

**MODELLING, CONTROL AND OPTIMISATION OF A DUAL CIRCUIT
INDUCED DRAFT COOLING WATER SYSTEM**

by

Cornelius Jacobus Muller

Submitted in partial fulfillment of the requirements for the degree

Philosophiae Doctor (Electronic Engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment and Information Technology
UNIVERSITY OF PRETORIA

August 2015

SUMMARY

MODELLING, CONTROL AND OPTIMISATION OF A DUAL CIRCUIT INDUCED DRAFT COOLING WATER SYSTEM

by

Cornelius Jacobus Muller

Promoter(s): Prof I. K. Craig
Department: Electrical, Electronic and Computer Engineering
University: University of Pretoria
Degree: Philosophiae Doctor (Electronic Engineering)
Keywords: modelling, verification, model predictive control, optimisation, energy, hybrid systems, advanced regulatory control, genetic algorithm, cooling water, dynamic simulation

The operation of any petrochemical plant requires the transfer of energy to and from the process, movement of material through the process piping and vessels and suppression of unwanted reaction or combustion. This is mainly achieved through the use of utilities which serve as auxiliary variables in plant operation and include electricity, steam, cooling water, hydrogen, nitrogen, compressed air and fuel gas.

These utilities have costs associated with them and it is therefore common practice to optimise on the use of these utilities. Other aspects to consider are the generation and transportation/transmission of these utilities to the point of consumption. These areas have not received much attention historically mainly due to the large differences that existed between product prices and utility costs. Recently, there has been a revival in the focus on efficient operation fuelled by global economic turbulence, higher energy cost, stricter environmental policies, dwindling fossil fuel supplies and the threat of climate change.

The losses encountered in the generation and transmission of utilities are in many cases substantial, especially in those utilities where releases to the atmosphere do not present significant safety or environmental risk such as compressed air, cooling water and steam

systems. Also in many cases, the transmission distance is substantial which gives rise to more losses and inefficiencies.

Furthermore, many electrical utility providers are using time-of-use electricity costs and maximum demand penalties. This necessitates scrutinising the exact time of consumption and peak consumption rather just the overall amount consumed.

This study explores the potential benefits of utility control and optimisation from a supply/generation point of view through the application of modern advanced control and optimisation techniques. The approach and results for a dual circuit cooling water system are illustrated.

The cooling water system is an example of a hybrid system where both continuous and discrete input variables are present. This complicates the formulation of control and optimisation solutions. A model of the system is developed and verified using real plant data. To improve model accuracy, a parameter estimation exercise is performed using a genetic algorithm as the optimisation platform.

Several control and optimisation schemes are then developed with varying degrees of complexity including advanced regulatory control (ARC), hybrid non-linear model predictive control (HNMPC) and economic hybrid non-linear model predictive control (EHNMPC). The ARC scheme uses classical advanced base layer control techniques, together with time and condition based switching logic, which do not require the use of a system model. The HNMPC and EHNMPC schemes make use of a system model in the formulation of the control solutions and use a genetic algorithm for optimisation.

The model is then used to evaluate the performance of the control and optimisation techniques in several simulation studies by comparing the results to that of a base case study. The results indicate that substantial energy and cost savings may be achieved without the need to install additional plant equipment.

OPSOMMING

MODELLERING, BEHEER EN OPTIMERING VAN 'N DUBBEL-BAAN GEÏNDUSEERDE-VLOEI VERKOELINGSWATERSTELSEL

deur

Cornelius Jacobus Muller

Promotor(s): Prof I. K. Craig
Departement: Elektriese, Elektroniese en Rekenaar-Ingenieurswese
Universiteit: Universiteit van Pretoria
Graad: Philosophiae Doctor (Elektroniese Ingenieurswese)
Sleutelwoorde: modellering, verifikasie, model voorspellende beheer, optimering, energie, hibriede stelsels, gevorderde regulerende beheer, genetiese algoritme, verkoelingswater, dinamiese simulاسie

Die suksesvolle bedryf van enige petro-chemiese aanleg is afhanklik van die oordrag van energie van en na die proses, beweging van materiaal deur die aanleg se pyp- en struktuurnetwerk en die onderdrukking van ongewenste reaksie of ontbranding. Hierdie doelwitte word hoofsaaklik bereik deur gebruik te maak van utiliteite wat as hulpveranderlikes dien en sluit onder andere elektrisiteit, stoom, verkoelingswater, waterstof, stikstof, saamgepersde lug en brandstofgas in.

Daar is koste aan sulke utiliteitstrome verbonde en dit is daarom algemeen om die benutting daarvan te optimeer. Ander areas om in ag te neem, is die opwekking en vervoer van hierdie strome tot by die verbruikspunt. Hierdie areas het histories nie soveel aandag geniet nie, hoofsaaklik as gevolg van die groot verskille wat tussen produkpryse en utiliteitskoste bestaan het. As gevolg van die onstuimighede in die wêreld ekonomie, hoër energiekoste, strengere omgewingswette, kwynende fossielbrandstofreserwes en drygende klimaatsverandering word daar egter tans nuwe klem op meer effektiewe aanlegsbedryf geplaas.

Die verliese wat aangetref word in die opwekking en vervoer van utiliteite is in baie gevalle

noemenswaardig, veral in gevalle waar verliese na die atmosfeer nie enige ware risiko's vir veiligheid of die omgewing inhou nie, byvoorbeeld in die geval van saamgepersde lug, verkoelingswater en stoom. In baie gevalle is die afstand waaroor die oordrag plaasvind groot wat verder tot verliese en ondoeltreffendheid bydra.

Verder maak heelwat elektrisiteitsverskaffers deesdae gebruik van tyd-van-verbruik tariewe en maksimum aanvraag boetes, waar die hoeveelheid energie wat gebruik word nie die enigste faktor is wat in ag geneem moet word nie, maar ook wanneer die energie gebruik word.

Hierdie studie ondersoek die moontlike voordele verbonde aan die beheer en optimering van utiliteite uit 'n opwekkings- en oordragsoogpunt, deur gebruik te maak van moderne beheer- en optimeringstegnieke. Die benadering en resultate word geïllustreer deur die toepassing daarvan op 'n dubbel-baan verkoelingswaterstelsel.

Die verkoelingswaterstelsel is 'n voorbeeld van 'n hibriede stelsel waar beide diskrete en kontinue insetveranderlikes teenwoordig is, wat die formulering van beheer- en optimeringsoplossings kompliseer. 'n Model van die stelsel word ontwerp en geverifieer deur gebruik te maak van ware aanlegdata. Ten einde die model se akkuraatheid te verbeter word 'n genetiese algoritme gebruik in 'n parameterpassingsoefening.

Verskeie beheer- en optimeringskemas word dan ontwikkel met wisselende grade van kompleksiteit, insluitend gevorderde regulerende beheer, hibriede nie-lineêre model voorspellende beheer en ekonomiese hibriede nie-lineêre model voorspellende beheer. Die gevorderde regulerende beheer maak gebruik van klassieke gevorderde basisvlak beheer tegnieke, tesame met tyd- en voorwaarde-afhanklike skakelingslogika, wat nie die ontwikkeling van 'n stelselmodel benodig nie. Die hibriede nie-lineêre model voorspellende beheer en ekonomiese hibriede nie-lineêre model voorspellende beheer skemas maak gebruik van 'n stelselmodel in die formulering van die beheeroplossing en gebruik 'n genetiese algoritme vir optimering.

Die model word daarna gebruik om die prestasie van die verskillende beheer- en optimeringstegnieke te evalueer in 'n verskeidenheid van simulasiestudies deur die resultate te vergelyk met die van 'n basisgeval. Die resultate dui daarop dat 'n noemenswaardige energie- en kostebesparing bewerkstellig kan word sonder dat addisionele toerusting geïnstalleer hoef te word.

LIST OF ABBREVIATIONS

AC	Alternating Current
APC	Advanced Process Control
ARC	Advanced Regulatory Control
BIP	Binary Integer Programming
CT	Cooling Tower
CV	Controlled Variable
CW	Cooling Water
CWS	Cooling Water Supply
CWR	Cooling Water Return
DCS	Distributed Control System
DSM	Demand-Side Management
DV	Disturbance Variable
EA	Evolutionary Algorithm
EHNMPC	Economic HNMPC
FF	Feed-Forward
GA	Genetic Algorithm
GRG	Generalised Reduced Gradient
HNMPC	Hybrid Non-linear MPC
LED	Light Emitting Diode
LP	Linear Programming
LSQ	Least Square
LTI	Linear Time Invariant
MIMO	Multi-Input Multi-Output
MINLP	Mixed Integer Non-Linear Programming
MIQP	Mixed Integer Quadratic Programming
MLD	Mixed Logical Dynamic
MPC	Model Predictive Control
MV	Manipulated Variable
NMPC	Non-linear MPC
NLP	Non-Linear Programming

PID	Proportional, Integral, Derivative
PLC	Programmable Logic Controller
POET	Performance, Operation, Equipment, Technology
PV	Pressure Valve
QP	Quadratic Programming
RCGA	Real Coded Genetic Algorithm
RLSQ	Recursive Least Square
RTO	Real-Time Optimisation
SISO	Single-Input Single-Output
SLP	Successive Linear Programming
SQP	Successive Quadratic Programming
TOU	Time-of-Use
TV	Temperature Valve
TW	Tempered Water
TWS	Tempered Water Supply
TWR	Tempered Water Return
VSD	Variable Speed Drive

ACKNOWLEDGEMENTS

I wish to express my gratitude to Sasol for making the data available for the model validation and providing funding for this study. I want to thank Mr. Mark Congiundi for his assistance in validating the model equations and his advice in general. Finally, I am very grateful to Professor Ian Craig for his guidance and support over many years.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION	1
1.1 Problem statement	1
1.1.1 Context of the problem	1
1.1.2 Research gap	3
1.2 Research objective and questions	3
1.3 Hypothesis and approach	4
1.4 Research goals	5
1.5 Research contribution	6
1.6 Overview of study	7
CHAPTER 2 LITERATURE STUDY	8
2.1 Chapter objectives	8
2.2 Energy management	8
2.2.1 Energy efficiency management	9
2.2.2 Energy cost management	10
2.3 Process control	13
2.3.1 Basic regulatory control	14
2.3.2 Advanced regulatory control	15
2.3.3 Optimal control	20
2.3.4 Model predictive control	21
2.3.5 Real-time optimisation	22
2.4 Optimisation	23
2.4.1 Genetic algorithm	26
2.5 System modelling	29
2.5.1 Dynamic and steady-state models	30
2.5.2 Linear and non-linear models	31



2.5.3	Time variant and invariant models	31
2.5.4	First principle and empirical models	31
2.5.5	Parametric and non-parametric models	32
2.5.6	Model examples	32
2.5.7	Energy modelling	32
2.6	Hybrid systems	34
2.7	Process equipment	36
2.7.1	Pumps	36
2.7.2	Fans	42
2.7.3	Heat exchangers	42
2.7.4	Control valves	43
2.8	Cooling water systems	44
2.8.1	Cooling tower theory	46
2.9	Conclusion	47
CHAPTER 3 APPROACH AND METHODS		48
3.1	Chapter overview	48
3.2	Process description	48
3.3	System modelling	52
3.3.1	Steady-state model	56
3.3.2	State-space dynamic model	66
3.3.3	Parameter estimation	69
3.3.4	Process modelling results	71
3.3.5	Model step tests	74
3.4	Control and optimisation	87
3.4.1	Base case	89
3.4.2	Advanced regulatory control	90
3.4.3	Hybrid non-linear model predictive control	94
3.4.4	Economic hybrid non-linear model predictive control	98
3.5	Conclusion	99
CHAPTER 4 SIMULATION RESULTS AND DISCUSSION		100
4.1	Chapter overview	100
4.2	Control and optimisation results	100

4.2.1	Base case	100
4.2.2	ARC results	104
4.2.3	HN MPC results	110
4.2.4	Economic HN MPC results	115
4.2.5	Case comparison	116
CHAPTER 5 CONCLUSION		125
5.1	Concluding remarks	125
5.2	Future scope	127
APPENDIX A MATLAB SIMULATION DETAILS		136
A.1	Simulink models	136
A.2	Matlab code description	140
A.2.1	Model verification	142
A.2.2	Parameter estimation using the ga function	149
A.2.3	Control and optimisation	150

LIST OF FIGURES

2.1	Example of a TOU tariff trend.	11
2.2	Tempered water pump performance curves.	37
2.3	Cooling water pump performance curves.	38
2.4	Performance illustration of pumps running in series on a system with a constant system curve.	40
2.5	Performance illustration of pumps running in parallel on a system with a constant system curve.	41
2.6	Valve characteristic curves with constant differential pressure across valve. . .	45
3.1	Dual circuit induced draft counter-flow cooling water system.	50
3.2	Simplified system representation.	53
3.3	Model disturbance inputs for validation (plant duty and wet-bulb temperature). .	71
3.4	Model tempered water temperature results (model response (solid line) versus the plant data (dotted line)).	74
3.5	Responses in f_{TW} , f_{CW} , and T_{TWS} (represented by the model matrix rows) to changes in U_{TW} and U_{CW} (represented by the columns).	76
3.6	Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in U_{TW} , U_{CW} (represented by the columns).	77
3.7	Responses in f_{TW} , f_{CW} , and T_{TWS} (represented by the model matrix rows) to changes in U_{CT} and OP_{TV} (represented by the columns).	78
3.8	Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in U_{CT} and OP_{TV} (represented by the columns).	79
3.9	Responses in f_{TW} , f_{CW} , and T_{TWS} (represented by the model matrix rows) to changes in OP_{PV} and Q_P (represented by the columns).	80
3.10	Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in OP_{PV} and Q_P (represented by the columns).	81

3.11 Responses in f_{TW} , f_{CW} , and T_{TWS} (represented by the model matrix rows) to changes in T_a and RH (represented by the columns).	82
3.12 Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in T_a and RH (represented by the columns).	83
3.13 Model disturbance inputs for first simulation (plant duty and wet-bulb temperature).	89
3.14 ARC scheme illustration.	91
3.15 HNMPC scheme illustration.	96
4.1 Controlled variables (Base case) for the first simulation.	102
4.2 Controlled variables (Base case) for the second simulation.	103
4.3 Controlled variables (ARC case) for the first simulation.	106
4.4 Manipulated variables (ARC case) for the first simulation.	107
4.5 Controlled variables (ARC case) for the second simulation.	108
4.6 Manipulated variables (ARC case) for the second simulation.	109
4.7 Controlled variables (HNMPC case) for the first simulation.	111
4.8 Manipulated variables (HNMPC case) for the first simulation.	112
4.9 Controlled variables (HNMPC case) for the second simulation.	113
4.10 Manipulated variables (HNMPC case) for the second simulation.	114
4.11 Controlled variables (Economic HNMPC case) for the first simulation.	117
4.12 Manipulated variables (Economic HNMPC case) for the first simulation.	118
4.13 Controlled variables (Economic HNMPC case) for the second simulation.	119
4.14 Manipulated variables (Economic HNMPC case) for the second simulation.	120
4.15 Power consumption (black line) and electricity cost (grey line) for the first simulation of the EHNMPC case.	121
4.16 Power consumption (black line) and electricity cost (grey line) for the second simulation of the EHNMPC case.	121
A.1 Simulink main page for the verification model.	137
A.2 Simulink main page for the base case and ARC model.	138
A.3 Simulink main page for the HNMPC and EHNMPC model.	139
A.4 Simulink CT bank.	140
A.5 Simulink CT.	141
A.6 Simulink CW pump.	142

A.7 Simulink CW plant.	143
A.8 Simulink TW pump.	144
A.9 Simulink TW plant.	145
A.10 Simulink CT fan switching logic.	146
A.11 Simulink CW pump switching logic.	146
A.12 Simulink TW pump switching logic.	147
A.13 Simulink Triggered Delayed Pulse logic.	147
A.14 Simulink PID logic.	148
A.15 Simulink mid-of-three logic.	148
A.16 Simulink solver settings.	149

LIST OF TABLES

3.1	Model parameters.	54
3.2	Optimised model parameters	73
3.3	Model correlation coefficients with and without parameter estimation.	73
3.4	CV and MV limits	97
4.1	Average power and energy consumption comparison.	123
4.2	Electricity cost comparison.	123
4.3	Constraint violation comparison.	124

CHAPTER 1

INTRODUCTION

This chapter presents the problem to be addressed in this study and clarifies the objectives this study aims to achieve. The research approach is also stipulated and a brief overview of the study is provided.

1.1 PROBLEM STATEMENT

1.1.1 Context of the problem

The process industry (including refineries and chemical plants) is a major global consumer of energy and even a small improvement in efficiency can have a significant impact on consumption worldwide. From a cost perspective, energy costs in refineries in the United States are estimated at between 50% and 60% of total cost, with 30% to 40% for chemical plants [1]. In South Africa the petrochemical industry contributed more than 20% of electricity consumption in manufacturing between 1993 and 2006 [2]. Historically, product prices far exceeded energy costs and energy optimisation was not considered a high priority. This has changed and energy management has received renewed focus in recent years due to increasing energy cost, dwindling fossil fuel supplies, the threat of climate change, struggling global markets, stricter environmental policies and energy shortages faced by some countries [3, 4, 5]. Furthermore, as this is a continuously changing environment, a solution that is infeasible at present may be feasible in a year or two, which means that optimisation opportunities should be re-evaluated regularly.

Whether for environmental reasons or purely financial, any company can benefit from run-

ning their processes more efficiently, especially with regard to energy consumption. In some instances processes are designed with energy efficiency in mind. This can have large long-term benefits though it comes at additional capital cost and more complex process dynamics (for instance by using process-to-process heat exchangers for heat integration and recovery). The norm though is the opposite and the subsequent opportunities for optimisation after commissioning can be attractive if the company is willing to pursue them as described for example in [6] and [7].

In the process industry energy is used in several forms typically through the use of utilities, including electricity, steam, tempered water, compressed air, fuel gas, hot oil, cooling water, and refrigerated water. Utilities are used as auxiliary variables to enable the addition and removal of energy to and from the system, to power process equipment, and to inhibit unwanted reaction. A reduction in the consumption of these utilities leads to a direct energy and cost saving. The less frequently considered areas for optimisation are the generation and transportation of these utilities. There are considerable energy losses in the generation and transmission of utilities [8, 9]. Some utilities are generated far from the process plant and have to be transported or transmitted there, resulting in more losses and potential inefficiencies. Another aspect to consider is the fact that inefficiencies and poor control of these utilities result in the need to run unnecessarily large buffer capacities to handle upset conditions, which typically means that the additional capacity is vented or purged during stable operating conditions. Therefore, control and optimisation strategies focussing on the optimal generation and transmission of utilities have clear benefit potential.

Although electrical energy can be used directly as a utility (for example to power electrical heaters) it is also typically used in the generation and transport of other utilities (for example in driving motors on pumps and compressors for cooling water or compressed air). Therefore, electrical energy is ubiquitous for almost all utilities. Most of the world's electrical power supply (more than 40%) is still generated from coal [10] which is an inefficient process with a typical efficiency of 35% from burning the coal to the point of consumption [11]. Therefore, the focus of this study will fall on electrical energy optimisation.

1.1.2 Research gap

The process industry is under constant pressure to improve efficiency and reduce cost and waste. To stay competitive, a company must find ways of reducing cost and improving efficiency simultaneously. Energy optimisation is a good candidate to achieve this without extensive capital expenditure, especially in the light of rapidly increasing electricity prices and the introduction of carbon taxes in many countries [4].

The latest research in energy efficiency improvements are mainly focussed on a few industries including power generation, the mining sector, smelting furnaces, paper and pulp plants, the cement industry, renewable energy sources, commercial buildings and smart grids [12, 13, 14, 15, 16, 17, 18, 19, 20]. There seems to be a shortage in the application of these energy efficiency focussed techniques in the process industry where the benefits could be substantial.

Therefore, the research gap that this work aims to address, is to provide a study illustrating the potential benefit of energy efficiency improvement initiatives in the petrochemical industry with particular attention to the generation and transportation of process utilities.

Furthermore, most published material on cooling water system models are very detailed and focus on the process design, which is not ideal for the purposes of controller and optimiser design [21, 22, 23, 24]. Therefore, a simplified model is required that sufficiently captures the behaviour of the system for simulation purposes and controller/optimiser design. This gap is also addressed in this research initiative.

1.2 RESEARCH OBJECTIVE AND QUESTIONS

Modern control and optimisation techniques aimed at energy efficiency improvement can be very powerful in industry, both for providing financial benefit and to reduce global energy consumption, resulting in reduced emissions. This will become more important in the future to ensure sustainable industry performance amid rising energy prices, diminishing fossil fuel supplies and more stringent environmental policies. Therefore, the main objective of this study is to show how to combine energy management and process control principles to achieve improved energy efficiency in the process industry with focus on utilities.

The following research questions summarise the focus areas of this study:

- Can the energy consumption model of the process in terms of discrete devices be combined with its dynamic process model for purposes of control and efficiency optimisation, especially considering the non-linear, interactive, and hybrid nature of the process?
- Is it possible to perform the optimisation in a single optimisation layer, considering the complexity and non-linearity of the combined model, or is it better to perform discrete and continuous control and optimisation in separate layers in the conventional hybrid control manner?
- Will APC or ARC be successful in optimising energy efficiency individually or is a combination of the two superior?
- How can time-of-use cost data be incorporated into the optimisation algorithm to perform demand side management of process units and utilities?

1.3 HYPOTHESIS AND APPROACH

The null hypothesis for this study is that the operating efficiency of a process can be improved by developing an energy model for the process and by using this model in the application of advanced control and optimisation to minimise energy consumption subject to process constraints.

The focus falls strongly on process units where there is a combination of continuous and discrete energy consuming equipment, for example heat exchangers (continuous) and pumps (discrete). These systems are referred to as hybrid systems. Examples of such systems are cooling water systems, distillation columns with fin fan banks for cooling, and conveyor systems. The presence of both discrete and continuous components significantly complicates controller and optimiser design [25, 26].

The optimisation itself is done using advanced regulatory control (ARC) and advanced process control (APC) techniques. Depending on the nature and complexity of the problem, the techniques could be used individually or in combination with each other. The chosen algorithm may have to consider both dynamic control issues and steady-state optimisation. If a process

does not contain much dynamics, an optimiser alone might suffice. Process plants, however, usually contain some significant dynamic behaviour which will require ARC/APC.

The solution is evaluated through a case study on a cooling water system. Firstly, a practical simulation model of the process is developed to represent the actual process. The model is then verified using real plant data and a benchmark base case of operation is established using the current base-layer control configuration which exists on the actual plant. Thereafter, the model is presented in state-space form to aid in controller and optimiser design. Different control and optimisation techniques are applied and tested using the verified model in simulation case studies to determine the efficacy of each approach in minimising energy consumption and/or cost while honouring process constraints. The results are then analysed to determine the energy efficiency improvement achieved compared to conventional base-layer control without optimisation.

1.4 RESEARCH GOALS

As mentioned, this research study aims to unify the fields of process control and energy management by using advanced process control and optimisation techniques with energy management as the core deliverable. Furthermore, different advanced process control techniques are applied to determine the relative success in optimising energy usage of each approach. In addition a generic guideline for energy optimisation on process units, which combines discrete and continuous energy components in the same optimisation problem, is described, which can be applied to a number of processes.

The goals can be summarised as follows:

- Successfully develop a model of the cooling water system suitable for simulation as well as control and optimisation design purposes.
- Utilise the model to create a simulation environment in order to perform simulation studies of the system.
- Develop several control and optimisation solutions aimed at reducing energy consumption and/or cost.

- Illustrate that the solutions are successful in optimising the energy consumption and/or cost of the system.

1.5 RESEARCH CONTRIBUTION

Refineries and chemical plants account for a significant portion of energy usage in the world. Therefore, a seemingly insignificant efficiency improvement in this industry can have a substantial impact on global usage and considerable savings could be realised by reducing consumption in these industries.

Therefore, the main contribution of this research is to add to the available knowledge base concerning energy efficiency improvement in the process industry, with emphasis on the utilities used in this industry. The contributions are in the fields of system modelling as well as advanced control and optimisation. In addition, the system under consideration is of a hybrid, non-linear, multi-variable nature and the solutions therefore form part of the active research areas of hybrid and non-linear multi-variable control [25, 26] as well as the field of energy management [4, 5, 12, 13, 15, 16, 19, 20, 27, 28].

The proposed solutions could be implemented without significant capital expenditure and will therefore be an attractive option for efficiency improvement in an industry under constant pressure for cost cutting and improved efficiency. This is further motivated by the introduction of carbon emission taxes in many countries with significant financial implications.

This study also presents a concise system model that is sufficient for process simulation aimed at control and optimisation studies. It also provides a convenient starting point for deducing simplified models for use in controller and optimiser design and implementation.

The demonstrated approach may be used on other process units, especially where discrete and continuous components are used, such as heater banks and distillation columns with fin fan banks. Furthermore, the optimisation can easily be extended by taking time-of-use data into account and thereby combining demand side management with this suggested approach as is illustrated.

The following publications have originated from this work:

- C. J. Muller and I. K. Craig, “Cooling Water System Modelling for Control and Energy Optimisation Purposes,” in: *Proceedings of the 19th IFAC World Congress*, Cape Town, pp. 3973–3978, September 2014.
- C. J. Muller and I. K. Craig, “Modelling of a dual circuit induced draft cooling water system for control and optimisation purposes,” *Journal of Process Control*, vol. 25, pp. 105–114, January 2015.
- C. J. Muller and I. K. Craig, “Energy reduction for a dual circuit cooling water system using advanced regulatory control,” *Submitted to: Applied Energy*, July 2015.
- C. J. Muller and I. K. Craig, “Economic Hybrid Non-linear Model Predictive Control of a Dual Circuit Induced Draft Cooling Water System,” *Submitted to: IEEE Transactions on Control Systems Technology*, November 2015.

1.6 OVERVIEW OF STUDY

This study specifically focuses on a dual circuit induced draft cooling water system. Chapter 2 provides a comprehensive literature study to illustrate the current knowledge base pertaining to energy management, process control, optimisation, cooling water systems and hybrid systems. Chapter 3 describes the methodology and approach followed in solving the problems associated with the different components of the study. The first part of the chapter covers modelling of the system which includes a simulation model to represent the actual plant and a state-space model to aid in control/optimisation development. The next part of the chapter focuses on the control and optimisation of the system in order to reduce energy consumption and/or cost while honouring system constraints and ensuring stable operation. Thereafter, Chapter 4 presents an analysis and discussion of the results, including a comparison of the newly developed control solutions to the conventional base-layer control (currently implemented on the process) to determine the efficiency improvement achievable with the application of advanced control and optimisation techniques. Finally Chapter 5 gives some concluding remarks and Appendix A contains more details on the platform specific solution implementation.

CHAPTER 2

LITERATURE STUDY

2.1 CHAPTER OBJECTIVES

This study utilises concepts that span several research fields, namely energy management, process modelling, process control and optimisation. The relevant literature in these fields is covered in this chapter. Furthermore, some theory on the relevant process equipment such as pumps, cooling towers and heat exchangers is also presented.

2.2 ENERGY MANAGEMENT

In broad terms, energy management refers to the process of identifying and implementing means for performing the same task with less energy and/or saving energy cost for a specified consumption.

Energy efficiency is a good measure of the effectiveness of an energy management initiative that focusses on energy usage reduction. The aim is to put as much of the supplied energy to good use in a system and therefore minimise waste. Furthermore, if less energy can be inserted into the system, while still achieving the desired result, a dual benefit is achieved. An example of the former is to insulate a water heater to prevent heat loss to atmosphere. An example of the latter is to use a heat pump instead of an electric element for water heating.

Initiatives to reduce the cost of energy for a specified consumption include, for example, power factor correction and demand side management (focussed on time-of-use and maximum

demand optimisation). Many companies are installing electricity metering systems to enable accurate usage monitoring and to motivate for future improvement projects.

In this study, process utilities have been identified as the vehicles for energy transfer. Although it may seem that energy efficiency improvement initiatives are embraced globally, many barriers still inhibit efficiency improvement opportunities from being implemented [4].

2.2.1 Energy efficiency management

Energy efficiency can be defined as the ratio of useful energy output to energy input of a system [14]. There are several models for describing energy efficiency of systems. Xia et al. classify energy efficiency into four categories namely performance, operation, equipment, and technology efficiency (POET) [14, 20, 28, 29, 30]. This model provides a unifying framework for energy efficiency. Some popular energy efficiency improvement initiatives, that fall mostly in the categories of technology and equipment efficiency, are the use of heat pumps, solar geysers, waste gas boilers, equipment insulation, heat integration, and motorised pulleys, which all affect the overall efficiency of a system. Control and optimisation also play important roles in the operations efficiency category as it mostly deals with the aspect of time coordination of system components.

As discussed in Section 1.1.1, when utility consumption/usage by the main process is minimised it results in a direct energy saving, which would improve the energy efficiency of the process that is using the utility. Another less direct area for optimisation is in the generation and transmission/transportation of utilities. Optimising the generation and transportation of utilities will also improve overall process efficiency. This may be achieved by improving the efficiency of the utility process (for example by switching equipment off when not required – operation efficiency), by using more efficient equipment (such as more efficient pumps – technology and equipment efficiency) or by making better use of existing equipment (for example by regular service intervals or by insulating lines – equipment efficiency).

Examples: In [20], several different operating scenarios are considered for a belt conveyor system and optimised for energy consumption by adjusting speed, feed rate and working time, while honouring process constraints. In [12], another conveyor system feeding a coal fired

power plant is optimised with regard to energy consumption (and cost). In [28], efficiency of a pumping system is improved by providing a means of determining the optimal pump sizing for a new system as well as optimal operating conditions for an existing pumping system, both falling in the category of operation efficiency.

2.2.2 Energy cost management

Energy cost management is less concerned with how much energy is consumed and more with the energy cost associated with the operation of a system over a period of time. This is mostly relevant to electrical energy and falls within the performance and operating efficiency categories of the POET framework. Two commonly used cost management techniques for electrical energy are power factor correction and demand side management.

2.2.2.1 Power factor correction

Power factor correction is used to reduce reactive power and minimise losses caused by cycling current to and from the utility supplier. Furthermore, by reducing the reactive power, the apparent power is reduced, which has a direct effect on the billed amount and allows the user to specify a lower notified maximum demand, thereby adding an additional saving. The power factor is a function of the equipment used and is typically solved by installing capacitor banks in the distribution station in the case of inductive industrial loads like inductive motors.

Although good results have been obtained by using this technique (and in some cases adaptive power factor correction techniques also cater for different plant operating scenarios), it is usually a once off improvement and does not cater for other inefficiencies in the way the plant is operated. Furthermore, as the plant conditions change and equipment is added, removed or replaced, the capacitor banks have to be modified as well [31].

2.2.2.2 Demand side management

Demand side management refers to the planning and implementation of energy usage manipulation techniques at the consumption or load side, in order to produce desired changes in the utility's load shape [27].

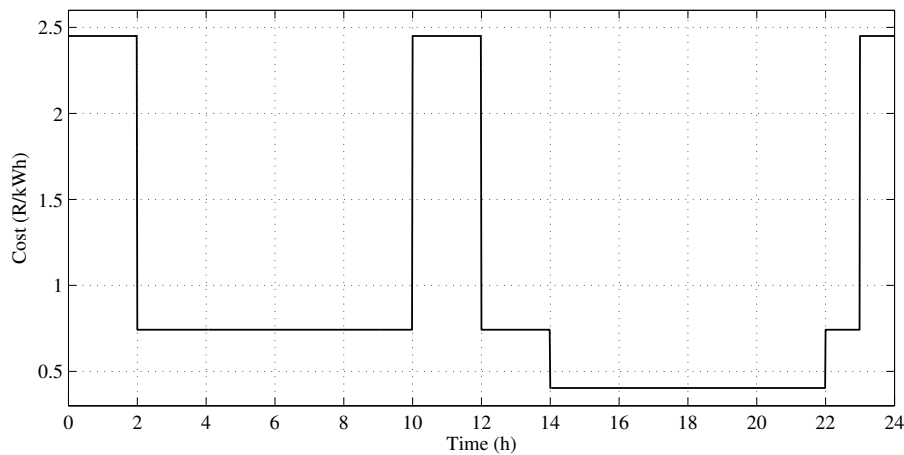


Figure 2.1: Example of a TOU tariff trend.

There are different motivations for applying demand side management, such as simplification of the generation process by spreading consumption more evenly over time or to economically schedule power consumption to reduce cost of consumption (which is addressed in this study). This differs from the traditional approach of controlling supply purely from a generation perspective based on the load on the grid and an assumed future consumer demand. Demand side management is typically encouraged by the utility supplier through one or more of a variety of techniques aimed at actively influencing the demand for electricity [13, 15, 27, 28].

One such technique is the enforcement of *time-of-use* (TOU) tariffs. This means that the cost for electricity is higher during certain times of the day, such as peak periods versus off-peak or standard periods. Therefore, at times when supply is under pressure, the rates are higher to encourage a more conservative usage from the consumers. A response to TOU is *load shifting*, when less energy is consumed during peak periods and more during off-peak periods. Rates may also be dependent on the day of the month, month of the year, and season. From a control and optimisation perspective, the time-of-use rate data can be included in the control and optimisation problem to allow the controller to apply load shifting where possible to minimise electricity cost for a specific production rate. Figure 2.1 shows an example of a TOU tariff over the course of a day.

Related to TOU is *maximum demand*, which requires a consumer to agree not to use more than a predefined amount of power at any time. Should the consumer not honour this

agreement, large penalties may be levied.

An effect of implementing time-of-use and maximum demand rates is that the load becomes more evenly spread over time. Therefore, consumption from the grid is reduced during peak periods and increased during off-peak periods. This is referred to as *peak shaving* or *peak clipping* and *valley filling*. Systems with energy storage capacity may, for example, be able to store energy during off-peak times for use in peak periods.

The supplier may also have the ability to manipulate the load directly through shutting down or cycling the customer's equipment remotely whenever the need arises, such as when supply is constrained. This is referred to as *direct load control* or *load shedding*. The reason for applying load shedding may also be to accomplish peak shaving. A related concept is *flexible load shape* where the quality of the supply is intentionally reduced in exchange for incentives on the demand side or to manage constraints on supply (for example by pulsing the supply to electric water heaters as opposed to providing continuous supply during certain times).

Capacity market programs refer to agreements where the consumer has to reduce loads according to predefined levels of contingency. If the customer fails to do this, fines may be applicable or the consumer might run the risk of being completely cut off by the supplier.

Strategic conservation is a method by which the utility supplier aims to influence the consumer's natural tendency to move to more energy efficient installations in order to accelerate it. Examples include rebate or subsidy programs for installing more energy efficient equipment (such as solar geysers or energy efficient light bulbs such as LED or fluorescent) and weatherproofing of existing equipment.

Strategic load growth aims to influence the type or shape of the demand growth and is related to valley filling. Realising that growth in demand is inevitable, being able to determine or influence the shape of this load growth can be very beneficial to both the supplier and consumer. Examples include the promotion of dual fuel heating systems (for example where gas is used during peak times and electricity is used during off-peak times) or reduced rates during off-peak times (related to TOU).

Several recently published studies take these methods into account in order to minimise energy consumption and/or cost. Model predictive control (MPC) and optimal control (dis-

cussed in the next section) lend themselves to intuitively including time-of-use data into the optimisation problem in order to reduce energy consumption and/or cost.

Examples: In [13], load shifting is applied to a run-of-mine ore milling circuit to reduce energy cost. In [15], demand side management is performed for a grid-connected battery-photovoltaic system, under TOU tariffs (with power selling), to minimise cost by optimising energy flow between the grid and the system. In [12], [19] and [20], belt conveyor systems are optimised in terms of energy cost, taking TOU rates into account. In [28], both open-loop optimal control and closed-loop model predictive control are used to optimise energy cost using TOU and maximum demand data for a pumping system.

2.3 PROCESS CONTROL

The field of process control is considered a mature field, yet thousands of publications are still produced every year. Among these, the number of articles focussing on energy efficiency in the process industry is surprisingly low. Although energy usage has been scrutinised in the process industry for many years, few studies have been done on techniques for energy optimisation specifically tailored for the process industry (although several companies are developing products and services to this end). Regardless of what source the energy is supplied from, the amount of energy consumed by the process should be controlled and optimised for efficiency. Fortunately, the field of process control contains a wealth of information and techniques which can be applied to various industries and processes.

There are many ways to save on energy consumption in the process industry though, in general, processes tend to run at far less than optimal levels. In some cases, energy efficiency is considered during the design phase of the plant and manifests itself in the form of process-to-process heat exchangers (or heat integration), advanced control techniques, inherently efficient equipment and insulation. All of these techniques require additional capital expenditure during construction and are therefore often omitted, with the onus on the operations team to implement optimisation initiatives after commissioning.

One popular approach is the application of advanced process control (APC) though it is seldom implemented purely for energy optimisation. Consequently, the model developed for the purposes of APC will typically contain some indirect energy data, but will not focus

entirely on energy usage and will therefore not be entirely sufficient for energy optimisation. This leaves room for the development of techniques which focus more on energy efficiency improvement. Furthermore, process utility systems such as cooling water systems have not enjoyed much attention, which leaves room for improvement. The following subsections describe some of the control techniques that could be used for improved efficiency.

2.3.1 Basic regulatory control

The concept of feedback control originated centuries ago, although it was not until the advent of modern computing that it became truly ubiquitous. Control now plays a vital role in the majority of technological fields, including telecommunication, automotive, aerospace, process control, mining and minerals, robotics, agriculture and medical technology to name a few [3]. Many complex control algorithms exist and many are specifically created for specific applications, though the foundation of regulatory control lies in the concept of negative feedback.

Negative feedback refers to the principle of comparing the current state of some primary variable (the controlled variable) to a desired value and making adjustments to another secondary variable (the control handle or manipulated variable) that has the ability to influence the behaviour of the former. The effect of the adjustment is then assessed by taking another measurement of the primary variable and repeating the process in order to minimise the difference (or error).

When the control handle only has two states (for example an electric element on a water heater having only on and off states), a typical mode of control is on-off control. Here, the control element is switched on or off depending on the direction of deviation of the controlled variable from the desired value (or set-point). On-off controllers typically have dead-bands defined to prevent toggling between states when the process variable is fluctuating slightly around the set-point. On-off control results in cyclic behaviour around the set-point. The smaller the dead-band, the tighter the control (though this comes at the cost of more regular switching of the control element). Examples are temperature control in homes, refrigerators, ovens and water geysers [22].

The workhorse in most control intensive industries is the Proportional, Integral, Derivative (or

PID) controller. The proportional term acts immediately on a change in error (in proportion to the error), the integral term acts on the actual value of the error, and the derivative term acts on the rate of change of the error. The algorithm has many variations in form and works best with linear time invariant (LTI) single-input single-output (SISO) processes of a modest model order. The direction of the controller is defined as either direct or reverse acting. When a controller has to increase its control output in order to increase the value of the controlled variable, it is a reverse acting controller. On the other hand, if the controller has to decrease its control output to increase the controlled variable, it is a direct acting controller [21, 22, 32, 33, 34].

2.3.2 Advanced regulatory control

There are several base-layer¹ control techniques that can be used to supplement the standard PID feedback controllers typically used in industry. This is required when processes exhibit behaviour other than those that the PID is effective at controlling. These may include processes that have interaction, non-linearities, higher order dynamics and/or dead time. Examples of typical advanced regulatory control (ARC) techniques are ratio control, feed-forward control, override selector control, cascade control, split-range control, valve position control and inferential control.

Using a combination of these techniques can result in a very powerful, robust and versatile control scheme. It is however still limited with regard to the degree of optimisation achievable, as none of these techniques make use of an optimisation algorithm and are mostly based on the assumption that the process demonstrates linear behaviour.

When considering plant-wide control [32, 33, 35, 36, 37], these techniques become invaluable. They provide the means for stabilisation, allow for constraint handling and may provide a certain extent of optimisation. Therefore, ARC is applicable to multiple layers of the control hierarchy (regulatory control, supervisory control and local optimisation) and aids in achieving the operational and economic objectives as set out in the design phase.

Some ARC techniques are briefly described below. Many texts exist for more detailed de-

¹Base-layer techniques refer to techniques that are implementable on the basic control system (do not require additional specialised software or hardware).

scriptions of these techniques – see for example [21], [22], [32], [38] and [39].

2.3.2.1 Cascade control

In cascade control, the output of one controller forms the set-point of another. This is typically used where one feedback controller (the master or primary controller) controls a slow primary process and another controller (the slave or secondary controller) controls a faster secondary process.

An example is a master level controller that manipulates the set-point of a slave flow controller. The main reasons for applying cascade control are linearisation of the process for the master controller and better disturbance rejection.

In the level-flow cascade example, disturbances may come in the form of pressure changes upstream or downstream of the flow valve, which affect the flow rate at a specific valve opening. The flow controller would be able to recover from such a disturbance quickly, whereas, in the case where the level controller acts directly on the valve, the disturbance would persist until the slower level process deviates from set-point before the controller would act to compensate for the change in flow rate. Furthermore, when the level controller is set to act directly on the valve, the flow response for a specific change in valve opening will typically vary as the valve moves from low lift to high lift or when process conditions change. This leads to non-linearities in the valve-to-level response. Once again, the flow controller would be able to compensate for these non-linearities on a much faster time scale than the level controller. Therefore, the request for a change in flow rate from the level controller would be honoured across the range of valve lift, thereby linearising the flow-to-level response.

For more information and examples on cascade control, see [21], [22], [32], [38] and [39].

2.3.2.2 Feed-forward control

With feed-forward (FF) control, a known disturbance variable is measured and the measurement (together with some characterisation of the process response to the disturbance) is used to directly adjust the manipulated variable. Therefore, action is taken pro-actively based on the measured disturbance rather than purely based on feedback control, resulting

from a deviation of the controlled variable from set-point. This is especially advantageous in processes with slow dynamics. If the characterisation of the disturbance effect is perfect, the feed-forward control would be able to prevent deviation from set-point completely. This is generally not the case in practice and therefore feed-forward control is usually implemented as a supplement to feedback control. Two types of feed-forward control exist namely additive and multiplicative feed-forward.

In the additive case, a change in the measured disturbance variable is scaled with a feed-forward gain value and added to the output of a feedback controller or sent directly to an actuator or slave controller. This is typically used in conjunction with a feedback controller. The aim is to cancel the effect of the disturbance variable on the controlled variable as soon as it is detected. The expected effect (disturbance model) is used to predict what the response in the controlled variable will be and the output of the controller (or the manipulated variable) is adjusted in such a way as to counter the effect of the disturbance. The accuracy of the disturbance model determines how effective the action is. Therefore, when the actual process differs from the model, the performance degrades. In the petrochemical industry, a lead-lag plus dead-time function is typically used to capture the disturbance model due to the fact that most of the processes can be fairly accurately approximated as first-order or ramp functions with possibly some dead-time. Determining the lead, lag, and dead-time values is trivial when both the process and disturbance models are first-order.

In the multiplicative feed-forward case, the disturbance signal is multiplied by another variable (typically the output of a feedback controller) to form the output signal sent to the actuator or slave controller. This is analogous to using a feedback controller that cascades to a ratio controller (discussed in the next section).

One very useful application of feed-forward control is the decoupling of two or more interacting feedback control loops. When feedback control loops interact, the action of one loop affects the controlled variable of the other and vice versa. Therefore, the system is no longer SISO but rather multi-input multi-output (MIMO). By adding feed-forward action from the output of one feedback controller to the other controller, the interacting actions are anticipated by each controller which allows the loops to be decoupled. This reduces the degrading effect of the interaction on the control performance.

For more information and examples on feed-forward control, see [21], [22], [32], [38] and [39].

2.3.2.3 Ratio control

Ratio control can be viewed as a form of feed-forward control. A disturbance variable (or wild variable) is measured and another variable (the manipulated variable) is adjusted in ratio to the former. This is typically used in blending of two product streams or controlling flow ratios in a variety of processes such as distillation and reaction. When the required ratio is not constant, the ratio controller set-point can be adjusted manually or by another controller, for example a feedback controller using a temperature or composition measurement. This cascading of a feedback and ratio controller is similar to the multiplicative feed-forward/feedback control combination mentioned above. First-order plus dead-time functions are sometimes used to shape the response of the ratio control in cases where the manipulated stream should not be changed immediately due to the process dynamics.

For more information and examples on ratio control, see [21], [32], [38] and [39].

2.3.2.4 Override selector control

When there is more than one controlled variable sharing a manipulated variable, override selector control can be applied to allow for transition of control from one controller to another. The most common variants of override selector control are high, low and middle selection. In most cases, override selector control is used for constraint handling, where one variable is the normally controlled variable and another is a constraint, that has to be honoured under certain conditions (at the temporary expense of the primary controlled variable).

An important consideration in the use of override selector control is the prevention of wind-up in the non-selected controller(s). This is usually achieved by allowing the output on the non-selected controller to float only a calculated distance away from the selected control output. The distance is calculated as the product of the gain and error of the non-selected controller. Therefore, the further the non-selected controller is from set-point, the larger the allowed deviation is. On the other hand, if the non-selected controller is approaching its set-point, it means that it should prepare for selection to be transferred to it and the allowed deviation

becomes progressively less until eventually it is equal to zero at the point of transfer (when the non-selected controller is at set-point). This allows for bump-less transfer of control from one controller to the next.

An example of the application of override selector control is in the control of fuel gas to a multi-burner furnace. To maintain the desired duty in the furnace, the flow of fuel gas would have to be maintained at set-point (assuming there is no change in calorific value of the gas) regardless of whether burners are being brought on-line or taken off-line. The constraint to be managed during these events is preventing the pressure from dropping below a point where the flame could be lost. Therefore, the normal controller would be the flow controller whereas the constraint handling controller would be a pressure controller with a set-point slightly above the minimum allowed pressure. When a burner is suddenly brought on-line, the flow resistance of the system decreases which would cause a temporary decrease in burner inlet pressure. At the same time the flow rate would temporarily increase causing the flow controller to cut back on the control valve. This causes a further decrease in burner inlet pressure with the possibility of losing the flame(s). Therefore, the override selector control would temporarily transfer control to the pressure controller which would prevent the control valve from cutting back too far (at the temporary expense of a higher flow rate). As soon as the transient response has settled out and the pressure starts increasing again, control would be transferred back to the flow controller.

For more information and examples on override selector control, see [21], [22], [32], [38] and [39].

2.3.2.5 Valve position/throttling control

Valve position control refers to using the output of one controller as the controlled variable of another controller with the ability of manipulating the operating point of the first controller. This is used for example where two control valves are installed in series in order to accurately control the flow in the line over a wide operating range. The first controller will manipulate one valve in order to control the flow-rate at set-point whereas the second controller will manipulate the position of the second throttling valve to ensure that the first controller's valve position remains close to a desired value (for example 50%).

This technique is also useful for handling throughput constraints on systems, where the valve position of the bottleneck control valve is used as an indication of when a throughput constraint is hit, and a valve position controller throttles back on the input to the system accordingly, or moves the point at which the production rate is set (typically through the use of an override selector). Therefore, this technique may be used to eliminate the need for complex logic for reconfiguring the control structure when certain constraints become active.

For more information and examples on valve position control, see [38] and [39].

2.3.3 Optimal control

Optimal control is a technique by which control inputs are calculated that will drive a process to a desired state while honouring system constraints and at the same time minimise cost (or maximize return). The cost is described in terms of an objective function, which can be formulated in a number of ways, depending on the criteria for optimisation (time, energy, cost, etc.). The algorithm uses a model of the process to predict its behaviour and uses this prediction as a constraint in the optimiser [40]. Optimal control is typically implemented as an open-loop control strategy over a finite time period.

Several studies have shown how optimal control may be used to improve the energy efficiency and/or cost in certain systems. Wu et al. demonstrated its use for optimising power flows in terms of cost and stability between a photovoltaic-battery system and the grid using TOU tariffs and sell back rates in an open-loop fashion [15]. In [20], several different cost functions are used with a belt conveyor system to optimise energy cost and consumption during different operating scenarios where the conveyor speed can be adjusted using a VSD (continuous optimal control). In [19], switched optimal control is performed using binary integer programming (BIP) with time-of-use rates to minimise the energy cost of a belt conveyor system in a colliery where the conveyors can only be on or off. In [12], another belt conveyor system feeding a coal fired power plant is optimised using both switched and continuous optimal control. In [28], an optimal control strategy is deployed on a pumping system to reduce energy cost using both TOU and maximum demand data.

2.3.4 Model predictive control

Many advanced control techniques exist though none have had more success in the process industry than Model Predictive Control (MPC). This technique has been implemented to the extent that the term APC has become analogous to MPC in the process industry [41]. Therefore, this study will focus on MPC as an APC technique.

From its first formulation in the 1970's [42] and application in the refining industry in 1980 [43], there has been a rapid growth in the application of MPC as well as in research by the academic community. Today, there are thousands of MPC controllers implemented globally in the process industry, with several companies developing proprietary versions of the algorithm accompanied by tools for modelling and performance monitoring [44].

MPC is a control technique where optimal control is applied in an iterative fashion. At each iteration, the problem is solved based on the latest measurement data and the past move pattern. The first move vector is implemented where-after the process is repeated using the latest process measurements. This allows for feedback and compensates for plant-model mismatch. Therefore, it is also sometimes referred to as receding horizon control because the control horizon retreats after every iteration. The popularity of MPC stems from its ability to include equality and inequality constraints directly in the control law, flexibility in formulating the objective function and defining the process model, the ability to accommodate multi-variable systems and its disturbance rejection ability due to its inherent feed-forward action [6, 16, 45, 46].

There is an abundance of literature on MPC, especially regarding application to the process industry and specifically for control of non-linear interactive processes like distillation. Although energy optimisation can easily form part of the objectives of MPC, there is limited literature focussing on the application of MPC specifically for energy management such as demand side management.

Examples: In a recent publication by Matthews and Craig, MPC was deployed as a means of demand side management for a run-of-mine ore milling circuit using time-of-use cost data [13]. In [15], MPC was used in the optimisation of a photovoltaic-battery system to minimise cost amid disturbances and compared to the open-loop optimal control case. In [16], MPC

is used with weather and occupancy forecasts to minimise the energy usage for a building control system while maintaining desired comfort requirements. In [28], an MPC strategy is deployed on a pumping system to reduce energy cost using both TOU and maximum demand data. Another application is in water treatment [18].

2.3.5 Real-time optimisation

Real-time optimisation (RTO) is a technique for determining the economic optimal steady-state target values for key process variables, typically serving as supervisory control for a lower level dynamic advanced controller such as MPC. This allows for a degree of decoupling between the economic and control objectives [13, 46].

Several MPC controllers may be guided by a real-time optimiser to further improve economic performance. MPC packages commonly make use of dynamic linear models and linear/quadratic optimisation techniques, which limit the amount of economic optimisation obtainable, especially when the process is non-linear. RTO on the other hand more often utilises a rigorous first principle steady-state model, which can either be linear or non-linear, to guide the steady-state behaviour of the plant closer to the true economic optimum. It is also not uncommon to use the steady-state gain matrix of the linear dynamic model as the process model for the RTO and use a non-linear cost function for optimisation.

By optimising this economic model and determining the corresponding process variable values, these values can be written as targets to an MPC controller. An MPC is typically more concerned with the dynamic performance of the system and optimising transient behaviour than with the long-term (or steady-state) behaviour. In other words, the MPC manages the dynamic control of the plant to maintain a desired operating point, whereas the RTO determines where the optimal operating point is. Although RTO conventionally uses first principle models such as that obtained from mass and energy balances, some applications use empirical models [47].

RTO usually operates at a much lower execution frequency than MPC and the base layer and requires the plant to be stable before changes can effectively be suggested to the lower levels. Therefore, RTO applications often require steady-state detection and variable filtering to suppress higher frequency dynamics and disturbances [22]. It may also require the estimation

of some model parameters to update the model at every iteration. The period between executions is a function of the process response time and the performance of the MPC. Similar to MPC, RTO has been studied thoroughly for application in the process industry though not always explicitly for energy optimisation. In cases where the layers of MPC and RTO are merged into a single economic dynamic optimisation problem, the solution is often referred to as Economic MPC (EMPC) [46].

Examples: In [13], RTO is used in combination with MPC to optimise the energy cost of a grinding mill. The RTO has an execution time of 30 minutes compared to a 10 second execution time for the MPC. A non-linear cost function was used for RTO and related to the linear dynamic plant model by using the steady-state gain values. In [6], RTO is used to write economically optimal set-points to an MPC controller for the optimisation of a fuel gas blending system using a first principle non-linear steady-state model.

2.4 OPTIMISATION

There are many different techniques for finding the optimal value of a function or system. The applicable technique depends on the nature of the system (linearity, continuity and/or number of dimensions). The most commonly used techniques follow the hill-climbing approach where the slope of the function is determined in one or another way (also referred to as gradient-based optimisation). By following the slope in a specific direction (depending on whether the performance function is looking for a minimum or maximum), the optimisation algorithm can eventually get sufficiently close to the optimum.

To solve an optimisation problem, several steps are required and, based on criteria checks, the correct optimisation algorithm must be chosen for a specific problem. The first step is to mathematically formulate the objective function together with the applicable constraints. The objective function captures the goal for the optimisation exercise (for example minimise cost, maximise throughput, minimise energy usage, minimise waste, maximise yield, or maximise runtime). The constraints usually take the form of inequality and equality equations and contain the model of the process and limits on variables. After formulating the objective function and constraints, several criteria may be applied in order to determine which method to use for solving the optimisation problem. These include the nature of the objective function, the nature of the constraints, the number of variables and the nature of the independent

variables [22].

If the problem does not include inequality constraints and the equality constraints can be eliminated by solving them for the dependent variables and substituting them into the objective function, the problem is classified as unconstrained. This allows for the use of unconstrained optimisation techniques. If these criteria are not met, constrained optimisation methods must be used [22].

If the problem is unconstrained, the number of variables have to be considered. If a single independent variable is used, single variable optimisation techniques may be used (for example indirect methods, interpolation methods or region elimination). If the problem is multi-variable, two classes of solutions are mainly used: those that require the calculation of the function gradient and those that do not. The methods that do not require calculation of the slope use iterative experimentation and are called direct methods, for example Nelder-Meade (or sequential simplex) and Powell's method. The methods that require the calculation of the function gradient may use partial differential equations or numerical methods and are called indirect methods. These include steepest descent, conjugate gradient, Newton's method and quasi-Newton methods. These techniques are also used in constrained problems [22].

When dealing with a constrained optimisation problem, the approach is more general. Here, the linearity of the objective function and constraint functions have to be considered. If the objective function and the constraints are linear, the most common method is linear programming (LP). The simplex method for solving LP problems provides high computational efficiency and scalability [22].

If the objective function and/or the constraints are non-linear, non-linear optimisation techniques are required, generally referred to as non-linear programming (NLP). In the special case where the constraints are linear and the objective function is quadratic, quadratic programming (QP) may be used which provides high computation efficiency. In the more general case where the constraints are non-linear (the process is non-linear) and/or the objective function is non-linear (apart from quadratic), other non-linear programming techniques have to be used. Some techniques iteratively linearise the problem including successive linear programming (SLP) and generalised reduced gradient (GRG). Other methods that deal with the non-linearities more directly include successive quadratic programming (SQP) and penalty

functions with augmented Lagrangian [22, 48, 49, 50, 51].

Gradient-based optimisation techniques work on the assumption that the system is continuous (or smooth). When a system contains discrete input variables, discontinuities are introduced which violate the condition of differentiability that most of the gradient-based techniques require [49, 50].

When all the input variables are discrete, integer programming may be used (such as branch and bound or cutting plane techniques). When only some of the inputs are discrete, the system is called a mixed integer or hybrid system. For these problems, the system either has to be presented/transformed in such a way as to allow the use of modified gradient-based or integer programming algorithms, or an optimiser capable of handling hybrid problems directly must be used.

Furthermore, if the system is non-linear, the number of available solvers decreases further and the problem is denoted as mixed integer non-linear programming (MINLP). The applicable method will also depend on whether the system is convex or non-convex and whether the discrete and continuous variables are separable or not. Examples of algorithms for solving MINLP problems include outer approximation, generalised benders decomposition, branch and cut, and extended cutting plane, though these are limited to convex problems [52, 53]. Examples of methods for solving these systems directly include the genetic algorithm (see Section 2.4.1) and particle swarm optimisation.

Therefore, the optimisation algorithm to be used depends on the nature of the performance function and its constraints. The process model is evaluated as part of the performance function evaluation (usually captured in the equality and inequality constraints) and therefore the nature of the process model plays a vital role in selecting the correct optimisation algorithm [22].

There are endless applications that require the use of an optimisation algorithm. The most applicable to this study are parameter estimation for model development and operational optimisation (for example through the use of MPC, optimal control or RTO as discussed in Section 2.3).

Examples: In [15], linear programming is used for solving optimal control and MPC problems

for power flow optimisation of a photovoltaic-battery system supplying a load and connecting to the grid, using time-of-use tariffs and sell-back rates. In [16], a quadratic program is used to minimise a linear-quadratic cost function for energy minimisation of a commercial building heating system subject to comfort constraints. In [20], parameter estimation is performed off-line and on-line by using least square (LSQ) and recursive least square (RLSQ) techniques respectively. After identifying the model, several operational optimisation cases are solved using non-linear programming (NLP). In [19], a belt conveyor system in a colliery is optimised using binary integer programming. In [12], both real-valued (continuous) and binary (switching) optimisation techniques are employed in different simulation studies for optimal control. In [28], mixed integer particle swarm optimisation is used for optimal sizing and control of a pumped system. In [52], a Newton-type NLP algorithm is used to solve an non-linear model predictive control (NMPC) problem for controlling a chemical reaction process.

2.4.1 Genetic algorithm

Of special interest to this study is the Genetic Algorithm (GA), which is a type of Evolutionary Algorithm (EA). Evolutionary algorithms derive their name from the fact that they are based on the theories of natural selection and genetics. They mimic the processes of nature to arrive at a set of solutions believed to be stronger/fitter than their counterparts, based on certain criteria.

An evolutionary algorithm typically evaluates a set of candidate solutions (individuals) against a performance criteria (here called the fitness function). It then selects the fittest of the current set of individuals (the current population). This is referred to as the *selection* process. Thereafter, it generates a new set of individuals (the next generation) based mainly on *recombination* of these individuals (the fitter the individual, the more likely it is to propagate its genes to the next generation). A degree of *mutation* is also introduced to allow for some randomness. Some members of the previous generation may be retained or the entire population may be replaced by the newly generated population. This process is then repeated a number of times until some criteria is met, such as a maximum number of generations being exceeded or some tolerance being reached. The main steps in the algorithm therefore are evaluation, selection, recombination (or crossover), mutation and replacement

[21, 54, 55, 56].

The genetic algorithm is one form of evolutionary algorithm. The vector of optimisation parameters represents an individual (also called a genome with the individual parameters referred to as genes). The algorithm in its canonical form typically uses a bit-string representation of the optimisation parameters and therefore discretises the parameters. The number of bits allocated for the discretisation influences the resolution of the optimisation problem. For continuous variables, the variables are simply discretised to the defined resolution. For discrete variables, this may result in a portion of the bit-string not being used and therefore some string combinations not being valid. This is one aspect that must be addressed in the specific formulation of the algorithm [57]. If the individuals are encoded using real numbers, the solution is called a real coded genetic algorithm (RCGA) [58].

After evaluating the individuals in the current population and determining the relative fitness of each individual with respect to the other individuals in the population, the selection process is carried out to form an intermediate population. The probability that an individual is selected for (or duplicated to) the intermediate population is typically proportional to its fitness. Several selection methods exist though most are based on some form of stochastic sampling where the probability of selection is biased towards the fitter individuals [57]. A fraction of the individuals in the current population with lower fitness values may also be designated as elite to be transferred directly to the new population.

After the selection process, recombination (or crossover) and mutation is applied to the intermediate population to create the next population (or generation). Recombination is achieved by swapping fragments of the bit-strings between individuals, thereby generating offspring with genes that are related to that of the parents. Randomly paired individuals in the intermediate population are recombined (or crossed over) based on a crossover probability. Therefore, some pairs will be recombined to form new individuals for the next generation whereas others will simply move to the next generation unaltered. If a pair is selected for recombination, a randomly selected single recombination point designates the crossover point and the bit-string values between the two individuals in the pair are swapped around this point [57].

After recombination, mutation may be applied to all bits in the population by flipping bits

based on the mutation probability (which is usually set very low). The degree of recombination is typically set much higher than that of mutation. In some cases, the mutation is performed by generating a new random bit (which means that the existing bit may remain unchanged even though it was selected for mutation), whereas in others the existing bit is flipped. After mutation and recombination is complete, the process is repeated with the newly created population. The algorithm normally replaces the entire current population with the new one at each iteration [21, 57].

Constraint handling is typically performed by incorporating penalty functions and/or by using selection, recombination and mutation algorithms that only result in feasible points [59].

Some advantages of genetic algorithms include [21, 57]:

- Very little initial state data is required.
- GAs are capable of handling different types of optimisation variables in the same problem formulation (such as discrete and continuous variables).
- Do not require gradient information.
- Constraints are easily incorporated.
- Due to the stochastic nature, GAs are also less likely to get stuck at local optima (more likely to converge on the global optimum).
- Can be very robust if set up correctly.

These advantages make GAs especially attractive for problems that

- are difficult to formulate mathematically,
- have strong non-linearities,
- have interactions between variables,
- contain discontinuities (such as absolute value extractions, maximum/minimum oper-

ators or discrete states) and are non-differentiable,

- are constrained,
- are mixed-integer (or hybrid),
- contain multiple local optima (non-convex),
- have uncertainty regarding the initial starting point,
- are time-variant and/or
- contain randomness or noise.

Therefore, the genetic algorithm is extremely versatile in that it can be used to solve problems from single variable unconstrained optimisation problems through to constrained mixed-integer non-linear optimisation, with minimal modification.

Some of the disadvantages are that these algorithms are very computationally intensive (which is less of a concern with modern computing power) and that optimality cannot be guaranteed. Therefore, on-line applications using evolutionary algorithms have been limited [21]. The stochastic nature also implies that the solutions are not exactly repeatable which may be a problem in some applications.

In this study, the genetic algorithm is used both for parameter estimation (Section 3.3.3) and control (Section 3.4.3).

2.5 SYSTEM MODELLING

Models are used to replicate/mimic the behaviour of a system. Two main categories are commonly considered, namely steady-state models, which do not change with time, and dynamic models concerning the transient behaviour of a system. Furthermore, models can be derived from first principles (using the laws of nature) or empirically (using behavioural data) or a combination of the two. Models are further classified as being linear or non-linear, time variant or invariant, discrete or continuous, parametric or non-parametric, and statistical or deterministic [60]. Therefore, any specific model may have many classifications.

One common use of system models is in process simulation. In most cases, researchers in the academic world do not have access to real processes (or access is limited). Therefore, models play an important role in system simulations. Simulations may also be used to test control schemes before implementation on the actual process. Another function of system models is controller design, especially for model based control algorithms such as optimal control and MPC.

Based on the sections below, there are limitless variations of models. Therefore, establishing the purpose and required accuracy of the model, before starting with the modelling, is important. Trade-offs exist between model complexity, accuracy, cost, time for development and versatility. In many cases a complex, first principle, non-linear model is used for process simulation with a reduced linear model used for controller design.

For the purposes of this study, deterministic process and energy models are the main focus points. The process model describes the behaviour and limitations of the process as a whole, whereas energy modelling aims at describing the overall energy consumption of the system. Some energy models are simple (such as the on/off status of an electric heater) whereas others may be more complex (such as the energy consumption of a multi-stage centrifugal compressor, where several factors influence the consumption, including guide vane setting, gas temperature, gas composition and flow rate.).

2.5.1 Dynamic and steady-state models

Dynamic models are aimed at capturing the behaviour of a system as it transitions from one state to another. This change in state might be caused intentionally by changing the range or set-point of a controlled variable or it may be caused by a disturbance acting on the system. These models indicate how a variable changes over time when exposed to a change in the system. There are various ways to capture this relationship, from an array of values (non-parametric), to explicit equations (parametric or first principle, usually representable in the form of differential equations with regard to time). As most petrochemical processes can be approximated fairly accurately with low order linear step response models, this is the type of model typically used for MPC and optimal control applications in this industry.

Steady-state models are formulated under the assumption that the dynamics of the system

may be omitted. These models are not concerned with the evolution of the system over time and assume that the system is balanced. Steady-state models are typically used in plant design, using mass and energy balances. Therefore, these models are also ideal for use in an RTO application. In its simplest form, a steady-state equation may be a constant gain value (or matrix) or a set of algebraic equations.

2.5.2 Linear and non-linear models

Linearity in a model refers mainly to two characteristic properties namely homogeneity and the principle of superposition. Homogeneity means that scaling the input to the system, scales the output with the same factor. The principle of superposition implies that a scalar combination of inputs, results in a scalar combination of outputs [22, 32, 34, 38]. When a system does not behave in this way it is said to be non-linear.

Although most real systems exhibit non-linear behaviour, it is in many cases reasonable to approximate the system as being linear or to linearise the system around an operating point, which allows for the use of linear system techniques for behavioural analysis and controller development, which is beneficial from a complexity point of view [33, 34]. Combining linearity with time invariance (discussed next) result in the linear time-invariant (LTI) class of models for which a wealth of mature theory exists.

2.5.3 Time variant and invariant models

Systems with repeatable behaviour at different points in time are referred to as time invariant. If a system's behaviour changes with time, it is said to be time variant. In other words, for a time invariant process, if an input is applied now, the behaviour will be the same as when the input is applied at any time in the future.

2.5.4 First principle and empirical models

Models may be derived from known physical principles like mass and energy balances, chemical reaction theory, and laws of physics, when the behaviour of the system is well understood and transparent (also referred to as physical models). In some cases the internal details of a system are unknown or the complexity of the system does not lend itself to reasonably de-

veloping a first principle model. In such cases, the model may be derived using experimental or historical data referred to as empirical modelling [22, 32, 38].

2.5.5 Parametric and non-parametric models

In some cases the type of behaviour that can be expected of a system is known (the form of the model) even though the exact behaviour depends on specifying (or fitting) some parameters. These parameters may be determined by consulting design data or physical dimensions of the system or be determined experimentally. In other cases the behaviour is unknown and empirical tests of the system are performed to gather sets of data describing the response of the system as a result of a known input to the system like a step or impulse. The former is referred to as parametric and the latter as non-parametric models.

2.5.6 Model examples

In [6], a first principle dynamic non-linear model of a fuel gas blending process is used for system simulation; a linear dynamic model with gain scheduling is used for MPC control and a non-linear steady-state model is used for RTO. Similarly, in [13], a non-linear dynamic model of an ore milling process is used for simulation whereas a linear dynamic model is used for MPC. The gain matrix of the linear dynamic model is then used for RTO. In [16], a linear parametric model is developed to describe the thermal capacity of a commercial building, which is then used in an MPC application and tested on a real building. In [20], an energy model of a belt conveyor system is developed from a combination of sources that is geared specifically for energy optimisation. The result is a four parameter parametric model where the parameter values may be determined from design data or empirically. A review of existing energy modelling techniques for belt conveyors is also presented.

2.5.7 Energy modelling

The electrical energy consuming elements in a system can fall either in the discrete (such as a pump, fan or heater) or continuous category (such as a pump or fan fitted with a variable speed drive or a heater with a variable power setting). The energy model can therefore contain a combination of these elements. Each piece of equipment has a status and a power function. The status can be on or off and can be represented by

$$u(t) = \begin{cases} 1 & \text{for component ON,} \\ 0 & \text{for component OFF.} \end{cases} \quad (2.1)$$

The power consumption (kW) of a component can then be denoted as

$$W(t) = u(t)w(t) \quad (2.2)$$

where $w(t)$ depends on the equipment design and whether it is discrete or continuous. In the simplest case, $w(t)$ will be constant such as the rated power of a simple water heater. In a more complex case, $w(t)$ may be representing the power curve of a centrifugal compressor where the power output depends on several factors such as flow, gas temperature and compression ratio.

Equation 2.2 may then be used to formulate functions for the total power consumption ($W_T(t)$ in kW) and cost ($C_T(t)$ in \$/h) for a system of N components, denoted as

$$W_T(t) = \sum_{i=1}^N W_i(t) \quad (2.3)$$

and

$$C_T(t) = \sum_{i=1}^N p(t)W_i(t) \quad (2.4)$$

where $p(t)$ denotes the rate in \$/kWh.

The cost function to be minimised for control and/or optimisation can then be formulated for energy cost and/or consumption as

$$J = \int_{t_0}^{t_f} (W_T(t) + C_T(t))dt. \quad (2.5)$$

Depending on the formulation of $w(t)$ for each component, Equation 2.5 can be used for

dynamic control through MPC or at a higher level in RTO. If the energy component dynamics are captured and the goal is to have dynamic control, using the running signals of the components as control handles, then Equation 2.5 can be combined with the dynamic process model into an MPC solution. If the objective is to leave dynamic process control to the MPC and only optimise on energy in the longer term, Equation 2.5 could be used for RTO only. MPC and RTO can also be combined so that the MPC includes the dynamic $w_i(t)$ functions and the RTO only uses steady-state power consumption calculations for long term energy optimisation by writing set-points to the MPC with regards to process variables that will allow the MPC to switch energy components on/off, taking care of the dynamic fluctuations introduced by the switching.

2.6 HYBRID SYSTEMS

As mentioned, all process units require energy to be added and removed and material to be moved to, from, and within the system. Adding and removing energy are typically achieved through heat exchangers and moving of material through pumps, fans, compressors and conveyors. Heat exchangers are continuous energy devices where a continuous range of energy values are added or removed, typically by adjusting the flow rate through one or both of the streams using control valves. Pumps, fans and conveyors are usually discrete energy devices with on and off states or a range of discrete speed settings. Models containing both discrete and continuous input variables are generally referred to as hybrid systems. The presence of both continuous and discrete components complicates modelling, control and optimisation [25].

The behaviour of most continuous devices such as valves and heat exchangers is well understood and can easily be demonstrated with a simple material flow calculation or duty calculation (assuming that temperature and flow measurements are available and the characteristics of the media are known). Modelling of the discrete devices is also fairly straightforward. A simplifying assumption is to treat the device as having a zero duty when off and a fixed duty when on. This will not suffice for this study as a more accurate result is desired. Therefore, the performance of the device at different operating points will be considered. For a pump, this can be achieved by considering duty equations or the pump energy curves (described in more detail in Section 2.7.1). To complicate matters, a device may be equipped

with a power modulating device such as a variable speed drive (VSD). This means that the device has both an on/off status and, when on, an adjustable continuous energy component. Therefore, a single device may have two handles, one discrete and one continuous.

A conventional way of dealing with hybrid systems is to approach it as two distinct layers, one for the discrete components and another for the continuous components. The bottom layer, for example, may be concerned with the continuous process and the top layer with the discrete part. In this way well-known techniques for continuous and discrete optimisation may be utilised for the different layers [12, 18, 19, 20, 61].

In some cases, however, it is not easy to distinguish between the layers due to the level of integration between the discrete and continuous components. Therefore, the need arose for a systematic method of combining these discrete and continuous components for modelling and control design. For example, stopping a pump might have a favourable effect on the energy consumption model, but will affect the flow rate of the pumped stream, which could have a degenerate effect on the plant and its economics. Therefore, it is important to be able to consider both aspects in the optimisation problem simultaneously.

In [25], a method is presented in which the hybrid process is described in terms of linear dynamic equations subject to linear inequalities, that involve continuous and discrete variables (denoted as Mixed Logical Dynamic (MLD) systems). The inequalities are obtained through manipulation of combinational logic (describing the process component relationships) and mixed integer quadratic programming (MIQP) is used as the optimisation engine. These linear inequalities, together with the continuous process model, are then used to formulate a model predictive control (MPC) strategy. In other examples, the hybrid system is treated in a more direct manner through the use of an optimiser capable of directly solving mixed-integer problems with discontinuities [28].

Several dependencies and constraints may be formulated for the combinations of components as described in [12] and [25] such as logical correlations, mass and energy balances, boundary conditions, and process constraints.

2.7 PROCESS EQUIPMENT

Energy consumption in process plants occur through several mechanisms. Examples of direct electrical energy use include pumps, compressors and heaters. Indirect use through heat transfer, using steam, cooling water or hot oil, is another common technique and is achieved with heat exchangers.

Regardless of the technique, energy is required to move material and to heat or cool it. Therefore, much of the energy consumption for a plant occurs through the use of utilities such as cooling water, refrigerated water, steam and fuel gas. Apart from the actual thermal energy consumed, electrical energy is required to transport the utility media through a plant by using pumps, fans, compressors and blowers. The energy consumed in these devices can be modelled based on equipment design and operating conditions.

Some of the common equipment types applicable to this study are discussed next, presenting only a sample of the landscape of process equipment.

2.7.1 Pumps

There are several types of pumps such as canned pumps, screw pumps, centrifugal pumps and positive displacement pumps. Of these, centrifugal pumps are most widely used in the process industry and will subsequently be used in this study. Centrifugal pumps are characterised by pump performance curves that indicate the head-versus-flow and power-versus-flow of a particular pump. For the pumps used in this study, these curves are shown in Figures 2.2 and 2.3. The operating point of the pump is at the intercept of the system curve and the pump curve. Adding a variable speed drive (VSD) to a pump means that multiple pump curves will exist for the different speeds.

The power consumption of a centrifugal pump (assuming it is in a running state) can be determined by duty equations or pump power curves. For example, with centrifugal pumps, the power requirement for a pump can be calculated by

$$w = \frac{\rho g H f}{\eta} \quad (2.6)$$

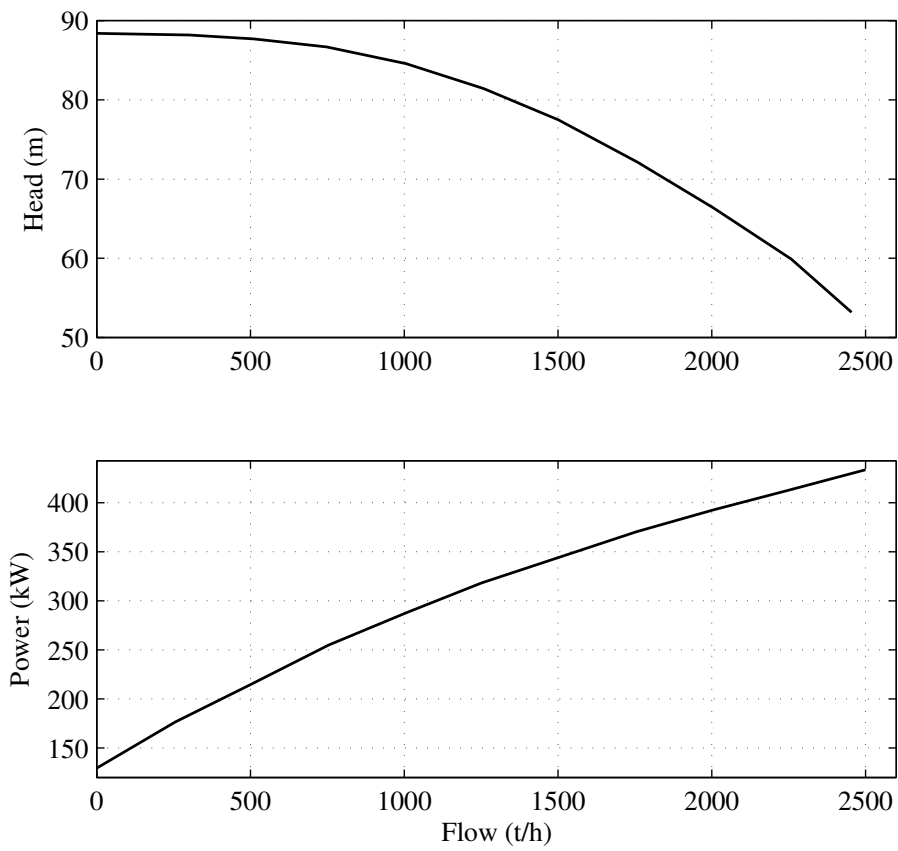


Figure 2.2: Tempered water pump performance curves.

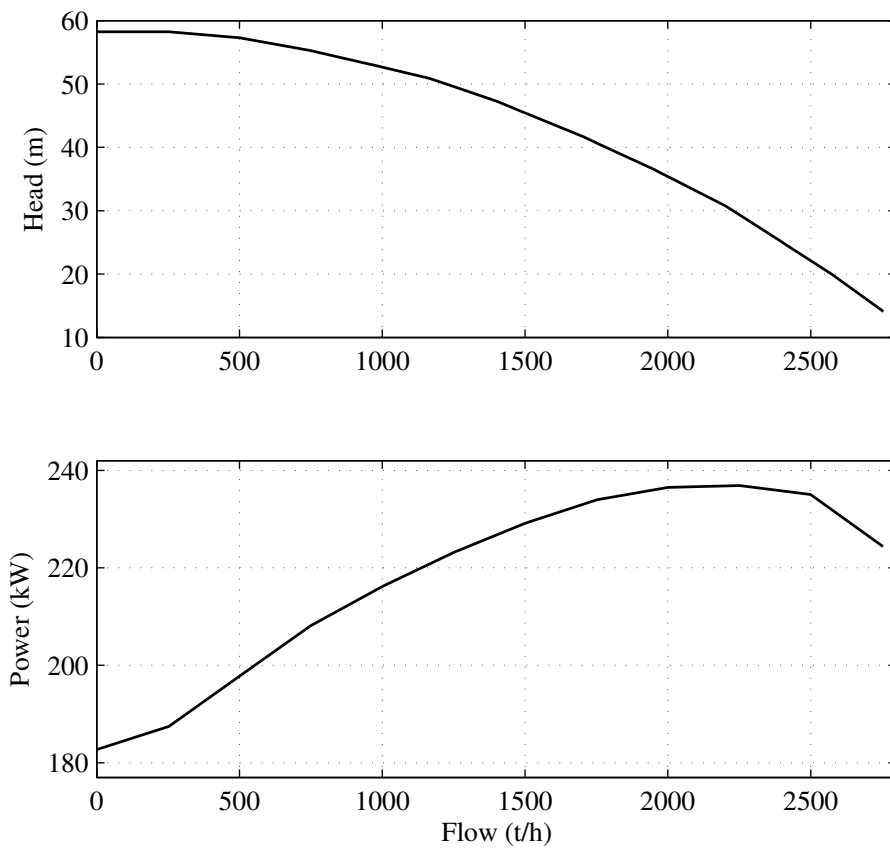


Figure 2.3: Cooling water pump performance curves.

where w is the pump power (kW), ρ is the specific gravity of the fluid (t/m^3), g is the gravitational constant, H is the head of liquid (i.e. potential energy) added by the pump (m), f is the flow rate (m^3/s), and η is the efficiency of the pump [28, 62].

As the flow rate increases, the head drops. The optimal point of operation is where the highest flow per unit of energy is obtained. Where energy curves are available, the data can be captured in lookup tables or fitted with polynomial functions to estimate the power consumption based on the flow rate.

2.7.1.1 Pumps in series

If a larger total head is required, pumps may be installed in series. When pumps are running in series, the total flow rate is carried by each pump. The additional pump will increase the achievable head, which will mean a higher flow rate than that achieved using a single pump, or the additional head might be required to simply overcome the static head of the system (such as pumping over a hill) [63].

Figure 2.4 illustrates the scenario of two identical pumps in series on a system with a constant system curve. The operating point, when running a single pump, is at A; when running the two pumps in series, the operating point moves to B on the combined pump curve. C shows how the operating point of the first pump changes when the second pump is running. Therefore, the increase in discharge pressure allows more flow to be pushed through the system than what is achievable with a single pump. Care must be taken not to increase the flow rate to beyond the rated flow rate of the pumps.

2.7.1.2 Pumps in parallel

For larger flow rates than those achievable with a single pump, pumps may be installed in parallel. When pumps are running in parallel, the total flow is distributed between the pumps. If the system has a constant flow coefficient (for example when no throttling element like a discharge valve is used), the slightly higher discharge pressure delivered by the combined pumps will result in a marginally higher flow rate than what would have been achieved using a single pump (assuming that the process flow coefficient is not so high that a single pump will run past its maximum capacity) [63].

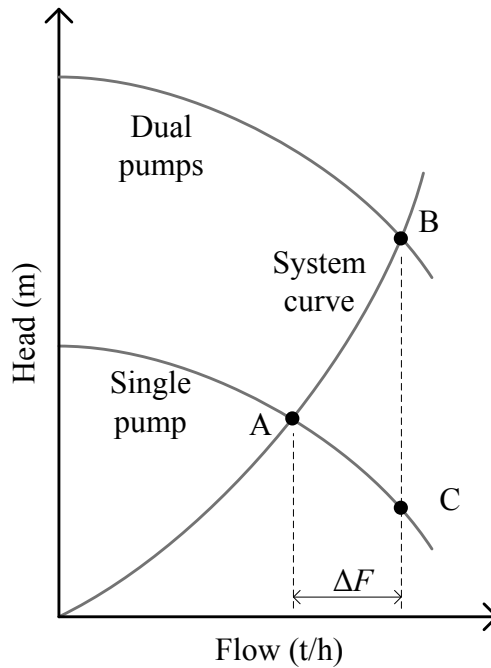


Figure 2.4: Performance illustration of pumps running in series on a system with a constant system curve.

Figure 2.5 shows an illustration of the operation of a single versus two identical parallel pumps on a system with a constant system curve. The operating point when running a single pump is at A. When running the pumps in parallel, the operating point moves to B on the combined pump curve. C shows the operating point of the individual pumps when running in parallel. Therefore, the progressive increase in discharge pressure (and the resulting increase in flow rate) becomes less with each additional pump that is added in parallel.

2.7.1.3 Pump efficiency

The pump efficiency is a measure of how effectively a pump can transfer its input energy (by motor or turbine) into the moving of liquid. Several components affect the overall efficiency (or the hydraulic efficiency denoted as $\eta_{hydraulic}$). These include the pump efficiency (η_{pump}), motor efficiency (η_{motor}) and coupling efficiency ($\eta_{coupling}$). The hydraulic efficiency is equivalent to η in Equation 2.6 and is calculated as

$$\eta_{hydraulic} = \eta_{pump}\eta_{motor}\eta_{coupling} \tag{2.7}$$

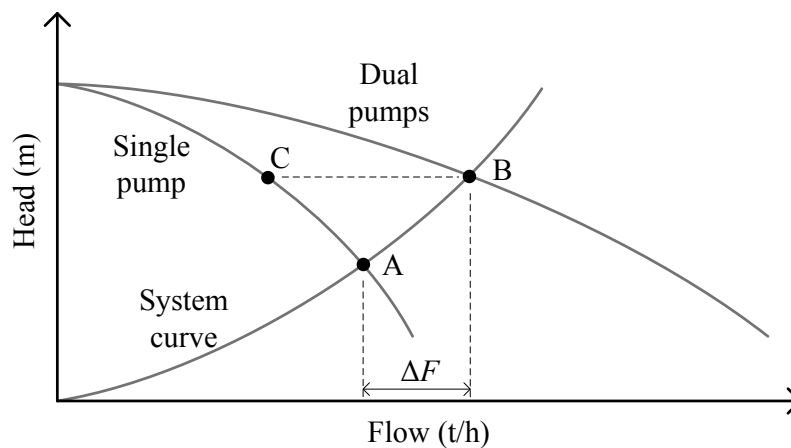


Figure 2.5: Performance illustration of pumps running in parallel on a system with a constant system curve.

where η_{pump} , η_{motor} and $\eta_{coupling}$ depend on design, equipment condition and operating point. When a pump is running at the optimal efficiency point, it is not only beneficial from an energy efficiency point of view, but also from an equipment maintenance perspective. The optimal efficiency point is at a specific flow rate for a constant impeller size or rotation speed.

The operation point (flow rate) is determined by the system curve (as discussed in Sections 2.7.1.1 and 2.7.1.2). Therefore, if the system design is such that the pump does not operate at acceptable efficiency, either the characteristic of the system has to be changed or the pump impeller or rotation speed has to be adjusted. Changing of the system characteristic may, for example, be achieved by using a control valve (though this introduces its own energy efficiency degradation). If the operating point is anticipated to be fixed, an impeller change can be considered. If not, the use of a variable speed drive (VSD) is an option (discussed in the next section). In accordance with centrifugal pump affinity laws, the relationship between impeller diameter/pump speed and the flow rate is linear whereas the power consumption has a cubed relationship. Therefore, if 25% less flow is required, the resulting effect on power will be a reduction of approximately 60% [62].

2.7.1.4 Variable speed drives

The pump curve shown in Figure 2.2 is based on a specific impeller size and pump rotation speed. The curve may be adjusted either by changing the impeller or by varying the rotation speed of the pump. The former requires a physical change of the pump internals and is a more permanent change to the pump capacity. The latter is a way of manipulating the pump capacity in real-time. This may be achieved by using a gearbox or a variable speed drive (VSDs). VSDs are more commonly used and the cost and size have drastically dropped in recent years. Typically a VSD changes the pump speed by manipulating the frequency of the alternating current (AC) power supply on an AC motor. Slowing down the motor speed (and therefore the pump rotation speed) has the same effect as reducing the impeller size.

2.7.2 Fans

Similar to pumps, fans are most commonly driven by electrical motors. Fans may be used in-line to transfer gaseous material through the process or, for example, be used to remove energy from a process stream by pulling/pushing air through a radiator assembly, through which the process stream is flowing. The fans used in this study are those situated on top of cooling towers and operate at a fixed speed. The fans provide an air draft through the cooling towers through which the warm cooling water return stream is sprayed.

2.7.3 Heat exchangers

A common requirement in the process industry is the ability to transfer thermal energy from one process stream to another. The most common way of achieving this is by using heat exchangers. The science behind heat exchange is well understood and, depending on the accuracy required, can be fairly easily modelled, using well-known heat exchange equations [22].

Heat exchanger efficiency is higher when the differential temperature between the two energy streams are higher. To obtain the higher differential temperature, however, might require more energy on the utility side, which contradicts the improved efficiency on the exchanger. Therefore, once again, an optimal solution must be found. Sometimes the cooling/heating medium flow can be adjusted using automatic control. In other cases, manual adjustment of

a hand valve may be required.

Several studies describe the optimisation of heat exchanger networks [64, 65].

2.7.4 Control valves

Control valves are flow modulating devices that are equipped with actuators and positioning devices (such as transducers or positioners), which allow them to be manipulated by a control signal. This allows for the flow rate of a process medium to be manipulated, which in turn provides the means of manipulating other related variables such as pressures, temperatures and levels.

Control valves are designed according to the required flow rates at different operating points and the required installed flow characteristic. Many different types of control valves exist, including ball, globe, butterfly (disk), rotating globe and gate. Three main classes of valve characteristics are found, namely linear, equal percentage and quick opening. Depending on the service and the type of valve, the valve characteristic is usually aimed at creating a linear installed characteristic.

If, for example, a valve is situated between two processes with constant pressures (such as two headers), a linear valve characteristic will cause a linear installed characteristic. On the other hand, if the valve is fed by a pump, with a pump curve similar to that shown in Figure 2.2, and discharging to a system of more constant pressure, the discharge pressure of the pump, and consequently the differential pressure across the valve, will drop as the flow rate increases. This means that the higher the flow rate, the smaller the effect of a valve opening change will become. Installing an equal percentage characteristic valve in this case, will compensate for this phenomenon by allowing smaller flow changes per valve opening change at the lower range of the valve opening, and larger changes at the upper range of valve opening. The size of the valve in terms of achievable flow rate is specified as a flow coefficient, C_v [32]. The basic flow equations is

$$f = C_v q(OP) \sqrt{\frac{\Delta P}{\rho}} \quad (2.8)$$

where f is the flow rate; $q(OP)$ is a function describing the flow characteristic based on

the valve opening (OP – lift or rotation as a fraction of total movement range); ΔP is the differential pressure across the valve and ρ is the specific gravity of the medium flowing through the valve. The characteristic equations for the three types are

$$q(OP) = OP \quad (2.9)$$

for linear valves,

$$q(OP) = R^{(1-OP)} \quad (2.10)$$

for equal percentage (where R is a design parameter determining the extent of the non-linearity of the characteristic) and

$$q(OP) = \sqrt{OP} \quad (2.11)$$

for quick opening valves.

An illustration of these flow characteristics is shown in Figure 2.6 where the flow versus valve opening is shown for a constant differential pressure over the valve. The solid line represents the linear characteristic, the dashed line that of a modified equal percentage valve and the dotted line that of a quick opening valve.

2.8 COOLING WATER SYSTEMS

As mentioned, various utilities are used in the process industry, including steam, hydrogen, nitrogen, air, refrigerated water, fuel oil, fuel gas and cooling water. Some of these may be imported from external sources and some generated on-site.

An example of a common utility system (which is also typically a hybrid system) is a cooling water system consisting of several pumps, control valves, heat exchanges and cooling towers (some equipped with fans, others relying on natural draft).

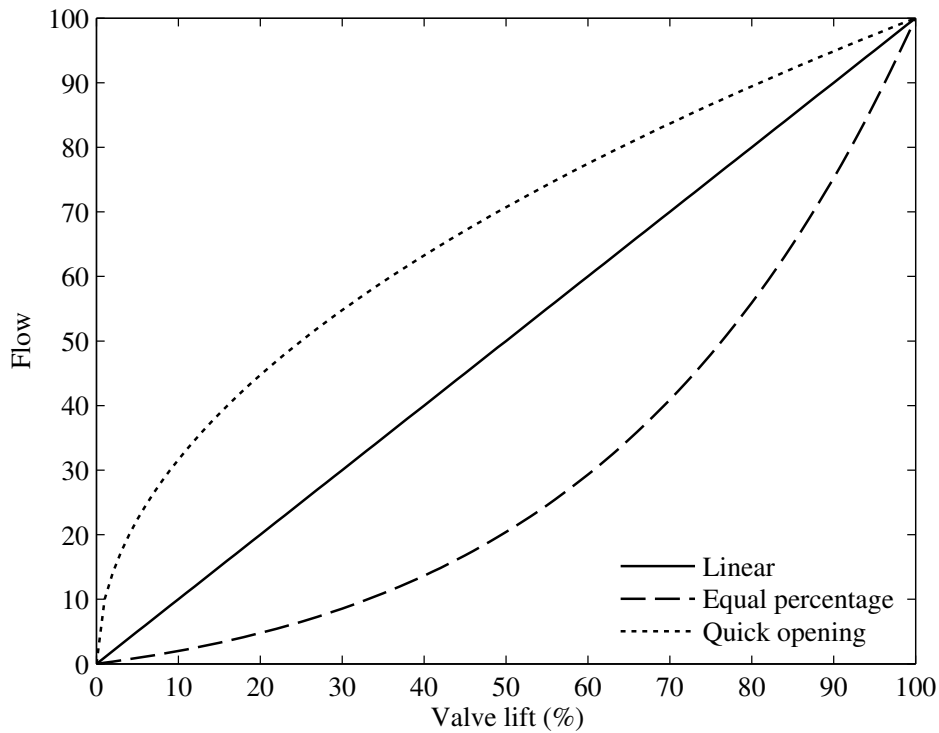


Figure 2.6: Valve characteristic curves with constant differential pressure across valve.

There are two categories namely open-loop and closed-loop. In open-loop systems, the water is pumped from a source like a river and, after heat is absorbed from the process heat exchanger network, the heated water is discharged back into the water source. Closed-loop systems recirculate the heated water to cooling towers where the heat is removed again. Air-cooled systems are also used in some cases where the water circuit is completely closed and the heat is removed from the water by means of a radiator (these are not that common in the process industry due to the amount of energy to be removed).

Detailed modelling of the cooling water process, such as is typically found in relevant literature, can become very complicated; therefore a simplified model is required for the purposes of basic simulation and controller/optimiser design.

The cooling water system considered in this study is a dual circuit induced draft system. The process description and system modelling is discussed in more detail in Chapter 3.

2.8.1 Cooling tower theory

Several types of cooling towers are found in industry and depend mainly on the amount of cooling required and the availability of water in the region. The most common type of cooling, when water is sufficiently available, is evaporative cooling, where the bulk of the cooling occurs through the partial evaporation of the cooling water as it comes into contact with an air draft [65]. Warm water is trickled or sprayed through the tower and cool air is pushed or pulled through the water. This causes some of the water to evaporate, which cools the remaining liquid portion through the extraction of the latent heat of vaporisation of the evaporated portion. The latent heat transfer is responsible for approximately 80% of the heat transfer, with the balance occurring through the sensible heat transfer [22, 62]. This loss of a portion of the water to atmosphere requires make-up water to be added to the system for replenishment.

The draft may be natural or mechanically induced and it may be cross-flow or counter-flow depending on the relative flow direction of the air versus the water [21]. The category of mechanical draft is further classified into forced draft or induced draft, depending on whether the fan pushes air through the tower from the air inlet, or whether it pulls air through the tower from the air outlet.

The efficiency of the tower is greatly affected by the water absorbing capacity of the air, which is gauged by the wet-bulb temperature of the air, which is affected by the relative humidity and the air temperature. The more water the air can absorb, the more latent heat transfer can occur, and the higher the cooling efficiency of the tower. Therefore, in humid conditions, the cooling capacity of a tower is less than in dryer conditions (under the same temperature and pressure). Furthermore, the contact time between the water and the air plays an important role. The taller the tower, the longer the contact time, the better the cooling efficiency.

The cooled water is pumped to the heat exchangers where heat is absorbed from the process (or in the case of this study the tempered water circuit which in turn absorbs heat from the process heat exchanger network).

One widely accepted theory for heat transfer in cooling towers is known as the Merkel equa-

tion, which proposes the enthalpy difference between the air and an assumed film of air around each water particle to be the main driving force for the cooling [22, 23]. Therefore, the heat transfer is viewed in terms of exchange between the air around the water particle and the air flowing through the tower. The air film around the water is at the same temperature as the water and saturated, whereas the air flowing through the tower is cooler and not saturated, which allows for the flow of energy from the water to the air flowing through the tower.

2.9 CONCLUSION

This chapter covered the necessary literature to lay the foundation for the rest of the study. In the next chapter, the specific cooling water system considered in this study is described and modelled. This is followed by a case study on the application of advanced process control and optimisation techniques, with focus on energy optimisation.

CHAPTER 3

APPROACH AND METHODS

3.1 CHAPTER OVERVIEW

This chapter consists of two parts. In the first part, a process description of a cooling water system is provided followed by a detailed discussion on the derivation of a model of the system. The model is then presented in a simplified discrete state-space form, as a collection of difference and algebraic equations, and dynamics are added to the model. Thereafter, the model is verified using plant data, and parameter estimation is performed using a genetic algorithm.

In the second part, different control and optimisation techniques are designed and several simulation studies are performed to determine the performance of the control schemes compared to that of the base case. These include an Advanced Regulatory Control (ARC) scheme, a Hybrid Non-linear Model Predictive Control (HNMPC) scheme, and an Economic HNMPC scheme.

3.2 PROCESS DESCRIPTION

Figure 3.1 shows an example of a cooling water system that uses two separate water circuits. The first one is called the tempered water (TW) loop, containing the water supplied to the plant heat exchanger network and is a closed, treated water loop. This circuit is equipped with a bank of pumps. The second circuit is the cooling water (CW) loop. This loop could contain untreated water (apart from standard dosing and filtering) from a water source such as a river or dam and is equipped with its own bank of pumps as well as the cooling towers

(CTs, in this case induced draft counter flow towers equipped with fans). Between the two water circuits is a bank of plate heat exchangers to transfer the heat collected by the tempered water through the plant heat exchangers to the cooling water circuit where it can be removed via the cooling towers.

The tempered water circuit is equipped with a temperature control valve (TV) that bypasses the tempered water side of the plate heat exchangers in order to provide a handle for temperature control of the tempered water supply temperature (T_{TWS}) to counter for a case where too much cooling is provided (such as during plant load reduction or rain spells).

Each of the pumps on the cooling water circuit is equipped with a discharge pressure control valve to protect against pump damage, by throttling back when the discharge pressure of a pump drops too low (corresponding to a high flow rate). Omitted from Figure 3.1 are side stream filters and a make-up water line, that links to the cooling water circuit for removing debris and topping up cooling water.

From a control perspective, the typical controlled variables are the tempered water supply temperature (T_{TWS}); the tempered water differential temperature (ΔT_{TW} , which is related to the plant duty and the tempered water flow rate); the total power consumption of the pumps and fans (W_T); and the electrical energy cost of the system (C_T).

Common disturbances that may act on the system include ambient temperature (T_a) fluctuations, changes in relative humidity (RH), equipment failures and load changes affecting the duty from the plant (Q_P).

The handles available to manipulate the plant are the valves and the on/off signals of the pumps and fans. In this example there are no variable speed drives (VSDs) on the pump or fan motors.

Regarding the pump banks and cooling towers, the system presents a large number of optimisation possibilities.

- Individual pumps may be optimised with regard to the flow rate through it (if equipped with a variable speed drive or flow control valve) by using the pump design data and empirical data to determine the most efficient operating point.

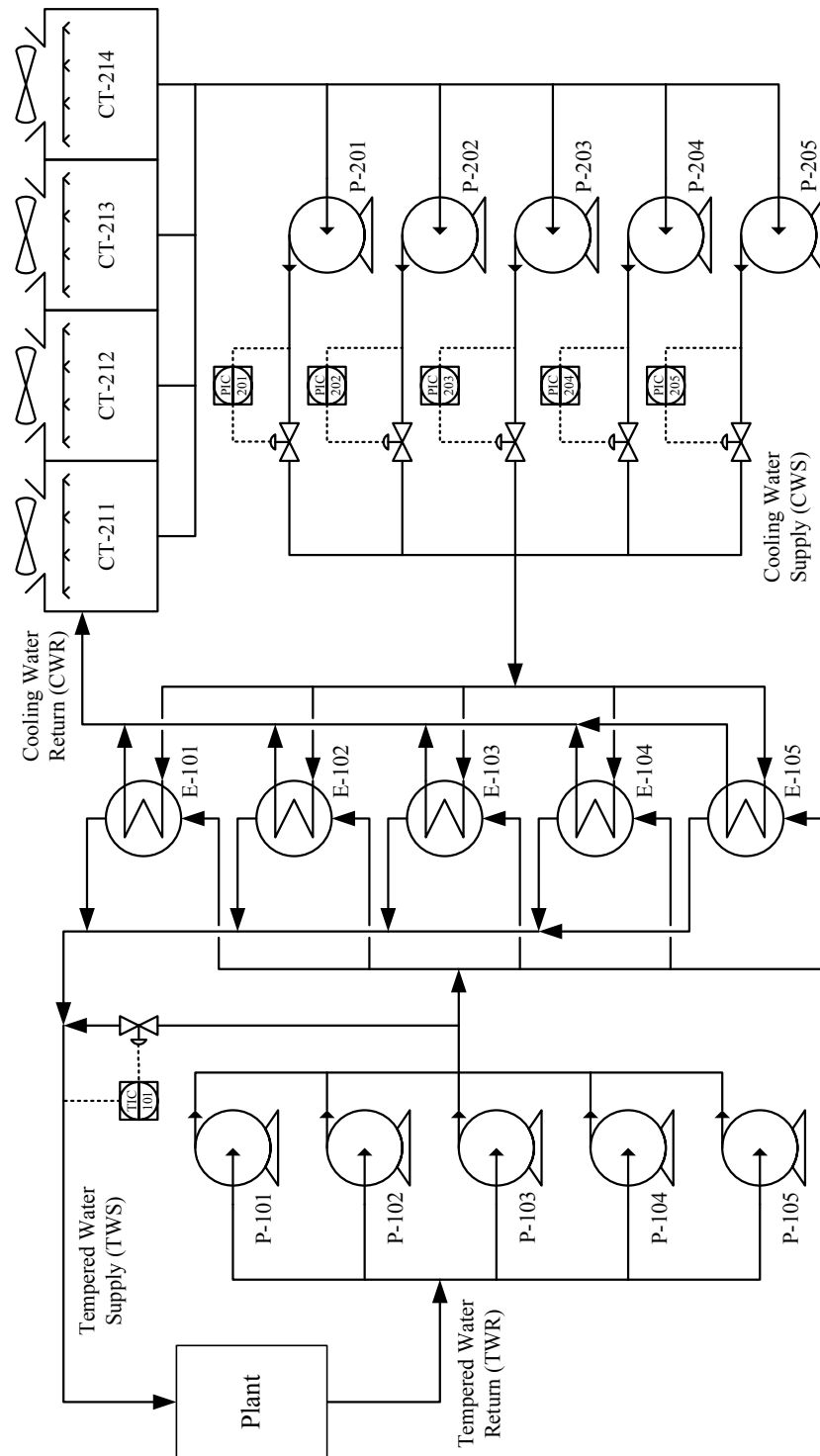


Figure 3.1: Dual circuit induced draft counter-flow cooling water system.

- A bank of pumps may be optimised with regard to the number of pumps running at any time.
- The water circuits may be optimised with regard to the required flow rate and temperature of the water.
- Each cooling tower can be optimised through the speed of the fan (if equipped with a variable speed drive) and flow of water through it.
- The cooling tower bank may be optimised with regard to the number of fans running and the temperature of the water exiting the towers.

The extent of the optimisation is limited through the following set of constraints:

- The number of tempered water pumps running (depicted as P-101 to P-105 in Figure 3.1) may not be less than two.
- The number of cooling water pumps running (depicted as P-201 to P-205 in Figure 3.1) may not be less than two.
- The number of cooling towers running (depicted as C-211 to C-214 in Figure 3.1) may not be less than one.
- The supply temperature of the tempered water to the plant (T_{TWS}) should not go below 26 °C or above 36 °C.
- The tempered water differential temperature ($\Delta T_{TW} = T_{TWR} - T_{TWS}$) should not exceed 6 °C.

The constraints on the pumps are for preventing pump damage by pushing the pumps past their design capacity. The lower limit on the tempered water supply temperature is to prevent operational difficulties on some equipment (like the refrigeration compressor system that requires a tempered water temperature at design) and, in the case of this specific plant, to prevent solidification of the product at low temperatures. The limit on the differential temperature of the tempered water is to ensure that the return temperature of the tempered water does not approach the process temperatures of the streams flowing through the plant

heat exchanger network, which would reduce heat transfer efficiency (or stop heat transfer). On the other hand, the lower the flow rate, the larger the temperature difference will be at a given duty, which would mean that fewer pumps are needed. Increasing the cooling water temperature (by reducing the number of cooling towers in operation or the number of cooling water pumps) will reduce power consumption, but will also cause higher tempered water supply temperature. If the users are temperature controlled, an increase in tempered water supply temperature will cause more flow demand, thereby countering the benefit of running less equipment on the cooling water side. Therefore, an optimal balance between flow rate and water temperature has to be determined.

3.3 SYSTEM MODELLING

This section details the development of a model of the cooling water system shown in Figure 3.1 as well as the validation of the model using real plant data [66]. The duty of the plant heat exchanger network varies with load and ambient conditions. Furthermore, the achievable supply temperature of the cooling water also varies with ambient conditions; therefore the efficiency of the cooling towers is not constant. One of the modelling assumptions is that pumps on both circuits and the cooling tower fans can be started and stopped at will without requiring manual isolation.

Figure 3.2 shows a simplified illustration of the process in Figure 3.1 where the system has been reduced to single pieces of equipment and the key variables are clearly indicated to aid in the derivation of the model equations. This simplified diagram is especially useful in the derivation of the flow and duty calculations described in Sections 3.3.1.2 and 3.3.1.3. The complete model is then assembled by combining the calculations of the individual devices.

The existing continuous control components include a temperature PID controller that manipulates the heat exchanger bypass on the tempered water side as well as the cooling water pump discharge pressure PID controllers. These controllers are used for constraint handling only (the constraints are not active during normal operation). Therefore, the temperature control valve will only open when the lower constraint on the tempered water supply temperature of 26 °C is violated. Similarly, the pressure control valves will only throttle back should one or more of the pumps deliver low discharge pressure corresponding to a flow rate that will violate the maximum flow constraint for the pump. Table 3.1 gives

a list of parameters used in the model with their descriptions, typical values and units. For improved readability, some of them will be repeated in the text.

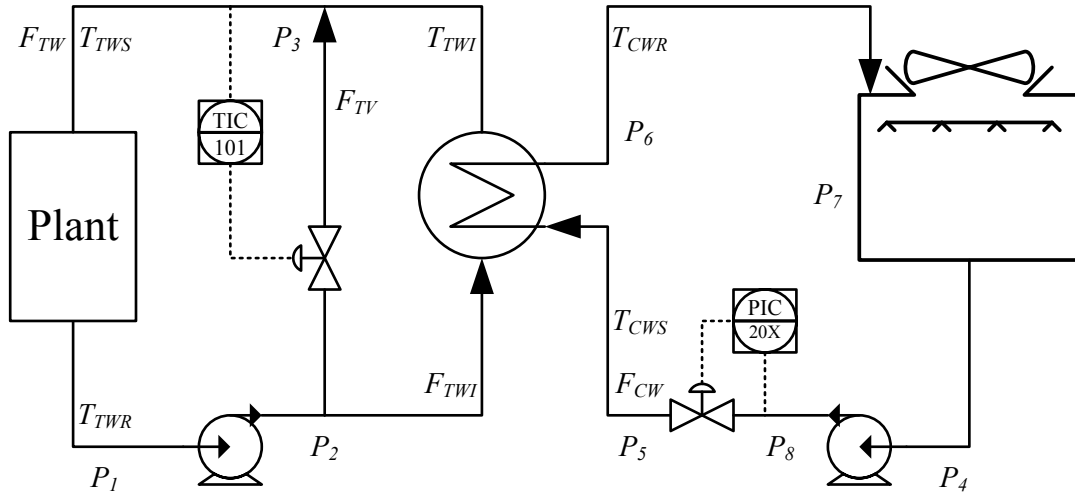


Figure 3.2: Simplified system representation.

In order to find a balance between accuracy and complexity for the model and to maintain a reasonable number of model parameters and variables [32], the model derivation was performed under the following simplifying assumptions:

- The pumps on each of the pump banks are identical and balanced in terms of receiving identical feed flow rates (when in a running state).
- The cooling towers are identical and balanced in terms of receiving identical feed flow rates (when in commission).
- The plant heat exchanger network has a constant flow coefficient (a constant system curve) as none of the exchanges are on automatic tempered water flow control (i.e. no control valves are changing the shape of the system curve on the plant side).
- A constant number of tempered water to cooling water heat exchangers are in use all the time (no commissioning or decommissioning of heat exchangers occur).
- The dosing system is not included in the model.

Table 3.1: Model parameters.

Variable	Description	Value	Units
τ_{TV}^f	Temperature control valve flow dynamic time constant	1/60	h
τ_{TWI}^f	TW intermediate flow dynamic time constant	1/60	h
τ_{TWI}^T	TW intermediate temperature dynamic time constant	4.86/60	h
τ_{TWS}^T	TW supply temperature dynamic time constant	3.87/60	h
τ_{TWR}^T	TW return temperature dynamic time constant	20.96/60	h
τ_{CW}^f	CW flow dynamic time constant	1/60	h
τ_{CWS}^T	CW supply temperature dynamic time constant	3.47/60	h
τ_{CWR}^T	CW return temperature dynamic time constant	8.25/60	h
P_s^{TW}	TW pump suction pressure	230	kPa-g
P_s^{CW}	CW pump suction pressure	20	kPa-g
K_{TW}	TW pump head correction factor	0.90	-
K_{CW}	CW pump head correction factor	0.58	-
ν	Vapour fraction CW flow for evaporative flow estimation	0.00153	%/°C
α	Approach of the cooling towers	10	°C
λ	Heat of vaporisation (for water in this case)	2260	kJ/kg
C_p	Specific heat (for water in this case)	4.18	kJ/kg.°C
C_{TV}	Flow coefficient of the temperature control valve	820	-
C_{PV}	Flow coefficient of the pressure control valves	200	-
C_{ETW}	Flow coefficient of the exchanger, TW side	737	-
C_{ECW}	Flow coefficient of the exchanger, CW side	237	-
C_{CT}	Flow coefficient of a single cooling tower	353	-
C_G	Flow coefficient of the plant heat exchanger network	417	-
R	Pressure control valve equal percentage coefficient	50	-
ρ	Specific gravity (for water in this case)	1	t/m ³
g	Gravitational constant	9.81	m/s ²
A	Heat exchanger area	100	m ²
U	Heat exchanger heat transfer coefficient	509	MJ/h.m ² .°C
W_{CT}	CT fan rated power	150	kW
c_c	Concentration cycles for CW circulation	3	-

- The side-stream filters are not included in the model.
- The combined suction pressure of the tempered water pumps is fixed at 230 kPa-g.
- The combined suction pressure of the cooling water pumps is fixed at 20 kPa-g (i.e. approximately 2m water level in the cooling towers).
- Heat addition by the work performed in the pumps is negligible.
- Heat exchange between the equipment and the surroundings is negligible (apart from that which occurs in the cooling towers).
- The same value is written to all the valves on the discharges of the cooling water pumps for the streams where the pumps are running and the valves are identical in size.
- When a cooling tower fan is switched off, the tower is automatically isolated from cooling water flow.
- The pump efficiencies are taken as constant values (i.e. not affected by flow rate).

As mentioned in Section 3.2, the controlled variables for the system are

- the tempered water supply temperature, T_{TWS} ,
- the tempered water differential temperature, $\Delta T_{TW} = T_{TWR} - T_{TWS}$,
- the power consumption of the system, W_T , and/or
- the electrical energy cost of the system, C_T .

The first two controlled variables are aimed more at constraint handling whereas the last two are aimed more at optimisation.

The model disturbance variables include

- the plant duty, Q_P (MJ/h),
- the ambient air temperature, T_a (°C), and

- the relative humidity of the air, RH (%).

The ambient air temperature and the relative humidity both affect the wet-bulb temperature of the air (T_{wb}) which affects the cooling efficiency of the cooling towers.

The model inputs are

- the tempered water pump running signals, $u_i^{TW} \in \{0, 1\}$,
- the cooling water pump running signals, $u_j^{CW} \in \{0, 1\}$,
- the cooling tower fan running signals, $u_k^{CT} \in \{0, 1\}$,
- the temperature control valve (TV) opening (OP_{TV}) (which has a linear characteristic), and
- the cooling water pump discharge pressure control valve (PV) openings (OP_{PV}) (which have equal percentage characteristics)

with $i = 1 \dots n_{TW}$, $j = 1 \dots n_{CW}$ and $k = 1 \dots n_{CT}$ where $n_{TW} = 5$ and $n_{CW} = 5$ are the numbers of tempered water and cooling water pumps and $n_{CT} = 4$ is the number of cooling tower fans. To reduce the number of input variables, the binary running signals are grouped together into discrete integer signals, representing the number of running pumps or fans such that $U_{TW} = \sum_{i=1}^{n_{TW}} u_i^{TW}$, $U_{CW} = \sum_{j=1}^{n_{CW}} u_j^{CW}$ and $U_{CT} = \sum_{k=1}^{n_{CT}} u_k^{CT}$.

3.3.1 Steady-state model

The next sections describe the derivations of the model elements mainly through the use of mass and energy balance concepts. At this stage only steady-state behaviour is considered.

3.3.1.1 Pump calculations

To determine the operating point of a pump in terms of discharge pressure, the pump performance curves are used (as illustrated in Figures 2.2 and 2.3). The performance curve indicates the head of liquid that the pump is delivering and the shaft power requirement at

a given flow rate. Alternatively, if the system curve is known, the operating point (flow and head) can be determined by the intercept between the pump curve and the system curve (see Section 2.7.1.2).

The pump curves can be captured in lookup tables for simulation purposes, although this is not ideal for use in control and optimisation. Therefore, the pump curves (both head and power curves) were modelled by fitting polynomial functions to the curve data [28]. Given a certain mass flow rate, f (t/h), these polynomial functions then generate the corresponding head, h (m), and power w (kW) for the pump. The liquid head is translated to discharge pressure, P_d (kPa-g), with $P_d = hg\rho + P_s$ where P_s is the suction pressure (kPa-g), $g = 9.81$ (m/s²) is the gravitational constant, and ρ (t/m³) is the specific gravity of the liquid (water in this case). The resulting polynomial functions are given in Section 3.3.2.

The pump running signals mentioned in the previous section indicate the number of pumps running on a bank. As mentioned, one of the modelling assumptions is that the running pumps on a bank are identical and perfectly balanced in terms of receiving the same flow, at the same suction pressure, and delivering the same discharge pressure. This significantly simplifies the modelling of the assembled system.

3.3.1.2 Flow calculations

With reference to Figure 3.2, the flow in the tempered water circuit is determined by the tempered water pumps' common discharge pressure (P_2) and the system characteristic (one of the modelling assumptions is that the suction pressure for the pumps is fixed). The system characteristic is determined by the flow coefficients of the temperature control valve (together with the fraction valve opening), the plant heat exchanger network and the cooling water heat exchangers' tempered water side. The total flow in the tempered water circuit is therefore calculated as

$$\begin{aligned}
 f_{TW} &= f_{TWI} + f_{TV} \\
 &= C_{ETW} \sqrt{\frac{\Delta P_{ETW}}{\rho}} + C_{TV} OP_{TV} \sqrt{\frac{\Delta P_{ETW}}{\rho}}
 \end{aligned} \tag{3.1}$$

and

$$f_{TW} = C_G \sqrt{\frac{\Delta P_P}{\rho}} \quad (3.2)$$

where f_{TV} is the mass flow through the temperature control valve (TV); f_{TWI} is the intermediate mass flow through the tempered water side of the heat exchanger bank; C_{TV} , C_{ETW} , and C_G are the flow coefficients of the temperature valve, the tempered water side of the heat exchangers, and the plant; OP_{TV} is the valve opening of the temperature valve; and $\Delta P_{ETW} = P_2 - P_3$ and $\Delta P_P = P_3 - P_1$ are differential pressures across the tempered water side of the heat exchangers (including the temperature valve) and across the plant heat exchanger network. $\Delta P_{TWP} = P_2 - P_1$ is the differential pressure across the pumps and is calculated from the pump curve fitted polynomial function (see Section 3.3.1.1). Furthermore,

$$\Delta P_P = \Delta P_{TWP} - \Delta P_{ETW} \quad (3.3)$$

and by equating (3.1) and (3.2) and substituting for (3.3) results in

$$\Delta P_{ETW} = \frac{\Delta P_{TWP}}{\left(\frac{C_{ETW} + C_{TV} OP_{TV}}{C_G}\right)^2 + 1} \quad (3.4)$$

which allows for the calculation of f_{TWI} and f_{TV} .

In the cooling water circuit, the flow rate is determined by the cooling water pumps' discharge pressure (P_8); the flow coefficients of the pressure control valves for the active pumps (together with the fraction valve openings – the assumption is that the same valve opening is written to all the valves); and the flow coefficients of the cooling towers and the cooling water heat exchangers' cooling water side. Once again, the assumption is that the combined suction pressure for all the pumps is constant. To determine the flow in the cooling water circuit, f_{CW} , the differential pressures in the cooling water circuit are broken down into the differential pressures:

across the pressure control valves,

$$\Delta P_{PV} = P_8 - P_5, \quad (3.5)$$

across the cooling towers,

$$\Delta P_{CT} = P_7 - P_4, \quad (3.6)$$

across the heat exchangers on the cooling water side,

$$\Delta P_{ECW} = P_6 - P_5, \quad (3.7)$$

across the cooling water pumps,

$$\Delta P_{CWP} = P_8 - P_4, \quad (3.8)$$

across the pumps and valves together,

$$\Delta P_{CW} = P_5 - P_4 \quad (3.9)$$

and noting that

$$\Delta P_{PV} = \Delta P_{CWP} - \Delta P_{CW}. \quad (3.10)$$

The flow equations are then written as

$$\begin{aligned} f_{CW} &= U_{CW} C_{PV} R^{(1-OP_{PV})} \sqrt{\frac{\Delta P_{PV}}{\rho}} \\ &= U_{CW} C_{PV} R^{(1-OP_{PV})} \sqrt{\frac{\Delta P_{CWP} - \Delta P_{CW}}{\rho}}, \end{aligned} \quad (3.11)$$

$$f_{CW} = U_{CT}C_{CT}\sqrt{\frac{\Delta P_{CT}}{\rho}}, \quad (3.12)$$

and

$$f_{CW} = C_{ECW}\sqrt{\frac{\Delta P_{ECW}}{\rho}} \quad (3.13)$$

where C_{PV} , C_{ECW} , and C_{CT} are the flow coefficients of the pressure control valves, the cooling water side of the heat exchangers and the cooling towers, OP_{PV} is the valve openings of the pressure control valves and R is the equal percentage design coefficient for the pressure control valves. The specific gravity $\rho = 1$ (t/m³) (for water) and will be omitted from the rest of the equations. Noting that $\Delta P_{CW} = \Delta P_{CT} + \Delta P_{ECW}$ and substituting with Equations 3.12 and 3.13 gives

$$\Delta P_{CW} = \frac{f_{CW}^2}{U_{CT}^2 C_{CT}^2} + \frac{f_{CW}^2}{C_{ECW}^2}. \quad (3.14)$$

Similarly, Equation 3.11 can be written as

$$\Delta P_{CW} = \Delta P_{CWP} - \frac{f_{CW}^2}{U_{CW}^2 C_{PV}^2 R^{2(1-OP_{PV})}} \quad (3.15)$$

and equating 3.14 and 3.15 results in

$$f_{CW} = \frac{C_{CT}U_{CT}C_{ECW}U_{CW}CV_{PV}R^{(OP_{PV}-1)}\sqrt{\Delta P_{CWP}}}{\sqrt{C_{ECW}^2U_{CW}^2CV_{PV}^2R^{2(OP_{PV}-1)} + C_{CT}^2U_{CT}^2U_{CW}^2CV_{PV}^2R^{2(OP_{PV}-1)} + C_{CT}^2U_{CT}^2C_{ECW}^2}}. \quad (3.16)$$

3.3.1.3 Duty calculations

Several processes require the addition of heat or removal of excess heat from a product stream, such as heat of reaction for exothermic processes or the condensing of an overhead vapour

stream in a distillation process. This energy transfer is typically achieved through the use of heat exchangers. In particular, the removal of energy is normally achieved through the use of cooling water or refrigerated water.

The accumulated thermal power to be removed by the cooling water system will be referred to as the plant duty, Q_P (MJ/h). After transferring the energy from the plant to the tempered water circuit through the plant heat exchanger network, it is transferred to the cooling water through the cooling water heat exchanger bank (shown as E-101 to E-105 in Figure 3.1). The heat is then removed from the cooling water in the cooling towers, mainly through the partial evaporation of the cooling water and to a lesser extent through heat transfer from the liquid water to the air stream.

In the cases of the plant and cooling water heat exchangers, the mechanism for the heat transfer (assuming that no phase change occurs) is referred to as sensible heat transfer (MJ/h) and is calculated by

$$Q = fC_p\Delta T \quad (3.17)$$

where f is the mass flow of the fluid (t/h), C_p is the specific heat of the fluid (kJ/kg°C), and ΔT is the differential temperature between the inlet and the outlet of the heat exchanger (°C). This holds for both the process and utility sides of the exchanger which, in the case of the cooling water heat exchangers, are both water streams with $C_p = 4.18$ kJ/kg°C.

In addition, the heat exchanger design plays a part in the heat transfer efficiency and its duty is determined by the heat transfer coefficient, U (MJ/hm²°C), the heat exchange area, A (m²), and the log mean temperature difference between the process and utility streams (°C) [22],

$$\Delta T_{lm} = \frac{(T_{TWR} - T_{CWS}) - (T_{TWI} - T_{CWR})}{\ln \left(\frac{T_{TWR} - T_{CWS}}{T_{TWI} - T_{CWR}} \right)} \quad (3.18)$$

where T_{CWS} and T_{CWR} are the cooling water supply and return temperatures; T_{TWI} is the tempered water intermediate temperature (the heat exchanger outlet temperature on the tempered water side); and T_{TWR} is the tempered water return temperature (refer to Figure

3.2). The heat exchanger duty is then given by

$$Q = UA\Delta T_{lm}. \quad (3.19)$$

The larger the heat exchange area, A , or the better the exchanger design (presented by U), the better the heat transfer.

T_{TWI} is calculated by substituting (3.17) (on the tempered water side) into (3.19) giving

$$T_{TWI} = \frac{T_{CWS} - \frac{Q}{f_{TWI}C_p} - e^K T_{CWR}}{1 - e^K} \quad (3.20)$$

with

$$K = \left(\frac{Q}{f_{TWI}C_p} - T_{CWS} + T_{CWR} \right) \frac{UA}{Q}. \quad (3.21)$$

The tempered water supply temperature is then calculated as

$$T_{TWS} = \frac{T_{TWI}f_{TWI} + T_{TWR}f_{TV}}{f_{TW}} \quad (3.22)$$

where f_{TWI} , f_{TV} , and f_{TW} are discussed in Section 3.3.1.2. It is clear that when the temperature control valve is closed, $f_{TW} = f_{TWI}$ in which case the system reduces to $T_{TWS} = T_{TWI}$.

The heat exchange in the cooling towers is a more complex matter and a simplified model was developed (compared to what is typically found in literature [21, 22, 23, 24]). This model uses the wet-bulb temperature (T_{wb}) of the air that is sent through the towers, which can be estimated using the ambient temperature, T_a , which is measured, together with the relative humidity (RH) measurement as follows [67]:

$$\begin{aligned}
 T_{wb} = & T_a \tan^{-1}(0.151977(RH + 8.313659)^{0.5}) \\
 & + \tan^{-1}(T_a + RH) - \tan^{-1}(RH - 1.676331) \\
 & + 0.00391838(RH^{1.5}) \tan^{-1}(0.023101.RH) \\
 & - 4.686035
 \end{aligned} \tag{3.23}$$

It is not practically possible to reach the wet-bulb temperature on the water exiting the cooling towers, as the main mechanism for heat transfer is the partial evaporation of the water into the air, which would cease at the wet-bulb temperature. The difference between the achievable T_{CWS} and T_{wb} is called the approach (α) and is a function of the tower design and the operating conditions. For example, if the tower is designed to spray water in finer droplets, the area for evaporation would increase, which would result in a smaller approach. Similarly, if the tower is designed to maximise contact time between the air and the water droplets, the evaporation would also be improved, once again leading to a smaller approach. The evaporation rate in the cooling towers (t/h) can be estimated using

$$f_e = \nu f_{CW}(T_{CWR} - T_{CWS}) \tag{3.24}$$

and

$$f_e = f_{mu} - f_b - f_d \tag{3.25}$$

where f_{mu} is the make-up water (added to the system on cooling tower level control), f_b is the blow-down (to prevent solids build-up in the system), and f_d is the drift loss through splashing and entrainment of the droplets in the air stream. Combining the estimations from [22] and [21] results in a vaporisation fraction $\nu = 0.00153$ (%/°C), $f_d = 0.001f_{CW}$ and $f_b = f_e/2$ (assuming three cycles of concentration, c_c). Equation (3.25) can then be used to approximate f_e . It does not, however, contain any reference to the ambient conditions in its current form. Therefore, a convergence term, $(T_{CWS}/(T_{wb} + \alpha))$, is introduced which serves to compensate for ambient condition changes (through its reference to the wet-bulb temperature) and also takes the approach into account. For example, if T_{wb} drops due to

a drop in RH on a dry day, it will result in a higher f_e which means more cooling due to the ability of the air to absorb moisture being higher. Equation (3.24) can then be used to calculate the new value of T_{CWS} ; and from the calculated values for T_{CWS} and T_{TWS} , the new values for T_{CWR} and T_{TWR} can be determined with the use of Equation 3.17.

The portion of the duty of the cooling towers due to the partial evaporation is calculated as

$$Q_e = f_e \lambda \quad (3.26)$$

where $\lambda = 2260$ kJ/kg is the approximate heat of vapourisation for water.

3.3.1.4 Power calculations

The power consumption of the system is required to be used as a control or optimisation variable, which may not be available as a direct measurement. Therefore, a power consumption model of the system is formulated based on the running statuses and power consumptions of the individual equipment. As discussed in Section 2.6, the running signal of the i th component is denoted by

$$u_i = \begin{cases} 1 & \text{for component ON} \\ 0 & \text{for component OFF} \end{cases} \quad (3.27)$$

and the power consumption by

$$W_i = u_i w_i \quad (3.28)$$

where w_i is the operating power of the i th component. In its simplest form, the model can be constructed by assuming a constant power consumption equal to the rated power of the equipment such that

$$w_i = W_i^{\max} \quad (3.29)$$

where W_i^{\max} denotes the rated power of the i th component obtained from the equipment data sheet [12]. It may be, however, that a more accurate model is required. If the power consumption data is available (for example from the power curves of the pumps) it can be calculated using the fitted polynomial functions (as discussed in Section 3.3.1.1) and flow rate (which is determined by the system operating point as discussed in Section 3.3.1.2). For this study, the rated power is used for the cooling tower fans and polynomial functions are used for the tempered water and cooling water pumps. The power functions are given in Section 3.3.2 (Equations 3.52 to 3.54). The total consumption at time t is

$$W_T(t) = \sum_{i=1}^N W_i(t) \quad (3.30)$$

with $N = n_{TW} + n_{CW} + n_{CT}$ and the total cost is

$$C_T(t) = W_T(t)p(t) \quad (3.31)$$

where $p(t)$ is the electricity price at time t . If there is a flat rate for the electricity and no time-of-use (TOU) tariff is applicable, $p(t)$ reduces to a constant [13].

For optimisation purposes, an objective function for energy optimisation may be formulated for time period $[t_0, t_f]$ as

$$J_W = \int_{t_0}^{t_f} W_T(t)dt \quad (3.32)$$

or for energy cost

$$J_C = \int_{t_0}^{t_f} C_T(t)dt. \quad (3.33)$$

3.3.2 State-space dynamic model

The model as described in 3.3.1 is a steady-state model, which does not contain information pertaining to the dynamic behaviour of the system. That might be sufficient for a real-time optimisation application though it will typically not be sufficient for use in a dynamic control application. Therefore, for the design of control and optimisation solutions for the system, as well as performing dynamic simulation, it is beneficial to represent the system in a dynamic form capturing the transients of the key process variables.

Due to its wide-spread use, a state-space representation was developed. To achieve this, dynamics were added to selected variables to represent the system states, whereas other variables were maintained as algebraic equations. To maintain a reasonable level of complexity, only the most important dynamic behaviours were included. To further simplify the model, it is assumed that each state variable's transient behaviour is described by a first order dynamic, which is the same for all its sources of change (inputs, disturbances, other states, etc.) with the gains differing according to the non-linear equations discussed in Section 3.3.1.

The model is based on the full system shown in Figure 3.1 and is an extension of the model in Section 3.3.1. It is presented in the form of a discrete time state-space model consisting of difference equations and algebraic equations. The states are

$$f_{TV}(k+1) = \frac{\Delta t C_{TV}}{\tau_{TV}^f} OP_{TV}(k) \sqrt{\frac{\Delta P_{ETW}(k)}{\rho}} + f_{TV}(k) \left(1 - \frac{\Delta t}{\tau_{TV}^f}\right) \quad (3.34)$$

$$f_{TWI}(k+1) = \frac{\Delta t C_{ETW}}{\tau_{TWI}^f} \sqrt{\frac{\Delta P_{ETW}(k)}{\rho}} + f_{TWI}(k) \left(1 - \frac{\Delta t}{\tau_{TWI}^f}\right) \quad (3.35)$$

$$f_{CW}(k+1) =$$

$$\frac{\Delta t}{\tau_{CW}^f} \left(\frac{C_{CT} U_{CT}(k) C_{ECW} U_{CW}(k) C_{VPV} R^{(OP_{PV}(k)-1)} \sqrt{\Delta P_{CWP}(k)}}{\sqrt{C_{ECW}^2 U_{CW}^2(k) C_{VPV}^2 R^{2(OP_{PV}(k)-1)} + C_{CT}^2 U_{CT}^2(k) U_{CW}^2(k) C_{VPV}^2 R^{2(OP_{PV}(k)-1)} + C_{CT}^2 U_{CT}^2(k) C_{ECW}^2}} \right) + f_{CW}(k) \left(1 - \frac{\Delta t}{\tau_{CW}^f}\right) \quad (3.36)$$

$$T_{TWS}(k+1) = \frac{\Delta t}{\tau_{TWS}^T} \left(\frac{T_{TWI}(k)f_{TWI}(k) + T_{TWR}(k)f_{TV}(k)}{f_{TW}(k)} \right) + T_{TWS}(k) \left(1 - \frac{\Delta t}{\tau_{TWS}^T} \right) \quad (3.37)$$

$$T_{TWR}(k+1) = \frac{\Delta t}{\tau_{TWR}^T} \left(\frac{Q_P(k)}{f_{TW}(k)C_p} + T_{TWS}(k) \right) + T_{TWR}(k) \left(1 - \frac{\Delta t}{\tau_{TWR}^T} \right) \quad (3.38)$$

$$T_{CWR}(k+1) = \frac{\Delta t}{\tau_{CWR}^T} \left(\frac{Q_P(k)}{f_{CW}(k)C_p} + T_{CWS}(k) \right) + T_{CWR}(k) \left(1 - \frac{\Delta t}{\tau_{CWR}^T} \right) \quad (3.39)$$

$$T_{CWS}(k+1) = \frac{\Delta t}{\tau_{CWS}^T} \left(T_{CWR}(k) - \frac{f_e(k)}{f_{CW}(k)\nu} \right) + T_{CWS}(k) \left(1 - \frac{\Delta t}{\tau_{CWS}^T} \right) \quad (3.40)$$

$$T_{TWI}(k+1) = \frac{\Delta t}{\tau_{TWI}^T} \left(\frac{T_{CWS}(k) - \frac{Q_P(k)}{f_{TWI}(k)C_p} - e^{K(k)}T_{CWR}(k)}{1 - e^{K(k)}} \right) + T_{TWI}(k) \left(1 - \frac{\Delta t}{\tau_{TWI}^T} \right) \quad (3.41)$$

with algebraic equations

$$K(k) = \left(\frac{Q_P(k)}{f_{TWI}(k)C_p} - T_{CWS}(k) + T_{CWR}(k) \right) \frac{UA}{Q_P(k)} \quad (3.42)$$

$$f_{TW}(k) = F_{TV}k + F_{TWI}(k) \quad (3.43)$$

$$\Delta P_{TWP}(k) = k_{TW}h_{TW}(k)\rho g \quad (3.44)$$

$$h_{TW}(k) = -(7 \times 10^{-6}) \left(\frac{f_{TW}(k)}{U_{TW}(k)} \right)^2 + 0.0036 \frac{f_{TW}(k)}{U_{TW}(k)} + 88.28 \quad (3.45)$$

$$\Delta P_{ETW}(k) = \frac{\Delta P_{TWP}(k)}{\left(\frac{C_{ETW} + C_{TVOPTV}(k)}{C_G} \right)^2 + 1} \quad (3.46)$$

$$\Delta P_{CWP}(k) = k_{CW} h_{CW}(k) \rho g \quad (3.47)$$

$$h_{CW}(k) = -(6 \times 10^{-6}) \left(\frac{f_{CW}(k)}{U_{CW}(k)} \right)^2 + 0.0005 \frac{f_{CW}(k)}{U_{CW}(k)} + 58.30 \quad (3.48)$$

$$f_e(k) = \left(\frac{c_c - 1}{c_c} \right) \left(\frac{f_{mu}(k)}{U_{CT}(k)} - 0.001 \left(\frac{f_{CW}(k)}{U_{CT}(k)} \right) \right) \left(\frac{T_{CWS}(k)}{T_{WB}(k) + \alpha} \right) U_{CT}(k) \quad (3.49)$$

$$Q_e(k) = f_e(k) \lambda \quad (3.50)$$

where Δt is the sampling time and $T_{WB}(k)$ is obtained from (3.23). The total power is given in the form

$$W_T(k) = W_{TW}(k) + W_{CW}(k) + W_{CT}(k) \quad (3.51)$$

where

$$W_{TW}(k) = U_{TW}(k) \left(-(2 \times 10^{-5}) \left(\frac{f_{TW}(k)}{U_{TW}(k)} \right)^2 + 0.1772 \frac{f_{TW}(k)}{U_{TW}(k)} + 131.7 \right), \quad (3.52)$$

$$W_{CW}(k) = U_{CW}(k) \left(-(8 \times 10^{-9}) \left(\frac{f_{CW}(k)}{U_{CW}(k)} \right)^3 + (2 \times 10^{-5}) \left(\frac{f_{CW}(k)}{U_{CW}(k)} \right)^2 + 0.0195 \frac{f_{CW}(k)}{U_{CW}(k)} + 182.9 \right) \quad (3.53)$$

and

$$W_{CT}(k) = U_{CT}(k) W_{CT}^{\max}. \quad (3.54)$$

The model therefore has eight state equations represented by (3.34) to (3.41), fourteen algebraic equations (of which some are intermediate variables) given by (3.42) to (3.54), five

inputs ($U_{TW}(k)$, $U_{CW}(k)$, $U_{CT}(k)$, $OP_{TV}(k)$ and $OP_{PV}(k)$), and three measured (or estimated) disturbances ($Q_P(k)$, $T_a(k)$ and $RH(k)$).

3.3.3 Parameter estimation

The model as described in Sections 3.3.1 and 3.3.2 was verified against plant data to ensure that the behaviour closely represents the actual process behaviour. A simulation study of the model illustrates a baseline behaviour of how the plant is normally operated, with all or most of the pumps and cooling towers in commission, except when failures occur.

Several key parameters were identified for the model. Initial parameter values were determined from design data (where available) and estimated from historical data and experience (where not available). The parameters include the flow coefficients (C_G , C_{ETW} , C_{ECW} , C_{CT} , C_{TV} and C_{PV}), the temperature time constants (τ_{TWS}^T , τ_{TWR}^T , τ_{CWS}^T , τ_{CWR}^T and τ_{TWI}^T), the heat exchanger heat transfer coefficient (U) and the pump head correction factors (K_{TW} and K_{CW}).

After running the model simulation with the initial values and comparing the simulated variables to that of the actual plant data (both visually and through the calculation of correlation coefficients), it became apparent that improved model accuracy was required. Consequently, an optimisation algorithm was implemented to determine a parameter set resulting in a better model accuracy.

The estimation parameters were selected based on perceived relative importance. The most important model dynamics are contained in the temperature equations and therefore the temperature time constants were chosen. The flow behaviour of the system is mainly affected by the pump performances and the flow coefficients of the individual system components and therefore these parameters were chosen. By omitting and adding parameters between estimation iterations, the sensitivity of the model quality on the parameters could be gauged. This ensured that all the selected parameters were significant enough to be included in the estimation problem with the aim of including as few parameters as possible.

Two optimisation exercises were performed. The first was to improve on the dynamic behaviour of the model and the second to improve on the model gains. The performance function

for the optimiser was changed accordingly to fulfil this purpose. For the first case, a sum of correlation coefficients between the plant data and the simulation results for the tempered water supply and return temperatures (T_{TWS} and T_{TWR}), the tempered water differential temperature, ΔT_{TW} , and the duty were used. For the second case, a least squared error approach was used between the simulated and plant data for the tempered water supply temperature and the tempered water differential temperature. The correlation coefficients and visual inspection of the data trends were used to gauge the improvement in model fit. Due to the absence of historised switching data for the cooling tower fans, they were taken to be constant at four fans running through the model validation exercise.

The parameter estimation problem is a constrained non-linear programming problem with only continuous input variables. The non-linearity is due to the non-linear process model (and the quadratic objective function for the second optimisation case). Gradient-based techniques were initially used for the least squared error parameter estimation (including SQP, interior-point and active set algorithms). Thereafter, a genetic algorithm was used, taking into consideration its ability to handle non-linear, constrained, multi-parameter problems directly and its higher probability of finding a global optimum, as opposed to settling at a local optimum (see Section 2.4.1 for more information on the genetic algorithm). The genetic algorithm yielded superior results compared to the gradient-based algorithms and was therefore selected.

To improve the convergence of the algorithm, upper and lower bounds were set for each parameter. The simulation was performed in Matlab using the existing `ga` function with a population size of 20 and a maximum number of generations of 50. The exercise was performed with specified initial conditions as well as randomly selected initial values. The random parameter selection provides more randomness to iterations of the optimisation problem which, further assists with preventing convergence to a local optimum. The rest of the `ga` parameters were left at the default values for problems with linear constraints (which include population type = double, cross-over fraction = 0.8, elite count = 2, mutation function = adaptive feasible and cross-over function = intermediate).

Figure 3.3 shows the input data to the model for the plant duty, $Q_P(t)$, (estimated using (3.17) with the measured flow rates and temperatures) and the wet-bulb temperature, $T_{wb}(t)$ (calculated using (3.23) with the measured ambient temperature, $T_a(t)$, and humidity, RH).

The input data represent a period of 6 days (144 hours) of plant operation at a one minute sampling interval during which significant load changes occurred.

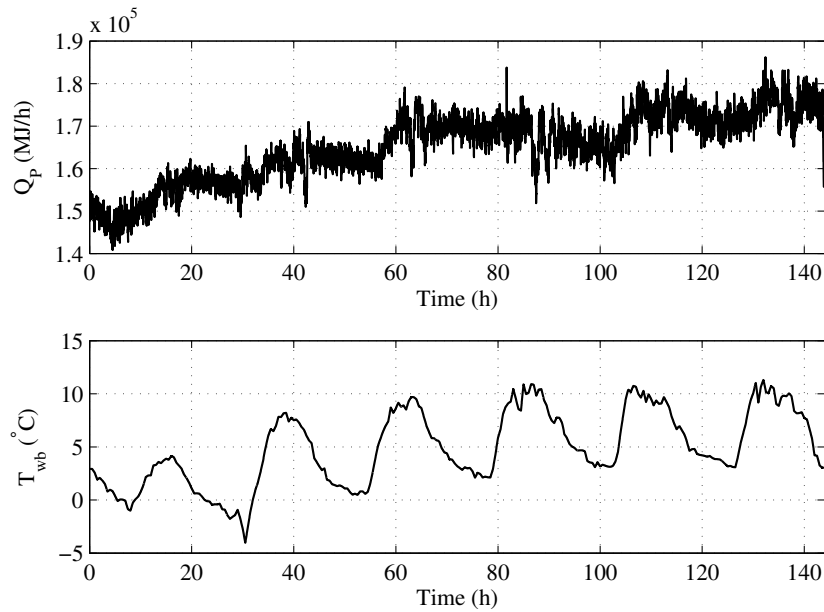


Figure 3.3: Model disturbance inputs for validation (plant duty and wet-bulb temperature).

3.3.4 Process modelling results

A simulation was performed for model verification using the model as discussed in Sections 3.3.1 and 3.3.2, with initial values based on design data, experience and historical data as mentioned in Section 3.3.3. The initial values are shown in Table 3.2.

The input data to the model for Q_P and T_{wb} are shown in Figure 3.3 in Section 3.3.3. Other inputs include the pump and cooling tower fan running signals U_{TW} , U_{CW} and U_{CT} . The verification data covers a period of 6 days (144 hours) of plant operation at a one minute sampling interval, during which significant load changes occurred. Some possible unmeasured disturbances that were not modelled, include heat exchanger isolation¹ and rain events.

The simulation yielded promising results using the correlation coefficients and squared errors between some of the plant and model outputs as a measure of similarity between the data

¹The effect of isolating a heat exchanger is twofold: Firstly it reduces the total heat exchange area which affects the stream temperatures on both loops. Secondly it affects the flow coefficient of the heat exchanger bank (on both the tempered water and cooling water sides) and therefore affects the flow rates in both loops.

sets and therefore the validity of the model.

The modelling was constrained by the following limitations:

- Limited process variable measurement data on the cooling towers (for example no flow or temperature measurement on the air through the towers and no pressure measurement on cooling water entering cooling towers).
- No cooling water heat exchanger commissioning/decommissioning data (not historised).
- No blow-down water flow measurement which requires its approximation using concentration cycles (see Section 3.3.1.3).
- Significant measurement noise on some variables which also affects calculated variables.
- Inability to perform step tests on the actual process.
- Limited tempered water pump switching data (only a few pump switching activities for the period of validation are observed).
- No cooling tower fan switching data (not historised).

As discussed in Section 3.3.3, several key parameters were estimated using a genetic algorithm to optimise firstly the correlation coefficients between plant and simulation data for the tempered water supply temperature, tempered water differential temperature and duty, and secondly the squared errors between the plant and simulation data for the tempered water supply temperature and the tempered water differential temperature.

With a population size of 20 and randomly selected initial values the genetic algorithm typically converges after 15 to 20 generations. The estimated (optimised) parameter values are given in Table 3.2. The upper and lower bounds that were used for the parameters during the estimation problem were set such that it would still yield realistic parameter values. Therefore, the estimated values are deemed to be a realistic representation of the plant. The step test results discussed in the next section further supports the legitimacy of the model behaviour.

Figure 3.4 shows the model response (the solid line) versus the plant data (the dotted line)

Table 3.2: Optimised model parameters

Parameter	Initial	Optimised
C_G	570	812
C_{ETW}	570	737
C_{ECW}	300	539
C_{CT}	220	193
C_{TV}	2000	912
C_{PV}	200	136
U	330	284
K_{TW}	0.90	0.45
K_{CW}	0.90	0.32
τ_{TWS}^T	3/60	6.06/60
τ_{TWR}^T	20/60	21.42/60
τ_{CWS}^T	5/60	7.39/60
τ_{CWR}^T	6/60	4.38/60
τ_{TWI}^T	3/60	8.00/60

for T_{TWR} and T_{TWS} using the optimised model parameters given in Table 3.2, on the data set used for the parameter fitting exercise. Q_e can be calculated using Equation 3.26 and reveals that it is indeed approximately 80% of the total duty as stated in Section 3.2.

The correlation coefficients for T_{TWR} , T_{TWS} , ΔT_{TW} and Q_e are shown in Table 3.3 for both the initial parameter and optimised parameter cases. A clear model accuracy improvement was achieved with the optimisation algorithm, especially on the tempered water temperatures.

Table 3.3: Model correlation coefficients with and without parameter estimation.

Variable	Initial	Optimised
T_{TWR}	0.6182	0.9339
T_{TWS}	0.4968	0.8924
Q_e	0.9020	0.9320
ΔT_{TW}	0.6309	0.7896

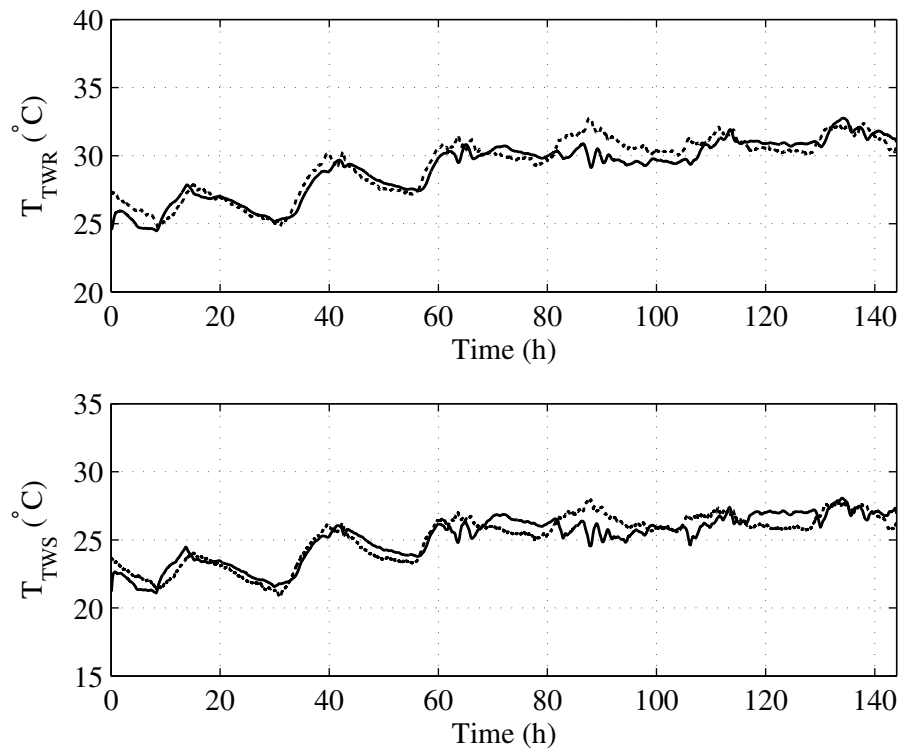


Figure 3.4: Model tempered water temperature results (model response (solid line) versus the plant data (dotted line)).

3.3.5 Model step tests

Figures 3.5 to 3.12 show the response of the model to input and disturbance variable step changes over the operating ranges with regard to some of the important process variables. The columns of the model matrices represent the manipulated/disturbance variables whereas the rows show the responses of the process variables. One manipulated/disturbance variable was changed at a time, while keeping the others constant at nominal values. The steps were made over a period of 10 hours, starting from steady-state conditions in each case and using typical move sizes. To illustrate the relative severities of the responses, the scales on the vertical axes were kept constant across each row (although the reference values may differ).

The process variables that were observed are: the tempered water flow rate (f_{TW}), the cooling water flow rate (f_{CW}), the tempered water supply temperature (T_{TWS}), the cooling water supply temperature (T_{CWS}), the total power consumption of the system (W_T) and the tempered water differential temperature (ΔT_{TW}).

Figures 3.5 and 3.6 show the responses for changes in the tempered water pump running signals (U_{TW}) and the cooling water pump running signals (U_{CW}). Figures 3.7 and 3.8 show the responses to changes in the cooling tower fan running signals (U_{CT}) and the temperature control valve opening (OP_{TV}). Figures 3.9 and 3.10 show the responses to changes in the cooling water pressure control valve openings (OP_{PV}) and the plant duty (Q_P). Finally, Figures 3.11 and 3.12 show the responses to changes in the ambient temperature (T_a) and relative humidity (RH).

The responses clearly illustrate the expected non-linear and interactive nature of the process though it also reveals some interesting dynamics and interactions.

3.3.5.1 Manipulated variable stepping

Starting with the tempered water pump running signals, U_{TW} , Figure 3.5 shows how the progressive increase in tempered water flow rate (f_{TW}) gets smaller as more pumps are brought on-line. This behaviour is in accordance with that of pumps running in parallel as discussed in Section 2.7.1.2, which contributes to the overall non-linearity of the system. As expected, there is no response in the cooling water flow (f_{CW}) from a change in the tempered water pumps as they are on physically separated circuits. A very interesting response is observed on the tempered water supply temperature (T_{TWS}), which increases as more pumps are brought on-line, rather than decreasing as would be expected. This is due to the decrease in tempered water differential temperature (ΔT_{TW} , see Figure 3.6), which affects T_{TWS} more drastically than the tempered water return temperature, T_{TWR} . Also evident from Figure 3.6 is that the cooling water supply temperature shows no response, as the assumption is that the energy transferred from the plant to the tempered water circuit, and then in turn to the cooling water circuit, does not change with a change in tempered water flow (the duty remains constant for the duration of the step changes in U_{TW}). As expected, an increase in tempered water pumps causes an increase in total power consumption, W_T , though the effect is also non-linear, as the incremental increase in power becomes smaller as more pumps are started. This is due to the inclusion of the pump power curves (see Figure 2.2) as opposed to simply using the rated power, which accounts for the fact that the power is related to the mass flow through the pumps, which becomes smaller per pump when running more pumps. The number of running pumps were increased in increments of one, starting from one, going

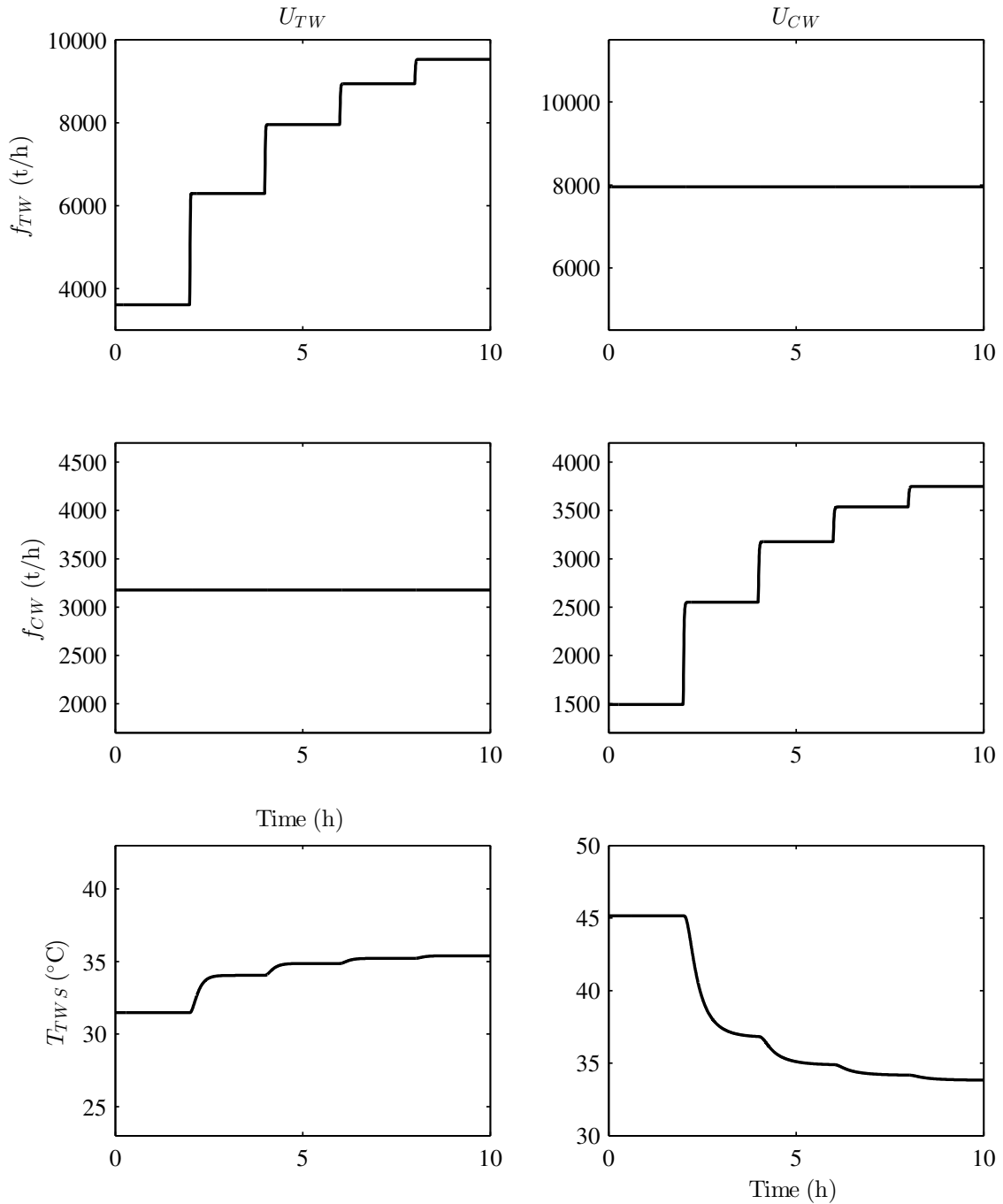


Figure 3.5: Responses in f_{TW} , f_{CW} , and T_{TWS} (represented by the model matrix rows) to changes in U_{TW} and U_{CW} (represented by the columns).

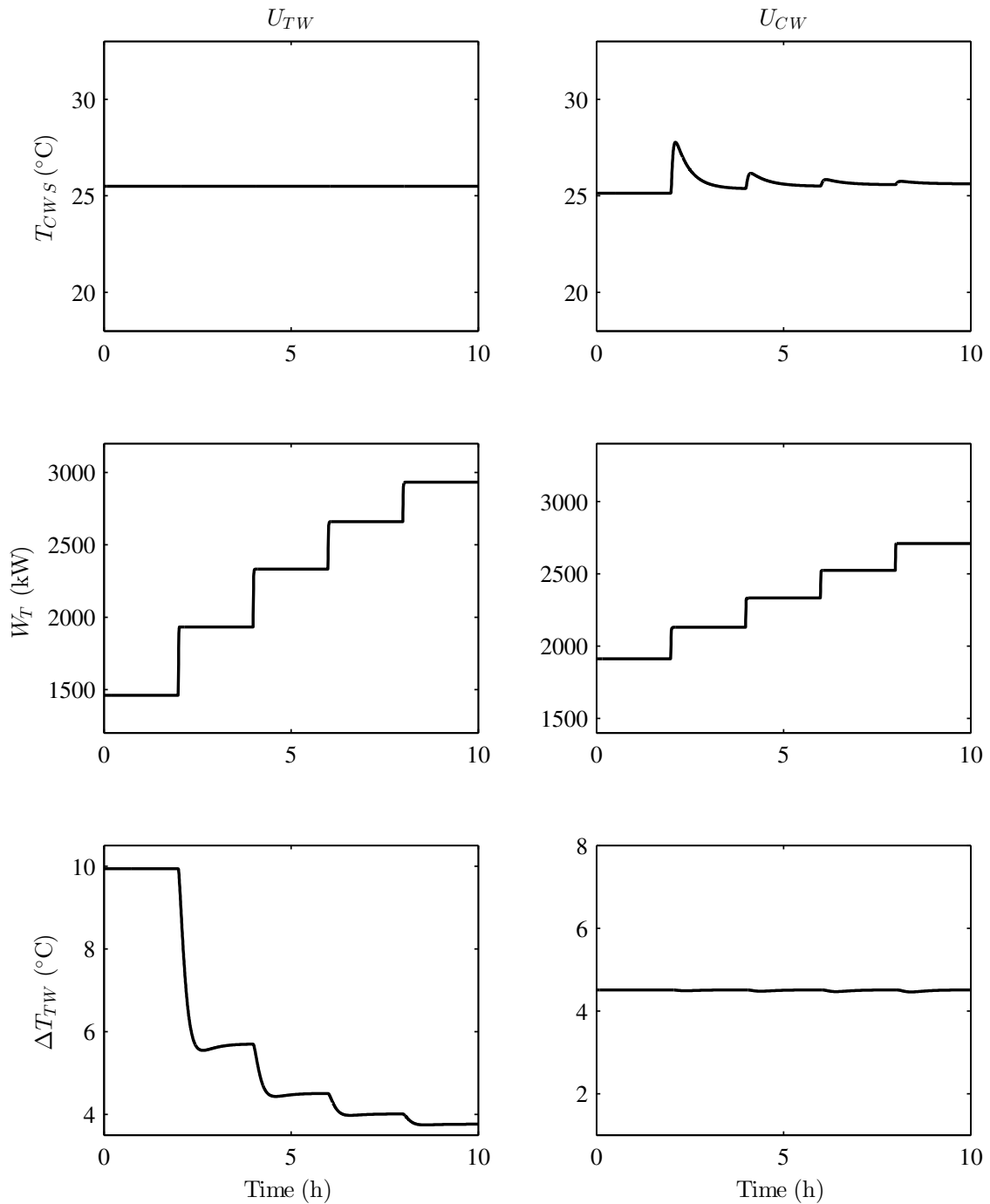


Figure 3.6: Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in U_{TW} , U_{CW} (represented by the columns).

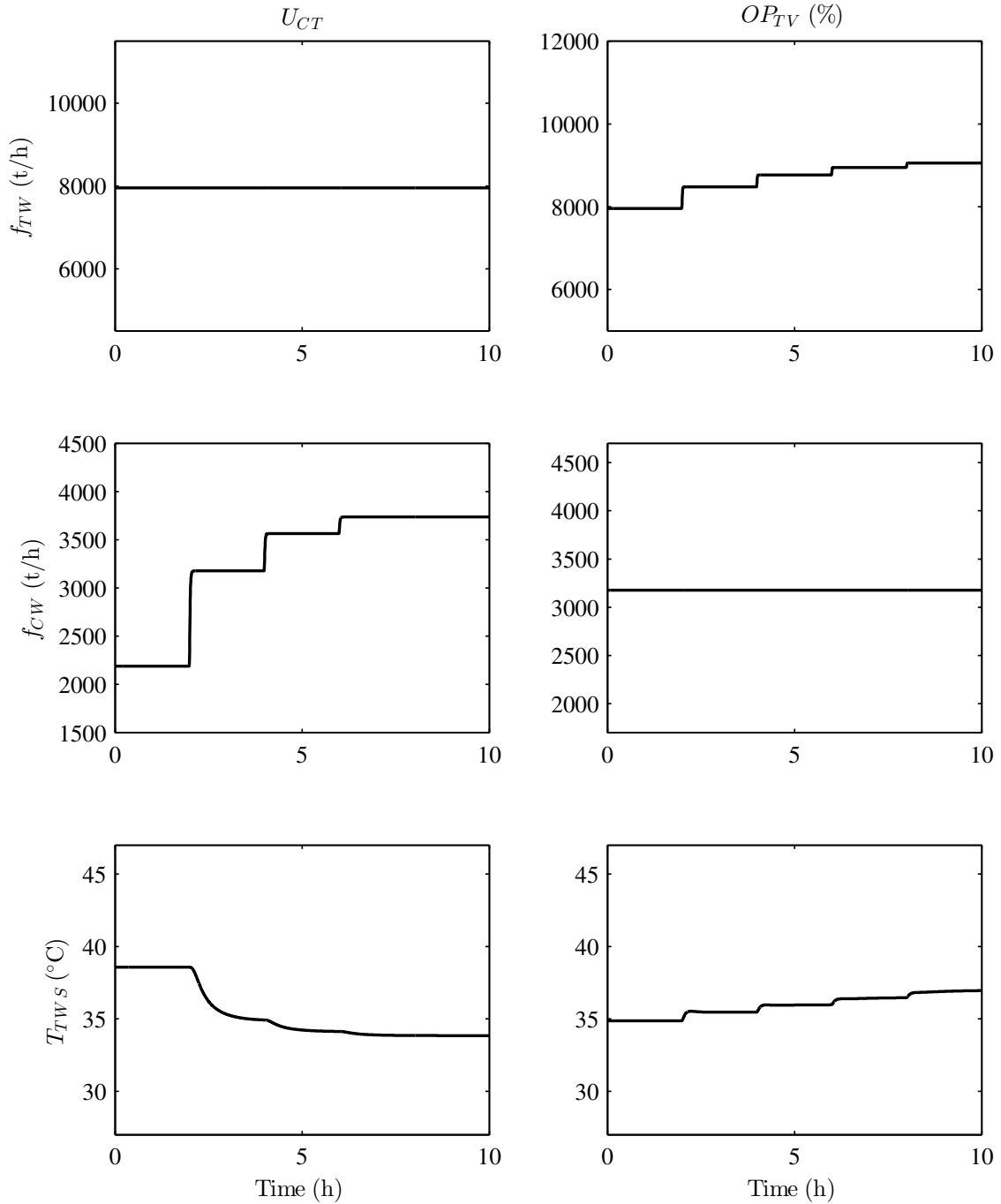


Figure 3.7: Responses in f_{TW} , f_{CW} , and $T_{TW,S}$ (represented by the model matrix rows) to changes in U_{CT} and OP_{TV} (represented by the columns).

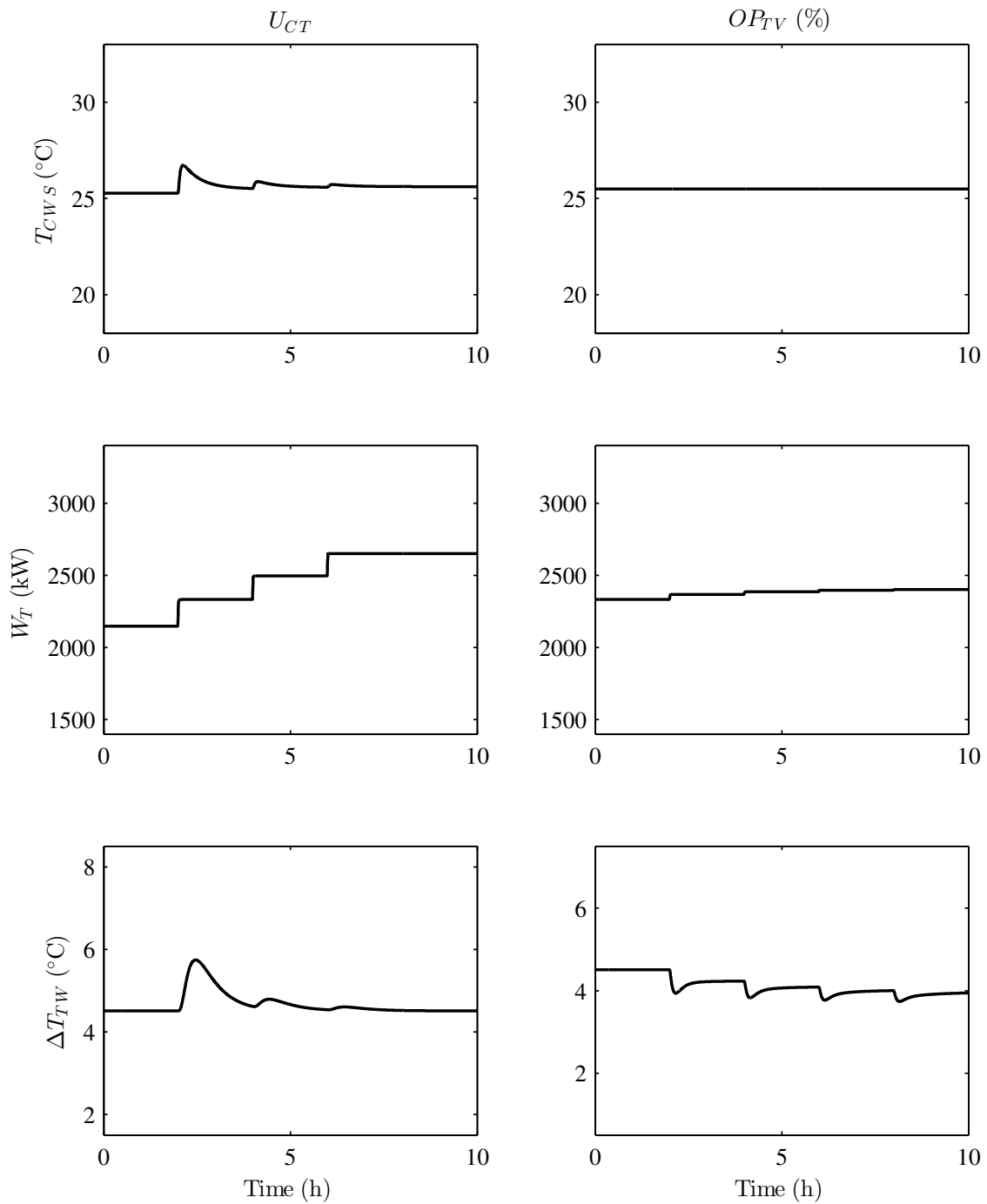


Figure 3.8: Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in U_{CT} and OP_{TV} (represented by the columns).

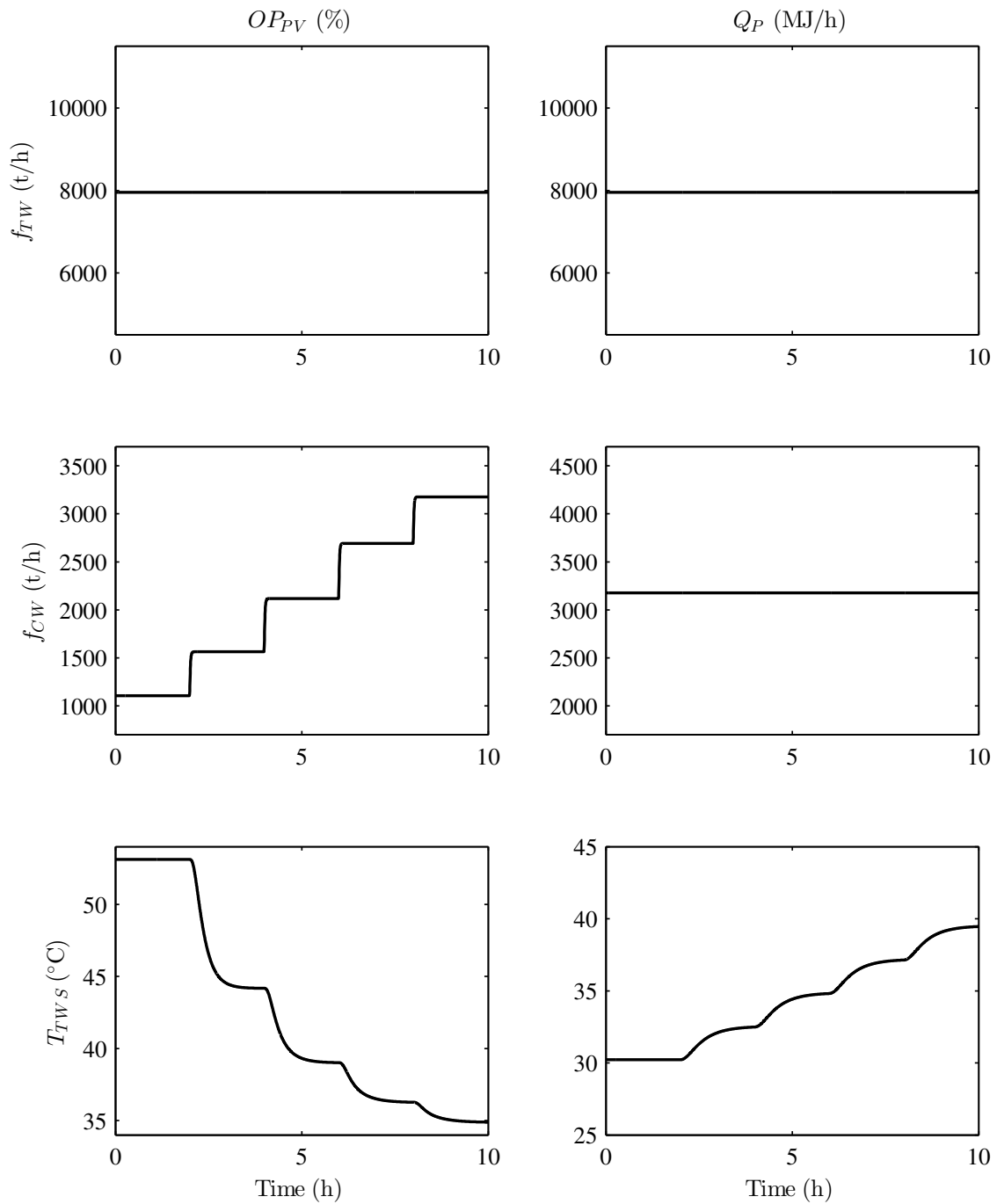


Figure 3.9: Responses in f_{TW} , f_{CW} , and $T_{TW,S}$ (represented by the model matrix rows) to changes in OP_{PV} and Q_P (represented by the columns).

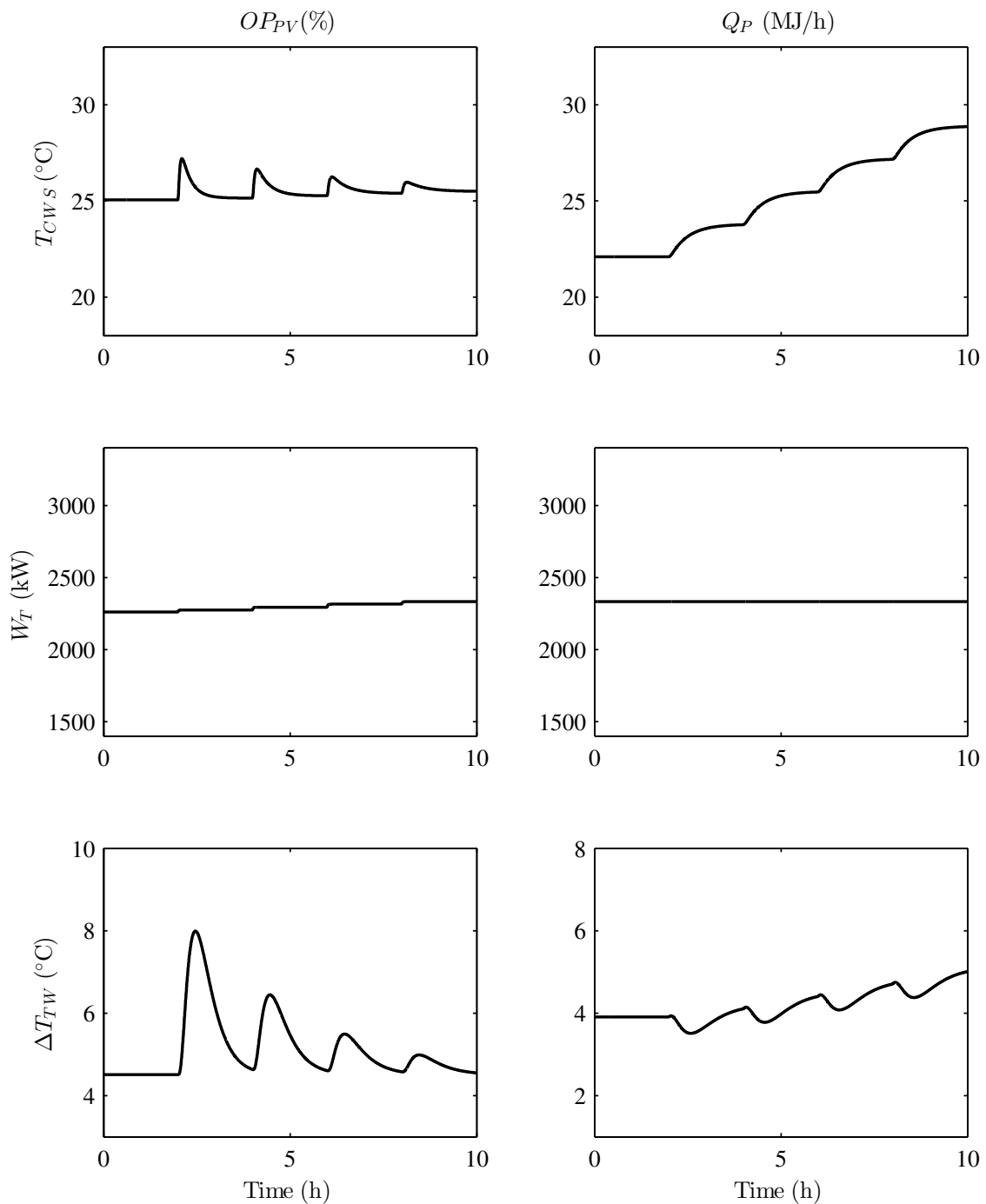


Figure 3.10: Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in OP_{PV} and Q_P (represented by the columns).

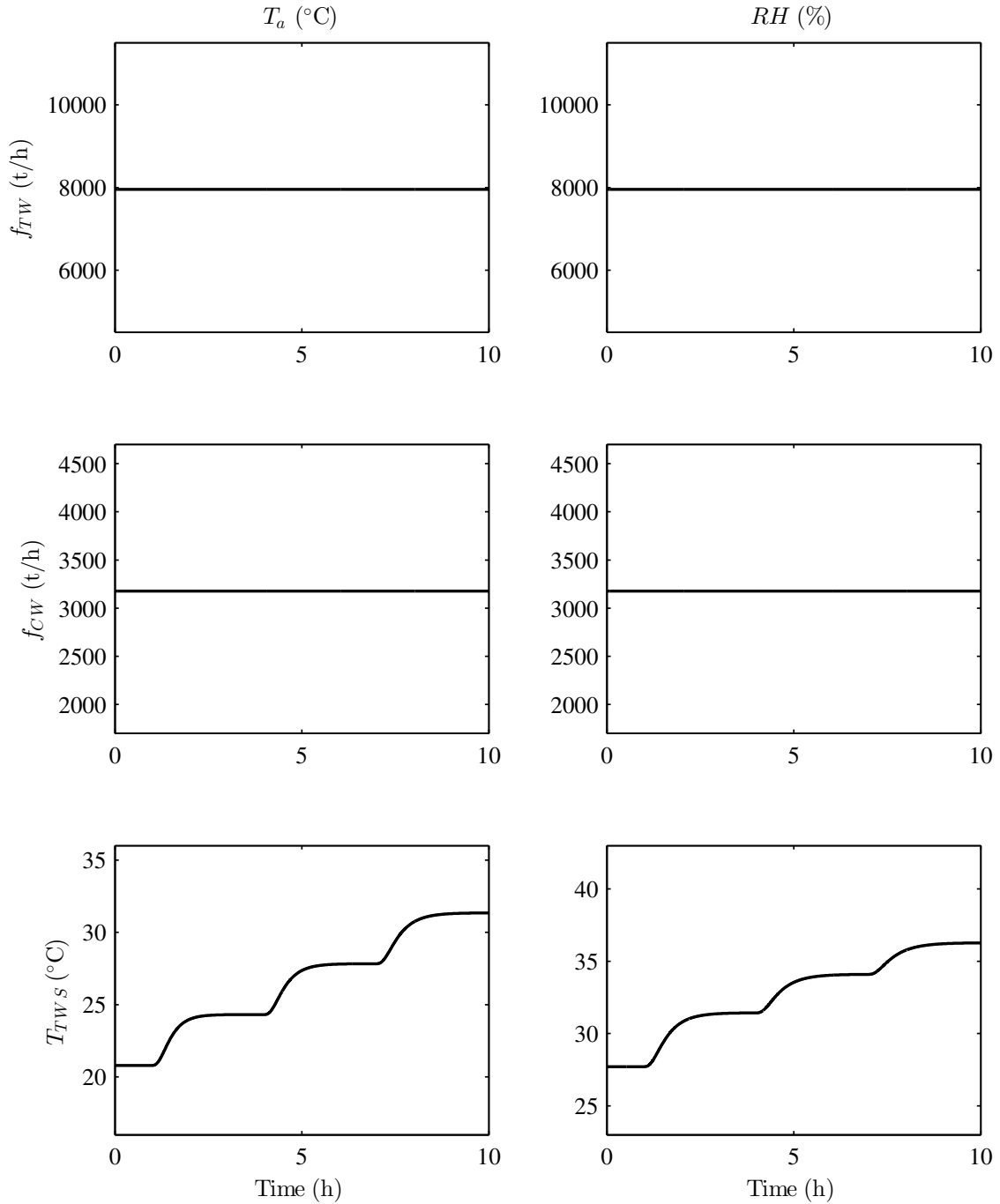


Figure 3.11: Responses in f_{TW} , f_{CW} , and T_{TWS} (represented by the model matrix rows) to changes in T_a and RH (represented by the columns).

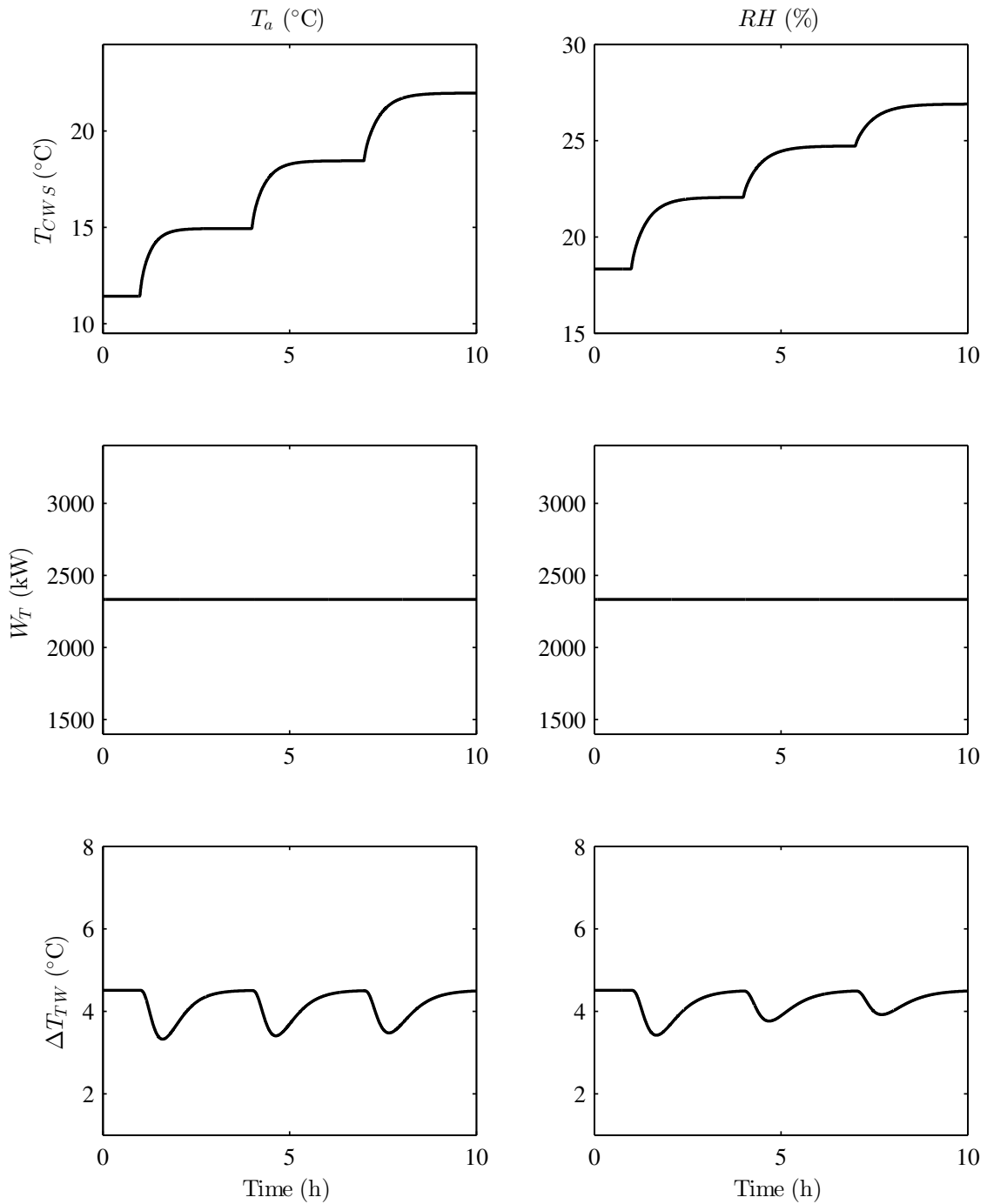


Figure 3.12: Responses in T_{CWS} , W_T , and ΔT_{TW} (represented by the model matrix rows) to changes in T_a and RH (represented by the columns).

up to five.

The cooling water pump running signals, U_{CW} , also exhibit a similar effect on the cooling water circuit flow rate as that of the tempered water pumps on the tempered water flow rate. The more pumps running, the smaller the progressive increase observed in flow rate (as seen in Figure 3.5). Also as expected, the number of cooling water pumps has no influence on the tempered water flow rate due to the physical segregation. There is a decrease in tempered water supply temperature, with an increase in cooling water pumps, as the increased cooling water flow provides for more effective removal of the energy from the tempered water circuit, as is expected. The cooling water supply temperature exhibits an interesting increase with an increase in running pumps (Figure 3.6). This is due to the increase in cooling water flow which results in the differential temperature of the cooling water reducing, similar to what is observed in the tempered water circuit. The total power increases in a similar non-linear fashion to that of the tempered water side (also refer to Figure 2.3). The tempered water differential temperature has a slight transient response with no steady-state gain (referred to as a zero-gain model). The number of pumps running were increased in increments of one, starting from one, going up to five.

The cooling tower fan running signals, U_{CT} , do not have an effect on the tempered water flow (Figure 3.7) as once again, the circuits are physically separated. The modelling was performed under the assumption that a tower is isolated from cooling water flow when the fan is not running (see Section 3.3.1). Therefore, when fewer cooling tower fans are running, fewer cooling towers are commissioned to receive cooling water flow, which increases the resistance of the circuit to flow. This is evident from the effect observed from the fan running signals to the cooling water flow. As more fans are switched on, the flow in the cooling water circuit increases as the system curve becomes flatter. Note that this is also a non-linear relationship, where the incremental increase in flow is smaller with each consecutive fan started/tower commissioned. As the system curve becomes progressively flatter, the operating point moves to the right and down on the combined pump curve, where the slope becomes steeper, resulting in the incremental decrease in pump discharge pressure becoming larger, with a smaller change in flow (refer to Figure 2.5). As expected, there is a decrease in tempered water supply temperature with an increase in cooling tower operation, though the effect is small compared to that of the number of tempered water and cooling water pumps. From Figure 3.8 it is again interesting to observe a slight increase in cooling water supply

temperature, which again is mainly attributed to the increase in cooling water flow in the circuit, resulting in a reduction in the cooling water differential temperature. The power in this case shows a linear increase due to the constant power consumption used for the fans, as opposed to a dynamic power curve, as is the case for the pumps. Again, the tempered water differential temperature only shows a slight transient with zero steady-state effect. The number of fans running were increased in increments of one starting from one going up to four.

Next, the tempered water temperature control valve opening (OP_{TV}) is considered. This variable was stepped in increments of 25% starting at zero going up to 100%. This valve bypasses the heat exchangers on the tempered water side and effectively reduces the total flow resistance in the tempered water circuit (flattens the system curve). This can be seen through the increase in tempered water flow as shown in Figure 3.7. Note that, although the valve has a linear characteristic, the installed characteristic exhibits non-linear behaviour. It would have been beneficial in this case to use a valve with an equal percentage characteristic to compensate for the non-linearity of the pumped system. Again, due to the circuit segregation, this valve does not affect the cooling water flow rate. A slight increase in tempered water supply temperature is observed, as expected, as bypassing the heat exchangers reduces the flow through them, reducing the energy transfer to the cooling water circuit. It is noteworthy that this effect is relatively small compared to that of the other manipulated variables, which puts into question the soundness of the original design, in which this valve is the only handle for manipulating the tempered water supply temperature. Similar to the tempered water pumps, this variable does not have a direct effect on the cooling water supply temperature as the duty is theoretically kept constant during the stepping of the valve opening (see Figure 3.8). There is a slight non-linear increase in total power consumption due to the increase in tempered water flow caused by the reduction in the system flow resistance (flattening of the system curve), which is also responsible for the observed reduction in tempered water differential temperature.

The cooling water pump discharge pressure control valve opening (OP_{PV}) was increased from 60% to 100% in increments of 10%. As these valves are situated on the cooling water circuit, there is no change in tempered water flow (see Figure 3.9). As can be expected, a very strong response is observed on the cooling water flow. The response is fairly linear, which is as a result of the equal percentage valve characteristic, which counters the non-linear pump

characteristic, resulting in a linear installed characteristic (see Section 2.7.4). This prominent response on the cooling water flow results in a strong response on the tempered water supply temperature. In fact, it is clear that the strongest handles for manipulating the tempered water supply temperature are the cooling water pump running signals and discharge control valves. Although the flow changes fairly linearly, the response on the temperature is non-linear. This is due to the reduction in cooling efficiency as the tempered water temperature approaches the cooling water temperature in the heat exchangers (refer to Equation 3.19). From Figure 3.10, a slight increase in cooling water supply temperature is observed, due to the reduction in cooling water differential temperature, caused by the increased cooling water flow. The increased flow also causes a slight increase in total power consumption. The tempered water differential temperature shows a transient increase with zero steady-state gain. The transient increase occurs due to the fact that, when the cooling water flow increases, the tempered water supply temperature decreases quickly, whereas the tempered water return temperature takes longer to change due to the plant dynamics.

3.3.5.2 Disturbance variable stepping

The plant duty (Q_P) was changed in steps of 10,000 from 130,000 to 170,000. From Figure 3.9, there is no change in either the tempered water or the cooling water flow rates with a change in duty, which is as expected. The duty change causes an increase in tempered water supply temperature, which is as expected, as well as a corresponding increase in cooling water supply temperature, as more energy is transferred to the cooling water circuit (see Figure 3.10). There is no change in the total power consumption as mass flow did not change. The tempered water differential temperature shows an interesting inverse response, with an eventual increase in differential temperature, caused by better heat transfer efficiency, when the tempered water circuit temperature is higher on average (the difference between the tempered water and cooling water temperatures is larger on average).

The next disturbance variable is the ambient temperature, T_a . Due to the increased time to steady-state for the ambient temperature, only three step changes were made as opposed to four. The ambient temperature was changed in increments of 6 starting from 12 up to 30. As shown in Figure 3.11, the ambient temperature does not affect the tempered water or cooling water flow rates directly. It does, however, have quite a significant effect on both the tempered

water supply temperature and the cooling water supply temperature (also see Figure 3.12). This is due to a reduction in the temperature difference between the cooling water return stream and the air flowing through the towers, which reduces cooling efficiency. There is no change in power consumption as the mass flow through the pumps did not change and no fans were switched. The tempered water differential temperature shows a negative transient response with zero steady-state gain. The negative transient is caused by the tempered water supply temperature increasing first (from the increase in cooling water supply temperature) before the tempered water return temperature eventually catches up.

Finally, the relative humidity, RH , was stepped in increments of 15 starting at 10 up to 55. As was the case with the ambient temperature, only three steps were made due to the long time to steady-state. Once again, from Figures 3.11 and 3.12, the flow rates in the circuits do not change, whereas the tempered water supply temperature and the cooling water supply temperature both increase. The increase in relative humidity causes an increase in wet-bulb temperature which means that the water absorbing potential of the air is reduced. This allows for less partial evaporation of cooling water, which reduces the cooling efficiency of the towers. The total power consumption remains constant and the tempered water differential temperature shows a negative transient response, with zero steady-state gain, for the same reasons mentioned above for the ambient temperature change.

In summary, the step test results confirm the non-linear interacting behaviour of the system. It appears that the strongest handle on the tempered water and cooling water temperatures is in fact the cooling water flow rate, which is most effectively manipulated using the cooling water pump running signals, U_{CW} , and pressure control valve opening, OP_{PV} . It further shows the notable disturbance effects that the plant duty, ambient temperature and relative humidity have on the system. It also confirms that the tempered water flow rate is the strongest handle for manipulating the tempered water differential temperature (the higher the flow rate the lower the differential temperature), which is most effectively manipulated using the tempered water pump running signals, U_{TW} .

3.4 CONTROL AND OPTIMISATION

This section illustrates how modern control and optimisation techniques may be applied to the system in order to optimise on the energy consumption and/or cost of operation, while

simultaneously maintaining the process within predefined limits.

Four cases are considered here. The first is a base case to illustrate unoptimised operation, where little or no intervention occurs in terms of optimising the system. In the second case, Advanced Regulatory Control (ARC) and conditional switching logic are implemented, with no specialised software required, other than the plant's existing control system (such as a Distributed Control System (DCS) or Programmable Logic Controller (PLC)) [68]. In the third case, Model Predictive Control (MPC) is applied as an Advanced Process Control (APC) technique, where a process model is required and a cost function is set up to minimise energy consumption. Due to the hybrid and non-linear nature of the system, as well as the fact that the system is not transformed or linearised for control purposes, the solution is referred to as Hybrid Non-linear MPC (HNMPC). The fourth case builds on the HNMPC case and includes Time-of-Use (TOU) tariffs for which the cost function is adapted to not optimise on energy consumption but rather on energy cost (referred to as Economic HNMPC, or EHNMPC). Cases three and four typically require the use of a dedicated server class machine and specialised software and therefore may require additional capital spending in order to implement these control solutions.

The manipulated variables (MVs, also referred to as handles or process inputs), disturbance variables (DVs, also referred to as feed-forward variables) and controlled variables (CVs, or process outputs) are given in Section 3.3.1. The T_{TWS} and ΔT_{TW} must be maintained within boundaries whereas the power consumption (and cost in the fourth case) must be minimised. As mentioned, the constraints on the controlled variables are

- $26^{\circ}\text{C} \leq T_{TWS} \leq 36^{\circ}\text{C}$ and
- $\Delta T_{TW} \leq 6^{\circ}\text{C}$.

For each case, two simulation studies are presented. The first simulation study covers an operating period of 7 days, using artificial data during which step-like and ramp-like changes are made to the plant load and sinusoidal disturbances are introduced on the humidity and ambient temperature, which both affect the wet-bulb temperature (and consequently the cooling efficiency of the cooling towers). The plant duty and wet-bulb temperature are shown in Figure 3.13. The plant duty is smoothed with a first order filter with a time constant of 30 minutes to more realistically mimic real plant rate changes. The sinusoidal changes in the

ambient temperature is on a 24-hour period with an amplitude of 5°C and baseline of 25°C to mimic diurnal temperature changes. The sinusoidal change in the relative humidity has a period of 7 days with an amplitude of 15% and a baseline of 35% to imitate a gradual change in the relative humidity over the course of a week.

For the second simulation, the plant data used for the model verification (as shown in Figure 3.3) was used as input data to illustrate how the control schemes could perform under a real plant operating scenario. The second simulation therefore runs over a period of 6 days (144 hours).

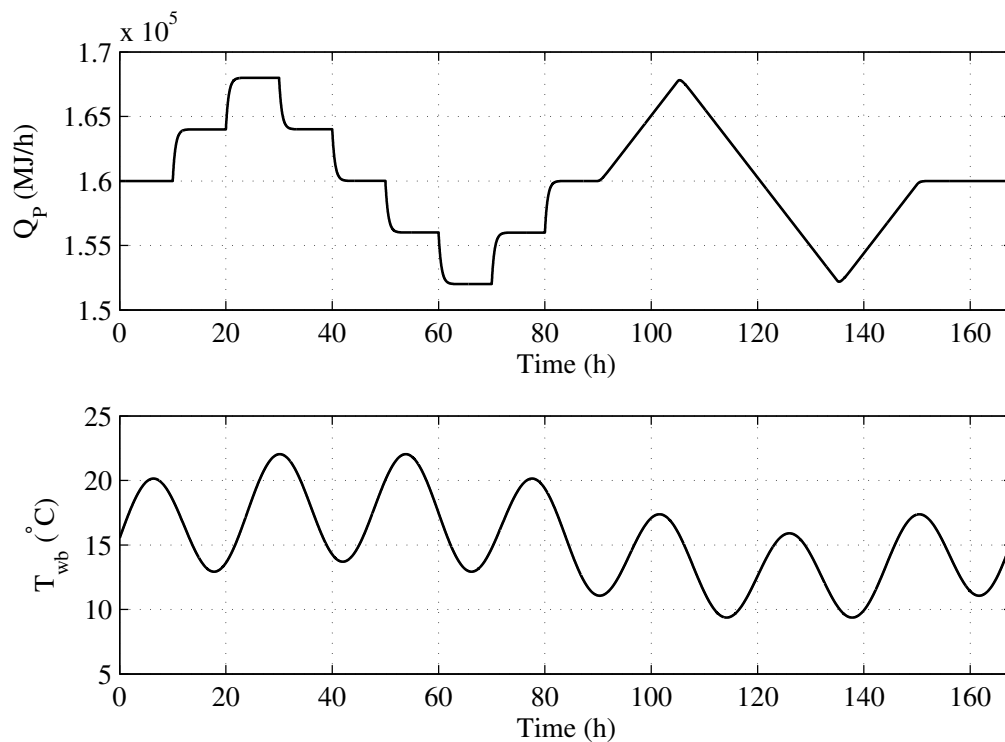


Figure 3.13: Model disturbance inputs for first simulation (plant duty and wet-bulb temperature).

3.4.1 Base case

To illustrate the possible benefit of optimising the cooling water system, a base case is established to represent the unoptimised operation of the plant, where little or no action is taken by operations personnel in terms of optimising the system. For the purposes of this study, the assumption is that one pump on each bank and one cooling tower is used as a spare/standby

and therefore four pumps are running on each bank with three cooling towers running at any time. The only automatic control active during this case is the temperature controller that bypasses the heat exchangers (TIC-101) with a set-point of 26°C. The cooling water pump discharge pressure controller valve openings (PIC-201 to 205 in Figures 3.1 and 3.2) are at 100% for the cooling water pumps that are running, and zero for those that are not running. No switching is performed on the pumps and cooling towers during the period of simulation. The simulation results are discussed in Section 4.2.1.

3.4.2 Advanced regulatory control

As discussed in Section 2.3.2, Advanced Regulatory Control (ARC) refers to a set of control techniques that are implementable on most modern control systems without the need to purchase additional software or hardware, but which typically go beyond the basic regulatory control provided by Proportional Integral Derivative (PID) controllers [32, 39]. In this case, these include cascade control, override selector control, and custom control algorithms in the form of conditional switching logic.

In comparison to the base case, some optimisation is performed by switching pumps and cooling tower fans on/off depending on the operating conditions and equipment availability. The temperature controller on the tempered water side is still operated in the same automatic mode with the same set-point value at 26°C.

The control scheme for this case is illustrated in Figure 3.14. The valves on the cooling water pump discharge lines are used for total cooling water flow control (FIC-101), replacing the pressure controllers (PIC-201 to 205) of the original set-up. For constraint handling on the tempered water supply temperature, two temperature controllers were added in a mid-of-three² override selector configuration to the flow controller desired set-point, f_{CW}^{SP} . This allows the flow set-point to be set by the temperature controllers, should a violation of the desired temperature range limits occur, while still allowing the operator to specify a set-point for normal conditions. During the temperature violation events, the flow set-point will be allowed to deviate from the desired set value (which was set at 3000 t/h during normal operation).

²The mid-of-three override selector selects the middle value of its three inputs as its output.

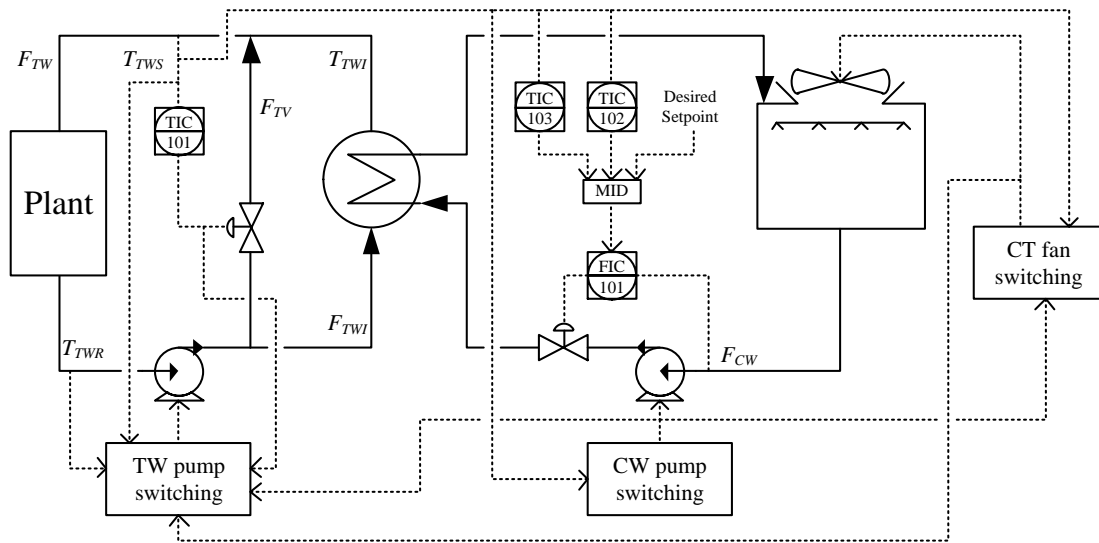


Figure 3.14: ARC scheme illustration.

The switching logic allows for the use of the discrete devices in the system as additional manipulated variables and therefore provides additional degrees of freedom. Based on the operational objective to minimise energy consumption, the switching logic is designed to minimise the number of equipment in a running state. This is done based on the specified process constraints and equipment limits, and are prioritised in terms of switching order.

In the plant-wide control framework [32, 33, 35, 36, 37], the operational objective is energy minimisation which is achieved by using the discrete and continuous control handles (the pumps, fans and valves) subject to the constraints on the tempered water supply and differential temperatures (the controlled variables) as well as the limits on the number of running equipment. The switching logic forms an integrated optimisation and control layer. The optimisation/control time scale separation is achieved through only performing equipment switching actions at defined intervals. Another perspective on the switching logic is that it provides a way to reconfigure the system when the active constraints move.

The generic approach for designing the switching logic is as follows (taking into account the steps for control structure design as suggested in [32], [33], [35], [36] and [37]):

1. Determine the operational objective that represents optimal operation (the objective function).

2. Establish the controlled variables for the process and their constraints.
3. Establish which discrete variables will be included in the switching scheme.
4. Determine the relationships between the discrete input variables and the controlled variables (for possible input/output pairing) and the objective function.
5. Determine the preferred switching order of the discrete input variables, especially where more than one discrete variable affects a particular CV.
6. Determine the limits on the discrete and continuous input variables. These may have to be re-evaluated continuously to account for changes in equipment availability.
7. Determine the desired default switching frequency (the amount of time a switching conditions must be active before a switching event will occur). Some conditions may warrant shorter delays or even immediate switching.
8. Determine the appropriate dead-bands for returning to range of the controlled variables to avoid chattering switching behaviour.
9. Formulate the switching rules for each discrete variable based on the information gathered above.

In some cases the system may be simple such as a bank of fin-fans used to control the pressure or temperature on a distillation column. In other cases (such as this one) the problem may be more complicated and include multiple discrete variables and/or process variables.

Following this approach for the cooling water system, the conditions for the switching logic are as follows:

1. The operational objective is to minimise energy consumption (the objective function to be minimised is the total energy consumption).
2. The controlled variables for the process are T_{TWS} and ΔT_{TW} and their constraints are as listed earlier in this section.
3. All the discrete variables will be included in the switching scheme (U_{CW} , U_{TW} and

U_{CT}).

4. All of the discrete variables affect T_{TWS} , whereas only U_{TW} has a considerable effect on ΔT_{TW} . All the discrete variables affect the power consumption. The nature/direction of these effects can be determined from Figures 3.5 to 3.8 and/or Section 3.3.2.
5. For T_{TWS} , the switching order is first U_{CW} , then U_{CT} and lastly U_{TW} . This is based on the relative effectiveness of each of the discrete variables on manipulating the T_{TWS} .
6. The limits on the discrete variable values are $2 \leq U_{CW} \leq 5$, $2 \leq U_{TW} \leq 5$ and $1 \leq U_{CT} \leq 4$. The value of OP_{TV} should remain below 95%. In this case, changes in equipment availability are not taken into account.
7. If a condition is active for 30 minutes or more on any of the variables, a switching action will be initiated.
8. The dead-band for returning to range on T_{TWS} is 2 °C and 1.75 °C for ΔT_{TW} .

The switching rules for each of the discrete variables are then as follows:

- Cooling water pump switching:
 - If ($T_{TWS} \leq 34$) and ($U_{CW} > 2$) for more than 30 minutes, switch a cooling water pump off.
 - If ($T_{TWS} > 36$) and ($U_{CW} < 5$) for more than 30 minutes, switch a cooling water pump on.
- Cooling tower fan switching:
 - If ($T_{TWS} > 36$) and ($U_{CW} = 5$) and ($U_{CT} < 4$) for more than 30 minutes, switch a cooling tower fan on.
 - If ($T_{TWS} \leq 34$) and ($U_{CW} = 2$) and ($U_{CT} > 1$) for more than 30 minutes, switch a cooling tower fan off.
- Tempered water pump switching:

- If $(\Delta T_{TW} < 4.25)$ and $(U_{TW} > 2)$ and $(T_{TWS} > 28)$ for more than 30 minutes, switch a tempered water pump off.
- If $(\Delta T_{TW} \geq 6)$ and $(U_{TW} < 5)$ for more than 30 minutes, switch a tempered water pump on.
- If $(T_{TWS} \leq 26)$ and $(U_{CW} = 2)$ and $(U_{TW} < 5)$ and $(OP_{TV} \geq 95\%)$ for more than 30 minutes, switch a tempered water pump on.

Note that the switching logic is only one of the several ARC techniques that comprise the complete ARC control scheme. The continuous control operates on a different (faster) time scale than the discrete switching logic.

The simulation results for the ARC case are discussed in Section 4.2.2.

3.4.3 Hybrid non-linear model predictive control

As discussed in Section 2.3.4, Model Predictive Control (MPC) is a form of Advanced Process Control (APC), where a model of the process is utilised to predict the future trajectory of the plant (based on past and present data). It then uses an optimiser to determine the control moves that will drive the plant to the desired operating point, in an optimal way, where optimality is defined by the formulation of a performance function. The process is repeated at every iteration after receiving the latest plant measurements. Only the first of the calculated control moves are implemented at each iteration, which is referred to as a receding horizon approach [6, 18, 32, 45, 69].

In the case of the cooling water system described in Section 3.2, the process is non-linear and of a hybrid nature (see Section 2.6), which requires the use of an optimiser that is capable of handling the non-linearities as well as the discontinuities caused by the discrete input variables³.

In this case, a genetic algorithm (GA) is used as the optimisation engine due to its ability to accommodate non-linear, mixed integer, constrained, multi-variable, interactive problems

³The alternative is to transform the system and linearise it into a form where the conventional linear continuous control techniques may be applied.

conveniently; and its robustness against converging on a local optimum (see Section 2.4.1). As discussed in Section 2.4.1, the genetic algorithm is a type of evolutionary algorithm (EA) where the processes of nature are used as inspiration for mathematical algorithms. The genetic algorithm evaluates a set of candidate solutions (individuals) based on the predefined objective function (referred to as the fitness function). The set of individuals (current population) is referred to as a generation. The fittest individuals of the current generation are then predominantly used to reproduce and form the next generation with some mutation to add randomness to the process. This continues until the maximum number of generations is reached or some tolerance between the average fitness and the fitness of the fittest individual is reached [21, 54, 55, 56, 57].

As with the parameter estimation discussed in Section 3.3.3, the existing Matlab `ga` function was used. For this part of the study, a population size of 10 with a maximum number of generations of 15 were chosen based on the typical convergence rate observed per iteration. For constraint handling, the algorithm minimizes a penalty function as opposed to the fitness function in which the penalty function includes a term for infeasibility. This is then combined with binary tournament selection which scrutinises pairing and selection based on the feasibility of the individuals [59]. Due to the mixed integer nature of the optimisation problem, specialised crossover and mutation functions are used together with a truncation procedure for integer restriction as described in [58].

The nature of this solution can be classified as hybrid non-linear model predictive control (HNMPC). The control scheme is shown in Figure 3.15.

The sizing parameters for the HNMPC problem are (apart from the GA parameters mentioned above) the iteration period ($T_{iter} = 30$ minutes or 0.5 hours), the prediction horizon ($n = 12$), the control horizon ($m = 4$), CV constraints, MV constraints and fitness function weights.

The fitness function used for the optimisation is a weighted sum of the squared constraint violations on the T_{TWS} and ΔT_{TW} (forming the quadratic part of the function) and the power consumption, W_T , (which forms the linear part of the function) over the prediction horizon period. To allow for wind-up detection for the flow controller (FIC-101), the valve opening value of the cooling water pump discharge valves is also included as a controlled variable,

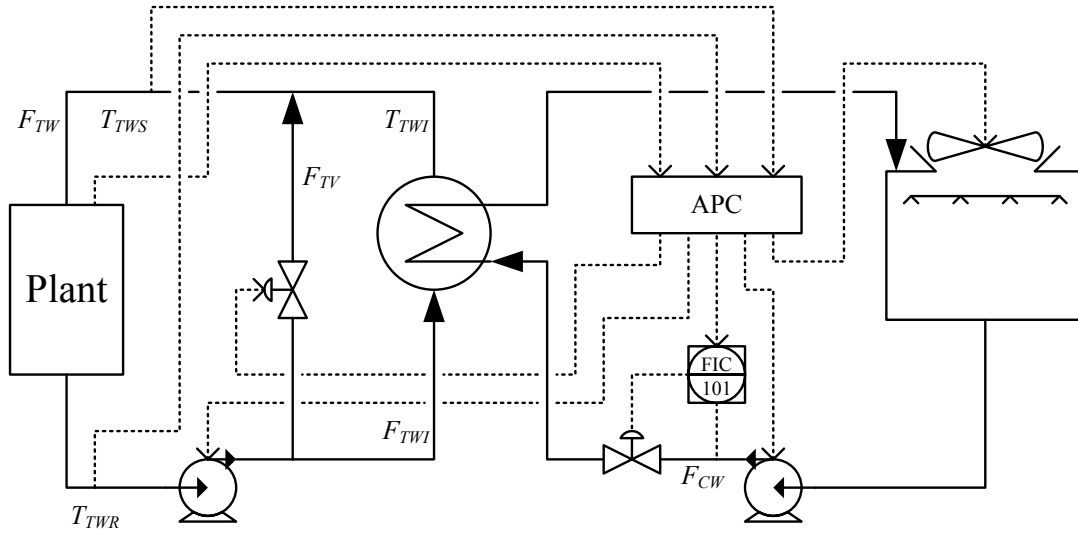


Figure 3.15: HNMPC scheme illustration.

which only comes into play when the flow controller set-point value and plant conditions result in the valves opening fully. The fitness function is as follows:

$$J = \frac{1}{j} \sum_{i=1}^j (Q_{T_{TWS}} E_{T_{TWS},i}^2 + Q_{\Delta T_{TW}} E_{\Delta T_{TW},i}^2 + Q_W W_{T,i} + Q_{PV} E_{OP_{PV},i}) \quad (3.55)$$

where j is the number of samples in the prediction period (the solver uses a sampling time of 30 seconds (1/120 hours) resulting in $j = 720$ over the 6 hour prediction time), $Q_{T_{TWS}}$, $Q_{\Delta T_{TW}}$, Q_W and Q_{PV} are the weighting variables and $E_{T_{TWS}}$, $E_{\Delta T_{TW}}$ and $E_{OP_{PV}}$ are the constraint violations for T_{TWS} , ΔT_{TW} and OP_{PV} (which are zero when operating within the constraints). The weighting values used in the simulations are $Q_{T_{TWS}} = 20$, $Q_{\Delta T_{TW}} = 15$, $Q_W = 10$ and $Q_{PV} = 100$.

The manipulated variable and controlled variable limits are shown in Table 3.4. Note that a safety margin is built into the controlled variable limits on the tempered water supply temperature and differential temperature, and that the maximum set-point for the cooling water flow is 4500 t/h, which was determined to be approximately the maximum achievable flow rate on the cooling water circuit with the maximum number of pumps and towers running. The power has no constraints and is used for optimisation only.

Table 3.4: CV and MV limits

Variable	Low limit	High limit
T_{TWS}	28	34
ΔT_{TW}	0	5.5
U_{CT}	1	4
U_{CW}	2	5
U_{TW}	2	5
f_{CW}^{SP}	500	4500
OP_{TV}	0	100

The optimisation algorithm uses the manipulated variable limits as hard constraints. Therefore, it will not move the manipulated variables outside these limits under any circumstances. The controlled variable limits are only used in the fitness function and are therefore considered soft limits. If the algorithm is unable to determine a set of manipulated variable values that maintain all controlled variables within limits, it will simply result in a poorly performing system, but will not cause infeasible solutions. This makes the optimisation very robust.

The optimisation problem is performed over the prediction horizon period (the product of the iteration time and the prediction horizon) which is 6 hours. This 6 hour simulation is run 10 times for each generation with 15 generations per iteration (therefore 150 times per iteration). With an iteration time of 30 minutes on a simulation period of 168 hours, the model is evaluated a total of 50,400 times over the course of the first simulation. Despite the large number of model evaluations, the solving time is still only a fraction of the iteration time⁴.

The simulation results for the HNMPC case are presented in Section 4.2.3.

⁴On a 2011 model Apple MacBook Pro with dual core 2.7 GHz Intel i7 processor and 8 GB RAM, the average solving time is approximately 90 seconds. The iteration time of the controller is 30 minutes. Therefore, even with modest computing power, the solution is realistically implementable

3.4.4 Economic hybrid non-linear model predictive control

The HNMPC solution illustrated above, optimises the total energy usage. With the assumption of a constant energy cost, this also translates into an optimisation in total energy cost. When Time-of-Use (TOU) rates and/or maximum demand concepts are applicable, the approach has to be changed if energy cost optimisation is to be achieved. Therefore, a trade-off now exists between total energy usage and total energy cost. This falls in the energy management category of Demand Side Management (DSM). The TOU rates used for this study is related to the actual rates applicable in South-Africa at the time of writing and are defined as:

$$p(t) = \begin{cases} p_o & \text{for } t \in [0, 6) \cup [22, 24) \\ p_s & \text{for } t \in [6, 7) \cup [10, 18) \cup [20, 22) \\ p_p & \text{for } t \in [7, 10) \cup [18, 20) \end{cases} \quad (3.56)$$

where t is the time of day in hours, $p_o = 0.4031$ R/kWh is the rate for off-peak time periods, $p_s = 0.7423$ R/kWh is the rate for standard hours and $p_p = 2.4503$ R/kWh is the rate at peak times. These rates are for the high demand season which is from June to August. The rates are different for the low demand season (September to May). R is the symbol for the South African Rand (currency). Figure 2.1 shows a graphical representation of Equation 3.56.

Due to the fact that the electricity rates are known exactly beforehand, they are treated differently to the disturbance variables, where the latest measurement is taken to be the constant value of that variable over the prediction horizon. Therefore, the energy cost optimisation is not only considered as a steady-state economic optimisation, but a truly integrated dynamic economic optimisation, which puts it in the class of Economic MPC [46]. Therefore, this case is called Economic Hybrid Non-linear MPC (EHN MPC). This control scheme is also illustrated by Figure 3.15, and the variable limits shown in Table 3.4 are still valid.

The fitness function for the HNMPC case above (Equation 3.55) is adjusted to include an energy cost term rather than the energy consumption:

$$J = \frac{1}{j} \sum_{i=1}^j (Q_{TWS} E_{TWS,i}^2 + Q_{\Delta TW} E_{\Delta TW,i}^2 + Q_C C_{T,i} + Q_{PV} E_{OPV,i}) \quad (3.57)$$

where C_T is the total electricity cost (as given in Equation 3.31), $Q_C = 4$ is the weighting value for the energy cost component and $p(i)$ is the energy tariff at sampling instant i (obtained from Equation 3.56 for the corresponding time of day).

The simulation results for the EHN MPC case are presented in Section 4.2.4.

3.5 CONCLUSION

This chapter provided a detailed description of the process and covered the methodologies for performing the system modelling, simulation and controller design. The next chapter provides the results obtained from these modelling and simulation studies.

CHAPTER 4

SIMULATION RESULTS AND DISCUSSION

4.1 CHAPTER OVERVIEW

In this chapter, the results from the simulation studies discussed in Chapter 3 are presented. The control and optimisation simulation studies are analysed to compare the performance of the ARC and APC cases to that of the base case.

4.2 CONTROL AND OPTIMISATION RESULTS

As discussed in Section 3.4, four cases were evaluated through simulation studies to illustrate their relative abilities to minimise energy consumption/cost while maintaining process operation within the predefined boundaries. Each case was subjected to two operating scenarios. For the first scenario, artificial disturbance data in the form of step, ramp and sinusoidal changes are used, as shown in Figure 3.13. For the second scenario, the actual plant data, used for the model validation, was utilised (see Figure 3.3). The results for each of the cases are presented below where-after a detailed comparison is presented.

4.2.1 Base case

Figure 4.1 shows the controlled variables for the first simulation. The tempered water supply temperature remains within limits through most of the simulation period, without the need to perform much control action, due to the fact that the average plant load and wet-bulb temperature are sufficiently high to benefit from the high flow rates in the circuits. As the simulation progresses, the wet-bulb temperature drops, which causes an overall drop in

tempered water supply temperature due to the increased cooling efficiency of the towers. At around 63 hours, the temperature cycle is clipped by the temperature controller which manages to prevent violation of the bottom constraint. This is also the case for the two subsequent cycles. At around 133 hours, the temperature control valve is fully open and a constraint violation occurs for about 8 hours when the temperature drops well below the lower limit. This illustrates the limited capability of the control scheme. The tempered water differential temperature is kept below the upper limit, not through active control, but rather by the conservative operation of four running tempered water pumps, which results in an unnecessarily large tempered water flow rate. Therefore, no optimisation is performed to drive towards the most economical operating point. The total power remains high throughout the simulation with four pumps on each circuit and three cooling tower fans in operation. The occasional increase in power observed occurs when the temperature control valve opens, which flattens the system curve, resulting in a slightly higher total flow through the tempered water pumps.

Figure 4.2 shows the controlled variables for the second simulation using actual plant data. In this case, although the average plant duty is slightly higher than in the first simulation, the wet-bulb temperature is about 10°C lower than the first simulation, which means that the cooling ability of the towers is significantly higher. This results in a much lower tempered water supply temperature; the temperature controller is unable to keep the temperature within limits and is fully open for most of the simulation. Furthermore, due to the fact that a simple PID controller with a fixed set-point is used, the controller still clips the cycles when they reach the lower limits, which prevents the temperature from moving further into the safety of its range, but rather results in the temperature remaining at the lower limit value. Once again the tempered water differential temperature is kept below the upper limit through the conservative operation, as four tempered water pumps are still used, which results in an unnecessarily large tempered water flow rate. No optimisation is performed and the total power remains high throughout the simulation, with an occasional decrease in power when the temperature control valve cuts back, resulting in a slightly lower total flow through the tempered water pumps.

Therefore, the base case is very limited, not only in terms of energy optimisation, but also in performing basic temperature control.

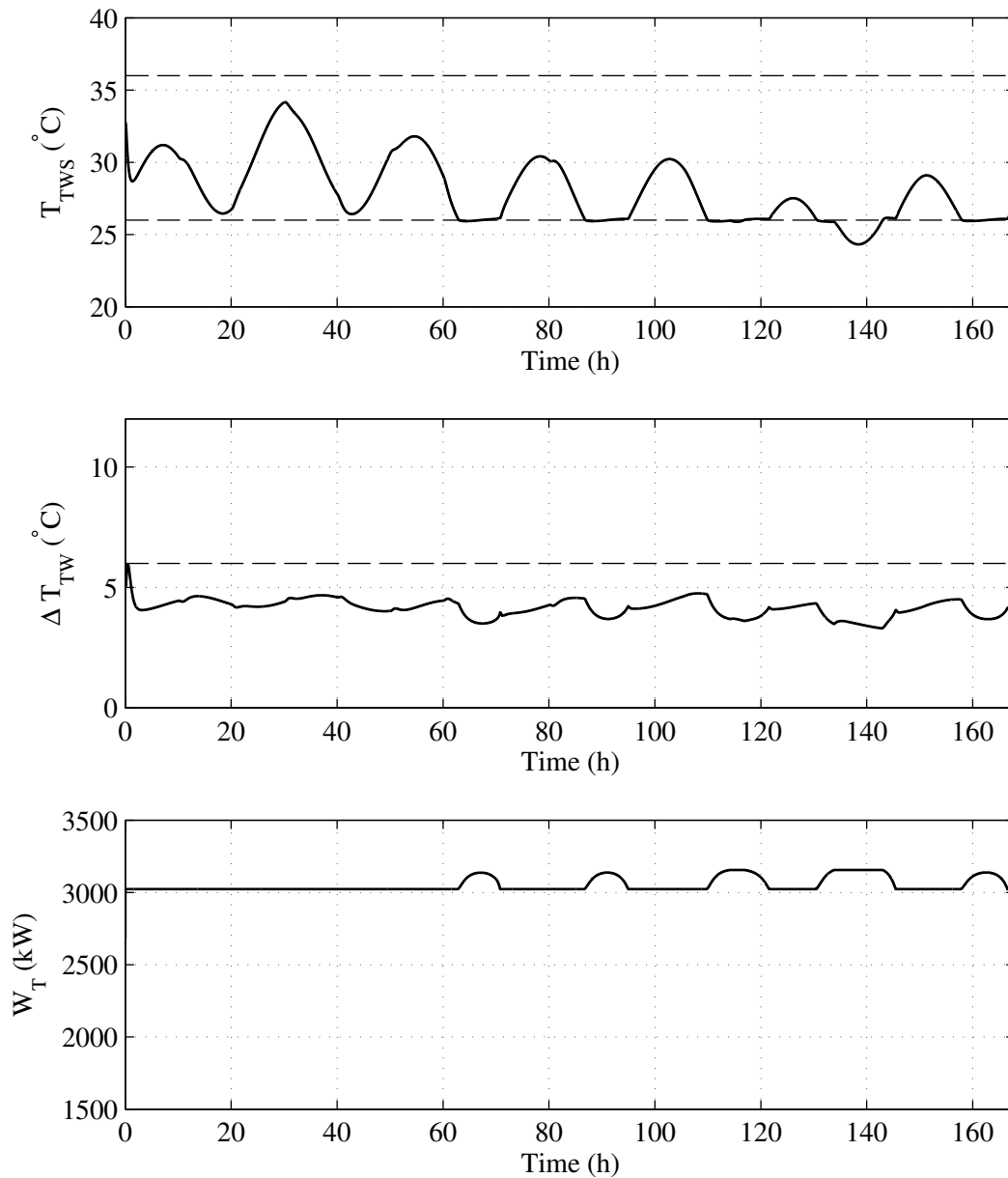


Figure 4.1: Controlled variables (Base case) for the first simulation.

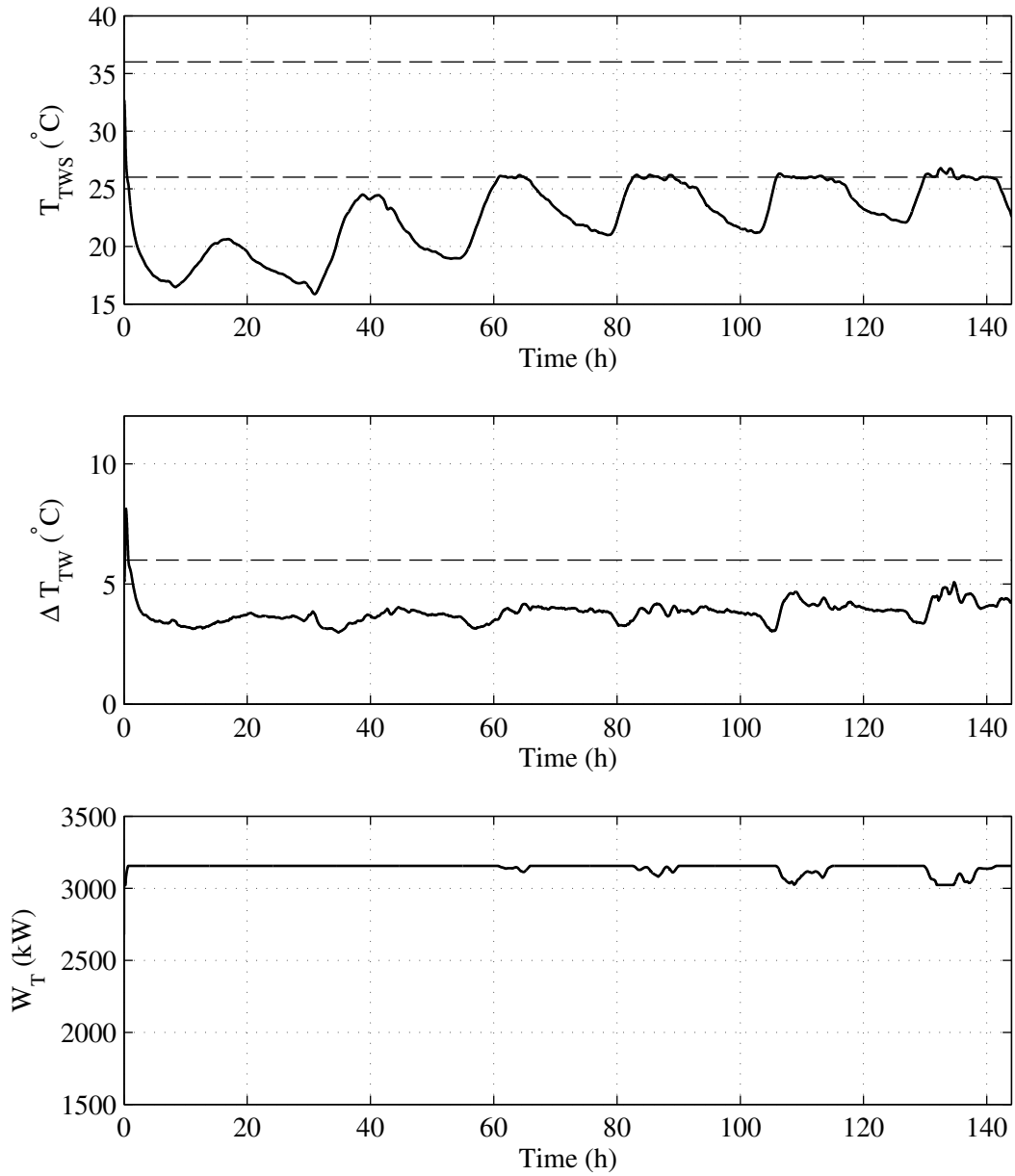


Figure 4.2: Controlled variables (Base case) for the second simulation.

4.2.2 ARC results

As discussed in Section 3.4.2, the ARC case builds on the base case in several ways, including: the use of the cooling water pump discharge pressure valves for flow control, rather than just under-pressure control; the addition of high and low temperature controllers with override selector control to additionally manipulate the cooling water flow rate when one of the constraints are approached; and the addition of switching logic to automatically switch pumps and fans on or off.

Figure 4.3 shows the controlled variables for the first simulation for the ARC case. The tempered water supply temperature is mostly maintained within constraints apart from some small short duration violations. The tempered water differential temperature is pushed toward the upper constraint and is mostly maintained below the upper limit. The largest difference observed, is in the power consumption, which is significantly reduced compared to the base case, due to the ability to switch pumps and fans off when the cooling capacity is not limited. Although not strictly considered a controlled variable, the cooling water pump discharge valve opening is also shown to illustrate when this indirect handle is wound up and when it is available to the flow and temperature controllers. Figure 4.4 shows the corresponding manipulated variables for the first simulation. It is clear that the number of cooling tower fans, cooling water pumps and tempered water pumps are actively pushed towards the minimum allowed values in order to minimise energy usage and are only increased when a process variable is about to violate a constraint. The tempered water supply temperature control valve remains closed during the entire simulation as the tempered water supply temperature is maintained close to the upper bound (the more economical area) whereas the temperature controller set-point is set equal to the lower limit. The last graph in Figure 4.4 shows the cooling water flow controller measured value as the solid line and the set-point as the dashed line. This illustrates that although a set-point of 3000 t/h is requested most of the time, it is typically not achieved when a low number of cooling water pumps and cooling towers are running. This corresponds to the valves saturating at 100% as shown in Figure 4.3. During the few instances where more cooling water pumps and cooling tower fans are running, the flow increases to a value where set-point tracking is achievable.

Figure 4.5 shows the controlled variables for the second simulation for the ARC case. The tempered water supply temperature is maintained within constraints much more effectively

than the base case with negligible constraint violations. The same can be said for the tempered water differential temperature, which is also driven towards the upper constraint when possible. The power consumption is lower for most of the simulation, resulting in an overall reduction in energy use. The tempered water temperature control valve is in operation during the first part of the simulation, allowing for its use in controlling the tempered water supply temperature at the lower bound. Figure 4.6 shows the corresponding manipulated variables. It is evident that very little switching is done on the tempered water pumps during the first almost 35 hours of the simulation. Referring back to the tempered water supply temperature plot in Figure 4.5, it indicates that the temperature is quickly brought to the lower limit at the start of the simulation and is then kept very close to this limit by the temperature controller. The result is that the temperature does not cross the dead-band setting of the switching logic and consequently switching does not occur until the dead-band is exceeded around 35 hours into the simulation. This presents a limitation for the optimisation of the ARC case as the dead-band is required to avoid chattering switching behaviour. Nevertheless, the results are much more satisfactory than that of the base case both in terms of energy reduction and in honouring process constraints. From the last graph in Figure 4.6 it is again evident that the cooling water flow controller can only track set-point (dashed line) as long as the valves are not saturated. Both cases illustrate that the application of advanced regulatory techniques results in a tremendous improvement in energy usage while still honouring constraints better than the base case.

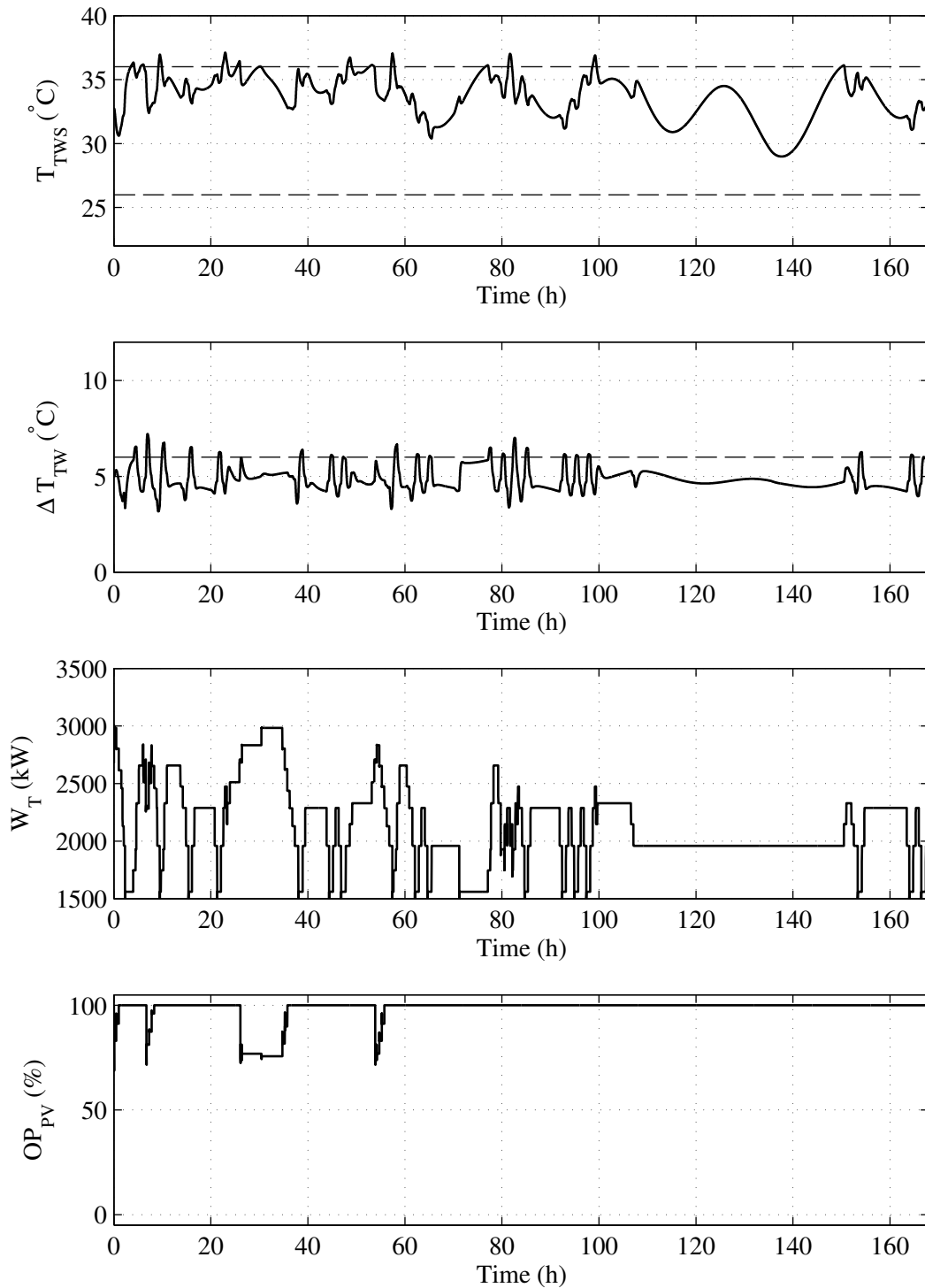


Figure 4.3: Controlled variables (ARC case) for the first simulation.

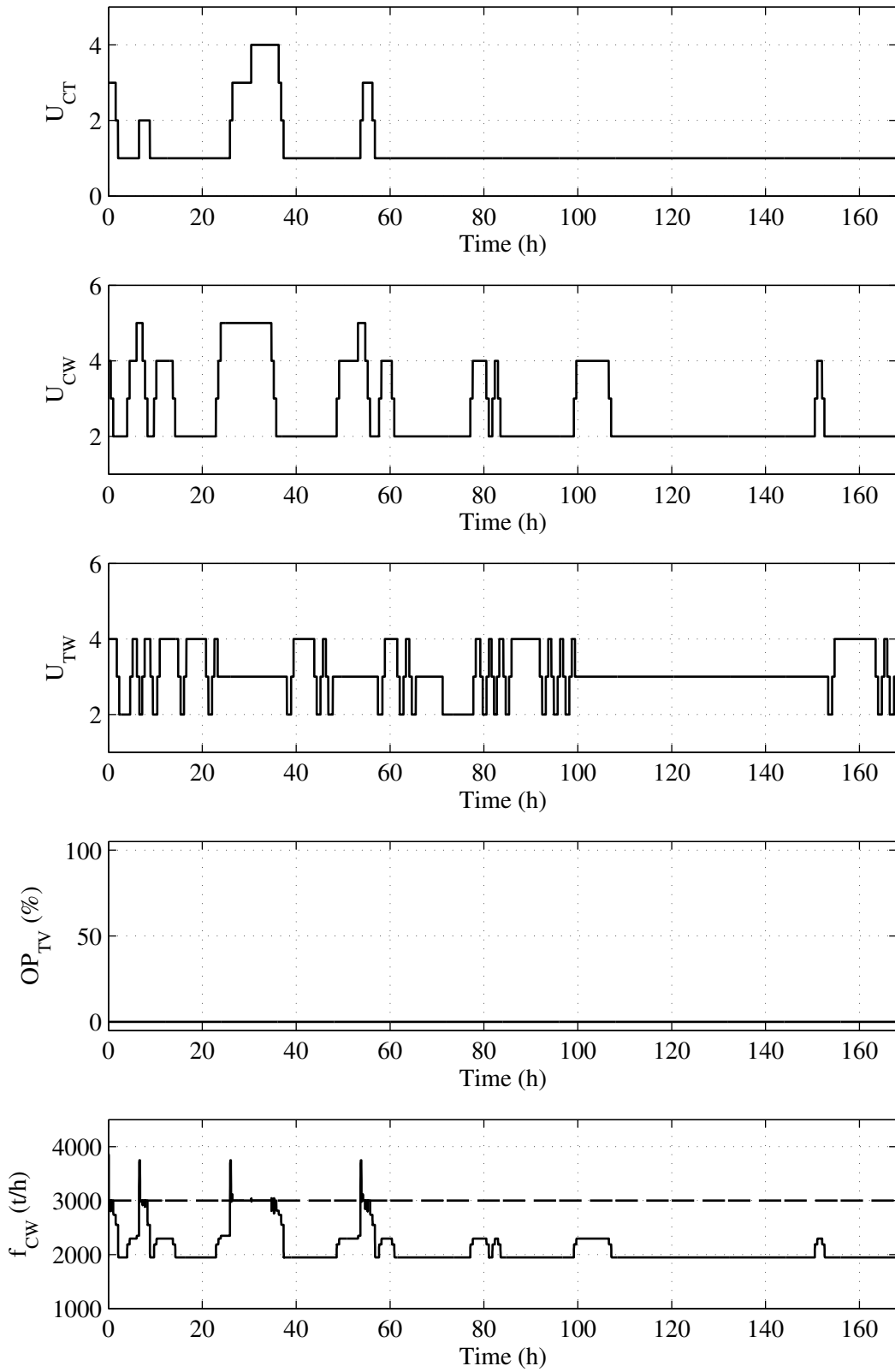


Figure 4.4: Manipulated variables (ARC case) for the first simulation.

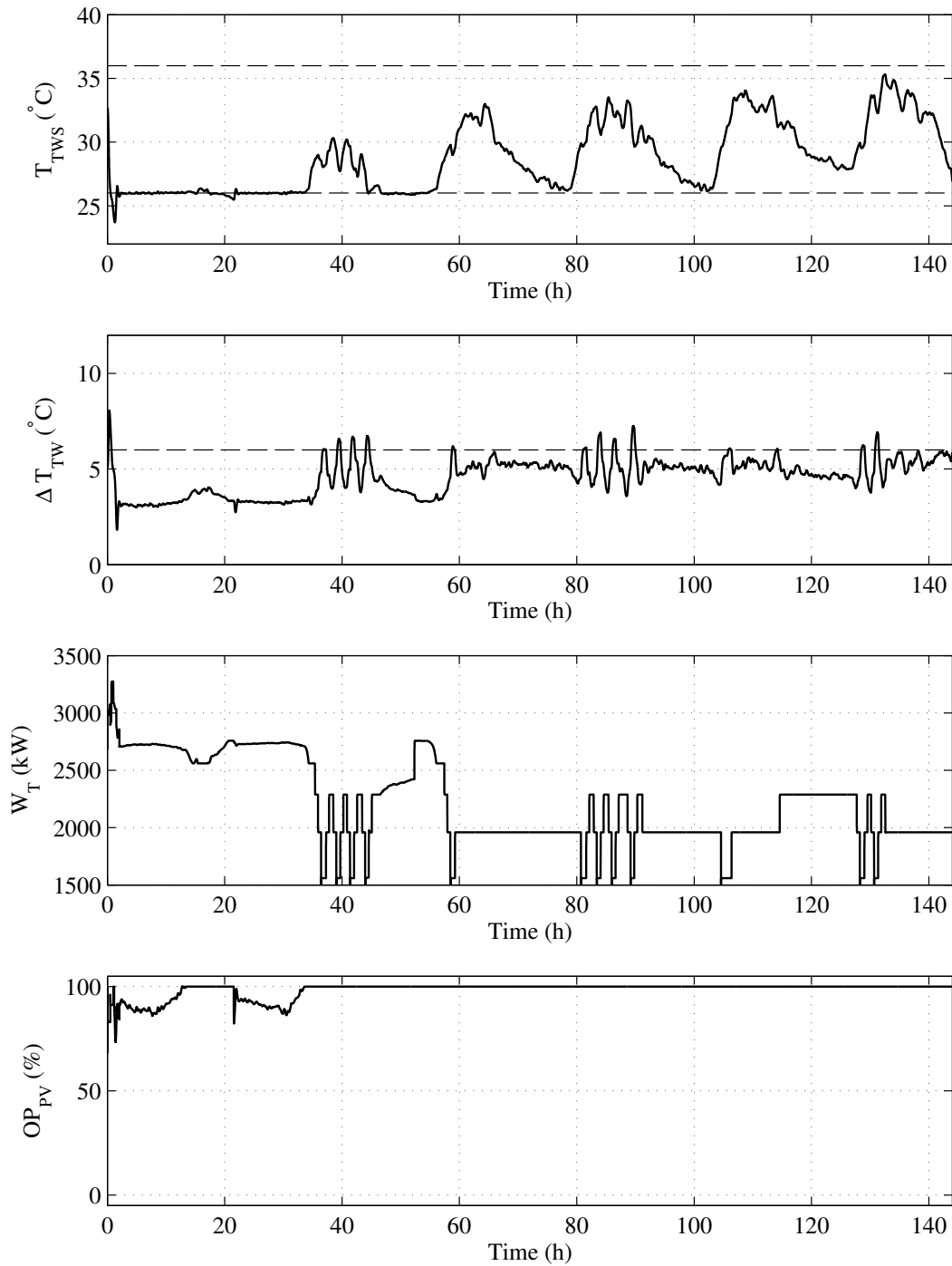


Figure 4.5: Controlled variables (ARC case) for the second simulation.

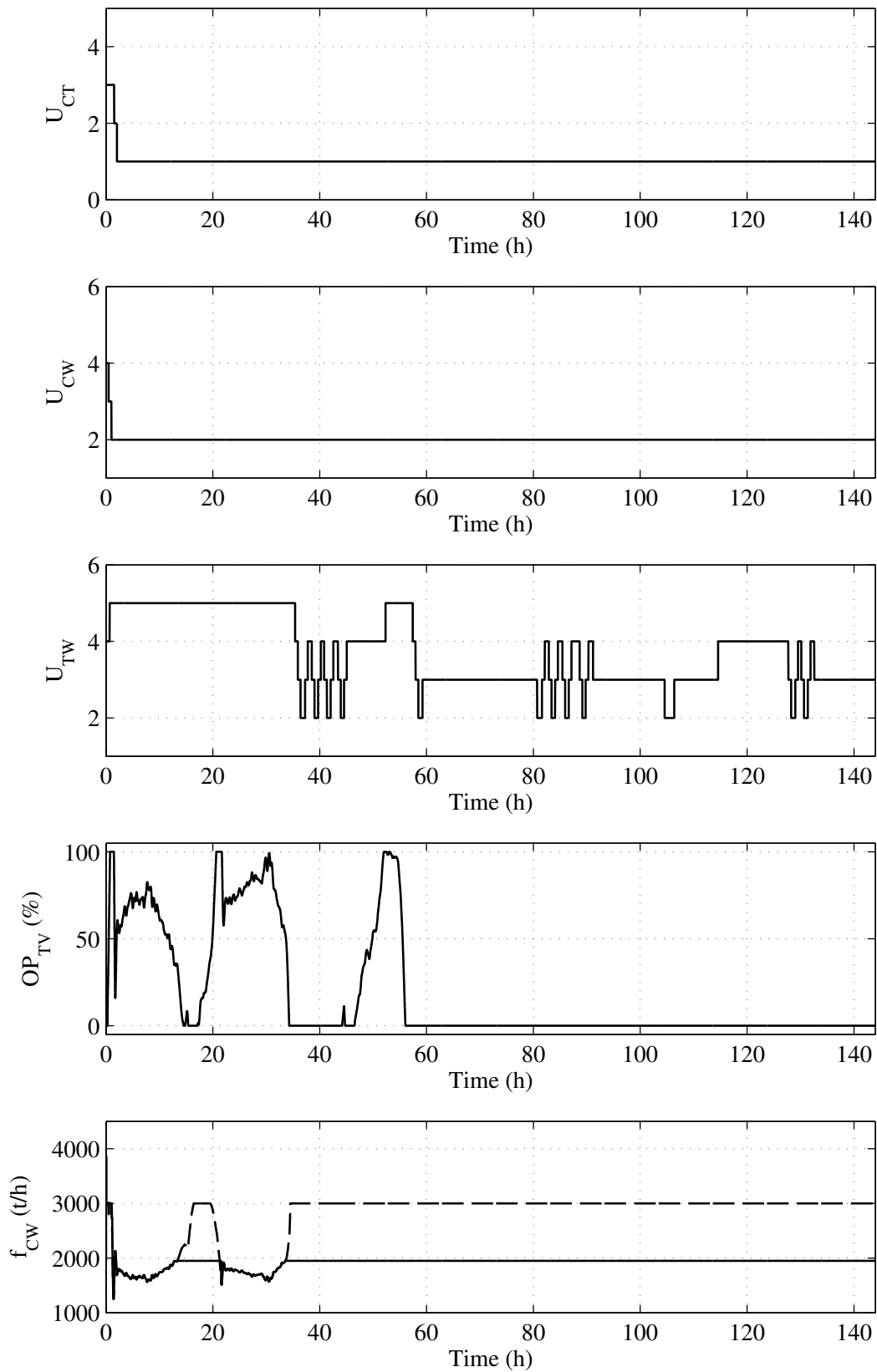


Figure 4.6: Manipulated variables (ARC case) for the second simulation.

4.2.3 HNMPC results

As discussed in Section 3.4.3, the hybrid non-linear model predictive control (HNMPC) case departs from the conventional base layer control and enters the domain of multi-variable control. Furthermore, in this case the hybrid and non-linear nature of the problem are directly dealt with in the HNMPC, which is beyond what is mostly encountered in industrial MPC applications.

Figure 4.7 presents the controlled variable values for the HNMPC for the first simulation case. The tempered water supply temperature is maintained within range with negligible constraint violation (superior to both the base case and ARC case). The tempered water differential temperature is driven toward the upper limit also with negligible constraint violation. The total power consumption is managed dynamically to produce a lower total energy usage than both previous cases. Whereas in the ARC case the cooling water pump discharge valve opening was included purely as an indication of controller wind-up, in this case it is included in the formulation of the performance function with a penalty only being incurred when the valve opening reaches 100%. In this way, the HNMPC knows that it should not increase the cooling water flow controller set-point further if the valve is already saturated at 100%. From the last graph in Figure 4.7 it is clear that this objective is achieved and the valve is prevented from saturating. Figure 4.8 shows the corresponding manipulated variable values. The superior control and optimisation results are achieved with visibly fewer switching activities compared to that of the ARC case. The pumps (where most of the energy is consumed) are actively driven to the minimum allowable numbers while still being able to run the maximum number of pumps when required to avoid constraint violation. The controller is capable of switching multiple devices on a bank in the same execution cycle compared to the sequential switching of the ARC case. This allows for superior disturbance rejection capability. The temperature control valve opening and cooling water set-point values (which are the continuous handles of the controller as opposed to the discrete running signals) are actively being manipulated to the desired values, with the flow controller set-point only being increased within the bounds of its valve saturation limits. This provides a means of dynamic MV limit adjustment.

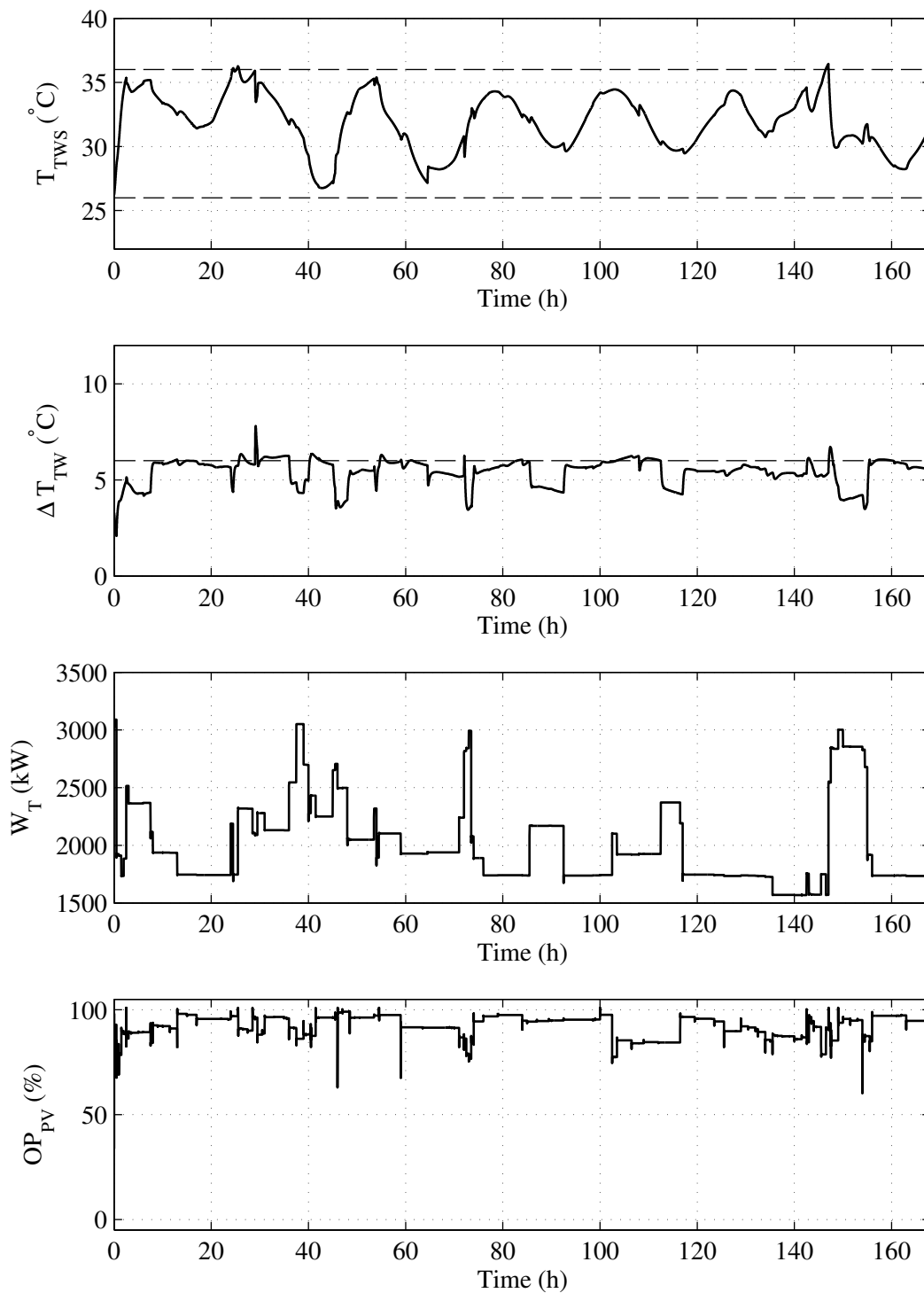


Figure 4.7: Controlled variables (HNMPC case) for the first simulation.

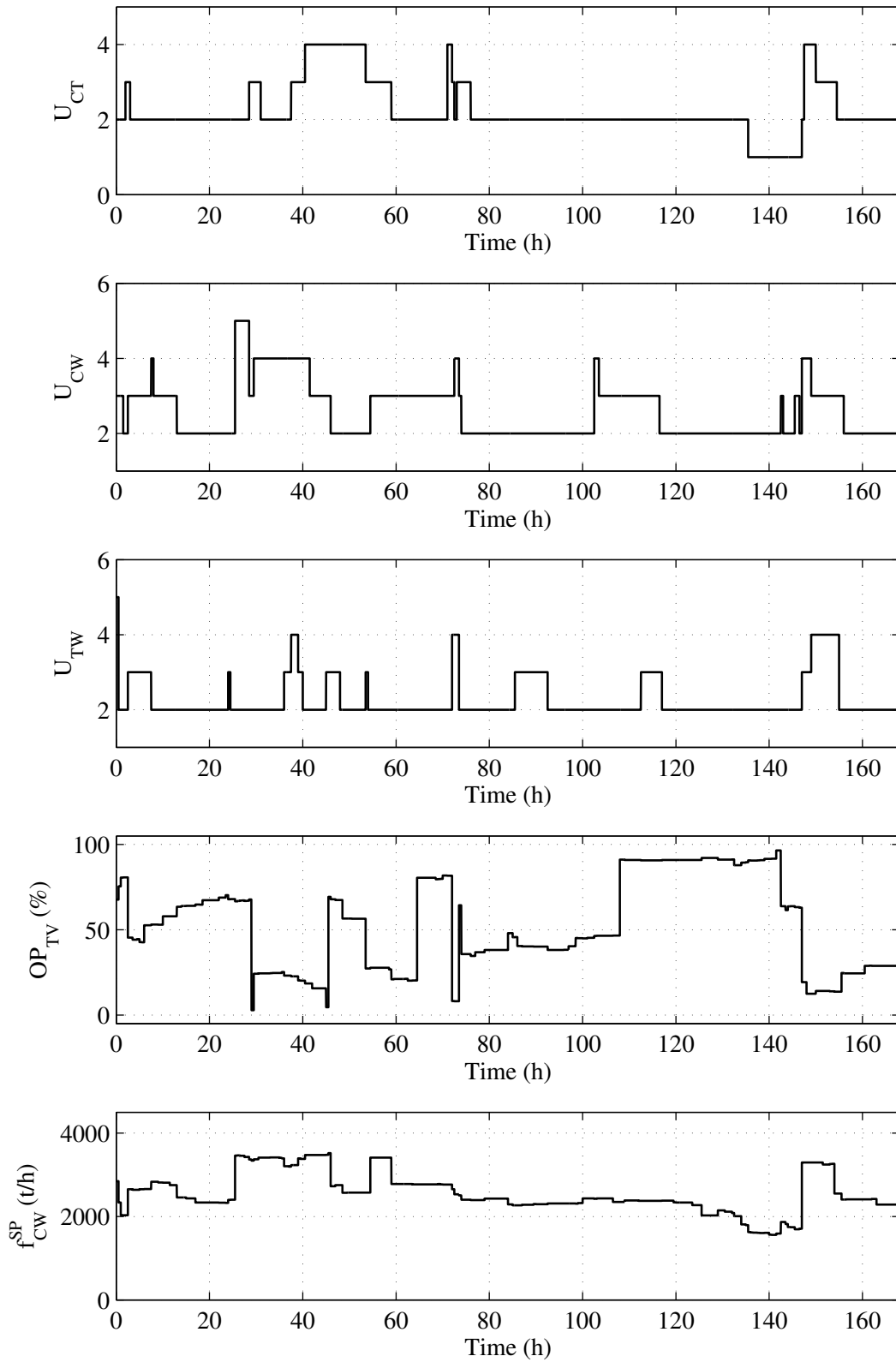


Figure 4.8: Manipulated variables (HNMPD case) for the first simulation.

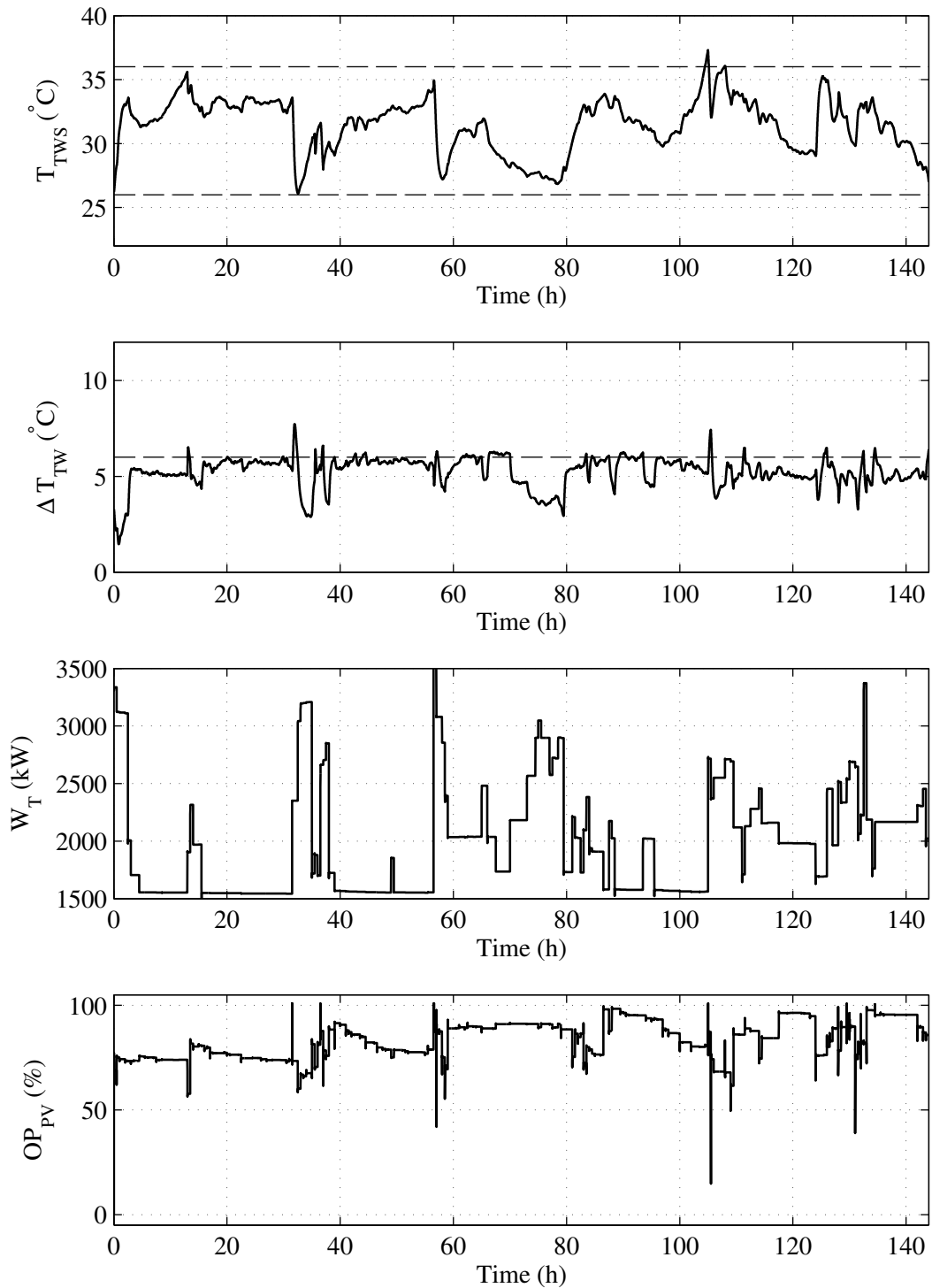


Figure 4.9: Controlled variables (HNMPC case) for the second simulation.

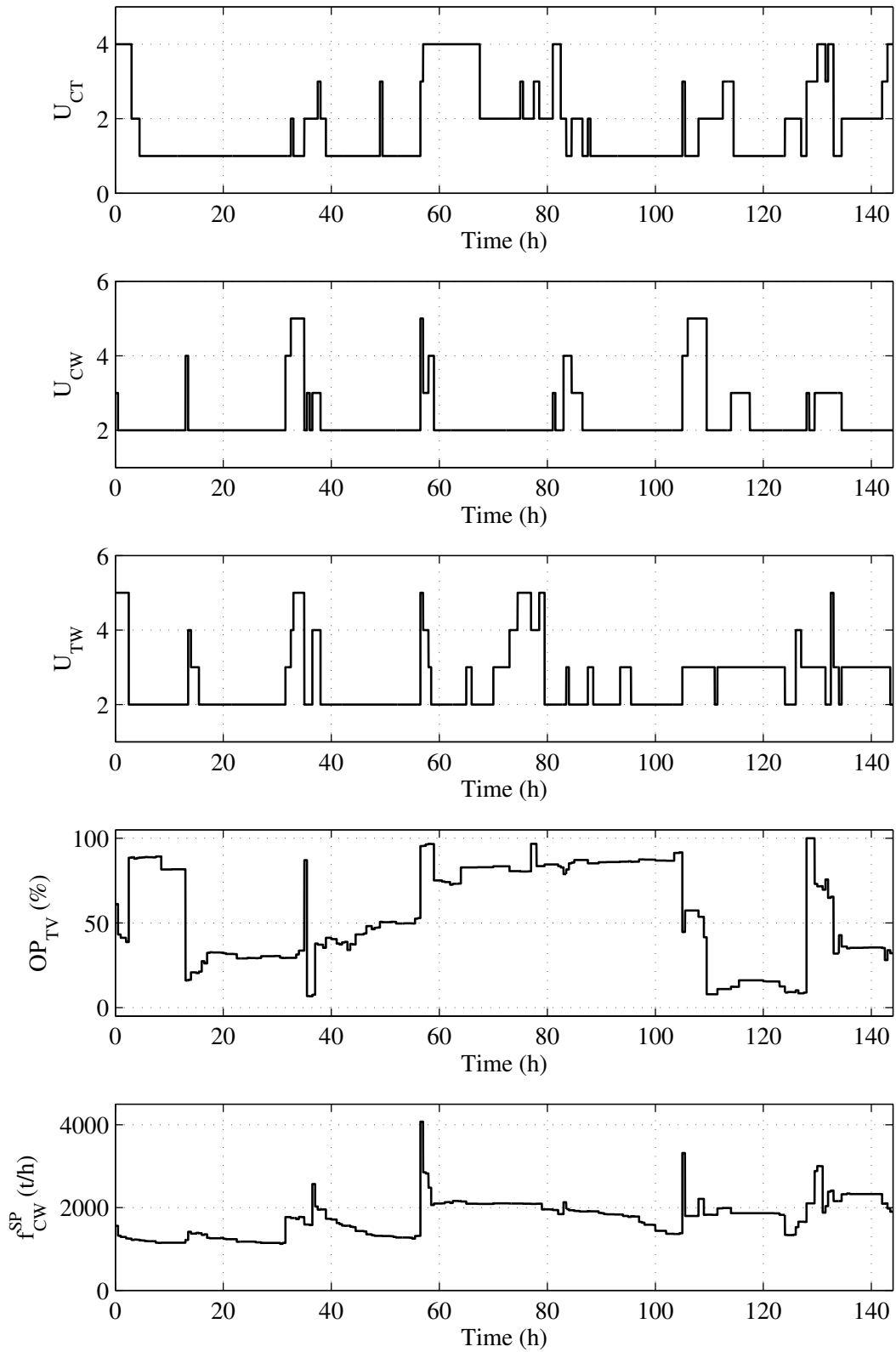


Figure 4.10: Manipulated variables (HN MPC case) for the second simulation.

Figure 4.9 presents the controlled variable values for the HNMPC for the second simulation case. The tempered water supply temperature is again maintained within limits with negligible constraint violations, which was a challenge for the base and ARC cases. The tempered water differential temperature is also being driven to the upper constraint, which allows for fewer tempered water pumps to be running. The total power consumption is actively minimised to provide much lower total consumption than the previous cases. The cooling water pump discharge valve opening is again prevented from saturating at 100% which provides wind-up detection for the flow controller set-point manipulated variable. The manipulated variable values for this case are shown in Figure 4.10. The pump and fan running signals are minimised whenever the process constraints allow it and the continuous handles are actively utilised. Notice how the cooling water flow controller only increases significantly when the number of running cooling water pumps and cooling tower fans are increased, which opens up room on the control valves.

Both simulations indicate that the HNMPC solution provides a further reduction in energy while also having superior constraint handling compared to the base and ARC cases.

4.2.4 Economic HNMPC results

As mentioned in Section 3.4.4, the Economic HNMPC case goes one step further by removing the assumption that the electricity cost is a constant and uses the time-of-use electricity rates to perform dynamic energy cost optimisation as opposed to simply minimising consumption. Therefore, if the main goal is to save cost, this scheme provides additional saving potential whereas if the main goal is to reduce consumption, the HNMPC case may be sufficient.

Figure 4.11 displays the controlled variable values for the EHN MPC case for the first simulation with Figure 4.12 showing the corresponding manipulated variable values. Figures 4.13 and 4.14 show the values for the second simulation. The constraint handling performance is comparable to that of the HNMPC case and so is its use of the manipulated variables. On face value the two cases seem to be performing similarly. From the calculated performance results presented in the next section, it is however clear that the goal of further reducing energy cost is indeed achieved by the EHN MPC in both simulation cases.

In addition, Figures 4.15 and 4.16 indicate the power consumption together with the electricity cost and illustrate how the EHN MPC controller cuts back consumption during high cost (peak) periods (within the limitations imposed by the process constraints). What is less evident is the increased consumption during cheaper off-peak periods. This can be explained by considering that the bulk of the off-peak period is during night time when the ambient temperature drops, thereby increasing cooling efficiency. As a result, the flow demands also decrease and therefore, although the electricity cost is lower, the demand is also lower, providing the controller with less incentive to increase cooling capacity by switching more pumps or fans on.

4.2.5 Case comparison

Apart from visual observation of the differences in the controlled and manipulated variables as presented in Sections 4.2.1 to 4.2.4, the cases are compared based on power and cost calculations. The average power consumption, total energy consumed and total energy cost for each case is calculated for the two simulations. The incremental differences in performance from one case to the next is also calculated to gauge the degree of change. The results are presented in Tables 4.1 and 4.2.

For the first simulation, the results suggest that the application of ARC can already provide an energy reduction of around 30% whereas the HNMPC and EHN MPC techniques provide additional 4% and 6% improvements respectively compared to the base case. The total electricity cost is also reduced by around 30% with the ARC and can be reduced by an additional 4% and 7% by the application of HNMPC and EHN MPC respectively.

The second simulation results reveal a similar benefit from ARC (an energy reduction of 29.6% with a 28.9% reduction in cost) with the additional benefit of being able to control the plant within the set of constraints not quantified here (see Figures 4.2 and 4.5). The HNMPC also achieves reductions similar to the first simulation (37% on energy and 35.2% on cost) and so does the EHN MPC (40.6% on energy and 41.6% on cost).

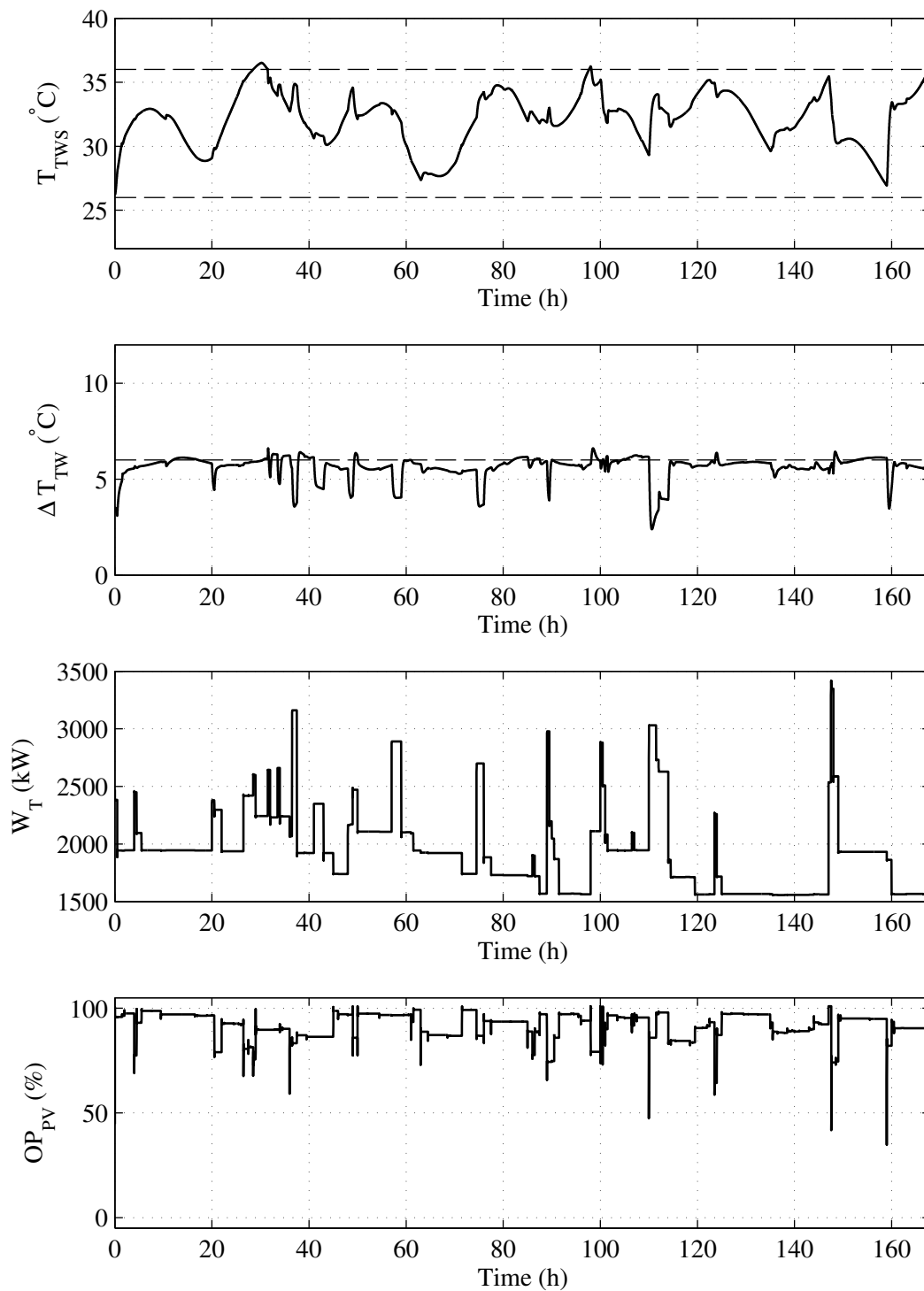


Figure 4.11: Controlled variables (Economic HNMPC case) for the first simulation.

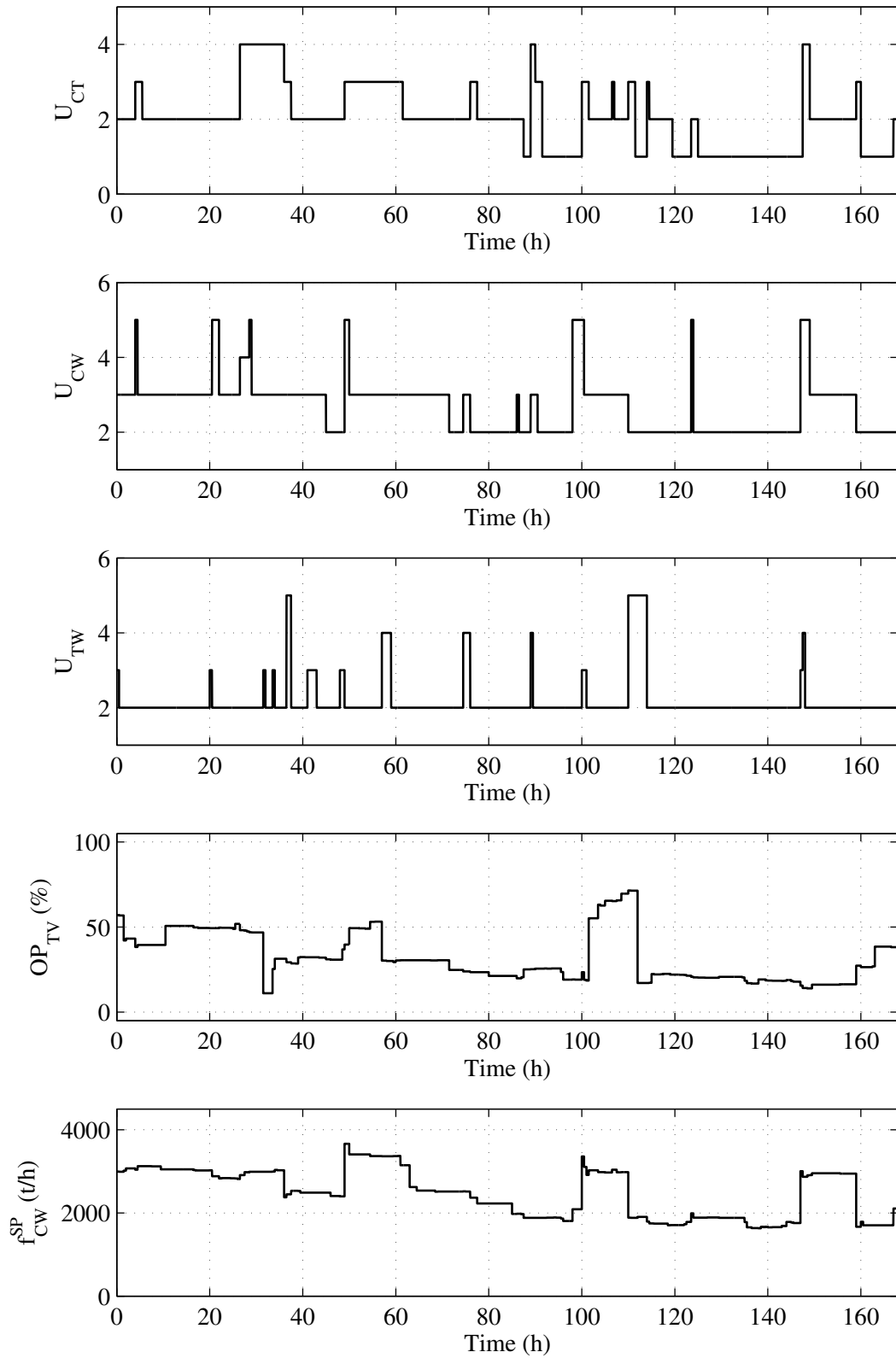


Figure 4.12: Manipulated variables (Economic HNMPC case) for the first simulation.

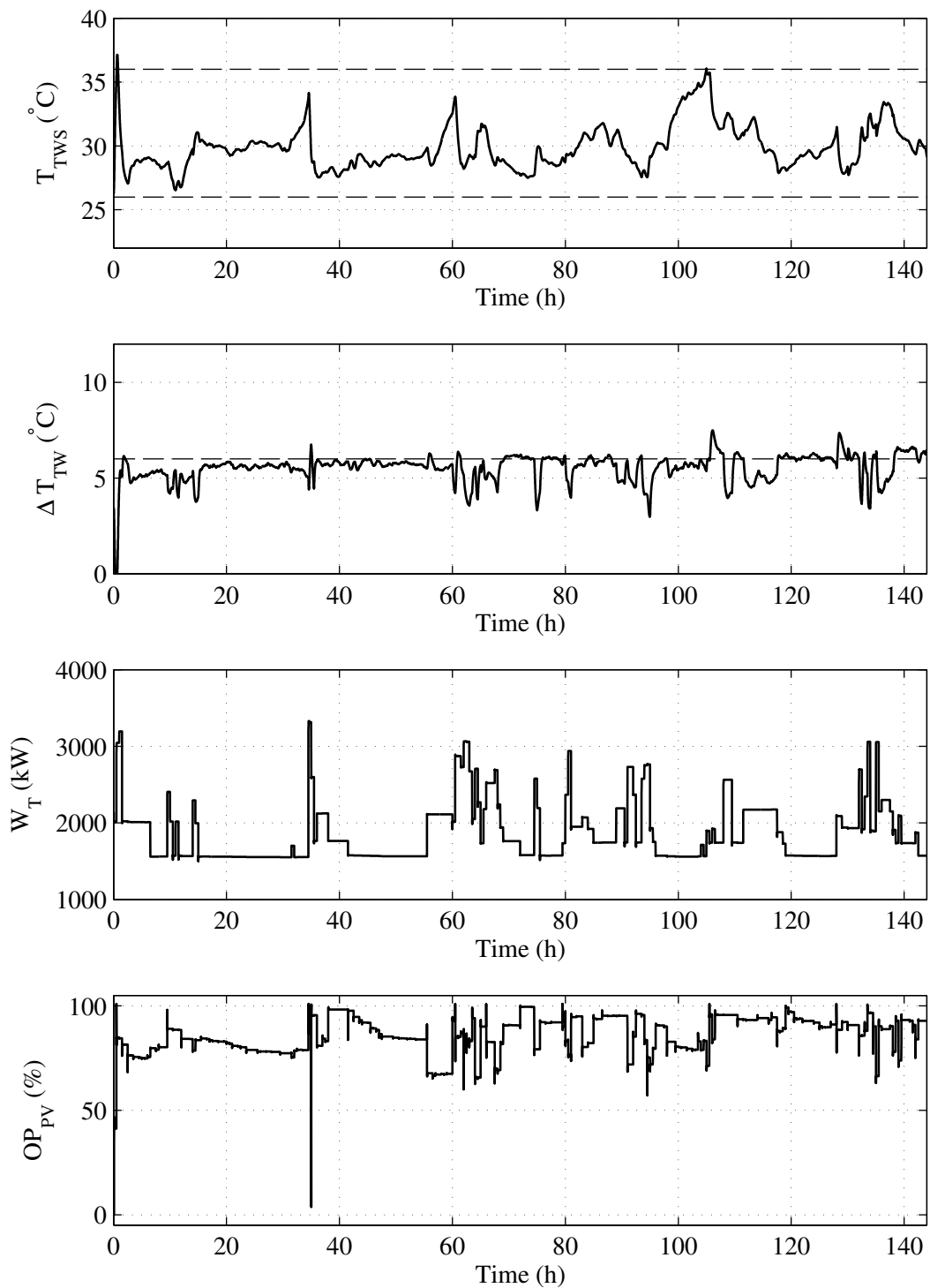


Figure 4.13: Controlled variables (Economic HN MPC case) for the second simulation.

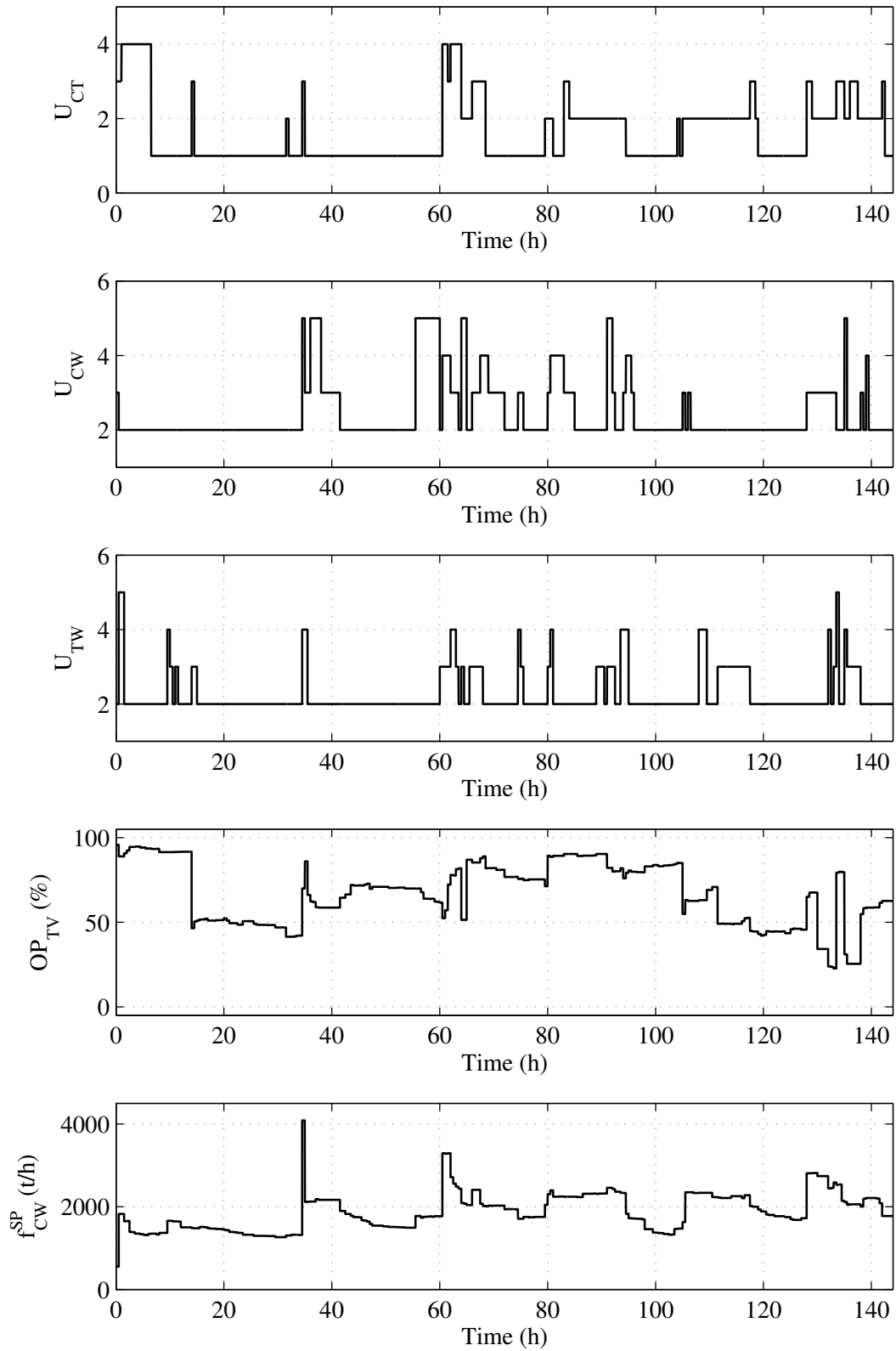


Figure 4.14: Manipulated variables (Economic HNMPC case) for the second simulation.

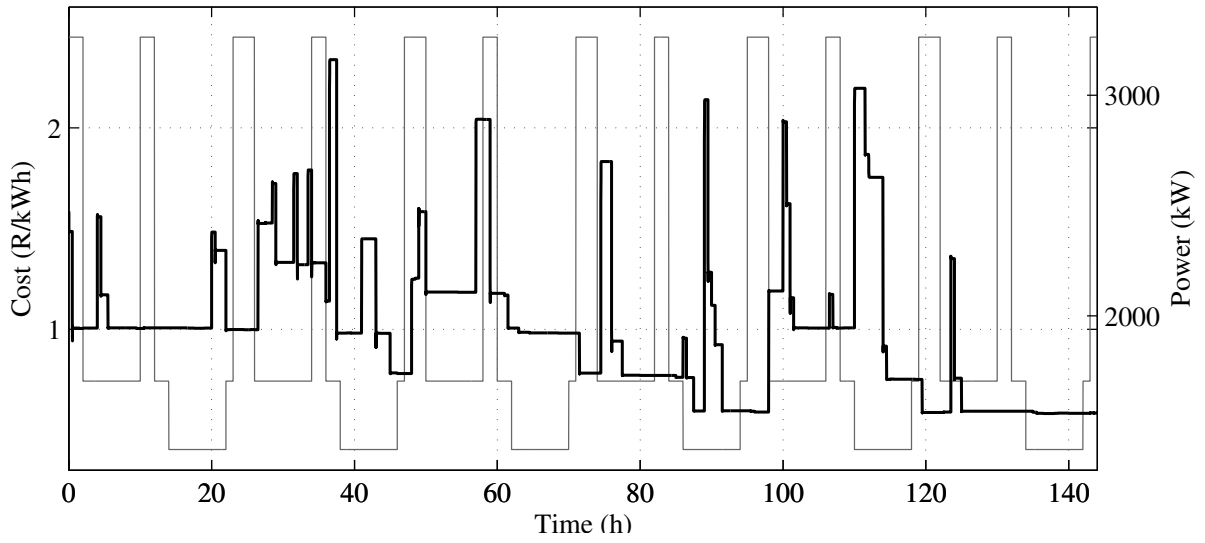


Figure 4.15: Power consumption (black line) and electricity cost (grey line) for the first simulation of the EHN MPC case.

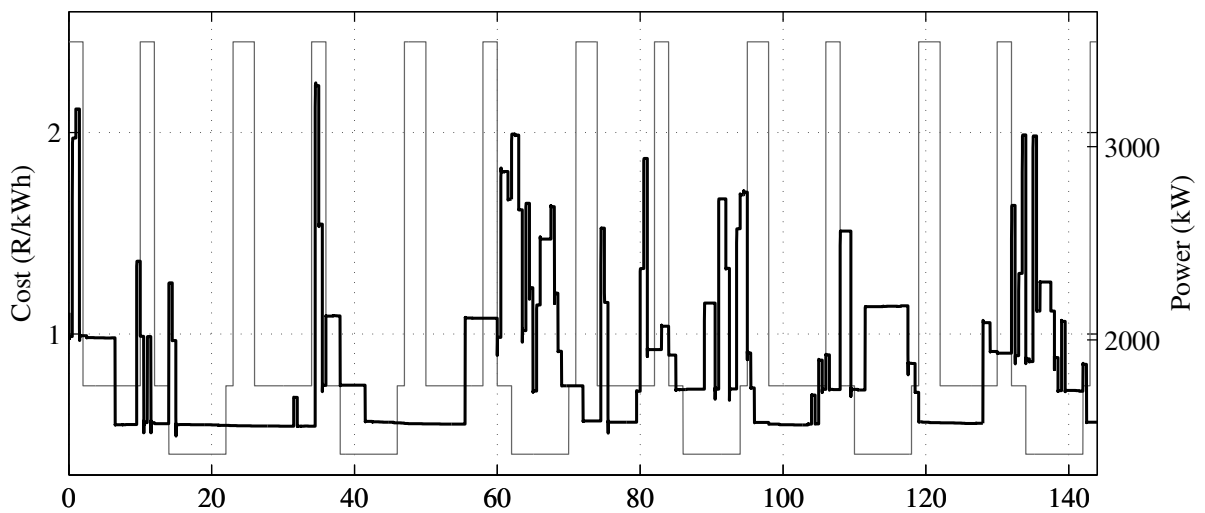


Figure 4.16: Power consumption (black line) and electricity cost (grey line) for the second simulation of the EHN MPC case.

To illustrate the relative constraint violations for the different cases, Table 4.3 contains violation indices for the tempered water supply temperature and differential temperature for the two simulations¹. It is clear that the violation of the tempered water supply temperature (which is the most important process variable) decreases as more advanced control is applied. The violation of the tempered water differential temperature on the other hand becomes progressively worse as the algorithms allow more violation in exchange for optimisation of the energy consumption/cost.

Overall, it is clear that the ARC outperforms the base case in all aspects and the HNMPC and EHNMPC similarly outperform the ARC scheme. Comparison of the HNMPC and EHNMPC reveals that the cases are closer in performance than the difference between the base and ARC cases. The aim of the EHNMPC is to produce a further saving in energy cost which it achieves. It is interesting to note that it does this while achieving a further reduction in energy consumption. It would however be possible under certain circumstances for the cost to be lower with a higher consumption. As explained in the previous section, this will not likely be the case for the cooling water system as the low cost off-peak period typically coincides with a reduced demand period due to the lower ambient temperature during night time.

¹Calculated by adding up the constraint violations over the simulation period and dividing by the number of samples

Table 4.1: Average power and energy consumption comparison.

Case	Average Power (kW)	Total energy (kWh)	Reduction from base	Incremental reduction
Simulation 1				
Base	3,052	512,736	-	-
ARC	2,130	357,840	30.2%	30.2%
HN MPC	2,001	336,168	34.4%	6.1%
EHN MPC	1,935	325,080	36.6%	3.3%
Simulation 2				
Base	3,142	452,448	-	-
ARC	2,213	318,672	29.6%	29.6%
HN MPC	1,981	285,264	37.0%	10.5%
EHN MPC	1,865	268,560	40.6%	5.9%

Table 4.2: Electricity cost comparison.

Case	Total Cost (R)	Reduction from base	Incremental reduction
Simulation 1			
Base	503,280	-	-
ARC	354,160	29.6%	29.6%
HN MPC	330,790	34.3%	6.6%
EHN MPC	316,860	37.0%	4.2%
Simulation 2			
Base	445,910	-	-
ARC	317,260	28.9%	28.9%
HN MPC	288,780	35.2%	9.3%
EHN MPC	260,610	41.6%	9.8%



Table 4.3: Constraint violation comparison.

Case	$e_{T_{rws}}$	$e_{\Delta T_{rw}}$
Simulation 1		
Base	0.0691	0.0000
ARC	0.0207	0.0208
HN MPC	0.0014	0.0256
EHN MPC	0.0058	0.0312
Simulation 2		
Base	3.5254	0.0047
ARC	0.0190	0.0244
HN MPC	0.0032	0.0242
EHN MPC	0.0013	0.0576

CHAPTER 5

CONCLUSION

5.1 CONCLUDING REMARKS

Utility generation and utility transmission are often overlooked as optimisation areas, although the potential for energy reduction could be substantial in some cases (as is shown in this study). The cooling water system considered here is a prime example of such a system.

The cooling water system model developed in this study provides valuable insight into the behaviour of the process. It illustrates the effects of the inputs and disturbances on the process variables and highlights the non-linear interactive nature of the system. The model can easily be customised to represent different sized systems and provides a robust simulation platform for studies of all sorts. Figures 3.5 to 3.12 illustrate the behaviour of the system visually, confirm that the system responds as expected and reveal some interesting less intuitive dynamics and interactions. The state-space model form described in Section 3.3.2 provides a convenient dynamic model format for designing control and optimisation solutions and also groups the model equations in a logical manner.

Model parameter estimation using an optimisation algorithm significantly improves model quality. The genetic algorithm is a convenient choice in this case due to its ability to handle complex, non-linear, constrained problems intuitively and its robustness against convergence to a local optimum. The parameter estimation results in adequate accuracy for the purposes of this simplified model as is shown by the correlation coefficients in Table 3.3, considering that the objectives of the model are firstly to provide a simple plant simulation platform,

and secondly to provide a starting point for the formulation of control and optimisation solutions.

It is difficult for a plant operator to determine whether the process constraints will be honoured or violated when switching pumps or cooling tower fans. This model allows the controller/optimiser to determine this and optimise the system while honouring the constraints.

A reduction in energy consumption and/or cost, by reducing the number of pumps and fans running during times where over-cooling is provided, is the main advantage of optimising this system. The application of advanced control techniques can be very effective optimisation vehicles for this purpose and a wealth of knowledge exists in these fields. For systems that contain continuous as well as discrete components, there are several approaches which may be considered for control and optimisation. Conventional approaches treat the discrete and continuous components as well as the dynamic versus economic optimisation in separate layers.

The ARC solution follows this approach by applying continuous control and discrete switching logic independently. Advanced base layer techniques such as override selector control aid the basic PID in providing constraint handling capabilities for the controlled variables, whereas time-based switching logic prioritises and facilitates switching activities of the discrete components, whenever the process conditions warrant it. The switching conditions are designed in such a manner as to minimise the number of running equipment, which results in the reduction of electrical energy consumption. The generic guidelines given in Section 3.4.2 allows for customisation to other optimisation criteria. Although the ARC approach yields a significant improvement compared to the base case, it is still limited in its optimisation capability.

The HNMPC approach unifies several of the layers by including the discrete and continuous components in a single control and optimisation application. This is accomplished by using the model developed in Section 3.3 to predict the response of the system to input and disturbance variable changes and using an optimiser to drive the process to a desired operating point, through the calculation of the required input sequence, based on the formulation of the cost function. The genetic algorithm is again used as the optimisation engine for the

same reasons mentioned above, but also due to its ability to handle both the discrete and continuous manipulated variables simultaneously, together with the discontinuities that the discrete variables create in the model.

The EHN MPC goes further to combine the dynamic and economic optimisation. This is achieved by including the predefined time-of-use electricity rates in the problem formulation and setting the performance function to optimise on the dynamic total electricity cost rather than simply using the energy consumption over the prediction horizon.

The combination of layers/objectives into a single solution in the HNMPC and EHN MPC cases yields superior results to the layered approach as can be seen from Figures 4.1 to 4.13 and Tables 4.1 and 4.2.

The results obtained from the control and optimisation simulation studies indicate that significant energy and cost saving may be realised with the use of modern control and optimisation techniques. These techniques do not necessarily require any physical changes to the plant equipment and are able to accommodate complex non-linear hybrid systems in elegant ways without having to linearise or transform the system model. Substantial benefits may be realisable by using the existing plant control system with the existing equipment without the need to install additional equipment such as variable speed drives.

5.2 FUTURE SCOPE

This study has not explored all the possibilities of the system and the control/optimisation techniques described. Possible future scope includes:

- Consider the case where one or more of the discrete components are fitted with variable speed drives.
- Further develop and refine the model to include for example air-flow estimations.
- Include other cost components in the EHN MPC case such as maximum demand charges.
- Consider the case where the plant heat exchanger network is not static but changes in terms of its flow coefficient.



- Introduce other disturbances into the system such as equipment failure and degradation, and plant trips.
- Extend these modelling, control and optimisation concepts to other utility systems of a similar nature.

REFERENCES

- [1] B. G. Lipták, *Distillation Control & Optimization*. Putman Media, 2007.
- [2] R. Inglesi-Lotz and J. N. Blignaut, “South Africa’s electricity consumption: A sectoral decomposition analysis,” *Applied Energy*, vol. 88, no. 12, pp. 4779–4784, December 2011.
- [3] I. K. Craig *et al.*, “Control in the process industries,” in *The Impact of Control Technology*, T. Samad and A. Annaswamy, Eds. IEEE Control Systems Society, 2011, available at <http://www.ieeecss.org/main/loct-report>.
- [4] J. Palm and P. Thollander, “An interdisciplinary perspective on industrial energy efficiency,” *Applied Energy*, vol. 87, no. 10, pp. 3255–3261, October 2010.
- [5] X. Xia and J. Zhang, “Mathematical description for the measurement and verification of energy efficiency improvement,” *Applied Energy*, vol. 111, pp. 247–256, November 2013.
- [6] C. J. Muller, I. K. Craig, and N. L. Ricker, “Modelling, validation, and control of an industrial fuel gas blending system,” *Journal of Process Control*, vol. 21, no. 6, pp. 852–860, July 2011.
- [7] N. L. Ricker, C. J. Muller, and I. K. Craig, “Fuel gas blending benchmark for economic performance evaluation of advanced control and state estimation,” *Journal of Process Control*, vol. 22, no. 6, pp. 968–974, July 2012.
- [8] M. S. Bhatt, “Energy audit case studies I—steam systems,” *Applied Thermal Engineering*, vol. 20, no. 3, pp. 285–296, February 2000.
- [9] R. Saidur, N. A. Rahim, and M. Hasanuzzaman, “A review on compressed-air energy use and energy savings,” *Renewable and Sustainable Energy Reviews*, vol. 14, no. 4, pp.

References

- 1135–1153, May 2010.
- [10] N. Tanaka and R. Wicks, “Power generation from coal: Measuring and reporting efficiency performance and CO₂ emissions,” International Energy Agency, Tech. Rep., 2010.
- [11] W. H. J. Graus, M. Voogt, and E. Worrell, “International comparison of energy efficiency of fossil power generation,” *Emergy Policy*, vol. 35, pp. 3936–3951, March 2007.
- [12] S. Zhang and X. Xia, “Optimal control of operation efficiency of belt conveyor systems,” *Applied Energy*, vol. 87, no. 6, pp. 1929–1937, June 2010.
- [13] B. Matthews and I. K. Craig, “Demand side management of a run-of-mine ore milling circuit,” *Control Engineering Practice*, vol. 21, no. 6, pp. 759–768, 2013.
- [14] X. Xia, J. Zhang, and W. Cass, “Energy management of commercial buildings – a case study from a POET perspective of energy efficiency,” *Journal of Energy in Southern Africa*, vol. 23, no. 1, pp. 23–31ges, February 2012.
- [15] Z. Wu, H. Tazvinga, and X. Xia, “Demand side management of photovoltaic-battery hybrid system,” *Applied Energy*, vol. 148, pp. 294–304, June 2015.
- [16] J. Široký, F. Oldewurtel, J. Cigler, and S. Prívará, “Experimental analysis of model predictive control for an energy efficient building heating system,” *Applied Energy*, vol. 88, no. 9, pp. 3079–3087, September 2011.
- [17] K. C. Edwards and D. P. Finn, “Generalised water flow rate control strategy for optimal part load operation of ground source heat pump systems,” *Applied Energy*, vol. 150, pp. 50–60, July 2015.
- [18] A. J. van Staden, J. Zhang, and X. Xia, “A model predictive control strategy for load shifting in a water pumping scheme with maximum demand charges,” *Applied Energy*, vol. 88, no. 12, pp. 4785–4794, December 2011.
- [19] A. Middelberg, J. Zhang, and X. Xia, “An optimal control model for load shifting - With application in the energy management of a colliery,” *Applied Energy*, vol. 86, no. 7–8, pp. 1266–1273, July–August 2009.

References

- [20] S. Zhang and X. Xia, “Modeling and energy efficiency optimization of belt conveyors,” *Applied Energy*, vol. 88, no. 9, pp. 3061–3071, September 2011.
- [21] B. G. Lipták, *Instrument Engineers’ Handbook*, 4th ed. Florida: CRC Press, 2006, vol. II: Process Control and Optimization.
- [22] D. W. Green, Ed., *Perry’s Chemical Engineers’ Handbook*, 4th ed. McGraw-Hill, 1997.
- [23] C. A. X. Marques, C. H. Fontes, M. Embiruçu, and R. A. Kalid, “Efficiency control in a commercial counter flow wet cooling tower,” *Energy Conversion and Management*, vol. 50, no. 11, pp. 2843–2855, November 2009.
- [24] —, “Energy efficiency in an industrial wet cooling tower through improved control,” *Computer Aided Chemical Engineering*, vol. 27, pp. 1449–1454, 2009.
- [25] A. Bemporad and M. Morari, “Control of systems integrating logic, dynamics, and constraints,” *Automatica*, vol. 35, no. 3, pp. 407–427, March 1999.
- [26] E. F. Camacho, D. R. Ramirez, D. Limon, D. Muñoz de la Peña, and T. Alamo, “Model predictive control techniques for hybrid systems,” *Annual Reviews in Control*, vol. 34, no. 1, pp. 21–31, April 2010.
- [27] C. W. Gellings, “The concept of demand-side management for electric utilities,” in *Proceedings of the IEEE*, vol. 73, no. 10. IEEE, October 1985, pp. 1468–1470.
- [28] H. Zhang, X. Xia, and J. Zhang, “Optimal sizing and operation of pumping systems to achieve energy efficiency and load shifting,” *Electric Power Systems Research*, vol. 86, pp. 41–50, May 2012.
- [29] X. Xia and J. Zhang, “Energy efficiency and control systems—from a POET perspective,” in *Proceedings of the IFAC Conference on Control Methodologies and Technology for Energy Efficiency*. Vilamoura: IFAC, March 2010.
- [30] —, “Energy audit - from a POET perspective,” in *Proceedings of the International Conference on Applied Energy*, Singapore, April 2010.
- [31] C. Alexander and M. Sadiku, *Fundamentals of Electric Circuits*, 2nd ed. McGraw-Hill,

References

- 2003.
- [32] D. E. Seborg, T. F. Edgar, and D. A. Mellichamp, *Process Dynamics and Control*, 2nd ed. Wiley, 2004.
- [33] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control Analysis and Design*, 2nd ed. Wiley, 2005.
- [34] R. C. Dorf and R. H. Bishop, *Modern Control Systems*, 10th ed. Pearson Prentice Hall, 2005.
- [35] S. Skogestad, “Control structure design for complete chemical plants,” *Computers and Chemical Engineering*, vol. 28, no. 1–2, pp. 219–234, January 2004.
- [36] T. Larsson and S. Skogestad, “Plantwide control—a review and a new design procedure,” *Modeling, Identification and Control*, vol. 21, no. 4, pp. 209–240, 2000.
- [37] J. J. Downs and S. Skogestad, “An industrial and academic perspective on plantwide control,” *Annual Reviews in Control*, vol. 35, no. 1, pp. 99–110, April 2011.
- [38] T. E. Marlin, *Process Control: Designing Processes and Control Systems for Dynamic Performance*. McGraw-Hill, 1995.
- [39] H. L. Wade, *Regulatory and Advanced Regulatory Control: System Development*. ISA, 1994.
- [40] D. E. Kirk, *Optimal Control Theory An Introduction*. Dover, 2004.
- [41] M. Darby and M. Nikolaou, “MPC: Current practice and challenges,” *Control Engineering Practice*, vol. 20, pp. 328–342, February 2012.
- [42] J. Richalet, A. Rault, J. Testud, and J. Papon, “Algorithmic control of industrial processes,” in *Proceedings of the 4th IFAC Symposium on Identification and System Parameter Estimation*. IFAC, September 1976, pp. 1119–1167.
- [43] C. Cutler and B. Ramaker, “Dynamic matrix control—a computer control algorithm,” in *Proceedings of the Joint Automatic Control Conference*, 1980.

References

- [44] S. Qin and T. Badgwell, “A survey of industrial model predictive control technology,” *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, July 2003.
- [45] E. F. Camacho and C. Bordons, *Model Predictive Control*, 2nd ed., ser. Advanced Textbooks in Control and Signal Processing. London: Springer, 2007.
- [46] M. Ellis, H. Durand, and P. D. Christofides, “A tutorial review of economic model predictive control methods,” *Journal of Process Control*, vol. 24, no. 8, pp. 1156–1178, August 2014.
- [47] M. Darby, M. Nikolaou, J. Jones, and D. Nicholson, “RTO: An overview and assessment of current practice,” *Journal of Process Control*, vol. 21, pp. 874–884, May 2011.
- [48] L. T. Biegler, “Advances in nonlinear programming concepts for process control,” *Journal of Process Control*, vol. 8, no. 5–6, pp. 301–311, October–December 1998.
- [49] —, “From nonlinear programming theory to practical optimization algorithms: A process engineering viewpoint,” *Computers and Chemical Engineering*, vol. 17, pp. S63–S80, 1993, Supplement 1.
- [50] L. G. Bullard and L. T. Biegler, “Iterated linear programming strategies for nonsmooth simulation: Continuous and mixed-integer approaches,” *Computers and Chemical Engineering*, vol. 16, no. 10–11, pp. 949–961, 1992.
- [51] L. O. Santos, P. A. F. N. A. Afonso, J. A. A. M. Castro, N. M. C. Oliveira, and L. T. Biegler, “On-line implementation of nonlinear mpc: an experimental case study,” *Control Engineering Practice*, vol. 9, no. 8, pp. 847–857, 2001.
- [52] T. Westerlund, F. Pettersson, and I. E. Grossmann, “Optimization of pump configurations as a minlp problem,” *Computers and Chemical Engineering*, vol. 19, no. 9, pp. 845–858, 1994.
- [53] L. Su, L. Tang, and I. E. Grossmann, “Computational strategies for improved minlp algorithms,” *Computers and Chemical Engineering*, vol. 75, pp. 40–48, April 2015.
- [54] K. J. Hunt, “Systems identification with genetic algorithms,” in *IEE Colloquium on*

References

- Genetic Algorithms for Control Systems Engineering*, May 1993, pp. 3/1–3/3.
- [55] P. J. Fleming and C. M. Fonseca, “Genetic algorithms in control systems engineering: A brief introduction,” in *IEE Colloquium on Genetic Algorithms for Control Systems Engineering*, May 1993, pp. 1/1–1/5.
- [56] C. M. M. Fonseca, “Multiobjective genetic algorithms with application to control engineering problems,” Ph.D. dissertation, Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK, 1995.
- [57] D. Whitley, “A genetic algorithm tutorial,” *Statistics and Computing*, vol. 4, no. 2, pp. 65–85, 1994.
- [58] K. Deep, K. P. Singh, M. L. Kansal, and C. Mohan, “A real coded genetic algorithm for solving integer and mixed integer optimization problems,” *Applied Mathematics and Computation*, vol. 212, no. 2, pp. 505–518, June 2009.
- [59] K. Dep, “An efficient constraint handling method for genetic algorithms,” *Computer Methods in Applied Mechanics and Engineering*, vol. 189, no. 2–4, pp. 311–338, June 2000.
- [60] W. L. Brogan, *Modern control theory*. Englewood Cliffs: Prentice Hall, 1991, p. 12.
- [61] X. Zhuan and X. Xia, “Optimal operation scheduling of a pumping station with multiple pumps,” *Applied Energy*, vol. 104, pp. 250–257, April 2013.
- [62] N. P. Lieberman and E. T. Lieberman, *A working guide to process equipment*, 3rd ed. McGraw-Hill, 2008.
- [63] R. L. Daugherty and J. B. Franzini, *Fluid Mechanics with Engineering Applications*, 7th ed. McGraw-Hill, 1977, pp. 524–526.
- [64] Y. Wang and Q. Chen, “A direct optimal control strategy of variable speed pumps in heat exchanger networks and experimental validations,” *Energy*, vol. 85, no. 1, pp. 609–619, June 2015.
- [65] K. Gololo, “Analysis, synthesis and optimization of complex cooling water systems,”

References

- Ph.D. dissertation, Faculty of Engineering, Built Environment and Information Technology, University of Pretoria, April 2013.
- [66] C. J. Muller and I. K. Craig, “Modelling of a dual circuit induced draft cooling water system for control and optimisation purposes,” *Journal of Process Control*, vol. 25, pp. 105–114, January 2015.
- [67] R. Stull, “Wet-bulb temperature from relative humidity and air temperature,” *Journal of Applied Meteorology and Climatology*, vol. 50, pp. 2267–2269, August 2011.
- [68] C. J. Muller and I. K. Craig, “Energy reduction for a dual circuit cooling water system using advanced regulatory control,” *Submitted to: Applied Energy*, July 2015.
- [69] D. Q. Mayne, “Model predictive control: Recent developments and future promise,” *Automatica*, vol. 50, no. 12, pp. 2967–2986, December 2014.

APPENDIX A

MATLAB SIMULATION DETAILS

The simulation studies, described in Chapters 3 and 4 for the model validation and control and optimisation, were implemented in Matlab, using several functions and code files. The modelling of the system was mostly done in Simulink with the Matlab code making reference calls to the Simulink models. Plant and simulation data was imported into Matlab from Excel sheets in some cases and created within Matlab in others. The different sections of this Appendix give more detail on the Matlab code and Simulink model development although the code itself is not presented. The code may be requested from the University of Pretoria if required.

A.1 SIMULINK MODELS

The model developed for the cooling water system was built in Simulink using a combination of function blocks including standard mathematical blocks, polynomial function blocks, transfer function blocks, and PID control algorithm blocks. Data sets and time stamped data were sent to the model from the Matlab workspace and some variables in turn were written back to the workspace for use in the Matlab programs. Three versions of the model were created: the first is used for the model verification using the imported plant data, the second is used to run the simulations for the base case and the ARC case and the third is used for the HNMPC and EHNMPC cases. The models are essentially identical apart from the interfaces to the Matlab workspace and some additional switching logic composite blocks for the ARC case.

Figures A.1 to A.3 show the main overview model pages for the three cases.

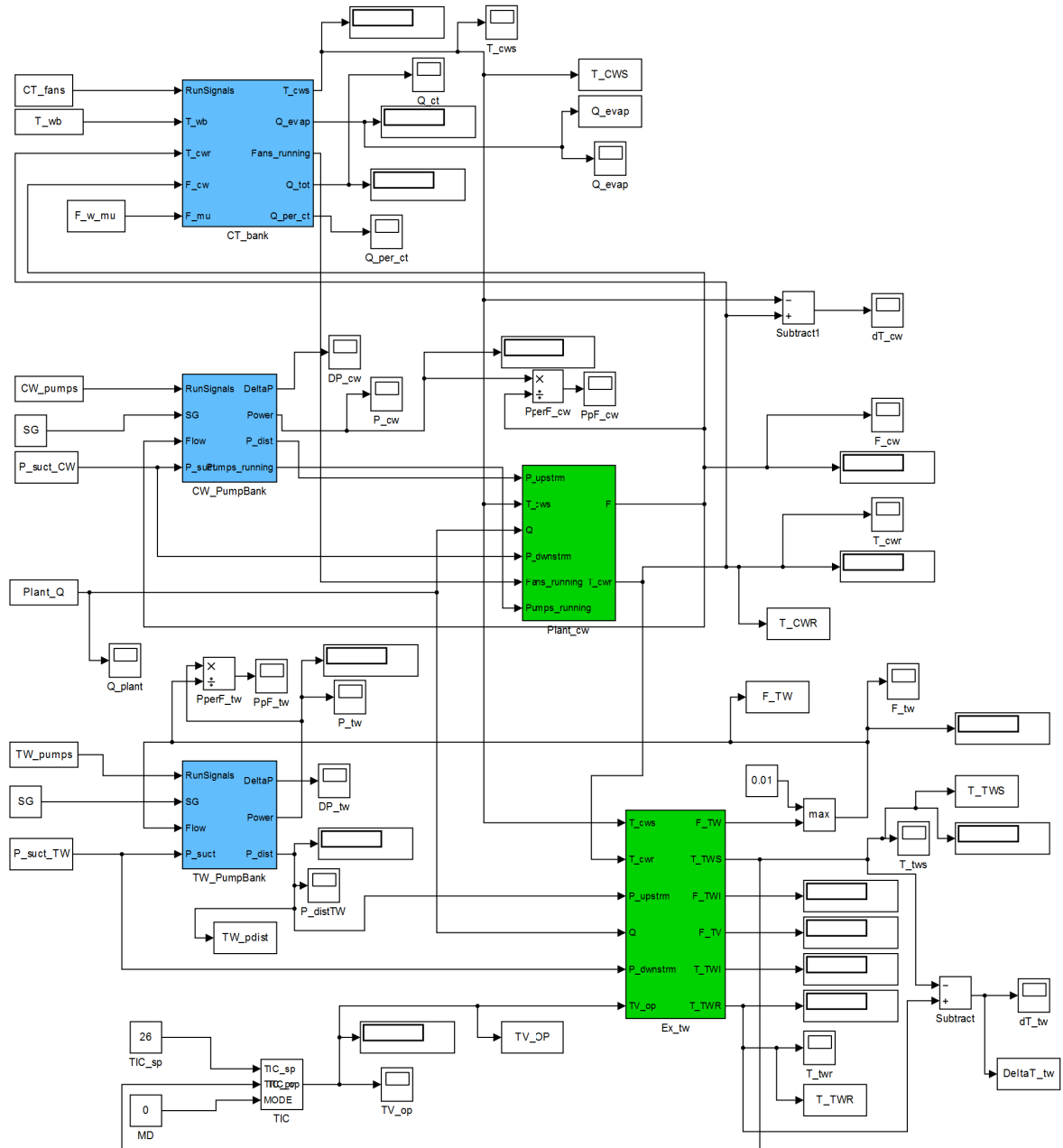


Figure A.1: Simulink main page for the verification model.

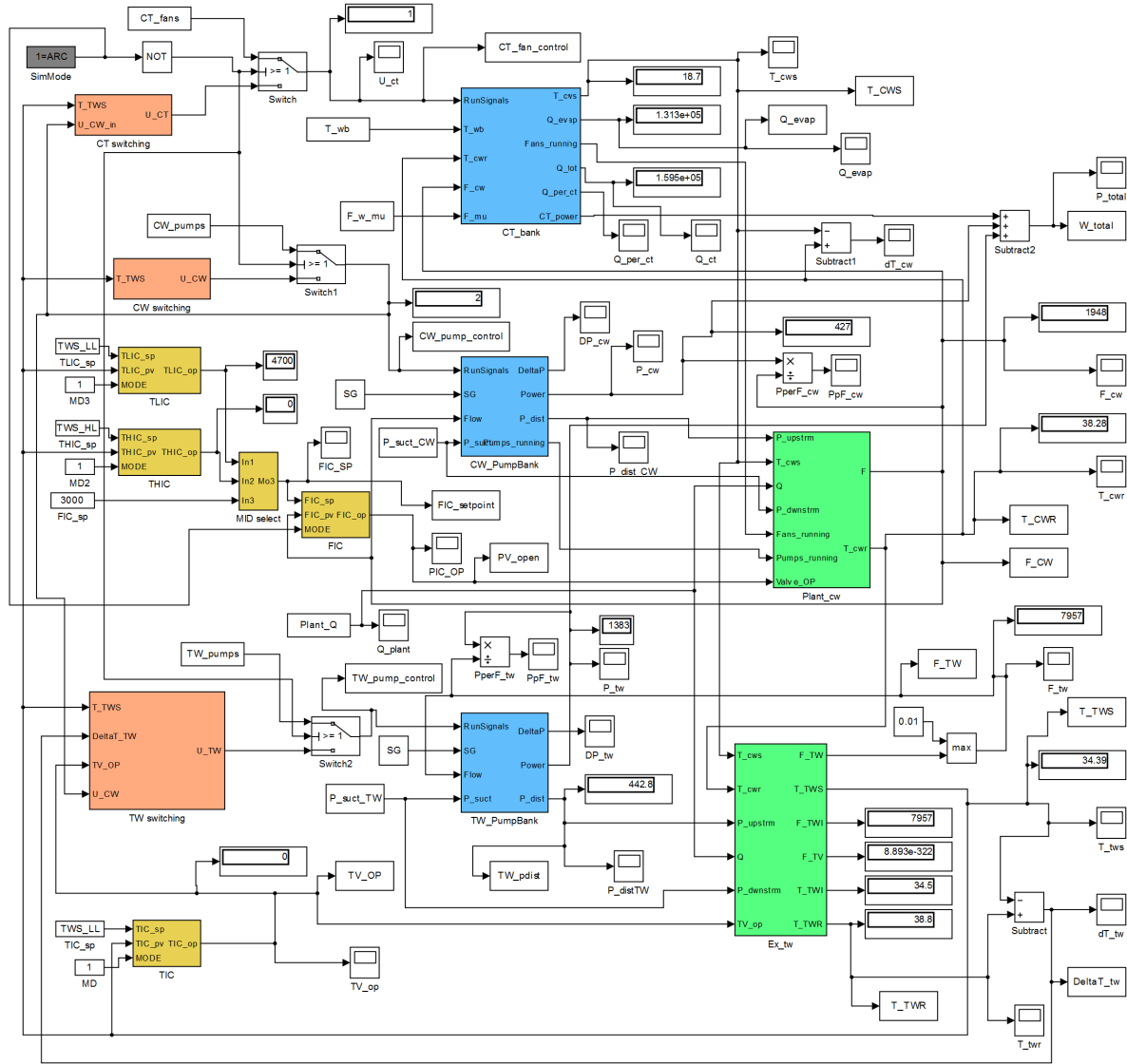


Figure A.2: Simulink main page for the base case and ARC model.

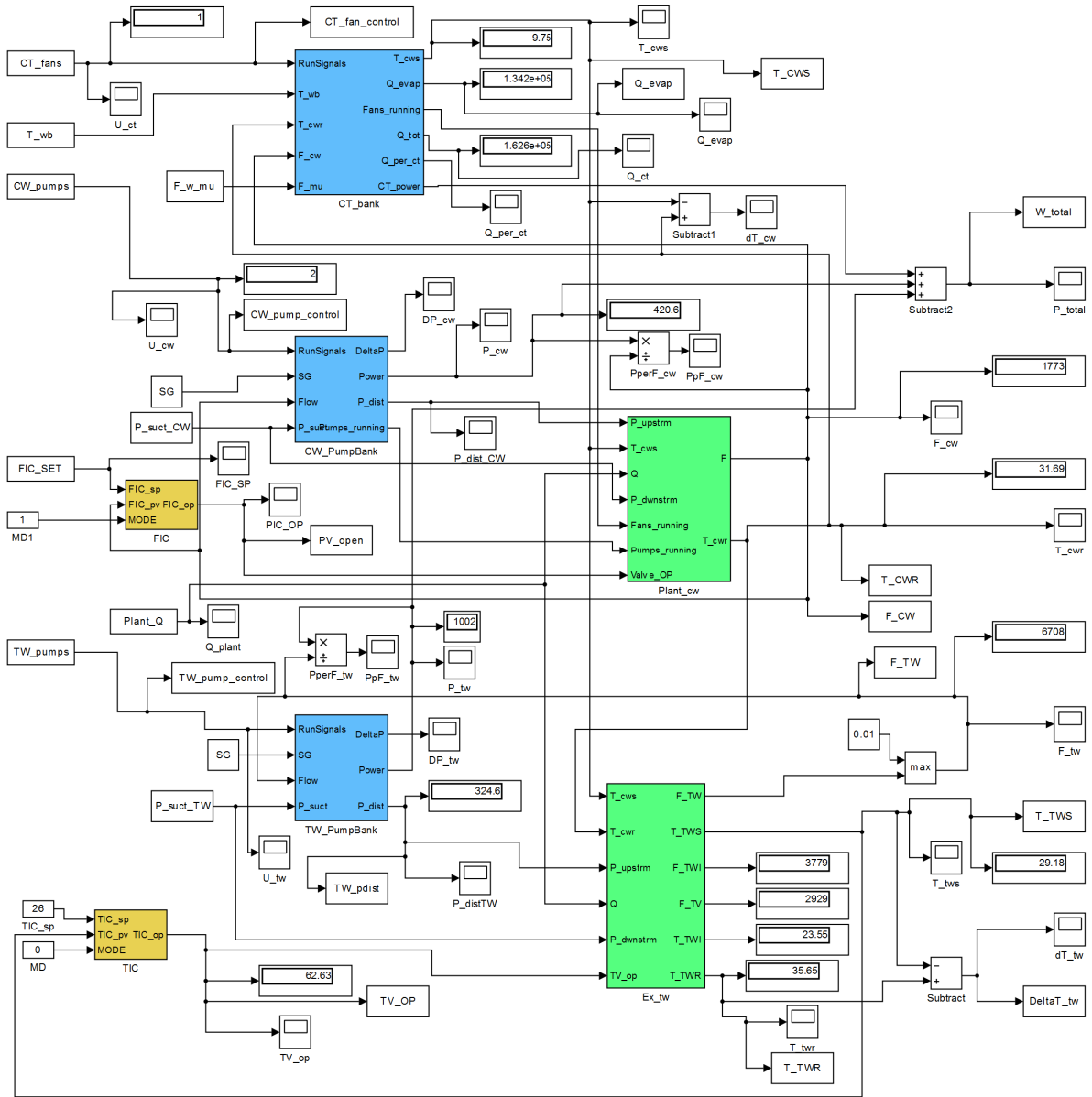


Figure A.3: Simulink main page for the HNMPC and EHN MPC model.

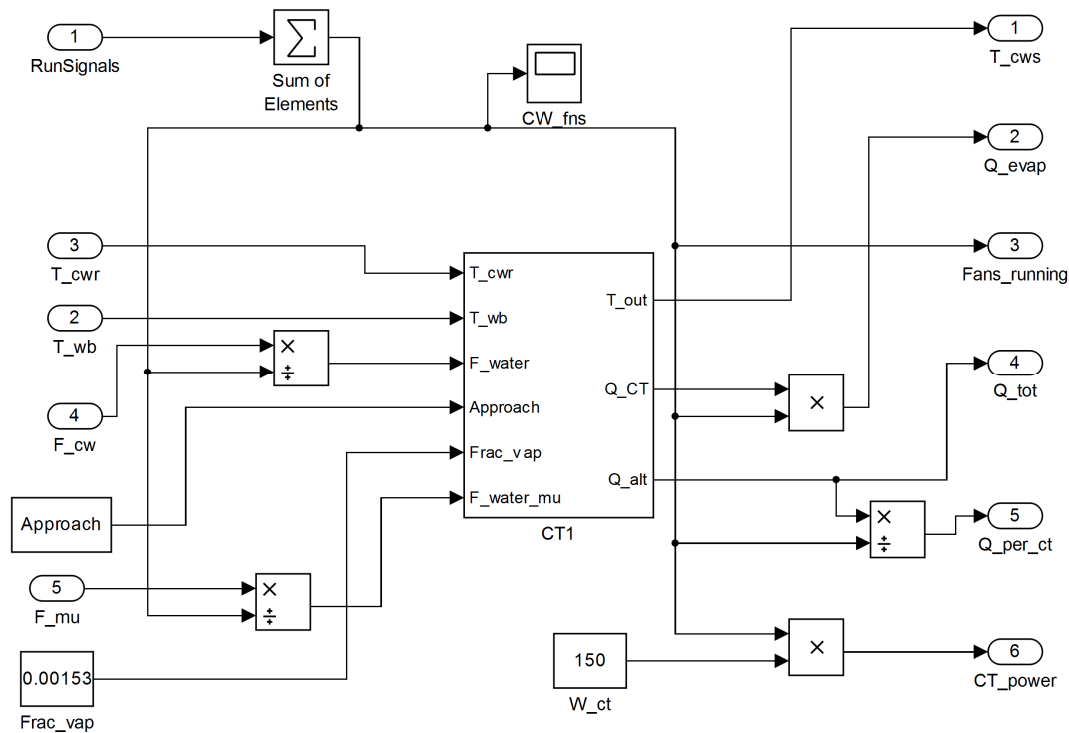


Figure A.4: Simulink CT bank.

The main pages contain some logic and data handling blocks with the rest of the calculations being performed in composite blocks. These are shown in Figures A.4 to A.12.

The solver settings are shown in Figure A.16 where it is indicated that the Bogacki-Shampine numerical integration technique was used for solving the model with a fixed step size of 1/120 hours (30 seconds).

A.2 MATLAB CODE DESCRIPTION

The following sections describe the approaches followed in implementing the solutions in Matlab for the modelling, verification, parameter estimation and the control and optimisation solutions. Apart from these main functions, several other Matlab files were created for plotting of graphs and performing miscellaneous tasks.

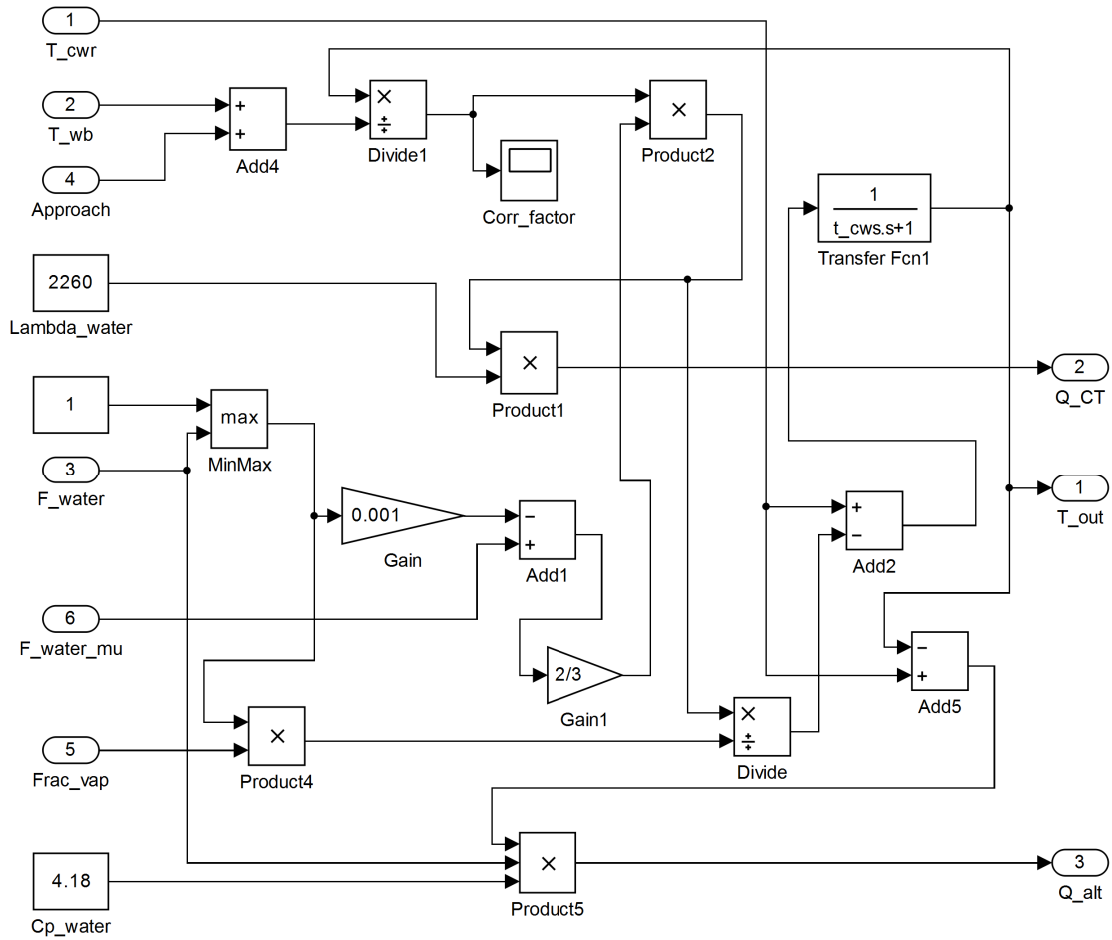


Figure A.5: Simulink CT.

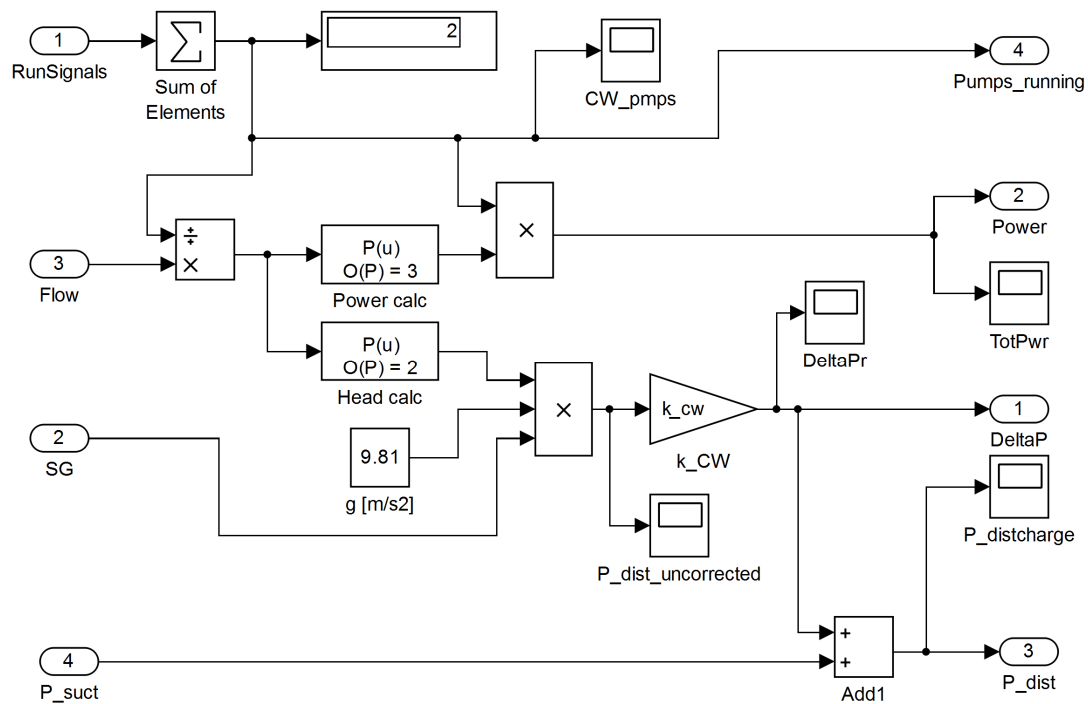


Figure A.6: Simulink CW pump.

A.2.1 Model verification

After importing the plant data from Excel files, the required calculations are performed to determine the wet-bulb temperature and approach. The simulation is then run for the simulated time of the plant data by running the Simulink model and feeding the data from the Matlab workspace, which include the plant duty, the pump and fan running signals, the calculated wet-bulb temperature and approach, the tempered water temperature controller valve opening and some model parameter and constants. Simulink then returns the simulation results for several of the process variables including the tempered water supply temperature, tempered water return temperature, tempered water differential temperature, cooling water supply temperature, cooling water return temperature and the evaporative duty. The correlation coefficients between the simulation and plant data are then calculated for the tempered water supply temperature, tempered water return temperature, tempered water differential temperature and duty to evaluate their similarities and the variables are trended together for visual inspection.

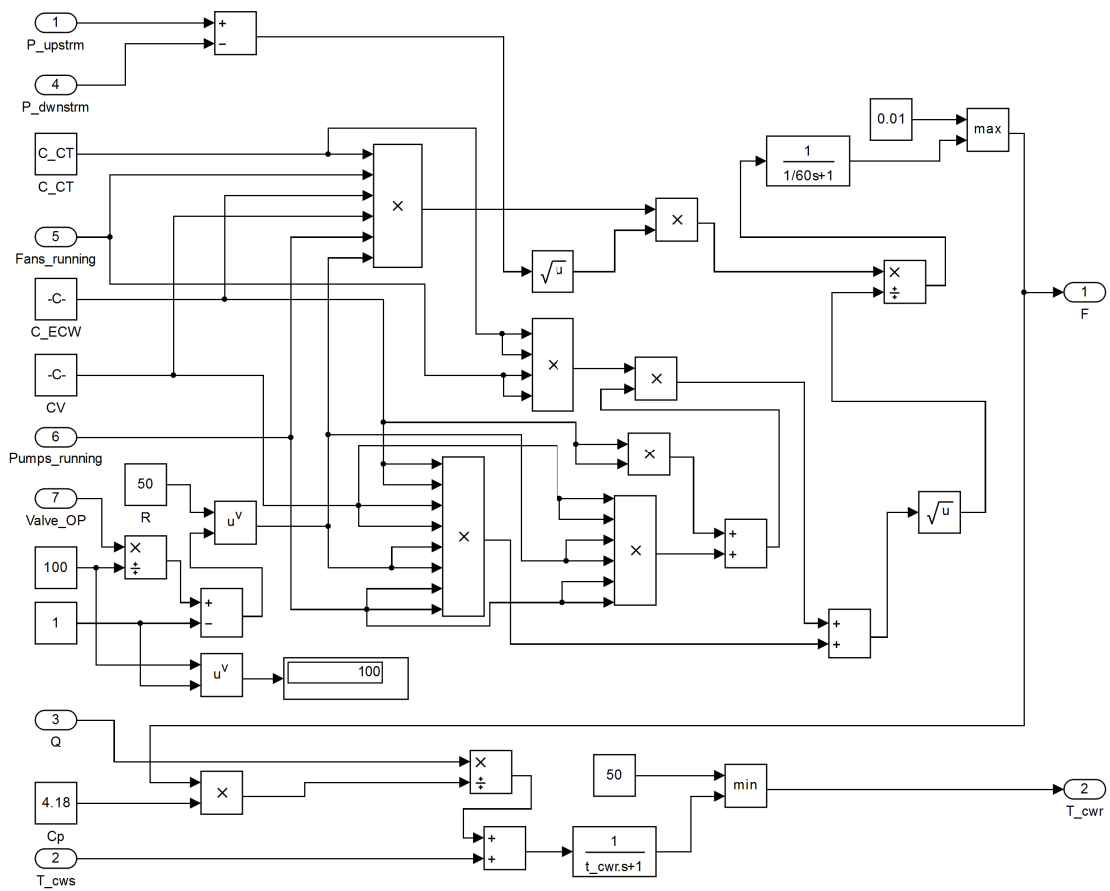


Figure A.7: Simulink CW plant.

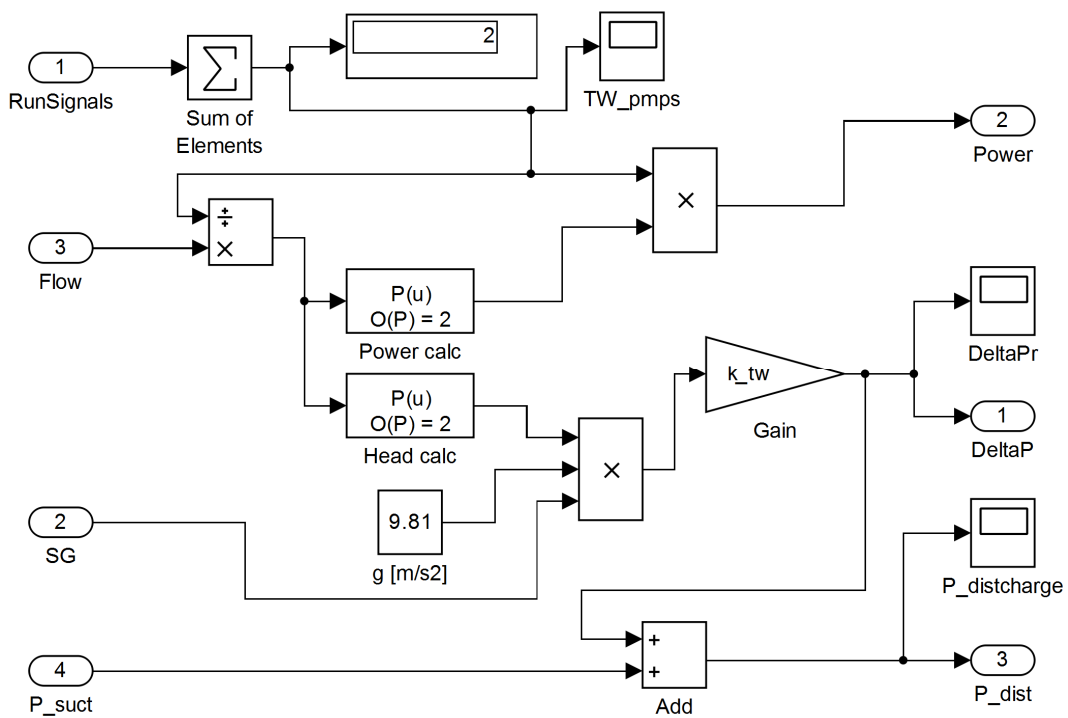


Figure A.8: Simulink TW pump.

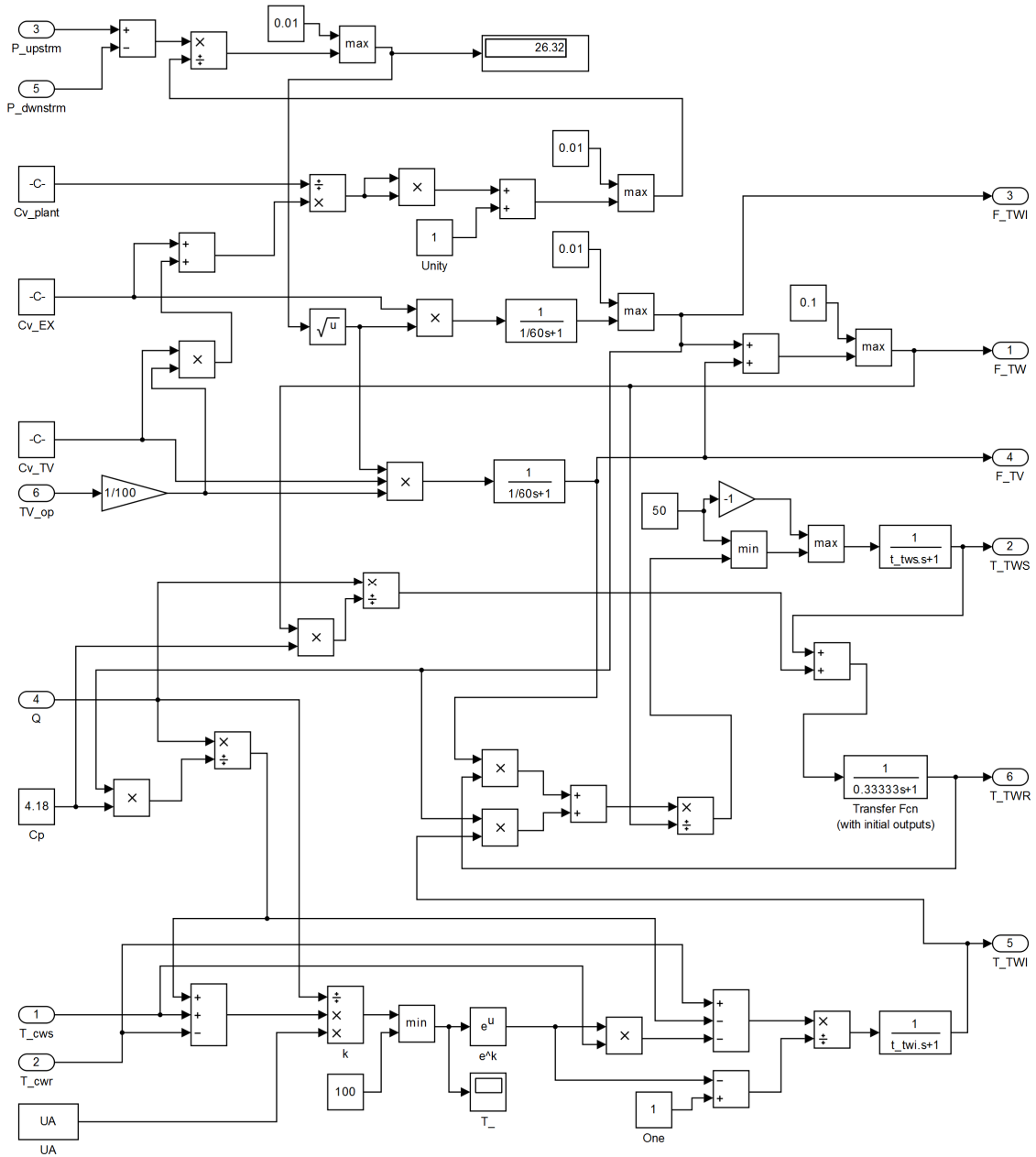


Figure A.9: Simulink TW plant.

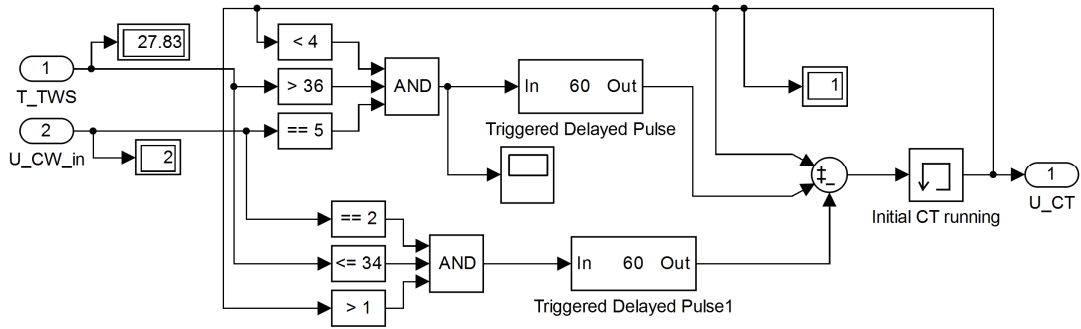


Figure A.10: Simulink CT fan switching logic.

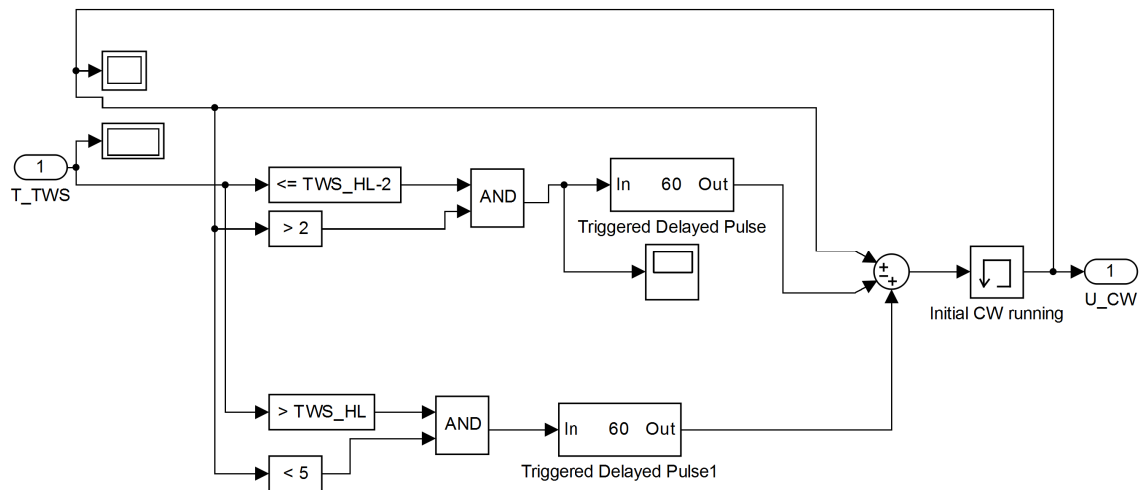


Figure A.11: Simulink CW pump switching logic.

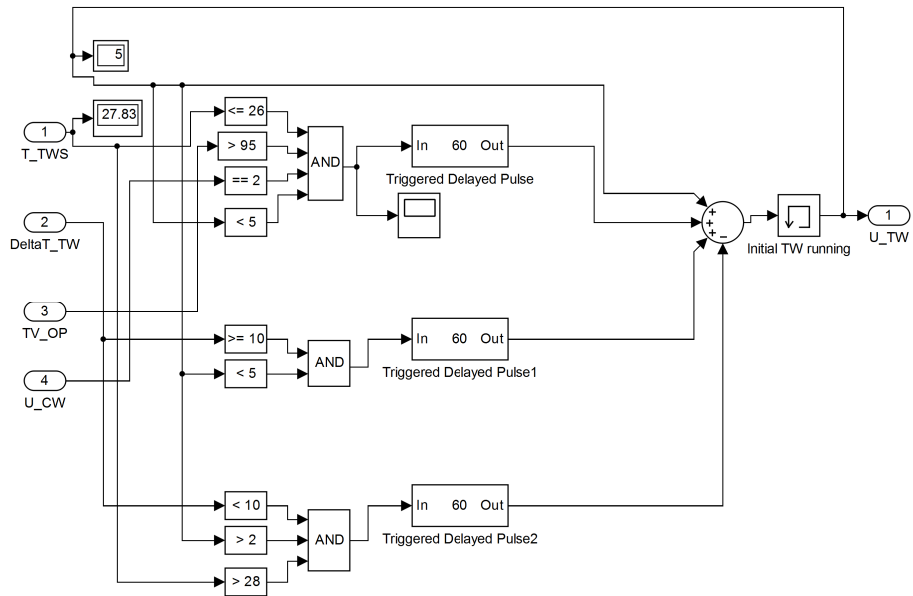


Figure A.12: Simulink TW pump switching logic.

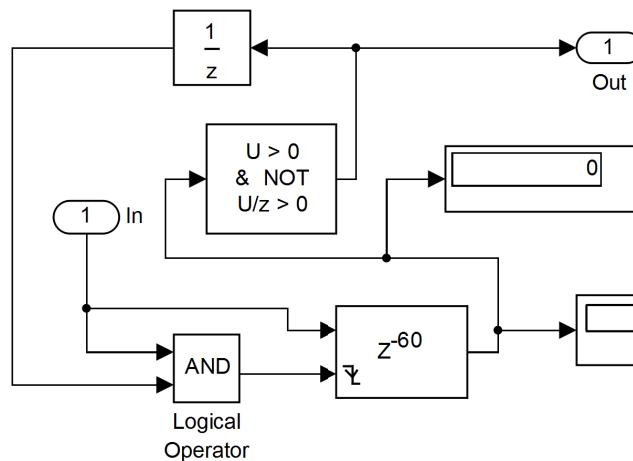


Figure A.13: Simulink Triggered Delayed Pulse logic.

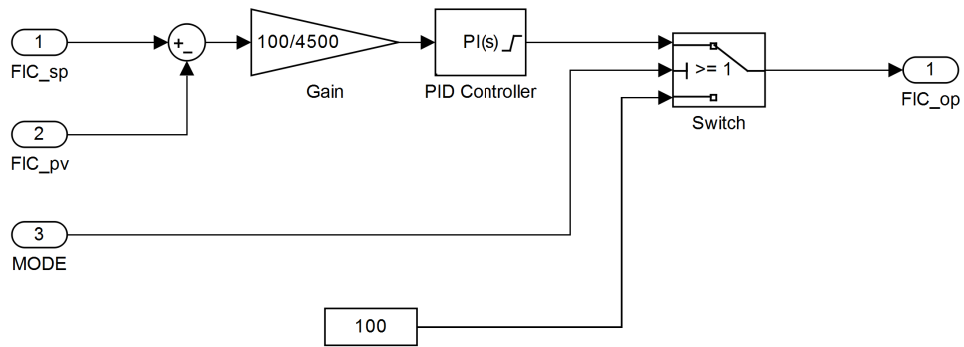


Figure A.14: Simulink PID logic.

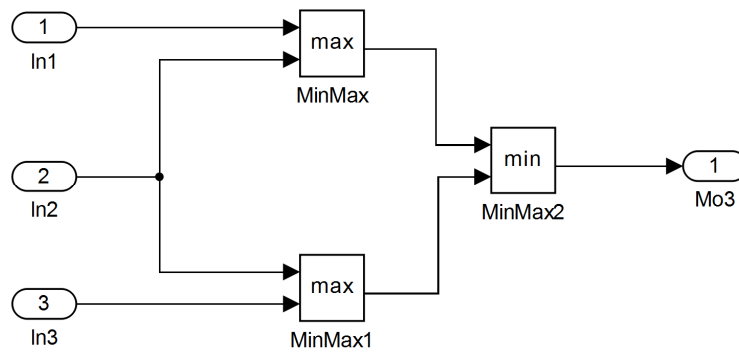
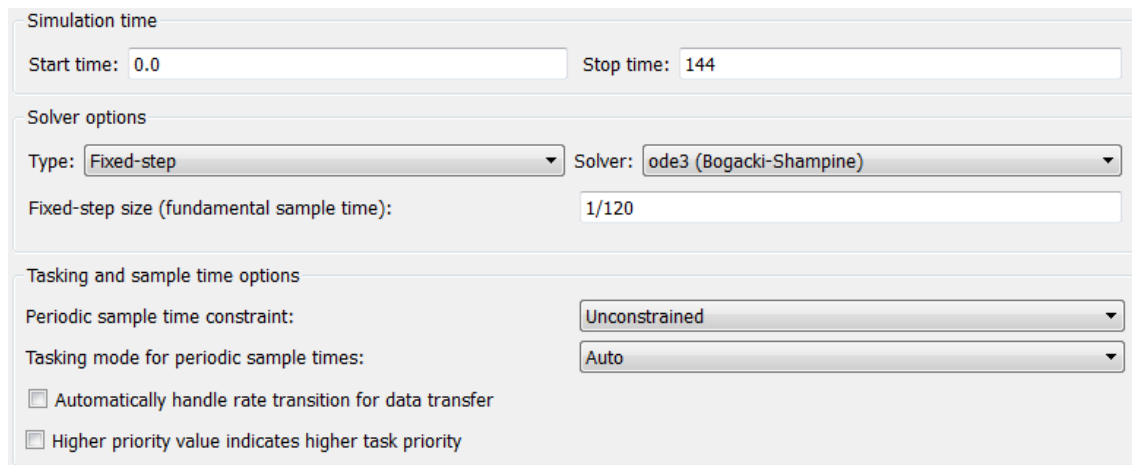


Figure A.15: Simulink mid-of-three logic.



The screenshot shows the Simulink solver settings dialog box. It is divided into three main sections:

- Simulation time:** Start time is set to 0.0 and Stop time is set to 144.
- Solver options:** The Type is set to 'Fixed-step' and the Solver is set to 'ode3 (Bogacki-Shampine)'. The Fixed-step size (fundamental sample time) is set to 1/120.
- Tasking and sample time options:** The Periodic sample time constraint is set to 'Unconstrained' and the Tasking mode for periodic sample times is set to 'Auto'. There are two unchecked checkboxes: 'Automatically handle rate transition for data transfer' and 'Higher priority value indicates higher task priority'.

Figure A.16: Simulink solver settings.

A.2.2 Parameter estimation using the `ga` function

For the parameter estimation, the verification approach described in Section A.2.1 is extended to execute repeatedly as part of an optimisation exercise. After setting upper and lower limits for the parameter set and specifying the genetic algorithm settings (population size, maximum number of generations and number of variables), the `ga` function call is performed with a pointer to the cost function. The cost function then executes the model using the parameter values chosen by the optimiser; the performance criteria is then evaluated with the data returned from the Simulink simulation. Two sets of performance criteria were used (in a mutually exclusive manner). The first set uses the negated sum of the correlation coefficients for the tempered water supply temperature, tempered water return temperature, tempered water differential temperature and duty. The second set uses the squared errors between the plant and simulation data for the tempered water supply temperature and tempered water differential temperature. The first set allows for the optimisation of the shape of the response whereas the second is used to achieve more accurate model gains. The results of the specified performance criteria calculation are then sent back to the optimiser which makes adjustments to the parameter set and repeats the process until the maximum number of generations is reached. The function then returns the set of optimised parameters to the workspace.

A.2.3 Control and optimisation

Four cases were considered namely base, ARC, HNMPC and EHNMPC. The base and ARC cases were implemented using the same Matlab file with the ability to switch between them by toggling a flag. The HNMPC and EHNMPC cases were also executed using the same file set where the main difference is in the definition of the cost function.

A.2.3.1 Base case and ARC case

The code starts off with the initialisation of some variables and importing of data files. After calculating the wet-bulb temperature for the simulation period and preparing the Simulink vectors, the model parameters and initial state are loaded. The Simulink model, which runs the simulation based on the input data provided, is then called. The resultant power, energy and cost values are then calculated followed by some trends. The type of simulation is set by toggling a flag value (determining whether to run the base or ARC case) and by setting the input data file (artificial versus plant data). The base and ARC cases run from a single Matlab file.

A.2.3.2 HNMPC case and EHNMPC case

Similar to the base and ARC case, the main code file starts off with the initialisation of variables and importing data files. Then, after calculating the wet-bulb temperature for the simulation period and preparing the Simulink vectors, the model parameters are loaded followed by the specification of the MPC parameters (prediction horizon, control horizon and iteration time). The initial state is then loaded and the model is run iteratively for the simulation period (one iteration time period at a time), each time only executing the first moves in the calculated control horizon before repeating the process.

During this step, several other Matlab functions are called. These define the initial manipulated variable values, run the optimisation problem (using the Matlab genetic algorithm function, `ga`), define the cost function and calculate the result and run the Simulink model.

The power, energy and cost values are then calculated followed by some trends. The type of simulation is set by adjusting the cost function used (determining whether the simulation is

an HNMPC or EHNMPC scenario) and by setting the input data file (artificial versus plant data).