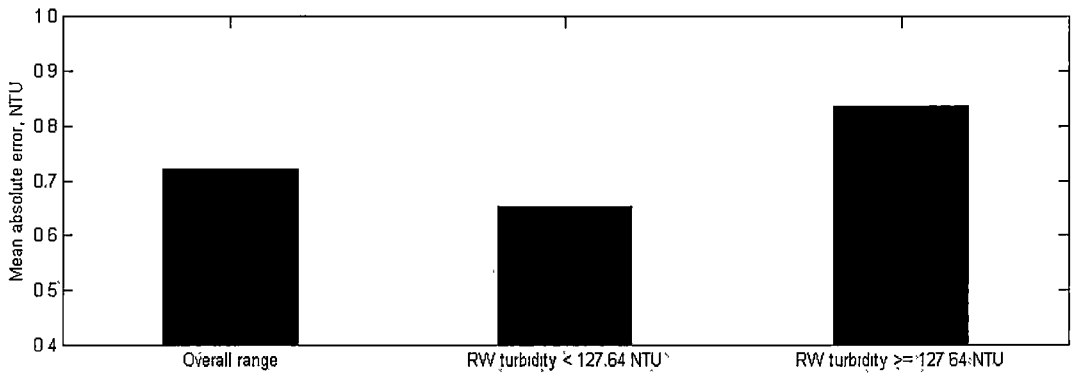
*r^2 is coefficient of determination, **Percentage error = (MAE/mean)*100

Table 7.3 Prediction performance of the intelligent control system, presenting the predictive performance of clarified water turbidity during the full-scale pilot plant test. The predictive performance was separately described in the "Overall range", "less than 127.6 NTU" and "more than 127.6 NTU" of raw water turbidity

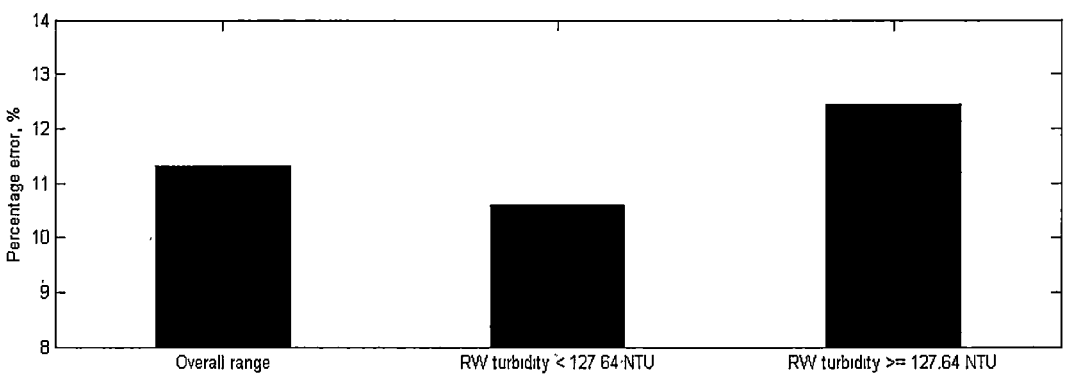
If the predictions were taken when the raw water turbidity value was less than 127.64 NTU, then the predictive performance was slightly improved. This was evidenced by the smaller MAE of 0.65 NTU and larger r^2 of 0.83. Given this better prediction, the predicted error was reduced to 11.07 percent. On the other hand, if the predictions were taken when raw water turbidity value was more than 127.64 NTU, then the predictive performance was less accurate. The associated MAE increased to 0.87 NTU and the ANN process model faced difficulty in recognising the trends of the clarified water turbidity (as evidenced by the decrease of r^2 to 0.69). The percentage error rose to 12.43 percent which was above the turbidimeter measurement error for an acceptable lower threshold (which was 10 percent). All of the associated statistical parameters are tabulated in Table 7.3. For comparison, the r^2 and MAE of the prediction over the range of raw water turbidity is graphically shown in Figure 7.11 a b, and c respectively. In each case, the prediction was shown for the overall turbidity range, and for turbidity within and outside of the historical data range.



a) Coefficient of determination (r^2)



b) Mean absolute error



c) Percentage error

Figure 7.11 Predictive performance, presenting the comparison of the predictive performance parameter in the separate ranges of "Overall range", "less than 127.6 NTU" and "more than 127.6 NTU"

a) Coefficient of determination (r^2), b) Mean absolute error , c) Percentage error

7.6.3 Clarified Water Turbidity

The clarified water turbidity of the intelligent clarifier (i.e. Clarifier number nine controlled by the intelligent control system) and the human clarifier (i.e. Clarifier number

ten controlled by human operators) are plotted in Figure 7.12. The associated statistic parameters are shown in Table 7.4.

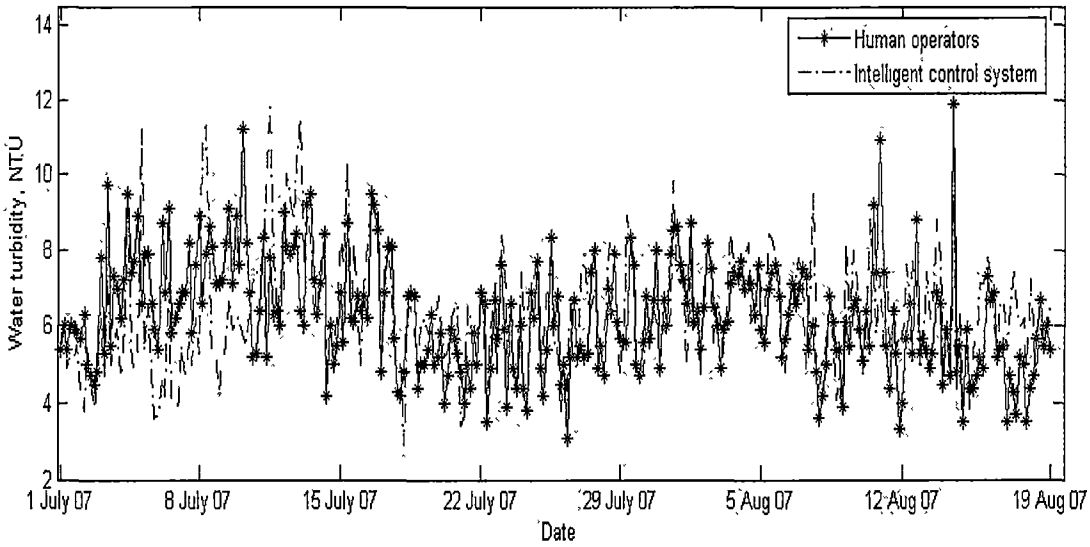
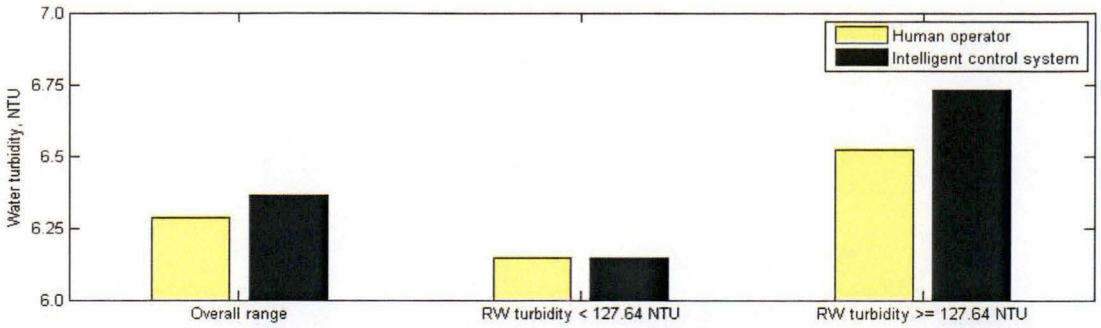


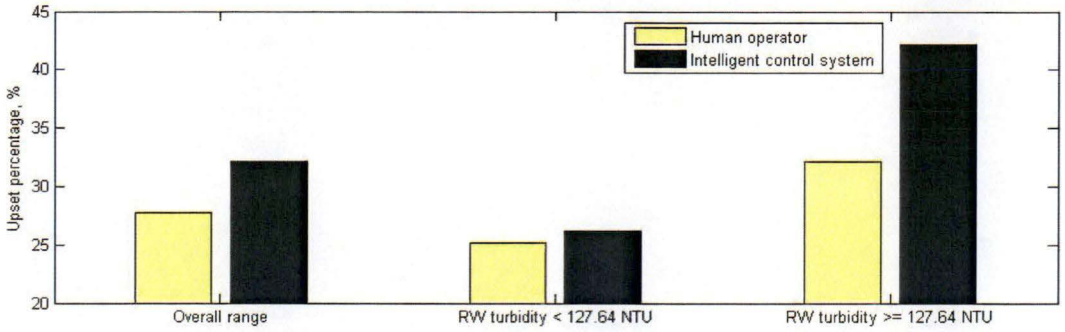
Figure 7.12 Clarified water turbidity; showing the clarified water turbidity values of the human clarifier and the intelligent clarifier during period of full-scale pilot plant test

For the overall range of raw water turbidity, both the intelligent control system and the human operators showed acceptable performance, since their overall means and standard deviations were 6.36±1.54 NTU and 6.28±1.47 NTU which were both less than the seven NTU target limit. The intelligent control system operated the clarifier with an upset number (the number of times the clarified water turbidity exceeded the operational target of seven NTU) of 95 runs from 296 runs (i.e. the upset percentage was 32.09 percent). On the other hand, the human operators performed better than intelligent control system by reducing the upset number to 82 from 296 runs (27.70 percent).

When the raw turbidity was less than 127.64 NTU, both human operators and the intelligent control system operated the clarifiers fairly well. Their associated means and standard deviations of clarified water turbidity showed almost identical values of 6.14±1.59 and 6.15±1.55 NTU, respectively. Both satisfy the operational target of seven NTU, with 187 runs in this range of raw water turbidity. Forty-seven runs (or 25.12 percent) upset the operational target when human operators operated the clarifier. On the other hand, 49 runs (or 27.70 percent) were upsets when the clarifier was operated by the intelligent control system. The associated clarified water turbidity and upset percentage in each range of raw water turbidity are shown in Figure 7.13 a and b.



a) Clarified water turbidity



b) Upset percentage

Figure 7.13 Turbidity removal performance, showing performance of the intelligent controller versus human operator in raw water turbidity ranges
a) Clarified water turbidity, b) Upset percentage

There were 109 runs in this turbidity range, where the raw water turbidity was higher than 127.64 NTU (i.e. beyond the range of mean ± standard deviation). Both the human operators and the intelligent control system operated the clarifiers satisfactorily. This was evidenced by clarified water turbidity with means and standard deviations of 6.52±1.23 and 6.73±1.46 NTU respectively. Although their raw water turbidity means were below the operational target of seven NTU, their associated upset percentages increased to 32.09 percent (i.e. 35 runs from 109 runs) for the human clarifier and 42.20 percent for the intelligent clarifier. Both results were higher than they were in the lower ranges of raw water turbidity. The performance of both the intelligent control system and the human operators are graphically shown in Figure 7.13.

An increased upset percentage indicated the existence of operational difficulties. The clarifiers in BKWTP were initially designed for maximum water turbidity of 100 to 150 NTU at normal plant flow rates. With such a high raw water turbidity range, the operators could only lessen the upset percentage by reducing the flow rate through the plant. However, the community demand for water has increased, so reducing the plant flow rate

could not be the preferred choice. As this constraint must be honoured, the operators were forced to optimise other control actions, and this was the main contributor to a raised percentage of upset values.

When the raw water turbidity increased beyond the training range of the intelligent control system (higher than 127.64 NTU), the upset percentage for intelligent control increased further than when using human operators. Unsuitable control actions occurred as the predictive performance of the process model decreased. When the predictions were employed beyond the range of the trained mean ±standard deviation of raw water turbidity, the predictive performance of the ANN model decreased because it had been trained inadequately in this range. It might be argued that a suitable prediction range should not be defined by using only raw water turbidity because the ANN clarifier process model was a multi dimension model in which the raw water turbidity was not the only input. However, the raw water turbidity was the input most closely related to the model's output (i.e. the clarified water turbidity). Its small coefficient of correlation, r of 0.29 is larger than any other inputs (Table 3.8). Additionally, raw water turbidity is obviously essential to successful clarifier operation (ASCE. & AWWA., 1990).

Raw turbidity range	**Parameters**	**Human operators**	**Intelligent control system**
Overall range	Clarified water turbidity Mean±Std (NTU)	6.28±1.47	6.36±1.54
	Upset number (run)	82	95
	Upset percentage (%)	27.70	32.09
Raw water turbidity less than 127.64 NTU	Clarified water turbidity Mean±Std (NTU)	6.14±1.59	6.15±1.55
	Upset number (run)	47	49
	Upset percentage	25.13	26.20
Raw water turbidity more than 127.64 NTU	Clarified water turbidity Mean±Std (NTU)	6.52±1.23	6.73±1.46
	Upset number (run)	35	46
	Upset percentage (%)	32.11	42.11

Table 7.4 Turbidity removal performance analysis, presenting performance of the intelligent controller versus the human operator in each range of raw turbidity.

7.6.4 Operational Cost

In the same manner, the operational costs will be discussed according to three ranges of raw water turbidity. The operational costs of varying raw water turbidity range are presented in Figure 7.14.

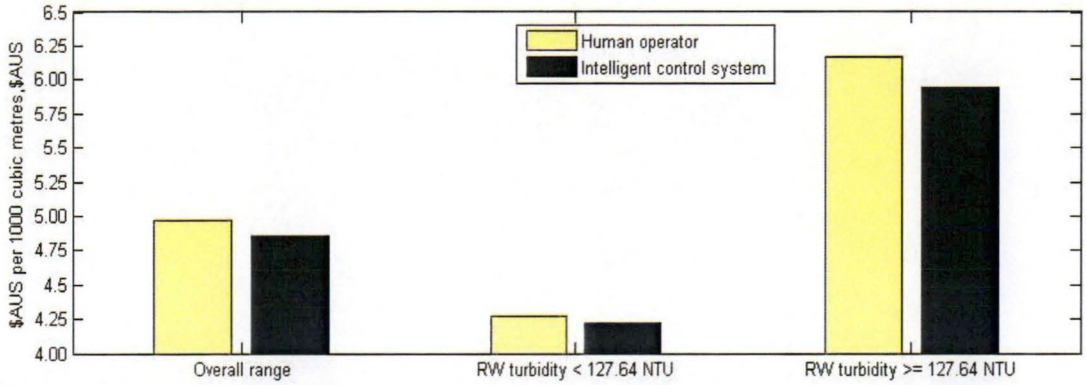


Figure 7.14 Operational cost, showing the comparison of operational cost between the intelligent controller and the human operator in different raw water turbidity ranges

For the overall range of raw water turbidity, the intelligent control system operated the clarifier at a cheaper cost compared with the human operators. Their operational costs (mean ± standard deviation) were $AUS 4.85±1.10 and $AUS 4.97±1.26 per thousand cubic metres respectively. Consequently, the intelligent control system could save the operational cost by as much as 2.44 percent.

When the raw turbidity was in the range of less than 127.64 NTU, the intelligent control system and human operators operated the clarifiers at almost the same cost. The human operators cost (mean ± standard deviation) $AUS 4.27±0.52 per one thousand cubic metres and $AUS 4.21±0.50 for the intelligent control system with a saving of only 1.25 percent.

However, when the raw water was in a higher range, more than 127.64 NTU, the intelligent control system could save the operation up to 3.71 percent. For one thousand cubic metres of clarified water, the human operators operated the clarifier at a mean cost of $AUS 6.17 and $AUS 5.94 for the intelligent control system. Their standard deviations are $AUS 1.25 and $AUS 1.00 respectively. The associated statistic parameters are tabulated in Table 7.5.

7.6.4 Operational Cost

In the same manner, the operational costs will be discussed according to three ranges of raw water turbidity. The operational costs of varying raw water turbidity range are presented in Figure 7.14.

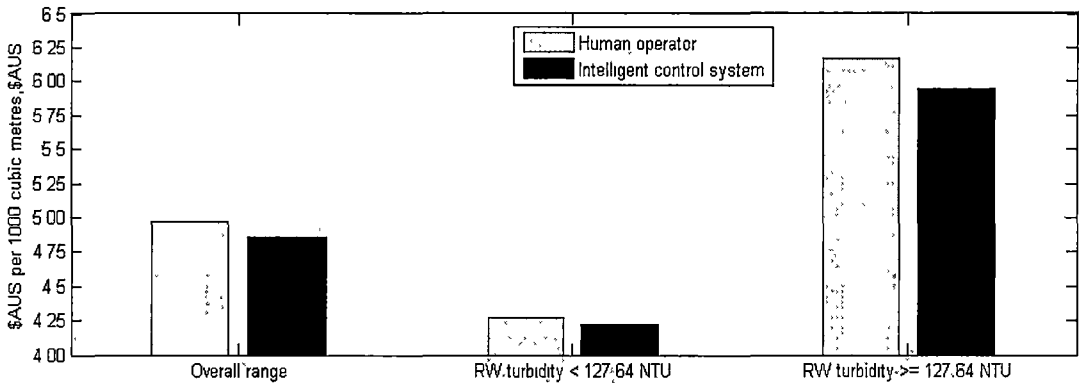


Figure 7.14 Operational cost, showing the comparison of operational cost between the intelligent controller and the human operator in different raw water turbidity ranges

For the overall range of raw water turbidity, the intelligent control system operated the clarifier at a cheaper cost compared with the human operators. Their operational costs (mean ± standard deviation) were $AUS 4.85±1.10 and $AUS 4.97±1.26 per thousand cubic metres respectively. Consequently, the intelligent control system could save the operational cost by as much as 2.44 percent.

When the raw turbidity was in the range of less than 127.64 NTU, the intelligent control system and human operators operated the clarifiers at almost the same cost. The human operators cost (mean ± standard deviation) $AUS 4.27±0.52 per one thousand cubic metres and $AUS 4.21±0.50 for the intelligent control system with a saving of only 1.25 percent.

However, when the raw water was in a higher range, more than 127.64 NTU, the intelligent control system could save the operation up to 3.71 percent. For one thousand cubic metres of clarified water, the human operators operated the clarifier at a mean cost of $AUS 6.17 and $AUS 5.94 for the intelligent control system. Their standard deviations are $AUS 1.25 and $AUS 1.00 respectively. The associated statistic parameters are tabulated in Table 7.5.

Raw turbidity range	Human operators	Intelligent control system
Overall range	4.97±1.26*	4.85±1.10* (Save up to 2.44%)
Raw water turbidity less than 127.64 NTU	4.27±0.52*	4.21±0.50* (Save up to 1.25%)
Raw water turbidity more than 127.64 NTU	6.17±1.25*	5.94±1.00* (Save up to 3.71%)

*$ AUS per one thousand cubic metres (based on 1 $ AUS = 28.00 Baht)

Table 7.5 Operational cost

When the raw water turbidity increased, the operational costs increased noticeably in both clarifiers, whether controlled by a human operator or by the intelligent control system. This was as expected since more chemical agents are used to precipitate the turbidity. Although this occurred quite commonly, it confirms that sensible control actions were applied.

It has so far been assumed that the chemical costs were the only contribution to the operational costs. The electricity and labour costs have been ignored. The alum unit cost was much cheaper than the polyelectrolyte unit cost. However alum was used in much larger amounts (10 to 75 mg/L) than that of polyelectrolyte (0.0 to 0.08 mg/L) (Table 3.7). Intelligent control reduced the total operational cost by reducing the alum dosage and increasing the polyelectrolyte dosage. Therefore employing this strategy would save operational costs to a figure below that of human operators in all raw water turbidity ranges.

7.6.5 Control Actions

Four control actions (alum dosage, polyelectrolyte dosage, turbine speed and sludge drainage rate) were used by the human operators and the intelligent control system to control their relevant clarifier. The changes of control actions were discussed in two scenarios: (i) their magnitudes were compared between human operator's and the intelligent control system's, and (ii) the coefficient of correlations (r) of these control actions to the raw water turbidity were used to check whether the set of control actions assigned makes sense.

Magnitude of control actions

In the first scenario, consideration was given to values over the overall range of raw water turbidity and then separately when the raw water was less than or more than 127.64 NTU. All the associated mean ±standard deviations and coefficients of correlation (r) to raw water turbidity in each range are shown in Table 7.6.

RW turbidity range	Controller	Alum dosage mg/L, (r)*	Poly. dosage mg/L, (r)*	Turbine speed RPM, (r)*	Sludge drainage rate sec /Hour, (r)*
Overall range	Human Operators	36.52±9.18 (0.92)*	0.037±0.012 (0.90)*	1.34±0.17 (-0.81)*	46.86±17.05 (0.80)*
	Intelligent control system	35.56±8.01 (0.94)*	0.039±0.013 (0.94)*	1.37±0.17 (-0.86)*	45.50±18.96 (0.78)*
RW turbidity ≤127.64 NTU	Human operators	31.42±3.84 (0.81)*	0.031±0.006 (0.71)*	1.42±0.12 (-0.69)*	37.43±8.28 (0.17)*
	Intelligent control system	30.95±3.57 (0.86)*	0.032±0.007 (0.81)*	1.45±0.14 (-0.71)*	36.23±6.83 (0.62)*
RW turbidity >127.64 NTU	Human operators	45.27±9.10 (0.88)*	0.049±0.013 (0.89)*	1.19±0.13 (-0.56)*	63.03±16.05 (0.80)*
	Intelligent control system	43.46±7.33 (0.89)*	0.052±0.01 (0.89)*	1.22±0.10 (-0.81)*	61.39±22.32 (0.65)*

*Coefficient of correlation (r) to raw water turbidity

Table 7.6 Control actions, showing the magnitude of control actions and their coefficient of correlation (r) to raw water turbidity during full-scale pilot plant test

- **Alum dosage**

Figure 7.15 presents the alum dosages which were fed by human operators and the intelligent control system during the period of the full-scale pilot plant test. For the overall range of raw water turbidity, the associated means and standard deviation of alum dosage for the human and the intelligent clarifier were 36.52 ± 9.18 and35.56±8.01mg/L, respectively. That was the human operator fed more alum than the intelligent controller by about 2.71 percent.

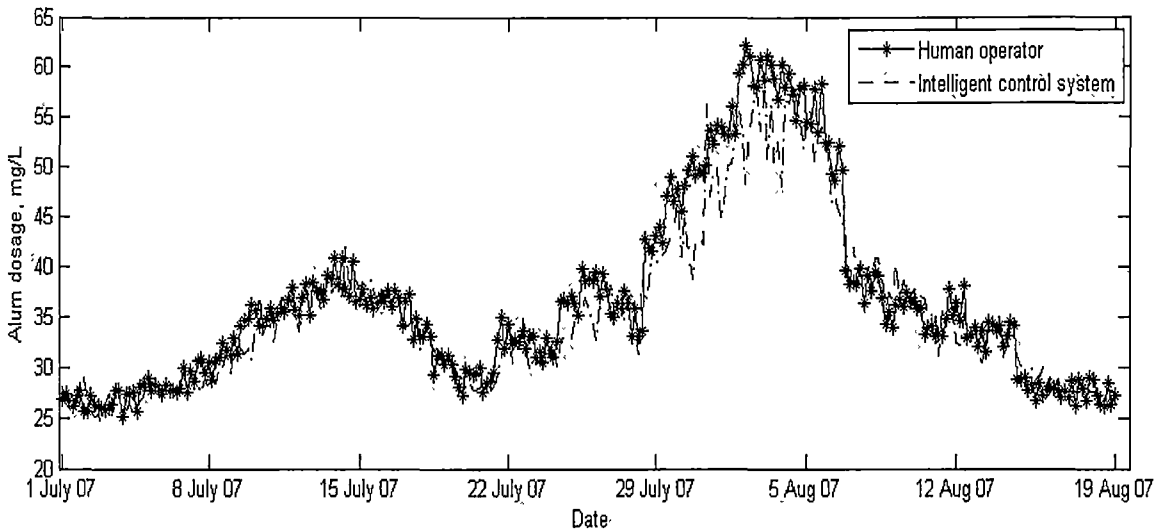


Figure 7.15 Alum dosages, showing that alum dosages were fed by the intelligent controller and the human operator during period of full-scale pilot plant test

The human and the intelligent controllers fed alum proportionally to the raw water turbidity with coefficients of correlation (r) of 0.92 and 0.94, respectively. Both correlations (r) being positive implied that the alum dosages changes in the same direction as raw water turbidity changes.

The human operator dosed more alum than the intelligent controllers by about 1.45 percent for the lower range of raw water turbidity (less than 127.64 NTU) as evidenced by their associated means ± standard deviations of 31.42 ± 3.84 and 30.95 ± 3.57 mg/L, for the human and intelligent clarifier respectively (Table 7.6).

When raw water turbidity was more than 127.64 NTU, the human operators used 4.24 percent more than intelligent controller. Their associated means ± standard deviations were 45.27±9.10 and 43.46±7.33 mg/L, for the human and intelligent clarifier respectively (Table 7.6).

Both methods of control fed alum in similar patterns regardless of the range of raw water turbidity received. Human operators fed alum *pro rata* to the changes of raw turbidity. This was evidenced by their positive signs of correlation coefficient (r) of 0.81 and 0.88 for the lower and higher ranges respectively, and corresponding correlations (r) of 0.86 and 0.89 for the intelligent control system.

- **Polyelectrolyte dosage**

Figure 7.16 presents the polyelectrolyte dosages used by human operators and the intelligent control system during the period of full-scale pilot plant test. Human operators dosed polyelectrolyte in the range of 0.02 to 0.070 mg/L and the intelligent control system dosed in the range of 0.024 to 0.068 mg/L.

Thus the intelligent control system used more polyelectrolyte than that of the human operators by about 5.41 percent since their means ± standard deviations, were 0.039 ± 0.013 and 0.037 ± 0.012 mg/L, respectively.

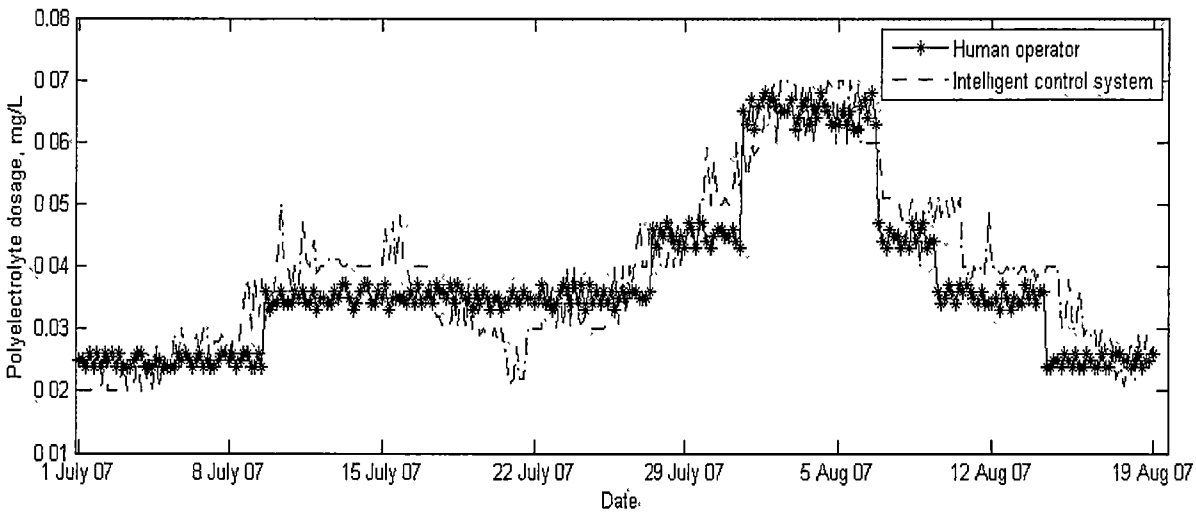


Figure 7.16 Polyelectrolyte dosages, showing the polyelectrolyte dosages were fed by the intelligent controller and the human operator during the period of the full-scale pilot plant test

For the lower raw water turbidity range, the human operators and the intelligent control system fed polyelectrolyte with the means ± standard deviations of 0.031 ± 0.006 and 0.032 ± 0.007 mg/L, respectively and similarly 0.049 ± 0.013 and 0.052 ± 0.01 mg/L for the higher range of raw water turbidity. Therefore, the intelligent controller dosed more polyelectrolyte than the human operators by about 3.23 percent in the lower turbidity range and 6.12 percent for the higher turbidity range.

Regardless of turbidity ranges, both controllers used polyelectrolyte in similar patterns. With changes of raw water turbidity, human operators dosed polyelectrolyte with correlation (r) of 0.71 and 0.89 for the lower and higher range of raw water turbidity

respectively, while corresponding figures for the intelligent controller were 0.81 and 0.89.

- **Turbine speed**

Figure 7.17 presents the values of turbine speed values which were assigned were in the range of 1.11 to 1.71 for human operators and 1.03 to 2.03 RPM as for the intelligent control system.

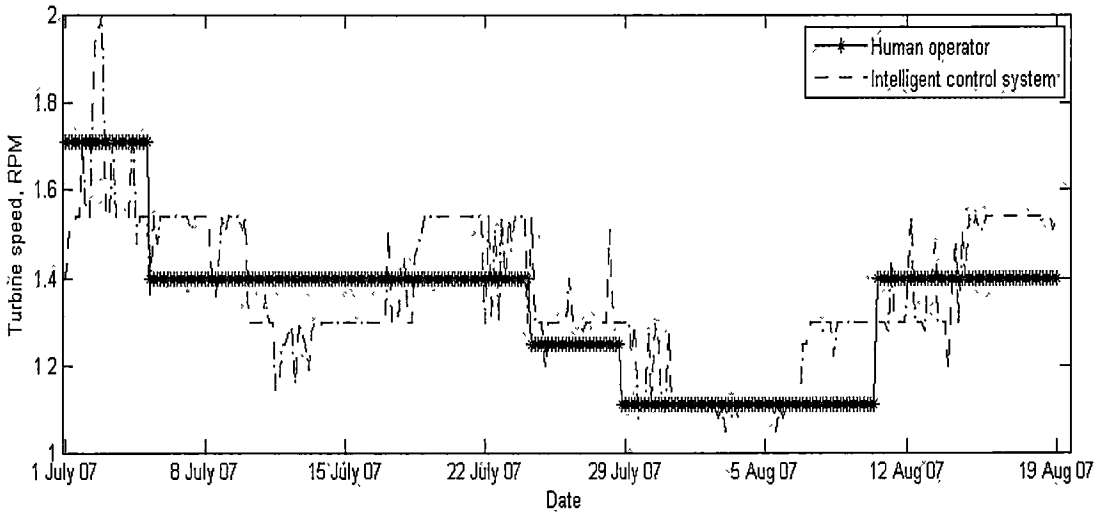


Figure 7.17 Turbine speed, showing the turbine assigned by intelligent controller and human operator during period of full-scale pilot plant test

For the overall range of raw water turbidity, the intelligent control system agitated the water by 2.24 percent more than the human operators do, as evidenced by the associated means ± standard deviations of 1.37 ± 0.17and 1.34 ± 0.17 RPM respectively. Their coefficients of correlation (r) to raw water turbidity were negative values of -0.81 (for human operator) and -0.86 (for intelligent controller). (Therefore changes in turbine speed are in the inverse direction to changes in raw water turbidity).

In the lower range of raw water turbidity, the human operator agitated the water more gently than the intelligent control system did by about 2.11 percent and 2.52 percent for the higher range of turbidity. In the lower range of raw water turbidity, their associated means ± standard deviations of turbine speeds were 1.42 ± 0.12 RPM and 1.45 ± 0.14 RPM which were assigned by human operators and the intelligent controller respectively, and that of 1.19 ± 0.13 RPM and 1.22 ± 0.10 RPM for the higher range of raw water turbidity.

Both human and intelligent controllers agitated the water in similar patterns regardless of turbidity ranges. Human operators assigned turbine speeds to change inversely with the changes of raw turbidity. This was evidenced by the same negative signs of correlation (r) of -0.69 and -0.56 for the lower and higher ranges respectively, and coefficients of correlation (r) of -0.71 and -0.81 for the intelligent control system.

- **Sludge drainage rate**

Figure 7.18 presents the sludge drainage rates which were assigned from the range of 30 to 80 sec/hour by human operators and 23 to 120 sec/hour by the intelligent control system.

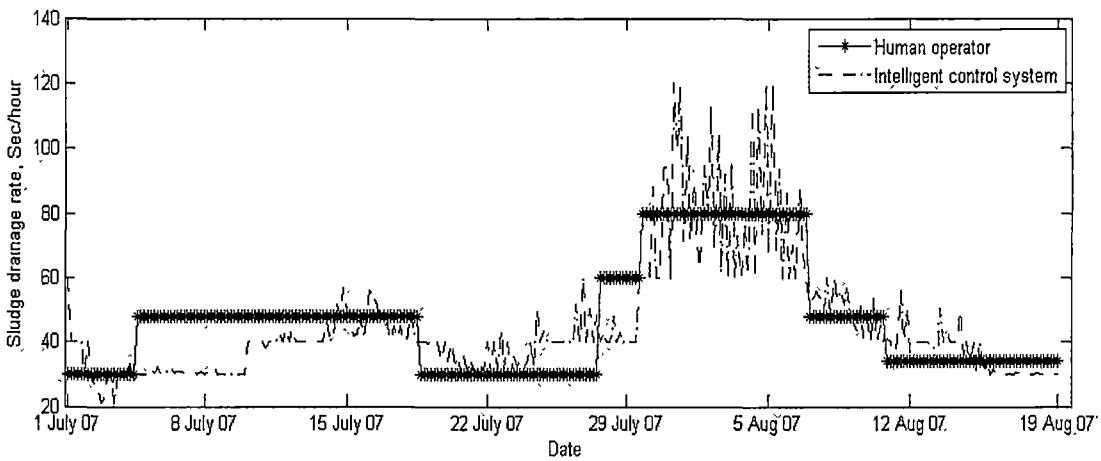


Figure 7.18 Sludge drain age rate

The human operates the clarifier by draining sludge at 2.90 percent more than the intelligent controller does. Their associated mean ± standard deviations are 46.86 ± 17.05 and 45.50 ± 18.96 sec/hour, respectively. Both the human operator and the intelligent control system changed the sludge drainage rate proportionally to raw water turbidity, with positive coefficient of correlations (r) of 0.80 and 0.78, respectively.

The human operators drained more sludge than the intelligent system controls did for both ranges of raw water turbidity. It was about 3.21 percent more for the lower range and 2.6 percent more for higher range of turbidity. Their associated means ± standard deviations were 63.03 ± 16.05 and 61.39 ± 22.32 sec/hour, respectively.

When the raw turbidity was in the lower range, human operators and the intelligent controller drained sludge in similar patterns proportional to raw water turbidity, with positive values of correlation (r) of 0.17 and 0.62, respectively. For the higher raw water turbidity range, human operators and the intelligent controller drained sludge directly in

relation to changes of raw turbidity. This was evidenced by the same positive signs of correlation (r) of 0.80 and 0.65 respectively.

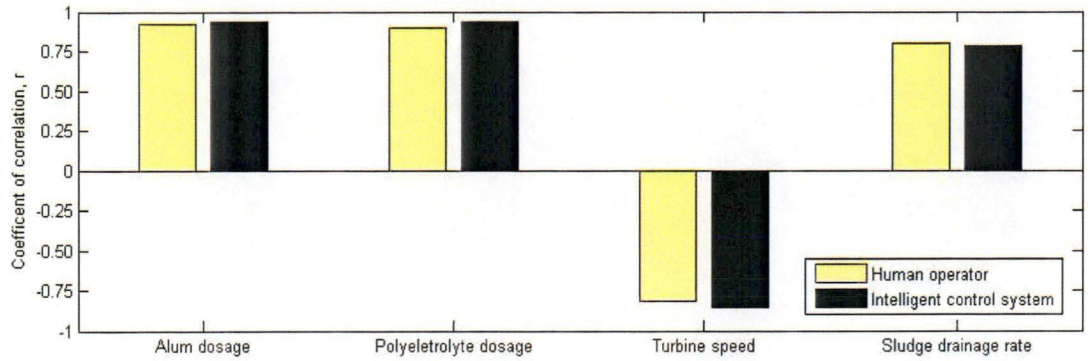
Regardless of the raw water turbidity range, the human operators assigned more alum dosage and less polyelectrolyte dosage, while the intelligent control system agitated with higher turbine speed but drained less sludge. It was known that higher chemical dosages and increasing turbine speed (within suitable ranges) enhanced floc amalgamation. The intelligent controller had to satisfy the operational cost constraint and thus it assigned a lower alum dosage which contributes more to operational cost than polyelectrolyte does. It compensated by dosing with more polyelectrolyte and using faster turbine speed. It also accepted more retained sludge by reducing the sludge drainage rate.

Rationality of control actions

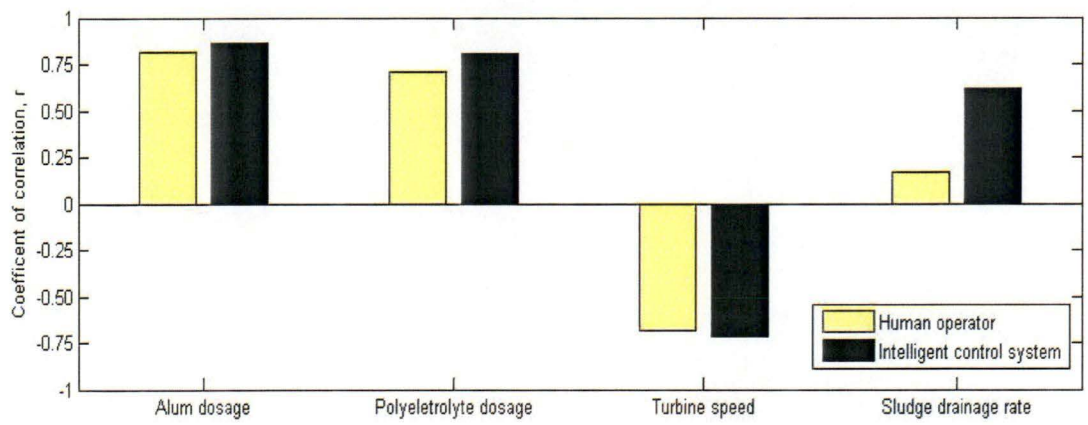
It was not straightforward to quantify how much each control action contributed to turbidity removal performance. Changing one control action might be compensated by changing another or a combination of the rest of the control actions. The interactions between each control action could be explained but not quantified in exact figures. In order to investigate how sensible the control actions were, the coefficients of correlation (r) of each one to the raw water turbidity were considered under two scenarios: (i) the overall range of raw water turbidity and (ii) two ranges of raw water turbidity when it was less or more than 127.64 NTU. Figure 7.19 presents the coefficient of correlations (r) to raw water turbidity of all four control actions assigned by the intelligent control system and human operators. The positive and negative sign of correlations (r) indicate the direction of each control actions relative to changes of raw water turbidity. The positive sign establishes a change in similar direction while the negative sign indicates an inverse relationship.

Regardless of raw water turbidity ranges (as illustrated in Figure 7.19 a, b and c), the pattern of control actions assigned by human operators and by the intelligent controller were very similar. The human operators and the intelligent control system both responded to the changes of raw water turbidity with a positive direction of change with chemical dosages and sludge drainage rate (noticed by the positive sign of correlation (r)) and with a negative direction with turbine speed (since their correlation signs are negative). For example, if raw water turbidity increased then both controllers added more chemical dosages and drained more sludge (positive sign of correlation) but slowed down the

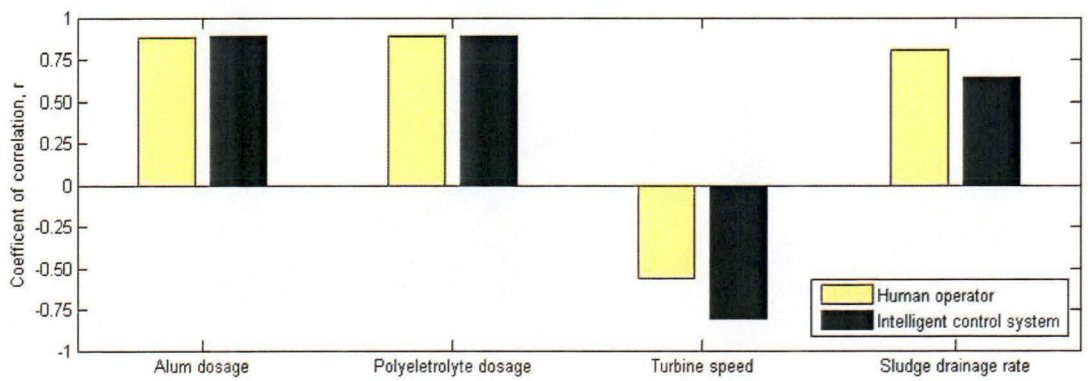
turbine speed (negative sign of correlation). Therefore, in practice these changes of control actions assigned by both forms of controllers were theoretically explicable (section 2.2).



a) Overall range of raw water turbidity



b) Raw water turbidity<127.64 NTU



c) Raw water turbidity≥127.64 NTU

Figure 7.19 Coefficients of correlation (r) between control actions and raw water turbidity: a) Overall range of raw water turbidity, b) Raw water turbidity<127.64 NTU and c) Raw water turbidity≥127.64 NTU

Although the intelligent controller assigned rational patterns, the magnitude of each control action was another important factor. Any mismatch of control actions provided by the intelligent controller might be contributed to by inaccuracy of prediction by the ANN process model, especially when the raw water turbidity was in the higher range of more than 127.64 NTU. The ANN process model had little experience in this region of high raw water turbidity which was beyond the training range. This would result in less predictive reliability and might drive the searching to a suboptimal solution space.

Additionally, both candidates operated near to, or over, the design limits since the clarifier was deigned for raw water turbidity up to 100 to 150 NTU. Also, only the human operators, but not the intelligent control system, had sludge concentrate information which could guide the selection of suitable control actions.

7.7 Chapter Conclusion

The main objective of the full-scale pilot plant test was to compare the performance of the intelligent control system and human operators. Clarifier number nine and Clarifier number ten were selected from twelve clarifiers all located in the first production line. Clarifier number nine was controlled by the intelligent control system and The Clarifier number ten was controlled by the human operators. Clarifier number nine was selected in order to avoid hydraulic bias since it is located at the middle of the production line and is of median size in the reaction well. Clarifier number ten gave a performance similar to number nine, as evidenced by the largest coefficient of determination (r^2) of 0.47 to clarified water turbidity of Clarifier number nine. Thus Clarifier number ten was used to represent the human operators' work during the full-scale pilot plant test.

For the overall range of raw water turbidity during the full-scale pilot plant test, the predictive performance of the clarifier process model was reliable. Since the percentage error (11.31 percent) was only marginally more than the lower turbidimeter measurement error (10 percent). The predictive reliability deteriorated when predictions were taken at the high turbidity range (more than 127.64 NTU.) when the percentage error rose to 12.48 percent. The ANN model might not have been sufficiently trained at this high range of turbidity which was above value of the mean plus the standard deviation of raw water turbidity available from past operational data. However, reliability remained acceptable due to the fact that its percentage error was only just above the lower limit of

turbidimeter measurement error of 10 percent. On the other hand, when prediction took place in the lower region of raw water turbidity (less than 127.64 NTU.), where the ANN model was well trained, the reliability was more acceptable, with a decrease in percentage error to 11.09 percent which was even closer to the turbidimeter measurement error.

In terms of turbidity removal performance, both human operators and the intelligent controller operated their clarifier equally well in the overall range of raw water turbidity with the mean ± standard deviation of clarified water turbidity of 6.28 ± 1.47 and 6.36 ± 1.54 NTU, respectively. Both achieved a standard of clarified water turbidity below the operational target of 7 NTU. However, the human operators operated the clarifier very slightly better than the intelligent controller with 1.60 percent lower mean clarifier water turbidity.

When the upset percentage was considered, it was found that the intelligent control system and human operators again showed similar performance in the lower range of turbidity. However in the high raw water turbidity range, the human operators exhibited superior performance to the intelligent control by lessening the upset percentage by 10 percent. That the upset percentage rose implies a performance deterioration by the intelligent controller. The performance of the intelligent control system was largely dependent on the predictive performance of the clarifier process model. In this high range of raw water turbidity, the process model might not be trained well enough. Predictive errors by the process model might result in misleading it to suboptimal solutions. Both human operators and intelligent controller used raw water qualities in common to achieve the optimal set of control actions. The human operators had guidance to the sludge concentration as one extra piece of information for control optimisation.

Although its turbidity removal performance was somewhat inferior, the intelligent control system operated the clarifier at a more economical cost than the human operators. The human operators paid about $AUS 4.97 per thousand cubic metres when the intelligent control system used only $AUS 4.85 per thousand cubic metres. The intelligent control system managed to economise on the chemical cost, though the electricity and labour costs were not accounted for in the operational cost. Due to the existence of a cost constraint, the optimal strategy involved more use of polyelectrolyte and less alum dosage because the alum was used in much greater amounts than polyelectrolyte even

though its unit cost was cheaper. During the period of the full-scale pilot plant test, the intelligent controller saved the operation 2.41 percent of cost. At this rate, if the intelligent control system were to be installed at all the clarifiers in BKWTP, the operational costs would save about $AUS 153000 per *annum* at the full production capacity of 3.5 million cubic metres per day. The reduction of labour cost for human operators is another reduction which was not included but employment would be reduced if this policy were to be introduced.

In respect to the of control action magnitudes, the intelligent control system reduced chemical use by selecting less alum dosage but more polyelectrolyte dosage to keep total cost low. To minimise the overall operational cost, the intelligent control system drove the control scheme to use less alum. For the physical control actions, the intelligent controller agitated more intensively to compensate for a lesser use of alum. With less alum used, there was not much sludge to drain, and the intelligent control system could assign a lower sludge drainage rate. The rationality of control actions was measured using coefficients of correlation (r) against raw water turbidity. In theory the intelligent control system responded to changes of raw water turbidity in a way similar to that of a human operator. Alum and polyelectrolyte dosages were assigned to alter in the same direction as changes of raw water turbidity but in an inverse direction for turbine speed and sludge drainage rate.

The feature of using more polyelectrolyte and less alum dosage in order to minimise the operational cost also exists in the simulation. Referring to the simulation results, the intelligent control system reduces the alum by about 4.1 percent and by much more than of 25 percent for the polyelectrolyte dosage (Section 6.6.4). The simulation is based on the training data (which was collected during the period of 2003 to 2004) when no occurrence of severe algal bloom phenomenon. After 2006 when sand mining in Chao Phraya River was banned, the water in the river became clear and algal bloom happen quite frequently and the human operator dealt with this problem by increasing the chemical dosages (especially polyelectrolyte dosage, section 7.3.4) so that a large difference of polyelectrolyte (which was found in the simulation) did not exist during period of the full-scale pilot plant test.

According to the pilot plant test results, it can be concluded that the intelligent control system is able to be used in practice since its clarified water turbidity satisfied the

operational target. It also proved that, even though the intelligent control system was designed to work with a SCADA system, it could also work with a manually controlled system in the case of BKWTP where delays due to manual entry of data exist. The predictive performance of the ANN process model was a significant factor in the performance of the intelligent controller. If the intelligent control system worked in the range that the ANN has previously experienced, the optimiser would be led to a suitable solution space.

Although the turbidity removal performance of intelligent control system was marginally less than that of the human operator, its clarified water turbidity still satisfied the operational target. In practice, the cheaper cost of operation is a major advantage with only a small compromise in poorer turbidity removal performance.

8.0 CONCLUSION AND RECOMMENDATIONS

8.1 Introduction

The overall goal of the research was to apply intelligent approaches to clarifier process modelling and control. Artificial neural network and Genetic algorithm techniques were merged to perform model predictive control of clarifiers. ANNs were used for modelling the clarifier process and GAs were used to optimise control actions in order to minimise the operational cost and clarified water turbidity and colour .

Clarifier process models were developed for different water treatment plants. The first case study is the Bryn Estyn Water Treatment Plant (BEWTP) in Hobart, Tasmania and the second is the Bang Khen Water Treatment Plant (BKWTP) in Bangkok, Thailand. They differ widely in environmental and operational conditions. BEWTP is fully operated under a SCADA system whereas BKWTP is manually operated. For both case studies, simulations of intelligent control were conducted. A pilot plant test was performed at BKWTP on one of the sixteen clarifiers. It was not possible to perform a pilot plant test at BWETP on their single clarifier due to their desire to keep control of clarifier operation. Also in BEWTP the full set of data concerning possible control actions was not available.

The first part of this chapter will outline the discussion and applications which are presented in the preceding chapters. The second part provides a summary of the major findings from the various chapters. Finally, recommendations for further study with respect to potential future directions are presented.

8.2 Discussion and Application Potential

A common misconception about intelligent approaches is that they are only applicable to large units with fully electronic linkages or those controlled under a SCADA system. BKWTP is about to install SCADA. However during the period of pilot plant testing, the intelligent control system was merged with manual control. This offered a good opportunity to not only compare the performance between the human operator and intelligent control system but also to assess the performance when the intelligent control system worked without SCADA.

The ANN process model replicates how the clarifier process works. Therefore the inputs to the model are the raw water qualities, control actions, and process variables. The outputs of the model were clarified water turbidity for the case of BKWTP. For BEWTP, two process models were used since clarified water colour and turbidity were the operational targets and set as outputs of the process models.

8.2.1 Data Manipulation and Analysis

Quality of input data to the ANN modelling was achieved through utilising extensive plant data records. Nonsense data was removed and substituted by interpolated data. The data filtering criteria used was set according to the suggestion of the expert operating staff. Most of the data was normally sampled every four hours (i.e. twice during a working shift). Data from different sampling rates was interpolated to a standard sampling rate of every four hours. This work is presented in Chapter 3.

8.2.2 Clarifier Process Modelling and Performance Evaluation

The clarifier process modelling is composed of three steps. The first one is input and output selection, and the second step is concerned with data clustering to ensure that training, validation and testing data sets are statistically equivalent. Finally, the model architectures are optimised using a trial and error based approach. After various candidate models were obtained, the predictive reliability and performance were evaluated for single time step and for multiple steps (long range) prediction. This is presented in Chapter 5.

The ANN process model replicates how the clarifier process works. Therefore the outputs of the model are clarified water turbidity in the case of BKWTP, and clarified water colour and turbidity for BEWTP. In common, the raw water qualities, control actions and process variables are used as the inputs for these models. However, the sludge drain rate data is not included as this data was unavailable.

To select the inputs for the ANN clarifier model, it is necessary to consider whether the type of inputs and their time lags are appropriate. To select the type of inputs, both the coefficient of correlation (r) between clarified water qualities and the inputs and also expert knowledge based on operator experience guided the choice of which types of inputs were suitable. However, the results showed that the coefficients of correlation

were small for both case studies and this reflected how complicated the relationships were. It was inappropriate to disregard those variables which have small correlation values since they are known to be significant to the clarified water qualities. Indeed all the operators in both BKWTP and BEWTP use these to guide the choice of control actions. Consequently all available variables were used for the input of the process models.

The choice of optimal time lag was guided by the Autocorrelation Function (ACF). Any lags which gave an ACF less than 0.50 were ignored. They were assumed to be insignificant and negligible. For both case studies, the maximum time lags were set to 20 hours (i.e. present time plus five time lags). In this way, maximum time lags were already defined and set as the maxima for further trial and error processes.

Before the ANN modelling process, training, validation and testing sets needed to be defined. To ensure that these three sets were statistically equivalent, a clustering technique based on Self-Organising Map (SOM) networks was successfully employed. The statistical hypothesis of T-test and F-test were applied to each variable with a maximum time lag to check the similarity of these three sets. The results revealed that they were statistically equivalent after a clustering method based SOM network had been applied. Their ranges were also checked to establish whether the validation and testing sets were subsets of the training set.

The optimal time lag, number of hidden layers and the number of neurons in each layer are problematic and difficult to assess. However these can also be obtained by trial and error processes. One and two hidden layers were used together with the number of neurons, which were varied from five to one hundred neurons (in increments of five neurons). Systematic trial and error processes were conducted for each time lag. The best architecture was the one that minimises the Mean Absolute Error (MAE) and maximised the coefficient of determination (r^2). All potential architectures were tested on the associated test sets. For the case of BEWTP, a temporal span of 8 hours and two layers with 95 neurons in the first and second hidden layers gave the best performance for predicting clarified water colour. For modelling the clarified water turbidity, two layers with 35 neurons in the first and 5 neurons in the second layer with a temporal span of 8 hours proved to be the best architecture. On the other hand, two layers of 20 neurons in

the first and 5 neurons in the second layer with a temporal span of 12 hours is optimal for prediction of clarified water turbidity in the case of BKWTP.

Not only a smaller MAE and a higher coefficient of determination were used to quantify the performance of the process model. The percentage errors in the predicted outputs were compared with the associated measurement errors. It appeared that all three ANN process models are reliable since their percentage error is just marginally greater than the associated measurement errors. In one particular case, the clarified water colour model of BEWTP, its predictive error is less than the measurement error.

The predictive performances were assessed not only for one step prediction, but assessment was also made of the multiple step (long range) prediction performance of all three process models. Their reliability and performance deteriorated severely after prediction of only one step since the prediction errors accumulated for every step.

Establishing of two process models for clarified water colour and turbidity each with optimal architectures and time lag for BEWTP and one process model for clarified water turbidity of BKWTP were the main contributions of Chapter 5, together with their performance evaluations. Also the testing sets obtained by SOM based clustering (which are statistically equivalent to their associated operational data) were included in this chapter. These testing sets were also used in the control simulation. As it had been shown that going beyond one step prediction results in unreliability, the control horizon was limited to one step of only four hours ahead.

8.2.3 Clarifier Intelligent Control Simulation and Performance Evaluation

The intelligent control system is composed of the clarifier process model, the Multiobjective Genetic Algorithm (MOGA) optimiser and the decision maker. The MOGA is developed from GAs but is able to optimise multiple objectives. The clarifier process models predict the clarified water qualities and pass these to the MOGA optimiser. Both of them are linked by the model predictive control scheme. In order to avoid ill posed problems, the control actions are bounded in accordance with raw water qualities. The control and predictive horizons are also set to only one time step of four hours ahead. The MOGA optimiser uses the predicted clarified water qualities to

optimise the control actions in order to minimise the water impurity and operational cost. Pareto-optimal fronts are generated iteratively until the pre-defined number of generations is reached. The last Pareto-optimal front is then assumed to be the best one. The decision maker selects the solution from the last Pareto-optimal front using the smallest Euclidean distance from the utopia point.

The optimal number of generations and population size are important to the success of the MOGA optimiser in order to generate good quality Pareto-optimal fronts. Other elements of MOGA such as the crossover probability value are guided by previous research. The Pareto-optimal front is qualified in terms of diversity and convergence by the Pareto Front Quality Index (PQFI). A smaller PQFI value implies a better convergence and diversity. The optimal number of generations and population size was obtained by trial and error processes by a systematic change of their values. For the BEWTP, the smallest PQFI was given by 200 generations and a population size of 80. For BKWTP case studies, 400 generations and 100 populations gave the smallest value of PQFI.

For the control simulation using the associated test set, the intelligent control system for BEWTP's clarifier minimises the operational cost, the clarified water colour and turbidity. The control actions are bounded according to raw water colour values in order to avoid ill posed problems. The intelligent control system works with predictive and control horizons of one four hour time step ahead. It was evident that the intelligent controller successfully controls the clarifier. The simulation results reveal that its clarified water colour and turbidity were below the operational targets, and its operational cost was considerably cheaper than that of human operators (by about 8.33. percent).

In comparison with the human operator, intelligent control gave better clarified water colour and turbidity as well as a lower percentage of upsets. It is hard to compare the magnitude of control actions since the human operator uses all the control actions but the intelligent control does not use the sludge drainage control (as historical data was not available). However, the intelligent controller uses a smaller quantity of chemicals but agitates the turbine faster. Using less chemicals satisfies the objective of minimising the operational cost, but impurity removal performance may be compromised by maintaining floc amalgamation through agitating the water more rapidly.

In the BKWTP simulation study, the control actions were optimised to minimise the clarified water turbidity and operational cost. Ill posed problems were also avoided by bounding of control actions in accordance with the raw water turbidity range. Simulation results showed that the turbidity removal performance satisfied the operational target. Its clarified water turbidity was less than seven NTU and slightly less than that of human operators. The intelligent controller also reduced the operational cost by 3.45 percent compared to human operators. In comparison with a human operator, the intelligent controller doses more polyelectrolyte but less alum in order to satisfy the objective of minimising the operational cost. It compensates for the smaller amount of alum (which is the major chemical agent) by introducing a higher turbine speed and keeping more sludge in the clarifier by reducing the sludge drainage rate.

8.2.4 Full-Scale Pilot Plant Test and Performance Evaluation

Testing of a Full-scale pilot plant was conducted at BKWTP for a period at the end of the early rainy season (July to August 2007). All model elements, the MOGA parameters, the optimal number of generations and populations were identical to those used for the BKWTP simulation. The control actions were bounded according to raw water turbidity values. The performance of the intelligent control system was compared with that of the human operators. The intelligent controller was installed at Clarifier number nine located at the centre of the first production line. This position was chosen to avoid the hydraulic bias. Clarifier number ten was chosen to represent human operators because of its similarity in performance to Clarifier number nine.

Fortunately, the raw water turbidity during testing the full-scale pilot plant is in the range of 70 to about 200 NTU. This range covered the training domain and at times also extended outside the training domain of the clarifier process model. This allowed measurement of the performance of the control system when the intelligent control lies in a well trained range and when it is less guided by previous experience. The training domain was defined according to the range of raw water turbidity from the operational data. If the raw water was in the range of mean ± one standard deviation of the raw water turbidity, then it was defined to be inside the training domain.

Both human operators and the intelligent control system satisfied the operational target of seven NTU, with closely matching performance. This was evidenced by the means of

clarified water turbidity of 6.28 and 6.36 NTU respectively. Although it had slightly worse turbidity removal performance, the intelligent controller reduced operational cost by about 2.44 percent. For the full capacity of 3.5 million cubic metres per day, this translates to savings of up to $AUS 153,000 per *annum*. Labour cost savings are not included but would also add to this figure.

It appeared that turbidity removal performance relies on the predictive performance of the process model. When the intelligent controller works beyond its training range (i.e. raw water turbidity > 127.64 NTU), its performance deteriorates and it becomes inferior to that of human operators. The upset percentage of human operators is 27.70 percent which is less than that of 32.09 percent of the intelligent controller in this high raw water turbidity range. On the other hand, if the intelligent controller works inside its training range (i.e. raw water turbidity < 127.64 NTU), its performance is more or less the same as that of the human operator. The associated upset percentages are 25.13 percent for human operators and 26.20 percent for the intelligent control system.

The rationale of control actions is examined by considering the sign of the coefficient of correlation (r) to the raw water turbidity. The control actions assigned by the intelligent control system are sensible since the changes assigned are theoretically explicable and linked to the changes of raw water. In considering the magnitude of its control actions, it is found that although the intelligent controller uses a lower dosage of alum in comparison with the human operator, it compensates for the smaller amount by using a higher polyelectrolyte dosage. A lower alum dosage lessens floc amalgamation. The intelligent controller then makes up for this by agitating faster and reusing more flocs by draining less sludge. The strategy of using more polyelectrolyte but less alum has the effect of minimising the operational cost. Although the unit cost of polyelectrolyte is higher than that of alum dosage, more alum is used than polyelectrolyte. Therefore, a reduction of the alum dosage is more economical to the operational cost than reducing the polyelectrolyte dosage.

8.3 Summary of Major Contributions

Although a few "intelligent" approaches to clarifier operation exist, most of them are only concerned with the application of ANNs to predict water quality. The major contributions of this work are in the establishment and implementation of intelligent

control systems as well as with assessing their performance in real practice. The intelligent control system architecture developed is a combination of ANN modelling and a MOGA optimiser. This is the first time that the intelligent approach fully mimics how human operators control the solid contact clarifier. Both chemical and physical control actions are used by the intelligent control system to minimise the clarified water turbidity and operational cost. This work affirms that the ANN modelling approach is indeed feasible for these particular case studies and is therefore proved to be applicable in a real world situation, and then develops new methods to provide effective intelligent control of clarifier operation. The contributions based on the result of this research work include the following

- ANN process models are successfully developed based on the particular operational data of each of these two case studies. Included in this is the unique use of the SOM that performs successfully with the operational data at hand.

- The optimal model architecture of each process model is obtained based on the operational data, and their time lags are 8 hours for BEWTP and 12 hours for BKWTP. To achieve the best prediction of the clarifier the process model is found to need both present and past (temporal) data. Although this was expected, this study shows exactly how long a period of data is needed to optimise the control actions.

- The prediction of the clarifier process model deteriorates when the prediction is taken outside the training domain. One step ahead (4 hour) prediction is reliable when compared with the measurement error. However, it is shown that it is not practical for the ANN modelling approach to be used for multiple steps (long range) prediction because errors accumulate over each time step.

- An intelligent control system with a combination of an ANN clarifier process model and GAs has been applied here for the first time in the water treatment industry. It successfully mimics how human operators control a clarifier. Both the physical and the chemical control actions are optimised in order to minimise clarified water turbidity and operational cost. It has been shown that the intelligent control system shows promise for practical application. Its turbidity

removal performance satisfies its operational target at a cheaper operational cost than human operators.

- In this thesis, the "ill posed problem" is avoided by the division of operational range according to raw water turbidity. This method has been shown to be effective in real life application.

- The performance of the intelligent control system depended upon the predictive performance of the clarifier process model. When the intelligent control system is required to work outside its training domain, it has been shown that its performance is less satisfactory than when working inside its training domain.

8.4 Recommendation and Further Study

This thesis presents a number of findings and new developments for clarifier control. However, it is not possible to explore several areas due to the limitation of case studies. To encourage the development of intelligent control in the water supply industry, there are some recommendations and suggestions for further study, as follows.

8.4.1 Alternative Input Variables

For the intelligent controller based ANN model, the success of the controller is sensitive to the type of input. If some parameters which have a strong relationship to the output are set as the input of the process model, then they will enhance the performance of the controller. Some potential inputs are outlined as follows:

Sludge concentration

Sludge concentration in the reaction zone of the solid contact clarifier has been known to be important to clarifier control. For example, too much sludge concentration implies chemical over dosage. It is the only one parameter providing the information in the reaction zone. Physically, the sludge is retained in the clarifier much longer than hydraulic detention time. This provides the evidences of how well the clarifier was operated from the previous time steps to the present. It contains time dependent information. However, it was unfortunate that sludge concentration data was unavailable in BKWTP case study. If the sludge concentration were one of the inputs of the model,

this would potentially benefit the predictive performance of the model by reducing the temporal span.

Filter run time

The purpose of the clarifier unit is to entrap the impurities (e.g. turbidity, colour, and algae) prior to the filters in order to ensure a long filter run time. If filter run time is shortened, this will result in a need to back wash the filter more often: this is very expensive. If the filter run time were to become an input of the model, it might help the clarifier process model to recognise this situation so to drive the optimiser to a better solution.

It would also be possible to set the filter run time as an output of the integrated process model which combines the clarifier and the filter. Rather than having separate process models for predicting clarified water turbidity and other impurities (which requires costly computational expenses). It may offer improvement if the filter run time is used as a single output of the integrated process model. This is because the one objective of the clarifier unit is to prolong the filter run time for as long as possible.

Water temperature in the clarifier

Any large clarifier faces a density current which may often cause turbidity leaks. This is always the case in BKWTP. The temperature gradient through the depth of clarifier tank is the main problem here. If the water temperature in the clarifier is monitored especially in the sedimentation zone, the temperature gradient can be estimated. If this information were to be input to the process model, it should be able to recognise the density current phenomenon as it occurs.

Streaming current detector

One of the main aims of using alum is to neutralise the charge on the particles of impurities. The process of charge neutralisation is sometimes referred to as coagulation. *Streaming current potential* is the charge on the particles after coagulation. This is measured using a streaming current detector (SCD). The streaming current potential can be directly used to determine the degree of coagulation and this is related to chemical dosage. Of course, alum is also consumed in increasing the size of charge-neutral particles by absorption. If streaming current potential were to be input to the process

model, it would help the model to recognise the patterns of chemicals being used. Thus it should enable the model to guide the optimiser to a better solution.

8.4.2 Intelligent Controller Performance Improvement

To improve the performance of intelligent controller (i.e. reduce upset percentage), it is necessary to improve the predictive performance of the process model. Process models with larger prediction range of input and prediction accuracy are preferable. (i) Larger prediction range (i.e. the process models which show prediction reliability over a larger input range of raw water turbidity) is required to ensure that the ANN process model predicts in the interpolation mode rather than extrapolation mode which is unfavourable to the ANN model itself. (ii) One way to improve the prediction accuracy is to balance the complexity and flexibility of the process model. However, it is a very hard task for one process model to be responsible for the whole prediction range. There are some interesting techniques relevant to these two strategies which can be shown as follows.

Adaptive Process Model

An adaptive feature in the process model should be able to enhance the predictive performance of the model. If the model were to face a new pattern of data which was out of the original range then the prediction mode was now on the extrapolation mode. It was unfavourable to the ANN prediction performance. This problem might be alleviated by using the adaptive training. The new data would be marked and stored. Finally, they will be used to retrain the model (offline). This kind of attempt should benefit the predictive performance since it enlarges the range of the inputs and kept the prediction in the favourite mode of interpolation.

Multiple Model Approach

The prediction performance of the model has a significant role in the performance of the intelligent control system, the use of only one model for the whole operational range may not be optimal. The ANN process model is trained to fit the trend of the data, but with only one model, it is hard to recognise some details in the data. This relates to the "complexity" and "flexibility" of the model. A multiple model approach is therefore worthy of investigation. However, before using a multiple model approach, the operational ranges of each model need to be defined by the boundaries of the inputs. This may not suit this research since the clarifier process models of both case studies have a

lot of inputs (i.e. large dimension). A multiple model approach could be investigated to assess whether this procedure would improve the performance of the intelligent control system.

8.4.3 Model Parsimony and Sensitivity Analysis

In this work, the process models are considered large dimension since there are a number of inputs. Perhaps reducing a number of inputs helps decreasing noise and minimises computational expenses. However, it has to ensure that predictive performance is maintained. Sensitivity analysis can be used to check which inputs should be eliminated. However, the suitable threshold value of sensitivity percentage is problematic and can possibly be set via the trial-error process. This would provide an opportunity to compare the predictive performance of the smaller model with the larger ones.

8.4.4 Plant-wide Intelligent Control

Since the performance of each treatment unit in a production line affects each of the other units, the intelligent control system could be perhaps applied to every unit of the water treatment plant or even to the whole treatment process. This would provide an opportunity to compare the plant-wide performance and its operational cost to the existing water treatment plant.

8.4.5 Practical recommendations

In practice, there are some useful recommendations for clarifier control such as mixing intensity (in the reaction zone). Mixing intensity is described in a dimensionless term of G*T. G is gradient velocity (sec^{-1}) which is a function of power input and shape of turbine blade and T is the time that water is retained in the reaction zone (i.e. hydraulic retention time (sec)). However, gradient velocity could not be found in this research since the power input was never recorded and the drag coefficient of this custom turbine was not available. This drag coefficient can only be obtained via hydraulic model testing in conjunction with dimensional analysis. It should be very interesting in the future works to discover how the intelligent control provides the mixing intensity (i.e. G*T) and the suitable range of G*T.

Sludge blanket height is the other interesting variable used to control the clarifiers. Typically, sludge blanket height has been used for turbidity leakage warning. If the

height of the sludge blanket reaches the setting limit (e.g. one meter from the surface of water), then it implies that the clarifier is losing its performance. This can be caused by improper dosage and mixing intensity. Unfortunately, there is no sludge blanket height sensor equipped at BKWTP. Further than turbidity leakage warning, if the sludge blanket height sensors were uniformly equipped through the clarifier, then the contour of sludge blanket would be available online and reflected the uniformity of the energy dispersion in the clarifier. This would provide an opportunity to investigate the response of intelligent control system via the contour of sludge blanket.

8.5 Conclusion

In the future, the water supply industry will be faced by increasingly more serious regulatory and environmental pressures to provide technically efficient and yet economical treatment. In response to these changes, new sensor technologies are continually being developed, evaluated and implemented in the industry. Manual control is expensive and out of date in practice. Manual control limits the performance of the online sensors and often results in human errors in the operation. Controllers need improvement in order to keep up with this modern sensor technology.

Intelligent technology has been applied sporadically in water treatment application over the last decade. The basic idea of the intelligent approach is to mimic how human operators solve the problem. The intelligent control system used in this study is the first one that fully mimics how a human operator controls the clarifier. All chemical and physical control actions are optimised to minimise the operational cost and the clarified water turbidity (and colour). The artificial neural network process model mimics how the operators predict clarified water qualities. Operator experience was imitated by the multi-objective genetic algorithm optimiser to optimise the control actions with respect to multiple objectives, which is similar to the way that human operators did their task by considering both the operational cost and the clarified water quality. Not only was simulation done by desktop technology but a pilot plant test was conducted for one of the case studies, BKWTP. The pilot plant results presented herein show that the intelligent control system can successfully control the clarifier in a real life situation. The turbidity removal performance satisfied the operational target at a cheaper operational cost compared with human operators.

In conclusion, with the successful results of the pilot plant test, both the operational targets of high water quality and cheaper operational cost are achieved. Intelligent control system is proved able to work in real practice. The existence of the intelligent control system in the water supply industry opens the door for an intelligent control system which frees the clarifier operation from human error and enables full utilisation of technology. This may overcome the limit of the manual skill of human operators. An intelligent control system that usefully employs the benefit of sensors and uses them to maximum capacity seems able to be used in place of the human operators. The skills of human operators could be usefully employed to supervise the intelligent control system and in system quality assurance. Intelligent technologies in the water supply industry can also be used to provide a strong platform for more sophisticated applications of plant-wide control and even applied to alternative applications such as wastewater treatment.

REFERENCES

ASCE. & AWWA. [1990], *Water treatment plant design*, 2 edn, McGraw-Hill, New York.

AWWA [1991], *Water Quality and Treatment : A Handbook of Community Water Supplies*, 5th edn, McGraw-Hill, New York

Baughman, D. R. & Liu, Y. A. [1995], *Neural networks in bioprocessing and chemical engineering*, Academic Press, San Diego.

Baxter, C. W., Zhang, Q., Stanley, S. J., Shariff, R., Tupas, R. R. T. & Stark, H. L. [2001], 'Drinking water quality and treatment: The use of artificial neural networks', *Canadian Journal of Civil Engineering*, vol. 28, no. 1, pp. 26-35.

Bhat, N. & McAvoy, T. J. [1990], 'Use of neural nets for dynamic modeling and control of chemical process systems', *Computers and Chemical Engineering*, vol. 14, no. 4-5, pp. 573-583.

Bishop, C. M. [1995], *Neural networks for pattern recognition*, Clarendon Press, Oxford.

Bowden, G. J., Maier, H. R. & Dandy, G. C. [2002], 'Optimal division of data for neural network models in water resources applications', *Water Resources Research*, vol. 38, no. 2, pp. 1-11.

Box, G. E. P. & Jenkins, G. M. [1976], *Time series analysis : forecasting and control*, Holden-Day series in time series analysis, Holden-Day, San Francisco
Brion, G. M., Neelakantan, T. R. & Lingireddy, S. [2001], 'Using neural networks to predict peak Cryptosporidium concentrations', *Journal of the American Water Works Association*, vol. 93, no. 1, pp. 99-105.

Burlingame, G. A., Pickel, M. J. & Roman, J. T. [1998], 'Practical applications of turbidity monitoring', *Journal AWWA*, vol. 90, no. 8, pp. 57-69.

Carter, S. [1999], 'The Sustainability Assessment of projects: An Artificial Intelligence approach, with application to Tasmania ', [PhD Thesis] University of Tasmania, School of Engineering.

Caudill, M. & Butler, C. [1990], *Naturally intelligent systems*, MIT Press, Cambridge, Mass ; London.

Chen, S. H., Jakeman, A. J. & Norton, J. P. [2008], 'Artificial Intelligence techniques: An introduction to their use for modelling environmental systems', *Mathematics and Computers in Simulation*, vol. 78, no. 2-3, pp. 379-400.

Cheng, F. Y. & Li, D. [1997], 'Multiobjective optimization design with pareto genetic algorithm', *Journal of Structural Engineering*, vol. 123, no. 9, pp. 1252-1261.

Chowdhury, S., Champagne, P. & McLellan, P. J. [2009], 'Models for predicting disinfection byproduct (DBP) formation in drinking waters: A chronological review', *Science of The Total Environment*, vol. 407, no. 14, pp. 4189-4206.

City Water Technology [2001], *Bryn Estyn water treatment plant process operations manual*, NSW, Australia.

Clesceri, L. S., Greenberg, A. E. & Trussell, R. R. [1989], *Standard methods for the examination of water and wastewater*, 17th edn, American Public Health Association, Washington, DC.

Coello Coello, C. A. & Lamont, G. B. [2004], *Applications of multi-objective evolutionary algorithms*, Advances in natural computation ; Vol. 1, World Scientific, Singapore ; Hackensack, NJ.

Coello Coello, C. A., Van Veldhuizen, D. A. & Lamont, G. B. [2002], *Evolutionary algorithms for solving multi-objective problems*, Genetic algorithms and evolutionary computation ; 5, Kluwer Academic, New York.

Dandy G. C. et al [1991], 'Use of artificial neural networks to forecast salinity in rivers', in Lingireddy, S. & G. M. Brion (eds), *Artificial neural networks in water supply engineering*, ASCE, Reston, Virginia, pp. 107-137.

Dayhoff, J. E. [1990], *Neural network architectures: an introduction* Van Nostrand Reinhold, New York.

Deb, K. [2001], *Multi-objective optimization using evolutionary algorithms*, 1st edn, Wiley-Interscience series in systems and optimization, John Wiley & Sons, Chichester ; New York.

Deb, K. & Goldberg, D. E. [1989], 'An investigation of niche and species formation in genetic function optimization', in *Proceedings of the Third International Conference on Genetic Algorithms*, Morgan Kaufmann Publishers Inc. , San Francisco, pp. 42-50.

Degrémont S. A. [1991], *Water treatment handbook*, 6th [English] edn, Degrémont Rueil-Malmaison, France

DeJong, K. [1975], 'An Analysis of the behaviour of a class of genetic adaptive systems', Ph.D thesis, University of Michigan.

Dellana, S. A. & West, D. [2009], 'Predictive modeling for wastewater applications: Linear and nonlinear approaches', *Environmental Modelling & Software*, vol. 24, no. 1, pp. 96-106.

Demuth, H. & Beale, M. [2000], *Neural Network Toolbox User's Guide (version 4)*, The Math Works, Inc., Natick, Massachusetts.

DeSilets, L., Golden, B., Wang, Q. & Kumar, R. [1992], 'Predicting salinity in the Chesapeake Bay using backpropagation', *Computers & Operations Research*, vol. 19, no. 3-4, pp. 277-285.

Economou, C. G., Morari, M. & Palsson, B. O. [1986], 'Internal model control: Extension to nonlinear systems', *Industrial & Engineering Chemistry, Process Design and Development*, vol. 25, no. 2, pp. 403-411.

Eshelman, L. J. & Schaffer, J. D. [1991], 'Preventing premature convergence in genetic algorithm by preventing incest', in *Proceedings of the Fourth International Conference on Genetic Algorithms,* , Morgan Kaufmann, San Diego, CA, pp. 115-122.

Filho, S. A., Silva, F., Lima, F., Lavrador, M., Maciel, B. & Jr, G. L. [2006], 'Using the lag of autocorrelation function in order to identify the anaerobic threshold during dynamic physical exercise', in *Computers in Cardiology*, Valencia, Spain, vol. 33, pp. 625-628.

Fleming, P. J. & Pashkevich, M. A. [2007], 'Optimal advertising campaign generation for multiple brands using MOGA', *IEEE Transactions on Systems Man and Cybernetics Part C*, vol. 37, no. 6, pp. 190-1201.

Flood, I. & Kartam, N. [1994], 'Neural networks in civil engineering. I: Principles and understanding', *Journal of Computing in Civil Engineering*, vol. 8, no. 2, pp. 131-148.

Fonseca, C. M. & Fleming, P. J. [1993], 'Genetic algorithms for multiobjective optimization: formulation, discussion, and generalization', in Forrest, S. (ed.), *Proceedings of the Fifth International Conference on Genetic Algorithms*, Morgan Kaufmann, Morgan Kaufmann, San Francisco, CA, pp. 416-423.

Freeman, J. A. & Skapura, D. M. [1991], *Neural network : algorithms, applications, and programming techniques*, Computation and neural systems series, Addison-Wesley, Reading, Massachusetts.

Fwa, T. F., Chan, W. T. & Hoque, K. Z. [2000], 'Multiobjective optimization for pavement maintenance programming', *Journal of Transportation Engineering*, vol. 126, no. 5, p. 367.

Garret, J. H., Ghaboussi, J. & Wu, X. (eds) [1992], *Neural Networks*, Expert Systems for Civil Engineers: Knowledge Representation, ASCE, New York.

Goldberg, D. E. [1989], *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley Pub. Co., Reading, Mass.

Goldberg, D. E. & Deb, K. [1991], 'A comparison of selection schemes used in genetic algorithms', in Rawlins, G. J. E. (ed.), *Foundations of genetic algorithms (FOGA-1)*, Morgan Kaufmann, pp. 69-93.

Goldberg, D. E. & Richardson, J. [1987], 'Genetic algorithm with sharing for multimodal function optimization', in *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application, Cambridge, Massachusetts*, L. Erlbaum Associates Inc., Hillsdale, NJ, , pp. 41-49.

Grefenstette, J. J. [1986], 'Optimisation of control parameters for genetic algorithms', *IEEE Transaction on Systems, Man and Cybernetic*, vol. 16, no. 1, pp. 122 -128

Hagan, M. T., Demuth, H. B. & Beale, M. H. [1996], *Neural network design*, PWS, Boston

Halhal, D., Walters, G. A., Ouazar, D. & Savic, D. A. [1997], 'Water network rehabilitation with structured messy genetic algorithm', *Journal of Water Resources Planning and Management*, vol. 123, no. 3, pp. 137-145.

Harnett, D. L. & Murphy, J. L. [1975], *Introductory statistical analysis*, Addison-Wesly, Reading, M.A.

Harris, C. J. [1994], *Advances in intelligent control*, Taylor & Francis, London ; Bristol, PA.

Hassibi, K. M. & Loparo, K. A. [1991], 'Model reference neurocontrollers based on feedback linearization', *IEEE International Joint Conference on Neural Networks*, vol. 2, no. 18-20, pp. 1813-1818.

Hassoun, M. H. [1995], *Fundamentals of artificial neural networks*, MIT Press, Cambridge.

Haupt, R. L., Haupt, S. E. & Haupt, R. L. [2004], *Practical genetic algorithms*, 2nd edn, John Wiley, Hoboken, N.J.

Haykin, S. [1999], *Neural Networks: A comprehensive Foundation*, 2nd edn, Prentice-Hall, New Jersey.

Hecht-Nielsen, R. [1989], 'Theory of the backpropagation neural network', *IJCNN International Joint Conference on Neural Networks, 18-22 June 1989, Washington, DC, USA, vol 1, pp. 593-605.*

Henson, M. A. [1998], 'Nonlinear model predictive control: Current status and future directions', *Computers and Chemical Engineering*, vol. 23, no. 2, pp. 187-202.

Hertz, J., Palmer, R. G. & Krogh, A. S. [1991], *Introduction to the theory of neural computation*, Santa Fe Institute studies in the sciences of complexity. Lecture notes ; v. 1, Addison-Wesley Pub. Co., Redwood City, California.

Holland, J. H. [1992], *Adaptation in natural and artificial systems : an introductory analysis with applications to biology, control, and artificial intelligence*, 1st MIT Press edn, Complex adaptive systems, MIT Press, Cambridge, Mass.

Horn, J., Nafpliotis, N. & Goldberg, D. E. [1994], 'Niched Pareto genetic algorithm for multiobjective optimization', in *Proceedings of the 1st IEEE Conference on Evolutionary Computation*, IEEE, Orlando, FL, USA, vol. 1, pp. 82-87.

Hudson, H. E. [1981], *Water clarification processes: practical design and evaluation*, Van Nostrand Reinhold environmental engineering series, Van Nostrand Reinhold Co., New York.

Hunt, K. J. & Sbarbaro, D. [1991], 'Neural networks for nonlinear internal model control', *IEE Proceedings : Control Theory and Applications*, vol. 138, no. 5, pp. 431-438.

Hunt, K. J. & Sbarbaro, D. [1995], 'Studies in artificial neural network based control', in London, UK, pp. 109-135.

Hunt, K. J., Sbarbaro, D., Zbikowski, R. & Gawthrop, P. J. [1992], 'Neural networks for control systems - A survey', *Automatica*, vol. 28, no. 6, pp. 1083-1112.

Ikonen, E. & Najim, K. [2002], *Advanced process identification and control*, Marcel Dekker, Inc., New York.

Irwin, G. W., Warwick, K., Hunt, K. J. & Institution of Electrical Engineers. [1995], *Neural network applications in control*, Institution of Electrical Engineers, London, U.K.

Jain, L. C. & Martin., N. M. [1999], *Fusion of neural networks, fuzzy sets, and genetic algorithm : industrial applications*, CRC press, Boca Raton, Florida.

James M. Montgomery Consulting Engineers Inc. [1985], *Water treatment principles and design*, Wiley, New York.

Kawamura, S. [1991], *Integrated design of water treatment facilities*, Wiley, New York.

Kazancioglu, E., Guangquan, W., Jeonghan, K. & B., S. [2003], 'Robust optimisation of an automobile valvetrain using a multiobjective genetic algorithm', paper presented to Proceedings of DETC'03, ASME, Chicago, Illinois, 2-6 Sept. 2003.

Kerri, K. D. [1996], *Water treatment plant operation : a field study training program*, 3rd edn, Foundation of California State University Sacramento, Sacramento.

Kilmasaukas, C. C. [1988], *NeuralWorks: An introduction to neural computing*, NeuralWare Inc, Sewickley, PA.

Kocjancic, R. & Zupan, J. [2000], 'Modelling of the river flow rate: the influence of the training set selection', *Chemometrics and Intelligent Laboratory Systems*, vol. 54, pp. 21–34.

Kohonen, T. [1982], 'Self-organized formation of topologically correct feature maps', *Biological Cybernetics*, vol. 43, pp. 59-69.

Kohonen, T. [1997], *Self-organising maps*, Springer-Verlag, Berlin, Heildelberg.

Kosko, B. [1992], *Neural networks and fuzzy systems : a dynamical systems approach to machine intelligence*, Prentice-Hall, Englewood Cliffs.

Kreyszig, E., Norminton, E. J. & Kreyszig, H. [2006], *Advanced engineering mathematics*, 9th edn, Wiley, New York.

Lachtermacher, G. & Fuller, J. D. [1994], 'Back-propagation in hydrological time series forecasting', in Hipel, A. H., et al. (eds), *Stochastic and statistic methods in hydrology and environmental engineering*, Kluwer Academic, Dordrecht.

Liu, J. C. & Wu, M. D. [1997], 'Fuzzy control of coagulation reaction through streaming current monitoring', *Water Science and Technology*, vol. 36, no. 4, pp. 127-134.

Maier, H. R. & Dandy, G. C. [1998], 'The effect of internal parameters and geometry on the performance of back-propagation neural networks : an empirical study', *Environmental Modelling and Software*, vol. 13, no. 2, pp. 193-209.

Maier, H. R. & Dandy, G. C. [2000], 'Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and...' *Environmental Modelling and Software with Environment Data News*, vol. 15, no. 1, pp. 101-124.

Maier, H. R., Morgan, N. & Chow, C. W. K. [2004], 'Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters', *Environmental Modelling & Software*, vol. 19, pp. 485–494.

Maren, A., Harston, C. & Pap, R. [1990], *Handbook of neural computing application*, Academic Press, San Diego, California.

Masschelein, W. J. [1992], *Unit processes in drinking water treatment*, Marcel Dekker, INC., New York.

Masters, T. [1993], *Practical neural network recipe in C++*, Academic Press, San Diego, California.

May, D. B. & Sivakumar, M. [2009], 'Prediction of urban stormwater quality using artificial neural networks', *Environmental Modelling & Software*, vol. 24, no. 2, pp. 296-302.

McCulloch, W. S. & Pitts, W. [1943], 'A logical Calculus of the ideas immanent in nervous activity', *Bulletin of Mathematical Biophysics* vol. 5, pp. 115-137.

McKinney, D. C. & Min-Der, L. [1994], 'Genetic algorithm solution of groundwater management models', *Water Resources Research*, vol. 30, no. 6, pp. 1897-1906.

Mendenhall, W. & Beaver, R. J. [1994], *Inroduction to probability and statistics*, Wadsworth, Belmont, CA.

Mills, P. M., Zomaya, A. Y. & Tadé, M. O. [1996], *Neuro-adaptive process control : a practical approach*, Wiley, Chichester ; New York.

Mirsepassi, A., Cathers, B. & Dharmappa, H. B. [1995], 'Application of artificial neural networks to the real time operation of water treatment plants', *Proceedings of the 1995 IEEE International Conference on Neural Networks*, vol. 1, pp. 516-521.

Moâsciânski, J. & Ogonowski, Z. [1995], *Advanced control with MATLAB and SIMULINK*, Ellis Horwood, London.

Murray-Smith, R. & Johansen, T. A. [1997], *Multiple model approaches to modelling and control*, Taylor & Francis, London, Bristol, PA.

MWA [2000], *Bang Khen water treatment plant ISO 9002 manual: Operational manual*, Bangkok, Thailand.

MWA [2006], *Metropolitan Waterworks authority (MWA) Annual Report 2006*, MWA, Bangkok, Thailand.

Nahas, E. P., Henson, M. A. & Seboro, D. E. [1992], 'Nonlinear internal model control strategy for neural network models', *Computers and Chemical Engineering*, vol. 16, no. 12, pp. 1039-1057.

Narendra, K. S. & Parthasarathy, K. [1990], 'Identification and control of dynamical systems using neural networks', *IEEE Transactions on Neural Network*, vol. 1, no. 1, pp. 4-27.

Negnevitsky, M. [2005], *Artificial Intelligence: A Guide to Intelligent Systems*, 2nd edn, Addison-Wesley, Harlow, England.

Norgaard, M., Ravn, O., Poulsen, N. K. & Hansen, L. K. [2000], *Neural networks for modelling and control of dynamic systems*, Springer-Verlag, London.

Obayashi, S., Takahashi, S. & Takeguchi, Y. [1998], 'Niching and elitist model for MOGAs', in *Parallel Problem Solving from Nature*, Springer Berlin New York.

Perendeci, A., Arslan, S., ɔelebi, S. S. & Tanyolă, A. [2008], 'Prediction of effluent quality of an anaerobic treatment plant under unsteady state through ANFIS modeling with on-line input variables', *Chemical Engineering Journal*, vol. 145, no. 1, pp. 78-85.

Pham, D. T. & Liu, X. [1995], *Neural networks for identification, prediction, and control*, Springer-Verlag, London ; New York.

Pizzi, N. G. [2005], *Water treatment operator handbook*, American Water Works Association, Denver, Colorado.

Pohlheim, H. [2005], *GEATbx version 3.7, Introduction to Evolutionary Algorithm, Overview, Method and Operators*, http://www.geatbx.com.

Pollack, A. J., Chen, A. S. C., Roy, C. H. & Goodrich, J. A. [1999], *Option for remote monitoring and control of small drinking water facilities*, Battelle Press, Columbus, Ohio.

Psaltis, D., Sideris, A. & Yamamura, A. A. [1988], 'Multilayered neural network controller', *IEEE control systems magazine*, vol. 8, no. 2, pp. 17-21.

Qi, M. & Zhang, G. P. [2001], 'An investigation of model selection criteria for neural network time series forecasting', *European Journal of Operational Research*, vol. 132, pp. 666-680.

Rawlings, J. B., Meadows, E. S. & Muske, K. R. [1994], 'Nonlinear model predictive control: a tutorial and survey', *Proc. Advanced control of chemical process*, pp. 185-197.

Reed, P., Minsker, B. S. & Goldberg, D. E. [2003], 'Simplifying multiobjective optimization: An Automated Design methodology for the nondomonated sorted genetic algorithm-II', *Water Resources Research* vol. 39, no. 7, pp. TNN 2-1-5.

Ripley, B. D. [1996], *Pattern recognition and neural networks*, Cambridge University Press, Cambridge, New York.

Riyaz, S., Audrey, C., Qing, Z. & Stephen, J. S. [2004], 'Advanced process control techniques for water treatment using artificial neural networks', *Journal of Environmental Engineering and Science*, vol. 3, p. S61.

Rodriguez MJ, West JR & Powell J [1997], 'Application of two approaches to model chlorine residuals in Severn Trent Water Ltd (STW) distribution systems ', *Water Science and Technology*, vol. 36, no. 5, pp. 317-324.

Rogers, L. L. & Dowla, F. U. [1994], 'Optimization of groundwater remediation using neural networks with parallel solute transport modeling', *Water Resources Research*, vol. 30, no. 2, pp. 457-481.

Romero, C. [1991], *Handbook of critical issues in goal programming*, Pergamon, Oxford.

Rumelhart, D. E., McClelland, J. L. & Asanuma, C. [1986], *Parallel distributed processing : explorations in the microstructure of cognition*, 2 vols., Computational models of cognition and perception, MIT Press, Cambridge, Massachusetts.

Saad, E. W., Caudell, T. P. & Wunsch Ii, D. C. [1999], 'Predictive head tracking for virtual reality', *IJCNN'99. Proceedings of the International Joint Conference on Neural Networks, Washington, DC, Vol. 6, pp 3933-3936.*

Sha, W. [2007], 'Comment on "Flow forecasting for a Hawaii stream using rating curves and neural networks" by G. B. Sahoo and C. Ray [Journal of Hydrology 317 (2006) 63-80]', *Journal of Hydrology*, vol. 340, pp. 119-121.

Shanks, R. L. [1978], *Water treatment plant design for practicing engineer*, Ann Arbor Science Publishers, Ann Arbor, Michigan.

Smith, M. [1993], *Neural Networks for statistical modelling*, Van Nostrand Reinhold, New York.

Srinivas, N. & Deb, K. [1994], 'Multi-objective function optimization using non-dominated sorting genetic algorithm', *Evolutionary Computation Journal* vol. 2, no. 3, pp. 221-248.

Steuer, R. E. [1986], *Multiple criteria optimization: theory, computation, and application*, Wiley series in probability and mathematical statistics. Applied probability and statistics, Wiley, New York.

Stone, M. [1974], 'Cross-validatory choice and assessment of statistical predictions', *Journal of the Royal Statistic Society series B*, vol. 36, pp. 111-147.

Sun, F., Chen, J., Tong, Q. & Zeng, S. [2009], 'Development and identification of an integrated waterworks model for trihalomethanes simulation', *Science of The Total Environment*, vol. 407, no. 6, pp. 2077-2086.

Swingler, K. [1996], *Applying neural networks : a practical guide*, Academic Press, London.

Tambo, N. & Watanabe, Y. [1979], 'Physical aspect of flocculation process - I. Fundamental treatise', *Water Research*, vol. 13, no. 5, pp. 429-339.

Tarassenko, L. [1998], *A Guide to Neural Network Computing Applications*, John Wiley & Sons, New York.

Tokar, A. S. & Johnson, P. A. [1999], 'Rainfall-runoff modeling using Artificial Neural Networks', *Journal of Hydrologic Engineering*, vol. 4, no. 3, pp. 232-239.

Tsoukalas, L. H. & Uhrig, R. E. [1997], *Fuzzy and neural approaches in engineering*, Adaptive and learning systems for signal processing, communications, and control, Wiley, New York

USEPA [1998], *Optimizing water treatment plant performance using the composite correction program*, Office of Ground Water and Drinking Water, Environmental Protection Agency, Cincinnati, OH.

USEPA [1999], *Guidance Manual for Compliance with the Interim Enhanced Surface Water Treatment Rule: Turbidity Provisions*, US Environmental Protection Agency, Washington.

Venema, R. S., Ypma, A., Nijhuis, J. A. G. & Spaanenburg, L. [1996], 'Neural control of artificial human walking: a structured approach', paper presented to Proceeding of Groningen International Student Conference on Information Technology, Groningen, Netherlands, February 1996.

Weigend, A. S., Huberman, B. A. & Rumelhart, D. E. [1990], 'Predicting the future: a connectionist approach', *International Journal of Neural Systems*, vol. 1, no. 3, pp. 193-209.

Widrow, B. & Hoff, M. E. [1960], 'Adaptive switching circuits', in *IRE WESCON Convention Record*, New York, pp. 96-104.

Zalzala, A. M. S. & Fleming, P. J. [1997], *Genetic algorithms in engineering systems*, IEE control engineering series ; 55, Institution of Electrical Engineers, London.

Zhang, G., Hu, M. Y., Patuwo, B. E. & Indro, D. C. [1999], 'Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis', *European Journal of Operational Research*, vol. 116, pp. 16-32.

Zhang, G., Patuwo, B. E. & Hu, M. Y. [1998], 'Forecasting with artificial neural networks: The state of the art', *International Journal of Forecasting*, vol. 14, pp. 35–62.

Zhang, G., Patuwo, B. E. & Hu, M. Y. [2001], 'A simulation study of artificial neural networks for nonlinear time-series forecasting', *Computers & Operations Research*, vol. 28, pp. 381-396.

Zhang, Q. & Stanley, S. J. [1997], 'Forecasting raw-water quality parameters for the north Saskatchewan River by neural network modeling', *Water Research*, vol. 31, no. 9, pp. 2340-2350.

Zhang, Q. & Stanley, S. J. [1999], 'Real-time water treatment process control with artificial neural networks', *Journal of Environmental Engineering*, vol. 125, no. 2, pp. 153-160.

Zhu, X. & Simpson, A. R. [1991], *An expert system for turbidity control advice in water treatment plant operations*, Department of Civil Engineering, University of Adelaide.

Zitzler, E. & Thiele, L. [1999], 'Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach', *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257-271.

Zurada, J. M. [1992], *Introduction to artificial neural systems*, West, St Paul.

APPENDIX A: BRYN ESTYN WATER TREATMENT OPERATIONAL DATA

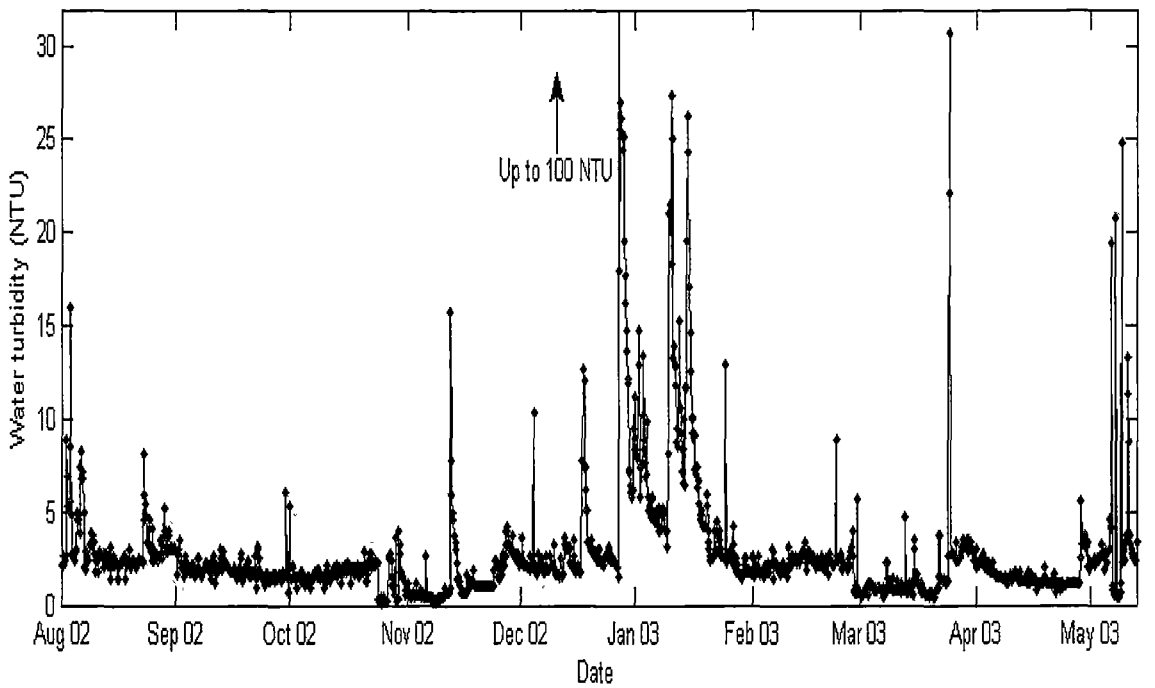


Figure A1 Raw water turbidity (BEWTP)

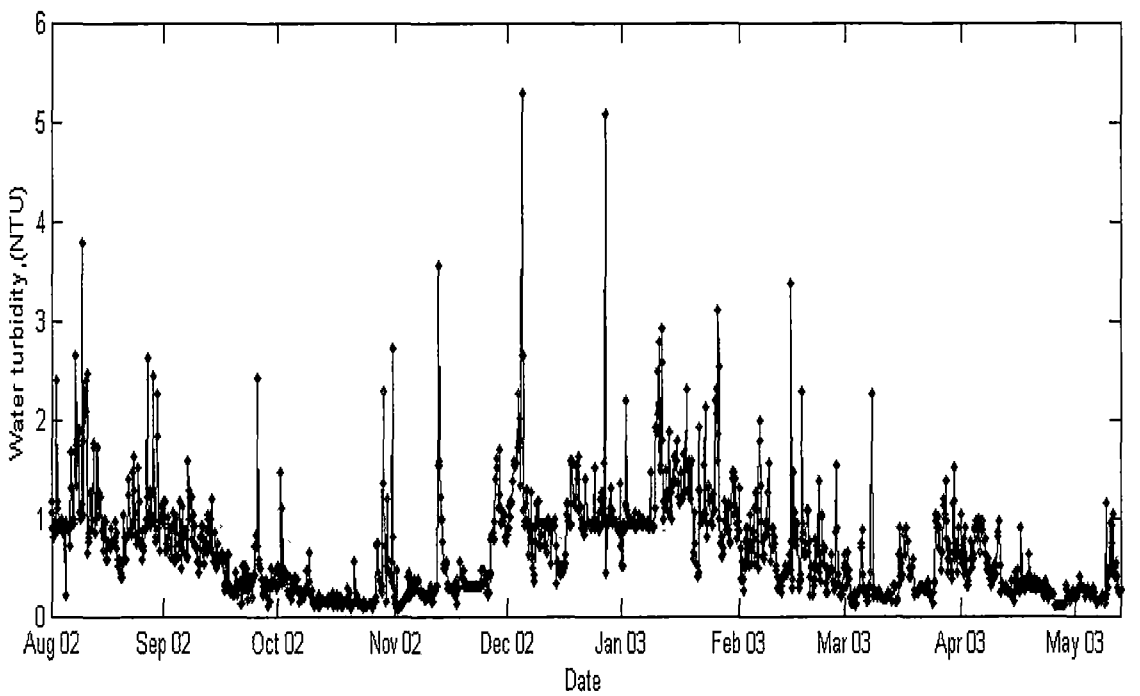


Figure A2 Clarified water turbidity (BEWTP)

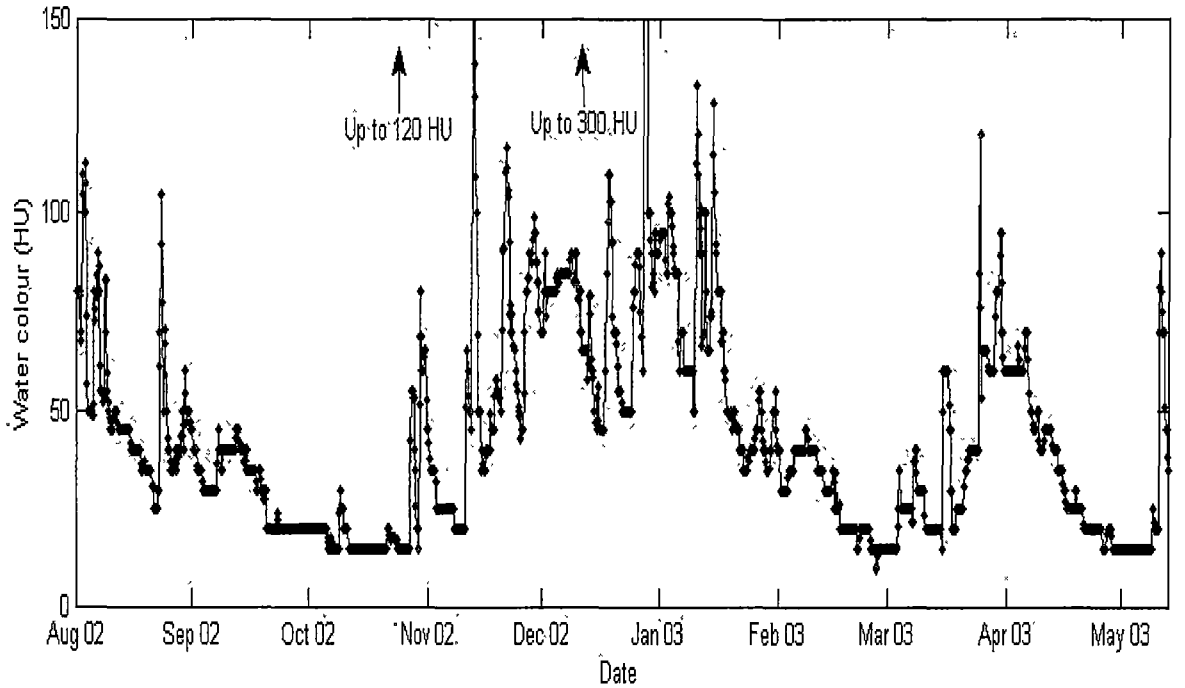


Figure A3 Raw water colour (BEWTP)

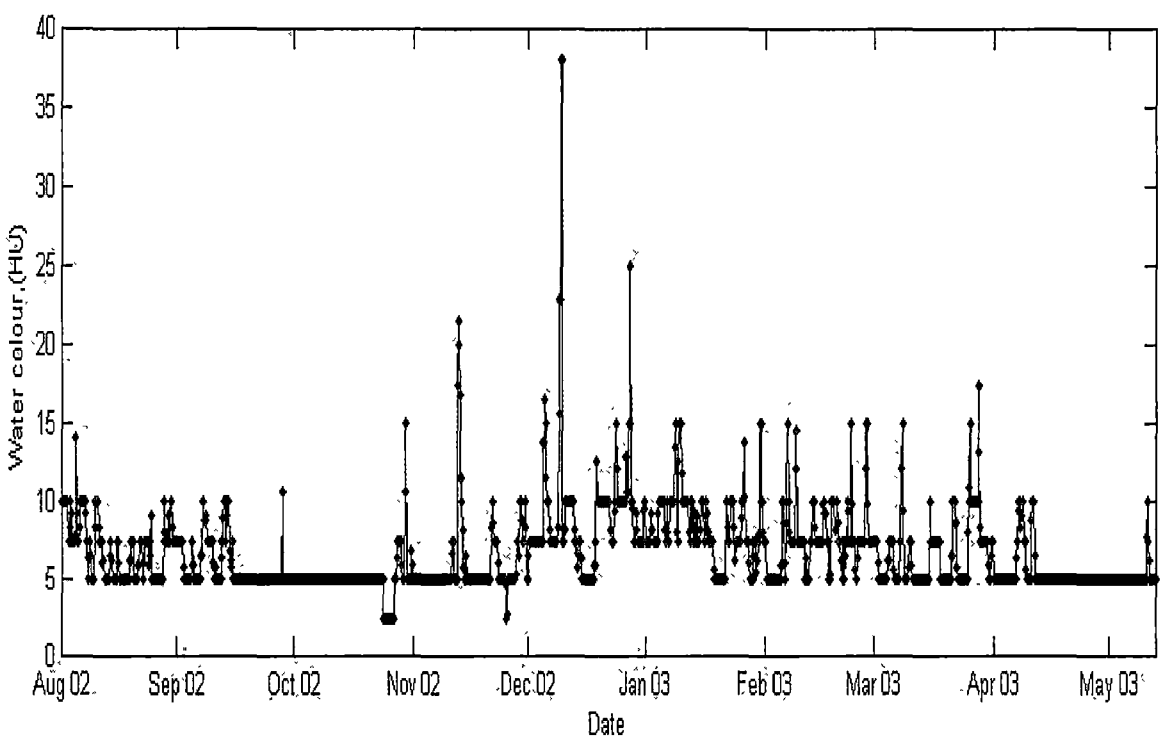


Figure A4 Clarified water colour (BEWTP)

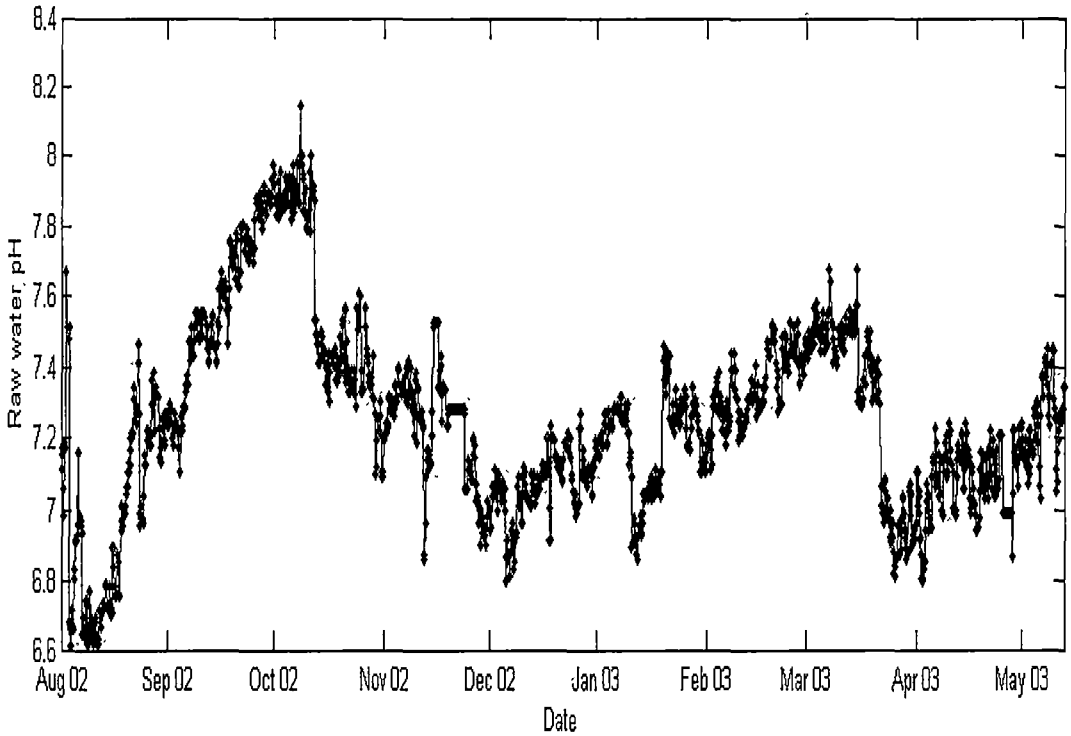


Figure A5 Raw water pH (BEWTP)

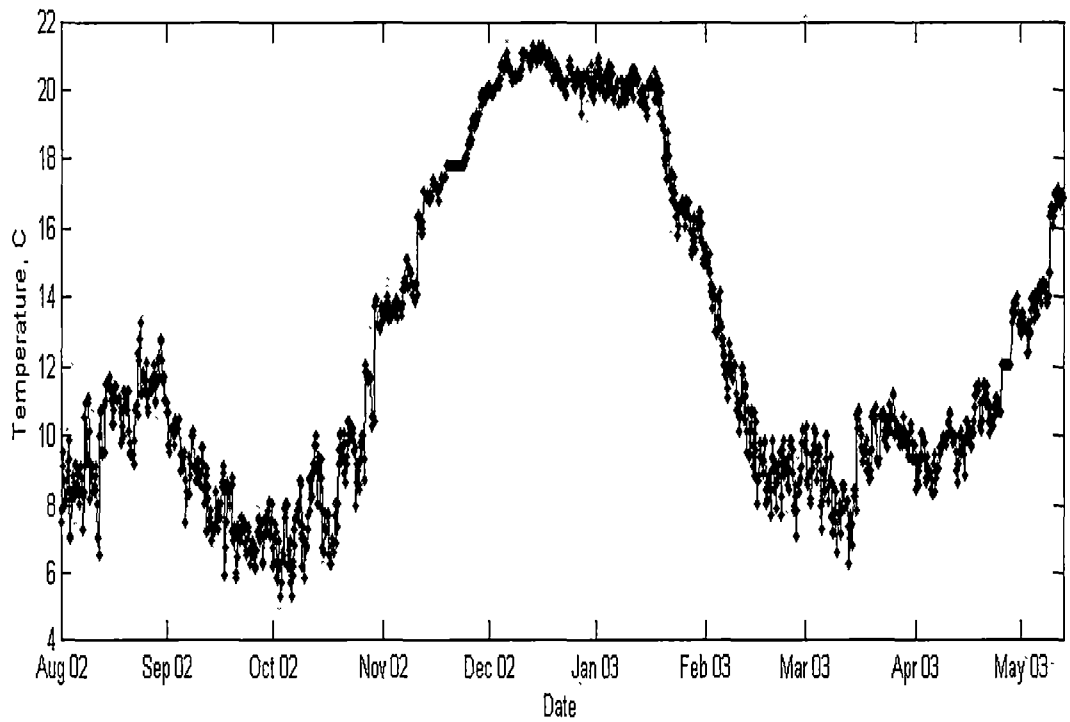


Figure A6 Raw water temperature (BEWTP)

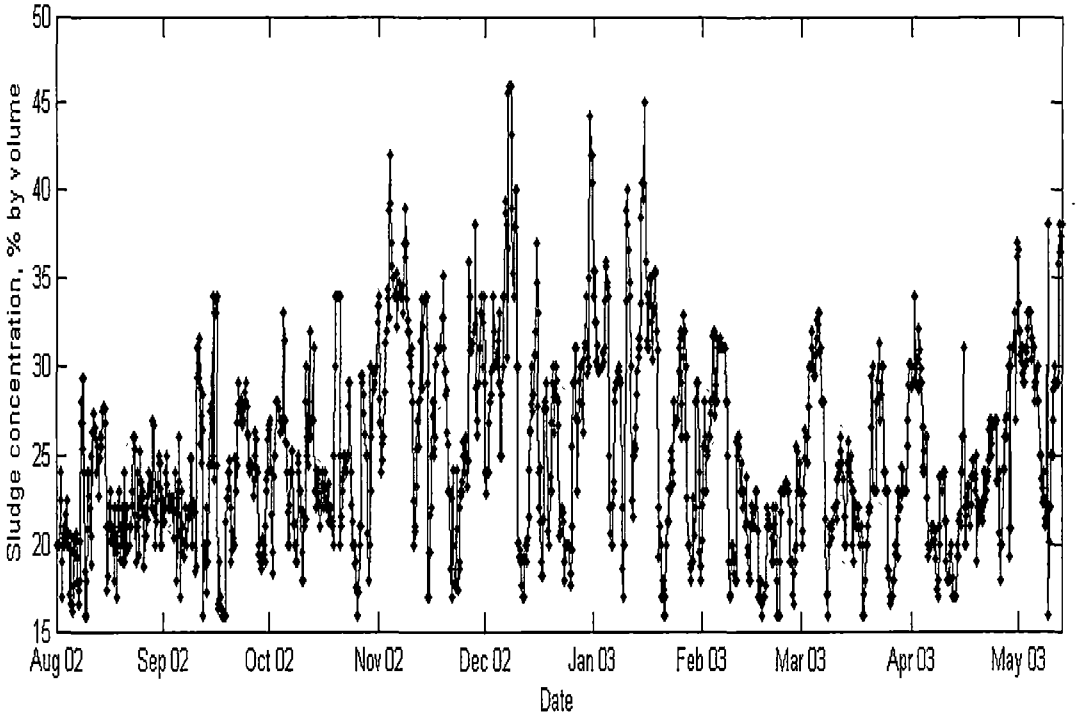


Figure A7 Sludge concentration (BEWTP)

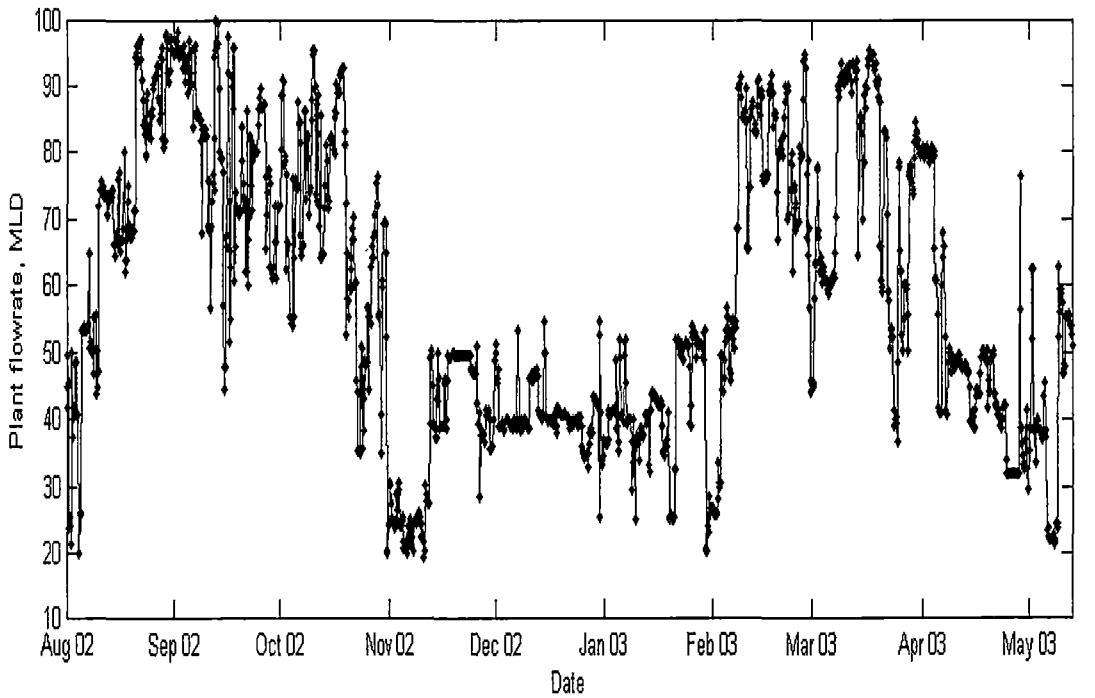


Figure A8 Plant flow rate (BEWTP)

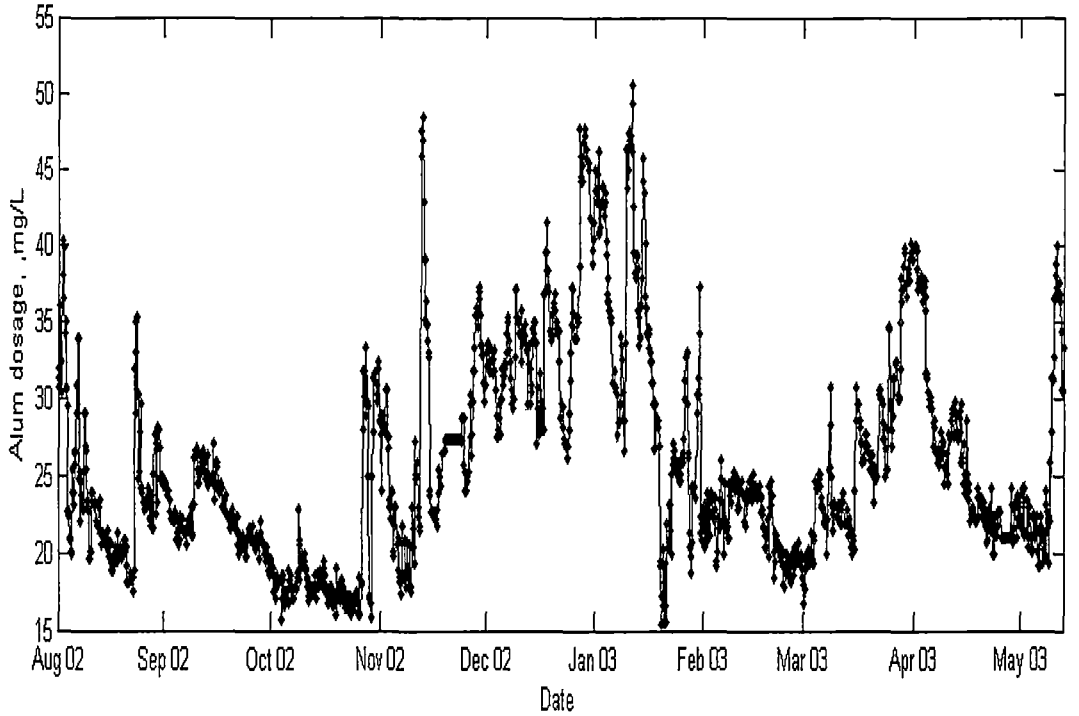


Figure A9 Alum dosage (BEWTP)

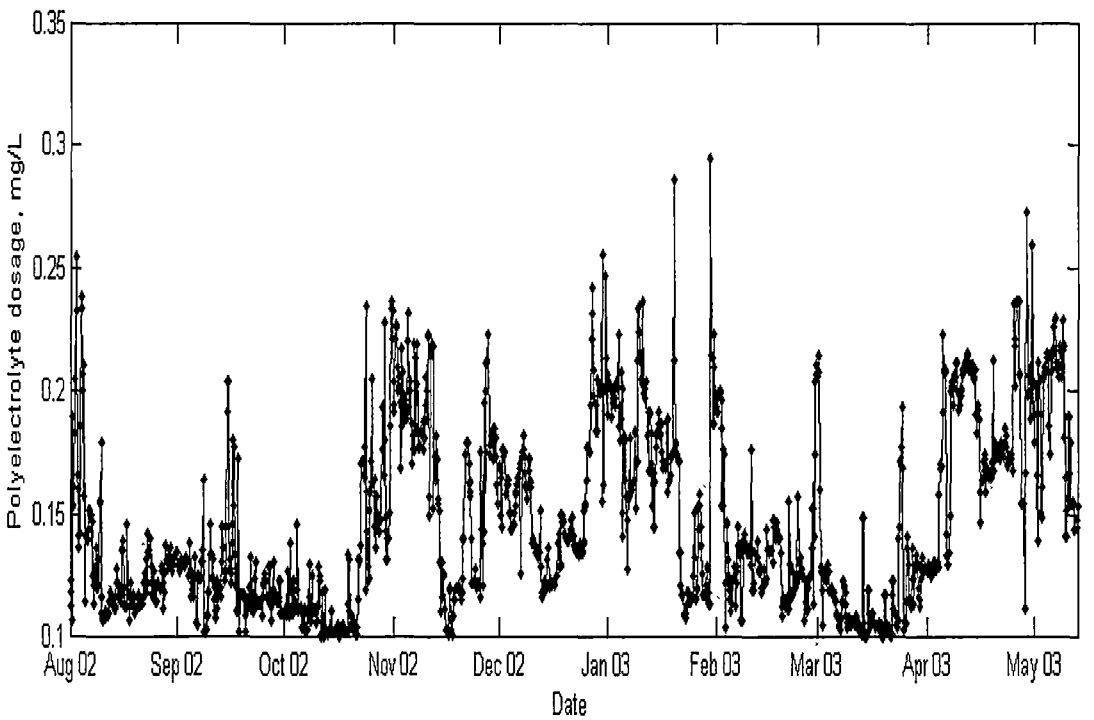


Figure A10 Polyelectrolyte dosage (BEWTP)

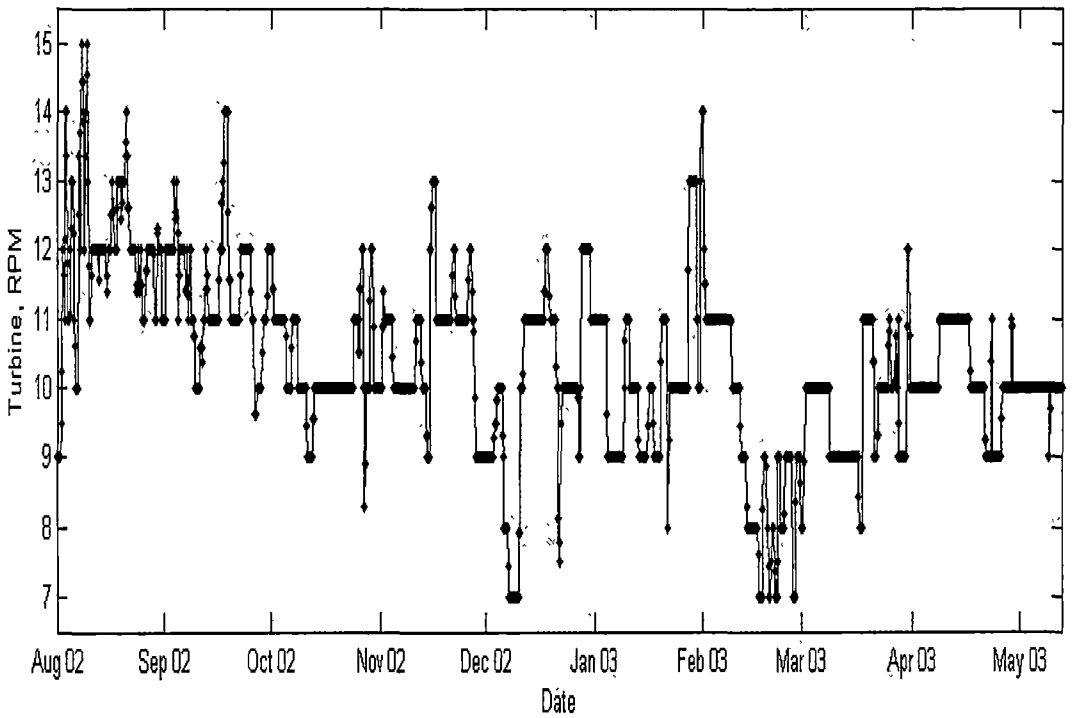


Figure A11 Turbine speed (BEWTP)

APPENDIX B: BANG KHEN WATER TREATMENT PLANT OPERATIONAL DATA

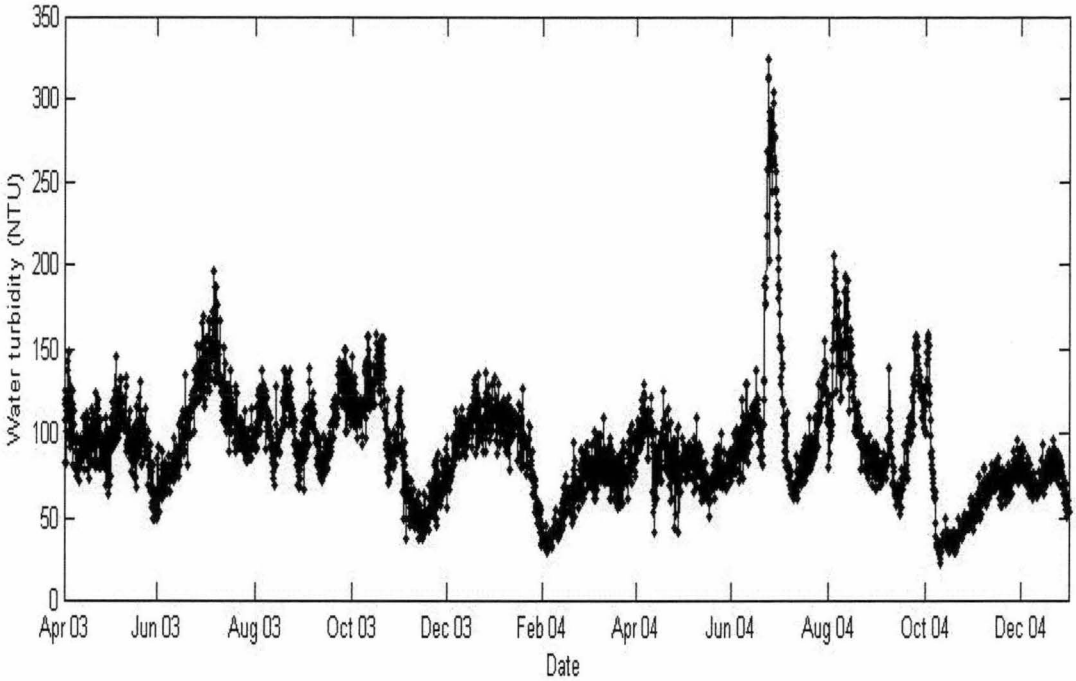


Figure B1 Raw water turbidity (BKWTP)

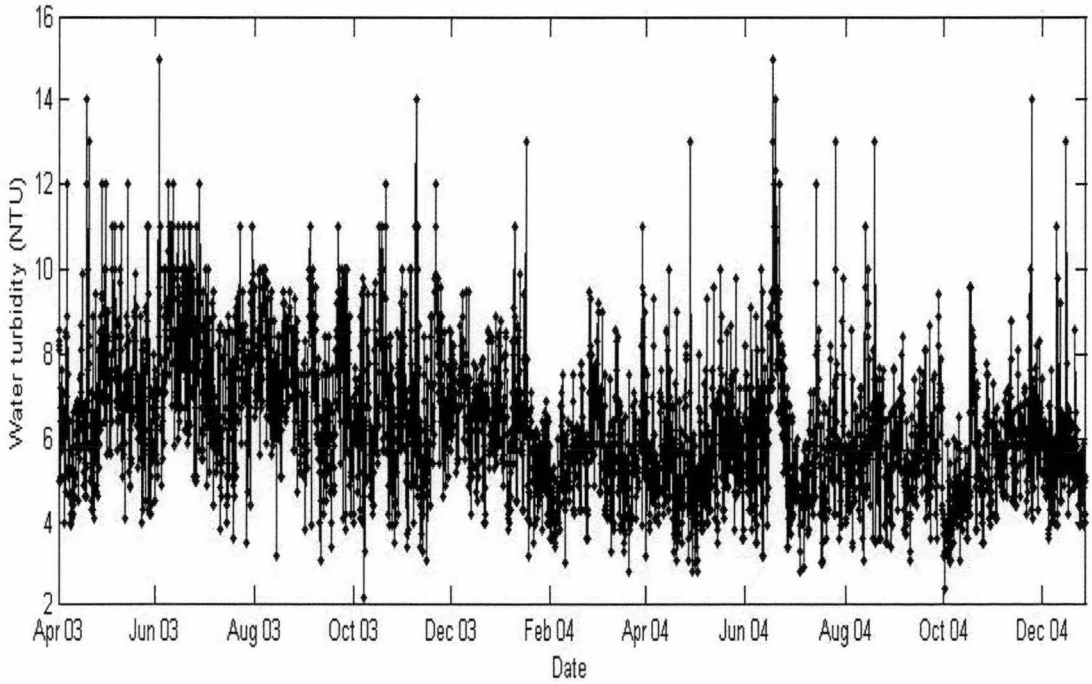


Figure B2 Clarified water turbidity (BKWTP)

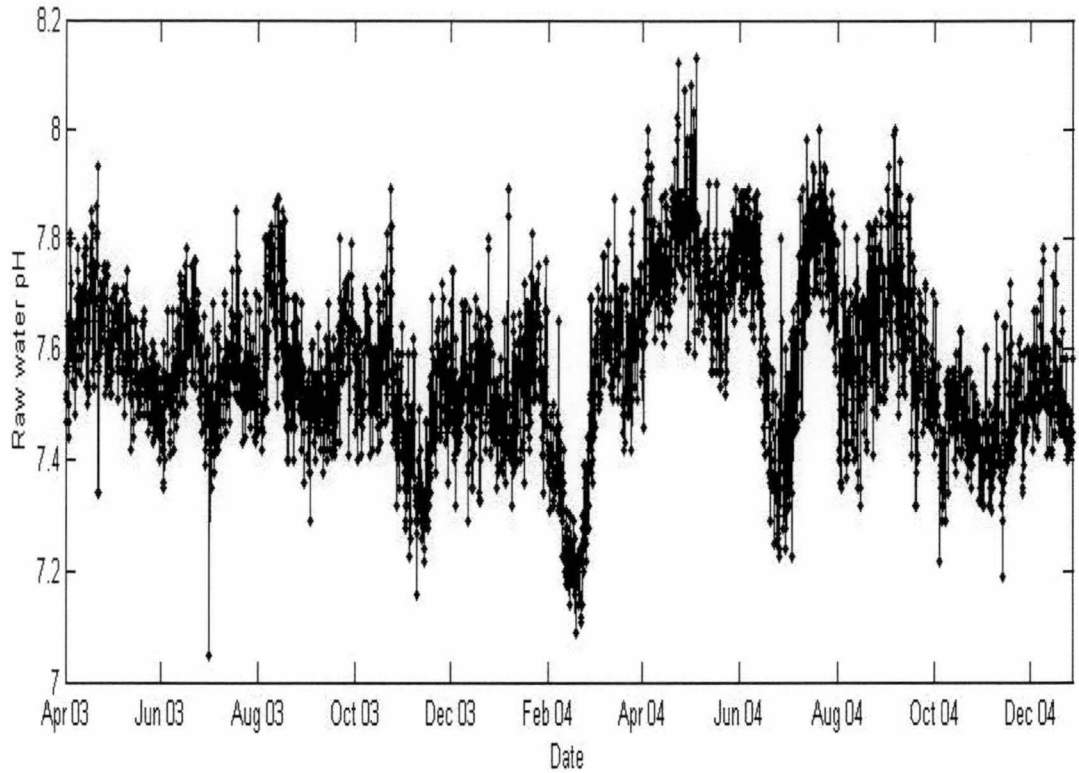


Figure B3 Raw water pH (BKWTP)

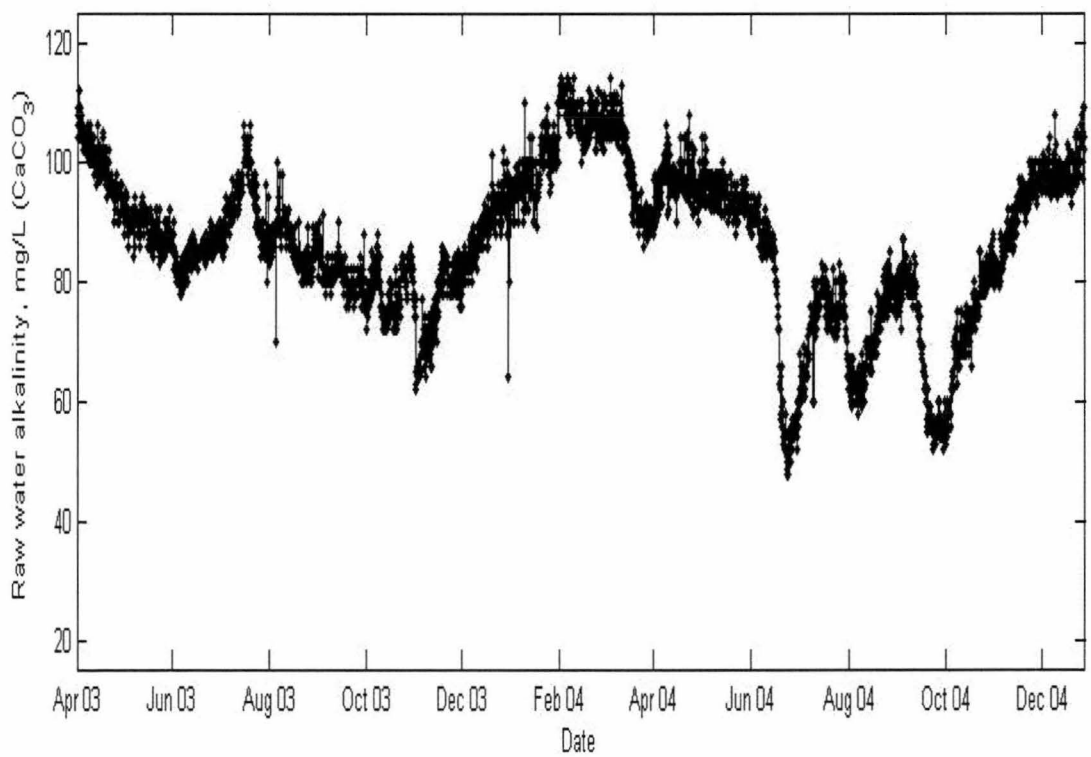


Figure B4 Raw water alkalinity (BKWTP)

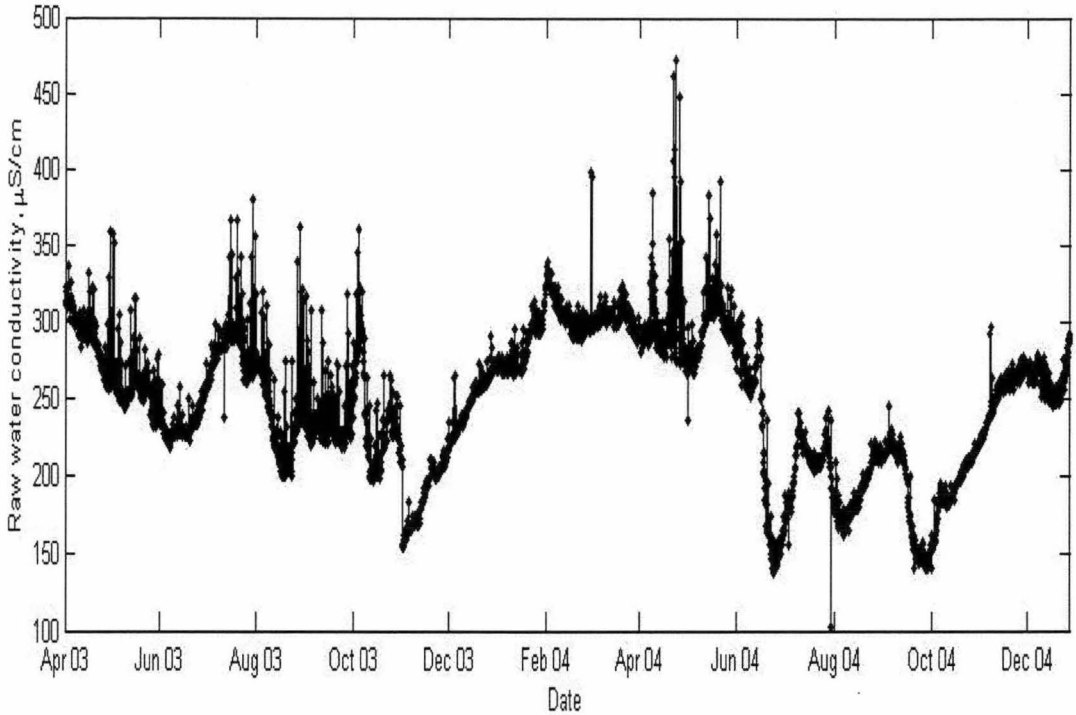


Figure B5 Raw water conductivity (BKWTP)

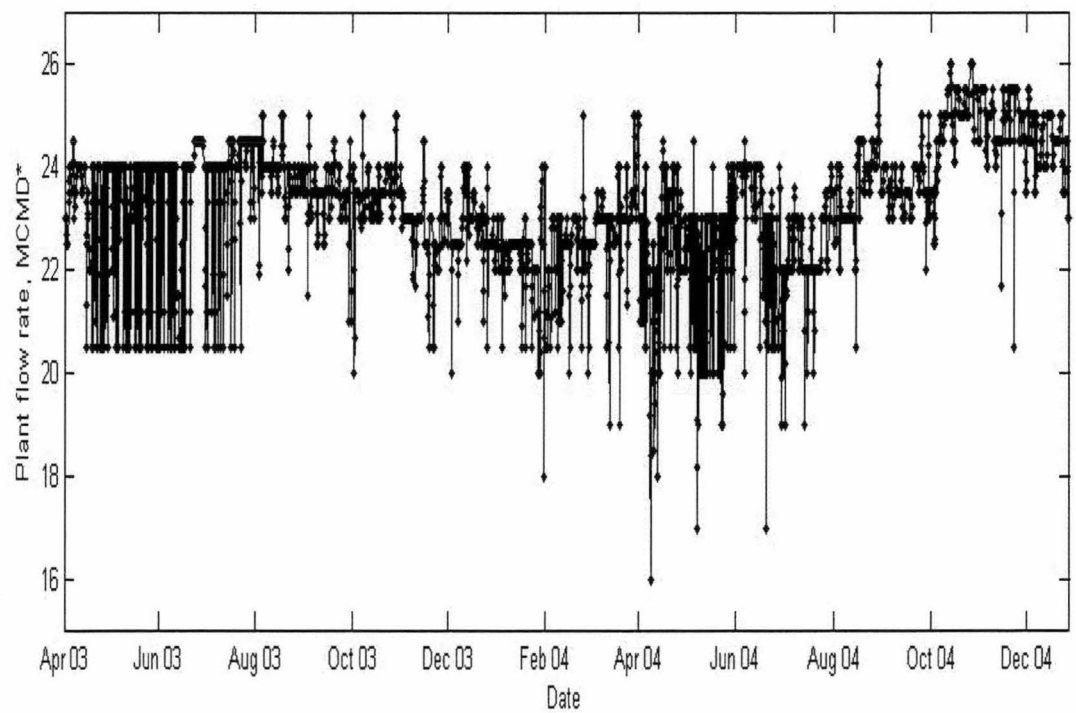


Figure B6 Plant flow rate (BKWTP)

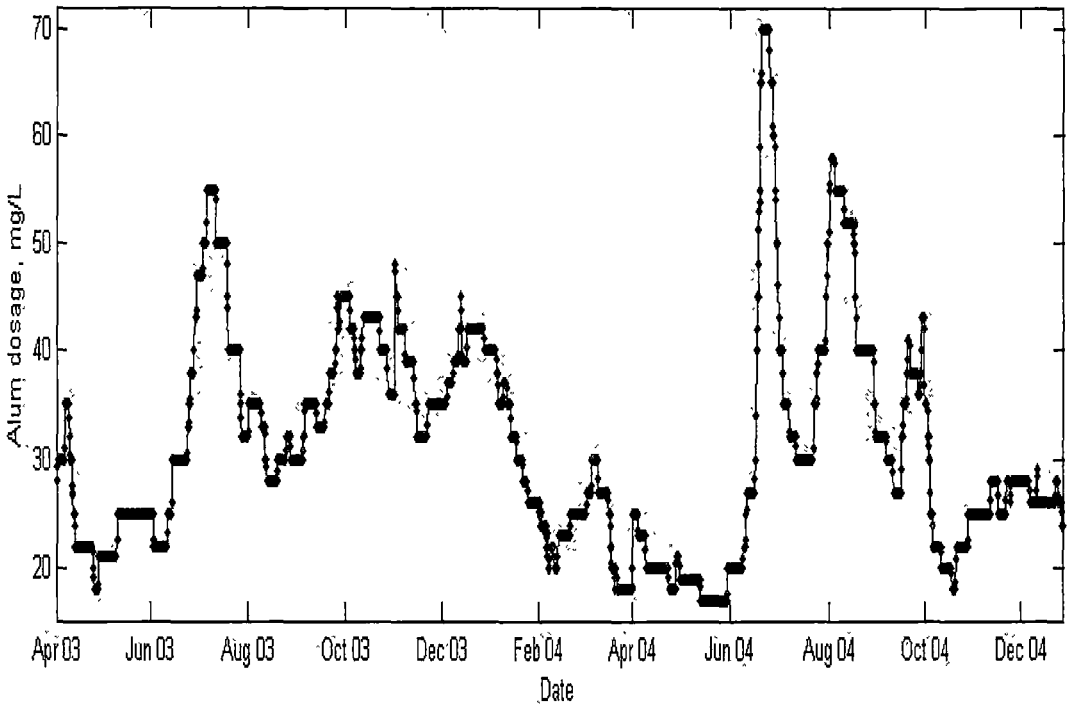


Figure B7 Alum dosage (BKWTP)

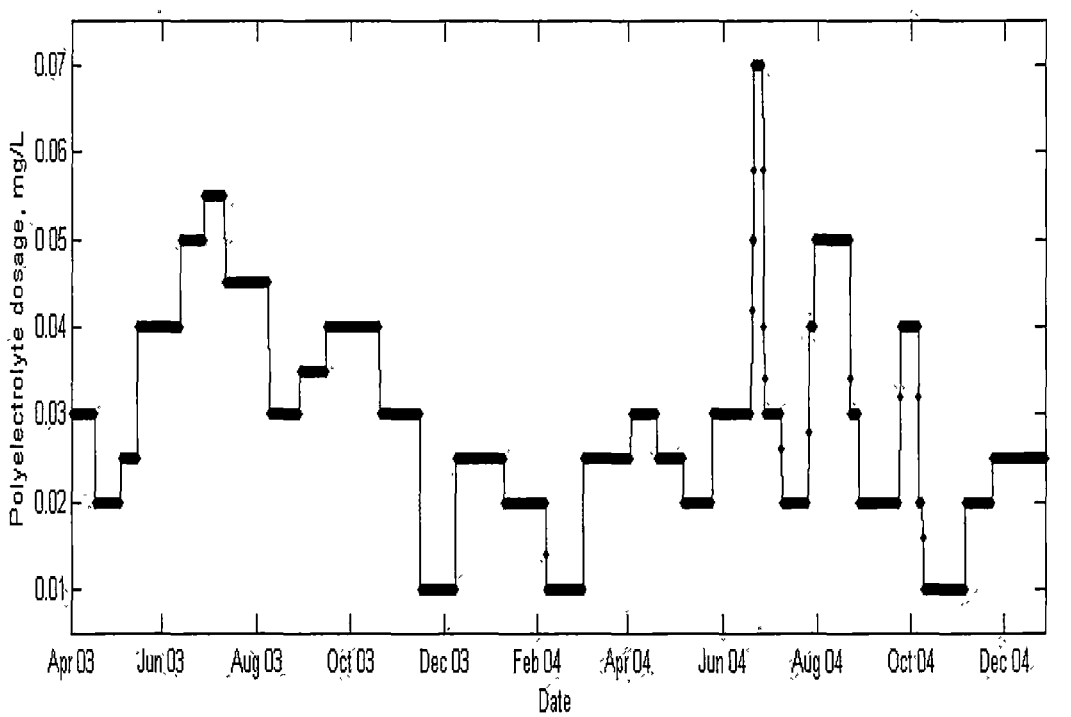


Figure B8 Polyelectrolyte dosage (BKWTP)

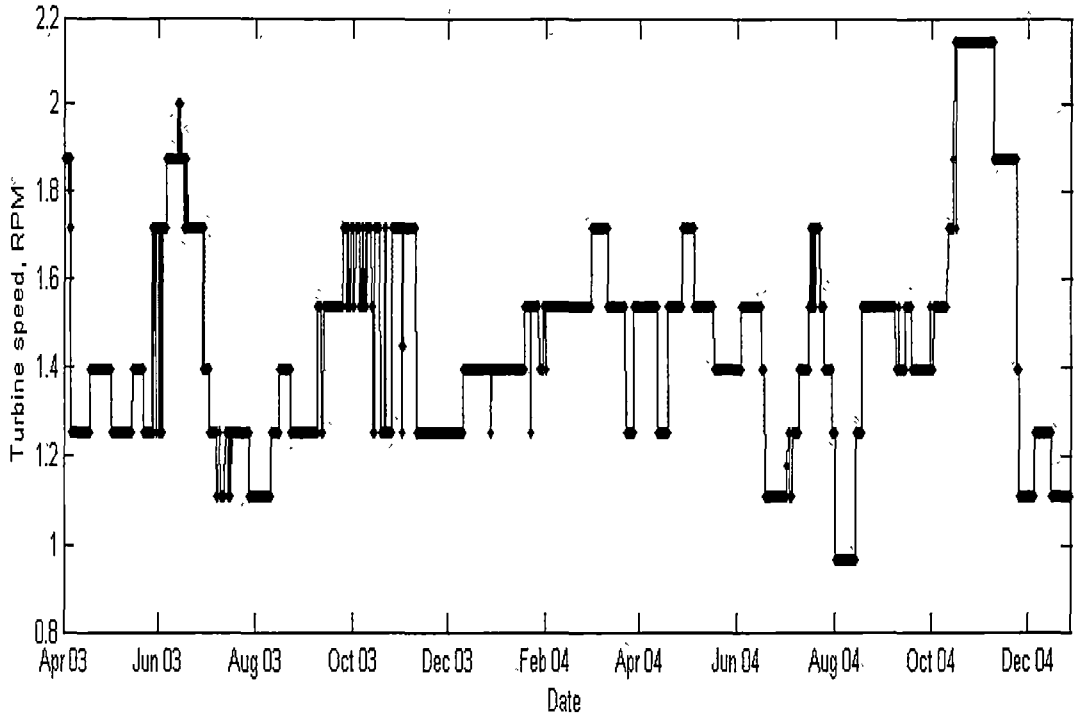


Figure B9 Turbine speed (BKWTP)

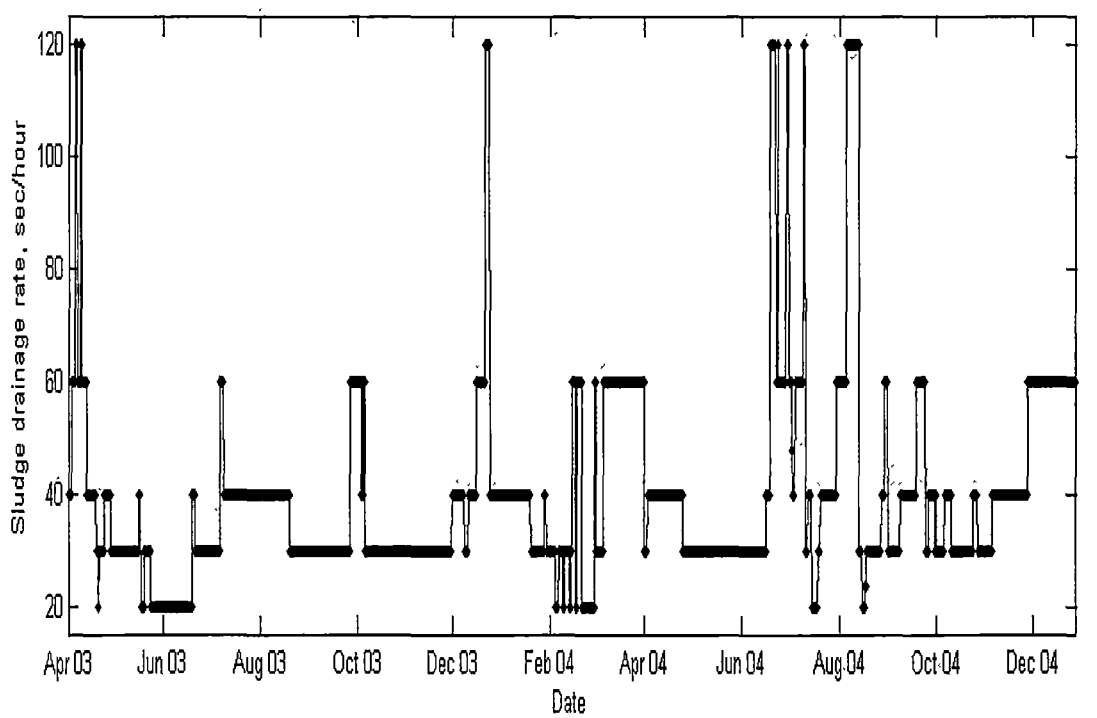


Figure B10 Sludge drainage rate (BKWTP)