Mixed integer linear programming for unit commitment and load dispatch optimisation

J van Niekerk

orcid.org/0000-0002-0761-959X

Dissertation submitted in fulfilment of the requirements for the degree *Master of Engineering in Industrial Engineering* at the North-West University

Supervisor: Prof SE Terblanche

Graduation ceremony May 2019

Student number: 23442212

Declaration

I, Jean-Pierre van Niekerk hereby declare that the entirety of the work contained in the electronically submitted thesis is my own, original work. I affirm that I am the sole author of the information contained in this document and that the sources used in the thesis were referenced duly in the document's bibliography. The publication and reproduction of the document content by the North West University of Potchefstroom will not result in the infringement of any third party rights. The entirety of the work was not submitted for the obtainment of any other qualification whatsoever and will only be submitted for the fulfillment of the requirements for the degree: Master of Engineering in Industrial Engineering.

Date of Declaration: March 28, 2019

Acknowledgments

The following institutions and people are to be acknowledged for their various contributions and support towards the completion of this research thesis:

- Prof SE Terblanche, at the North West University of Potchefstroom, for his assistance and guidance with the completion of this manuscript. Also for providing access to the commercial solver Cplex, enabling the author to develop and solve the unit commitment and load dispatch problem for power utilities.
- The subject matter experts in the South African power generation industry who provided the author with insight into the dynamics associated to the optimisation problem to be solved and how such a model is applied in a production environment.
- Annelie van Niekerk, my amazing wife, for her unconditional love and continued support during the two years of study. Her words of wisdom and encouragement made it possible to attempt and overcome this daunting task.
- Pierre and Christine van Niekerk as well as Jannie and Linda Du Plessis, my wonderful parents, for their love and motivation. They formed a significant support structure during this study period and without their self-sacrifice, this achievement would not have been possible.

Finally and most importantly, I thank God my savior for the ability He has bestowed upon me to develop, learn and grow. The completion of this thesis was only possible by the grace of God and as a result of the knowledge, wisdom, and insights obtained from Him alone. *Soli Deo Gloria!*

Abstract

The objective of solving the unit commitment and environmental economic load dispatch problem (UCEELD) for power utilities is to minimise the overall operational cost associated to power generation, while optimising the utilisation of natural resources. Power generation scheduling is however not a simplistic process as a multitude of aspects needs to be considered such as aging infrastructure, stringent emissions legislation, operational limitations and aligning base load with peaking station's scheduling. Apart from the financial objective, the optimisation problem is also focused on meeting the forecasted load demand of the power grid in an attempt to prevent grid instabilities. The intricacy of the scheduling and resource allocation process is significantly increased when a large power grid such as South Africa's grid is considered. Given the magnitude and complexity of the problem, a mathematical optimisation model was developed in this thesis applying mixed integer linear programming (MILP) as formulation technique and a commercial solver known as Cplex to obtain a proven global optimal solution to the mentioned problem. Specific emphasis was applied in using MILP instead of literature defined heuristic methods as these methods are not able to guarantee proven optimal solutions. Provided the nonlinearity of the UCEELD problem, the technique of piecewise linear approximation using binary variables were applied to linearise the nonlinear aspects of the problem with the aim of applying the MILP formulation. For the purpose of this thesis, only thermal, hydro and pumped storage generating technologies were considered for optimisation.

The contributions of this thesis were towards developing a realistically sized UCEELD model using MILP with the aim of incorporating the model into the production environment. Modeling contributions include the addition of thermal generation water consumption into the model objective function and incorporating stochasticity to the production model. The computational results provided in the thesis are based on the data obtained from a realistically sized power generation utility containing 98 thermal, 8 hydro and 6 pumped storage generating units. The model verification results confirmed that the proposed model is able to solve the optimisation problems accurately with the model response being as expected. From the validation results, it is observed that the proposed UCEELD MILP model is able to solve a realistically sized model to proven optimality within 44 minutes. A data handling tool comprising of a graphical user interface is also proposed to improve data acquisition, processing and incorporation into the optimisation model as well as interpretation thereof. The development of the UCEELD MILP model allows power utility management to effectively perform strategic decision-making within a short time frame to allow the optimisation of overall operational costs.

Keywords

Mixed integer linear programming, unit commitment and environmental economic load dispatch, optimisation, exact methods, piecewise linear approximation, and stochastic constraints.

Contents

List of Figures

List of Tables

List of Abbreviations

AC: Aspiration criteria

ACO: Ant colony optimisation

AEL: Atmospheric environmental license

AGC: Automatic generation control

BFP: Boiler feed pumps

BIP: Binary integer programming

CEELD: Combined economic emissions load dispatch

CO: Carbon monoxide

CO_2: Carbon dioxide

CV: Calorific value

DC: Direct current

ELD: Economic load dispatch

ESP: Electrostatic Precipitator

FD: Forced draught

FFP: Fabric filter plant

FGD: Flue gas desulphurisation

GB: Gigabyte

GL: Gig liters

GO: General overhaul

GUI: Graphical user interface

HD: High definition

HNN: Hopfield neural network

HP turbine: High pressure turbine

ID: Induce draught

IP: Integer programming

IP turbine: Intermediate pressure turbine

IR: Interim repair

LB: Lower bound

LP: Linear programming

LP turbine: Low pressure turbine

MAED: Multi-area economic dispatch

MCR: Maximum continuous rating

MILP: Mixed integer linear programming

MIP: Mixed integer programming

NLP: Non-linear programming

NO_2: Nitrogen dioxide

NP: Non-deterministic polynomial-time

O_2: Oxygen

PA: Primary air

PF: Pulverised fuel

PM: Particulate matter

PSO: Particle sward optimisation

PSS: Pumped storage scheduling

QIP: Quadratic integer programming

SA: Secondary air

SA: Simulated annealing

SO_2: Sulphur dioxide

STHS: Short term hydro scheduling

STHTS: Short term hydro-thermal scheduling

TL: Tabu List

UC: Unit commitment

UCEELD: Unit commitment and environmental economic load dispatch

UCELD: Unit commitment and economic load dispatch

UP: upper bound

VBA: Visual basic for applications

VP: Valve point

Chapter 1

Introduction

1.1 Background and rationale

Power utilities are companies in the electric power industry which engages in the generation and distribution of electricity. These utilities may include public or investor-owned entities. Such an entity which exists within South Africa is known as Eskom. This utility is the largest publicly owned power utility in South Africa and not only supplies electricity to the South African grid, but also distributes electricity to neighboring countries. The total installed capacity of this utility is approximated at 45 389 MW (Downs, 2012). Eskom uses a variety of methods for power generation with the aim of satisfying the grid demand. These methods include base load stations such as coal fired and nuclear units as well as peaking stations which include, but are not limited to, gas turbines, hydro, and pumped storage units. Base load stations refer to units that are used to maintain the minimum level of demand with uninterrupted supply. Peaking stations denote to units that are only occasionally utilised when the grid is either constrained or during peak hours of the day.

A general problem faced by power utilities in an effort to optimise resources and lower costs given a grid demand, is to decide which generating units to commit and the output load they should be dispatched at. This problem is known as the *unit commitment scheduling and load dispatching* process. The schedule development process is further complicated by factors such as:

1. Aging infrastructure that results in reduced operating efficiencies;
2. progressively stringent emissions legislation; and
3. operational limitations preventing unit loading adjustments (Hadji *et al.,* 2015).

Another factor to consider is when to commit peaking stations, in conjunction with base load stations, during scheduling and dispatching (Salama *et al.,* 2014). In order to comprehend the complexity of this decision-making process, a detailed discussion is provided on the challenges faced and how it influences resource utilisation and capital expenditure for a power utility such as Eskom.

1.1.1 Aging infrastructures

When generating unit commitment schedules for base load stations, such as coal fired units, the power utility needs to focus on each unit's operating efficiency in order to obtain the optimal utilisation of the available resources, at the lowest possible cost. The operating efficiency is governed by the unit's design envelope and determines the amount of coal (kg/s) consumed per megawatt generated (Li *et al.,* 2008). Depending on the design methodology used, a coal fired station's operating efficiency can range from anything between 30% - 37%. Note however that a units' operating efficiency can be reduced significantly as a result of aging infrastructure causing a decline in equipment performance. This reduction in efficiency is directly related to an increase in coal consumption (and possibly water utilisation), and subsequently a rise in the utilities' capital expenditure. The rate of deterioration of each coal fired unit is independent from one another as it is influenced by factors

such as the quality and frequency of maintenance execution per unit. Long term solutions to reduce a power utilities' coal consumption and increase the operating efficiency include an improvement of the aforementioned maintenance aspects. This will, however, be a costly initiative and is not always practical due to financial and time constraints. An alternative approach entails the commitment and dispatch of coal fired units by considering each unit's operating efficiency. Commitment of the most efficient units prior to the less efficient units can result in a significant decrease in coal consumption and consequently a reduction in the utilities' overall capital expenditure (Gunda & Acharjee, 2011). Another factor to consider independent from a unit's efficiency is the coal cost.

Due to aging infrastructure, coal fired units are also prone to increased water consumption. Each unit's water consumption increases with its individual rate of equipment deterioration (i.e. pipe leakages and passing valves). Although the cost of water usage does not have as large an impact as coal consumption, it still contributes to the overall operating costs incurred by the utility. In conjunction with the cost element of increased water consumption, another aspect that needs to be emphasized is the fact that South Africa is a semi-arid, water stressed country. It is stipulated by Wassung (2010) that South Africa has a limited fresh water supply of 13 227 million m^3 with a current water demand of 12 871 million m^3. When referring to the industrial sectors of South Africa, the quantity of water used by Eskom (excluding Medupi and Kusile) is estimated at approximately 273 million m^3 per annum. This accounts for 2% of the total fresh water available in South Africa (Vasanie, 2004). Water demand is constantly increasing due to the growth in human population and the expansion of industrial sectors. Still *et al.* (2008) stated that to meet the increasing demand of water supply in 2050, alternative methods will be required to reduce the commercial demand for water and to become more efficient in the use of available water resources. This can be addressed by implementing alternative water saving methods such as deciding which coal fired units to schedule depending on its water usage. This can reduce the water consumption and will enable the utility to manage its water resources more effectively and ensure that a reduction in water usage is realised. In turn, it will also assist the South African community with the management of its scarce resources. When implementing such a decision-making method the units with low water consumption will be committed prior to scheduling units with high water consumption.

1.1.2 Stringent emissions legislation

An additional factor which needs attention when developing commitment and dispatch schedules for coal-fired units; is the environmental impact. Electricity generation by a coal-fired power utility results in the production of emissions which has a negative impact on the environment. Depending on the plant conditions and the load at which each unit is loaded, it can result in either a decrease or increase in emission production. The maximum emission rates (Particulate matter, SO_x and NO_x) for a solid biomass combustion installation (i.e. a thermal power station) is governed by the *Atmospheric Emissions License* (AEL). The emission limits are provided in Section 43 of the National Environmental Management Act, Act number 39 of 2004 which is summarised below:

1. *Particulate matter* (PM): 50 mg/Nm^3
2. SO_2: 500 mg/Nm^3 and
3. NO_x: 750 mg/Nm^3.

As per the AEL, all coal-fired power stations (existing and newly built) must comply with the above-mentioned limits. Eskom has obtained postponement for existing power stations to comply to these legislative regulations for PM and NO_x emissions until the end of 2020, and until 2050 for SO_2 emissions (South-Africa, 2004). Myllyvirta (2014) reported that due to the postponement of all existing stations within Eskom's fleet in meeting the required emission limits, Eskom is allowed to emit an excess of 560000 tons of PM, 28 million tons of SO_2 and 2.9 million tons of NO_x. In allowing Eskom this leeway, it negatively affects the agricultural environment, health of humans and other living organisms and results in the degradation of the structural integrity of metals and buildings (Rall, 1974; Treshow, 1980; Winner *et al.,* 1985). This margin was granted to Eskom since it will

take some years to improve the existing stations to meet the new emission targets. It will require stations to improve ESP and bag filter performance, construction of FGD systems, switching fuel qualities and optimising SO_3 injection rates.

Until the above-mentioned has been achieved by Eskom, the power utility can make use of optimal commitment and load dispatch units scheduling as an alternative method in an attempt to reduce the effect of the current emissions (Mandal *et al.,* 2015). By committing units with the best emission performance prior to committing bad performing units, Eskom can minimize its environmental impact and ensure that the AEL is adhered to at all times. Implementing the foregoing will assist with optimising the utilisation of the natural resources for power generation (reducing water and coal usage), while minimizing the production of undesirable by-products.

1.1.3 Operational limitations

When developing a dispatch and commitment schedule for coal fired units, it is the responsibility of the system operator to not only consider the aspects as mentioned throughout Sections 1.1.1 and 1.1.2, but to also take into consideration operating limitations. This include, but is not limited to, factors such as the minimum-up and downtime of each unit, the rate at which each unit can respond to load change instructions, prohibited operating zones at which a unit might not be allowed to operate within, due to reduced component reliability, as well as outage schedules preventing units from being committed to the grid (Jadoun *et al.,* 2015; Li *et al.,* 2008; Norouzi *et al.,* 2014; Yang *et al.,* 2012). Each one of these aspects has a significant influence on the unit commitment and load dispatch scheduling process and therefore cannot be omitted by the system operator. By including the mentioned constraints in the decision-making process, will not lead to any financial gain or improved resource utilisation as mentioned in the above sections, but will ensure that no unit is operated outside its allowable operating envelope. If however any of these aspects are omitted, it can lead to unsafe and unstable operating conditions, augmented mechanical wear on already damaged components and in worst case might lead to multiple unit trips. It is therefore apparent that by including these aspects into the system operator's decision-making process, significant financial losses may be prevented.

1.1.4 Aligning base load and peaking stations' scheduling

In addition to the above-mentioned aspects, the system operator needs to align the commitment and dispatch schedule of peaking stations, such as hydro and pumped storage stations, with that of the base load stations (Chen, 2008). Although peaking technologies are more cost efficient in comparison to base load stations it has a limited power capacity which can be supplied to the grid (Salama *et al.,* 2013) and cannot be schedule uninterruptedly. Restricted power capacity in the case of hydro stations is governed by the department of water affairs. The amount of water discharged in a certain period is limited in order to reduce the effect it has on the downstream ecosystem. For pumped storage stations, limited power capacity is a result of the upper reservoir design volume constraints. Though these limitations exist, it is still beneficial for a power utility to incorporate these technologies into the generation schedule for the period at which the peaking stations' capacity is available, as it will result in significant financial gain (i.e. substantial saving in coal consumption).

The inclusion of peaking stations into the generation schedule does, however, increase the complexity of the decision-making process significantly. The reason being is that the system operator needs to decide when to commit and dispatch each peaking unit at a time where the most financial gain will be obtained. In addition, the system operator also has to ensure that each peaking unit contains enough capacity to supply the grid with sufficient power at the time of commitment. If the power capacities supplied by the peaking stations are not enough due to poor decision-making and capacity management, and the base load stations cannot be deployed timeously due to operating constraints, it may result in grid instabilities and even lead to a grid collapse. This will not only affect the power utility negatively but will also have an irreversible financial impact on the South

African economy.

1.1.5 Power generation schedule

To develop an optimal generation schedule where the available units are committed and dispatched with minimal operational cost, and optimal resource utilisation, the aspects mentioned in Sections 1.1.1 to 1.1.4 need to be considered. Failing to include one of the mentioned factors into the schedule development process, will result in an unrealistic and impractical operating schedule being obtained. Using an incorrect schedule for unit commitment and load dispatch decision-making, will either lead to suboptimal utilisation of the available resources or it can lead to substantial financial losses being incurred by the utility. It is therefore imperative that the system operator, when developing the generation schedule, incorporate all necessary factors in order to obtain a realistic and representative schedule.

1.2 Research scope

In the subsequent section, specific focus is set on the problem statement that needs to be investigated, a discussion is provided on the purpose of the study and a summary is given regarding the research objectives to be addressed during the development of the thesis.

As motivated in the previous sections, effective unit commitment and load dispatch scheduling is essential in, ensuring the minimisation of a utilities capital expenditure, as well as optimising the utilisation of natural resources. Power generation scheduling is however not a simplistic process as a multitude of aspects (as mentioned throughout Section 1.1) needs to be considered, prior to deciding which unit to commit and at what load the unit needs to be dispatch at. The intricacy of the scheduling and resource allocation process is significantly increased when a large power grid such as South Africa's grid is considered. This is due to the exponential increase in the data acquisition and computing power which is required to solve such a power generation scheduling problem. The magnitude of the problem and the complexity associated with the scheduling process will make it practically impossible to create a power generation schedule using manual computations. This process will be time-consuming as well as labor-intensive and will have a high probability of obtaining sub-optimal solutions.

1.2.1 Research problem statement

To address the above concerns, the researcher will need to develop a mathematical model capable of deciding which units to commitment and at what load each unit needs to be dispatched at, as to minimise capital expenditure and improve resource utilisation. The model will also need to be capable of automatically generating an excel based generation schedule from the obtained results, to provide the system operator with a consolidated tool when performing strategic decision-making. A graphical example of the above statement is provided in Figure 1.1.

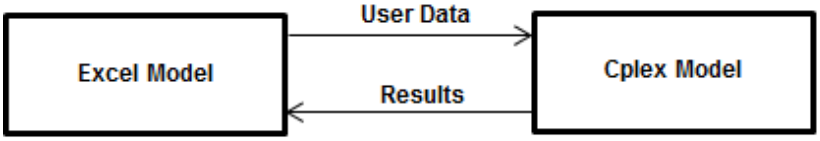

Figure 1.1: Computational process diagram

The aim of the consolidated excel generation schedule is to save time and provide users who are not proficient with Cplex, the capability of utilising and interpreting the generated schedule. By implementing such a model, it will promote the reduction of capital expenditure and workforce requirements, prevent unnecessary scheduling errors, and guarantee the satisfaction of the grid demand

in an optimal and efficient manner within the power utility production environment.

In practice, as well as in the literature, the application of heuristic approaches is very popular towards solving the unit commitment and load dispatch scheduling problem. These approaches typically generate random solution populations which at times only give the local maxima or minima as results. This is not ideal, as the definite optimal point of the solution needs to be obtained (Ashfaq & Khan, 2014) to prevent significant capital losses. The solution algorithms associated with heuristic methods are generally intertwined with the mathematical model which prevent the method from being applied to any problem instance without requiring some programming alterations. This leads to majority of the heuristic methods being incapable of accurately solving variations of the original problem instances it was designed for. Given the foregoing reasons, mixed integer linear programming (MILP) will be applied in this study in an attempt to solve the mentioned scheduling problem and obtain a global optimal solution for the different scenarios.

1.2.2 Research purpose

The purpose of this study is to solve the unit commitment and load dispatch problem by employing MILP technology for computing optimal generation schedules, aimed in reducing a power utilities' capital expenditure and improving its resource utilisation. In addition, aspects such as, aging infrastructure, environmental legislation, and operational constraints, as well as base load and peaking commitment co-ordination will be incorporated into the model formulation, as to guarantee its relevancy to real power generation applications. The study is, however, limited to coal fired, hydro and pumped storage power generation and do not take into consideration technologies such as nuclear, gas and wind turbines, and photovoltaic plants.

The anticipated contributions that will be made throughout this study will include the following:

1. Follow a linearisation model approach when formulating the UCEELD problem in order to simplify the nonlinear model complexity. By implementing the preceding, it will allow the researcher to easily incorporate stochastic scenarios into the optimisation model.

2. Develop a comprehensive UCEELD model which incorporate thermal, hydro and pumped storage generating units simultaneously. Throughout literature, it could not be identified that such a comprehensive model already exist. Literature models include either one or two of the above generation technologies and seldom include all three technologies in such extensive detail.

3. Applying the linearisation methodology to the UCEELD problem allows the researcher to determine the integrality gap obtained from the solution. By evaluating the aforementioned, one can determine if an exact global solution has been obtained. Using heuristic methods to solve similar problems does not allow the user to evaluate the integrality gap, and could lead to local optimums.

4. Adding demineralised water consumption rates to the thermal generating units' objective function, allows the model to not only optimise fuel cost or emissions production as was done in literature, but also allows the reduction of water usage in thermal power stations.

5. Adding outage constraints to the UCEELD problem allows the model to discount the units that are on outage and provides the user with an interactive and dynamic model. Throughout literature, the preceding constraints were not added to the proposed committment and load dispatch models.

1.2.3 Research objectives

The study objectives entail the following:

1. Provide a thorough survey of i) the mechanics associated to both exact and heuristic solution algorithms, and ii) a review of the mathematical advances stated in literature pertaining to the solution of the *unit commitment and environmental, economic load dispatch problem* (UCEELD) for power utilities.

2. Develop a data handling tool comprising a graphical user interface (GUI), to allow for ease of data acquisition, processing, and incorporation of user defined data inputs into the optimisation model.

3. To formulate a unit commitment and load dispatch MILP model in an interactive development software platform, which will be capable of reducing a power utilities' capital expenditure and improving its resource utilisation.

4. Verify the correctness and validate the effectiveness of the MILP model in obtaining optimal results when applied to realistic problem instances derived from real power generation data. Also evaluate the influence that different parameters, data inputs and modeling constraints might have on the solution time (i.e. model complexity) and modeling results.

5. Suggest follow-up work related to the work completed in this dissertation, worth being investigated when conducting future studies associated with UCEELD problem optimisation.

1.3 Research methodology

The research methodology provides an overview of the important literature topics which needs to be reviewed and the methods to be applied by the researcher in order to address the research objectives. The research methods are systematically outlined in order to address each of the objectives as stated in Section 1.2.3. Details regarding the foregoing is provided throughout Sections 1.3.1 and 1.3.2.

1.3.1 Literature review

In this dissertation, the literature review is presented in two parts and is covered by both Chapters 2 and 3. In Chapter 2, the focus is set on exact solution methods such as linear and integer programming. A detailed analysis is provided regarding the standard formulation of a linear programming problem and the primal simplex solution method is discussed. The mathematical approach to the simplex method is investigated with a detailed elaboration of the theory behind the method. The simplex method forms the basis from which many solution algorithms are derived from (Winston & Goldberg, 2004; Corneujols & Tutuncu, 2007). In addition, the standard formulation of both integer and mixed integer programming problems are provided, and the methodology of linearising optimisation problems by using logical modeling with binary variables are discussed. An overview of the branch-and-bound method applied to mixed integer problems is also provided in this section, with specific focus on the mathematics behind the solution method (Winston & Goldberg, 2004; Corneujols & Tutuncu, 2007). In Chapter 2, a section on heuristic solution methods where alternative approaches, specifically used to solve the UCEELD problems, are included with only a brief discussion of each (Mandal *et al.,* 2015; Senthil & Manikanda, 2010; Basu, 2005).

In Chapter 3, a detailed discussion is provided regarding the power generation operational overview and design considerations for coal-fired (Govidsamy, 2013), hydro (Ferreres & Font, 2010), and pumped storage stations (Antal, 2004). This entails a brief introduction to each power generation technology, an overview of the mechanical components utilised for power generation and challenges that are experienced daily, which contribute to unit commitment and load dispatch scheduling difficulties. In addition to the preceding, a literature review of the UCEELD problem is provided with a comprehensive analysis on the solution methods applied by other researchers to solve this problem. The literature review is utilised as a means of analyzing the research advances made in this field of study, and to provide insight into the application of MILP to power generation scheduling optimisation (Borghetti *et al.,* 2008; Norouzi *et al.,* 2014).

1.3.2 Research method(s)

To address the project objectives as outlined above, the subsequent research methodology is applied:

1. To gain fundamental knowledge of the subject matter under consideration, a comprehensive literature and theoretical background study are performed, regarding optimisation solution methods and its applicability to UCEELD problems. By investigating the foregoing, a detailed understanding can be obtained concerning the advances made in the specific field of study. The study will also assist in identifying which technical aspect needs to be considered when developing a MILP UCEELD solution model, as to obtain satisfactory results, which is comparable to existing models.

2. After identifying the solution methodology to be applied to the UCEELD problem, a data handling tool will be developed by means of using the software package Microsoft Excel (visual basics). This tool will be used to filter, process, and arrange input data in such a manner as to ensure compatibility to the optimisation software. The data handling tool will be equipped with a GUI to improve the efficiency of the data handling process. Design data obtained from realistic power generation scenarios will be used as inputs to this tool.

3. The software package, Cplex (IBM Corp, 2015), will be used as a solver to solve the UCEELD problem. Aspects identified throughout the literature review will be included into the problem formulation so as to ensure the comprehensiveness of the model. The model will be structured in a dynamic manner to allow the inclusion and/or exclusion of different constraints (parameters) to identify the effect each one has on the outcome of the model results. The optimised results obtained from each problem instance will be made available to the data handling tool to enable further processing and interpretation.

4. The results obtained from each model scenario will be compared and a conclusion drawn to determine the effect different scenarios will have on a utilities' power generation schedule. Finally, a comprehensive model will be proposed which can be utilised for optimising the commitment and dispatch decision-making process of coal-fired, hydro, and pumped storage stations to reduce capital expenditure and improve resource utilisation.

1.4 Provisional chapter division

The introduction is followed by five chapters, a bibliography, and several appendixes. In Chapter 2, a discussion is provided regarding general optimisation solution algorithms with focus on both exact and heuristic solution methods. In this chapter, a literature review pertaining to the mathematical reasoning and the theory behind exact methods such as linear and integer programming models as well as heuristic methods which include alternative programming models is provided. Simplistic examples are included to gain a fundamental understating of the computational dynamics behind each method.

An introduction to the fundamentals of power generation is depicted in Chapter 3, with an overview explaining the essence of the UCEELD problem as well as its applicability. Information is provided regarding the terminology and the technical aspects associated with coal fired, hydro and pumped storage power generation. The preceding is included to gain an understanding of the dynamics associated with each technology, and to provide insights into the problem to be solved. A literature review of the optimisation techniques generally applied to the UCEELD problem is presented. In addition, focus is set on the application of MILP and how it can be applied in solving power generation scheduling problems.

Chapter 4 presents the proposed mathematical construction of the MILP model applied to the UCEELD problem in order to minimise the operational cost of a power utility. In this chapter a number of basic notations are discussed in order to ensure the simplicity of the model derivation. The

model objectives and constraints are also defined with descriptive reasoning behind each constraint, clarifying why it was added to the optimisation model. Examples are provided on the mechanics which exist between the different constraints and how each contributes to the computation of an optimal solution.

Computational results obtained from the MILP model are provided in Chapter 5 as a means of validating and verifying the effectiveness of the proposed model in solving different arrangements of the UCEELD problem. Different UCEELD problem instances are considered for the purpose of the dissertation with focus on how the model results and solution time vary with the addition of constraints and or data set sizes. The applicability of the proposed mathematical MILP model to solve realistic instances of the UCEELD problem is also discussed. The model input data required to solve the UCEELD problem arrangements are elucidated, and a brief overview is provided on the GUI designed to simplify data acquisition and processing.

The dissertation ends in Chapter 6 with a final summary of the work conducted, a conclusion on important findings and an overview of the contributions made to the field of study. Suggested future work related to UCEELD problems is also proposed.

Chapter 2

Optimisation Solution Approaches

Numerous mathematical optimisation approaches are employed by researchers in solving different variations of the UCEELD MILP problem. These approached can be characterized as either exact or heuristic solution algorithms. Depending on the prerogative of the researcher and the aim of the research study, either one of these approaches can be applied. There are however some distinct differences between the two approaches that need to be considered prior to implementation, with the specifics of each approach being summarised in Table 2.1.

Table 2.1: Solving the MILP Problem

Exact Approaches	Heuristic Approaches
Implemented as branch-and-bound/cut algorithms within commercial solvers such as Cplex, Gurobi, etc.	Custom build algorithms based on mathematical techniques such as particle swarm optimisation, neural networks, etc.
Able to solve problems up to proven optimality	Optimality of solution cannot be proven
Comprise on the ability to compute the distance of intermediate solutions from optimal solutions	Capable of comparing intermediate solutions to bound from exact relaxation
Computationally expensive when applied to medium to large scale problems	Able to provide relatively accurate solutions within a short time period.
Distinct separation between the model and solver	Model changes might necessitate solver changes as well

When evaluating the differences between the two approaches, it becomes apparent that the exact solution approach will be more suitable to implement with the aim of satisfying the research objectives as stated in Section 1.2.3. In this study, focus is set in obtaining a proven optimal solution within a reasonable time frame using a commercial solver known as Cplex. Although exact solution approaches will be the emphasis for the remainder of the study, a brief overview of the heuristic approaches is also provided in this chapter as to ensure that the reader obtains a broad understanding of the available literature.

Exact solution methods

Exact solution approaches are algorithms utilised to solve optimisation problems to proven optimality. In the case of linear programming, an example of an exact approach is the primal simplex method, and in the case of integer or mixed integer programming, the branch-and-bound method applies. Although a multitude of exact solution methods exists within the optimisation domain, the focus of this literature review is limited to the simplex and brand-and-bound methods, since the commercial solver used to generate the results in this study are based on these two solution approaches.

2.1 Linear programming methods

Linear programming (LP) is an optimisation method, developed during the 19 hundreds, aimed at providing optimal solutions to problems, where the objectives and constraints of the problems are linear functions of the decision variables. The development of the linear programming theory was to enable users to make optimal decisions when faced with complex situations. The initial recognition of the linear programming type of problem was made by economists in the 1930s while developing methods to allocate available resources to reduce financial expenditure. Significant progress was made since then to improve the practical applications and theoretical development of the solution methodology. One of these advances was made in 1939 by L.V. Kantorovich a member of the soviet Union together with T.C. Koopmans from the US. They obtained a Nobel Prize in the field of economics for their contribution to the economic interpretation of linear programming focused on resource allocation. During the 1940s, L. Kantrovich, Johan von Neumann and George Dantzig created the mathematical subfield of linear programming which layed the foundation for Dantzig to invent the primal simplex method while working at the US Air Force (Dantzig, 2002). Although many other LP solution methods had been developed over the past few years, the primal simplex method still remains the most popular and effective method in solving LP problems. Additionally, the primal simplex solution method also forms the basis from which majority of the LP solution methods as well as integer and mixed integer solution methods were derived from. During the early ages linear programming application was primarily focused on solving problems in the petroleum refineries, manufacturing, food-processing and engineering design industries. However, this methodology has developed to such an extent that it can be applied to any industry imaginable to solve complex strategic decision-making problems (Rao, 2009).

In mathematical terms, a LP problem entails the optimisation of an objective function $\boldsymbol{c^T x}$ by means of finding a vector $\boldsymbol{x} \in \mathbb{R}^n$ where constraints $\boldsymbol{Ax} \leq \boldsymbol{b}$ are satisfied. The column vectors can be defined as $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$ and matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. A feasible solution to such a problem can be defined as a solution where vector $\boldsymbol{x^*} \in \mathbb{R}^n$ whereas an optimal solution would be where vector $\boldsymbol{x^*}$ results in $\boldsymbol{c^T x^*} \geq \boldsymbol{c^T x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Note that the preceding is defined as an optimal solution for a maximisation problem. When considering a minimisation problem, vector $\boldsymbol{x^*}$ would result in $\boldsymbol{c^T x^*} \leq \boldsymbol{c^T x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. In order to apply a solution method such as the primal simplex method to obtain an optimal solution to $\boldsymbol{c^T x}$, it is required to rewrite the LP problem in its standard form.

2.1.1 LP standard form formulation

The standard form of the linear programming problem can be portrayed by using two methods that include both the scalar or matrix forms. For more in-depth information regarding the LP standard forms and transformation rules, technical information complied by researcher such as Rao (2009), Corneujols & Tutuneu (2007), Lewis (2008), Winston & Goldberg (2004) and Feige (2011) can be sited. The scalar standard form is portrayed below using equations (2.1) - (2.3).

Minimise objective function

$$f(x_1, x_2, ..., x_n) = c_1 x_1 + c_2 x_2 + ... + c_n x_n \tag{2.1}$$

subjected to constraints:

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + ... + a_{1n} x_n &= b_1 \\ a_{21} x_1 + a_{22} x_2 + ... + a_{2n} x_n &= b_2 \\ &\vdots \\ a_{w1} x_1 + a_{w2} x_2 + ... + a_{wn} x_n &= b_w \end{aligned} \tag{2.2}$$

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ &. \\ &. \\ x_n &\geq 0 \end{aligned} \tag{2.3}$$

where $\boldsymbol{a}_{ij}$, $\boldsymbol{b}_i$ and $\boldsymbol{c}_j$ (i = 1,2,...,w; j = 1,2,...,n) are known constants, and $\boldsymbol{x}_j$ are defined as the decision variables with $\boldsymbol{x} \in \mathbb{R}^n$. Simplified to matrix form, the LP problem can be depicted as follow:

Minimise objective function

$$f(\boldsymbol{x}) = \mathbf{c^T}\boldsymbol{x} \tag{2.4}$$

subjected to constraints:

$$\begin{aligned} \mathbf{A}\boldsymbol{x} &= \mathbf{b} \\ \boldsymbol{x} &\geq 0 \end{aligned} \tag{2.5}$$

where

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ b_w \end{bmatrix}; \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{bmatrix}$$

and

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ . & & & \\ . & & & \\ a_{w1} & a_{w2} & ... & a_{wn} \end{bmatrix}$$

Alternatively, a condensed notation of the linear programming standard matrix form could be represented by $min\{\boldsymbol{c^T}\mathbf{x} : \boldsymbol{A}\mathbf{x} \leq \boldsymbol{b}, \mathbf{x} \in \mathbb{R}^n\}$.

If faced with a linear programming problem which is not stated in standard form, there are some transformation rules which needs to be applied in order to convert the problem to standard form. When referring to standard from, it entails the adherence of the LP problem to the following requirements:

1. An objective function which is of the minimisation type;
2. All decision variables are non-negative; and
3. All constraints are of the equality type.

To express any non-standard linear programming problem in the standard form as mentioned above, the subsequent transformation rules can be applied:

1. A linear programming maximization function, can be transformed to a minimisation function by means of multiplying the maximization function with a negative value throughout (i.e. -1). Refer to the below example:

   The maximization function

   $$f(\boldsymbol{x}) = -\mathbf{c^T}\boldsymbol{x}$$

is equivalent to the minimisation of

$$f(\boldsymbol{x}) = \mathbf{c^T}\boldsymbol{x}$$

Consequently, any linear programming problem in non-standard form can be written in standard form (minimisation type) by applying the preceding principle.

2. If an optimisation constraint is stated as an inequality constraint of the "greater than or equal to" type such as:

$$\mathbf{A}\boldsymbol{x} \geq \mathbf{b}$$

the constraint can be transformed to its equality form by means of subtracting a non-negative slack variable as depicted below:

$$\mathbf{A}\boldsymbol{x} - \mathbf{s} = \mathbf{b}$$

If however, the optimisation constraint is stated as an inequality constraint of the "less than or equal to" type such as:

$$\mathbf{A}\boldsymbol{x} \leq \mathbf{b}$$

the constraint is required to be converted to its equality form by means of adding a non-negative slack variable:

$$\mathbf{A}\boldsymbol{x} + \mathbf{s} = \mathbf{b}$$

By applying a slack variable to an inequality constraint, the user is able to transform the constraint to its standard form. Note however that the type of inequality constraint present in the LP problem will determine if a slack variable needs to be subtracted or added to facilitate the transformation.

3. When considering optimisation problems within the engineering field, it is apparent that the decision variables usually represent some physical dimensions of the non-negative type. However, there are cases where decision variables may be unrestricted in sign (can either take on positive, negative or zero values). In such cases, these unrestricted variables needs to be written as the difference between two non-negative decision variables. In example, if variable x_n is unrestricted in sign, it can be converted to standard form by means of applying the following:

$$\boldsymbol{x_n} = \boldsymbol{x_n^{'}} - \boldsymbol{x_n^{''}}$$

where

$$\boldsymbol{x_n^{'}} \geq 0 \; and \; \boldsymbol{x_n^{''}} \geq 0$$

Implementing variables $\boldsymbol{x_n^{'}}$ and $\boldsymbol{x_n^{''}}$, the user is able to convert the unrestricted decision variables into bounded non-negative decision variables that correspond to the prerequisites for the LP standard form formulation.

After applying the LP standard form transformations to non-standard form LP problems, the user is able to utilise a multitude of LP solution methods to obtain an optimal solution. One such method is the Primal Simplex algorithm.

2.1.2 Primal simplex method

Literature regarding the primal simplex method, compiled by researchers such as Corneujols & Tutuneu (2007), Winston & Goldberg (2004), Hua (1990) and Gartner (1995) were investigated with the aim of ascertaining knowledge regarding the mathematical derivation behind the simplex method. In these references the initialization point of the simplex method starts with a set of equations which includes the objective function and problem constraints portrayed in canonical form. Therefore, let us consider a general LP problem in standard form, focused on minimising objective function *f(**x**)* subjected to constraints:

$$\begin{aligned} a_{11}x_1 + ... + a_{1n}x_n + 1s_1 + 0s_2 + 0s_w &= b_1 \\ a_{21}x_1 + ... + a_{2n}x_n + 0s_1 + 1s_2 + 0s_w &= b_1 \\ &. \\ &. \\ a_{w1}x_1 + ... + a_{wn}x_n + 0s_1 + 0s_2 + 1s_w &= b_w \\ x_1...x_n, s_1...s_w \geq 0, \quad x \in \mathbb{R}^n, s \in \mathbb{R}^w \end{aligned} \tag{2.6}$$

The standard scalar LP problem as depicted above is converted to matrix form in order to derive the simplex solution method. By rewriting the LP problem in matrix form , the variables depicted in equation (2.6) can be represented as vectors and matrices (similar to what was mentioned in the LP standard form formulation section). Additional to the already defined vectors and matrices, the slack variables added to the LP problem can be denoted by variable s which is a w-dimensional column vector:

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ . \\ . \\ s_w \end{bmatrix}$$

Let $\boldsymbol{I}$ signify a $w \times w$ identity matrix. Considering the preceding, the LP problem equality constraints can be condensed to the subsequent:

$$\begin{bmatrix} \mathbf{A}, \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{s} \end{bmatrix} = \mathbf{b}, \quad \begin{bmatrix} \mathbf{x} \\ \mathbf{s} \end{bmatrix} \geq 0$$

Alternatively, the equality constraints for the LP problem can be written as

$$\begin{bmatrix} \mathbf{B}, \mathbf{N} \end{bmatrix} \begin{bmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{bmatrix} = \mathbf{B}\mathbf{x_B} + \mathbf{N}\mathbf{x_N} = \mathbf{b},$$

with the subscripts "B" and "N" referring to the basic and nonbasic variables respectively. In order to solve the equality constraint in terms of x_B, both sides of the equation is multiplied with $\boldsymbol{B}^{-1}$ to obtain the following:

$$\mathbf{x_B} + \mathbf{B}^{-1}\mathbf{N}\mathbf{x_N} = \mathbf{B}^{-1}\mathbf{b}$$

Solving for variable x_B, a simplified equation to calculate the basic variables are formulated. The subsequent equation is the first step in the simplex algorithm which is used to choose an initial basis.

$$\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x_N}$$

The same principle can be applied to the objective function by means of using the basis partitioning. Prior to partitioning, the objective function is set equal to zero, to obtain an equivalent representation of the initial objective function used for the derivation of the simplex method:

$$\mathbf{Z} = \mathbf{c}\mathbf{x} \Leftrightarrow \mathbf{Z} - \mathbf{c}\mathbf{x} = 0$$

By setting variable $\boldsymbol{c} = [\boldsymbol{c_B}, \boldsymbol{c_N}]$, and substituting it into the above equation, the objective function is altered in order to obtain:

$$\mathbf{Z} - \begin{bmatrix} \mathbf{c_B}, \mathbf{c_N} \end{bmatrix} \begin{bmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{bmatrix} = 0$$

Multiplying the two matrices and substituting variable $\boldsymbol{x_B}$ into the above equation, the following sequence of equations are obtained:

$$\mathbf{Z} - \mathbf{c_B x_B} - \mathbf{c_N x_N} = 0$$
$$\mathbf{Z} - \mathbf{c_B}(\mathbf{B^{-1}b} - \mathbf{B^{-1}N x_N}) - \mathbf{c_N x_N} = 0$$

By applying mathematical factorization to solve for $\boldsymbol{c_B B^{-1} b}$, a solution is obtained of the form:

$$\mathbf{Z} - (\mathbf{c_N} - \mathbf{c_B B^{-1} N})\mathbf{x_N} = \mathbf{c_B B^{-1} b}$$

Vector $(\boldsymbol{c_N} - \boldsymbol{c_B B^{-1} N})$ stated in the objective function is know as the reduced cost, due to the fact that the cost coefficients $(\boldsymbol{c_N})$ are reduced by the cross effect induced by the basic variables $(\boldsymbol{c_B B^{-1} N})$. Using the above equation, the effect of adjusting a nonbasic variable on the objective function can be determined. The answer obtained from the vector calculation after substituting the matrix values into the equation, the user is able to determine which non-basic variable needs to enter the basis. The variable to enter is the one for which the value $(\boldsymbol{c_N} - \boldsymbol{c_B B^{-1} N})$, is the most negative (for maximization problem it is the most positive). This calculation process forms part of the second step of the simplex method.

Additional to the foregoing, vector $(\boldsymbol{c_N} - \boldsymbol{c_B B^{-1} N})$ is also utilised as a stop criteria for the primal simplex method. Meaning, when considering a linear maximization problem where

$$(\mathbf{c_N} - \mathbf{c_B B^{-1} N}) \leq 0$$

the solver will abort as it will be indicative of an optimal solution. If however a minimisation problem is considered, the algorithm will continue until the subsequent is satisfied:

$$(\mathbf{c_N} - \mathbf{c_B B^{-1} N}) \geq 0$$

where after the simplex method will terminate to provide the user with an optimal solution. If however these stop criteria has not yet been satisfied, the user will need to perform a ratio test in order to select which existing basic variable will be required to leave the basis. This is done by implementing the following:

$$Ratio = \frac{\mathbf{B^{-1}b}}{\mathbf{B^{-1}N x_N}}$$

The variable that will be selected to leave the basis is the one for which $\frac{\boldsymbol{B^{-1}b}}{\boldsymbol{B^{-1}Nx_N}}$ is the minimum after the matrix values have been substituted into the equation and an answer has been obtained. Note however that only positive values are considered to leave the basis as negative values are indicative of extreme points. Also take into consideration that if the ratio test delivers only negative values, the system is unbounded. The preceding forms the third and final step in the primal simplex algorithm. After calculating the variable which is required to leave the basis, the simplex calculation process is re-initiated using the newly obtained bases as inputs. The solution algorithm then iterates through the mentioned steps until an optimal solution is obtained. Refer to Figure 2.1, for a detailed flow diagram of the complete primal simplex algorithm applied to a minimisation problem.

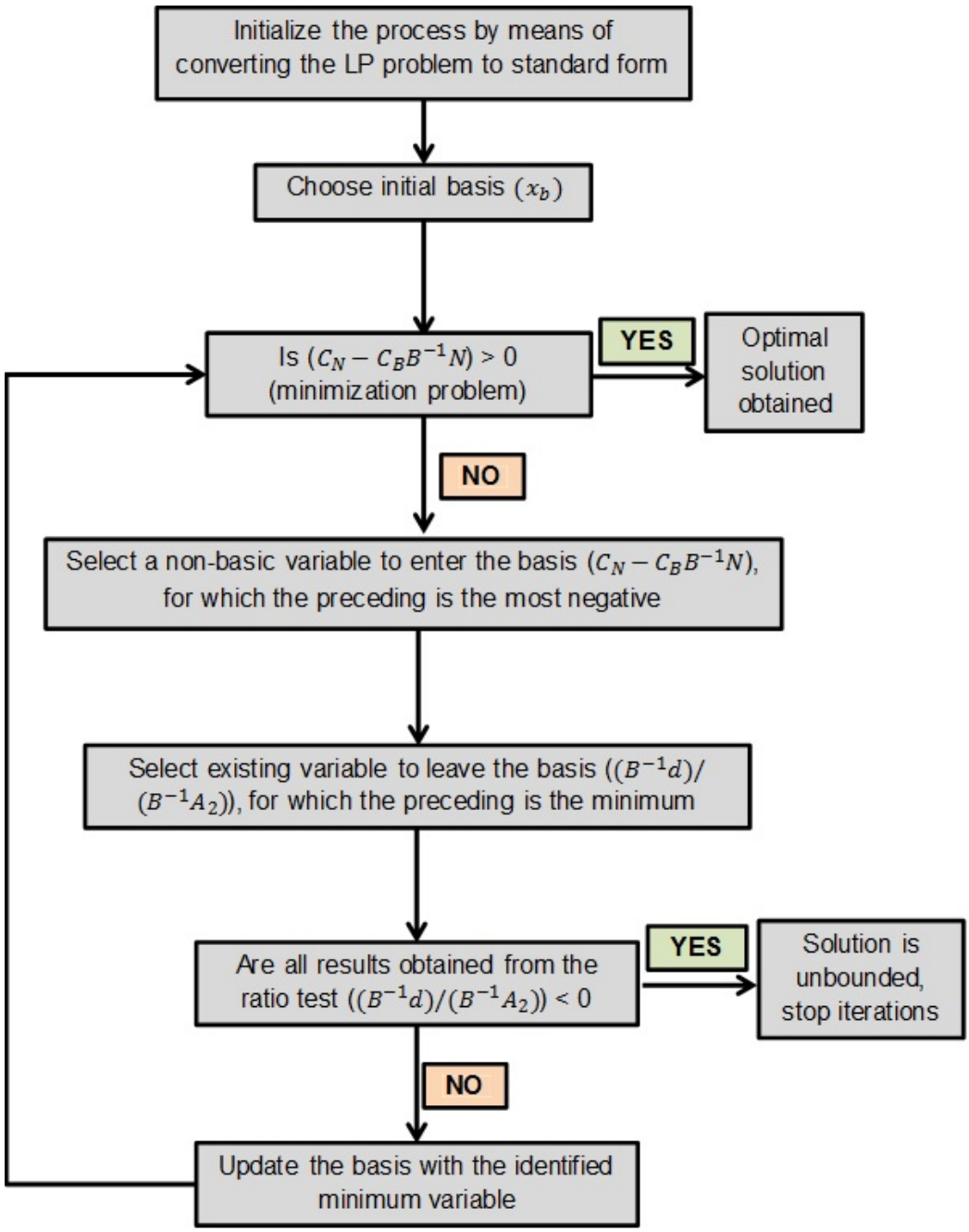

Figure 2.1: Iterative flowchart for the primal simplex algorithm applied to a minimisation LP problem

Considering the flow diagram and the primal simplex method derivation depicted above, the algorithm can be summarised in four steps which entails:

1. Covert the linear programming problem to standard form and chose an initial basis. Compute the basic feasible solution from the standard form (if not infeasible).

2. Evaluate the basic feasible solution and determine whether it is optimal. If optimal, terminate the computational process. If however the basic feasible solution is not optimal, then compute the nonbasic variable which will become a basic variable.

3. Thereafter, compute the basic variable which will become the nonbasic variable. These variables then need to be used to compute a new basic feasible solution which might provide a better solution.

4. Evaluate the new basic feasible solution to determine if it is optimal. If not optimal refer to step 2 and iterate through the above algorithm until an optimal solution has been obtained.

2.2 Integer programming methods

In the linear optimisation technique considered throughout Section 2.1.1, the decision variables were assumed to be continuous. The former refers to the capability of the variables to take on any real value. In many optimisation problems, using continuous variables are entirely appropriate and the problem is allowed to have fractional solutions. For example, if one considers engineering problems such as boiler lengths, plate thickness or even project timelines, each of the preceding may be allowed to take on fractional values such as 90.7 m, 2.30 mm and 3.25 hours respectively. However, there are practical applications where fractional solutions might not be physically meaningful or neither practically interpretable. Such solutions would entail for example 1.3 workers on a project, 1.8 boilers in a thermal plant or 1.5 tires on a motor vehicle. The fractional solutions would not bear any meaning, and the solution would need to be rounded off to the nearest integer value in order to provide clarity to the solution. Although rounding off a solution might be possible, in the majority of cases some constraints will be violated and the solution obtained from the objective function might be very far from the original solution. In order to avoid such difficulties, the optimisation problem can be formulated and solved as an integer programming problem (Rao, 2009). By constraining variables to only take on integer values, the solutions obtained by solving the optimisation problem will be practically meaningful and interpretable. Solution algorithms developed to address such problems were the branch-and-bound algorithms proposed by Land and Doig in 1960, as well as the cutting plane algorithm which was formulated by Gomory in 1958 (Genova & Guliashki, 2011). Integer programming (IP) problems can be divided into a multitude of categories of which include all-integer, binary integer and mixed integer programming problems. Note that these categories can also be subdivided into linear and nonlinear problems.

In mathematical terms, an IP problem entails the optimisation of an objective function $\boldsymbol{c^T x}$ by computing an integer vector $\boldsymbol{x} \in \mathbb{Z}^{\mathbf{n}}$ where constraints $\boldsymbol{Ax} \leq \boldsymbol{b}$ are satisfied. By dropping the integrality restrictions, the IP problem is again formulated as a LP problem which is known as the LP relaxation of the IP problem. A feasible solution to an IP problem can be represented by $\boldsymbol{x^{IP}} \in \mathbb{Z}^n$ with the objective function value being $\boldsymbol{z^{IP}} = \boldsymbol{c^T x^{IP}}$ whereas the optimal solution to a LP relaxation problem would be $\boldsymbol{x^{LP}} \in \mathbb{Z}$ with associated objective function $\boldsymbol{z^{LP}} = \boldsymbol{c^T x^{LP}}$. When considering a maximization problem, $\boldsymbol{z^{IP}}$ and $\boldsymbol{z^{LP}}$ refers to the lower bound (LB) and upper bound (UB) respectively. The opposite is true for a minimisation problem. In order to apply an integer solution method such as the branch-and-bound algorithm to obtain an optimal solution to $\boldsymbol{c^T x}$, the IP and or MIP problem needs to be written in standard form.

2.2.1 Integer programming standard form formulation

As discussed in the introduction to Section 2.2, an integer programming problem can take on various forms, for instance, all-integer, binary integer and mixed integer programming problems. In this section only a technical summary is provided. Studies done by Chen *et al.*(2010), Mallach (2015) and Galati (2010) can be sited for a more detailed explanation regarding the standard forms. For the purpose of defining the standard forms, the objective function and constraints are expressed as linear functions. As an introduction to the branch-and-bound solution algorithm, the standard forms for each of these problems are stated below in matrix form. The IP problem is stated in standard form by using vectors $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$ and a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$. The aim of the problem is to find an optimal solution by identifying a vector $\boldsymbol{x} \in \mathbb{R}^n$ which will:

Minimise the objective function

$$f(\mathbf{x}) = \mathbf{c^T x} \tag{2.7}$$

subjected to constraints:

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ \mathbf{x} &\in \mathbb{Z}_{+}^{n} \end{aligned} \tag{2.8}$$

Note that in this form, the optimisation problem may only assign discrete positive values to variables $\boldsymbol{x}$. If in the preceding IP problem, all the variables are constrained to binary values of type $\mathbb{B} = \{0, 1\}$ then the IP problem is transformed to a binary integer programming problem (BIP).

In a MIP problem, variables ($\boldsymbol{x}$) are allowed to take on either real or integer values. A MIP problem may also include variables (y) of the binary type to allow for decision making modeling. Considering the above, it is apparent that a MIP problem can take on a multitude of forms depending on the nature of the problem. Consequently, the standard formulation of the MIP problem can be portrayed as a combination of the LP, IP and BIP problems as depicted above. In the standard MIP form, vectors $\boldsymbol{c} \in \mathbb{R}^n$ and $\boldsymbol{p} \in \mathbb{R}^n$ are used to represent the objective function whereas vectors $\boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{c} \in \mathbb{R}^m$ in combination with matrices $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ and $\boldsymbol{D} \in \mathbb{R}^{n \times m}$ represent the equality constraints. The aim of the MIP problem is to find an optimal solution by identifying vectors $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{B}^m$ which will:

Minimise the objective function:

$$f(\mathbf{x}) = \mathbf{c^T x} + \mathbf{p^T y} \tag{2.9}$$

subjected to constraints:

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ \mathbf{Dy} &= \mathbf{c} \\ \mathbf{x} \geq 0, \quad \mathbf{x} &\in \mathbb{Z}^q \times \mathbb{R}^{n-q} \\ \mathbf{y} \in \mathbb{B}^m \; &where \; \mathbb{B} = \{0, 1\} \end{aligned} \tag{2.10}$$

In the case of q corresponding to a value of 0, $\boldsymbol{x}$ will be assigned any real number whereas if q equals n then $\boldsymbol{x}$ will take on only discrete values.

Depending on the optimisation problem at hand, the user needs to rewrite the problem in any one of the above standard forms in order to apply a solution method such as the branch-and-bound algorithm. When faced with nonlinear integer problems, methods such as binary integer logical formulation can be utilised to develop a linear approximation of the problem, prior to utilising MILP solution algorithms to solve the problem. By applying the logical formulation methodology as discussed, nonlinear problems can also be transformed to the above integer linear standard forms. A detailed discussion regarding nonlinear linearisation is provided throughout Section 2.2.2.

2.2.2 Logical modeling with binary variables (linearisation)

In contrast to linear programming, where the implementation of the primal simplex method is used to solve majority of the LP problems, in *nonlinear programming* (NLP) there are multiple methods which can be applied in solving NLP problems. Throughout literature it has been stated that although a multitude of methods exists, each method can only be applied to certain optimisation problems, as the effectiveness of each method was identified to be isolated to only specific problem instances. When applying NLP solution methods to problems outside of its development scope, the methods' effectiveness are expected to reduce significantly. Due to the evolving environment of NLP, continuous improvement is required in order to be able to adapt to the different problem instances (Corneujols & Tutuncu, 2006). It is also important to note that the complexity of the combined IP-NLP problems (also known as quadratic integer programming (QIP)) are much greater in comparison to single LP or IP problems. The increased complexity results in additional computational power required to obtain an optimal solution. In order to reduce the complexity of the QIP problem,

one can transform the problem into a MILP problem by utilising logical modeling with binary variables, as mentioned in Section 2.2.1. By implementing the preceding principle, it reduces the computational power and time required to obtain an optimal solution significantly. Note however that such a solution will however only be an approximated optimal solution as a result of the linearisation methodology applied. In order to linearise an NLP problem one of the following three methods can be utilised:

1. Piecewise linear approximation using continuous and binary variables
2. Piecewise linear approximation using binary variables
3. Piecewise linear approximation using function gradients and binary variables

Note however that all three methods make use of binary variables to effectively linearise the NLP problem to a MILP problem with some methods utilising the binary variables in the objective functions and others in the problem constraints. The form in which an NLP problem is generally represented prior to applying the piecewise linear approximation transformation, is depicted below:

Minimise the objective function

$$z = f(x_1, x_2, ..., x_n) \tag{2.11}$$

subjected to constraints:

$$\begin{aligned} g_m(x_1, x_2, ..., x_n) &= b_m \\ g_m(x_1, x_2, ..., x_n) &\geq 0 \end{aligned} \tag{2.12}$$

where $f(x_1, x_2, ..., x_n)$ represents a nonlinear objective function of any polynomial form. Equation (2.12) is defined as the problem constraints which are able to take on either linear or nonlinear forms. Variables $(x_1, x_2, ..., x_n)$ and b_m are subsets of real numbers $\mathbb{R}^n$. For the subsequent sections, $g_m(x_1, x_2, ..., x_n)$ are assumed as linear constraint.

1. Piecewise linear approximation using continuous and binary variables
To explain the methodology of nonlinear linearisation using non-negative continuous variables (Williams, 2009; Winston & Goldberg, 2004), the exponential objective function as depicted in Figure 2.2 is considered. Note that the objective function can take on any nonlinear form and is only considered as exponential for explanatory purposes. In Figure 2.2, the exponential function is set equal to $2x^2$. In order to linearise the mentioned function it is evaluated at different grid points. The points $(\boldsymbol{x_j})$ used for analysis is not necessarily required to be evenly spaced, but is however required to span the entire range of the area under investigation. In Figure 2.2, the range of interest is represented between $\boldsymbol{x_j} = 0, ..., 3$ with the grid being divided into equal spaced intervals of 1. By dividing the grid in three equally sized functions, the user is able to obtain values of $f_j(\boldsymbol{x_j}) = (f_1(x_1), ..., f_3(x_3))$ corresponding to the integers $\boldsymbol{x_j} = (0, ..., 3)$. Refer to Figure 2.2 for a visual representation of the above stated.

By joining the evenly distributed points $f_j(\boldsymbol{x_j})$ into one equation, it provides the user with a piecewise linear approximation of the nonlinear objective function $2x^2$. In order to combine the points, non-negative continuous variables $(\boldsymbol{\lambda_j})$ needs to be assigned to each given point. These weights are multiplied with each point $f_j(\boldsymbol{x_j})$ and summated in terms of index j to obtain the linear approximated objective function of the original nonlinear function. An example of the linearised objective function using non-negative continuous variables is depicted in equation (2.13).

$$z = \sum_{j=1}^{j=n} [\lambda_1 f_1(x_1) + \lambda_2 f_2(x_2) + ... + \lambda_j f_j(x_j)] \tag{2.13}$$

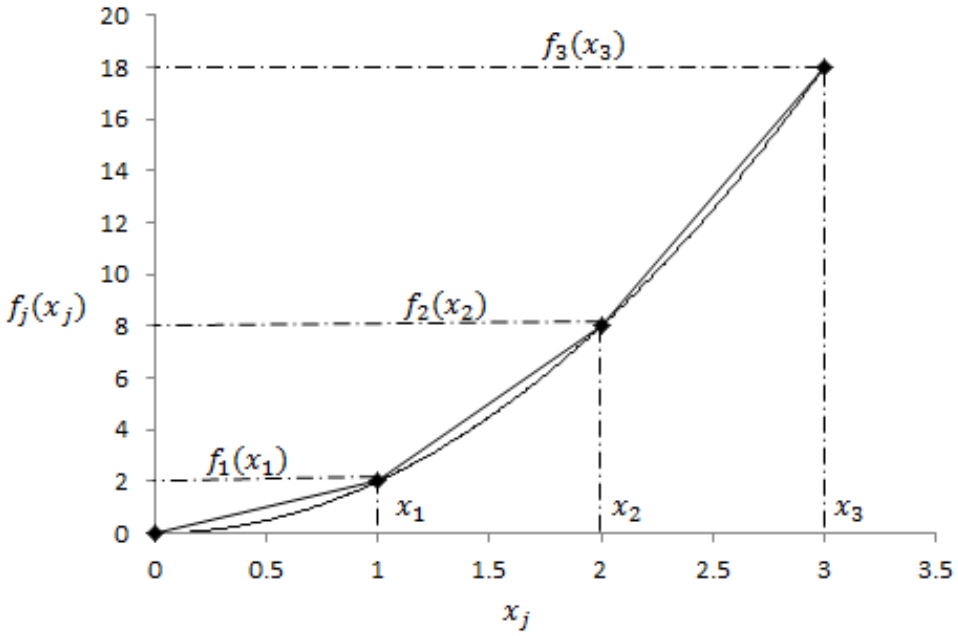

Figure 2.2: Linearisation on nonlinear objective functions using continuous and binary variables

In order to adapt the inequality constraints with the non-negative continuous variables as stipulated above, each one of the variables $\boldsymbol{x_j}$ also needs to be multiplied by λ_j and summated in terms of index j. The summated term is equivalent to $\boldsymbol{x_j} = \sum_{j=1}^{j=n}[\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_j x_j]$ which is substituted into function g_m to obtain the linear function value associated to variable $\boldsymbol{x_j}$. The preceding is then set equal to or less than variable $\boldsymbol{b_i}$ in order to constrain the summation function to an upper limit as noted in equation (2.14):

$$g_m(\sum_{j=1}^{j=n}[\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_j x_j]) \leq b_i \quad (i = 1, 2, ..., m) \tag{2.14}$$

In addition, to ensure that the summation of the non-negative continuous variables equal a value of *1*, equation (2.15) needs to be added to the model. By incorporating this constraint, variables $\boldsymbol{\lambda_j}$ are only allowed to be assigned fractional values less than or equal to one. The reason for constraining variables $\boldsymbol{\lambda_j}$ between *0* and *1* is because it needs to represent the percentage weighting assigned to each variable $f_j(\boldsymbol{x_j})$ and $\boldsymbol{x_j}$. Allowing a fractional value greater than 1 to be assigned to $\boldsymbol{\lambda_j}$ will not satisfy the definition of percentage values and would lead to erroneous results being obtained.

$$\lambda_1 + \lambda_2 + ... + \lambda_j = 1 \quad (j = 1, 2, ..., n) \tag{2.15}$$

Using this linearisation methodology, the user is required to add constraints to the model which will impose the condition that at most only two adjacent $\boldsymbol{\lambda_j}$ variables can take on non-zero values. If for example adjacent variables $\boldsymbol{\lambda_j}$ and $\boldsymbol{\lambda_{j+1}}$ are assigned non-zero values, these variables will represent a point on the linear approximation line which corresponds to the j^{th} and $j + 1^{th}$ grid points. Variables $\boldsymbol{\delta_j}$ are defined as binary variables which is incorporated to assist with the logical decision making process of deciding which two non-negative continuous variables to select (assign values of 1) and which to ignore (assign values of 0). When index j is at either a value of 1 or n, different constraints will apply as these index points are at the lower and upper points of the nonlinear curve. In order to implement the above methodology, the constraints stacked in equation (2.16) needs to be incorporated into the model. By doing this, the user effectively assigns the model with the ability to apply the theory of interpolation.

$$\begin{aligned} \boldsymbol{\delta_j} &\leq \boldsymbol{\lambda_j} \quad (j = 1) \\ \boldsymbol{\delta_j} &\leq (\boldsymbol{\lambda_j} + \boldsymbol{\lambda_{j-1}}) \quad (j = 2, 3, ..., n-1) \\ \boldsymbol{\delta_j} &\leq \boldsymbol{\lambda_j} \quad (j = n) \end{aligned} \tag{2.16}$$

To ensure that the above constraints hold true, equation (2.17) needs to be added to the model in order to allow only one $\boldsymbol{\delta_j}$ variable to be assigned a value of *1*. By only assigning one $\boldsymbol{\delta_j}$ variable with a non-zero value, it will activate only one of the constraints contained within equation (2.16) to ensure that the logical decision making process is done correctly. Without incorporating the mentioned constant will allow the model to select multiple grid points which in turn will result in unrealistic answers being obtained.

$$\delta_1 + \delta_2 + ... + \delta_j = 1 \quad (j = 1, 2, ..., n) \tag{2.17}$$

The constraints contained within equation (2.18) is added to the piecewise linearisation model to define variables $\boldsymbol{\lambda_j}$ as non-negative variables and variables $\boldsymbol{\delta_j}$ as binary variables.

$$\begin{gathered} \boldsymbol{\lambda_j} \geq 0,\ (j = 1, 2, ..., n),\ \lambda \in \mathbb{R}^n \\ \boldsymbol{\delta_j} = \{0; 1\},\ \delta \in \mathbb{B}^n \end{gathered} \tag{2.18}$$

Also note that the linear approximation method will either over or underestimate the nonlinear function's true value when implementing the above algorithm, but the offset can be reduced by refining the grid under consideration. By refining the grid, the computational time required to obtain the optimal solution will most definitely be increased.

2. Piecewise linear approximation using binary variables

Another methodology which can be applied to linearise a nonlinear objective function is known as piecewise linear approximation using binary variables (as defined by the author). For explanatory purposes, the same exponential objective function is utilised as was discussed in Section 2.2.2 part 1 to explain the reasoning behind the mentioned method. Contrary to what was discussed in Section 2.2.2 part 1, this methodology only utilises binary variables for its linearisation process and does not incorporate non-negative continuous variables. In the previous method, the functionality of interpolation was built into the model by using both the non-negative continuous variables and the binary variables. However in this method, the interpolation functionality is not available and the model will only be able to select from a number of data points using the binary variables. The number of data points will be dependent on the grid size division. Due to the linearisation methodology used in this method, the optimisation algorithm will only have a limited amount of data points to chose from. When compared to the previous method, this method will have far less options to chose from in order to obtain an optimal solution. For example, in Figure 2.3 the grid is divided into seven data points with the x-axis being represented by $\boldsymbol{x_j}$ and its associated y-axis by $f_j(\boldsymbol{x_j})$. When using the binary variables, the model will only be able to choose from one of the seven data points. Refer to Figure 2.3 for a visual representation.

Although the objective function in the initial method (Figure 2.2) is only divided into four data points (three linear equations), the methodology applied enables the model to interpolate between the different points resulting in a significant increase in options to chose from. Note however that the method discussed in Section 2.2.2 part 1 would maybe provide the user with a more optimal solution due to its ability to automatically generate multiple data points, but this method will most definitely require increased computational power and calculation time to obtain the desired solution. When faced with large scale problems, the piecewise linear approximation methodology using binary variables might be a better method to implement as the algorithm will be able to provide an approximated optimal solution within a reasonable time frame, contrary to the method discussed in Section 2.2.2 part 1.

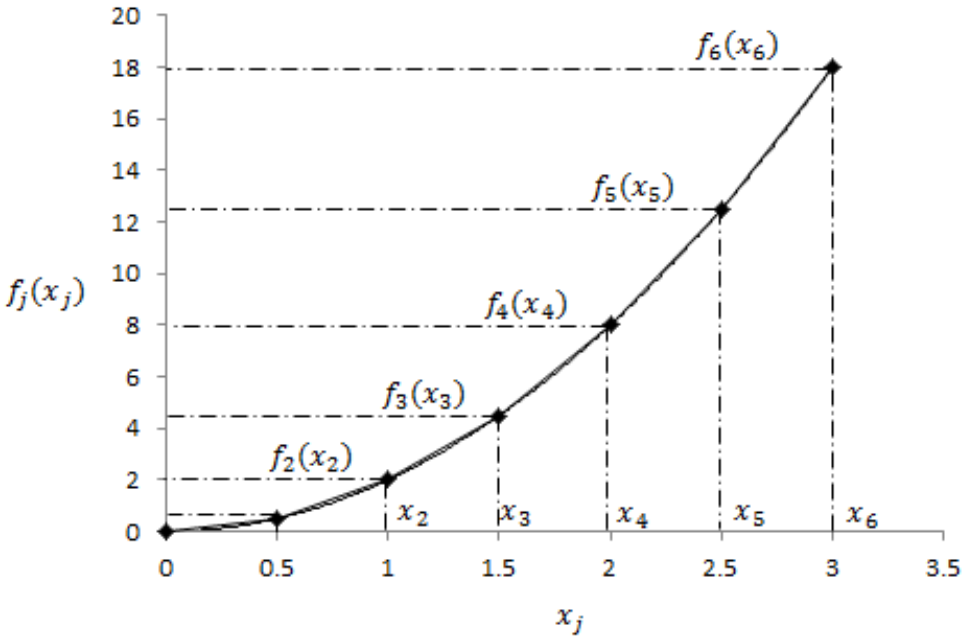

Figure 2.3: Linearisation on nonlinear objective functions using binary variables

To develop a linearised approximation of the nonlinear objective function $2x^2$, using binary variable formulation, a binary decision variable needs to be assigned to each grid point $f_j(\boldsymbol{x_j})$. By multiplying each grid data point with its associated binary variable ($\boldsymbol{\delta_j}$) and summating the multiplication term of each data point in terms of index j, a linear approximation of the nonlinear objective function can be obtained. An example of the linearised objective function using binary variables formulation is depicted in (2.19) below:

$$z = \sum_{j=1}^{j=n} [\delta_1 f_1(x_1) + \delta_1 f_2(x_2) + ... + \delta_j f_j(x_j)] \tag{2.19}$$

Similar to Section 2.2.2 part 1, the inequality constraints need to be adapted in terms of binary variables $\boldsymbol{\delta_j}$ by means of multiplying variables $\boldsymbol{x_j}$ with binary variables $\boldsymbol{\delta_j}$. The preceding is done to relate the constraints to the linearised objective function. The foregoing multiplication functions are summated in terms of index j and substituted into function g_m. The summated term, prior to its substitution, is equivalent to $\boldsymbol{x_j} = \sum_{j=1}^{j=n} [\delta_1 x_1 + \delta_2 x_2 + ... + \delta_j x_j]$. Function g_m is again set equal to or less than variable $\boldsymbol{b_i}$ in order to constrain the inequality constraint to an upper limit as denoted in equation (2.20). Note however that by implimenting the binary variables in equation (2.20), the model will only be allowed to select one variable $\boldsymbol{x_j}$ which corresponds to the data point $f_j(\boldsymbol{x_j})$ as selected in equation (2.19). The remaining variables will be assigned 0 values and will be discounted from the calculation process.

$$g_m(\sum_{j=1}^{j=n} [\delta_1 x_1 + \delta_2 x_2 + ... + \delta_j x_j]) \leq b_i \quad (i = 1, 2, ..., m) \tag{2.20}$$

For the above statement to hold true, equation (2.21) needs to be incorporated into the optimisation model. By summating variables $\boldsymbol{\delta_j}$ in terms of index j and setting the function equal to a value of *1*, it allows only one binary variable to be assigned a non-zero value. If the subsequent constraint is not added to the model, the model will be allowed to assign multiple binary variables non-zero values which in turn will result in incorrect results being obtained from the model.

$$\delta_1 + \delta_2 + ... + \delta_j = 1 \quad (j = 1, 2, ..., n) \tag{2.21}$$

Equation (2.22) needs to be added to the piecewise linearisation model in order to declare $\boldsymbol{\delta_j}$ as binary variables.

$$\boldsymbol{\delta_j} = \{0; 1\}, \ \delta \in \mathbb{B}^n \tag{2.22}$$

The model accuracy and computational time required to obtain an optimal solution will be determined by the amount of data points within which the grid is discretised when using the binary variable formulation method. When the grid is discretised into more data points, the model accuracy and solution time will increase and vice versa. It is therefore the prerogative of the user to determine the importance of the model accuracy and to determine if an approximated optimal solution will suffice. Depending on the conclusion, grid discretisation intervals can be determined.

3. Piecewise linear approximation using function gradients and binary variables

The third method that can be applied to linearise a nonlinear objective function is known as piecewise linear approximation using function gradients (Bradley *et al.,* 1977). Similar to the method discussed in Section 2.2.2 part 1, this method also utilises both non-negative continuous variables and binary variables to linearise the nonlinear function. However the difference between the two methods is that in this method the nonlinear objective function is divided into three evenly spaced linear functions ($y = mx + c$) from which the gradient (m) of each linear function is used to develop the piecewise linear objective function, whereas in Section 2.2.2 part 1 the nonlinear function was divided into multiple data points and the theory of interpolation was applied to obtain the mentioned function. Refer to Figure 2.4 for a graphical representation of the function gradient method. When using the theory of function gradients, the nonlinear function does not necessarily need to be divided into evenly spaced sections, but for explanatory purposes the x-axis is divided into even lengths in Figure 2.4. Also, take note that the nonlinear function intersects the y-axis at *0* in Figure 2.4. In mathematical terms, variable c used in the linear function formulation will be equal to zero and won't be included in the objective function. If however the nonlinear function intersects the y-axis at any other non-zero point, variable c will be equal to the value $f_j(\boldsymbol{x_j})$ of the intersection and will have to be incorporated into the objective function.

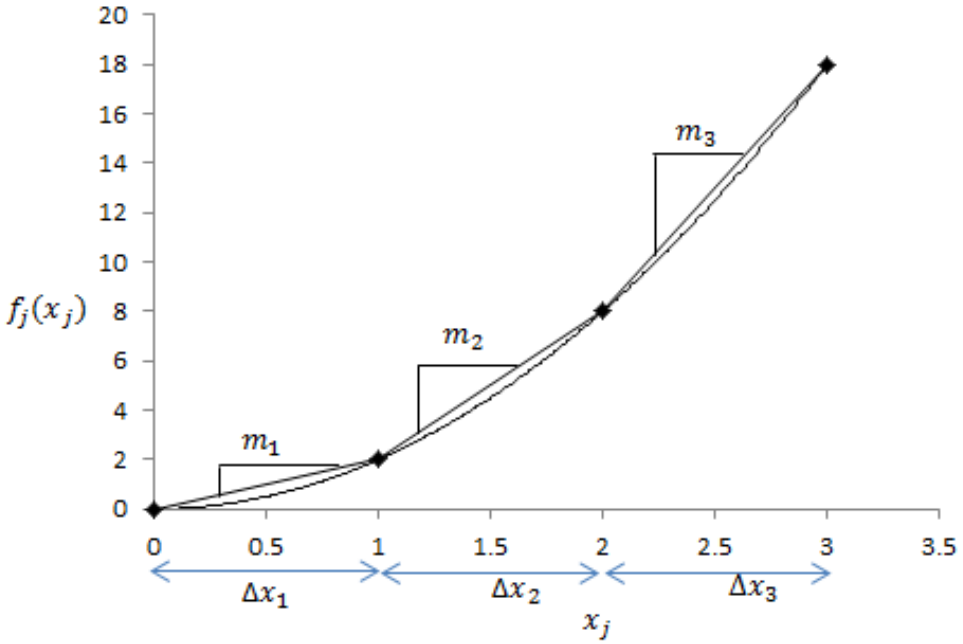

Figure 2.4: Linearisation on nonlinear objective functions using function gradients and binary variables

Given the aforementioned, the nonlinear objective function can be transformed to a linear approximation by means of multiplying each function gradient $\boldsymbol{m_j}$ with non-negative continuous variables $\boldsymbol{\lambda_j}$ and summating the multiplication functions in terms of index j. If however there exist a y intersect $f_j(\boldsymbol{x_j})$ of a non-zero value, the intersect value also needs to be added to the objective function in equation (2.23). Given the function in Figure 2.4, no y intersect exist that results in this value being omitted from the objective function.

$$z = \sum_{j=1}^{j=n} [m_1\lambda_1 + m_2\lambda_2 + ... + m_j\lambda_j] \tag{2.23}$$

Due to the mathematical formulation used in this method, $\boldsymbol{\lambda_j}$ effectively represents variable x in the general linear function $y = mx + c$. Therefore, in order to rewrite the x-axis inequality constraints of the stipulated problem in terms of the function gradients method, variables $\boldsymbol{\lambda_j}$ can be summated in terms of index j, substituted into function g_m and set equal to or less than variable $\boldsymbol{b_i}$. The summated term, prior to its substitution, is equivalent to $\boldsymbol{x_j} = \sum_{j=1}^{j=n}[\lambda_1 + \lambda_2 + ... + \lambda_j]$. Refer to equation (2.24) for the mathematical representation of the mentioned formulation.

$$g_m(\sum_{j=1}^{j=n}[\lambda_1 + \lambda_2 + ... + \lambda_j]) \leq b_i \quad (i = 1, 2, ..., m) \tag{2.24}$$

In order to enforce the model to be able to perform correct logical decisions, constraints needs to be incorporated to ensure that whenever $\lambda_2 > 0$ then $\lambda_1 = \Delta x_1$ and when $\lambda_3 > 0$ then $\lambda_1 = \Delta x_1$ and $\lambda_2 = \Delta x_2$, and so on. Without incorporating such constraints, the model will be able to minimise the objective function by assigning λ_2 a non-zero value, if m_2 has the smallest gradient, while maintaining variables λ_1 and λ_3 equal to 0. Although this might provide the smallest variable cost, the results would be incorrectly calculated. In order to enforce the preceding conditional constraints, binary variables ($\boldsymbol{\delta_j}$) needs to be incorporated in to the model. The binary variables are added to assist with the logical decision making process of deciding which conditional constraints to activate or discount from the model as to obtain an optimal feasibly solution. When index j is at either a value of 1 or n, different constraints will apply as these index points are at the lower and upper points of the nonlinear curve. In order to implement the above methodology, the constraints stacked in equation (2.25) are used in the model. When for example variable $\boldsymbol{\lambda_j}$ is set equal to a non-zero value by the model where index $j = n$, all binary variables $\boldsymbol{\delta_j}$ of index $j = 1...n$ will be assigned values of 1 to ensure the conditional constraints are satisfied. If however only λ_2 is set equal to a non-zero value, δ_1 will be the only binary variable assigned a value of 1 with the others binary variables remaining at 0 values.

$$\begin{gathered} \Delta x_j \boldsymbol{\delta_j} \leq \boldsymbol{\lambda_j} \leq \Delta x_j \quad (j = 1) \\ \Delta x_j \boldsymbol{\delta_j} \leq \boldsymbol{\lambda_j} \leq \Delta x_j \boldsymbol{\delta_{j-1}} \quad (j = 2, 3, ..., n-1) \\ 0 \leq \boldsymbol{\lambda_j} \leq \Delta x_j \boldsymbol{\delta_{j-1}} \quad (j = n) \end{gathered} \tag{2.25}$$

The constraints in (2.26) are added to the piecewise linearisation model to define variables $\boldsymbol{\lambda_j}$ as non-negative variables and variables $\boldsymbol{\delta_j}$ as binary variables.

$$\begin{gathered} \boldsymbol{\lambda_j} \geq 0, \ (j = 1, 2, ..., n), \ \lambda \in \mathbb{R}^n \\ \boldsymbol{\delta_j} = \{0; 1\}, \ \delta \in \mathbb{B}^n \end{gathered} \tag{2.26}$$

In summary, depending on the size and complexity of the nonlinear optimisation model and the availability of the computational time, either one of the piecewise linearisation models can be implemented to simplify a nonlinear problem to a integer-linear form. The linearisation is executed in order to apply solution algorithms such as the branch-and-bound method to obtain optimal solutions. Take into consideration some piecewise linearisation models will require additional solving time in comparison to others as a result of its mathematical formulation and the user will need to determine the amount of time available for solution prior to the implementation of the different methods.

2.2.3 Branch-and-bound method for mixed integer linear problems

Literature compiled by Corneujols & Tutuneu (2007) and Winston & Goldberg (2004) were investigated in order to obtain an understanding of the branch and bound method. For more detail, the

foregoing references can be cited. The branch-and-bound method is based on the principle of partitioning a set of feasible solutions into smaller subsets. These subsets are then individually evaluated in order to obtain an optimal solution. To explain the detailed methodology behind the branch-and-bound algorithm, a mixed integer linear programming problem *(MILP)* of the minimisation type is considered. The initial step in the branch-and-bound method is to solve the LP relaxation problem LP_0 by means of dropping all integer constraints from the *MILP*. Generally, LP_0 is solved by using the simplex method with the solution to LP_0 being defined as $\boldsymbol{x^{LP}}$. When solving LP_0 by means of applying an LP solution algorithm, and the solution ($\boldsymbol{x^{LP}}$) obtained is integer, then the *MILP* problem is solved and the branch-and-bound algorithm terminates. If, however, the solution to LP_0 is infeasible, the *MILP* will also be infeasible and will again result in the termination of the branch-and-bound algorithm. Only if there exists a component within $\boldsymbol{x^{LP}}$ which does not contain an integer value, the branch-and-bound method initiates the second step in the algorithm. Note however if a variable is allowed to have fractional values, the branch-and-bound algorithm will not be applied to this variable. The algorithm will only be applied to variables which are constrained to be integer. In the second step, the lower bound *(LB)* of *MILP* is set equal to the objective function value obtained for LP_0, and the upper bound *(UB)* takes on the value of infinity. After defining *LB* and *UB*, the feasible regions as depicted by LP_0 needs to be partitioned into two separate sub-problems. In order to partition LP_0, a variable $\boldsymbol{x_r^{LP}}$ contained within $\boldsymbol{x^{LP}}$ for which the integrality constraint is violated, is identified and divided into two sub-problems, say, LP_1^r and LP_2^r . This is done by adding constraint:

$$\boldsymbol{x_r} \leq [\boldsymbol{x_r^{LP}}]$$

to the parent problem (LP_0) to obtain LP_1^r, as well as constraint:

$$\boldsymbol{x_r} \geq [\boldsymbol{x_r^{LP}}]$$

to obtain LP_2^r. As mentioned above, LP_0 is defined as a parent node whereas sub-problems LP_1^r and LP_2^r can each be denoted as child nodes. The preceding process is known as branching. After creating child nodes, the LP relaxation problem of each is solved and the solutions obtained are checked for integer feasibility. If the integrality constraints are violated and the objective function value obtained is less than the upper bound Z, the node is added to the waiting list. Such a node is know as a dangling node. If however all the integrality constraints are satisfied, the child node becomes the new integer solution. If the objective value (Z_p) obtained from the solution is less than the current upper bound (Z), Z_p will become the new upper bound so that $Z = Z_p$. Else, if the objective value is higher than the current upper bound, the node is removed from the waiting list. For explanatory purposes, refer to Figure 2.5. Given the example portrayed in Figure 2.5, we note that branching was done on parent node (LP_0) as discussed above, where after the LP relaxation problem of both child nodes LP_1^r and LP_2^r were solved using an LP solution algorithm. In order to elaborate on the remaining steps in the branch-and-bound algorithm, the results obtained from node LP_1^r is assumed to be non integer values and is placed in the waiting list whereas node LP_2^r is discarded from the waiting list as the node provided non-optimal results.

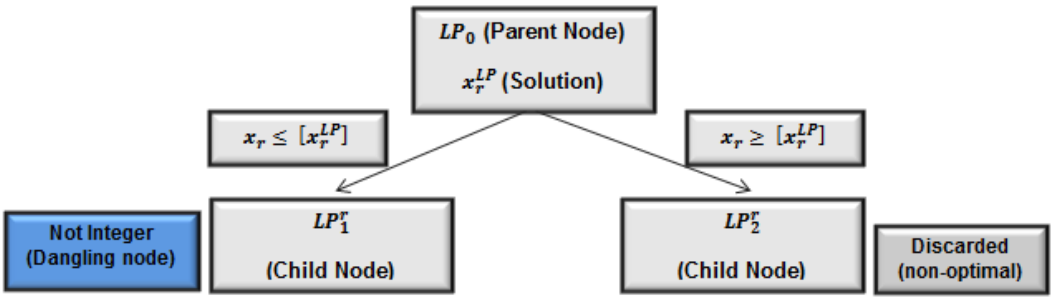

Figure 2.5: Branching process used in the branch-and-bound algorithm

The next step is to select the next sub-problem from the dangling nodes in the waiting list on which branching needs to be performed. For the above example, branching is done on node LP_1^r. This process is termed node selection. During node selection, the node for which the sub-problem has the minimum objective function value needs to be selected for branching. Again the LP relaxation problem of each branched node is solved to obtain a solution. The mentioned evaluation criteria are then utilised to determine if the solution will become the next incumbent integer solution, if further branching is required or if the node needs to be eliminated from the computational process. If the waiting list is empty, the branch-and-bound algorithm will terminate as no node will be available for selection. In such a case, the optimal MILP solution will then be the last identified incumbent solution. It might also be that no incumbent solution was obtained during the computational process which would indicate that the MILP problem has no feasible solution. Refer to Figure 2.6 for a representation regarding the node selection process.

After selecting node LP_1^r, branching is performed to sub-divide LP_1^r into child nodes LP_3^r and LP_4^r. Applying the algorithmic methodology discussed, it is identified that child node LP_3^r satisfies all integrality constraints and that the results obtained from this node becomes the incumbent solution. The same process is applied to LP_4^r but the node is identified to be infeasible and is consequently discarded from the calculation process.

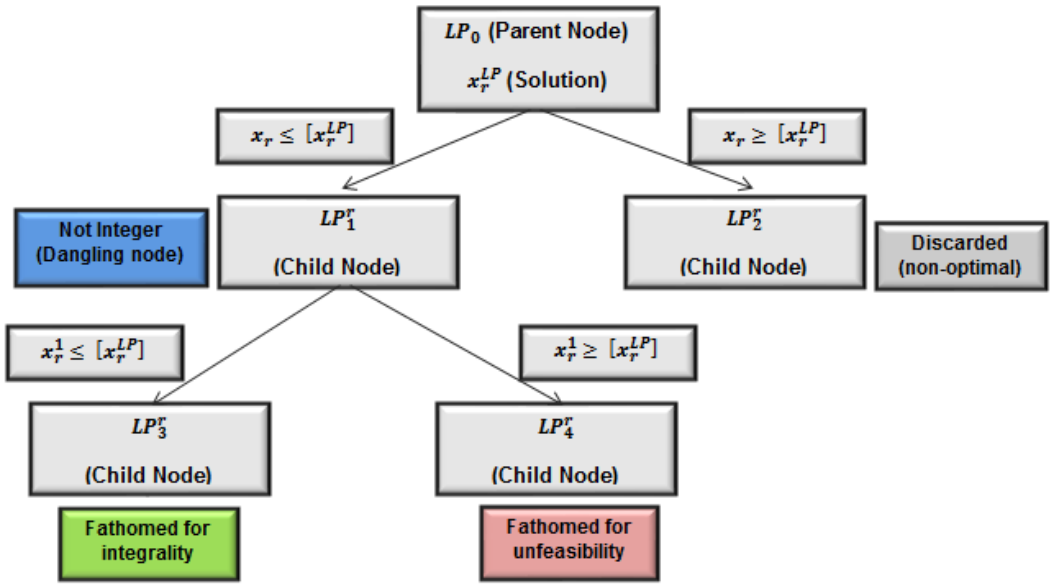

Figure 2.6: Node selection process used in the branch-and-bound algorithm

After traversing the waiting list for dangling nodes, it is determined that all the nodes were processed and that the list is empty. Therefore, the incumbent solution LP_3^r becomes the optimal solution to the minimisation problem where after the branch-and-bound algorithm will terminate. Refer to Figure 2.7, for a detailed flow diagram of the complete branch-and-bound algorithm applied to a minimisation problem as discussed.

In summary, the principle that underpins this algorithm is to subdivide a complex problem into sub-problems of progressively decreasing dimensions in order to solve the problem easier (applying the principle of divide and conquer). Also note that when confronted with either IP or MIP problems, this algorithm can be applied to obtain a mathematically proven optimal solution. When discussing the unit commitment and load dispatch optimisation problem throughout Chapters 4 to 5, this solution method must be kept in mind as the commercial software used to solve the mentioned problem implements this method.

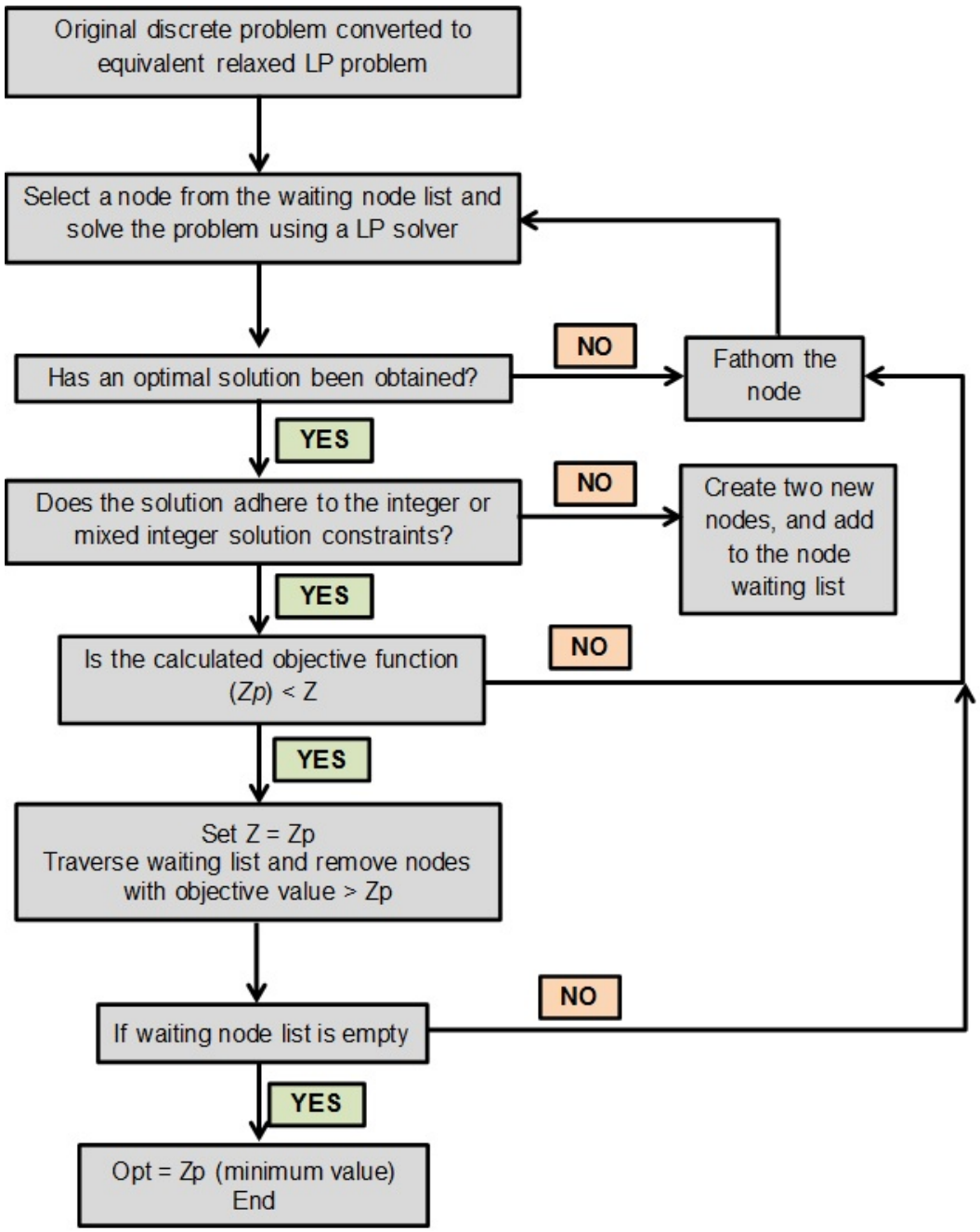

Figure 2.7: Iterative flowchart for the branch-and-bound algorithm applied to IP or MIP problems

Heuristic solution methods

Heuristic solution methods or algorithms are designed to solve complex optimisation problems faster and more efficient than traditional exact solution methods. However, in order to increase the computational speed of these algorithms, a compromise is made to sacrifice precision, accuracy and solution optimality. Generally, heuristic algorithms are employed when only an approximated solution is required and where exact solutions would be computationally too expensive. These algorithms are usually problem specific and when applied to different problem instances, they typically do not provide satisfactory results. Metaheuristic algorithms, are similar to general heuristic algorithms, but they are designed to be compatible with a multitude of problem instances and provide adequate results when applied to various scenarios. However, at times these algorithms are only able to generate local optimum values (maxima or minima) which deviate from the global optimal solutions. General heuristic methods include, but are not limited to:

1. Particle swarm optimisation
2. Tabu search
3. Hopfield neural network
4. Simulated annealing
5. Ant colony optimisation

Throughout the subsequent sections, a high level overview of the principles from which each method originate, is provided.

2.3 Particle swarm optimisation

The particle swarm optimisation algorithm (PSO) was developed by Kennedy and Eberhart in 1995. This method was based on the concept of mimicking the behavior of individual agents or particles within a group or swarm. The PSO algorithm provides an optimisation tool which is population-based where the agents or particles adjust their position within a certain time period. In order to obtain an optimal solution using the PSO algorithm, the particles are allowed to fly around a multidimensional search space during which the position of each particle is adjusted according to the particles own and its neighboring particle's experience (Li *et al.,* 2008). In mathematical terms, we define a d-dimensional search space, as well as the velocity and position of the i^{th} particle in the swarm which is represented by vectors $V_i = (v_{t1}, v_{t2}, ..., v_{td})$ and $x_i = (x_{t1}, x_{t2}, ..., x_{td})$ respectively. Vector $pbest_i = (pbest_{t1}, pbest_{t2}, ..., pbest_{td})$ is defined in order to record the previous best position of the i^{th} particle. Index $gbest_{td}$ is implemented in the algorithm to log the position of the best particle among all of the other particles in the swarm.

To calculate the modified position and velocity of a particle from the current velocity and the distance from $pbest_{td}$ to $gbest_{td}$, the following equation can be utilised:

$$v_{td}^{k+1} = w \times v_{td}^{k} + C_1 \times r_1 \times (pbest_{td} - x_{td}^{k}) + C_2 \times r_2 \times (gbest_{td} - x_{td}^{k}) \tag{2.27}$$

With $d = 1, 2, ..., N_g$ and $i = 1, 2, ..., N_p$. Note that N_g is the number of elements in a particle whereas N_d denotes the number of particles in a swarm. Variable r_i is a uniform random number within the range $[0; 1]$, variables C_1 and C_2 represents acceleration constants, w is the swarm inertia or also know as weight parameter and lastly, v_{td}^{k} is the velocity of particle i at an iteration k. Constants C_1 and C_2 are implemented in equation (2.27) to pull each particle to either positions $pbest_{td}$ or $gbest_{td}$. In order to change the position of each particle in a swarm, the updated velocity can be utilised as depicted in equation (2.28).

$$x_{td}^{k+1} = x_{td}^{k} + v_{td}^{k+1} \tag{2.28}$$

To reduce the iterations required to obtain an optimal solution, when using the PSO algorithm, the correct inertia weight (w) needs to be selected in order to obtain a balance between the local and global explorations. In order to calculate the inertia weight, the following equation can be applied:

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}} \times iter \tag{2.29}$$

With $iter_{max}$ and $iter$ defined as the maximum and current number of iterations, respectively. The preceding is a high level overview of the classical PSO algorithm. Note however that this algorithm has also been applied in combination with other methods in literature to improve the performance of the PSO algorithm (Mandal *et al.,* 2015).

2.4 Tabu search

The tabu search method was developed by F. Glover during 1989 and it is applied as a metaheursitic local search method. The method was designed to have a flexible memory in order to save and employ information regarding the historical searches to explore the search space for an optimal solution (Naama *et al.,* 2013). The computational principle of the tabu search method is based on the hill-climbing algorithm. However, the difference is that if a current solution is identified, the algorithm is allowed to move out of the mentioned area although it might negatively affect the objective function at first, in the hope that it will eventually lead to a better solution. The tabu search algorithm can be divided into two main components which consist of the tabu list (TL) and the aspiration criterion (AC), with the *TL* being divided into the following three subsections

1. Forbidding: In this phase the search algorithm checks the moves which was executed by the search algorithm and prevent the algorithm from executing moves which were already explored during previous executions and were deemed undesirable.

2. Freeing: The preceding process is focused on which solutions satisfy the aspiration criteria and is allowed to exit the tabu list. In this phase, the algorithm also tracks in what instance this freeing action takes place.

3. Short term memory: This strategy is included into the algorithm in order to update the tabu list, based on the memory structure formulated during the forbidding and freeing phases.

The tabu list is employed to store a list of all the moves that are not permitted to be applied to a current solution and to record the frequency, recency and move direction of each decision. The purpose of the *AC* is to identify when a move is allowed to be free after satisfying a specific move criterion. The procedure of the Tabu search algorithm can be summarised by the following steps:

1. Initialize the algorithm by selecting an initial solution i from S and set variables $K=0$ and $i^* = i$

2. While the stopping criteria is not yet met, set $K = K + 1$ and generate a subset say (V^*) of solution $N(i,K)$ which will either hold one aspiration condition or violate one tabu condition.

3. Chose the best solution say j in subset V^* and replace the initial solution with the mentioned solution $i = j$.

4. If it is found that $f(i) \leq f(i^*)$ then the algorithm will set $i^* = i$

5. Thereafter the aspiration criteria and tabu list are updated and the algorithm is stopped if the stop criteria are met. Otherwise, if the criteria are not met revert back to step 2.

The tabu search algorithm will stop if there is no feasible solution, if the maximum allowable number of iterations has been exceeded or if an optimum solution has been obtained (Joshi, 2018).

2.5 Hopfield neural network

During the early 1980s, John Hopfield developed the neural network method know as the Hopfield model. The model consist of multiple neurons with all neurons being both input and output neurons. Each neuron is connected in both directions to the neighboring neurons (Benyahia *et al.,* 2008). In this method, the input is applied simultaneously to all neurons in order for the mathematical computation to take place. The output obtained from each neuron is then again supplied to all the neurons as inputs where the computational process is re-initiated. This process continues until a stable state is obtained, where after the model terminates to provide the network output (Dash, 2013). The Hopfield neural network (HNN) is generally applied in solving combinatorial optimisation problems. The HNN system consist of different elements of which include parallel input and output channels as well as a significant amount of interconnectivity between the neural network elements.

In practice, either the discrete or continuous Hopfield neural network is utilised. In the discrete model, each processing element or neuron can take on either a value of 0 or 1, which is represented by V_i^0 and V_i^1 respectively. The inputs (U) to the discrete model can either be obtained from an external source (I) or from the neighboring neurons (V_j). Equation (2.30) is indicative of the inputs to neuron i, with variable T_{ij} presenting the interconnective conductance between the input (i) and output (j) neurons:

$$U_i = \sum_{j=1}^{N} T_{ij} V_j + I_i \tag{2.30}$$

The input to each neuron is sampled randomly by means of the following two rules.

$$\begin{aligned} V_i &= 0, \text{if } U_i \leq \theta_i \\ V_i &= 1, \text{if } U_i \geq \theta_i \end{aligned} \tag{2.31}$$

The continuous Hopfield model is based on continuous variables. The output values obtained from neurons i is within a range $V_i^0 \leq V_i \leq V_i^1$ with the input-output function being a continuous increasing function of input U_i to neuron i. Each neuron's dynamic characteristics can be represented by the subsequent differential equation, with variable T_{ij} being the self-connection conductance of the i^{th} neuron.

$$\frac{dU_i}{dt} = \sum_{j=1}^{N} T_{ij} V_j + I_i \tag{2.32}$$

Each neuron's output in the continuous model is given by (2.33), with variable $f_i(U_i)$ being the input-output function:

$$V_i = f_i(U_i) \tag{2.33}$$

The energy function used in the HNN model is defined in (2.34). Note that when taking the time derivative of the energy function, a negative function is obtained. It is for this reason that in the computational process of the HNN model, the algorithm will always result in the energy function being gradually reduced as to converge the function to a minimum value.

$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij} V_i V_j - \sum_{i=1}^{N} I_i V_i \tag{2.34}$$

Depending on the type of problem that needs to be solved, either the discrete or continuous model can be applied to solve the optimisation problem. The results obtained from the mentioned algorithms are however not proven optimal. Exact methods will need to be implemented to obtain proven optimal solutions (Benhamida *et al.,* 2018).

2.6 Simulated annealing

Another heuristic optimisation algorithm is the simulated annealing (SA) method. This algorithm was proposed by S. Kirkpatrick *et al.* in 1983. The theory on which the SA algorithm is based, is the annealing process of a molten metal (Amhamad & Shrivastava, 2016). In this process, a hot metal at a high temperature is cooled down at a slow incremental rate as to ensure that thermal equilibrium is reached at each stage of the cooling process. The annealing process is terminated when the energy of the system reaches a global minimum value. This principle is applied to complex nonlinear combinatorial minimisation problems in order to obtain the global minimum. Note that this algorithm will not always be able to provide a proven global minimum due to its heuristic structure. Exact methods will need to be employed to determine the optimality of the solution obtained from the simulated annealing method. In the SA algorithm, a temperature parameter T

is defined which is slowly reduced during the optimisation process. At each point T, an iterative procedure is executed. In the iterative process, a Gaussian probabilistic distribution function is used to obtain a trail solution from a current solution. If the cost associated with the trail solution is less than that identified for the current solution, then the trial solution is accepted by the algorithm and implemented to generate other trial solutions. If however the trail solution is greater than the current solution, it is only accepted when the transition probability of acceptance ($P(T)$), as determined by the Boltzmann distribution, is greater than a random number ranging between $0-1$. Variable $P(T)$ is defined as:

$$P(T) = e^{\frac{-\Delta F}{T_t}} \tag{2.35}$$

In the SA algorithm, each trail solution which is identified at a temperature T, is tested for an appropriate time period in order for the algorithm to reach thermal equilibrium. If not enough time is assigned for the testing process, there exists a high probability that only a local minimum will be obtained. The reduction in temperature in order to reach equilibrium is done by implementing the subsequent geometric function:

$$T_{t+1} = \alpha T_t \tag{2.36}$$

The mentioned iterative process is repeated until no significant improvement in the objective function is noted where after the computational process will terminate. Similar to the tabu search algorithm, the process can also terminate when the maximum number of iterations have been reached. Also note that the SA algorithm is able to obtain a global solution as the calculation process allows the SA solution to jump out of a local solution by accepting deteriorated solutions (Sasikala & Ramaswamy, 2010).

2.7 Ant colony (swarm) optimisation

The Ant colony Optimisation (ACO) model was initially proposed by Marco Dorigo in 1992. The calculation methodology on which this method is based, was inspired by the behavior of ant colonies while scouring food. During the process of searching for food, ants usually try to identify the shortest path between their nest and the food source to minimise their effort. When the mentioned path has been identified by a group of ants, they will utilise a hormone know as pheromone to indirectly communicate with one another the logical decisions which were made to identify the path. This is done to enable the other ants to access the same route without any difficulty. The path which the ants will choose to use will be directly linked to the intensity of the pheromone. If a certain path is not used regularly, the pheromone trail will evaporate over time and the ants will lose interest. However, when the pheromone trail is strong, the path will be more frequently used and as a result will grow faster as compared to the less traveled paths (Afroozi *et al.,* 2014).

When the preceding is applied to optimisation problems, one can derive that the best solutions will have the highest intensity of pheromones and therefore will have an increased probability to be chosen. This principle can be applied in the optimisation environment to solve combinatorial optimisation problems. When implementing the ACO algorithm, artificial ants will be allowed to travel through the search space (from node to node) in order to obtain a solution. The movement of the artificial ants will be highly dependent on the previous actions which are stored in the algorithm's memory using a specific data structure. The pheromone intensities are only updated when the artificial ants have traveled through all the nodes with high intensities indicating good paths and vice versa. In order to prevent the ants from being stuck in local optima solutions, the intensity of the pheromones is progressively reduced by evaporation. In mathematical formulation, the probability of an ant k, which is at node i, to select a new node j to move to, can be calculated by means of equation (2.37). In this equation J_i^k represents the nodes of which ant k has not yet visited, variable $\tau_{ij}(t)$ denotes the concentration intensity of the pheromones at edge (i, j) in the t^{th} iteration. Variable n_{ij} represents the visibility of the different paths whereas variables α and β are

implemented to control the importance of the visibility versus the pheromone intensity (Khodia *et al.,* 2014).

$$p_{ij}^{k}(t) = \frac{[\tau_{ij}(t)]^{\alpha}[n_{ij}]^{\beta}}{\sum_{j \in J_i^k}[\tau_{ij}(t)]^{\alpha}[n_{ij}]^{\beta}}, \textit{If } j \in J_i^k \textit{ and } 0 \textit{ if } j \notin J_i^k \tag{2.37}$$

After completing an iteration, the pheromone intensity on each edge or path is updated by means of implementing Equation 2.34 with p being representative of the pheromone persistence and $\Delta\tau_{ij}$ being equal to the solutions identified at each iteration t.

$$\tau_{ij}(t) = pt_{ij}(t) + \Delta\tau_{ij}(t) \tag{2.38}$$

The function utilised to calculate $\Delta\tau_{ij}$ is provided below, with index n referring to the number of artificial ants being considered in the problem:

$$\Delta\tau_{ij} = \sum_{k=1}^{n} \Delta\tau_{ij}^{k}(t) \tag{2.39}$$

The quantity $\Delta\tau_{ij}^{k}(t)$ represents the amount of pheromone which is added on edge (i, j) per unit of length at the end of each iteration t, by the k^{th} ant. Note that if $(i, j) \in T^k(t)$, then $\Delta\tau_{ij}^{k}$ (t) is calculated by $\frac{Q}{L^k(t)}$. If however $(i, j) \notin T^k(t)$ then $\Delta\tau_{ij}^{k}(t)$ will equal a value of zero. $T^k(t)$ is indicative of the tour selected by an artificial ant, $L^k(t)$ is the length of the tour and Q is a user defined constant. Similar to the previous heuristic methods, the ACO algorithm will be iteratively executed until a solution is obtained and all constraints are satisfied before terminating the process (Afroozi *et al.,* 2014).

2.8 Summary

A technical overview of both exact and heuristic methods was provided in this chapter with specific focus on the linear and integer programming algorithms and their associated standard forms. A detailed discussion regarding the mathematics related to the simplex and branch-and-bound algorithms were provided as to ensure the correct modeling approach is applied by the author when formulating the UCEELD MILP problem in Chapter 4. The reason the mentioned methods are emphasised in this thesis is because the commercial software, Cplex, implements these exact approaches as solvers in the background. A high-level overview of the heuristic methods generally applied to the unit commitment and load dispatch problems such as the particle swarm optimisation, tabu search, hopfield neural network, simulated annealing, and ant colony optimisation methods were also discussed. No in-depth mathematical analysis was conducted on the heuristic methods, but a brief explanation of the principles on which each method is based, were given. In the subsequent chapter, an overview is provided on the technical aspects associated with coal fired, hydro and pumped storage power generation technologies. A detailed literature review on the solution methods applied by researchers to solve power utility optimisation problems is also provided.

Chapter 3

Technical Background

Power generation technologies operational overview

In this section the engineering basics behind coal fired, hydro and pumped storage power generation are provided as an introduction to the unit commitment and load dispatch optimisation problem. Details are provide on the different subsystems used within the mentioned power generation technologies and how these systems interact with one another in order to generate power. Note that technologies such as nuclear, photovoltaic and wind power are not considered for the purpose of this dissertation as mentioned in Section 1.2.2 and will not be elaborated on in the following chapter.

3.1 Coal fired power generation

The primary function of coal fired or also know as thermal power stations are to transform demineralised water into superheated steam using coal as energy source (Govidsamy, 2013). In order to allow a coal fired power station to generate high enthalpy superheated steam, different processes are implemented which include feedwater and condensate, boiler, turbine and auxiliary systems such as portrayed in Figure 3.1. There is however other subsystem also present in a coal fired station, but the scope of the dissertations is limited to the above mentioned to provide a high level overview of the technology.

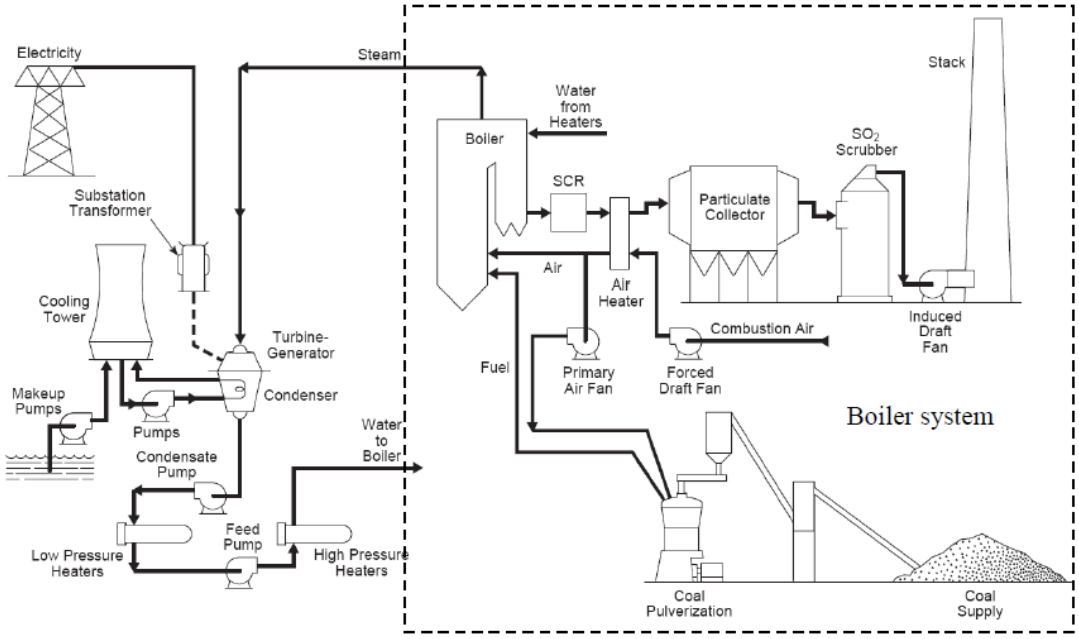

Figure 3.1: Overview of a general coal fired power generation process (Govidsamy, 2013)

The principle on which coal fired power generation is based is known as the Rankine cycle. This is a thermodynamic cycle where a working fluid such as demineralised water is both vaporized and condensed while flowing through different sets of processes. A basic Rankine cycle consists of a boiler, turbine, condenser and compressor (Koretsky, 2013). To significantly increase the efficiency of a basic Rankine cycle, feed water heaters are also added to the process to recover some latent heat of vaporization which was used during the vaporization process **(process 2 - 3)** (Jestin, 2017). The working fluid is transported to the boiler by means of boiler feed pumps (also known as the compression **process 1 - 2**). Energy used for power generation in the turbine is produced by means of pulverised combustion in the boiler. The energy is transferred from the boiler to the turbine by means of heat transfer tube banks knows as superheater, reheater and economizer sections **(process 3 - 5 and 6 - 7)**. Depending on the design of a coal fired power station, the station can either consist of a single high pressure (HP) turbine or multiple HP, intermediate pressure (IP) and low pressure turbines (LP). The amount and size of turbines installed will also govern the rated megawatt output of the power station **(process 5 - 6 and 7-8)**. After passing the turbine, energy needs to be transferred between the system and the surroundings to ensure that the system is returned to its original state in order to complete the cycle. The energy transfer is induced by means of heat expulsion in the condenser **(process 8 - 1)** and auxiliary power input to the compressor. The preceding can be represented graphically by means of a temperature-entropy diagram as depicted in Figure 3.2 (Wu *et al.,* 2014). This diagram is an overview of a sub-critical Rankine cycle, similar to the principle on which some South African power stations are designed on. In the South African context, coal fired power stations are used as base load stations to regulate and maintain the grid frequency stable.

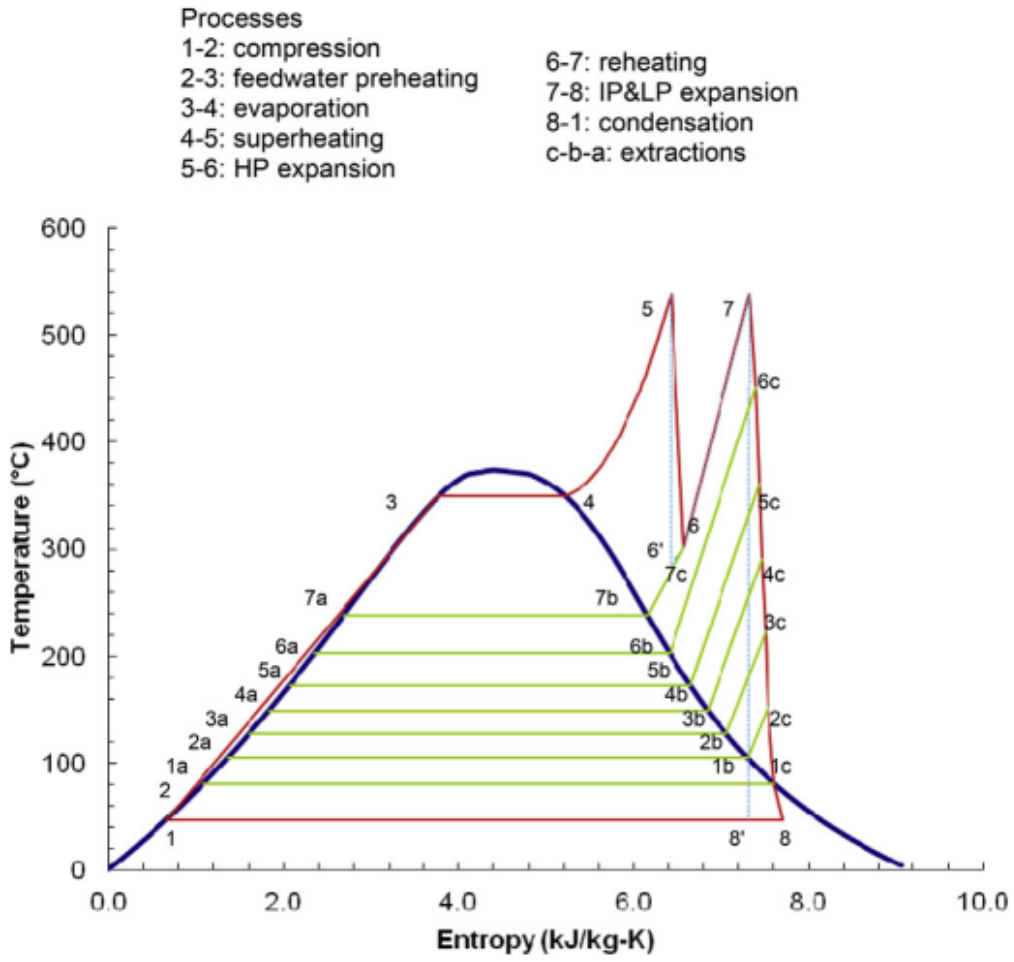

Figure 3.2: T-S Diagram of a typical sub-critical Rankine cycle (Wu *et al.,* 2014)

A more detailed analysis of the different processes depicted in the Rankine cycle follows in the sub-

sequent sections with an elaboration on some of the design considerations. Note that the equations depicted in the subsequent sections are provided for understanding the principles behind the power generation process only and was not explicitly used in formulating the UCEELD problem.

3.1.1 Compression system

The initial step in the Rankine cycle as depicted in Figure 3.2, is to raise the pressure of the working fluid to the desired turbine inlet pressure setpoint. In order to achieve a pressure increase, boiler feed water pumps (BFP's) are added to the power generation process (Teleman, 2016). Depending on the design conditions of the turbine, different BFP's may be utilised to satisfy the pressure increase required. The work necessary to increase the fluid to the desired pressure is calculated by means of equation (3.1). Note that variables $\hat{h_{2s}}$ and $\hat{h_1}$ are representative of the isentropic outlet enthalpy and actual inlet enthalpy of the working fluid, $\dot{m}$ depicts the mass flow of the fluid whereas n_C presents the isentropic efficiency of the BPF's (Kapooria *et al.,* 2008).

$$W_{12} = \frac{\dot{m}(\hat{h_{2s}} - \hat{h_1})}{n_C} \tag{3.1}$$

Depending on the efficiency variations, the work required per pump may differ. To determine the isentropic efficiency, the following equation is used with variable $\hat{h_2}r$ representing the actual enthalpy at the outlet conditions (Kapooria *et al.,* 2008):

$$n_{12} = \frac{\hat{h_{2s}} - \hat{h_1}}{\hat{h_{2r}} - \hat{h_1}} \tag{3.2}$$

The inlet conditions to the BFP's are usually single phase saturated by design as majority of BPF's are not able to handle two-phase mixtures (Koretsky, 2013). Due to the fact that the molar volume of a liquid is much less than that of a vapor, the power required by the BPF's to compress the fluid is only a small fraction of the total power generated by the turbine. Generally, the power used for the fluid compression is termed auxiliary power. After being compressed, the fluid is in the subcooled phase due to its high pressure and low temperature. BPF's generally consist of multiple stages from where spray water can be extracted at any intermediate pressure to control the steam temperature. The next step in the Rankine cycle is to preheat the feedwater prior to it entering the boiler.

3.1.2 Feedwater and preheating

The purpose of installing feedwater heaters in a power generation process is twofold. The first reason for installing feedwater heaters is to reduce the operational cost required to produce power (amount of coal consumed) and secondly to prevent thermal shock to any metal components in the boiler when the working fluid is introduced back into the system. Depending on the design of the turbine, the inlet temperature of the steam required can range from anything between 500 $°C$ to 540 $°C$. Adding feedwater heaters to the cycle, would reduce the ΔT which exist between the inlet and outlet of the boiler and will result in less primary energy (coal) being consumed to meet the desired temperature and pressure setpoints. Depending on the design of the power station, multiple HP and LP heaters may be present in the cycle (Jestin, 2017). HP heaters are designed to use steam extractions after the HP turbine as heating source whereas LP heaters would utilise steam extractions from either the IP or LP turbines as heating medium (Pieterse, 2016). Implementing the mentioned components into the Rankine cycle would mean that less energy will be lost at the cooling towers and resultantly an increase in the overall thermal efficiency of the station would be gained (Wu *et al.,* 2014). At times, the feedwater heaters may be exposed to tube leaks or fouling which will reduce the efficiency of both the heaters and the overall power generation process and result in more coal being consumed for a megawatt generated. To determine the effect of plant defects on the heater efficiency, equation (3.3) can be used (mass and energy balance) with $E_{feedwater}$ representing the energy contained in the feedwater, $E_{bledsteam}$ the energy in the extraction steam and $E_{heaterflashbox}$ the energy in the

feedwater entering the flashbox of the heater from the previous heater:

$$n_{heater} = \frac{E_{feedwater}}{E_{bledsteam} + E_{heaterflashbox}} \tag{3.3}$$

When comparing the performance of two identical generating units, it might be that one would be able to consume less coal than the other due to more efficient feedwater heaters (less defects) being present in the cycle. After exiting the feedwater heaters, the working fluid enters the heat transfer regions of the boiler.

3.1.3 Evaporation, superheating and reheating systems

The boiler consists of different heat transfer regions that include the evaporation, superheater and reheater sections (Wu *et al.,* 2014). In these sections, the energy released by the combustion process of the fuel (coal) is transferred to the working fluid (steam). The fluid is heated from liquid to saturation phase in the economizer section whereas in the remaining sections the fluid is heated to the superheated vapor phase. The mechanisms by which heat is transferred to the economizer and reheater sections are by means of convective and conductive heat transfer. In the superheater sections, the mentioned mechanisms are also present, but with radiation heat transfer being the primary source. The quantity of radiation, convection and conduction heat transferred to the different boiler sections can be calculated by means of applying equations such as the Stefan-Boltzmann's Law (3.4), Newton's law of cooling (3.5) and Fourier's law (3.6) respectively (Welty *et al.,* 2008). In the mentioned equations, variable A refers to the area of heat transfer, dT to the temperature difference across the material, T to the flame temperature, ϵ to the heat uptake capability, σ represents the Stefan-Boltzmann constant, variable h_c depicts the heat transfer coefficient of the process, k refers to the thermal conductivity of the material and lastly s depicts the material thickness.

$$q = \epsilon\sigma T^4 A \tag{3.4}$$

$$q = h_c A dT \tag{3.5}$$

$$q = kA\frac{dT}{s} \tag{3.6}$$

It is apparent from critically evaluating the above equations that, depending on the cleanliness of the heat transfer sections, which influences the value of variable k, the boiler's heat pick-up capability will either increase or decrease. If a boiler is filthy it will result in less heat being transferred to the steam and consequently lead to a reduction in the boiler's thermal efficiency performance. The loss of energy as a result of ineffective heat transfer in the boiler is termed dry fluegas losses (van Rooyen, 2014). Generally, systems such as sootblowers are installed to manage the boiler's cleanliness, but due to equipment unavailability at times, the heat transfer sections cannot be cleaned sufficiently. A reduction in the boiler's efficiency as a result of the foregoing, will lead to more coal being consumed to maintain the boiler at the desired operating conditions. Another method to determine the rate of heat transfer at each section of the boiler other than the above mentioned, is by applying the enthalpy calculation method (Koretsky, 2013). This method can also be used to determine the extent of tube fouling when comparing the difference in design data with the answers obtained from equation (3.7).

$$Q_B = \dot{m}(\hat{h_{out}} - \hat{h_{in}}) \tag{3.7}$$

Similar to what was mentioned in Section 3.1.2, if two identical generating units are compared to each other, one unit might be able to outperform the other by consuming less coal, given that its boiler cleanliness is managed more effectively. After the steam exits the superheater and reheater sections of the boiler, it is transported to the HP, IP and LP turbines respectively.

3.1.4 HP, IP and LP expansion system

The superheated steam enters the HP turbine through governor valves which are used to regulate the inlet pressure to the turbine. As the steam moves through the turbine, its energy content is reduced by means of expansion and cooling while producing work (Sampath, 2015). The rate of work produced is dependent on the mass flow of steam entering the turbine and its associated operating conditions. To calculate the work produced, the first law of thermodynamics can be applied with the assumption that the potential and kinetic energy and heat transfer through the turbine casing is negligible. Refer to (3.8) for the mathematical formulation with variables h_5 and h_{6s} representing the inlet and outlet enthalpy of the steam flowing through the HP turbine, respectively, and n_T referring to the isentropic efficiency of the turbine (Kapooria *et al.,* 2008).

$$W_{56} = \dot{m}(\hat{h_5} - \hat{h_{6s}})n_T \tag{3.8}$$

The amount of work produced by the turbine is, however, highly dependent on the turbine efficiency which can be influenced by plant defects such as deposits on the balding. This implies that due to plant defects, less megawatt will be produced by the HP turbine per kilogram of coal consumed by the boiler. In order to determine the isentropic efficiency, (3.9) can be applied (Jestin, 2017). Note that variable h_6 depicts the actual outlet enthalpy of the steam leaving the HP turbine.

$$n_{56} = \frac{\hat{h_5} - \hat{h_{6r}}}{\hat{h_5} - \hat{h_{6s}}} \tag{3.9}$$

In the HP turbine, the steam enters in the superheated state in order to prevent condensation during the expansion and cooling process. If the steam entered the turbine in saturated state, however, liquid would form during the generation process. Liquid formation would result in wear of the turbine blades and consequently long term action will need to be taken to replace the turbine. After exiting the HP turbine, the steam can either be routed to the condenser or to the boiler depending on the design of the coal fired power station. In Figure 3.2, it is apparent that an IP and LP turbine is installed in the process under consideration. Prior to entering the remaining turbines, the steam is first superheater in the reheater section of the boiler. Note that the reheated steam contains less energy than that of the steam entering the HP turbine due to the reduced pressure and mass flow available. After being reheated, the steam expands through the IP and LP turbines in order to produce work where after the process fluid is routed to the condenser (Sampath, 2015). To calculate the isentropic efficiencies and work produced by the IP and LP turbines, the same equations can be used as mentioned above while considering the change in inlet and outlet conditions.

3.1.5 Condensate system

In the condenser, the working fluid changes from saturated vapor to saturated liquid. Phase change of the working fluid is induced by means of circulating low temperature cooling water through the tubes of the condenser to extract the latent heat of vaporization (van Rooyen, 2014). The temperature difference between the steam and the cooling water initiates the heat transfer process which continues until equilibrium is reached. Note that the phase change occurs at a constant pressure. Vacuum pumps known as ELMO pumps are connected to the condenser to maintain the process under vacuum which allows the fluid to flow from the boiler to the condenser. The heat expulsion is however dependent on the cleanliness of the condenser, and will reduce as the condenser tubes are exposed to fouling. A reduction in the condenser's heat transfer capability will directly influence the condenser's efficiency (n_{81}) as calculated by equation (3.10) (Kapooria *et al.,* 2008). In the equation below, variable T_{ci} depicts the cooling water inlet temperature, T_{co} the cooling water outlet temperature and T_{sat} the saturation temperature of the water at the specified operating pressure of the condenser.

$$n_{81} = \frac{T_{co} - T_{ci}}{T_{sat} - T_{ci}}100 \tag{3.10}$$

If the outlet temperature of the cooling water is able to reach the saturation temperature of the water in the condenser, the efficiency will equate to a value of 100%. This is, however, an ideal situation and in reality plant defects and inefficiencies will result in lower condenser efficiencies being obtained. Lower efficiencies will result in the condenser operating at increased backpressures and will consequently require the unit to consume more coal for each megawatt generated (Jestin, 2017). If two units need to be compared to one another with regards to cost efficiency, the unit with the condenser efficiency less than the other will consume more coal per megawatt generated and will be considered more expensive. The heat energy extracted from the condenser can be estimated by means of using the same principle as was stated in (3.7).

3.1.6 Milling and combustion system

The starting point of the combustion system in a coal fired power station is the milling plant. This system consists of mills or also know as pulverisers, pulverised fuel (PF) transportation piping and burners. Mills are utilised to grind coal to the desired particle size distribution (into PF) prior to being dried and transported by means of primary air (PA) to the burners. The transportation process occurs through the PF piping, with the piping connecting the burner inlets to the mill outlets (Govidsamy, 2013). In some designs, classifiers are installed after the mill outlets, prior to the PF piping, to prevent too large particle sizes from entering the PF piping as it might result in PF settling and consequently pipe blockages (Afolabi, 2012). In addition to the mass flow of coal and primary air entering each burner, secondary air (SA) is added at each burner to provide enough air for complete combustion. Ensuring that each burner receives sufficient amount of air, it will assist to produce an even distribution of heat through the furnace. The quantity of SA introduced to each burner is controlled by means of auxiliary air dampers. Mixing of the PF, PA and SA occurs in front of the burners. The fuel-air mixture enters the boiler where after the combustion of coal occurs and by-products such as flue gas and ash particles are formed. The flue gas enters the heat transfer regions as mentioned in Section 3.1.3 and acts as heating medium for the steam.

3.1.7 Draft group and airheater system

The draft group system consists of three types of fans which include the PA fans, forced draft fans (FD) and Induced draft fans (ID). The PA fans are installed to supply the milling system with enough air to transport and dry the PF. These fans impart a high static pressure onto the PA system to enable the coal-air mixture to overcome the system resistance from the mills to the burners. The FD fans are responsible for supplying the burner corners with SA which is the main source of combustion air. The FD fans take suction from the ambient surroundings. Depending on the design of the draft group system, the PA fans can either take suction from the FD fan discharge or from the ambient air. Normally there are two FD and two PA fans installed per boiler. After the flue gas (i.e. SO_x, NO_x, O_2, CO_2 and CO) and ash particles are produced in the combustion chamber, the ID fans are utilised to suck these by-products through the boiler. The ID fans are designed to maintain the boiler under vacuum by controlling the boiler operating pressure slightly below atmospheric pressure (or 100 Pa). This is done to ensure safe operating conditions and prevent the combustion flames from propagating outside of the furnace walls (Govidsamy, 2013).

Air heaters are added to the draft group system to improve boiler efficiency (Manivel *et al.,* 2017). The aim of the air heaters are to extract heat from the warm flue gas exiting the boiler and transfer it to the cold air entering the air heaters, with the cold air being supplied from both the FD and PA fans. Depending on the design implemented, there may only be rotary SA heaters installed or a combination of both rotary SA and tubular PA heaters. The SA heaters are generally regenerative type heaters where metal plates are heated by the flue gas when the gas passes through the plates. The SA air is heated by the mentioned plates when it passes through the heater during each rotation. In the tubular air heaters, flue gas passes through the tube side in order to heat the metal structure. The air flows around the heated tubes, inside the PA heater casing from where the PA is heated (Swart, 2016). After heating, the PA is transported to the milling system whereas the SA is sent to

the burner corners.

3.1.8 Net power generation and cycle efficiency

In the Rankine cycle, net power generated by a coal fired unit is defined as the power produced by the HP, IP and LP turbines subtracted by the amount of auxiliary power utilised by the boiler feed pumps. However, in reality, other components such as the mills and fans also consume quite a significant amount of auxiliary power and this needs to be taken into consideration when computing the net power production. The net work obtained from a generating unit can be calculated by means of equation (3.11).

$$W_{net} = W_{turbine} - W_{auxiliary} \tag{3.11}$$

The cycle efficiency or also known as thermal efficiency is a dimensionless performance measure used to evaluate the effectiveness of the power generation process. The cycle efficiency is defined as the fraction of how the energy which is added to the boiler by means of coal combustion, has been converted to useful energy (Anjali & Kalivarathan, 2015). Refer to equation (3.12) for the calculation of the thermal efficiency.

$$n_{thermal} = \frac{W_{turbine} - W_{auxiliary}}{Q_B} \tag{3.12}$$

If for example any of the above mentioned plant sections does not perform as per its specification, it will result in a reduction in the thermal efficiency performance and consequently, more heat energy will be required to maintain the net power generation at the required setpoint value. This, in turn, will result in an increase in the operational cost of the utility which is undesirable.

3.1.9 Auxiliary systems

Additional to the plant sections discussed throughout Sections 3.1.1 to 3.1.8, auxiliary systems such as coal, ash, water treatment, hydrogen, auxiliary cooling, fuel oil, compressed air and fire water plants exist within a thermal power station (Palanichamy, 2015). Each of the mentioned plant sections is either responsible for the supply of primary resources such as coal, water, hydrogen and fuel oil to the boiler or is installed with the purpose of handling the ash by-products originating from the combustion process. Although the auxiliary plant sections form an integral part of the entire power generation process, these plant sections generally do not contribute to significant operational financial losses. Even though there might be instances where these systems contribute to the mentioned losses as a result of extended unit downtime caused by auxiliary system unreliability and unavailability, such cases rarely occur. Factors which have more of a direct impact on the day-to-day operational financial losses, are the boiler and turbine plant defects (as mentioned in the preceding sections) which leads to power station partial load losses. Note that some of these defects can only be addressed during extended maintenance opportunities which include interim repairs (IR) or general overhauls (GO). It is for this reason that the focus in this literature review is rather on the partial load losses instead of constraints that might be encountered on the auxiliary systems.

3.1.10 Power station partial load loss constraints

When considering coal fired generating units, partial load losses can be defined as defect induced conditions resulting in the temporary reduction of a unit's power output capability, with the limitations being present until the defects are either addressed or resolved. The major partial load losses power stations are generally faced with include air heater leakages, poor emissions performance, high condenser backpressure, and subgrade coal qualities. Note however that there are other load losses additional to those mentioned, but they occur less frequently.

1. Air heater leakage related load losses
Air heater leakages are caused by the deterioration of the PA and SA heaters. The mechanisms for PA heater deterioration include both fly ash erosion and sulfuric acid corrosion. These mechanisms cause the wall thickness of the heat transfer tubes in the PA heaters to reduce until holes start to form in the tubes. With holes present in the heat transfer tubes and the tubes being under vacuum conditions induced by the ID fans, the high pressured PA leaks into the flue gas. This increases the gas volume flowing to the ID fans. The escalation in flue gas volumetric flow imparts additional strain onto the ID fans, as the fans need to try and maintain the furnace pressure at the required set point given the rise in gas flow. In order to process the additional load, the ID fans are required to increase its output. Similar is true for the PA fans as these fans need to counteract the effect of the lost PA. To address the preceding, the PA fans are required to increase its volumetric output to maintain a constant mass flow of air to the milling plant. Deterioration on SA heaters are caused by air heater blockages and tearing of rubber seals. A reduction in the SA heaters$'$ performance mainly affects the ID fans as blockages increase the system resistance to which the ID fans are exposed to and the tearing seals increase the additional amount of SA being introduced into the flue gas stream. The mentioned conditions require the ID fans to increase its volumetric output in order to maintain effective furnace pressure control.

Initially, the PA and ID fans have enough additional capacity to counteract the increased loads. However, as the PA and SA heaters deteriorate, the rate of air heater leakages steadily increase to such an extent that the ID and PA fans eventually run out of capacity (operating with fan vanes at 100% open). When the ID fans are saturated, the unit operators are obligated according to the fossil fuel firing regulations to deload the unit with a few megawatts in order to bring the ID fans into control range (vane positions $\leq$ 90%). This is to ensure the unit remains capable of controlling its furnace pressure below atmospheric pressure. The PA fans are however allowed to operate at 100% vane positions as these fans are only required to supply a set quantity of PA to the milling plant and is not used to control any process fluctuations. However, if the PA leakage is high enough the PA fans will not be able to supply the required amount of PA to the milling plant and will place the unit at risk of PF pipe blockages. In such a case, the unit operator is again required to deload the unit until a sufficient amount of PA is supplied to the milling plant. Although the FD fans will also be required to increase its volumetric output with the deterioration of the SA heaters, usually these fans are designed with far more capacity than what is required and the fans remain well below 80% vane positions.

2. Emissions related load losses
Poor emissions performance are generally caused by mechanical defects on the electrostatic precipitator (ESP) plant. The mechanical defects may include but are not limited to blocked precipitator hoppers, choked or defective precipitator conveyors, poor performing collector and discharge electrodes and unavailability of the rapping systems. Blocked hoppers are usually caused by defective hopper heaters or closed knifegates. Process related aspects contributing to reduced emissions performance include boiler sootblower unavailability (as mentioned in Section 3.1.3), resulting in ineffective heat pick-up in the boiler causing high backend temperatures. An increase in the temperatures entering the ESP's causes the volumetric flow through the system to increase, contributing to a reduction in gas residence time and consequently a decline in ESP efficiency. Air heater leakages also negatively affect the performance of the ESP's as an increase in volume flow to the ID fans, through the ESP's, has the same effect as mentioned for the temperature influence and causes a decrease in ESP efficiency. If any of the above mentioned defect conditions are present on the plant, it will hinder the unit's ability to maintain its emissions below the legislative limits as stated in Section 1.1.2. In such a case, the unit operator is obligated to reduce the unit's output to maintain the emissions below the desired limits. If however the extent of the defects are too severe, sometimes even load reduction does not have enough of an impact to lower the emissions production. The only remaining option is to request unit downtime to conduct partial ESP plant repairs. Although temporary repairs will assist with the emissions performance, extended maintenance interventions such as IR's and GO's

are required to completely resolve the issues.

3. Condenser backpressure related load losses
The performance of a condenser will deteriorate due to factors such as air ingress, tube fouling or scaling, low cooling water flow rates, defective vacuum extraction pumps, and condenser tube leaks. The mentioned aspects will reduce the condenser's efficiency and resultantly lead to an increase in the unit's backpressure. As depicted in Section 3.1.5, increased backpressures will lead to additional coal being consumed per megawatt generated. With an increase in coal consumption, extra PA and SA will need to be supplied to the boiler to prevent sub-stoichiometric combustion from occurring. The introduction of additional volumetric air flow to the boiler will result in further strain being imparted on the ID fans. The preceding will require the ID fans to increase its volumetric output in order to maintain stable process control. As noted in Section 3.1.10 part 1, the unit operator will be required to take a load loss if the high condenser backpressure results in ID fan saturation as to ensure the fans remain in control range. Depending on the design of the condenser plant, some generating units also have controls implemented to automatically deload the unit when the condenser backpressure exceeds the allowable limit. Deloading will continue until the unit is within an allowable operating range. The reason such logics are implemented is to prevent saturated steam from pushing back to the earlier stages of the LP turbine, as there exists a possibility that the steam will start to condense and form water droplets in the upper stages. The mentioned stages are not designed to handle water droplets and eventually, the droplets will start causing physical damage to the blades. High condenser backpressures will also induce additional stresses on the turbine balding which will lead to stress corrosion cracking. Similar to the above mentioned, the defect induced conditions will require either IR or GO maintenance opportunities to resolve the issues. However, if not resolved the affected units will be subjected to load losses and will not be able to attain full load operation.

4. Coal quality related load losses
Generally, a power station is designed to operate within a certain coal quality band of say 14 MJ/kg - 21 MJ/kg. These qualities are used as a guideline to determine the capacity required for the design of the PA, SA and ID fans as well as the milling plant and its associated components. However, in instances where coal qualities below the design band are supplied to the station, it causes equipment such as the milling plant and PA fans to run above its design load, when the unit is operating at maximum continuous rating. This is due to the fact that the equipment needs to increase its feed of air and coal (with a lower heat content) to the boiler as to maintain a constant heat input into the boiler. Operating equipment above its maximum capacity is both detrimental to the plant and can lead to unsafe operating conditions. In order to prevent the foregoing from occurring, unit operators are required to take load losses to maintain the milling plant and PA fans within its design operating ranges as specified by the fossil fuel firing regulations. In order to counteract the effect of poor coal qualities, better coal sources can be identified and the coal obtained from the mentioned sources can be distributed to the different units. This is however not always practicable due to financial and coal constraints and consequently, the units are required to run with load losses. Note however that the strain induced on the milling plant and PA fans are also aggravated by poor equipment performance as mentioned in the foregoing sections.

3.2 Hydro power generation

Hydro power generation is the process where electricity is generated by means of using water energy to drive turbines, which are connected to generators. In this process, the kinetic energy contained within the flow of water, originating either from a dam or river, is converted to rotational energy in a turbine. The rated power output of a hydro unit is dependent on the turbine design and available kinetic energy. This type of technology is widely implemented where geographic and climatic conditions are suitable as the power contained by the reservoirs of a hydro unit is highly dependent on the rain patterns or volume of the upstream catchment areas. The generation process is

environmentally friendly as it uses a renewable energy source which does not produce any pollutants. Hydro power stations are generally used for peak load regulation as it is quick to commit and is capable of responding rapidly to load variations. The construction cost associated with this kind of technology is however expensive as it sometimes requires complex tunneling and building of dams (Jacobs, 2017). Refer to Figure 3.3 for a high-level overview of the hydro power station technology.

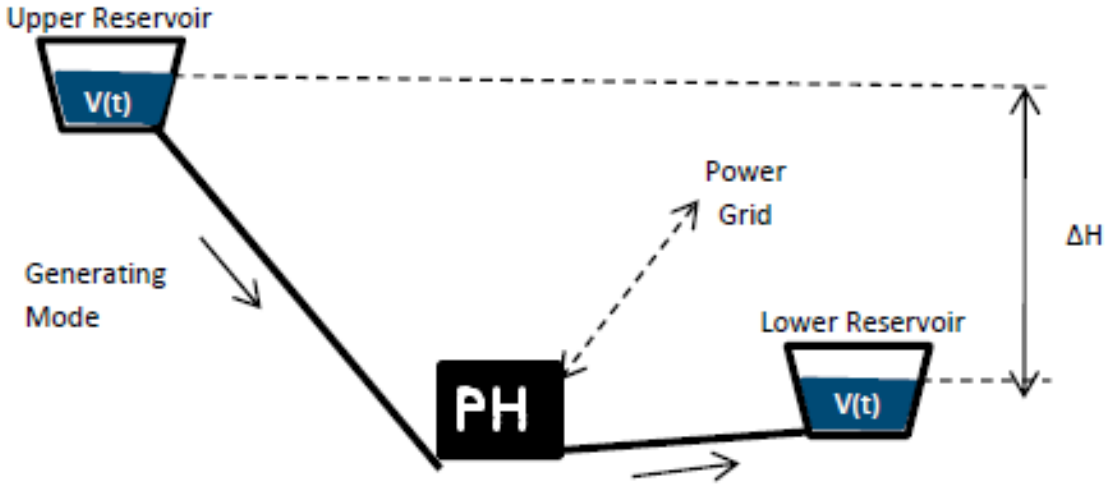

Figure 3.3: Overview of hydro power generation process (Antal, 2014)

The next section gives an overview of the equipment, systems, and layout of a general hydroelectric station. Note however there is a multitude of variations and combinations of equipment as each hydro station is tailored to the topography and environmental surroundings in which it is situated. For the purpose of this dissertation, the focus will only be set on medium to high head storage reservoirs, with run-of-the-river stations being excluded from the study.

3.2.1 Upper reservoir

The upper reservoir, or also known as the dam, is generally erected in a watercourse to hinder natural water flow and force a rise in the water level upstream of the dam. The construction of the dam may be of such sort that the water level is several hundred meters higher than the initial river flow level. The high water level is to ensure enough water head is stored upstream of the hydro units as to obtain maximum profit gain of the contained potential energy (Ferreres & Font, 2010). The dams are usually designed to be able to discharge water from the dam to either satisfy downstream requirements (via spillways) or to supply the required quantity of water to the generating units. When designing the upper dams, different mechanisms such as pushing and tipping forces need to be taken into consideration as these forces can result in the destruction of the dam structure. Different permanent and temporary dam configurations include embankments, gravity, arch, buttress, coffer and rubber dams. Headworks known as spillways or sluice ways are generally installed on these dams to control the amount of water contained in the upper dams and protect against overfilling by directing the water directly into the downstream river (Jacobs, 2017).

3.2.2 Intake works

The intake structure is the connection between the dam and the pentstock system. These structures are generally constructed by using reinforced concrete in order to guarantee structural integrity. The intake is formed by means of excavation on the reservoir bed at the inlet of the penstock. At times it is necessary to excavate below the penstock inlet as to allow silt to accumulate below the discharge level (Jacobs, 2017). In some underground installations, circular tower intakes are implemented which consist of circular vertical intakes used to direct water to the turbines. Trash or coarse screens, floating booms or trash racks are installed at the intake to prohibit debris such

as leaves, plants or logs from entering the penstock (Bratko, 2013). To determine the extent of screen blockage, differential pressure transmitters are installed on the screens to prevent extensive blockages. After the screens, control gates are installed to isolate the penstock from the reservoir either during emergency or normal operating conditions.

Penstocks are large transportation ducts or tunnels used to convey the water from the reservoirs to the turbines. The penstocks are usually bored through solid rock faces and reinforced with either concrete or steel depending on the pressures and rock conditions involved. As the penstocks get closer to the turbine, the amount of static water pressure increases. With an increase in static pressure the support, tunnel lining and sealing arrangements are increased. The penstocks are usually equipped with surge tanks to either damper or absorb any oscillations or pressure shock waves induced by emergency conditions (Jeyalalitha, 2008). Main inlet valves are installed at the outlet of the penstocks, prior to the turbines. The purpose of the valves is to isolate the turbines form the penstocks, in case undesirable conditions arise. The type of main inlet valves installed may either be butterfly valves or ball valves depending on the pressures to which it is exposed to. Depending on the design of the turbine, water exiting the main valve can either be distributed horizontally or vertically to the turbine by means of spiral casings. The mentioned casings are connected to the penstocks via the main valves and expansion joint compensator. The expansion joints are added to prevent expansion forces from being transferred to the turbines (Jacobs, 2017).

3.2.3 Power house

The power house is the area where the turbines, generators, and transformers are contained. Within the power house different types of turbine designs can be installed of which can either include reaction or impulse turbines. The different types of impulse turbines entail the Pelton, Turgo, and Crossflow turbine whereas the reaction turbines comprise of the Francis and Kaplan turbine (Lombard, 2010). The principle on which the impulse turbines are based is to induce rotation by means of using high velocity jets of water to impact the turbine buckets connected to the turbine wheels. The velocity of the water stream reduces to a value of almost zero after transferring its kinetic energy to the impluse turbine, where after the water is sent to the tailrace. In contrast to the impulse turbines, the reaction turbines are completely filled with water. Rotation is induced in these turbines by means of a pressure difference across the turbine runners. The pressure difference pushes the water through the runners and as a result, causes the turbines to start rotating (Jacobs, 2017). In both scenarios, the net power which these hydro turbines are capable of generating, is dependent on the available system head (H_n), the water density (p), gravitational constant (g), volumetric flow rate of the water (Q) and the turbine efficiency (n). The turbine efficiency is estimated by taking into account the manometric, mechanical and volumetric efficiencies in the system. Depending on the efficiency variation of the entire system, the net power generated by each hydro power turbine will vary accordingly. Also, note that the volumetric flow of water entering the turbines are controlled by governor valves. Refer to equation (3.13) for the calculation of the hydro turbine net power generation (Ferreres & Font, 2010).

$$P_{net} = QH_npgn \tag{3.13}$$

In order to accurately estimate the head available for power generation, head losses such as the following need to be taken into consideration (Ferreres & Font, 2010):

1. Head losses due to pipe friction, using the Darcy Weisbach principle and the Moody diagram to calculate these losses.

2. Head losses due to trash racks, using the formula developed by Kirschmer as calculation method to estimate the losses.

3. Head losses due to sudden contractions in process piping leading to the turbines, as estimated by the coefficient of contraction method.

4. Head losses due to inlet valves as depicted on the valve's performance data sheets.

The turbine-generators can either be installed vertically or horizontally. In the horizontal installation, the machines are equipped with journal bearings used to maintain alignment of the generator and to support the shaft during rotation. With the vertical installations, the machine shaft is used to transfer the rotational energy directly to the generator from the turbine. The generator consists of a rotor within which multiple DC powered electro-magnetic elements are contained. A stator with electrical windings surrounds the rotor and is used to induce electrical energy during rotation. Other designs exist such as older vertical machines where the magnets contained in the rotor are energized by a slip ring and brush type exciter. However, in the modern era, a static exciter is rather used instead of the slip ring arrangement. Thrust bearings are also installed on the turbine-generator to support the weight of the entire system whereas guide bearings are used to ensure correct alignment of the entire rotor system (Jacobs, 2017). Transformers are also installed in the power house to step-up the generated voltage to enable the power to be sent through the transmission lines to the grid (Zoi, 2013).

3.2.4 Tailrace

After the water exits the turbine system, it is either transported in an open channel or duct to the watercourse, dam or reservoir. These channels or ducts act much like spillways with the water being directed to the watercourse with as minimum erosion as possible. When considering underground stations, the water is initially transported through tunnels before being discharged into the tailrace. In both arrangements, the water introduced back into the watercourse is at its minimum level of energy and results in the water entering the system at a very low velocity. Draft tubes are installed at the discharge of the turbine prior to entering the tailrace with the purpose of creating a siphon or vacuum effect that sucks the water from the turbine. By implementing the draft tube, the effective head over the turbine is increased to allow a rise in the turbine's power output. Note that the tailrace has been designed as the lowest point of the entire hydro system in order to ensure that all of the potential and kinetic energy stored in the water is used during the generation process (Jacobs, 2017).

3.3 Pumped storage power generation

Pumped storage schemes are similar to hydro generation technologies, with the difference being that the turbines of a pumped storage scheme have the capability of functioning as both turbines and centrifugal pumps, whereas hydro stations only have standard turbine installations. Pumped storage units are designed to act as mechanical batteries in that these units are able to generate power using the same water repetitively. As a result of the water recycling capability, these units are less dependent on the water flow of the upper catchment area or the rainfall conditions. Identical to hydro stations, pumped storage units consist of both upper and lower reservoirs. During low demand periods, pumped storage schemes are operated in pumping mode in order to covey the water from the lower to the upper reservoir. When the electricity demand increases, the water is released from the upper reservoirs to drive the turbine generators and produce power (Prasad *et al.,* 2013). After generation, the water is discharged to the lower reservoir where it is ready to be conveyed back to the upper reservoir by means of pumping. Pumped storage stations are designed to be net consumers of electricity as a result of the pumping functionality of the facility. Generally, these stations will consume up to 30% more electricity than what it generates (Hunt *et al.,* 2014). When operating in pumping mode, the generator acts as a motor with the turbine and rotor rotating in the opposite direction as compared to the direction of rotation in generating mode. to prevent excessive starting currents on the generator in pumping mode, a pony motor is installed to the machine shaft to assist the pump turbine to accelerate to the required operating speed, where after the generator takes over. The same motor is utilised during shutdown operation to brake the unit. The reasoning behind constructing pumped storage units are not because of its energy efficiencies, but because of the technologiesábility to supply instantaneous power to the grid under constrained

condition. Refer to Figure 3.4 for a high-level overview of the pumped storage power generation technology (Jacobs, 2017).

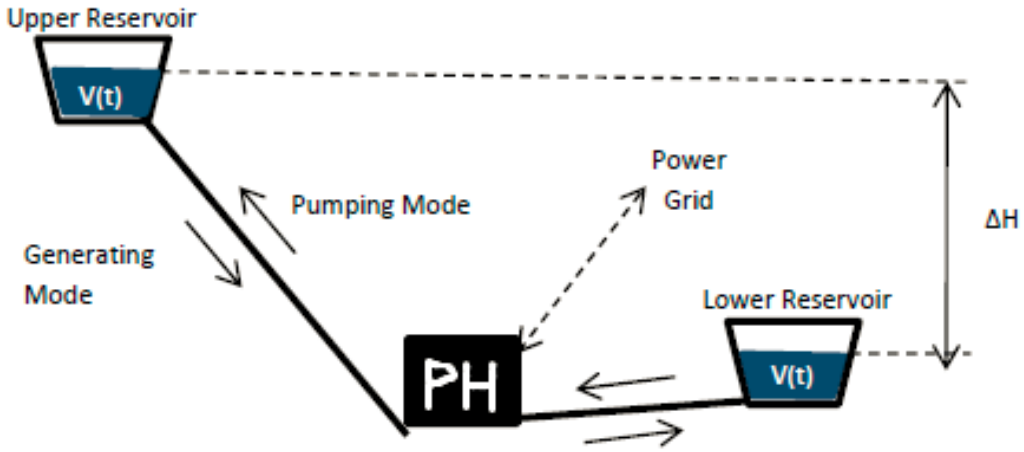

Figure 3.4: Overview of pumped storage power generation process (Antal, 2014)

Considering the above layout, it is apparent that the design layout of a pumped storage station is similar to the previously mentioned hydro stations. The major difference between these two technologies is, however, the pumping functionality and the turbine generator design. Pumped storage units can be installed with two different turbine configurations of which include a separate turbine and pump or a combined turbine/pump setup (Norang, 2015). An advantage to the separate turbine and pump configuration is that it minimises the transition time required to switch between pumping and generating mode. However, installing such a system would require additional mechanical and electrical equipment as well as a larger power house structure resulting in a rise in capital expenditure. Contrary to the foregoing, combined configurations require less expensive structures but the transition time between pumping and generating modes increase significantly. In addition to the different configurations, pumped storage units can either be designed as single speed or adjustable speed pumping units. Note that variable speed pumping units are more efficient in comparison to single speed units due to the elimination of control valve/gate throttling to control the water discharge rates. Depending on the pump efficiencies of a pumped storage station, the units would require approximately 90 minutes of pumping for each hour of full load generation in order to refill the upper reservoirs. This means that a pumped storage station would only be able to generate power for 9 hours within a 24 hour period and therefore cannot be utilised as a base load station (Jacobs, 2017). Therefore, this type of technology is generally employed for peak demand regulation.

3.4 UCEELD problem formulation

Power utilities nationally and internationally are faced with the problem of preparing a short term production schedule aimed at deciding which generating units to commit to the grid and at what output level, as to satisfy the power demand at the cheapest possible cost. This problem does not include maintenance planning as such planning forms part of the medium term planning process. The maintenance schedule does however serve as input to the short term production planning schedule. The formulation of this schedule consists of many complex variations which are technically challenging to solve. As a result of energy not being able to be stored in the power grid, it is required that the production schedule ensure an exact match between the power demand and production at each time period of the study horizon. For short term planning, it is also required to consider the hourly forecasted power demand as input when formulating the production schedule. The generating units considered in the production schedule are required to be modeled as being able to supply a certain amount of power, taking into consideration limiting factors such as mentioned throughout Section

3.1. Other factors that also need to be considered when producing such a schedule include, but are not limited to, ramp rate constraints, minimum on and off times required after start-up and shutdowns, operating reserve availability, prohibited operating regions, outage schedule constraints, and environmental legislation. Factors only isolated to hydro and pumped storage units include volumetric water flow and reservoir level management constraints. Considering the above mentioned parameters, each utility is required to generate a short term production schedule which will lead to the lowest amount of operational cost being incurred.

In the literature, different types of optimisation problems exist within the power generation contexts including, but not limited to, the economic load dispatch (ELD) (Sayah & Zehar, 2006), combined economic emissions load dispatch (CEELD) (Manteaw & Odero, 2012), multi-area economic dispatch (MAED) (Secui *et al.,* 1984), unit commitment (UC) (Yahya *et al.,* 2015), unit commitment and economic load dispatch (UCELD) (Santillan *et al.,* 2016), short term hydro scheduling (STHS) (Borghetti *et al.,* 2008), short term hydro-thermal scheduling (STHTS) (Salama *et al.,* 2013) and pumped storage scheduling (PSS) (Chen, 2008) problems. If however valve point effects are taken into consideration, the notation VP is added to the above stated acronyms (Ghasemi, 2013; Azzam *et al.,* 2014). These problems involve the optimisation of a power utilities' performance in order to reduce operational cost and improve resource utilisation while considering different aspects of the generation process. In this dissertation, the focus is on combining the dynamics of the above problems into a single model (excluding VP effects). This model will be defined as the unit commitment and environmental economic load dispatch optimisation problem (UCEELD). The UCEELD model will be applied to thermal, hydro and pumped storage generating units. The mathematical formulation will be discussed in more detail throughout Chapter 4. Each of the mentioned problems encompasses different dynamics and therefore various solution methods can be applied in solving these problems. Refer to Section 3.5 for a detailed discussion of the different solution methods used in literature to solve the listed problems.

3.5 Review of power generation optimisation solution methods

The most rudimentary form of the problem under consideration is the ELD optimisation problem. In this problem, the primary objective is to minimise the operational cost of thermal power stations by means of analyzing the fuel cost associated with each individual coal fired unit and dispatching the units accordingly. In recent years, a multitude of optimisation methods was developed in an attempt to solve the ELD problem. The two main categories within which these methods can be divided are exact and heuristic methods. The exact methods include integer programming (Dillon *et al.,* 1978), lambda iteration (Suman *et al.,* 2016), direct search methods (Chen, 2007), dynamic programming (Lowery, 1996) and quadratic programming (Papageorgiou & Fraga, 2007). Heuristic methods on the other hand include algorithms such as the particle swarm optimisation (Gaing, 2003), Hopfield neural network (Park *et al.,* 1993), evolutionary programming (Jayabarathi *et al.,* 2005), genetic algorithm (Walters & Sheble, 1993), tabu search (Senthil & Manikanda, 2010), differential evolution (Ghasemi *et al.,* 2016), biogeography-based optimisation (Bhattacharya & Chattopadhyay, 2010) and cuckoo search algorithm (Basu & Chowdhury, 2013). The last few years, a lot of focus has also been drawn to emission reduction within thermal power stations using optimal load scheduling, as a result of other solutions requiring significant amount of time to execute. It is for this reason that in order to address both the economical and environmental aspects, the CEED problem was formulated. To solve the CEED problem, researches proposed different techniques. The first method implemented was to treat the economic aspect of the CEED problem as a single objective, with the emissions only being considered as a problem constraint. Another approach is to combine both the economic and emissions aspects into one objective function and thereafter using a user defined weighting parameter to define each aspect's importance (Jeddi & Vahidinasab, 2014). Solution methods applied to the CEED problem by researchers include the non-dominated sorting genetic algorithm (Abido, 2003), bacteria foraging algorithm (Panigrahi *et al.,* 2011), differential evolution algorithm (Lu *et al.,* 2011), hybrid neuro-fuzzy system (Chaturvedi *et al.,* 2008), evolutionary algorithm (Abido, 2006), virus

optimisation algorithm (Liang & Juarez, 2014), harmony search algorithm (Jeddi & Vahidinasab, 2014) and the fuzzy clustering-based particle swarm algorithm (Agrawal, 2008).

Depending on the power grid design, utilities can either be connected to single or multi-area grids. In single-area grids, one power utility is connected to the grid to meet the power demand. This is typically considered during the formulation of the ELD and CEED problems. In multi-area grids, different power utilities operate interconnected from one another within a competitive environment. A multi-area grid is designed to improve grid reliability, stability, reserve sharing and reduce cost production. Within such a grid, certain utilities can be excluded from the power generation process, if for example the utilities' electricity production price is too expensive or a surplus of supply exist. The competitiveness of a multi-area grid promotes operational efficiency improvements and innovation between independent power utilities. Multi-area problems are increasingly more complex in comparison to single-area problems, due to the multitude of additional decision making tasks that a model needs to perform in order to optimise the operational cost of the power grid. Within literature, the multi-area power grid optimisation problem is abbreviated as MAED (Secui *et al.,* 2015). This problem still utilises the fuel cost functions of the individual coal fired units as objective functions and attempts to minimise the operational cost, while considering multiple grid areas. Algorithms proposed to solve the MAED problems include the Dantzig-Wolf decomposition approach (Romano *et al.,* 1981), linear programming (Desell *et al.,* 1984), evolutionary programming using tie-line constraints (Jayabarathi *et al.,* 2000), direct search method (Chen & Chen, 2001), neural networks using new nonlinear optimisation methods, differential evolution with time varying evolution, improved particle swarm optimisation (Zhu, 2003) and artificial bee colony optimisation (Secui *et al.,* 2015).

When considering the ELD, CEED and MAED problems, an assumption is made that the utilities' generating units are already committed to the power grid, and that the load at which the units are required to be dispatched, is the only uncertainty which needs to be optimised. However, in reality, the UC problem forms the initial phase of the production schedule development process and needs to be executed prior to solving the mentioned problems. The results obtained from the UC model, will provide the user with an indication of which units to commit to the power grid. The preceding problem is also coupled with the ELD problem in order to address both the commitment and dispatch of generating units simultaneously and is known as the UCELD problem. The aim of the UC and UCELD problems is to reduce the accumulated start-up, shutdown and operational cost of a power utility by managing the commitment (on/off selection) of the units accordingly. General algorithms applied by researchers to solve the UC and UCELD problems include exact methods such as the branch-and-bound, dynamic programming, Newton's method, Lagrangian relaxation method, lambda iteration method and mixed integer programming (Afzal & Madhav, 2017) whereas heuristic methods include simulated annealing, artificial neural networks, tabu search, particle swarm optimisation (Hadji *et al.,* 2015), chaotic ant swarm algorithm (Cai *et al.,* 2010), bacterial foraging algorithm, evolutionary programming, waterdrop algorithms and bio-geography based optimisation (Afzal & Madhav, 2017).

Although the preceding models address majority of the challenges faced by power utilities when developing a short term production schedule, they do not take into consideration the commitment and dispatch of peaking stations such as hydro and pumped storage generating units. The ELD, CEED, MAED and UCELD problems only focus on coal fired generating units and the optimisation of its operational activities. The management process of hydro station scheduling is defined as the STHS problem. When interconnected with the coal fired power stations, the scheduling process is redefined as the STHTS problem. The objective of these two problems is mainly to critically evaluate the generation schedule obtained from the department of water affairs and dispatch the available hydro units in such a manner as to minimise the operational cost incurred by the power utility. To solve the hydro scheduling problems, researchers have applied algorithms that include classical methods such as linear programming (Chang & Waight, 1999), maximum principle, network flow

programming (Sjelvgen *et al.,* 1983), dynamic programming (Yang & Chen, 1989), decomposition (Habibollahzadeh & Bubenko, 1986) and nonlinear programming (Ikura & Gross, 1988). Artificial intelligence concepts such as differential evolution (Jayabarathi *et al.,* 2007), simulated annealing (Wong & Wong, 1994; Basu, 2015), genetic algorithms (Gil *et al.,* 2003), particle swarm optimisation (Sun & Lu, 2010) and neural networks (Basu, 2003) were also applied in solving the SHTS and STHTS problems. The problem of optimising the commitment and dispatch of pumped storage power stations, is known as the PSS problem. In this problem, the objective is to maintain the required amount of energy storage in the upper reservoirs of the pumped storage station. This is enforced by means of switching the pumped storage units between generating and pumping modes. The reason for setting the energy management of the upper reservoir as objective, is to ensure that enough energy reserves are present during peak demands as to allow the units to be available for generation dispatch. Similar solution methods as was applied to the hydro scheduling problem are applied by researchers in solving the PSS problem (Chen, 2008).

In analyzing the above solution methodologies applied to the various power generation optimisation problems, it is apparent that heuristic, also called artificial intelligence methods have captured the interest of researchers. The reason being is because heuristic methods can be utilised in solving nonlinear, nonconvex, combinatorial and large scale problems within a reduced time period when compared to exact methods. However, when applying heuristic methods a compromise needs to be made with regards to solution accuracy as these methods are generally only able to obtained local optimum values. These methods are usually also quite complex and they behave stochastically. By implementing exact methods to solve power generation optimisation problems, an increase in the solution time needs to be expected as the model will attempt to find a definite global optimum solution (Ashfaq & Khan; 2014). Depending on the criticality of the problem and accuracy required, either heuristic or exact methods can be applied. As stipulated in Chapter 1, it is critical that a global optimal solution is obtained when solving the UCEELD problem as less of an accurate solution will result in an undesirable increase in operational cost. It is for this reason that MILP programming will be implemented in this dissertation in an attempt to solve the UCEELD problem.

3.6 Application of MILP to power grid optimisation

Prior to applying MILP to the nonlinear UCEELD problem, linearisation principles as discussed in Section 2.2.2 needs to be used to linearise the nonlinear functions. These functions include the fuel cost functions associated with the coal fired units and the water consumption functions related to the pumped storage units. For the purpose of this dissertation, the piecewise linear approximation methodology using binary variables (Section 2.2.2 part 2) will be used to linearise the mentioned nonlinear functions. The preceding method will be used in an attempt to reduce computational time by means of preventing the optimisation model to generate its own quantity of data points using the interpolation functionality. The reason this method is followed, is because of the UCEELD problem being of type NP-hard. This means that with an increase in the number of data points from which the model is able to make decisions, the computational time and complexity will increase exponentially. By applying the piecewise linear approximation methodology mentioned in Section 2.2.2 part 2, the user is able to dictate the interval sizes within which the nonlinear functions are divided and consequently the number of data points available for decision making. Dividing the nonlinear functions into large sized intervals, will assist the MILP model to be able to obtain a global optimal solution, within a reasonable amount of time. Although the linearisation methodologies as discussed in Sections 2.2.2 part 1 and part 3 might provide more accurate results, the model computational time will definitely increase. Provided that the study horizon considered in the UCEELD problem is a twenty four hour period, the model will need to be able to solve within a period of less than twenty four hours. Therefore, the decision was taken to try and keep the possible amount of scenarios at a minimum by implementing the methodology described in Section 2.2.2 part 2. In the subsequent chapter, the mathematical formulation of the UCEELD problem in MILP format is provided with a detailed analysis of the objective function and its associated constraints.

3.7 Model formulation and experimental scope

In the sections to follow the MILP UCEELD model will be developed. The objectives of the model will entail the reduction of the power utilities' capital expenditure and optimising its resource utilisation. The model will take into consideration thermal, hydro and pumped storage generating technologies and will include the following constraints:

1. Power balance, operating reserve availability, unit commitment considering up and downtimes, ramp rate capabilities, prohibited operating regions, outage schedule consideration, environmental aspects, interconnected multi-area power flow between neighboring areas.

2. Water balance for both hydro and pumped storage units

3. Volumetric flow rate range selection for pumped storage units

4. Power demand, Fuel consumption, and Emissions production stochasticity

In order to evaluate the performance of the MILP UCEELD model after development, the following experiments will be conducted:

1. Model verification tests. The preceding tests will consist of 12 problem instances which will be used the evaluate the model's response when only certain mathematical constraints are considered. The verification process is initially performed using a base model consisting of 3 thermal units. The base model is then augmented to a 6 thermal, 1 hydro and 2 pumped storage unit case study.

2. Model validation tests. The aforementioned tests will consist of 2 problem instances. The initial problem instance is concerned with evaluating the model's performance by comparing the model results with the performance of an actual thermal power station. For the purpose of the initial instance, only 6 thermal units will be utilised. The second problem instance will be focused on determining the capability of the MILP UCEELD model in solving a realistically sized power utility optimisation problem. The realistically sized model will consist of 98 thermal, 8 hydro and 6 pumped storage generating units. In conjunction with the preceding, an analysis is also conducted to determine the scalability of the MILP UCEELD model and the influence the scalability has on model solution time.

Chapter 4

Mathematical Model Formulation

4.1 General notation

In the subsequent chapter the mathematical formulation of the UCEELD problem, as discussed in Chapter 3, is presented with a detailed interpretation on the dynamics of the objective function and its associated constraints. However, prior to the formulation of the model some general notations are provided as introduction.

In order to account for the number of geographical areas incorporated into the UCEELD model, the index set $\mathcal{N}$ is defined with $n \in \mathcal{N}$ referring to a specific geographical area. The graph $G(\mathcal{N}, \mathcal{A})$, with arc set $\mathcal{A}$ is used to accommodate the modelling of power flow from geographical areas $i \in \mathcal{N}$ to $j \in \mathcal{N}$, with $(i, j) \in \mathcal{A}$.

Let $\mathcal{H}$ denote the index set of thermal and hydro units considered in the UCEELD problem, with $h \in \mathcal{H}$ referring to a specific unit within the collection. The subset $\mathcal{H}(n) \subseteq \mathcal{H}$ denotes the index set of all thermal and hydro units belonging to the geographical area $n \in \mathcal{N}$.

The collection of pump storage stations considered in the model is denoted by the index set $\mathcal{S}$. The subset $\mathcal{S}(n) \subseteq \mathcal{S}$ denotes the index set of all pump storage stations belonging the the geographical area $n \in \mathcal{N}$. The reservoir volume for each pump station $s \in \mathcal{S}$ is given by the input parameter v with each reservoir volume being defined by index $v \in \mathcal{V}$. The set of pump storage units installed at each pumped storage station $s \in \mathcal{S}$ is defined by the index set $\mathcal{U}(s)$.

The study horizon within which the generating units need to satisfy the forecasted grid demand is denoted by the index set $\mathcal{T}$. Each time period $t \in \mathcal{T}$ represents one hour of the twenty four hour study horizon.

The proposed optimisation problem that will address all the modeling requirements of the UCEELD problem, is a nonlinear programming problem. In order to incorporate logical decision-making in the model, a linearization approach is followed. More specifically, to be able to solve the UCEELD within an exact framework, e.g. the branch-and-bound approach, the resulting optimisation problem is required to be a mixed integer linear programming problem. The approach followed in this chapter to linearize the objective function and constraints involves the creation of a set of discretised function points (see Section 2.2.2 part 2). The index set $\mathcal{J}$ is used to represent these discretised function points.

In order to capture the variability observed in many of the input parameters of the UCEELD, for instance, grid demand, fuel costs, and emissions, a scenario based approach is followed. More specifically, the index set $\mathcal{P} = \{1, 2, \ldots, |\mathcal{P}|\}$ represents the possible future realisations and each stochastic input parameter will be indexed by a scenario $p \in \mathcal{P}$.

4.2 Model objective

The mathematical formulation of the UCEELD problem's objective function is provided in (4.1). The objective is to minimise the operational cost of the generating units, available to be committed, to meet the grid demand. The different types of power generating units considered in the optimisation model include coal fired, hydro and pumped storage units. Note that although various generating units are considered in this model, no operational cost is assigned to the hydro and pumped storage units. The reason for this is because these units utilise the potential and kinetic energy contained within the water reservoirs and or dams installed upstream to generate electricity. As a result, no fuel is consumed for power generation by these type of generating units. There are however fixed maintenance costs which can be assigned to each of these units but this specific cost is not taken into consideration for the purpose of this study. Analyzing the above mentioned it is apparent that only the coal fired units are assigned operational costs within the optimisation model. The operational cost of the coal fired units are however threefold and include coal consumption, water utilisation and start-up/shutdown costs (Fossati, 2012; Jeddi & Vahidinasab, 2014). Note however that throughout literature, the water usage cost has not yet been included in the objective function for coal fired stations, but will be added to the dissertation as a contribution.

In (4.1) auxiliary variables F and W, which will be discussed in more detail below, are representative of the coal consumption and water utilisation costs associated with the operation of the coal fired units. The start-up cost of a coal fired unit $h \in \mathcal{H}$ is given by $c_h^{(u)}$, whereas the shutdown cost is given by $c_h^{(d)}$. Binary decision variables w_{th} and x_{th} are introduced to keep track of the on and off selection of each coal fired unit. If variable w_{th} is assigned a value of 1, the unit is committed and the start-up cost will be considered in the objective function. During unit commitment, variable x_{th} assumes a value of 0 which prevents the shutdown cost to be considered. The opposite is true when the unit is decommitted, then w_{th} will take on a value of 0 and x_{th} will take on a value of 1.

$$\min \quad (F + W + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{h \in \mathcal{H}(n)} (c_h^{(u)} w_{th} + c_h^{(d)} x_{th})) \tag{4.1}$$

The cost associated with the station's coal consumption is determined by two factors which include the quality of coal obtained from the mines (dependent on the calorific value (CV) and ash content) and the thermal efficiency at which each unit is operating. A Unit with a low thermal efficiency will consume more coal for a given megawatt generated in comparison to a unit with a high efficiency, which will result in a higher operating cost being incurred for that specific unit. By incorporating the coal cost of each coal fired unit, the model will attempt to schedule the units with the lowest coal consumption/ coal cost while satisfying the remaining constraints. This cost is calculated by implementing equation (4.2).

The value $c_{jh}^{(f)}$ in (4.2) is the fuel cost associated with a discretised function point $j \in \mathcal{J}$, for a coal fired unit $h \in \mathcal{H}$. As mentioned in Section 3.4, the linearisation methodology as discussed in Section 2.2.2 part 2 has been applied to derive the discretised fuel cost data from each coal fired unit's nonlinear cost curve. The fuel cost is directly dependent on the load a unit is required to operate at. The binary decision variable z_{pthj} is used in the model to select the load at which each coal fired and hydro unit is required to be committed, if available for selection. If z_{pthj} is assigned a value of 1, unit $h \in \mathcal{H}$ is committed for service during time period $t \in \mathcal{T}$, at a load corresponding to the discretised level of $j \in \mathcal{J}$. This in turn will also determine the coal consumption of the given unit. The remaining load ranges for the unit will be assigned a value of 0 as the unit can only be assigned a single load point during each time period. The input parameters r_p, $q_{hp}^{(f)}$ and $q_{hp}^{(e)}$ are considered in the model to induce a stochastic influence to the optimisation problem. Parameter $q_{hp}^{(f)}$

denotes the expected variability in fuel consumption whereas $q_{hp}^{(e)}$ relates to the possible variability in emissions production for coal fired units. Parameter r_p represents the expected power demand variability.

$$F = \sum_{p\in\mathcal{P}} \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} \sum_{h\in\mathcal{H}(n)} \sum_{j\in\mathcal{J}} r_p q_{ph}^{(f)} c_{hj}^{(f)} z_{pthj} \tag{4.2}$$

To enforce a single load point selection in (4.2), constraint (4.3) is added to the model. The variable z_{pthj} is summated in terms of index j to ensure that only one load point may be selected per unit from the set of discretised load points provided to the model via the user input. The binary decision variable y_{th} is introduced to indicate whether a unit $h \in \mathcal{H}$ is in operation during time period $t \in \mathcal{T}$, or not.

$$\sum_{j\in\mathcal{J}} z_{pthj} = y_{th}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.3}$$

The second cost to consider is demineralised water consumption. To produce demineralised water, chemical cleaning processes are required which is used to extract the unwanted ions/ impurities from the water. If one unit consumes more demineralised water in comparison to another unit due to plant defects, the total operating cost for the specific unit will increase. By implementing this cost in the objective function, the model will attempt to reduce the total water consumed by the coal fired units via the unit scheduling process. The water cost for each unit is calculated by (4.4). The input value $c_h^{(w)}$ in this equation denotes the rand value associated with the water consumed by a coal fired unit $h \in \mathcal{H}$. Thus, if a unit is on load and operational, variable y_{th} will be assigned a value of 1. The opposite is true when the unit is off. By multiplying the value $c_h^{(w)}$ with the decision variable y_{th}, water cost will only be considered in the objective function when a coal fired unit is committed to the grid.

$$W = \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} \sum_{h\in\mathcal{H}(n)} c_h^{(w)} y_{th} \tag{4.4}$$

The third cost to consider for coal fired units, is the cost associated with the start-up and shutdown of the units. During start-up and shutdown fuel oil is utilised to provide flame ignition and or combustion support until pulverised fuel is either introduced into or taken out of the boiler. Depending on the type of fuel oil and quantity consumed, each unit will have a cost linked to its start-up and shutdown operation. Incorporating the above mentioned to the objective function will govern the extent to which the model will allow operation changeover between the units.

The fact that there are no costs assigned to the hydro and pumped storage units as mentioned previously, means that these units are considerably cheaper in comparison to the coal fired units and the optimisation model will want to commit these units continuously to satisfy the grid demand. To prevent the preceding from occurring, constraints are utilised. For pumped storage units, the power dispatch is governed by the available power contained in the upper and lower reservoirs. This is determined by the level in each reservoir. Providing a level setpoint to the upper reservoir, at a specific time in the study horizon, the pumped storage units will switch between generating and pumping modes to satisfy the given setpoint and prevent continuous unit commitment. For hydro units, the power dispatch is governed by a dispatch schedule obtained from the department of water affairs. This schedule considers the water availability in the dams as well as the effect the water release will have on the downstream ecosystem (to prevent floods). Therefore, the model allows the hydro and pumped storage units only to be scheduled during a fixed timeframe, which is usually during peak demand periods when considering the South African context, to prevent continuous unit commitment. The preceding constraints will be discussed in detail in the sections to follow.

4.3 Satisfying the power grid demand

The main objective of the optimisation model, apart from the financial objective of cost minimisation, is to meet the forecasted load demand of the power grid. The preceding is implemented in the model as an inequality constraint, as seen in (4.5). This constraint is known as the power balance. In the mentioned constraint, the difference is taken between the forecasted grid demand and the power generated by coal fired, hydro and pumped storage units of the resident area under consideration as well as the power imported from and or exported to neighboring areas. Effectively it is an energy balance of the entire power grid, incorporated into the model to ensure that the power supply will ultimately equal the power demand. Term $P^{(T)}$ in (4.5) denotes the power supplied by both coal fired and hydro generating units whereas the second term, $P^{(F)}$ denotes the power flow between geographical areas. The third term contained within (4.5), refers to the power generated or consumed by pumped storage power stations and is represented by variable $p^{(p)}_{ptsu}$. Lastly, term $p^{(d)}_{ptn}$ in (4.5) refers to the grid demand for both resident and neighboring countries. The grid demand is a forcasted data set which is used as an input to the optimisation model. This data set is the driving force for both unit commitment and load dispatch decision making. The optimisation model will commit a number of units at a given load to satisfy the grid demand. Chen (2008), Ni & Guan (1999) and Jadoun (2015) provides information regarding the power balance constraints mentioned in literature.

$$0 \leq P^{(T)} + P^{(F)} + \sum_{s \in \mathcal{S}(n)} \sum_{u \in \mathcal{U}(s)} p^{(p)}_{ptsu} - p^{(d)}_{ptn} \leq 60.0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{4.5}$$

To calculate the value of $P^{(T)}$, (4.6) is added to the model. In this equation, input parameter $p^{(t)}_{hj}$ refers to the discretised operating envelope (load range) at which each coal fired and hydro generating unit is able to operate within. Parameter $p^{(t)}_{hj}$ is multiplied with the binary decision variable z_{pthj} to select an operating point to which the given unit needs to be loaded provided the discretised operating envelope ($j \in \mathcal{J}$) and the remaining operating constraints considered in the model. By summating the multiplication function in terms of index $h \in \mathcal{H}(n)$ and $j \in \mathcal{J}$ the total power generated by the committed coal fired and hydro units are obtained for the resident area.

$$P^{(T)} = \sum_{h \in \mathcal{H}(n)} \sum_{j \in \mathcal{J}} p^{(t)}_{hj} z_{pthj} \tag{4.6}$$

The power flow between geographical areas as determined by $P^{(F)}$ can be either power imported from or alternatively exported to a neighboring country. The reason why multi-area power flow is incorporated into the optimisation model is because the probability exists that a neighboring country will be able to supply power to a resident country at a far cheaper rate, in comparison to what the given country will be able to generate it for. If this is the case, it will be more economically feasible to import a fraction of the power required to satisfy the grid demand instead of producing power by using more costly resources. For explanatory purposes we compare coal fired units with hydro and or pumped storage units. If a resident country is only dispatching coal fired units to meet the grid demand, it will result in a high cost being incurred by the resident country as coal fired power generation is rather expensive. If however there is a possibility to import hydro and or pumped storage generated power from a neighboring county to satisfy a third of the grid demand, while using coal fired units for the remaining two thirds, this will reduce the cost significantly. For this reason, the functionality of multi-area power flow is incorporated into the optimisation model. It is however important to note that the foregoing ratio of the amount of local power generation versus power importation at which an optimal solution will be obtained is dependent on various constraints incorporated into the optimisation model which needs to be satisfied and cannot be determined by only considering power generation cost.

Variable $P^{(F)}$ is divided into two mathematical terms. The first term computes the power imported from a neighboring country whereas term two accounts for the power exported by the resident country. Refer to (4.7) for the mathematical formulation. By taking the difference between the imported and exported power, $P^{(F)}$ will either take on a positive or negative value depending on the magnitude of each term. By summating $p^{(f)}_{pt(i,j)}$ in terms of index $(i,j) \in \mathcal{A}$ for import power and $p^{(f)}_{pt(j,i)}$ in terms of index $(j,i) \in \mathcal{A}$ for export power, the power flow per area can be calculated.

$$P^{(F)} = \sum_{(i,j)\in\mathcal{A}} p^{(f)}_{pt(i,j)} - \sum_{(j,i)\in\mathcal{A}} p^{(f)}_{pt(j,i)} \tag{4.7}$$

The power generated by pumped storage power stations ($p^{(p)}_{ptsu}$) is dependent on the mode of operation the pumped storage unit is selected to operate on. As mentioned, pumped storage units can either operate in generating mode or pumping mode. When in generating mode, the unit supplies power to the grid whereas in pumping mode the unit consumes energy from the grid. The detailed calculation of $p^{(p)}_{ptsu}$ is depicted throughout equations (4.8) and (4.9).

Equation (4.8) is effectively a small scale power balance of a pumped storage power station which is incorporated into the global power balance, as seen in (4.5). In constraint (4.8), variable $p^{(p)}_{ptsu}$ is defined as a rational number decision variable representing the power flow either from or to a pumped storage station. The optimisation model can assign either positive or negative values to this variable depending on the mode of operation. Positive values will be assigned to $p^{(p)}_{ptsu}$ when a unit is providing power to the grid. The contrary is true when a unit is consuming power. Variable $L^{(P)}$ in (4.8) refers to the power generated by a pumped storage unit when selected to generating mode with the detailed calculation of $L^{(P)}$ provided in (4.9). Similar to $p^{(t)}_{hj}$, input parameter $l^{(p)}_{suvj}$ refers to the discretised operating envelope (load range) at which each pumped storage generating unit is able to operate within when in generating mode. In order to select the active operating point in the identified envelope, $l^{(p)}_{suvj}$ is multiplied with binary decision variable f_{ptsuvj}. The purpose of variable f_{ptsuvj} is identical to variable z_{pthj}, but consist of more dimensions and is only applicable to pumped storage units. By summating the multiplication function in (4.9), in terms of index $v \in \mathcal{V}$ and $j \in \mathcal{J}$, the power generated by each active pumped storage unit is obtained. Note that in reality, the power generated by a pumped storage unit is dependent on the water head contained in the upper reservoir and may vary with water level fluctuations. In order to model this phenomenon, one can either make use of variable load ranges at a fixed volumetric flow rate or use a fixed load range at variable volumetric flow rates, given varying upper reservoir water volumes. An example is provided to explain the reasoning.

If the volume of an upper reservoir changes over time, the available power output that a pumped storage unit can supply will change given a fixed volume flow through the turbine. If this reasoning is followed, one can supply the model with a fixed volumetric flow range (i.e. $0m^3/h$ - $10m^3/h$) as user input and given different reservoir volumes (i.e. $100m^3$,$200m^3$ or $300m^3$) the model should select a power output from a selection of load ranges (i.e. (0 MW - 100 MW), (0 MW - 120 MW) or (0 MW - 140 MW)) associated to the upper reservoir volume and volumetric flow rate at time period *t*. On the other hand, if the volume of an upper reservoir changes over time, the available power output which a pumped storage unit can supply, can be maintained constant by varying the volumetric throughput through the turbine. By following this reasoning, one can supply the model with a fixed load range (i.e. 0 MW - 120 MW) as user input and given different reservoir volumes (i.e. $100m^3$,$200m^3$ or $300m^3$) the model should select a volumetric throughput from a selection of volumetric flow rates (i.e. ($0m^3/h$ - $10m^3/h$), ($0m^3/h$ - $8m^3/h$) or ($0m^3/h$ - $6m^3/h$)) associated to the upper reservoir volume and load operating point at time period *t*. Note that for the purpose of this study, the second reasoning was applied to optimise the operation of the pumped storage units. Variable LP_{nehvj} is supplied a fixed load range from where the model needs to analyze the upper reservoir volume and select the volume flow form a set list (Section 4.12.1) corresponding to the upper reservoir volume and load operating point. As the reservoir volume changes, the volumetric

flow rate through the turbine, selected by the model, will also change to satisfy the load operating point at which the unit is dispatched.

Input parameter $p_{su}^{(w)}$ depicts the power consumption when in pumping mode. Parameter $p_{su}^{(w)}$ is assigned a fixed power consumption per unit to reduce model complexity. Variable $p_{tsu}^{(r)}$ is a binary decision variable used to track the time a pumped storage unit is operational in pumping mode. By multiplying $p_{su}^{(w)}$ with $p_{tsu}^{(r)}$, it provides the model with the functionality to only consider pumped storage power consumption when in pumping mode. If $p_{tsu}^{(r)}$ is assigned a value of 0, the power consumption term will not be considered. A 0 value will be indicative of a pumped storage unit operating in generating mode. The contrary is true when $p_{tsu}^{(r)}$ is assigned a value of 1.

$$p_{ptsu}^{(p)} - L^{(P)} + p_{su}^{(w)} p_{tsu}^{(r)} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.8}$$

$$L^{(P)} = \sum_{v \in \mathcal{V}} \sum_{j \in \mathcal{J}} l_{suvj}^{(p)} f_{ptsuvj} \tag{4.9}$$

It is imperative to note that in practice, the power generated must always equal the grid demand. If the power generation falls short of or exceeds the grid demand, it can result in grid instabilities and consequently a grid collapse. Although the preceding is a prerequisite, an error term/ allowable error range is added to the power balance constraint to improve solving time. This is incorporated via an inequality constraint. For explanatory purposes, the error range is defined as $0 \leq error \leq 60$ in (4.5). The error range refers to the allowable deviation at which the power generated may deviate from the power demand. In the model formulation, the power generated is allowed to exceed the power demand with the magnitude provided in the error range. This error range is applied to the power balance due to the discretisation methodology implemented in the optimisation model. By only allowing the model to select unit load ranges from a predefined discretised list, it reduces the possibility of an equality constraint from being satisfied as the increments into which the data is divided might be too broad, and do not allow for an exact solution. These increments can however be refined, but by providing more data inputs to the model, the complexity as well as solution time will most definitely increase accordingly. For this reason, an inequality constraint is implemented into the model with an error term. The size of the allowable error range is however dependent on the magnitude of the problem. The addition of the error term does not affect the credibility of the model results, but does however only provide an approximation as to what an optimal commitment and load dispatch schedule will be, within a reasonable time-frame. The model will always schedule the available units to either satisfy or exceed the grid demand. For this reason, the results obtained for each unit load will need to be normalized with respect to the power demand, by the system operator, to ensure power supply is never exceeded and grid stability is maintained at all times. It is the prerogative of the system operator to either trade accuracy for solving time or vice versa depending on the allowable error range provided to the model. As mentioned, this error range may change depending on the problem complexity, magnitude and desired solving time and needs to be managed by the system operator to obtain the desired results.

4.4 Operating reserve availability

Apart from meeting the grid demand, the optimisation model needs to ensure that enough units and or power capacity is available to the system operator within a short period of time in case there is a disruption to the normal power supply. This is known as operating reserves. These reserves are crucial in ensuring that the day-ahead unit commitment and dispatch schedule will be able to withstand any unforeseen variation in load profiles or equipment failures. Operating reserves, also known as spinning reserves can be assigned to either one of the units considered in the optimisation model, given that the remaining constraints are still satisfied. When considering multi-area

power system capabilities, operating reserves for an area may include importing power from other systems or retracting power which was provided to a neighboring country in an attempt to maintain a stable grid frequency during unforeseen occurrences. This is however not explicitly modeled by an optimisation constraint, but is indirectly considered by means of (4.5). In Mahor *et al.* (2009), Zhou (2010) and Ghasemi *et al.* (2016) various approaches are depicted of the reserve availability constraints implemented in the literature.

To guarantee reserve availability, operating constraints are added to the model by means of (4.10) - (4.15). The amount of operating reserves required for each area (reserve demand) in the optimisation model is provided as a fraction of the total grid demand. The fraction is a user defined parameter and its magnitude is dependent on the extent of load variability or uncertainties expected for the planning horizon. Refer to (4.10) for the mathematical formulation of the operating reserve demand. In this equation, grid demand ($p^{(d)}_{ptn}$) is multiplied with the reserve demand fraction ($g^{(d)}$) to obtain the desired operating reserve demand ($s^{(d)}$) for each area. The total operating reserves available for each area must satisfy the reserve demand in order for the optimisation model to find a feasible solution.

$$s^{(d)} = p^{(d)}_{ptn}(g^{(d)}/100) \tag{4.10}$$

To calculate the operating reserves available to the system operator, (4.11) is incorporated to the model. The calculation entails the addition of the available operating reserves for each coal fired and hydro unit ($s^{(t)}_{pth}$ summated in terms of index $h \in \mathcal{H}(n)$), together with the available operating reserves for each pumped storage unit ($s^{(p)}_{ptsu}$ summated in terms of index $s \in \mathcal{S}(n)$ and $u \in \mathcal{U}(s)$). The methodology utilised to calculate $s^{(t)}_{pth}$ and $s^{(p)}_{ptsu}$ will be discussed in the subsequent sections. Note that reserve demand is calculated per area and not per Unit. For this reason the summation of all operating reserves available per unit, if the unit is selected to maintain a reserve capacity, is taken to ensure the reserve constraints per area are satisfied.

$$\sum_{h\in\mathcal{H}(n)} s^{(t)}_{pth} \quad + \sum_{s\in\mathcal{S}(n)} \sum_{u\in\mathcal{U}(s)} s^{(p)}_{ptsu} = s^{(d)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{4.11}$$

In order to calculate the available reserve for each coal fired and hydro unit, (4.12) is added to the model. In this equation, the difference is taken between the design power output of a unit ($d^{(t)}_h$), also know as the unit's maximum continuous rating (MCR), and the current power output of a unit ($p^{(t)}_{hj}$). The result of this calculation provides an indication as to the remaining capacity of a unit that is still available to be committed, if an unforeseen event occurs. Input parameter $d^{(t)}_h$ is multiplied with binary decision variable y_{th} to ensure that the equation is only in effect when the unit is committed. If the unit is operational, variable y_{th} will be assigned a value of 1. If however the unit is decommitted, this variable will take on a value of 0 and the constraint will not be considered. The same principle applies with regards to the multiplication of input parameter $p^{(t)}_{hj}$ with binary decision variable z_{pthj} which is discussed throughout Section 4.3 in detail. Variable $s^{(t)}_{pth}$ in (4.12) is a decision variable defined as a rational number which is only capped to positive values. The optimisation model can assign any positive value to this variable as long as the constraints are satisfied. If a unit is required to maintain a certain reserve quantity, a value greater than 0 will be allocated to variable $s^{(t)}_{pth}$ for the selected unit. Note that although all units will be able to maintain reserve capacities, only selected units will be assigned the task to meet the reserve demand. This selection will be primarily dependent on the unit's operating cost. Therefore some units will be allowed to be loaded to MCR operation as they are not selected by the optimisation model to maintain a reserve capacity (generally cheaper units). Other units will be loaded at lower load ranges because of the reserve requirements (more expensive units) and will not be allowed to operate at full load conditions. Loading units to MCR operation will result in variable $s^{(t)}_{pth}$ being assigned a 0 value as no reserve capacity will be required by the affected units.

$$s_{pth}^{(t)} \leq (d^{(t)} y_{th} - \sum_{j \in \mathcal{J}} p_{hj}^{(t)} z_{pthj}), \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.12}$$

Equations (4.13) is added to the optimisation model to prevent the operating reserve assigned to each unit, to exceed the unit's load ramp-up capability. Equation (4.13) is applicable to coal fired and hydro units only. Although a unit might be operating at minimum generation and still have ample capacity to contribute to the grid, the unit will only be able to provide power to the grid at the designed ramp-up capability, given an unforeseen event occurs. The design ramp-up capability of a thermal unit is incorporated in (4.13) by means of input paramter $r_h^{(u)}$. This variable is multiplied with binary decision variable y_{th} to prevent the equation from being considered when a unit is not operational. Equations (4.13) will only influence the model's decision making when the difference between a unit's design and operating power output is greater than its design ramp rate capability. If however, the difference in output is less than the ramp rate capability, (4.12) will govern the decision making for both coal fired and hydro units.

$$s_{pth}^{(t)} \leq r_h^{(u)} y_{th}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.13}$$

The same principle is applied in (4.14) to calculate the available reserves for each pumped storage unit as was used in (4.12). Input parameter $d_{su}^{(p)}$ refers to the design power output of a pumped storage unit and $p_{ptsu}^{(p)}$ to the operating output. The difference is taken between these two parameters to determine the remaining capacity of a pumped storage unit that is still availability to be committed, if necessary. Parameter $d_{su}^{(p)}$ is multiplied with binary decision variable $t_{tsu}^{(r)}$ in the mentioned equation to track unit operation, similar to (4.12). This is to prevent the constraint from being considered when a unit is decommitted from the grid. Variable $s_{ptsu}^{(p)}$ is a decision variable defined as a rational number which is capped to positive values. If a unit is required to maintain a certain reserve quantity, a value greater than 0 will be allocated to this variable. The preceding will be done until enough reserves are assigned to the different units to satisfy the reserve demand.

$$s_{ptsu}^{(p)} \leq (d_{su}^{(p)} t_{tsu}^{(r)} - p_{ptsu}^{(p)}), \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.14}$$

Lastly, constraint (4.15) is added to the optimisation model to prevent the operating reserve assigned to each pumped storage unit, to exceed the unit's load ramp-up capability. Input parameter $r_{su}^{(u)}$ used in (4.15) refers to the design ramp-up capability of a pumped storage unit. This parameter is multiplied with binary decision variable $t_{tsu}^{(r)}$ to prevent the constraint from being considered when a unit is not operational. Constraint (4.15) will only influence the model's decision making when the difference between a unit's design and operating power output is greater than its design ramp rate capability. If however the difference in output is less than the ramp rate capability, (4.14) will govern the decision making for the pumped storage units.

$$s_{ptsu}^{(p)} \leq r_{su}^{(u)} t_{tsu}^{(r)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.15}$$

4.5 Unit commitment considering up and downtimes

The unit commitment problem is a large mathematical optimisation problem which entails the co-ordination of power generating units to achieve a common target, while considering factors such as reliability, financial, regulatory and unit operating constraints. For the purpose of this study, the common target to be satisfied as mentioned in Section 4.2 is to reduce financial expenditure while still satisfying the power grid demand. In the unit commitment problem, the optimisation model must evaluate the grid demand via the power balance, and perform an on/off selection to ensure

enough units are committed to satisfy this demand. Only thereafter, the model can assign a load to which each unit need to be dispatched to. In accordance with unit scheduling, parameters such as minimum up and down time need to be considered.

The minimum up and downtime constraints are however primarily applicable to coal fired units as hydro and pumped storage units can be committed to the grid within a few minutes. When considering coal fired units, the duration required for a unit start-up or shutdown can range from anything between 1 hour to 24 hours depending on the operational condition of the unit and the problems encountered when executing these activities. During start-up there are certain operating parameters, such as temperatures and pressures, which needs to be satisfied as well as operational tests and checks which needs to be performed before grid synchronization is authorized. When a coal fired unit is desynchronized from the grid, time is required to systematically shutdown mechanical and C&I equipment in an attempt to safely remove fuel from the combustion process. By excluding the mentioned constraints from the model, it will reduce the complexity of the problem to be solved, but will most certainly provide an unrealistic unit commitment schedule which will not be practicable to apply to real life events. For this reason, the minimum up and downtime constraints have been added to the optimisation model, although it increases the model complexity significantly. The mathematical formulation for the unit commitment constraints associated with coal fired and hydro units as well as pumped storage units are depicted throughout (4.16) to (4.28). Note that (4.16) - (4.22) are applied to coal fired and hydro units whereas (4.23) - (4.28) are related to pumped storage units. The research done by Chang *et al.* (2001), Tseng *et al.* (2000), Tuffah & Gravdahl (2013), Ashan *et al.* (2018) and Borghetti *et al.* (2008) provide some insight into the formulation of the above mentioned constraints.

Equation (4.16) has been added to the optimisation model to track the operation of both coal fired and hydro units, to effectively provide a unit commitment schedule for each unit. As discussed in Section 4.2, w_{th} and x_{th} in (4.16) are binary decision variables used to monitor the on/ off selection of a unit whereas y_{th} was added to calculate the time a unit was operational, after receiving the on selection. Variable $y_{(t-1)h}$ in (4.16) still fulfill the same purpose as y_{th} but only references the operational state of a unit at time period t-1 instead of period t. If a unit is off at the start of the study horizon and only receives an on selection in period t, say for example t = 2, then variable y_{th} will be assigned a value of 1 and $y_{(t-1)h}$ will take on a 0 value for the period under consideration.

$$w_{th} - x_{th} = y_{th} - y_{(t-1)h} \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.16}$$

To explain the functionality of the constraint, an example is provided. If a unit is committed to the grid, variables w_{th} and y_{th} will be assigned a value of 1 whereas variables x_{th} and $y_{(t-1)h}$ will take on 0 values. The equality constraint will be satisfied with both sides of the equation equaling 1. The opposite is true when a unit is decommitted, variables w_{th} and y_{th} will be assigned 0 values whereas variables x_{th} and $y_{(t-1)h}$ will each take on a value of 1. As a result the equality constraint will be satisfied with both sides of (4.15) equaling a value of -1. Lastly, if a unit is not considered for selection during the study horizon, all variables in (4.15) will be assigned 0 values to prevent the unit from being committed.

Equation (4.17) is added to the model to prevent simultaneously on/off selection of a single generating unit in the same time period. If w_{th} is assigned a value of 1, the model will maintain x_{th} equal to 0. Only when w_{th} equals 0, thereafter x_{th} will be allowed to be assigned a value of 1. Note that (4.17) is defined as an inequality constraint as the possibility exist where both w_{th} and x_{th} equates to 0. It will occur when no selection is made regarding the operation of a generating unit. The unit is not required to switch on or off, and will remain in the same mode of operation as was selected in the previous time interval (either committed or decommited to the grid).

$$w_{th} + x_{th} \leq 1 \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.17}$$

To add the functionality of minimum up and downtime to the model, (4.18) - (4.21) are incorporated. Note however that this problem is twofold. The model first needs to evaluate the unit commitment schedule of the previous study horizon. Thereafter it needs to ensure that a unit is not allowed to switch operation in the current study horizon, if the minimum up or down-times have not yet been satisfied for a unit in the previous study horizon. The preceding is enforced by equations (4.18) and (4.20). Secondly, after the previous study horizon's constraints have been satisfied, the model must evaluate the current study horizon and ensure the minimum up and downtime constraints are adhered to in this time period. Equations (4.19) and (4.21) provide this functionality to the model. These constraints are however only applicable to coal fired units and will not be considered for hydro units due to the fast response time.

Input parameters $t_h^{(u)}$ and $t_h^{(u1)}$ in (4.18) denotes the time a unit was operational for the current as well as the previous study horizon. The difference between $t_h^{(u)}$ and $t_h^{(u1)}$ provides the model with the remaining time the unit is required to be operational. Parameter $t_h^{(u1)}$ is multiplied with variable y_{1h} to ensure that (4.18) is only considered when the unit is operational at the start of the study horizon. Equation (4.19) is implemented to enforce the minimum up-time of a generating unit in the remaining stages of the study horizon. In this equation $t_h^{(u)}$ is multiplied with w_{th} to ensure that the constraint is only in effect when a unit is operational. The time a unit is operational is tracked by summating variable y_{th} in terms of index $t \in \mathcal{T}$ for each unit. This time must exceed the predefined operational time, before an off selection will be allowed for the given unit.

$$\sum_{\substack{t\in\mathcal{T}\\ t\geq 1\\ t\leq(t_h^{(u)}-t_h^{(u1)})}} y_{th} \geq t_h^{(u)} - t_h^{(u1)} y_{1h}, \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.18}$$

$$\sum_{\substack{k\in\mathcal{T}\\ k\geq t\\ k\leq(t+t_h^{(u)}-1)}} y_{kh} \geq t_h^{(u)} w_{th}, \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.19}$$

Input parameters $t_h^{(d)}$ and $t_h^{(d1)}$ in (4.20) depicts the downtime of a unit in the current as well as in the previous study horizon. Taking the difference between $t_h^{(d)}$ and $t_h^{(d1)}$, gives an indication as to the remaining downtime required per unit before operation change over is allowed. Parameter $t_h^{(d1)}$ is multiplied with (1 - y_{1h}) to ensure (4.20) is excluded from the model, if the unit was operational at time period 1 of the study horizon. If however the unit was decommitted during hour 1, only then (4.20) will be considered. Similar to (4.18), this equation is only applicable to the initial stages of the study horizon. Equation (4.21) is added to the model to ensure that the minimum downtime of a generating unit is adhered to for the remaining stages of the study horizon. Parameter $t_h^{(d)}$ is multiplied with x_{th} to prevent the constraint from being considered when a unit is operational. By summating (1 - y_{th}) in terms of index $t \in \mathcal{T}$ for each unit, the model is able to track the downtime per unit. The time a unit is decommitted must exceed the predefined downtime, before an on selection will be permitted by the model.

$$\sum_{\substack{t\in\mathcal{T}\\ t\geq 1\\ t\leq(t_h^{(d)}-t_h^{(d1)})}} 1 - y_{th} \geq t_h^{(d)} - t_h^{(d1)}(1 - y_{1h}), \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.20}$$

$$\sum_{\substack{k\in\mathcal{T}\\ k\geq t\\ k\leq(t+t_h^{(d)}-1)}} 1 - y_{kh} \geq t_h^{(d)} x_{th}, \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.21}$$

Equation (4.22) is added to the model to provide the user with the functionality of specifying a predefined number of coal and or hydro units that are required to be operational during the study horizon. To calculate the number of units selected for operation, variable z_{pthj} is summated in terms of index $h \in \mathcal{H}(n)$ and $j \in \mathcal{J}$. The answer obtained from this calculation must either equal or exceed the value stated on the right hand side of the equation. For explanatory purposes, the right hand term is assigned a value of 2. This means that at least two units need to be committed to the grid in order to satisfy the equation. Note however that this equation is not compulsory for unit commitment, but provides the user with the ability to run test examples with different number of units committed to the grid. If for example on load maintenance or tests are required on various units for a given study horizon (i.e. air heater leakage tests, duct leakage inspections etc.) this equation can be used to force the model to commit the affected units. By doing this, it might result in a suboptimal solution being obtained. In order to exclude this constraint from the model, a value of 0 needs to be assigned to the right hand side term of (4.22).

$$\sum_{h \in \mathcal{H}(n)} \sum_{j \in \mathcal{J}} z_{pthj} \geq 2 \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{4.22}$$

When considering pumped storage technology, minimum up and downtime are not incorporated into the model due to its fast reaction time. It is however, necessary to incorporate scheduling constraints which will allow a pumped storage unit to switch between its different operating modes (either generating or pumping mode) in order to satisfy the water balance constraints. In order to provide the optimisation model with the functionality to schedule pumped storage units with regards to its generating and or pumping modes, (4.23) - (4.28) are added to the model. Equations (4.23) and (4.24), similar to (4.16), are used to track the on/off selection as well as the operational time of a pumped storage unit in either generating or pumping mode. Both equations are utilised to provide the system operator with a detailed commitment schedule, specifying selected operating modes and commitment durations assigned to each generating unit in the study horizon. This schedule is crucial when performing either day or week-ahead planning.

Variables $p_{tsu}^{(n)}$ and $t_{tsu}^{(n)}$ are binary decision variables incorporated into (4.23) and (4.24) to track the commitment (on selection) of a pumped storage unit in either pumping or generating mode. If for example, variable $p_{tsu}^{(n)}$ is assigned a value of 1, and $t_{tsu}^{(n)}$ a value of 0, it will be indicative of a pumped storage unit operating in pumping mode. The opposite is true when operating in generating mode. Variable $t_{tsu}^{(n)}$ will be assigned a value of 1 and $p_{tsu}^{(n)}$ a value of 0. Variables $p_{tsu}^{(f)}$ and $t_{tsu}^{(f)}$ are applied to track the decommitment (off selection) of a generating unit from either pumping or generating mode. If a pumped storage unit is decommitted from pumping mode, variable $p_{tsu}^{(f)}$ will be assigned a value of 1 and $t_{tsu}^{(f)}$ a value of 0. When decommitted from generating mode the contrary applies. Variables $p_{tsu}^{(r)}$ and $t_{tsu}^{(r)}$ are also binary decision variables added to (4.23) and (4.24) to track the time a unit was operational, either in pumping mode or in generating mode. Binary decision variables $p_{(t-1)su}^{(r)}$ and $t_{(t-1)su}^{(r)}$ fulfill the same purpose as $p_{tsu}^{(r)}$ and $t_{tsu}^{(r)}$, but only references the operational state of a unit in time period t-1 instead of period t. Depending on the mode of operation, either $p_{tsu}^{(r)}$ or $t_{tsu}^{(r)}$ can be assigned a value of 1 or 0.

$$p_{tsu}^{(r)} - p_{(t-1)su}^{(r)} - (p_{tsu}^{(n)} - p_{tsu}^{(f)}) = 0, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.23}$$

$$t_{tsu}^{(r)} - t_{(t-1)su}^{(r)} - (t_{tsu}^{(n)} - t_{tsu}^{(f)}) = 0, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.24}$$

The fact that the purpose of (4.23) and (4.24) are identical to (4.16), but only applied to different power generating technologies, no in-depth explanation will be provided as to the functionality of the mentioned equations and how the equation results needs to be interpreted. Refer to (4.16) for

a detailed explanation on the mechanics of these equations, when applied to different model instances.

To prevent the simultaneous on/off selection of a single pumped storage generating unit, (4.25) and (4.26) are added to the model. Equations (4.25) and (4.26) are related to pumping and generating modes respectively. The mathematical formulation of the mentioned equations are depicted below:

$$p_{tsu}^{(n)} + p_{tsu}^{(f)} \leq 1, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.25}$$

$$t_{tsu}^{(n)} + t_{tsu}^{(f)} \leq 1, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.26}$$

When variable $p_{tsu}^{(n)}$ is assigned a value of 1, pumping mode is initiated and the unit will transport water from the bottom reservoir to the upper reservoir. Variable $p_{tsu}^{(f)}$ will assume a value of 0 during the mentioned mode of operation as (4.25) will not allow both variables to take on values of 1. This is to prevent illogical scheduling as a unit cannot receive an on and off selection at the same time. The same principle applies to (4.26), but this equation governs the logical scheduling of a pumped storage unit when in generating mode. Note that both equations are defined as inequality constraints to allow the model to assign 0 values to the mentioned variables when no operation mode change over is initiated.

Another factor to consider when scheduling pumped storage units is the fact that a single unit is not allowed to pump and generate at the same time. To prevent simultaneous pumping/generation selection, (4.27) and (4.28) are incorporated into the model.

$$t_{tsu}^{(n)} + p_{tsu}^{(n)} \leq o_{su}^{(p)}, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.27}$$

$$t_{tsu}^{(r)} + p_{tsu}^{(r)} \leq o_{su}^{(p)}, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.28}$$

For explanatory purposes, an example is provided to explain the functionality of the above equations. If a pumped storage unit is selected to provide power to the grid (i.e. selected to generating mode) variables $t_{tsu}^{(n)}$ and $t_{tsu}^{(r)}$ will be assigned a value of 1. Both equations (4.27) and (4.28) will prevent variables $p_{tsu}^{(n)}$ and $p_{tsu}^{(r)}$ to be selected, and as a result, maintain these variables at 0 values until pumping mode is initiated by the model. At this stage, the pumping variables will take on values of 1 whereas the generating variables will switch to 0 values. Under no circumstance will these constraints allow the model to simultaneously activate both pumping and generating modes. Similar to (4.25) and (4.26), these equations are formulated as inequality constraints to allow the model to assign 0 values to these variables if no operation mode changeover is required for a given unit. Note that $o_{su}^{(p)}$ is a binary user defined input parameter which provides the model with information regarding the outage schedule for each pumped storage unit. If a unit is scheduled for an outage, parameter $o_{su}^{(p)}$ will be assigned a 0 value by the user. As a result, the left hand side of (4.27) and (4.28) will be forced to 0 and prevent scheduling of the affected unit.

In conjunction with the preceding two constraints, (4.29) is added to the model to prevent any unit from operating in generating mode whilst other units connected to the same upper reservoir are selected to operate in pumping mode. If any unit is selected to pumping mode (i.e. $p_{tsu}^{(r)} = 1$) variables $t_{tsu}^{(r)}$ and $t_{tsu}^{(n)}$ will be prohibited to take on a value of 1 until pumping has been completed. Variable M is a user defined input parameter that needs to take on a very large positive value to ensure that equation (4.29) holds true when a unit is operating in pumping mode.

$$\sum_{u \in \mathcal{U}(s)} L^{(P)} \leq M(1 - p_{tsu}^{(r)}), \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.29}$$

4.6 Ramp rate capability

When referring to ramp rate capability, it is indicative of the allowable maximum up or downward change in power output per time period (i.e. *MW/min*) of a unit. In other words, a unit will not be able to increase or decrease its power output at a rate which exceeds the maximum up or downward ramp limits. For application to the optimisation horizon, the ramp rate capability of each unit is provided in terms of *MW/h* via user input. Depending on the power generation technology under consideration, each ramp rate may differ. Also, when incorporating ramp rate capabilities into the optimisation model, both start-up/shutdown and on-load ramp rates must be considered. Equations (4.30) - (4.37) are added to the model to incorporate the mentioned functionality. Different formulations of the foregoing is presented by Borghetti *et al.* (2008), Hedman *et al.* (2009) and Espana *et al.* (2012) throughout their research.

Equation (4.30) governs the maximum allowable ramp-up capability of both coal fired and hydro units. In this equation, input parameter $p_{hj}^{(t)}$ is multiplied with binary decision variable z_{pthj} and summated in terms of index $j \in \mathcal{J}$ to select the optimal operating point of a unit for time period *t*. In conjunction with the preceding, $p_{hj}^{(t)}$ is multiplied with $z_{p(t-1)hj}$, and again summated in terms of index $j \in \mathcal{J}$ to determine the optimal load selection of a unit for time period *t-1*. Subtracting the two multiplication functions from one another as depicted in (4.30), provides the user with the rate (*MW/h*) at which a unit increased load from time period *t-1* to *t*. Variable $R^{(U)}$ is incorporated into (4.30), via user input, to prevent the load increase of a unit from exceeding its maximum allowable ramp capability. Depending on the operational state of a unit, either during start-up or normal operation, variable $R^{(U)}$ can be assigned different values. Refer to (4.31) for a detailed discussion regarding the calculation of $R^{(U)}$. By adding $R^{(U)}$, a unit is allowed to increase its load with either less than or equal to the quantity specified by the user. Note however that a unit will not be allowed to undergo a load step change which exceeds $R^{(U)}$. If a unit is not selected for operation, variables z_{pthj} and $z_{p(t-1)hj}$ will be assigned 0 values and prevent the constraint from being considered. The contrary is true when the unit is committed for operation.

$$(\sum_{j\in\mathcal{J}} p_{hj}^{(t)} z_{pthj} - \sum_{j\in\mathcal{J}} p_{hj}^{(t)} z_{p(t-1)hj}) \leq R^{(U)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.30}$$

Variable $R^{(U)}$ is divided into two terms in (4.31). The first term denotes the allowable maximum ramp-up capability of a unit during normal operation whereas term two portrays the ramp-up capability during unit start-up. Parameter $r_h^{(u)}$ is multiplied with binary decision variable $y_{(t-1)h}$ to ensure that the first term is only considered when the unit is committed for operation (i.e. when variable $y_{(t-1)h} = 1$). If a unit is decommitted, term one will be discounted from the calculation as $y_{(t-1)h}$ will equal 0. Term two will only come into affect when a unit is selected for start-up. Parameter $r_h^{(us)}$ is multiplied with w_{th} in term two. If selected for start-up, w_{th} will be assigned a value of 1 and consequently, $R^{(U)}$ will equal the value of $r_h^{(us)}$. Note however that $R^{(U)}$ cannot equal the sum of both terms in (4.31), as only one term can be in effect at a time. As stated, the active term will be dependent on the mode of operation of the unit.

$$R^{(U)} = r_h^{(u)} y_{(t-1)h} + r_h^{(us)} w_{th} \tag{4.31}$$

Following the same mathematical reasoning as depicted in (4.30), equation (4.32) is incorporated into the model to govern the maximum allowable ramp-down capability of both coal fired and hydro units. In (4.32) the operating load of a unit at period *t-1* is subtracted with the load at period *t* to obtain the rate (*MW/h*) at which a unit decreased its load during each hour. Variable $R^{(D)}$ denotes the maximum allowable ramp down capability of each unit and is supplied to the model via user input. Similar to $R^{(U)}$, $R^{(D)}$ also consist of two terms. Term one denotes the ramp down capability of a unit when operational ($r_h^{(d)}$), whereas term two is indicative of the shutdown ramp capability of

a unit when decommittment from the grid ($r_h^{(ds)}$). Term one will be active when a unit is operational (y_{th} = *1*, x_{th} = *0*) whereas term two will be considered in the optimisation model when a unit is selected for shutdown (y_{th} = *0*, x_{th} = *1*). Due to the similarities of (4.32) and (4.33) (ramp-up capability) with (4.30) and (4.31) (ramp down capability) respectively, we will not elucidate on the below equations. Refer to the above paragraph for a more detailed explanation on the mathematical formulation and mechanics of such equations.

$$(\sum_{j\in\mathcal{J}} p_{hj}^{(t)} z_{p(t-1)hj} - \sum_{j\in\mathcal{J}} p_{hj}^{(t)} z_{pthj}) \leq R^{(D)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, h\in\mathcal{H}(n) \tag{4.32}$$

$$R^{(D)} = r_h^{(d)} y_{th} + r_{ny}^{(ds)} x_{th} \tag{4.33}$$

When considering pumped storage ramp rate capabilities, the same principle applies as was mentioned during the discussion for coal fired and hydro units. Variable $p_{ptsu}^{(p)}$ denotes the power flow from a pumped storage unit in time period t whereas $p_{p(t-1)su}^{(p)}$ references the power flow in time period *t-1*. Subtracting the variables as noted in (4.34), the rate *MW/h* at which a pumped storage unit increased its load from time period *t-1* to t is obtained. Similar to variable $R^{(U)}$ in (4.30), $R^{(UP)}$ is incorporated to limit the ramp-up rate of a pumped storage unit below the maximum allowable rate.

$$(p_{ptsu}^{(p)} - p_{p(t-1)su}^{(p)}) \leq R^{(UP)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n), u\in\mathcal{U}(s), v\in\mathcal{V} \tag{4.34}$$

The calculation of variable $R^{(UP)}$ is divided into two sections identical to (4.31) and for this reason we will not elucidate on the below equations. Parameter $r_{su}^{(u)}$ depicts the ramp up rate under normal operating conditions whereas $r_{su}^{(us)}$ denotes the allowable start-up ramp rate of each pumped storage unit.

$$R^{(UP)} = r_{su}^{(u)} t_{(t-1)su}^{(r)} + r_{su}^{(us)} t_{tsu}^{(n)} \tag{4.35}$$

Equations (4.36) and (4.37) are incorporated into the optimisation model to account for the ramp down capability of pumped storage units. The mentioned equations use the same mathematical reasoning as mentioned in (4.32) and (4.33). In (4.36), variable $R^{(DP)}$ signifies the maximum ramp down capability of a pumped storage unit. The foregoing is calculated by applying (4.37). In this equation, parameters $r_{su}^{(d)}$ and $r_{su}^{(ds)}$ refers to the ramp rate capability of a pumped storage unit under normal operating and shutdown conditions respectively. Due to the similarities of these equations with that explained for coal fired and hydro units, no further exposition will be provided.

$$(p_{p(t-1)su}^{(p)} - p_{ptsu}^{p)}) \leq R^{(DP)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n), u\in\mathcal{U}(s), v\in\mathcal{V} \tag{4.36}$$

$$R^{(DP)} = r_{su}^{(d)} t_{(t-1)su}^{(r)} + r_{su}^{ds)} t_{tsu}^{(f)} \tag{4.37}$$

By incorporating the ramp rate capabilities into the optimisation model, it increases the problem complexity significantly. The reason for this is the fact that the model needs to compare the results of different time periods in order to satisfy the constraints. As a result, the solution time increases. However failing to include this into the model, the results obtained from the model will allow units to operate outside its design base. Consequently, the system operator will not be able to use these results for real time power dispatch.

4.7 Prohibited operating regions

Generating units may have certain regions where operation is either impossible or undesired due to either process instabilities, physical limitations of mechanical components or arranged testing which needs to be conducted at predefined load ranges. Refer to Section 3.1.10 for examples of the mentioned limitations. These prohibited operating regions produce discontinuities in the load curves, of the affected units, since the units are forced to operate over or under certain specified limits to maintain safe and or stable operation. In order to account for a unit's prohibited operating ranges (4.38) (applicable to coal fired and hydro units) and (4.39) (applicable to pumped storage units) are incorporated into the model. Researchers such as Jeddi & Vahidinasab (2014), Mandal *et al.* (2015), Pan *et al.* (2017) also proposed different methodologies as to how prohibited operating regions can be incorporated into the power generation optimisation models. Note however that in the UCEELD model as developed in this dissertation, only the logical modeling of the above researchers was followed with the constraints being restructured to fit into the MILP framework.

To prevent scheduling of a coal fired or hydro unit at a prohibited operating point, the summation of binary decision variable z_{pthj} in terms of index $j \in \mathcal{J}$, is set equal to a value of 0 where index j is constrained between ranges $pr_h^{(l)}$ and $pr_h^{(h)}$. The span of the range is dependent on user input. Input parameters $pr_h^{(l)}$ and $pr_h^{(h)}$ denotes the lower and upper ranges at which a unit is forbidden to operate. For explanatory purposes, an example is provided. Say a unit's operating envelop (0 MW - 600 MW) is divided into 15 data points (40 MW/point). If parameters $pr_h^{(l)}$ and $pr_h^{(h)}$ is set equal to 4 (160 MW) and 8 (320 MW) respectively, it will prevent a unit from being dispatched at a load which falls within the range of data points between 4 (160 MW) - 8 (320 MW), as variable z_{pthj} is forced to a value of 0. For this example, a unit will however still be allowed to be dispatched within the load range 1 (40 MW) - 3 (120 MW) and 9 (360 MW) - 15 (600 MW) if committed to the grid.

$$\sum_{\substack{j\in\mathcal{J}\\ j\geq pr_h^{(l)}\\ j\leq pr_h^{(h)}}} Z_{rptnyj} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.38}$$

The same principle is applied to prevent scheduling of pumped storage units at a prohibited operating point. Binary decision variable f_{ptsuvj} is summated in terms of index $j \in \mathcal{J}$ and set equal to 0 where j ranges between $pr_{su}^{(l)}$ and $pr_{su}^{(h)}$. Parameters $pr_{su}^{(l)}$ and $pr_{su}^{(h)}$ depicts the lower and upper ranges at which a pumped storage unit is prohibited to operate.

$$\sum_{\substack{j\in\mathcal{J}\\ j\geq pr_{su}^{(l)}\\ j\leq pr_{su}^{(h)}}} f_{ptsuvj} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s), v \in \mathcal{V} \tag{4.39}$$

If no prohibited operating regions exist for either coal fired, hydro or pumped storage units the functionality can be eliminated. This is achieved by setting input parameters $pr_h^{(l)}$ and $pr_h^{(h)}$ equal to 0 for coal fired and hydro units whereas parameters $pr_{su}^{(l)}$ and $pr_{su}^{(h)}$ can be set equal to 0 for pumped storage units. By implementing the foregoing, (4.38) and (4.39) will be discounted from the model and the entire operating envelope of the affected unit will be considered for optimisation.

4.8 Outage schedule considerations

Every system operator, at some point, is responsible to take a generating unit offline temporarily for either maintenance or system upgrades. The time a unit is down, is known as an outage opportunity. The duration of an outage is dependent on the magnitude of the work that needs to be performed, but can range from anything between 12 to 120 days. When considering coal fired, hydro

and pumped storage generating technologies, it is important to note that a combination of these units can either be scheduled for an outage during the same time period, or during time periods offset with a few weeks/months from one another. This is governed by an outage schedule which ensures that enough units are available at all times to meet the grid demand. Throughout literature, researchers do not include outage schedule constraint functionalities into the commitment and dispatch optimisation problems as they assume that only the available units will be considered when performing the optimisation computations. However, in this dissertation the outage schedule will be considered in order to improve the versatility of the model.

The process of determining the optimal outage schedule is, however, a separate optimisation problem, which will not be addressed throughout the scope of this project. The assumption is made that an optimal outage schedule has already been obtained from the system operator. In order to enforce unit downtime via the existing outage schedule, Equation 4.40 is incorporated into the model by means of an inequality constraint. Variable y_{th} is set equal to or less than binary input parameter $o_h^{(t)}$. Parameter $o_h^{(t)}$ refers to the optimal outage schedule as obtained from the system operator. If $o_h^{(t)}$ is set equal to 1 by the user, this will be indicative of a unit that is available for commitment. Variable y_{th} will be allowed to take on a value of either 0 or 1 depending on the decision making of the optimisation model. If however a unit is scheduled for an outage, $o_h^{(t)}$ will be set equal to 0 by the user and the unit will be discounted from the optimisation process.

$$y_{th} \leq o_h^{(t)} \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.40}$$

If the outage schedule is altered by the system operator, the inputs to the model will also need to be updated accordingly. The preceding is required to ensure the model considers the correct combination of generating units when performing load commitment.

4.9 Environmental dispatch

Environmental load dispatch is a set of constraints applicable to only coal fired generating units, which is aimed to reduce both emissions production and water utilisation (make-up). Considering the Air Quality Act, it is apparent that the air emissions limits set out for coal fired units are becoming increasingly stringent. In order to comply to the prescribed limits, coal fired utilities can either allow for large scale maintenance to be conducted on electrostatic precipitations (ESP)/ fabric filter plants (FFP) to improve current performance or install new technologies such as flue gas desulphurization (FGD) plants. These are however long term solutions which will require significant capital expenditure. The same is true when considering water consumption for coal fired units. To reduce water consumption, long term solutions will entail large maintenance interventions which will require a significant amount of time. South Africa is an arid country and immediate short term interventions are required to prevent water shortages at communities neighboring coal fired units.

In the interim, coal fired utilities can incorporate air emissions and water consumption limits into a load dispatch optimisation model to reduce both emissions production and water consumption. By doing so, the model will commit the coal fired units in such a manner as to maintain the units below the prescribed limits. Note that emissions production is load dependent whereas water make-up is load independent. Excessive water make-up is reliant on the quantity of water leaks, incorrect blowdown cycles, passing valves on the plant and or open cooling water evaporation. Researchers such as Ghasemi (2013) and Manor*et al.* (2009) propose different ways of incorporating the emissions legislation into the model. As mentioned in Chapter 3, these limitations can either be incorporated into the model as another objective function or as a constraint. For the purpose of this dissertation, the constraint methodology is utilised. The water consumption limits have not yet been mentioned in literature and are added into the UCEELD model to improve its accuracy when compared to reality. Equations (4.41) and (4.42) are used to enforce both emissions production and water consumption limits in the optimisation model. Input parameter $e_{hj}^{(f)}$ in the above equation refers to

the discredited envelope of emissions emitted as a result of the load range at which each coal fired unit is able to operate within. Note that the emissions concentration (mg/Sm^3) is either influenced by unit load, coal quality or poor plant performance. Parameter $e_{hj}^{(f)}$ is multiplied with parameter $q_{ph}^{(e)}$ to induce a stochastic influence to the optimisation problem as previously mentioned. Term $e_{hj}^{(f)} q_{ph}^{(e)}$ in (4.41) is multiplied with binary decision variable z_{pthj} and summated in terms of index $j \in \mathcal{J}$ to select the emissions emitted by a coal fired unit as a result of the load at which the unit is dispatched. The multiplication function is set equal to or less than the emissions limit ($e^{(l)}$) as stated by the Air Quality Act. By doing so, the loading of each coal fired unit is prohibited to exceed a load range which will result in the contravention of the emissions limit. Note that this constraint is only considered when a unit is committed for operation.

$$\sum_{j \in \mathcal{J}} (q_{ph}^{(e)} e_{hj}^{(f)} z_{pthj}) \leq e^{(l)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.41}$$

Equation (4.42), is implemented to prevent coal fired units, situated in different geographical areas, from exceeding the maximum allowable water usage limits as prescribed per area. If a coal fired unit is committed to the grid in say geographical area 1, the water consumed by the affected unit will be accounted for by means of (4.42). If however the unit is decommitted, the opposite will apply. By multiplying parameter $q_h^{(w)}$ with binary decision variable y_{th} and summating the function in terms of index $h \in \mathcal{H}(n)$, the total water consumption as a result of the active coal fired units are obtained for each area. The multiplication function is divided with a fixed cost amount in $Rand/Volume$ to convert parameter $q_h^{(w)}$ to $Volume/Time$. This is done to ensure that the multiplication function of (4.42) is in the same SI units as that of parameter $w_n^{(m)}$. The cost associated with the water usage ($Rand/Volume$) is taken as the cost required to purify and treat the water before it enters the power generating process. Parameter $w_n^{(m)}$ depicts the water usage limit assigned to each geographical area. The multiplication function is set equal to and or less than $w_n^{(m)}$. By doing so, the amount of coal fired units committed per geographical area, will be dependent on the water limitations stipulated by the user for each area as well as the water consumption per unit as a result of poor plant performance.

$$\sum_{h \in \mathcal{H}(n)} ((q_h^{(w)} y_{th})/1505.49) \leq w_n^{(m)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{4.42}$$

4.10 Interconnected multi-area power flow

As stipulated in Section 4.3, variable $p_{pt(i,j)}^{(f)}$ denotes the power flow between geographical areas. When considering the foregoing, it is imperative to incorporate an upper and lower limit functionality to specify the maximum and minimum amount of power allowed to be exchanged between two geographical areas. It is incorporated into the optimisation model by means of (4.43) and (4.44) respectively. In literature Jadoun *et al.* (2015) and Pandit *et al.* (2015) also incorporate similar constraints which is defined as tie-line constraints, used to govern the maximum and minimum power flow between areas. In this dissertation, the same analogy is utilised.

In (4.43), variable $p_{pt(i,j)}^{(f)}$ is set equal to and or less than a user defined upper limit (b). This upper power flow limit will vary according to the agreement between power utilities situated in different geographical areas. If for example, a resident area does not have enough capacity to supply the demand, an agreement will be made with neighboring utilities to supply power equal to the shortfall of the resident area, and the upper limit will be updated accordingly. If however no power supply is required from a neighboring utility, the upper limit can be set equal to 0, which will ensure that (4.43) is discounted form the optimisation process.

$$p_{pt(i,j)}^{(f)} \leq b, \quad p \in \mathcal{P}, t \in \mathcal{T}, (i,j) \in A \tag{4.43}$$

The same principle applies when considering (4.44), but it is only focused on the user defined lower limit. Note that variable $p^{(f)}_{pt(i,j)}$ is set equal to and or greater than 0, to prevent a negative power flow between two geographical areas. This limit can either remain equal to 0 or be adjusted to match the agreement between two power utilities.

$$p^{(f)}_{pt(i,j)} \geq 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, (i,j) \in A \tag{4.44}$$

These two equations thus govern the allowable amount of power that may be distributed by the optimisation model between the different geographical areas.

4.11 Hydro generating units

As mentioned in Section 4.2, power generation of hydro units are governed by the dispatch schedule obtained from the department of water affairs and is provided as a user input to the optimisation process. This schedule allows the commitment of hydro generating units to the grid for only a certain time period, in order to manage the water levels downstream of each hydro station. In order to enforce time dependent unit commitment for each hydro unit, (4.45) is incorporated into the optimisation model. Researchers such as Chen (2008) and Chan *et al.* (2001) proposed different constraints to consider when incorporating hydro units into the power generation optimisation problems. Note however that a lot of the constraints mentioned by these researchers are not incorporated into the hydro model as a steady river/ dam level is assumed given the short study horizon considered in the UCEELD model. For the pumped storage units discussed in Section 4.12, a more detailed model is developed to account for the water level variability.

Input parameter $w^{(h)}_{hj}$ in (4.45) represents the discretised envelope of hydro water consumption per unit as a result of the load range at which each hydro unit is able to operate within. Note that $w^{(h)}_{hj}$ is multiplied with binary decision variable z_{pthj} and summated in terms of index $j \in \mathcal{J}$ to select the water consumed by a hydro unit as a result of the load at which the unit is dispatched. By setting the multiplication function equal to or less than 0, the user forces downtime of the affected hydro unit, as per the dispatch schedule. In doing this, variable z_{pthj} is assigned a value of 0 in order to ensure the feasibility of (4.45).

It is, however, important to note that this equation is only applicable to certain hydro units, given a specific time horizon. The units and time period affected by (4.45) are specified by constraining indexes $h \in \mathcal{H}(n)$ and $t \in \mathcal{T}$, and can be updated by the user as the dispatch schedule changes. In this example, (4.45) will only be in effect when the model's calculation process reaches time period $t >= 5$ to $t <= 10$. During this time, hydro units ranging from $h >= 99$ to $h <= 105$ will be prevented from being committed to the grid in an attempt to minimise the disruption of the downstream ecosystem. After the predefined time period has passed, the optimisation model is then again allowed to select the affected hydro units for generation. Note however that during time period $t >= 5$ to $t <= 10$, units $h <= 99$ and $h >= 105$ will still be available for commitment. Note that the aforementioned is however only an example and in reality, indexes $t \in \mathcal{T}$ and $h \in \mathcal{H}(n)$ can be constrained as the user deems it necessary.

$$\sum_{j \in \mathcal{J}} w^{(h)}_{hj} z_{pthj} \leq 0, \quad p \in \mathcal{P}, t \in \mathcal{T} : 5 \leq t \leq 10, n \in \mathcal{N}, h \in \mathcal{H}(n) : 99 \leq h \leq 105 \tag{4.45}$$

Equation (4.46) is added to the optimisation model, in case unstable operating conditions are identified for a hydro unit, when operating within its design load range. This equation is used to limit the water consumed by the affected hydro unit and prevent the unit from operating outside its stable and or safe operating envelope. Say for example a hydro unit's design load range is between 0 MW to 120 MW. It might be that above 100 MW unstable operating conditions are experienced. Unstable conditions can arise as a result of plant defects and or outstanding maintenance. Considering

the foregoing, the user will not want this unit to be scheduled above 100 MW, by the optimisation model, as it may result in undesired outcomes. To enforce this principle and prevent loading above 100 MW, (4.46) is applied.

$$\sum_{j\in\mathcal{J}} w_{hj}^{(h)} z_{pthj} \leq w_h^{(hm)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{4.46}$$

In order to limit hydro power generation, the multiplication function in (4.46) is set equal to or less than the maximum allowable water consumption limit ($w_h^{(hm)}$). In the above example this will be equal to the quantity of water consumed at 100 MW. The water consumption limit is specified by the user and will prevent the hydro unit from operating outside its desired load range. If however no limitations are required on the power generation of a hydro unit, parameter $w_h^{(hm)}$ needs to be set to a value greater than the maximum continuous rating of the affected unit.

4.12 Pumped storage generating units

Pumped storage units act as a battery and or energy storage device to the power grid and are utilised for power generation during peak periods. Being a cheap source of power generation, the optimisation model will attempt to schedule these units at full load for the entire study horizon. This is however not practicable as the water levels of the upper reservoirs only contain a limited water capacity available for power generation. If not managed correctly, these reservoirs will run out of capacity after a few hours of operation. In order to prevent this from happening, pumped storage units are forced by the system operator to occasionally switch between generating and pumping mode, to maintain the upper reservoirs at a predefined level setpoint. The preceding is usually initiated during lower grid power consumption, typical during early mornings or late nights, where coal fired units are capable of satisfying the grid demand. This is done to replenish the upper reservoir water levels to ensure that there are enough pumped storage generation capacity available during each peak period. In order to describe the process of pumped storage power generation mathematically (4.47) to (4.60) are incorporated into the optimisation model. Borghetti *et al.* (2008) together with Ni & Guan (1999) presents detailed models used to simulate the dynamics of pumped storage stations when incorporated into the dispatch and commitment schedules.

4.12.1 Water balance

As mentioned in Section 4.2, the power generated by pumped storage stations is governed by the available power contained in the upper and lower water reservoir for each station. It is, therefore necessary to model the level of each reservoir (upper and lower), and track the changes to the levels as a result of a unit either operating in generating and or pumping mode.

To track the water volume of the upper reservoir of each pumped storage station, (4.47) is incorporated into the model. Variables $v_{pts}^{(u)}$ and $v_{p(t-1)s}^{(u)}$ in (4.47) denotes the water volume of each upper reservoir at time periods *t* and *t-1* respectively. Variable $v_{p(t-1)su}^{(fl)}$ which is summated in terms of index $h \in \mathcal{H}(n)$ is representative of the quantity of water either flowing out of (when in generating mode) or into each reservoir (when in pumping mode). The detailed calculation of $v_{p(t-1)su}^{(fl)}$ will be provided in the sections to follow. When a pumped storage unit is in generating mode, $v_{p(t-1)su}^{(fl)}$ will take on a positive value. This will be indicative of water flowing out of the reservoir. The opposite is true when in pumping mode. In generating mode, $v_{p(t-1)s}^{(u)}$ will be reduced by the value of $v_{p(t-1)su}^{(fl)}$. The answer to the subtraction function will then be captured by $v_{pts}^{(u)}$ and be logged by the optimisation model as the new reservoir volume for time period *t*. The preceding is an iterative process which will be repeated up and to the end of the study horizon. This is done to keep track of the upper reservoir volume during each time period as power is generated or consumed. This

equation is effectively a mass balance of the water flowing out of and into each upper reservoir. Refer to (4.47) for the mathematical formulation.

$$v_{pts}^{(u)} = v_{p(t-1)s}^{(u)} - \sum_{u \in \mathcal{U}(s)} v_{p(t-1)su}^{(fl)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n) \tag{4.47}$$

In order to prevent pumped storage units from continuously being scheduled in generating mode, due to it being a cheap source of power generation, (4.48) is added to the optimisation model. Input parameter $v_{ts}^{(ue)}$ in (4.48) is user defined and depicts the upper reservoir water volume set-point (minimum allowable volume) at time period t. Note that $v_{ts}^{(ue)}$ is multiplied with binary input parameter $o_{su}^{(p)}$ representing the outage schedule for each pumped storage unit. If a unit is scheduled for an outage, $o_{su}^{(p)}$ will be assigned a value of 0. The opposite applies if the unit is available for generation. By multiplying these two parameters with one another, the user ensures that this equation will only be considered when units in a pumped storage station are available for generation.

By providing the upper reservoirs with a setpoint, (4.48) will force the optimisation model to switch the pumped storage units between generating and pumping modes in order to meet the desired water volume at the predefined time period. In (4.48) the time period at which the water volume setpoint must be reached is set at t = 6 of the study horizon. Note however that at hour 7 in the study horizon and beyond, the model will be allowed to commit the pumped storage units for generation continuously without having to consider any setpoint. The reason for this is the fact that the volume setpoint has already been met at hour 6 of the study horizon. If the remaining period of the study horizon is too long, say for example from hour 7 to hour 24, it might cause the reservoirs to run out of capacity. The reason for this being, is that the pumped storage units will be scheduled at full load from hour 7 to 24 which will result in the upper reservoirs being completely drained. To prevent the preceding from occurring, the user can adjust the time period at which the volume setpoint needs to be satisfied to a subsequent hour in the study horizon, say for example hour 15. This will force the optimisation model to maintain sufficient capacity in the upper reservoirs for the remaining periods of the study horizon, by switching between generating and pumping modes during hour 1 to hour 15.

It is for this reason that the system operator needs to analyze the available capacity of the reservoirs on a continuous basis. By doing so, the operator will guarantee that a realistic setpoint (maximum allowable water volume ($v_{ts}^{(ue)}$)) is provided at the correct time period (lowest grid demand), to ensure enough generation capacity is available for the current as well as subsequent study horizons at peak demand.

$$V_{pts}^{(u)} \geq V_{ts}^{(ue)} o_{su}^{(p)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.48}$$

Lastly, to effectively track the water balance of the upper reservoir, using (4.47) and (4.48), the user needs to initialize variable $v_{pts}^{(u)}$ at time period $t = 1$. Without providing an initial reservoir volume, the upper reservoir water balance calculation will fail to execute. The initialization functionality is added to the model by means of (4.49). In this equation variable $v_{pts}^{(u)}$ is set equal to the user defined input parameter $v_{ts}^{(ui)}$ (initial upper reservoir water volume) at time period 1 of the study horizon. This value needs to correspond to the actual water volume as contained in the upper reservoir of each pumped storage station. Providing the model with an inaccurate input value, will result in erroneous commitment and dispatch schedules being obtained for each pumped storage unit.

$$v_{pts}^{(u)} = v_{ts}^{(ui)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n) \tag{4.49}$$

The same mathematical principle, which was used to track the upper reservoir water volume, is incorporated to track the water volume of the bottom reservoir of each pumped storage station.

This functionality is incorporated into the model by means of (4.50) and (4.51). In (4.50), variables $v^{(l)}_{pts}$ and $v^{(l)}_{p(t-1)s}$ denotes the water volume of each bottom reservoir at time periods t and *t-1* respectively. Contrary to (4.47), variable $v^{(fl)}_{p(t-1)su}$ is added to $v^{(l)}_{p(t-1)s}$ in (4.50), to obtain the new reservoir volume at time period t. The reason for this being is that when a pumped storage unit is in generating mode ($v^{(fl)}_{p(t-1)su}$ takes on a positive value), water will be consumed from the upper reservoir and transported to the bottom reservoir, resulting in a rise in the bottom reservoir water volume. The magnitude of volume rise in the bottom reservoir will be directly dependent on the amount of water consumed from the upper reservoir. As a result, the amount consumed from the upper reservoir will need to be added to the volume of the bottom reservoir to ensure conservation of the mass balance. The opposite is true when in pumping mode. Considering the dynamics of the bottom reservoir, the user is not required to provide a water volume setpoint as was done for the upper reservoirs in (4.48). By controlling the upper reservoir volume, it will indirectly govern the bottom reservoir volume at each time period t as these reservoirs are directly linked to one another. Effectively, (4.50) will respond to any changes in the upper reservoir volume. Similar to (4.47), the calculation is an iterative process which will repeat until the end of the study horizon is reached. By doing this, the dynamics in the bottom reservoirs can be tracked accurately to ensure realistic results are obtained from the model when operating in either generating or pumping modes.

$$v^{(l)}_{pts} = v^{(l)}_{p(t-1)s} + \sum_{u \in \mathcal{U}(s)} v^{(fl)}_{p(t-1)su}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n) \tag{4.50}$$

Equation (4.51), similar to (4.49), is incorporated to initialize variable $v^{(l)}_{pts}$ in (4.50) at time period $t = 1$. In (4.51) variable $v^{(l)}_{pts}$ is set equal to the user defined input parameter $v^{(li)}_{ts}$ (initial bottom reservoir volume) at time period 1 of the study horizon. By providing an acceptable starting value, the model will be able to track the fluctuations in the water balance of the bottom reservoir accurately.

$$v^{(l)}_{pts} = v^{(li)}_{ts}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n) \tag{4.51}$$

In order to accurately track the water flow through both upper and lower reservoirs, as depicted in the above sections, variable $v^{(fl)}_{ptsu}$ is incorporated into the model. This variable is representative of the quantity of water either flowing out of or into each pumped storage unit at time period t of the study horizon. In (4.52), the detailed mathematical calculation for $v^{(fl)}_{ptsu}$ is provided. Note that the calculation is done using the bottom reservoir as reference point. Meaning that $v^{(fl)}_{ptsu}$ takes on a positive value when a pumped storage unit is operating in generating mode (water flowing into bottom reservoir calculated by term 2 of (4.52)) and a negative value when in pumping mode (water flowing out of bottom reservoir calculated by term 3 of (4.52)). In term 2 of (4.52), parameter $v^{(fl)}_{suvj}$ refers to the discretised envelope of water consumed as a result of the load range at which each pumped storage unit is able to operate within and due to the variability of the upper reservoir volume over time. Given a fixed operating load, the water consumption for a pumped storage unit can vary as a result of the variability of the upper reservoir volume. For example, if a pumped storage unit can be dispatched at a load range between 0 MW - 120 MW with an upper reservoir volume of 20 m^3, the equivalent volumetric flow rate through the turbine would range between 0 m^3/s - 2 m^3/s. If however the same pumped storage unit is dispatched at an upper reservoir volume of 10 m^3, the equivalent volumetric flow rate through the turbine would need to range between 0 m^3/s - 4 m^3/s to maintain a load range between 0 MW - 120 MW. If the volumetric flow rate is not adjusted, the operating load range of the pumped storage unit will drop to between 0 MW - 90 MW as a result of the reduced potential energy in the upper reservoir.

Infinite examples such as this exist, however, to reduce the model complexity, it was assumed that there are only two possible upper reservoir volume ranges (consequently two volumetric flow rate

ranges which are contained in $v_{suvj}^{(fl)}$) that the model is able to chose from, for a fixed load range. By allowing more data points to be considered (i.e. more reservoir volume ranges), it will definitely increase the model's accuracy, but will become impractical to solve due to increased model solution time. Given the two reservoir volume ranges, the model will utilise the results obtained from the water balance to determine the upper reservoir volume at time period t. After identifying the reservoir volume, f_{ptsuvj} will be assigned a value of 1 for index $v \in \mathcal{V}$ corresponding to the identified volume. The remaining sections of index v will be assigned values of 0. Input parameter $v_{suvj}^{(fl)}$ is then multiplied with binary decision variable f_{ptsuvj} and summated in terms of index $v \in \mathcal{V}$ and $j \in \mathcal{J}$ to select the water consumed by a pumped storage unit as a result of the load at which each unit is dispatched and the volume contained in the upper reservoir. Note that the multiplication function is subtracted from $v_{ptsu}^{(fl)}$ as a result of the bottom reservoir being used as reference point.

In term 3, input parameter $p_{su}^{(q)}$ denotes that quantity of water consumed from the bottom reservoir when a pump storage unit is operated in pumping mode. For this application, it is assumed that the turbine of each pumped storage unit consumes a constant water quantity when in pumping mode. Parameter $p_{su}^{(q)}$ is multiplied with binary decision variable $p_{tsu}^{(r)}$ to activate the multiplication function (term 3) when a unit is switched to pumping mode. When in pumping mode, $p_{tsu}^{(r)}$ will be assigned a value of 1 and vice versa when switched to generating mode. By setting (4.52) equal to 0, it ensures that $v_{ptsu}^{(fl)}$ will take on the value of either term 2 or term 3, depending on which term is active for the time period under consideration. Take note that both terms cannot be active at the same time. A pumped storage unit can either be in generating mode or pumping mode in a single time period, not both. Refer to Section 4.5 for a detail explanation regarding this statement.

$$v_{ptsu}^{(fl)} - \sum_{v \in \mathcal{V}} \sum_{j \in \mathcal{J}} v_{suvj}^{(fl)} f_{ptsuvj} + p_{su}^{(q)} p_{tsu}^{(r)} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.52}$$

By accurately modeling the dynamics of the upper and lower reservoir volumes using the equations set out in Section 4.12.1, it will ensure that the pumped storage units considered in the optimisation model will be committed to the grid at time periods and dispatched at loads which will correspond to periods of high grid demands to prevent frequency incidents.

4.12.2 Volumetric flow rate range selection

As mentioned throughout Sections 4.3 and 4.12.1, it is assumed that there will only be two volume ranges in which each upper reservoir volume will be allowed to fluctuate within. In order to determine within which volume range each upper reservoir falls, the water balance is utilised. Applying the results obtained from variable $v_{pts}^{(u)}$ in the water balance, the actual upper reservoir volume is obtained. By utilising (4.53) and (4.54) in conjunction with $v_{pts}^{(u)}$, the model can determine within which predefined reservoir volume range the current reservoir volume falls. Input parameter r_{sv} in both equations are representative of the predefined volume ranges for the upper pumped storage reservoirs. The volume ranges are divided into range (1 - 2) and range (2 - 3). The values contained in each volume range is determined by user input.

Parameter r_{sv} is multiplied with binary decision variable $d_{ptsuv}^{(m)}$ in (4.53) to track the upper bound of the top reservoir volume. If for example $v_{pts}^{(u)}$ falls within range (1 - 2), variable $d_{ptsuv}^{(m)}$ will be assigned a value of 1 for index $v \in \mathcal{V}$ corresponding to point 1 of range (1 - 2). Point 1 will then form the upper limit of the range within which the reservoir volume falls. Note that in this example, variable $d_{ptsuv}^{(m)}$ will be assigned 0 values in point 2 and 3 of index v. Term 2 in (4.53) is added to ensure that the equation remains feasible although a pumped storage unit is decommitted from the grid. If this term was emitted from the equation, an infeasible solution will be obtained for (4.53) and the model will not be able to solve correctly.

$$\sum_{v\in\mathcal{V}} r_{sv}(d^{(m)}_{ptsuv} + (1 - t^{(r)}_{tsu})) \geq v^{(u)}_{pts}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.53}$$

Similar to (4.53), input parameter r_{sv} in (4.54) is multiplied with binary decision variable $d^{(n)}_{ptsuv}$ to track the lower bound of the top reservoir volume. For the above example where the reservoir volume falls within range (1 - 2), variable $d^{(n)}_{ptsuv}$ will be assigned a value of 1 for index $v \in \mathcal{V}$ corresponding to point 2 of range (1 - 2). Point 2 will form the lower limit of the range within which the reservoir volume falls. Point 1 and 3 in index v will be assigned 0 values for variable $d^{(n)}_{ptsuv}$. If however a pumped storage unit is operated in generating mode for a prolonged period of time, it might result in the volume range moving from range (1 - 2) to range (2 - 3) due to the reduction in reservoir volume. Consequently, binary variables $d^{(m)}_{ptsuv}$ and $d^{(n)}_{ptsuv}$ will be updated by the model, by assigning a value of 1 for both points 2 and 3 in index v respectively. The remaining points in index v will be assigned 0 values.

$$\sum_{v\in\mathcal{V}} r_{sv} d^{(n)}_{ptsuv} \leq v^{(u)}_{pts}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.54}$$

In order to force variables $d^{(m)}_{ptsuv}$ and $d^{(n)}_{ptsuv}$ to track the upper and lower bounds within which the top reservoir volume fluctuates, (4.55) is incorporated into the model. Variable $d^{(n)}_{ptsuv}$ is subtracted from $d^{(m)}_{ptsu(v-1)}$ and is set equal to a value of 0. By equating the subtraction function to 0, it ensures that both binary decision variables are assigned either a value of 1 or 0 simultaneously, and prevent these variables from being assigned different values in time period t. If for example $d^{(n)}_{ptsuv}$ is assigned a value of 1 for index v corresponding to point 2 of range (1 - 2), $d^{(m)}_{ptsuv}$ will be forced by (4.55) to assign a value of 1 to point 1 of range (1 - 2) to ensure the equation's feasibility. By using the results obtained from this equation, the model will utilise range (1 - 2) as the volume range within which the upper reservoir volume operates for time period t.

$$d^{(m)}_{ptsu(v-1)} - d^{(n)}_{ptsuv} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.55}$$

Variable f_{ptsuvj}, which is used for load and volumetric flow decision making, is linked to the lower reservoir volume range by means of (4.56). In this equation f_{ptsuvj} is set equal to $d^{(n)}_{ptsuv}$. By implementing this equation, the model will prevent a value of 1 to be assigned to f_{ptsuvj} if the upper reservoir's volume of the affected unit does not fall within the limit prescribed by f_{ptsuvj} and $d^{(m)}_{ptsuv}$. For example, if the reservoir's volume falls within the (1 - 2) volume range, the model will only be allowed to assign values of 1 to f_{ptsuvj} in index v corresponding to point 2 of range (1 - 2). In (4.56), f_{ptsuvj} is set equal to the lower bound of the volume range and therefore point 2 is used as reference. The equation can however be updated by the system operator by setting f_{ptsuvj} equal to the upper bound which will force the model to use point 1 of volume range (1 - 2) as reference.

$$\sum_{j\in\mathcal{J}} f_{ptsuvj} = d^{(n)}_{ptsuv}, \quad p \in \mathcal{P}, t \in T, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s), v \in \mathcal{V} \tag{4.56}$$

It is, however, important to note that the process of volumetric flow rate range selection is only applicable if a pumped storage unit is operating in generating mode. If at any time, a pumped storage unit is either decommitted from the grid or switched to pumping mode, (4.53) to (4.55) needs to be discounted from the model. In order to achieve the foregoing, (4.57) and (4.58) is added. Note that in these equations both variables $d^{(n)}_{ptsuv}$ and $d^{(m)}_{ptsuv}$ are set equal to binary decision variable $t^{(r)}_{tsu}$. Variable $t^{(r)}_{tsu}$, as mentioned in Section 4.5, is used to track the on/off selection of pumped storage units in generating mode. Therefore, if a pumped storage unit is selected to generating

mode, $t_{tsu}^{(r)}$ will be assigned a value of 1. This, in turn, will allow (4.57) and (4.58) to come into effect and start tracking the volume range within which a pumped storage unit operate within, when in generating mode. The contrary is true when a unit receives an off selection. $t_{tsu}^{(r)}$ will be assigned a value of 0, and (4.57) and (4.58) will be discounted from the decision making process.

$$\sum_{v \in \mathcal{V}} d_{ptsuv}^{(m)} = t_{tsu}^{(r)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{4.57}$$

$$\sum_{v \in \mathcal{V}} d_{ptsuv}^{(n)} = t_{tsu}^{(r)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s)c \tag{4.58}$$

By including the content discussed throughout 4.12 into the optimisation model, it most definitely increases the complexity as well as the solution time. Although this is the case, these constraints cannot be omitted from the model as it will result in an unrealistic and inaccurate commitment and dispatch schedule being obtained for the pumped storage units.

4.13 Stochastic elements

Stochastic elements or variables refer to data which might have a random probability distribution or pattern that may be analyzed statistically. Such variables also exist in the unit commitment and load dispatch optimisation problem. Due to the uncertainty contained within these variables, it affects the model's decision making process significantly. Therefore it needs to be incorporated into the optimisation model to ensure the results obtained will be comprehensive and accurate. These variables include power grid demand variability and coal fired fuel consumption variability. Both these variables are predicted values derived from historical data. It is however imperative to note that no prediction can be done without some degree of uncertainty and therefore stochasticity need to be incorporated into the model. A detailed discussion is provided on each in the sections to follow. The principle of two-stage stochasitc modeling as presented in Papavasiliou & Oren (2013) and Shiina (2004) is applied in the UCEELD model formulation. The model first needs to perform its decision making process on the commitment of the units given the stochastic nature of the forecasted demand, fuel cost and emissions production. Only thereafter, the dispatch process can be initiated in order to develop the optimal generation schedule.

4.13.1 Power demand variability

In order to consider the stochasticity of the predicted power grid demand, index p is added to the optimisation model as discussed in the introduction to Chapter 4. By using index p, input parameter $p_{ptn}^{(d)}$ as depicted in Section 4.3, is modified to contain a stochastic index. By doing this, parameter $p_{ptn}^{(d)}$ can be provided with input data sets containing different scenarios of grid demands which might occur in time period t. Note that this is not contradictory to what was discussed in Section 4.3, but only contributory to make the model more relevant to actual conditions. In Section 4.3, $p_{ptn}^{(d)}$ was defined as a single data set (effectively index $p = 1$) whereas now it is introduced as a multidimensional data set (effectively index $p > 1$). Therefore, multiple data sets (scenarios) can be contained within $p_{ptn}^{(d)}$ for different indexes of p. Say for example index p ranges from 1 to 3. This will mean that variable $p_{ptn}^{(d)}$ will consist of three sets of possible grid demand data, which will need to be satisfied by the model, via unit commitment and load dispatch, in order to obtain a feasible solution. Note however that the range for index p is user defined and can be updated as the user deems necessary. If however the range for index p is increased, the model complexity and solution time will be greatly amplified. Index p is also added to the remaining variables in the model to account for the power grid demand stochasticity.

Another factor to consider when introducing power demand variability (i.e. different load scenarios), is the influence it has on the total fuel cost calculated in the objective function. According to historical data, each grid demand scenario has a certain probability to occur during each time period t. The weight assigned to each scenario as a result of its probability to occur, will directly influence the total fuel cost as calculated in the objective function. This is because the fuel cost per scenario will be proportioned according to this same probability. Note that the summation of all the probabilities should provide the user with an answer of 1. To incorporate this effect accurately, (4.2) is multiplied with input parameter r_p. This parameter proportions the fuel cost ($c_{hj}^{(f)}$) in (4.2) with a weight equivalent to the probability for each power demand scenario to occur. Which ever scenario carries the most weight, will be the primary driver for the decision making process of the optimisation model. In other words, the fuel cost used to satisfy the grid demand with the highest probability to occur, will have the biggest influence on the model's commitment and dispatch process. For example, if there are three grid demand scenarios, it might be possible that the probability is 0.3 for scenario 1, 0.5 for scenario 2 and 0.2 for scenario 3 to occur. This means that the probability that the optimisation model will need to satisfy the power grid demand contained in data set 2 of index p (while consuming an equivalent amount of fuel) is the most likely to occur. Therefore this scenario will outweigh the other scenarios and will be the primary driver for the decision making process. However, the model decision making cannot be solely based on scenario 2 as scenario 1 and 3 will also have an effect on the final outcome, even though it is less likely to occur. The model needs to take all possible scenarios as supplied by the system operator into consideration in the objective function in order to provide a consolidated answer, considering each scenario. By incorporating this functionality into the optimisation model, the system operator will be able to obtain the best possible commitment and load dispatch schedule, provided a certain level of uncertainty.

4.13.2 Fuel consumption variability

Coal fired fuel consumption stochasticity is incorporated by means of adding index p to all relevant variables in the optimisation model and by multiplying (4.2) with input parameter q_{ph}. This parameter is used to simulate different fuel consumption scenarios (by means of multiplying (4.2) with different fractions) that might exist. Fuel consumption variability of a coal fired unit may be due to changes in plant conditions and or plant performance. Depending on the input variability range obtained from the system operator, q_{ph} can contain multiple scenarios with which the fixed coal usage, depicted by input parameter $c_{hj}^{(f)}$, can be adjusted to model fuel cost stochasticity. By incorporating this functionality into the model, it increases the model complexity significantly. The reason for this is that, for each possible fuel consumption scenario, the model needs to solve the optimisation problem by iterating through the different load scenarios as mentioned in Section 4.13.1, as well as compute through each index t, n, h, j for coal fired and hydro units and through each index t, n, s, u and v for pumped storage units. By adding this functionality, it creates a stacked effect (something similar to a *for loop*) which increases the model size exponentially. The model, therefore, needs to iterate through all of the possible combinations in order to minimise the objective function. Depending on the number of combinations, the model's solving time will increase until it becomes impractical to solve.

4.13.3 Emissions production variability

Emissions production variability is incorporated into the mathematical model by means of adding index p to all relevant variables and multiplying (4.10) with input parameter $q_{ph}^{(e)}$. Emissions production variability of a coal fired unit may be a result of poor performing plant equipment or burning sub-grade coal with high ash and sulfur content. Depending on the inputs obtained from the system operator after analyzing the mentioned aspects, $q_{ph}^{(e)}$ may contain multiple scenarios with which the fixed emissions production ($e_{hj}^{(f)}$) can be adjusted. These scenarios are applied to model the emissions production stochasticity for a specific time period. The preceding ensures the model's adherence to the legislative emissions restrictions when committing and dispatching units, given a wide band of

possible variability. By doing so, the utility's chances of incurring a section 30 legal contravention are eliminated. This is done by ensuring the units are scheduled is such a manner as to handle any margin of emissions production variability as defined by the system operator. This makes the model more robust as it is capable of accounting for dynamic process fluctuations. This does however also increase the model complexity similar to what was explained in Section 4.13.2.

4.14 Transmission line losses

In addition to the constraints discussed throughout Sections 4.2 to 4.13, transmission line losses can also be taken into consideration when developing the UCEELD model. This will entail the commitment and dispatch of generating units while considering the length and capacity of the various transmission line networks as it can contribute to some extent of megawatt losses. The incorporation of these constraints are however complex and results in the problem again taking on a nonlinear form. The results obtained from (Bhattacharay & Chattopadhyay: 2010; Jeddi & Vahidinasab: 2014) shown that the transmission losses calculated by their models were only equivalent to approximately 1% of the total grid's generation capacity, which is relatively negligible. Incorporating these losses into the UCEELD model would increase the complexity of the problem as well as the computational time required to obtain an optimal solution. Given the foregoing reasons, it was decided that the transmission line losses would not be incorporated into the scope of this dissertation.

Chapter 5

Model Results and Interpretation

In this chapter the computational results obtained from the MILP UCEELD optimisation model, outlined in Chapter 4, is presented. The data obtained from the model is technically evaluated by means of verification and validation. The process of model verification is concerned with the determination of the model accuracy and establishing if the model's response is as expected. The verification process also entails the evaluation of the mathematical coding to assess the correctness thereof. The validation process is concerned with the ability of the optimisation model to solve realistic real-world problems with specific focus on how accurate the results represent reality. In many cases, the process of verification and validation are linked to one another and for this reason, both these evaluation techniques are addressed in this section.

The verification process is initially performed using a base model consisting of 3 thermal units. The base model is augmented to a 6 thermal, 1 hydro and 2 pumped storage unit case study as the model verification progresses. The purpose of utilising a small scale model is to simplify the process of critically evaluating the influence of each equality and inequality constraint on the model dynamics. Using an extensive model to perform the verification process will unnecessarily complicate the interpretation of the model results. The validation of the MILP UCEELD model is performed by means two problem instances. The first problem instance entails the evaluation of the model's performance by comparing the model results with the performance of an actual thermal power station. For the first problem instance, only 6 thermal generating units are considered. The second problem instance entails the solution of a 98 thermal, 8 hydro and 6 pumped storage unit case study. The size of the second problem instance used for the validation process corresponds to that of a realistic sized problem similar to what is found in South Africa's power grid. In Addition to the preceding, an analysis is performed to evaluate the effect model scaling has on solution time to determine if it will be possible to incorporate the optimisation model into a real life production environment.

The mathematical MILP UCEELD model is solved by using the commercial solver, Cplex (IBM Corp, 2015). The user defined input data is introduced into Cplex by means of linking the model to a Microsoft Excel visual basics tool, equipped with a graphical user interface aimed at simplifying the data acquisition process. A detailed explanation is provided in Section 5.5 on the workings of the data handling tool. The solution algorithm implemented by Cplex software to solve the MILP UCEELD problem is the branch-and-bound method. The input data used to simulate the different generating units are derived from a consistent data set obtained from a power utility named Eskom. In order to prevent the distribution of confidential design information, the data was adapted by the author. Effort was however made to ensure that the data is still a realistic representation of reality.

5.1 Study horizon and power demand

In order to perform the validation and verification process, a 24 hour (1 day) study horizon is considered with the study horizon being divided into hourly intervals. A random power demand curve is populated by the author for the verification process, as a way of testing the model response

with the inclusion of different constraints (Figure 5.1). When performing the validation process a realistic day ahead hourly demand curve is incorporated into the model in order to simulate reality (Figure 5.26). In both curves hour 1 corresponds to 01:00 am and hour 24 represents 24:00 pm when using the digital time format.

5.2 Model data

The input data used during both the validation and verification experiments are structured similarly to the data portrayed in Tables 5.1 to 5.6 with the quantity of data varying for each problem instance. The first three tables contain data related to 6 thermal units and 1 hydro unit whereas the data in Tables 5.4 to 5.6 are related to 2 pumped storage units. The below data is an extract of the inputs used to perform the data verification process as stipulated in Section 5.3 of this dissertation. The input data depicted in Tables 5.1 to 5.3 are consolidated using the VBA graphical user interface as discussed in Figure 8.1, Appendix B. Similar to the preceding, the data in Tables 5.4 to 5.6 are inserted into the excel user sheet using the VBA graphical user interface depicted in Figure 8.2, Appendix B.

The data in columns 2 to 4 of Table 5.1 refers to the exponential curve coefficients associated to the coal consumption of each unit. Similar is true for columns 5 and 6 which depicts the linear curve coefficients associated with each units emissions production. Columns 7 to 10 contains data regarding each units' minimum and maximum load ranges as well as the up and down ramp rates linked to each unit.

Table 5.1: Thermal and hydro units input data

Unit	Coal (a)	Coal (b)	Coal (c)	Emis (a)	Emis (b)	Min Load (MW)	Max Load (MW)	Ramp Up (Mw/h)	Ramp Down (MW/h)
1	-0.00002	0.12	-6.02	0.13	2.48	220	396	120	120
2	0.00004	0.047	13.09	0.16	5.81	90	400	120	120
3	-0.00001	0.097	-0.20	0.13	4.10	90	400	120	120
4	-0.00005	0.15	-5.87	0.08	5.82	90	396	120	120
5	-0.00013	0.20	-8.54	0.04	3.87	90	370	120	120
6	-0.00003	0.13	-9.51	0.05	2.74	120	520	140	140
7	0	0	0	0	0	90	150	120	120

In Table 5.2, the first 5 columns contain data relating to the start-up and shutdown fuel oil consumption rates for each unit as well as the cost per ton of coal and fuel oil. Columns 6 and 7 contains data linked to the minimum and maximum prohibited load ranges for each of the generating units. The last two columns refer to outage schedule data and water usage rates for each unit.

Table 5.2: Thermal and hydro units input data continued

Unit	Start-up Fuel Oil (Ton)	Shutdown Fuel Oil (Ton)	Coal Cost (R/Ton)	Fuel Oil Cost (R/Ton)	Min Prohib Load	Max Prohib Load	Unit Outage	Water Usage (a)
1	11	14	430	3000	0	0	0	0.068
2	15	12	400	7760	0	0	0	0.018
3	13	16	415	1250	12	15	0	0.028
4	32	25	380	3000	0	2	1	0.019
5	25	19	395	2864	0	0	1	0.020
6	40	37	470	3791	0	0	1	0.099
7	0	0	0	0	0	0	1	0

Table 5.3 is the last table associate to the thermal and hydro units. The data depicted in columns 2 to 4 provides information regarding the up and down time of each unit as well as the time each unit has either been operational or out of commission. Additional information contained within this table also includes start-up and shutdown ramp rates, maximum hydro flow rates and lastly linear coefficients pertaining to the hydro unit flow curves.

Table 5.3: Thermal and hydro units input data continued (1)

Unit	Up Time	Down Time	Time Running	Time Down	Start Ramp (MW/h)	Shutdown Ramp (MW/h)	Max Hydro Flow (m^3/s)	Hydro Flow (a)	Hydro Flow (b)
1	1	1	0	0	150	150	0	0	0
2	1	1	0	0	150	150	0	0	0
3	1	1	0	0	150	150	0	0	0
4	2	2	0	0	150	150	0	0	0
5	1	2	0	0	150	150	0	0	0
6	1	2	0	0	150	120	0	0	0
7	1	1	0	0	150	150	224000	0.57	-0.89

Table 5.4 is related to pumped storage data such as the minimum and maximum load ranges for each pumped storage unit, up and down ramp rates, minimum and maximum prohibited operating regions, the initial top and bottom reservoir volumes and lastly the final top reservoir volume setpoint for each reservoir.

Table 5.4: Pumped storage units input data

Unit	Min Load (MW)	Max Load (MW)	Up Ramp (MW/h)	Down Ramp (MW/h)	Min Prohib Load	Max Prohib Load	Initial Top Res Vol (GL)	Final Top Res Vol (GL)	Initial Bottom Res Vol (GL)
1	30	250	100	100	0	0	15.5	16.5	7.5
2	30	250	100	100	0	0	13.5	16.5	7.5

Table 5.5 summarises the exponential volumetric flow curve coefficients related to each pumped storage unit. Each unit is linked to the main reservoir and depending on the reservoir volume, the pumped storage unit flow rate curve will change (i.e. affected by the available head in each upper reservoir).

Table 5.5: Pumped storage units input data continued

Unit	Reservoir Volume Number	Pumped Storage (a)	Pumped Storage (b)	Pumped Storage (c)	Reservoir Volume (GL)
1	1	-2E-16	3.6568	-2E-13	27.5
1	2	-0.00004	2.3289	-0.0076	17.5
1	3	-0.00009	1.001	-0.0152	7.5
2	1	-2E-16	3.6568	-2E-13	27.5
2	2	-0.00004	2.3289	-0.0076	17.5
2	3	-0.00009	1.001	-0.0152	7.5

The last table contains data pertaining to the start-up and shut down ramp rates, pump volumetric flow and power consumption rates as well as the outage schedule for each pumped storage unit.

Table 5.6: Pumped storage units input data continued (1)

Unit	Start-Up Ramp (MW/h)	Shut down Ramp (MW/h)	Pump Flow (GL/h)	Pump Power (MW)	Unit On Outage
1	150	150	1.2	30	1
2	150	150	1.2	30	1

Note that the above tables are utilised as examples to provide the reader with insights into the input data used for the MILP UCEELD model. As previously mentioned, the amount of data required will be dependent on the size and complexity of the problem instance to be solved.

5.3 Model verification

During model verification, 12 different problem instances are considered with the first instance (Section 5.3.1) being the most rudimentary and the last (Section 5.3.12) comprising of the complete model as outlined in Chapter 4. In order to evaluate the dynamics of the MILP UCEELD model, the results obtained from the initial problem instance is considered as the base case model. As the problem instances progress additional constraints are added to the base model. The results obtained from the altered models are compared to its neighboring instances in order to determine the effect each constraint has on the model dynamics. Evaluating the change in unit loading every time a constraint is added to the model, provides the reader with the ability to discern between the different instances and the influence each constraint has on the model dynamics.

5.3.1 Single area coal fired unit dispatch - base model

The initial problem instance, only takes into consideration the coal cost associated with the thermal units as encompassed by ((4.1) - (4.3)) and the power balance constraints ((4.5) - (4.6)). The base model discounts the water consumption, start-up, and shutdown costs included in (4.1) as well as the unit commitment variable (y_{th}) in (4.3). Excluding the commitment variable prevents the model from performing unit commitment decision making. Input parameters r_p and q_{ph} are assigned values of 1 to eliminate the stochastic functionality of the fuel cost in (4.2). When considering the power balance, only the first term ($P^{(T)}$) and fourth term ($p^{(d)}_{ptn}$) are taken into consideration with the remaining terms being omitted. The remaining constraints are excluded from the base model problem instance. In this example, only 3 thermal units are considered, with the results obtained from solving the base model being portrayed in Figure 5.1.

In analyzing the results obtained in Figure 5.1, it is apparent that thermal unit 2 is the least expensive unit to dispatch with thermal unit 3 and unit 1 being more costly. Given the low operational cost associated to unit 2, the unit is dispatched at 250 MWh - 370 MWh output for most of the study horizon, with the load being reduced at the start of hour 21 due to the decline in power demand. Unit 3 being the second cheapest unit, is dispatched during peak demand periods (hours 1 to 4 and 8 to 16) at loads of between 340 MWh - 390 MWh where after the unit is deloaded to its minimum allowable generation as a result of the reduced grid demands. Unit 1 is mostly operated at 90 MWh for majority of the time due to the operational cost linked to the unit. It is only during time periods 1 to 2 and 11 to 12 where the unit is required to increase its load to satisfy the grid demand. The cumulative operating cost incurred by the utility, given the study horizon, equates to a value of R 2 060 377 in total. The stated operating cost will be used as baseline in evaluating the remainder of the problem instances. Note however that it is expected for the operational cost to increase as constraints are added to the model. The reason being is because the model is forced to satisfy operational constraints other than the primary goal of cost minimisation.

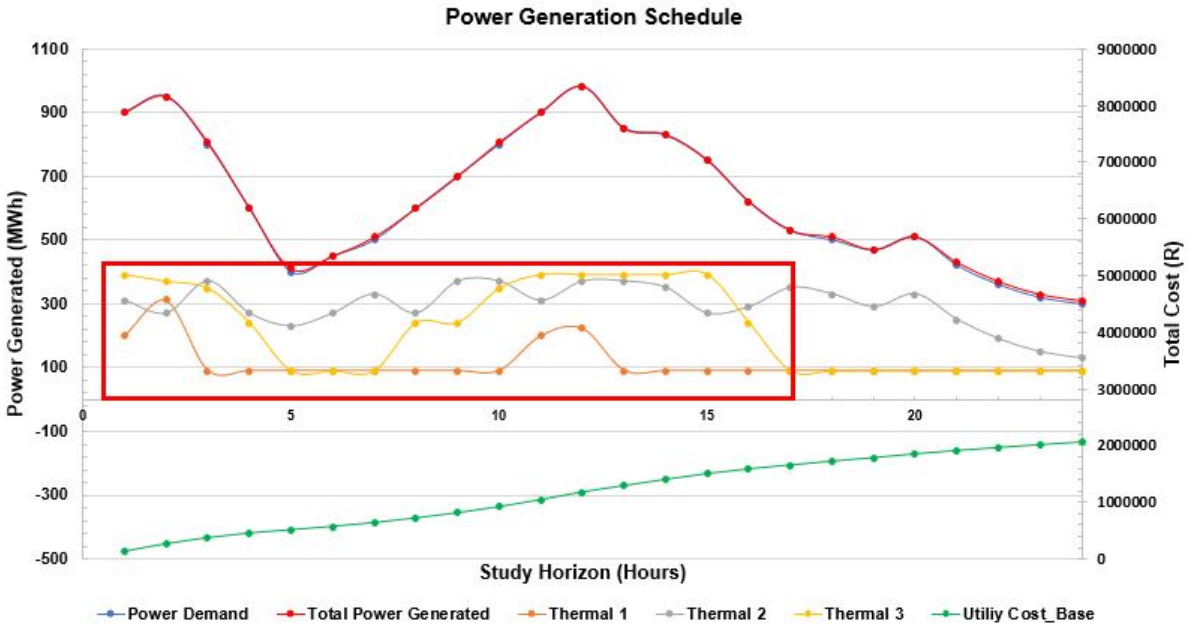

Figure 5.1: Single area coal fired unit dispatch - base model

5.3.2 Emissions limitations

Problem instance two entails the inclusion of (4.41) to the base model as explained in Section 5.3.1. In this equation parameter $q_{ph}^{(e)}$ is assigned a value of 1 to prevent the stochastic functionality of emissions production from being taken into consideration for the thermal units. The emissions limits ($e^{(l)}$) assigned to each unit for the given problem instance were 50 mg/sm^3, 60 mg/sm^3 and 55 mg/sm^3 respectively. Figure 5.2 is representative of the results computed by the optimisation model.

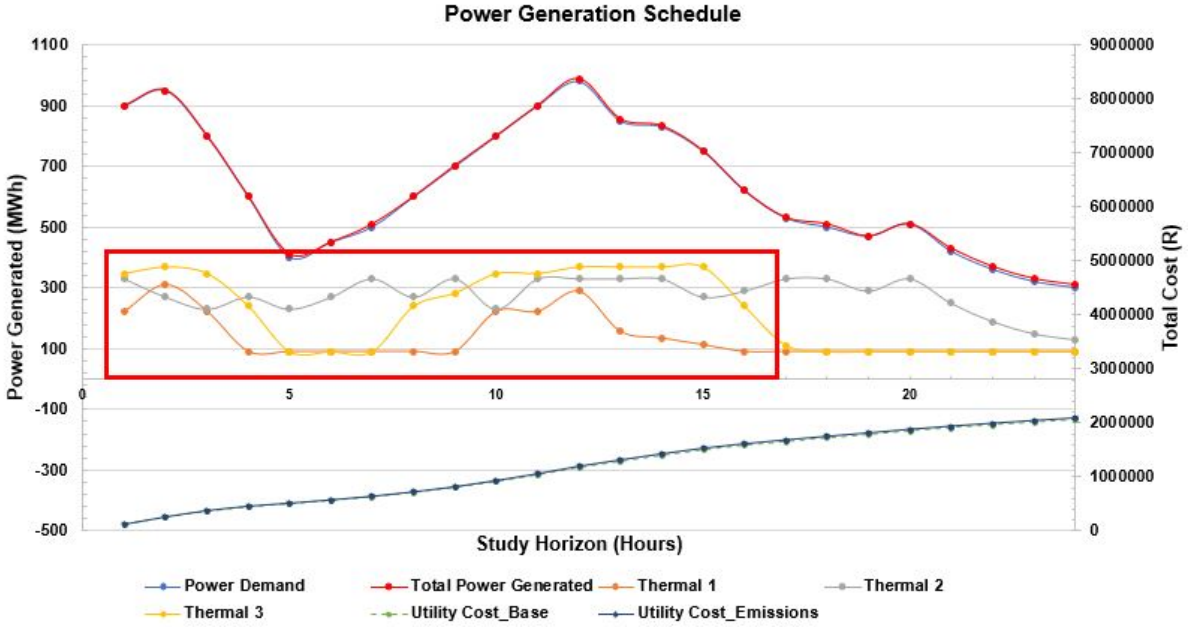

Figure 5.2: Emissions limitation influence on model results

Comparing the model results depicted in Figure 5.2 with that of Figure 5.1, it is apparent that the addition of the emissions limitation constraints definitely resulted in the alteration of the load dispatch schedule. Specific emphasis needs to be placed on the data encompassed in the red block. In order to prevent thermal unit 2 and 3 from exceeding the emissions limits, the model deloaded

these units to below the allowable limits. Unit 2 was not allowed to exceed a load output of 330 MWh in this problem instance which is much lower in comparison to the 370 MWh which was generated in instance 1. The maximum output of Unit 3 was also reduced with approximately 20 MWh, from 390 MWh to 370 MWh, as a result of the emissions limitations. In order to counteract the load reduction of unit 2 and unit 3, the load of thermal unit 1 was increased above 90 MWh at time periods 3, 10 and 13 - 15 to ensure the satisfaction of the power balance. The cumulative cost incurred by the utility after considering the emissions limitation constraints equates to a value of R 2 092 207. In comparison to the initial problem instance, the operating cost of instance 2 increased slightly. Meaning that if a utility is required to manage the emissions production by means of load dispatch decision making, it will result is higher operating costs being experienced.

5.3.3 Prohibited operating regions

Problem instance three involves the addition of (4.38) to the model mentioned in Section 5.3.2, with the purpose of the constraint being to account for prohibited operating regions. These regions entail the load ranges at which the units are not allowed to be dispatched at. In this problem instance only thermal unit 3 was constrained at the lower ($pr_h^{(l)}$) and upper ($pr_h^{(h)}$) load ranges corresponding to the discretised load data points 12 and 15 respectively. Meaning that although the design operating envelope of the unit is specified at 90 MWh - 390 MWh, the unit is only allowed to be dispatch between an operating range of 90 MWh - 304 MWh as a result of the prohibited regions. Prohibited operating regions can be induced by factors such as mentioned in Section 3.1.10. The influence of adding (4.38) to the model can be noted in the red block of Figure 5.3.

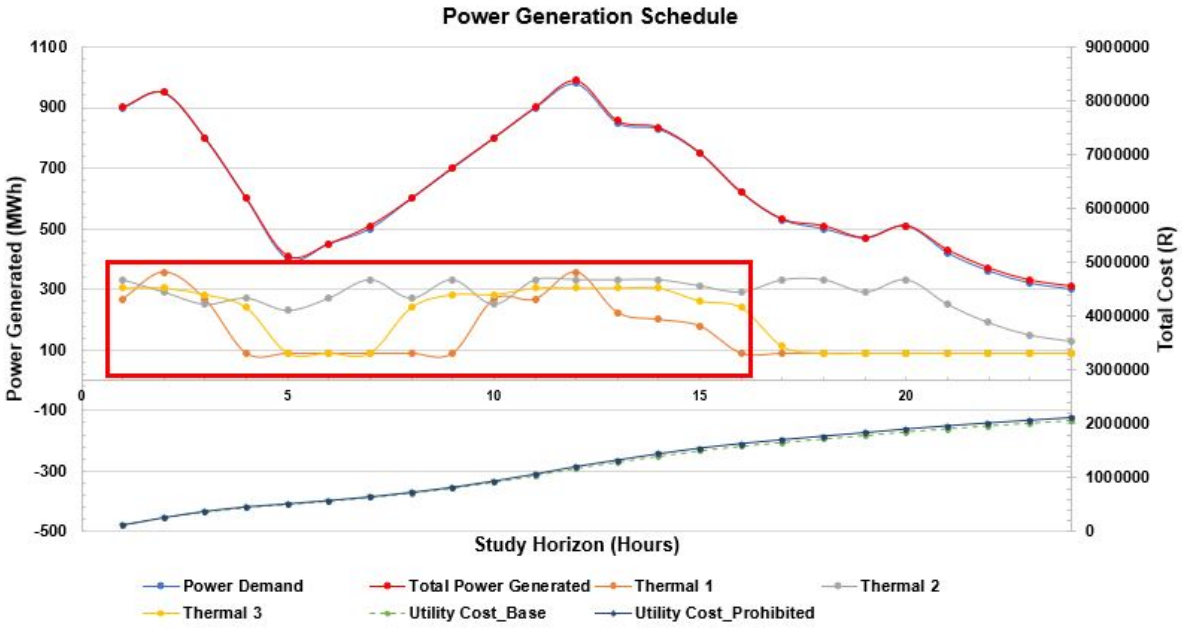

Figure 5.3: Prohibited operating regions' influence on model results

When evaluating the data portrayed in Figures 5.2 and 5.3, it is clear that the addition of (4.38) resulted in a change to the unit load dispatch schedule. Considering the fact that thermal unit 3 is prohibited to operate above 304 MWh, the model reduced the output of the unit significantly. Unit 3's maximum load output was reduced from 370 MWh in problem instance 2 to 304 MWh in the current instance. In order to account for the reduction in megawatt output from unit 3, the model increased the maximum power generated by thermal unit 1 from 290 MWh to 356 MWh. In this problem instance, the generation schedule for unit 2 remained relatively constant with only a few small changes. In adding (4.38) to the model, the cumulative cost increased from R 2 092 207 to R 2 105 511. Scrutinizing the increase in operational cost, it is apparent that if a unit is subjected to prohibited operating regions induced by plant defects, it can result in additional cost being incurred.

It is therefore beneficial for a utility to immediately attend to any defect conditions before it results in loading constraints on a unit as it can lead to significant cost savings.

5.3.4 Unit commitment

The unit commitment functionality is incorporated into the optimisation model by means of adding (4.16) to (4.22). Due to the interconnectivity of these constraints, they need to be lumped together into the model and cannot be individually added to evaluate the influence of each. Input parameters $t_h^{(u)}$ and $t_h^{(d)}$ in (4.18) and (4.20) are assigned values of 1 for the purpose of this problem instance whereas parameters $t_h^{(u1)}$ and $t_h^{(d1)}$ are allocated 0 values. Equation (4.22) is set equal to or greater than a value of 0 to ensure that at least 1 unit remains committed to the grid during the study horizon. The remaining variables in (4.16) to (4.22) are decision variables and will take on either values of 0 or 1 depending on the commitment schedule generated by the optimisation model. Terms 3 and 4 in (4.1) are incorporated into the model in order to account for the cost incurred by the utility as a result of unit start-up and shutdown operation. For explanatory purposes, the fuel oil cost associated to the shutdown of unit 3 was assumed to be quite low to demonstrate the effect of the unit commitment constraints. The results obtained from the computational process are portrayed in Figure 5.4. The unit commitment constraints added in problem instance 4 are an extension to the model discussed in Section 5.3.3.

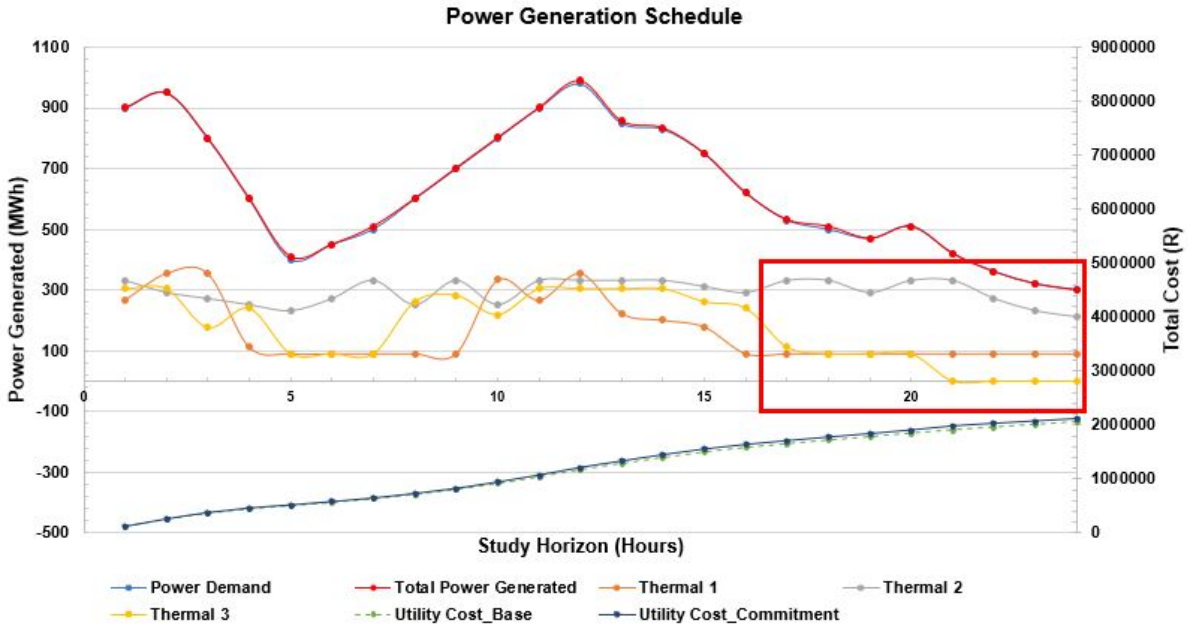

Figure 5.4: Unit commitment influence on model results

Adding the unit commitment constraints to the optimisation model, allows the model to decommit a unit from the grid instead of running the unit at its minimum allowable output load, given a low power demand. In analyzing Figure 5.4, it is apparent that the model prescribed decommitting thermal unit 3 from the grid at hour 21 of the study horizon instead of running the unit at 90 MWh as seen in Figure 5.3. The decommitment of unit 3 resulted in the reduction of the cumulative operating cost from R 2 105 511 to R 2 093 305 in the study horizon. By shutting down unit 3, the model is able to reduce the capital expenditure associated with the operating cost. In order to counteract the megawatt lost from unit 3's shutdown, the loading on unit 2 is increased. The unit's megawatt output increased from below 200 MWh at hour 21 in problem instance 3 to above 200 MWh in instance 4. The increase in load output is substantiated when evaluating the trends for unit 2 in Figures 5.3 and 5.4 respectively. The loading of the three thermal units prior to hour 21 remained relatively constant with only small changes in the load output being noted.

5.3.5 Ramp rate limits

To evaluate the effect of the ramp rate limit constraints on the commitment and load dispatch schedule, (4.30) to (4.33) are added to the model discussed in problem instance 4. Input Parameters $r_h^{(u)}$ and $r_h^{(d)}$ in (4.31) and (4.33) are assigned ramp rate values of 100 MW/h whereas parameters $r_h^{(us)}$ and $r_h^{(ds)}$ are allocated values of 150 MW/h. The mentioned ramp rate values are however only incorporated into the model for explanatory purposes with the aim of determining the effect of the constraints on the model results. Realistically, the ramp rates may vary from 100 MW/h to 900 MW/h depending on the technology used. Refer to Figure 5.5 for a graphical representation of the model results, after incorporation of the ramp rate limitations.

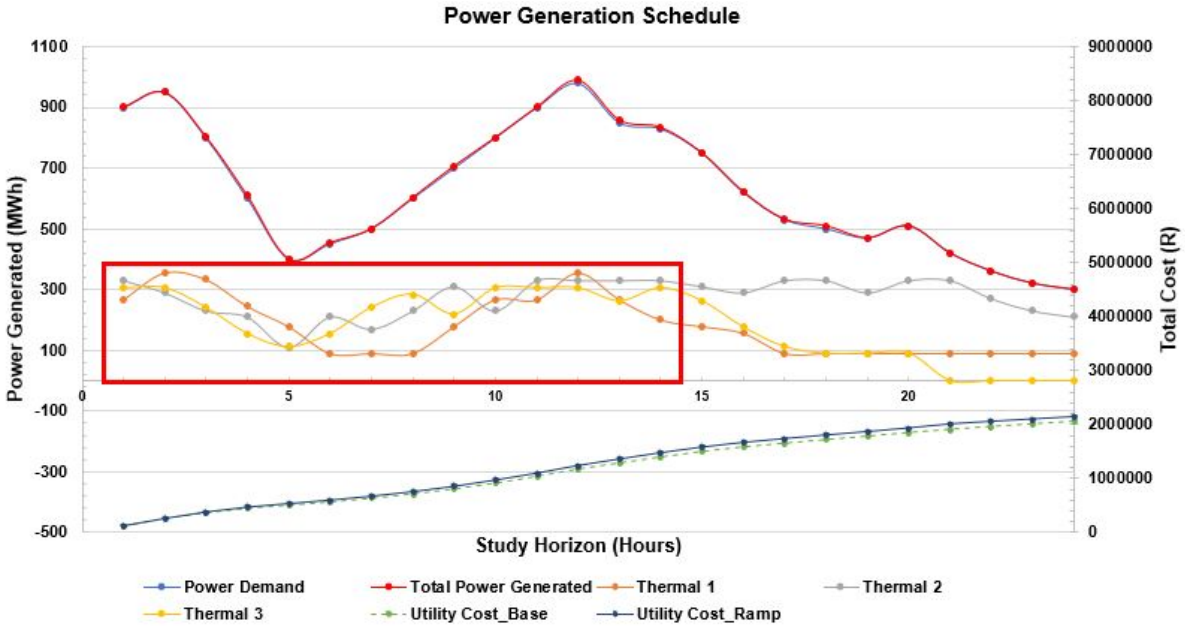

Figure 5.5: Ramp rate limit influence on model results

From the results portrayed in Figure 5.5, it can clearly be noted that the inclusion of the ramp rate limitations affects both the unit loading and the cumulative operating cost. The rate at which the megawatt loading is allowed to change for thermal units 1 to 3 in time period 1 to 14, results in a smoother load profile curve. For example, unit 1 in Figure 5.4 is allowed to reduce its load from 112 MWh at hour 4 to 90 MWh in hour 5. After incorporating the ramp rate limitations, the unit is prohibited to deload at such a fast pace and only reaches 90 MWh in hour 7 of the study horizon as seen in Figure 5.5. The same principle applies to the remaining units and therefore the loading schedule is significantly altered. The loading of unit 2 is reduced at the early stages of the study horizon due to the delayed response of the other two units in reducing its output to 90 MWh. As mentioned previously the delayed response is induced by the ramp rate limit constraints. The cumulative cost also increases from R 2 093 305 in problem instance 4 to R 2 127 382 in instance 5 as the units are forced to operate closer to reality.

5.3.6 Multi-area power flow functionality

As outlined in Chapter 4, multi-area power flow is the process of power importation from a neighboring country or exportation by a resident country. In order to incorporate the multi-area power flow functionality into the optimisation model depicted in Section 5.3.5, term $P^{(F)}$ in (4.5) as well as (4.43) and (4.44) are included into the model. For the purpose of this problem instance, an allowable power flow of 50 MWh was specified from area 1 (resident area) to area 2 (neighboring area). Power flow from area 2 to area 1 was prohibited. The preceding is specified in the model by means of variable b. Identical to that of area 1, area 2 consist of 3 thermal units (numbered as unit 4 to unit 6). Therefore, problem instance 6 comprise of a 6 thermal unit problem with 3 units being

assigned to each area. Depending on the coal and fuel oil cost consumed per area, the model will either export or import power to satisfy the power grid demands of both areas in the most optimal and cost effective way possible. However, given the restrictions specified by the author, the model will only allow power flow in one direction and discount the above mentioned. The results obtained from the optimisation model for area 1 and 2 are depicted in Figures 5.6 and 5.7 respectively with the data of interest being marked in the red blocks.

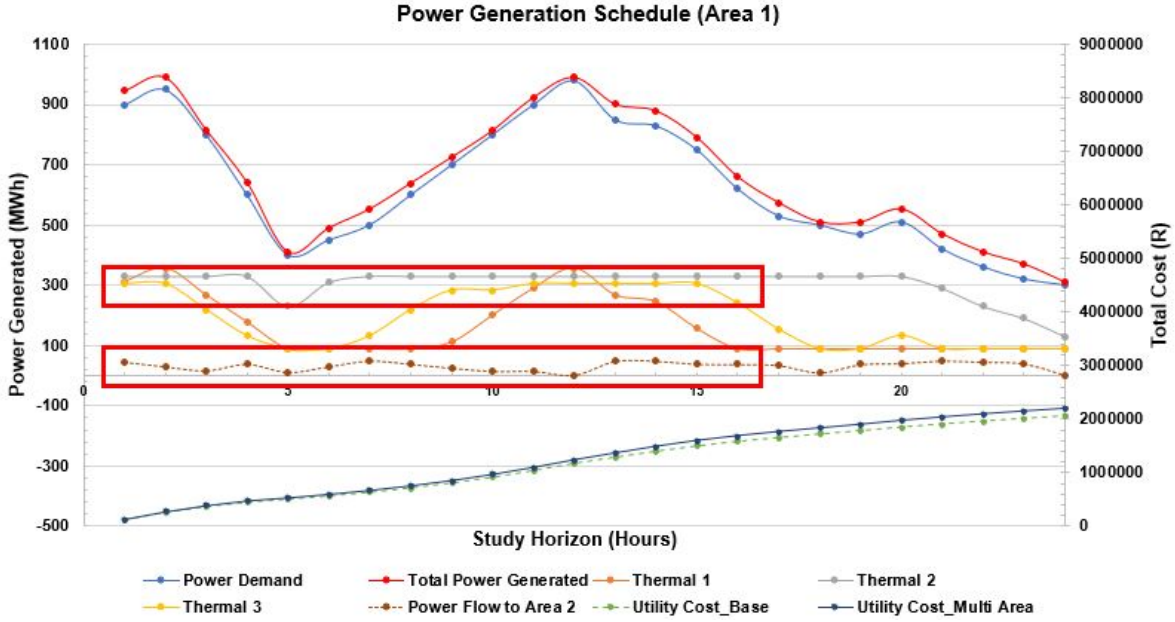

Figure 5.6: Multi-area power flow area 1, influence on model results

In analyzing the results obtained for area 1, it is apparent that the multi-area power flow constraints included in the model, forced an increase in the loading of the units in area 1. The preceding statement can be substantiated by comparing the results obtained from problem instance 6 with that of instance 5. In Figure 5.6, unit 2 is dispatched at its maximum allowable load of 330 MWh for majority of the time, much higher than its previous loading. The output of unit 3 has increased during the early hours of the study horizon with the unit not being shutdown after hour 21, as was seen in instance 5. Instead, unit 3 is loaded at 90 MWh from hour 21 to 24. The optimisation model decided to export power to area 2 as a result of the power flow link specified by the user. The power exportation from area 1 to area 2 does not exceed a value of 50 MWh for the instance under consideration (emphasized in the bottom red block of Figure 5.6). The operating cost incurred for area 1 accumulated to a value of R 2 127 382, R 63 527 higher than the previous value of R 2 191 010. The increase in operating cost can be attributed to the additional power that had to be generated for the power exportation to area 2.

Figure 5.7 is representative of the commitment and dispatch schedule generated by the optimisation model for area 2. From the results portrayed in this figure, it is apparent that thermal unit 5 is the most expensive to dispatch and therefore power importation is rather considered. It is only at hour 7 when the model decides to start unit 5 and dispatch the unit to 120 MWh in preparation for the peak demand. After hour 19, both units 5 and 6 are shutdown as a result of the reduced load demand. Note however that during these hours power is still imported from area 1 to satisfy the grid demand. The power production from unit 4 is less expensive than the other units and therefore the unit is loaded at close to full load throughout the study horizon. It is only during low demand periods where the unit is deloaded to an output of between 170 MWh to 220 MWh. The cumulative operating cost incurred by the utility in area 2 equates to a value of R 2 518 119. Note that although power is imported to reduce the utilities' operating cost, the cost is still higher than that of area 1 as a result of the high rate of coal consumption to which the units are subjected to. High coal consumption can be credited to plant inefficiencies or low coal qualities.

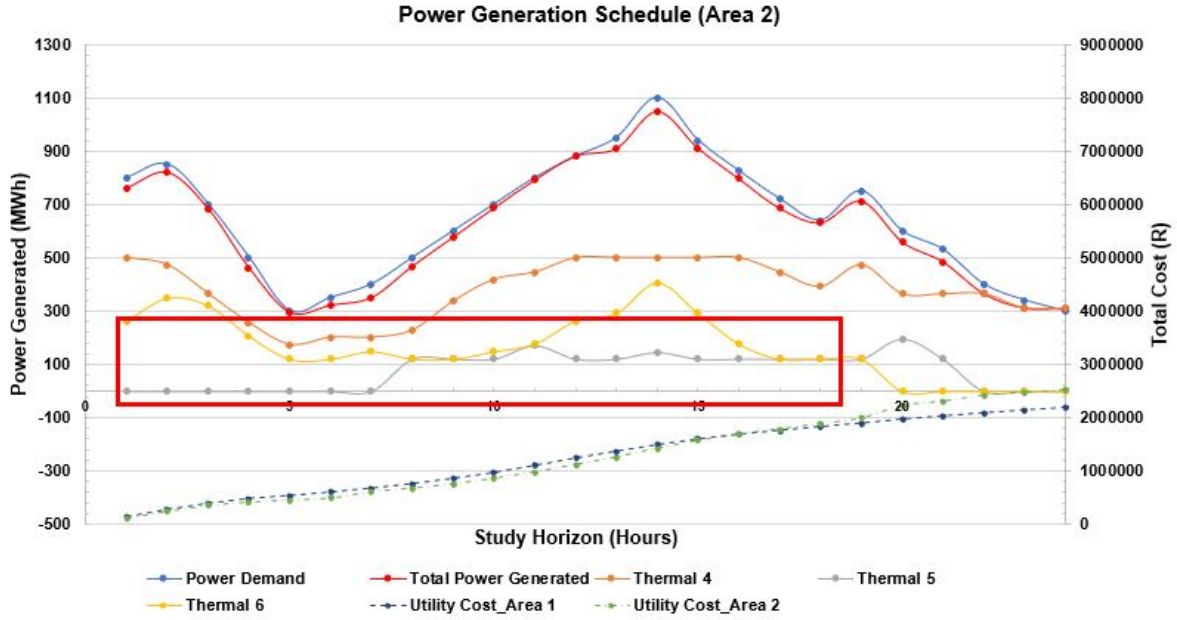

Figure 5.7: Multi-area power flow area 2, influence on model results

5.3.7 Outage schedule

To investigate the influence of unit downtime on the computational results, Equation 4.40 is added to the model. The user defined input parameter $o_h^{(t)}$ is assigned values of 1 for units 2 and 3 in area 1 as well as for units 4 to 6 in area 2. The only unit that is excluded from the computational process with the purpose of simulating an outage opportunity, is unit 1 in area 1. To force unit downtime, parameter $o_h^{(t)}$ is assigned a value of 0 for unit 1. Figures 5.8 and 5.9 depict the influence unit downtime may have on the model results. Note that in this example the allowable power flow from area 1 to area 2 was specified at 500 MWh and vice versa. The reason why this was done was to allow the model to obtain a feasible solution. If the power flow was maintained at 50 MWh from area 1 to area 2, the units in area 1 would not be able to satisfy the grid demand. Consequently, the model would provide an infeasible solution as feedback.

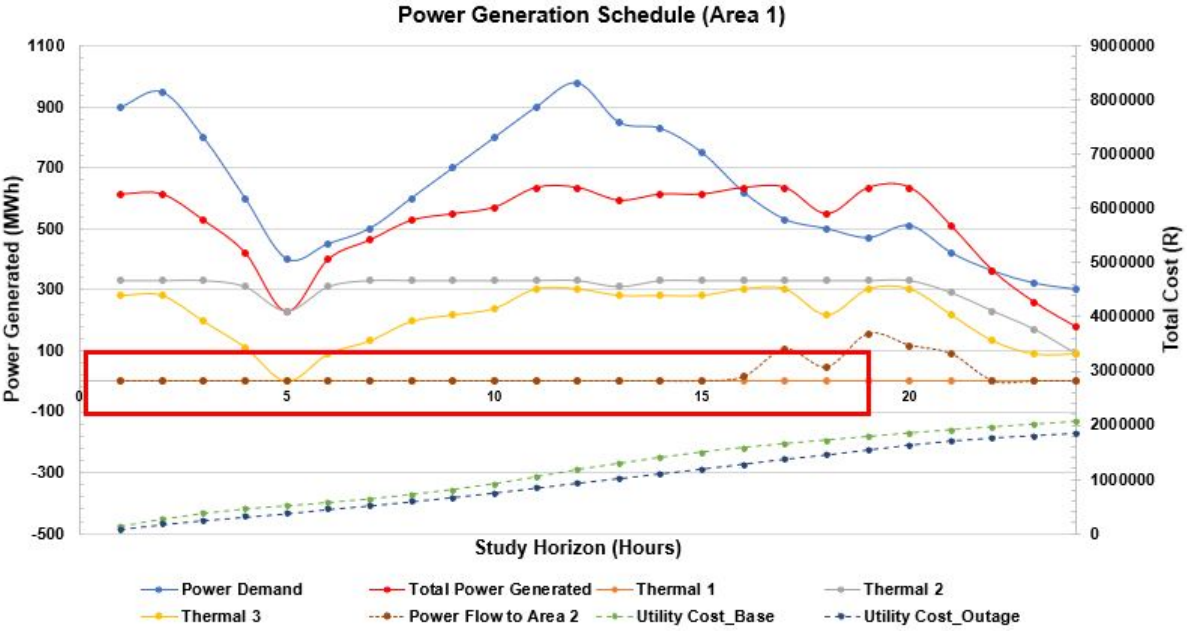

Figure 5.8: Outage schedule added to area 1, influence on model results

The results portrayed in Figure 5.8 provides a clear indication of the outage constraint influence. In the red block unit 1 is shown to be scheduled at 0 MWh for the study horizon as it is booked for an outage. Given that unit 1 is not available, unit 2's loading is still maintained at 330 MWh for majority of the time. The power output of unit 3 is also increased in order to account for the reduction in unit 1's output. Note however that the two units of area 1 are incapable of supplying enough power to satisfy the grid demand as seen when comparing the power demand and generated curves in Figure 5.8 with one another. The power generated is significantly less than that of the demand. This is indicative of a power generation resource shortage in the area which will lead to the power demand not being satisfied. It is for this reason that a significant amount of power is imported from area 2, although it is more expensive, to satisfy the grid demand of area 1. The quantity of power imported can be seen in Figure 5.9. The cumulative cost incurred by the utility in area 1 is significantly less than that of the previous problem instance as a result of unit 1's downtime. The area's operating cost equates to a value of R 1 821 875.

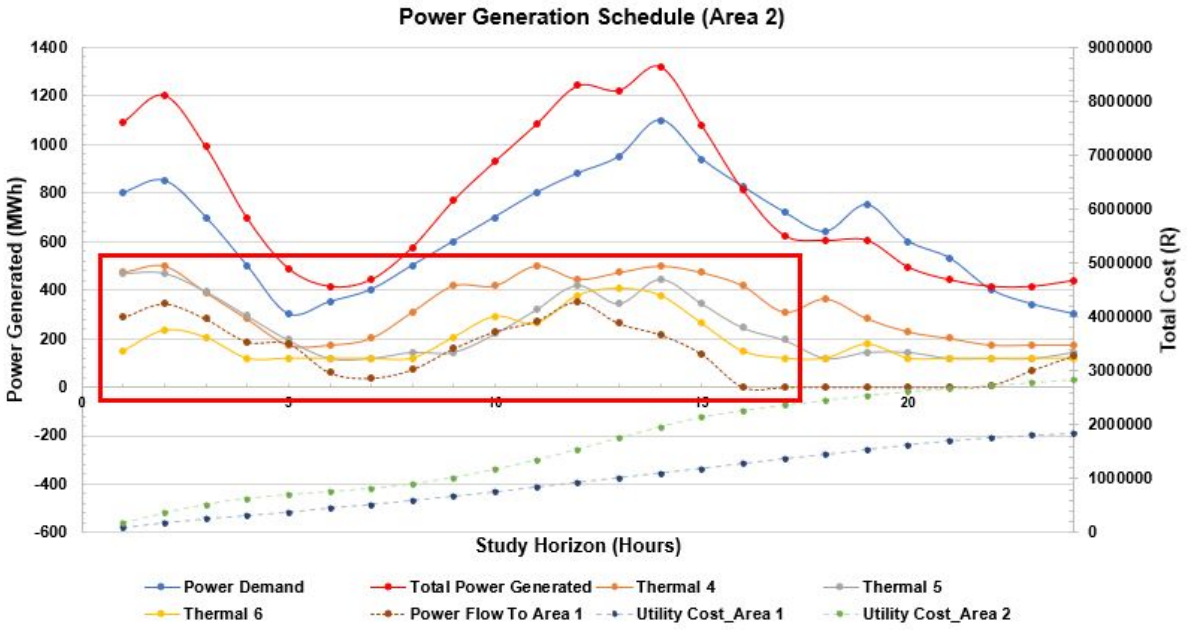

Figure 5.9: Reaction of area 2 to outage schedule, influence on model results

The loading of area 2's units are significantly higher in comparison to the previous problem instance as these units are required to supply area 1 with enough power to meet the grid demand. Analyzing the power generated versus demand curve, it is apparent that more power is being generated by the units in area 2 than what is required. The difference being the power which needs to be exported to area 1. The cumulative operating cost incurred by area 2 equates to a value of R 2 840 452, which is significantly higher than the previous instance. The reason being that area 2 needs to generate additional electricity to accommodate the shortage in area 1.

5.3.8 Water consumption limitations

In problem instance 8, water consumption limitation constraints are added to the model in order to evaluate the effect these constraints have on the model results. Equation (4.4) is added to the objective function depicted in Chapter 4, with parameter $c_h^{(w)}$ representing the rand per hour value associated with the demineralised water consumed by each thermal unit. Equation (4.42) is also incorporated into the model with parameter $w_n^{(m)}$ being assigned a value of 0.12 ML/h per area. The purpose of (4.42) as mentioned in Section 4.9 is to prevent the various areas from exceeding its demineralised water usage targets by scheduling the units optimally. In this instance the outage limitations induced on unit 1 of area 1 is omitted to allow the model to determine which units to decommit from the grid in an attempt to prevent excessive water consumption. Figures 5.10 and

5.11 portrays the results obtained for areas 1 and 2 respectively after solving the water constrained optimisation problem. Note that the allowable power flow between areas 1 and 2 in the mentioned problem instance still remains at 500 MWh in both directions.

Adding the water consumption limitation constraints to the model resulted in a significant change to the commitment and dispatch schedule generated by the optimisation model. This is confirmed by comparing the results depicted in Figures 5.8 and 5.9 with that of Figures 5.10 and 5.11. In the previous problem instance unit 1 was decommitted from the grid as a result of the outage constraint. However, in the current instance unit 1 is again committed to the grid together with units 2 and 3. The units contained in area 1 are loaded at relatively high outputs as the model requires a significant amount of power to be exported to area 2. The reason being is because thermal unit 5 in area 2 was decommitted from the grid for the entire study horizon to prevent the area from exceeding the water usage limitations. Although two units are still active in area 2, they are incapable of supplying enough power to meet the grid demand during peak periods. Also note that power generation in area 1 has been identified to be much cheaper in comparison to area 2 and therefore power is imported from area 1 even during lower demand periods, as it reduces the overall operational cost incurred by the utility in area 2. The cumulative cost incurred by area 1 during the study horizon equates to a value of R 2 553 627. The operating cost for area 1 is significantly more in comparison to the previous instance as additional power generation is required in this area for the exportation to area 2. The preceding can be noted when evaluating the power generation versus demand curve in Figure 5.10.

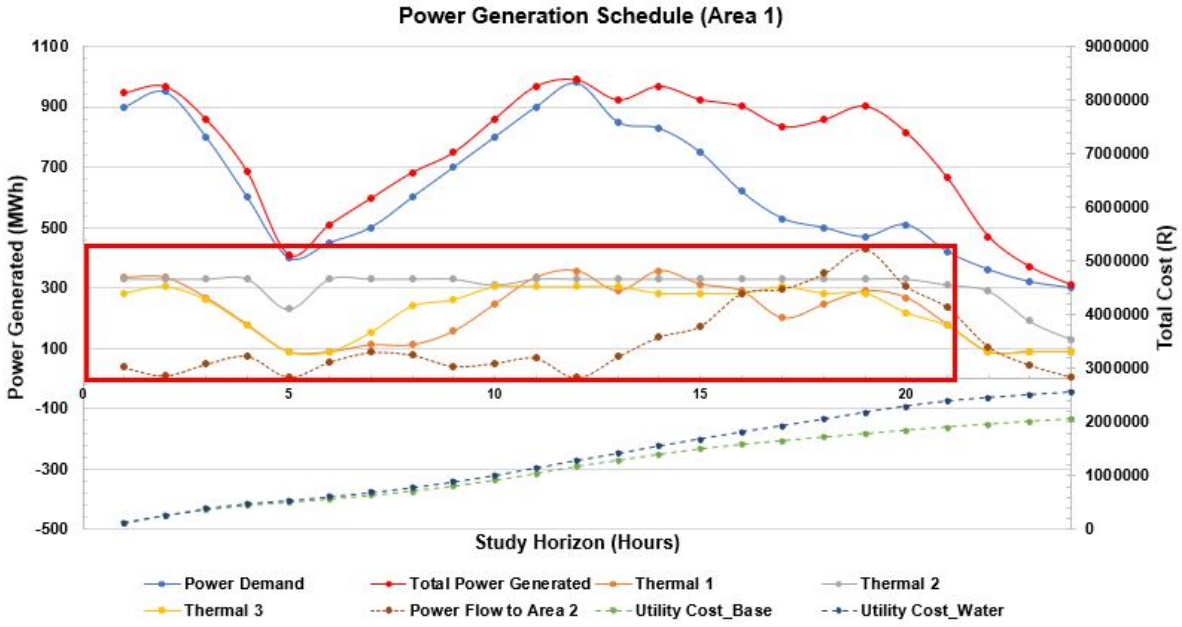

Figure 5.10: Water consumption limitations in area 1, influence on model results

Evaluating the data depicted in Figure 5.11, it is clear that the units in area 2 are relatively expensive to dispatch and therefore they are not utilised to their full capacity to meeting the power demand. During hours 4 - 10 and 18 - 24 unit 6 is loaded at its minimum allowable load. Unit 4 is maintained at similar loads for periods 5 - 7 and 18 - 24. The remaining power required to meet the demand is supplemented by importing power from area 1. The maximum amount of power that was imported from area 1 to 2 in a single hour equated to a value of 430 MWh. From the data portrayed in Figure 5.11, it has been concluded that unit 5 was decommitted from the grid to prevent area 2 from exceeding the water usage limitations. Note that unit 5's water usage is not the most severe when analyzing the water consumption of the units contained in area 2. Therefore, the unit was not only shutdown as a result of its water usage. The model evaluates fuel cost, water consumption, and fuel oil usage simultaneously as well as emissions performance to determine which unit to decommit. In this case unit 5 was the unit which had to be shutdown as to optimise the overall model results.

The cumulative operating cost incurred by the utility in area 2 for the current problem instance is equivalent to R1 829 146 which is much less than the R 2 840 452 reported in the previous instance. The reason why the cost reduced significantly from problem instance 7 to 8 is because of the decommitment of unit 5 in area 2 as well as the power that is imported from area 1 to area 2. Note however that if the water usage of the available units exceed the limitations set out by (4.42), and the capacity of the units does not allow the decommitment of a unit as there will not be enough power to meet the grid demand, the model will not be able to provide a feasible solution and the computational process will be terminated.

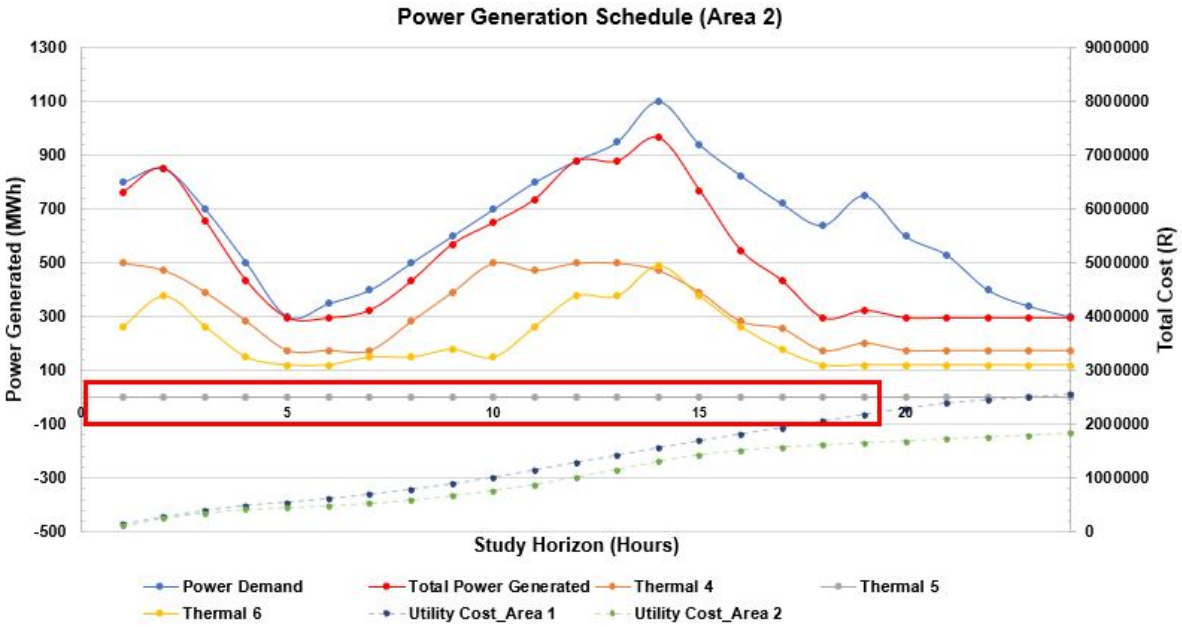

Figure 5.11: Water consumption limitations in area 2, influence on model results

5.3.9 Spinning reserve functionality

The spinning reserve functionality is incorporated into the optimisation process by means of adding (4.10) - (4.13) to the model, with term 2 of (4.11) ($s^{(p)}_{ptsu}$) being omitted. The reason for the exclusion is because the term is representative of the reserves required by pumped storage units. The preceding will however be taken into consideration in Section 5.3.11. Parameter $g^{(d)}$ in (4.10) is assigned a value of 5%. This parameter ensures that the amount of operating reserves maintained by the available thermal generating units equate to a value of 5% of the total power demand for the affected area. Note that only the reserve schedule results have been reported in this section to portray the model's capability of tracking the reserve availability. Refer to Figure 5.12 for a detailed representation thereof.

In analyzing the reserve schedule depicted in Figure 5.12, it is apparent that the model is capable of tracking the available reserves throughout the study horizon. The reserve quantities are switched between the various units in both areas 1 and 2 in order to ensure that enough reserve capacity is maintained. The quantity of reserves required per hour will fluctuate with the change in power demand as it is defined as a percentage of the total demand. For example, in hour 7 a reserve of 25 MWh needs to be maintained whereas in hour 14 it increases to 42 MWh and so on. Note that the model does not select one single unit to manage the operating reserves, but rather switches the responsibility between the various units in an attempt to optimise the commitment and dispatch schedule.

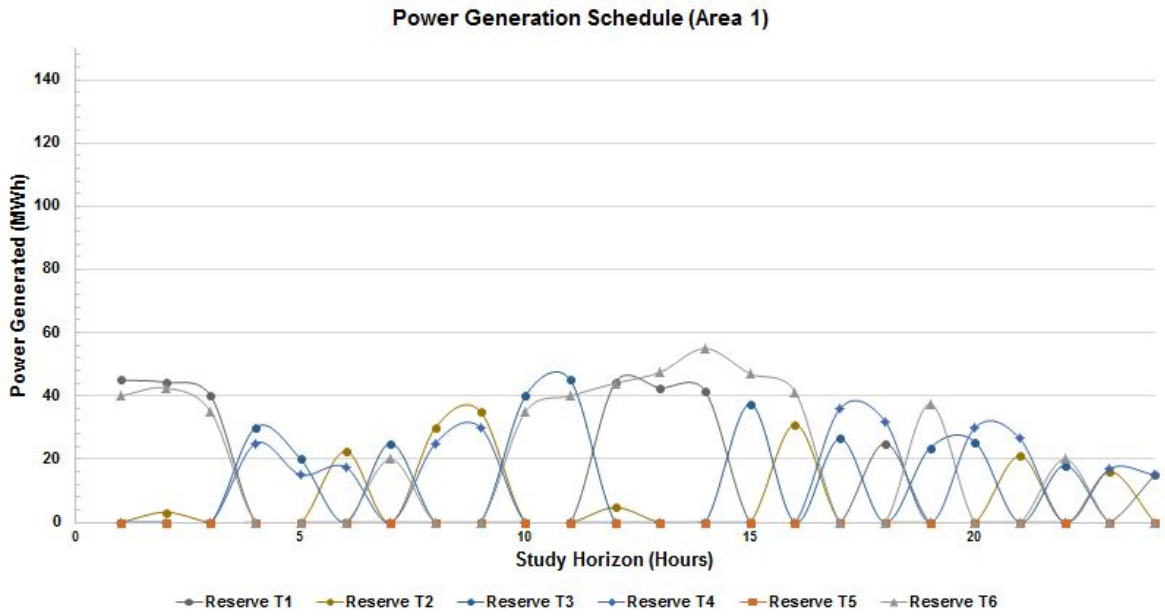

Figure 5.12: Spinning reserve requirements in area 1, influence on model results

5.3.10 Hydro generating units

Problem instance 10 entails the inclusion of a hydro generating unit to area 2. The foregoing is achieved by adding (4.45) and (4.46) to the model as well as the necessary input data into the excel acquisition tool. The prohibited operating time for the hydro unit as mentioned in Section 4.11 was specified from hour 1 to 7 in (4.45), with the unit being allowed to be committed to the grid after hour 7. Variable $w_h^{(hm)}$ in (4.46) has been set equal to 62 m^3/s to prevent the hydro unit from exceeding its maximum allowable operating limits. The model results obtained after the inclusion of the hydro unit is depicted in Figure 5.13 for area 1 and Figure 5.14 for area 2. Keep in mind that thermal unit 5 in area 2 still remains decommitted from the grid as a result of the water limitation constraints added in Section 5.3.8.

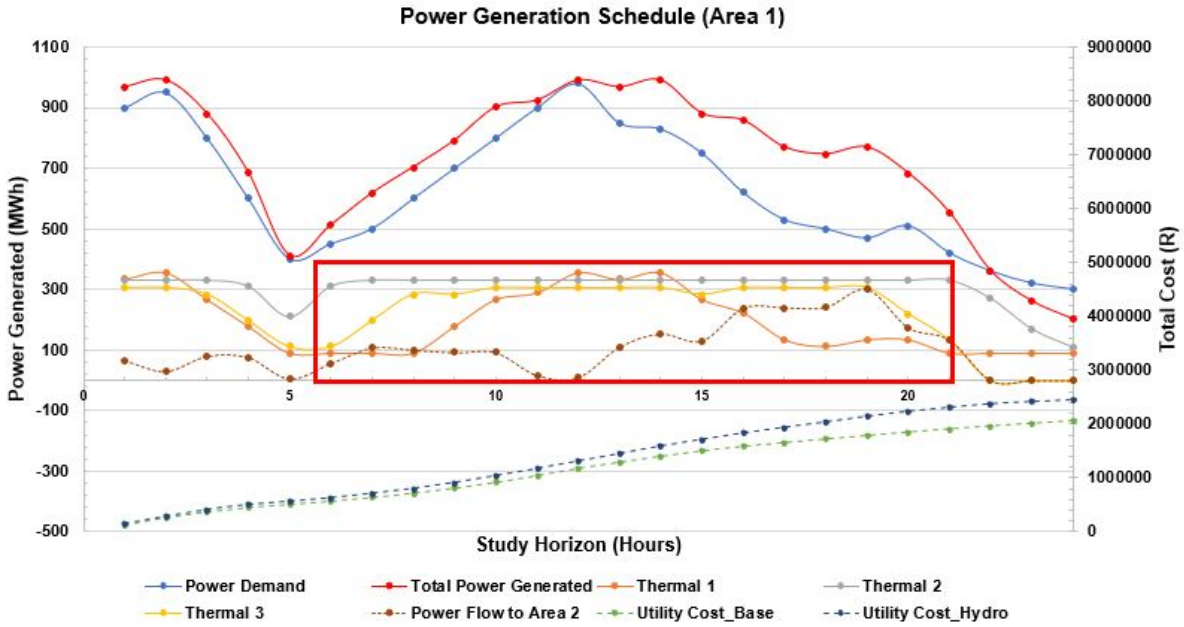

Figure 5.13: Reaction of area 1 to hydro unit inclusion, influence on model results

The loading of units 2 and 3 in area 1 did not change significantly with the inclusion of the hydro

unit in area 2. The only major change was the decommitment of unit 3 from the grid at hour 22 of the study horizon. After shutting down unit 3, the model requested power to be imported from area 2's hydro unit to supplement the capacity lost from unit 3. The hydro unit has no cost associated with its power generation and therefore it is beneficial to import power to area 1 using the hydro unit as source. Note however that power importation from area 2 to area 1 was only considered during the last hours of the study horizon where the grid demand in area 2 was low enough to make the power importation feasible. At low loads the power generation of thermal unit 6 in area 2 has been identified to be cheaper than that of unit 3 in area 1. The preceding does not hold true for high load conditions. From a cost perspective, it makes sense to rather shutdown unit 3, maintain unit 6 at minimum load and import power from area 2's hydro unit to substitute the power shortage in area 1, given low load conditions in area 2 and hydro unit availability. In the initial hours between 1 and 7, the power flow was directed from area 1 to area 2 as the hydro unit was not yet on load. It was only after hour 7 that the hydro unit was committed (hydro power generation portrayed in the red block of Figure 5.14). After hydro commitment the power flow was still directed from area 1 to area 2 to support the high power demand in area 2 and to counteract the cost of operating thermal units 4 and 6 at high loads. As a result of the hydro unit being available for generation, the loading of unit 1 was seen to reduce over selected areas of the study horizon as the power required to be exported to area 2 diminished. The power exported to area 2 reduced from a maximum of 430 MWh in the previous problem instance to 300 MWh. The maximum allowable power which the hydro unit was able to generate was equivalent to a value of 106 MWh. The cumulative operating cost incurred by area 1 equated to a value of R 2 432 626 which is less than the cost mentioned in Section 5.3.8 for area 1. A reduction in the operating cost was expected as a result of the decommitment of unit 3 and the power imported from area 2 during the last hours of the study horizon.

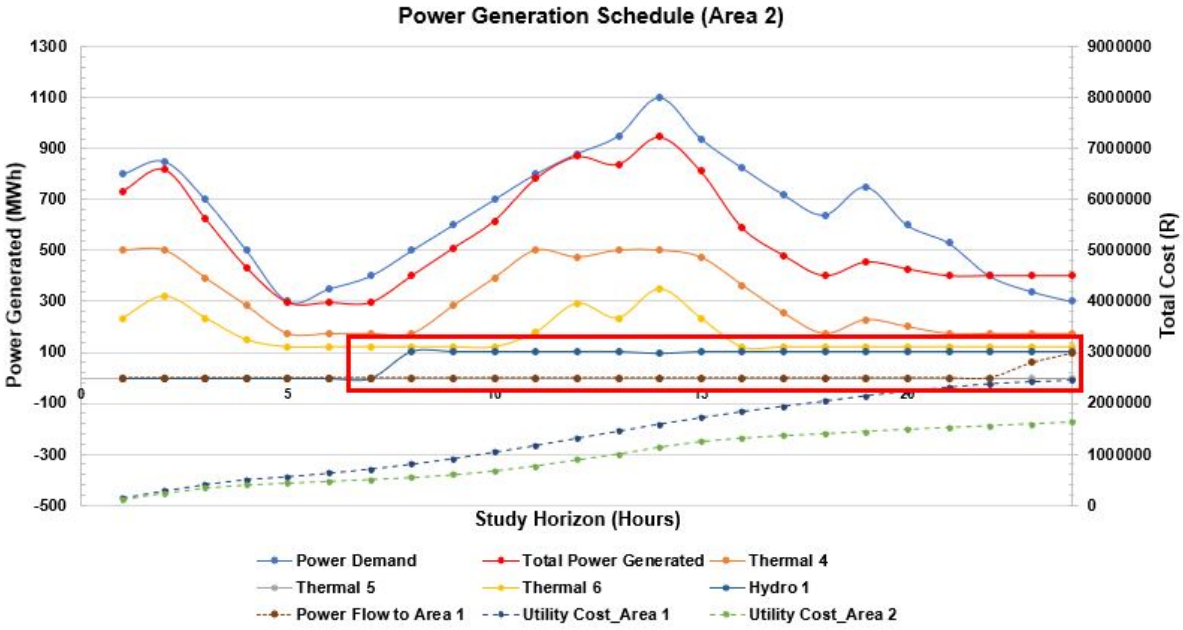

Figure 5.14: Adding a Hydro unit in area 2, influence on model results

Except for the last few hours (16 - 24) of the study horizon where unit 6's load remained constant, the unit's load also reduced, shortly after commitment of the hydro unit to the grid. The reason for unit 6's load reduction, is because it is cheaper to substitute the power generated by a thermal unit with that of a hydro unit. The loading of unit 4 remained relatively unchanged in comparison to the previous problem instance. The cumulative operating cost associated with area 2 decreased to a value of R 1 622 709 which is less than the operating cost stated in problem instance 8 of Section 5.3.8. The reduction in operating cost can be attributed to the inclusion of the hydro unit in area 2.

5.3.11 Pumped storage generating units

The aim of problem instance 10 is to determine the effect, the inclusion of pumped storage units will have on the unit commitment and load dispatch schedule. For the purpose of the problem instance, only two pumped storage units are added to the model with one being assigned to area 1 and the other to area 2. By adding the foregoing units, the model is expanded to a 6 thermal, 1 hydro and 2 pumped storage unit optimisation problem. Similar to the unit commitment constraints, the interconnectivity of the pumped storage constraints requires all the equations to be lumped together into the model simultaneously in order to evaluate its influence. Individually adding the pumped storage constraints will lead to infeasible and erroneous results being obtained. To include the pumped storage units into the optimisation model, the following equations had to be added to the Cplex coding:

1. The power balance constraints which includes term 2 ($p_{ptsu}^{(p)}$) of (4.5) as well as (4.8) and (4.9) with input parameter $p_{su}^{(w)}$ in (4.8) being set equal to 30 MWh for each unit.

2. Spinning reserve constraints comprising of term 2 ($s_{ptsu}^{(p)}$) in (4.11) and (4.14) to (4.15) with parameter $r_{su}^{(u)}$ being assigned a value of 100 MW/h for each unit.

3. Unit commitment constraints including (4.23) to (4.29), with the outage input parameter $o_{su}^{(p)}$ being assigned a value of 1 for both pumped storage units.

4. Ramp rate capability constraints consisting of (4.35) to (4.37) where parameters $r_{su}^{(u)}$, $r_{su}^{(us)}$, $r_{su}^{(d)}$ and $r_{su}^{(ds)}$ being set equal to 100 MW/h, 150 MW/h, 100 MW/h and 150 MW/h for each unit respectively.

5. Prohibited operating region constraints are added by means of (4.39), with no pumped storage unit being subjected to forbidden regions in the current problem instance.

6. Lastly, water balance and reservoir volume flow constraints are added to the model by means of (4.47) to (4.58).

The results obtained from solving the optimisation model specified in problem instance 10, is portrayed throughout Figures 5.15 to 5.17 with the data points of interest being encompassed in the red blocks. The power flow between areas 1 and 2 was maintained at a value of 500 MWh in both directions and unit 5 in area 2 remains decommitted as a result of the water usage limitations. The power generated by the pumped storage units as well as the upper and lower reservoir volumes is depicted in Figure 5.17 to prevent data cluttering in Figures 5.15 and 5.16.

The addition of the pumped storage unit in area 1 resulted in a reduction in both the thermal unit loading (at certain time intervals) and the cumulative operating cost incurred by the utility in area 1. When comparing the model results portrayed in Figure 5.13 with that of Figure 5.15, the foregoing statement is substantiated. Knowing that the pumped storage unit in area 1 supplied additional power to the grid, the model decided to decommit unit 3 at hour 19 instead of 22 as seen in the previous problem instance. The quantity of power exported from area 1 to area 2 also reduced as a result of area 2 receiving loading support from the pumped storage unit added in that area. Although the power flow from area 1 to area 2 decreased, there is still a significant amount of power being exported to provide loading support and reduce the operational cost incurred by area 2's thermal units.

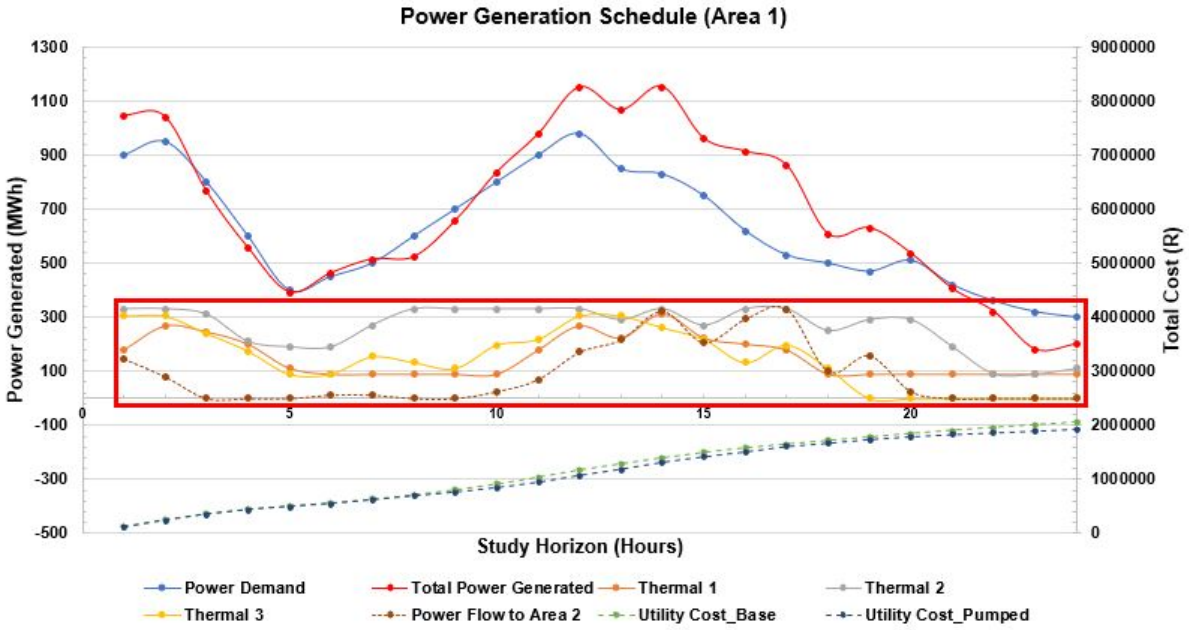

Figure 5.15: Pumped storage unit added to area 1, influence on model results

Combining the influence of the power exportation to area 2 and the addition of the pumped storage unit loading support, the model is able to reduce the power output of the thermal units in area 2 significantly. The foregoing can be noted in Figure 5.16. Note that the loading of the thermal units in area 1 did not decrease significantly in the early hours of the study horizon as the pumped storage units were mostly operating in pumping mode. The reason being was to ensure that the pumped storage units achieve its upper reservoir volume setpoint at hour 5 of the study horizon. During the early hours, power was also imported from area 2 and exported to area 2 in order to provide sufficient power supply to the pumped storage units while operating in pumping mode. After incorporating the pumped storage unit to area 1, the cumulative operating cost reduced from R 2 432 626 to R 1 911 656 as a result of the negligible cost associated to the pumped storage power generation.

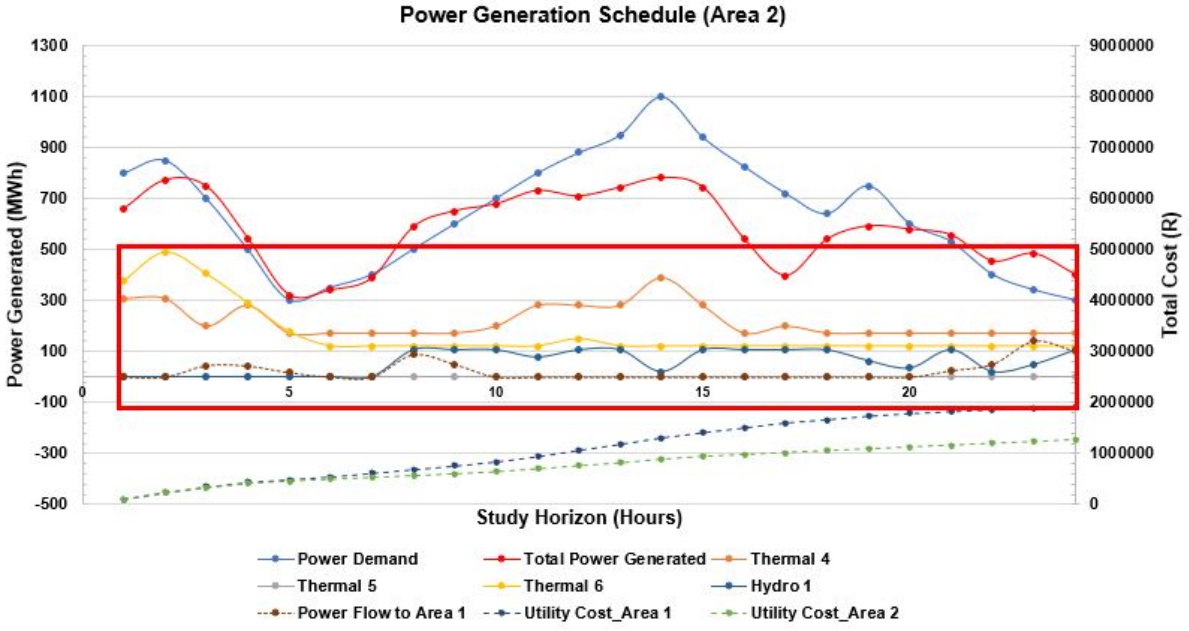

Figure 5.16: Pumped storage unit added to area 2, influence on model results

Comparing the results depicted in Figure 5.14 with that of Figure 5.16, it is clear that the loading of both units 4 and 6 in area 2 reduced significantly after adding the pumped storage unit to area 2.

During hours 6 - 9 and 16 - 24 both the thermal units in area 2 were loaded close to the minimum allowable power generation which corresponds to 120 MWh for unit 6 and 175 MWh for unit 2. The hydro unit in area 2 was also dispatched at 105 MWh in hour 8, similar to the previous problem instance, with the load of the unit fluctuating to maintain reserve power at certain hours of the study horizon. Comparable to the pumped storage unit in area 1, the unit in area 2 was mostly operated in pumping mode during hours 1 - 5 to ensure that the upper reservoir volume setpoint at hour 5 was satisfied. After hour 5, the pumped storage unit was dispatched at varying loads in order to regulate the grid demand. When relating the cumulative operating cost of the current and previous instance with one another, it is apparent that the cost reduced from R 1 622 709 to R 1 274 078 as a result of the pumped storage units being able to provide cheap electricity. Note that the cumulative operating cost of area 2 is cheaper than that of area 1 as a result of the additional power that had to be generated by area 1 for the exportation to area 2. Refer to the power demand versus power generated curves in both Figures 5.15 and 5.16 to evaluate the above statement. In Figure 5.15, the power generated is generally higher than the power demand whereas in Figure 5.16 the opposite is true. The mentioned is indicative of the power flow dynamics between the two areas.

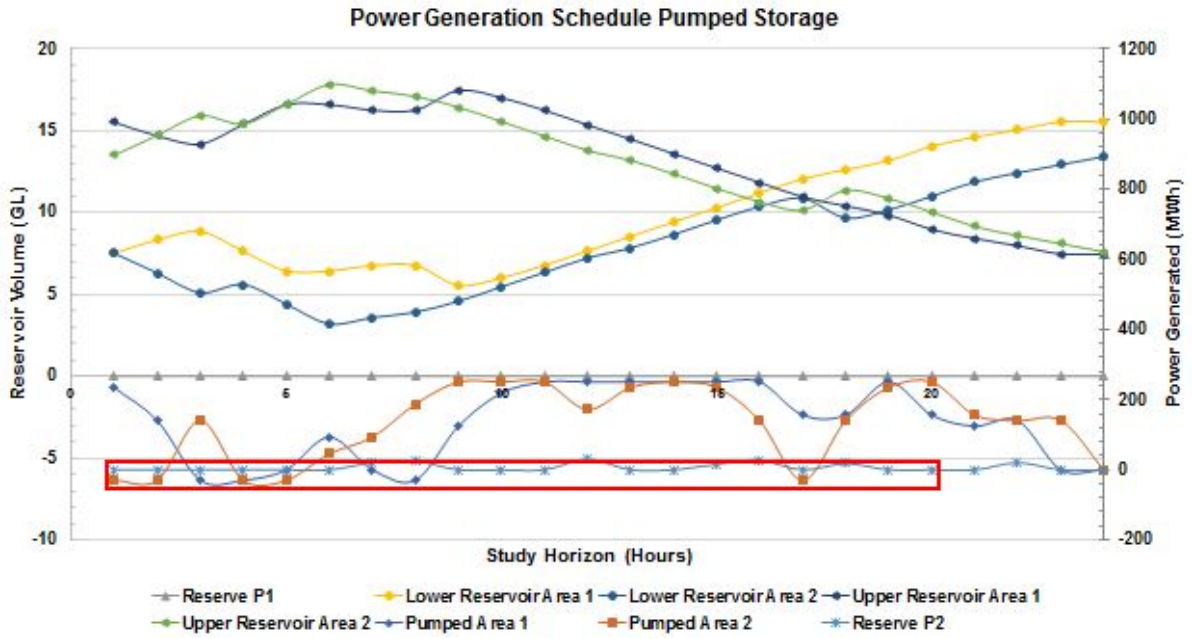

Figure 5.17: Pumped storage unit dynamics

Figure 5.17 represents the dynamics of the pumped storage units for the period under consideration. Given the data portrayed in the above figure, it is clear that the volume flow of water to and from the upper and lower reservoirs are accurately tracked by the model. From the above results, one can attain that the upper reservoirs' levels increased from hour 1 to 5 (when the units operate in pumping mode) as a result of the level setpoints provided to the model. The upper reservoir level setpoint of reservoir 1 and 2 were each set at 16.5 Giga Liters (GL) and had to be reached at hour 5. The contrary is true for the bottom reservoirs as water was extracted from the lower reservoirs to fill the upper ones. After hour 5, the units were switched to generating mode as seen in the load profiles and the upper levels started to drop. From hour 6 to 24, the upper reservoir levels slowly drop, with short periods of pumping being allowed to assist with load regulation during low demands and preventing the reservoirs from dropping below its minimum allowable volume limits. Note that the periods of pumping is emphasized in the red block.

5.3.12 Stochastic variable incorporation

Stochasticity is incorporated into problem instance 12 by means of adding two arrays to index p for variable $p_{ptn}^{(d)}$, with the aim of modeling variability in the grid power demand. Input parameter

$q_{ph}^{(f)}$ is multiplied with $c_{hj}^{(f)}$ to induce fuel cost variability whereas $q_{ph}^{(e)}$ is multiplied with $e_{hj}^{(f)}$ to account for possible emissions production changes. For the purpose of evaluating the influence of parameter stochasticity, two stochastic scenarios are considered where the power demand, fuel cost and emissions production data are adjusted. The two scenarios are incorporated into the model by means of setting index p equal to a value of 2. The initial scenario assumes that the mentioned parameters remain constant as was inserted into the excel acquisition tool. The second scenario considers a change in the forecasted power demand by a fraction of 0.8. The fuel cost and emissions production variables are augmented with fractions as depicted in Tables 5.7 and 5.8 respectively. By increasing the number of stochastic scenario's, the model complexity is seen to increase exponentially. It is for this reason that only two scenarios are considered to maintain the model simplicity and ensure ease of data interpretation. Only the results attained for stochastic scenario 1 is presented in this section as it can be compared to the results obtained from the model in Section 5.3.11. Data comparison is required to evaluate the effect model stochasticity has on the unit commitment and dispatch schedule. Providing the results obtained from stochastic scenario 2 will not contribute to the model verification process and therefore it is omitted from this section.

Table 5.7: Fuel cost stochasticity

Scenario	Area	Unit	Fraction of fuel cost change ($q_{ph}^{(f)}$)
2	1	1	1.12
2	1	2	0.99
2	1	3	1.08
2	2	4	0.91
2	2	5	0.74
2	2	6	1.05

Table 5.8: Emissions production stochasticity

Scenario	Area	Unit	Fraction of emissions production change ($q_{ph}^{(e)}$)
2	1	1	1.09
2	1	2	1.21
2	1	3	1.00
2	2	4	0.72
2	2	5	1.04
2	2	6	0.89

To ensure that the results obtained from the stochastic model can be compared with the previous models, no changes were made to the input parameters and constraints. This means that the changes identified in the unit loading schedule can only be attributed to the consideration of the stochastic influence. In this problem instance the model first needs to determine which units to commit prior to dispatching the units for each stochastic scenario. The model is not allowed to change the unit commitment decision during each stochastic scenario, but needs to maintain the commitment decision fixed when loading the units. Therefore, the unit commitment schedule needs to be able to provide the optimal results while considering the variability of the different process parameters. The results obtained from the stochastic model is portrayed throughout Figures 5.18 to 5.20.

Evaluating the data depicted in Figures 5.15 and 5.18, it can be concluded that the unit commitment schedule remained unchanged for the thermal units in area 1. The optimisation model prescribed that all three thermal units remain on load for the entire study horizon. Although there were no changes to the commitment schedule, there were some alterations to the load dispatch schedule. The loading profile of thermal unit 1 remained relatively unchanged with slight changes during the peak demand periods. The power output assigned to unit 3 increased during hours 10 to 20 of the study horizon, with the unit being decommitted in hour 21 instead of hour 19 as depicted

in the previous problem instance.

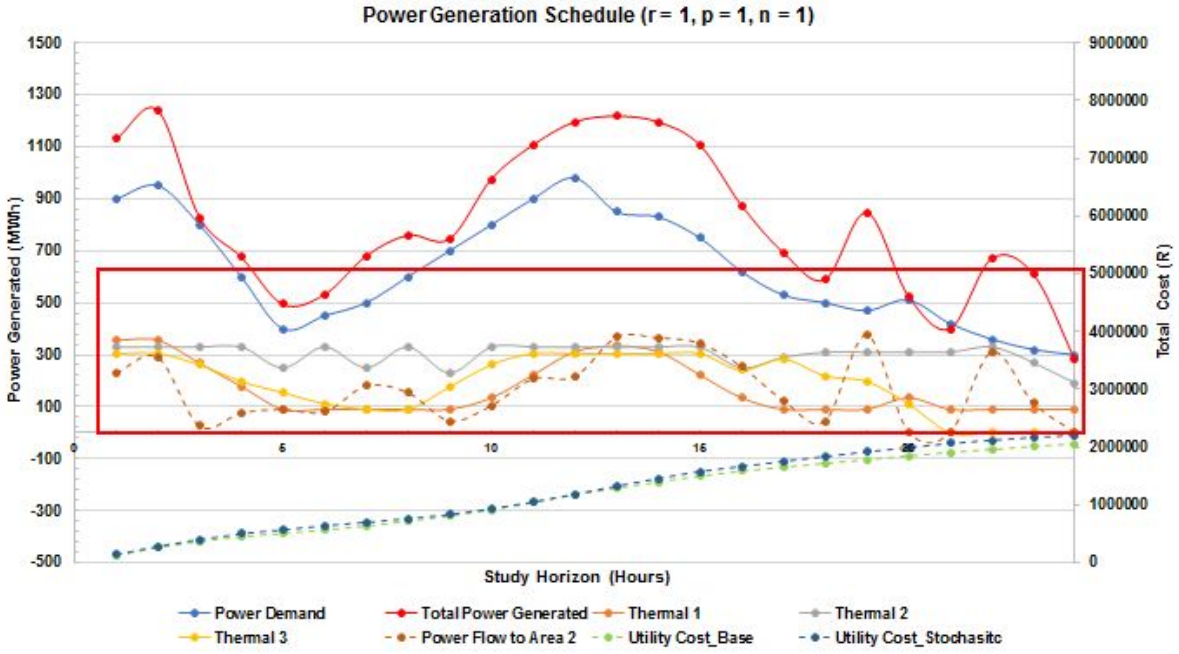

Figure 5.18: Stochastic variables added to area 1, influence on model results

The most significant change in the loading profile is noted for unit 2 where the unit is dispatched at close to full load operation for the entire study horizon. The power exported from area 1 to area 2 also increased significantly due to the cost associated with thermal power generation in area 2. In comparing the power exported in Figure 5.15 with that of Figure 5.18, it is apparent that additional power is required by area 2 in hours 3 - 9 and 22 - 24 of the study horizon. The cumulative operational cost of area 1 increased to a value of R 2 208 928 from the previous problem instance. The increase in operational cost can be linked to the additional power that is exported to area 2. The power dispatch of the pumped storage units in both areas does not have a significant impact on the loading of the thermal units in area 1, but does however result in the alteration of the load dispatch schedule in area 2.

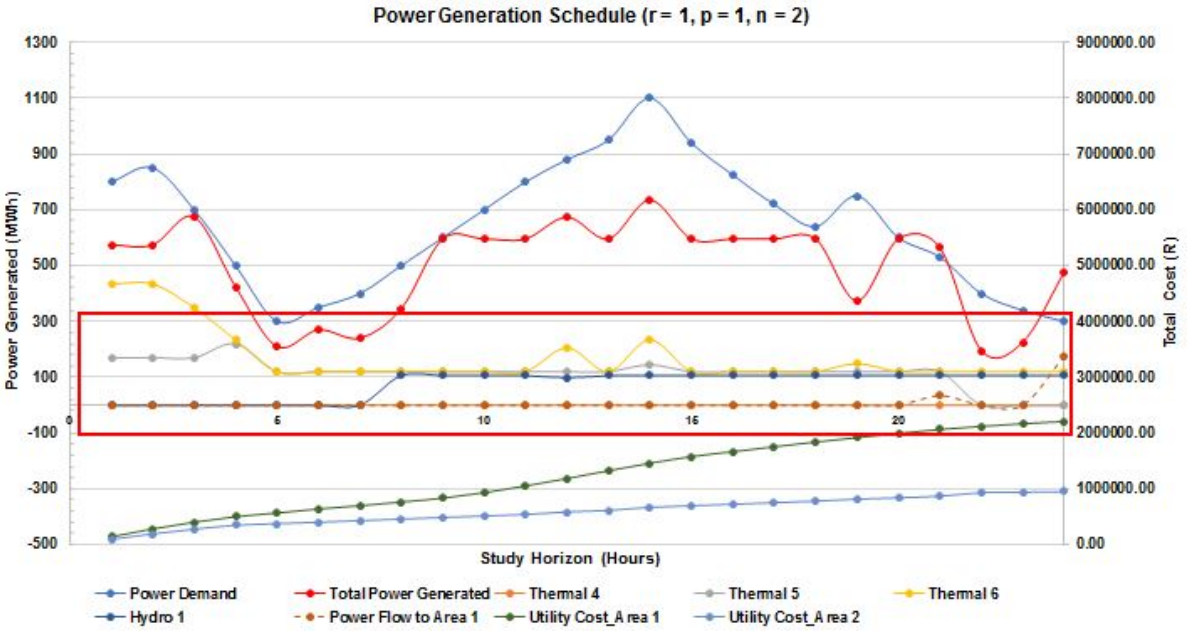

Figure 5.19: Stochastic variables added to area 2, influence on model results

Scrutinizing the results obtained from Figure 5.19 and comparing it to the previous model results as portrayed in Figure 5.16, it is noted that both the unit commitment and load dispatch schedules were changed as a result of adding a stochastic element to the model. In the previous problem instance unit 5 was decommitted from the grid whereas in the current problem instance the model decided to shutdown unit 4 and commit unit 5 instead. The reason why the specific decision making was made is because unit 4's fuel cost was only adjusted by a fraction of 0.91 in the second stochastic scenario whereas unit 5's fuel cost was reduced significantly with a fraction 0.74. The preceding infers that unit 4 will be able to generate power at a cheaper rate than that of unit 5 provided that the fuel cost stochasticity holds true on the day of generation. The loading dispatch schedule assigned to the thermal units in area 2 has also changed in the sense that the units were loaded at close to minimum generation for majority of the study horizon. It is only during the first 4 hours when the pumped storage units and hydro unit were not yet committed, where the loading of units 5 and 6 was maintained above minimum generation. Similar to the previous problem instance, the hydro unit was dispatch during hour 8 at a load of 106 MWh. As a result of the thermal units being maintained at very low loads in area 2, the cumulative operating cost dropped from R 1 273 078 in Section 5.3.11 to R 950 718. The drop in operational cost is maintained in area 2 by means of importing power from area 1.

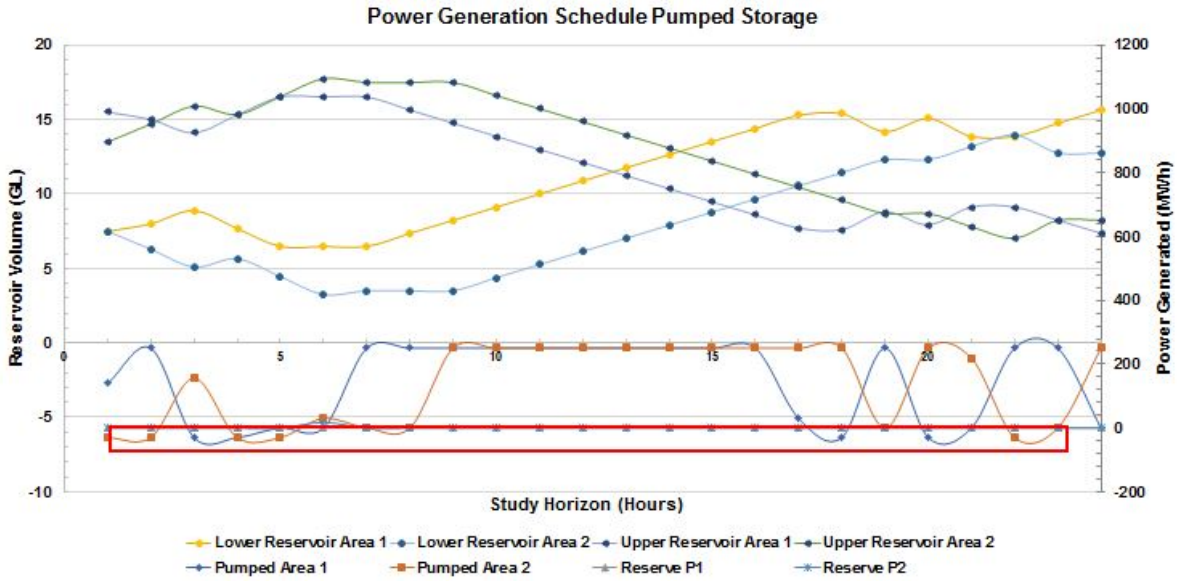

Figure 5.20: Pumped storage unit dynamics

Figure 5.20 provides information regarding the commitment and dispatch schedule developed for the pumped storage units contained in areas 1 and 2. Similar to the previous problem instance, the pumped storage units primarily operated in pumping mode during the initial hours of the study horizon (hour 1 - 6) in order to satisfy the upper reservoir setpoints. After hour 6, the pumped storage units were dispatched to approximately 250 MWh in order to support the power grid during the peak hours. The duration for which the pumped storage units were dispatched at 250 MWh increased in comparison to the data depicted in Figure 5.17. Note that only after hour 16, the model allowed the units to start switching between pumping and generating modes in order to assist with load regulation during the low grid demands and preventing the upper reservoirs from falling below its lower volume limits. When operating in pumping mode, it was noted that the upper reservoir volumes increased while the lower reservoir volumes decreased. The opposite was seen when the pumped storage units were operated in generating mode. Similar to the previous example, the periods of pumping is accentuated in the red block of Figure 5.20.

5.4 Model validation

The process of model validation entails the solution of two problem instances. The first problem instance is concerned with evaluating the model's performance by comparing the model results with the performance of an actual thermal power station. The purpose of this problem instance is to determine if the model will be capable of improving the overall power station's loading requirements in order to reduce capital expenditure and improve resource utilisation. For the purpose of the aforementioned problem instance, a 6 thermal unit problem is considered. The second problem instance consists of solving a realistic sized power grid problem containing 98 thermal, 8 hydro and 6 pumped storage units. The purpose of the second validation process is to determine the capability of the model in solving a large scale problem as well as the accuracy of the results obtained. To solve the comprehensive realistic sized power grid problem, constraints (4.1) to (4.58) as depicted in Chapter 4, are considered in the optimisation model. In conjunction with the preceding, the validation process also entails the analysis of the effect model scaling has on the solution time to determine if the proposed model is suitable for implementation in a production environment. Model scaling entails the process of expanding the optimisation problem by increasing the number of units or constraints and evaluating the increase in solution time as a result of the foregoing. In Section 5.4.2, 8 different model scaling problem instances were considered of which can be noted in Figure 5.25.

5.4.1 Comparing model results with actual thermal unit loading

In order to compare the model results with actual thermal unit loading requirements, a 6 thermal unit power station case study was considered. Given that only thermal units were considered for the mentioned problem instance, both hydro and pumped storage constraints were excluded from the model computations. The design generation capacity of the thermal units considered in this example can be noted in Figure 5.21 (green line) which equates to a value of 4116 MW. Although the station is capable of generating 4116 MW as per design, plant constraints and defect conditions prohibit the units from attaining full load operation resulting in the station only being able to supply a limited amount of megawatts. The preceding is classified as the available load the power station can supply and is portrayed by the red line in Figure 5.21. Plant constraints include aspects such as primary air heater leakages, ID fan constraints, and milling plant constraints. The blue line in Figure 5.21 depicts the load demand which the units are required to satisfy in the given problem instance.

The results obtained after solving the 6 thermal unit problem instance are portrayed in Table 5.9. Note that the model results are compared to the actual unit loading as determined by the system operator. When evaluating the loading requirements for each of the 6 units, it is apparent that there are some similarities between the model results and the actual unit loading. It can however clearly be noted that during certain time periods, the loading requirements as proposed by the model differs quite drastically from the actual unit loading. For example, when comparing the loading requirements for unit 2 at hour 1 of the study horizon, it is identified that the model proposed a unit loading of 675 MW, whereas the actual unit loading was set at 576 MW. Similar is true when comparing the results for unit 5, hour 1. The model proposed that unit 5 be loaded to a megawatt output of 386 MW instead of the actual loading of 558 MW. It is also apparent from the below data that when the units are loaded close to full load operation, the margin with which the model can optimise the unit loading decreases as all the units will be required to operate at full load to satisfy the load demand. The foregoing will result in the model results and actual unit loading not greatly deviating from one another. However, when the load demand decreases, the margin for optimising the loading requirements improves greatly. See Table 5.9 for a detailed comparison of the model results with the actual loading requirements as instructed by the system operator.

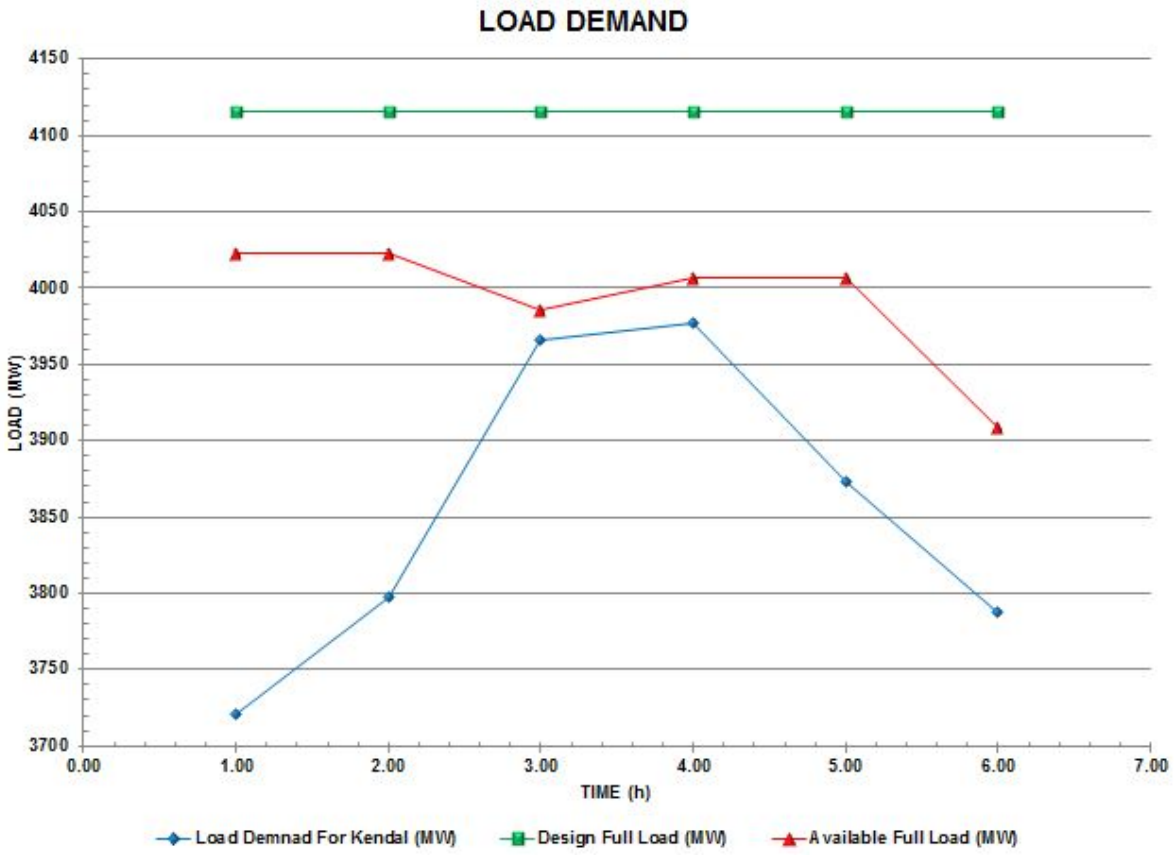

Figure 5.21: Load demand for first validation problem instance

Table 5.9: Model versus actual loading requirements

Hour	U1 Mod	U1 Act	U2 Mod	U2 Act	U3 Mod	U3 Act	U4 Mod	U4 Act	U5 Mod	U5 Act	U6 Mod	U6 Act
1	675	679	675	576	675	644	675	621	386	558	640	640
2	686	679	686	638	686	661	686	641	420	535	640	639
3	675	679	686	676	686	679	649	649	629	642	640	638
4	686	679	686	676	686	679	649	643	629	659	640	638
5	686	679	686	667	686	667	686	632	527	628	640	598
6	644	678	686	671	686	669	649	640	533	582	588	544

By comparing the proposed fuel cost obtained from the model results with the fuel cost incurred by the power station when using the actual unit loading requirements, provides a means of evaluating the effectiveness of the model in improving the overall performance of the thermal power station. The preceding analysis can be seen in Figure 5.22 with the blue line representing the model fuel cost. The orange line depicts the actual power station fuel cost. The green line in the below figure is indicative of the expected cumulative rand value saving that the power station would have received, if the units were loaded according to the model results instead of using the actual loading requirements. When evaluating the results portrayed in Figure 5.22, it is clear that if the proposed model loading requirements were applied by the system operator instead of using the actual loading requirements, a definite rand value saving would have been noted for the 6 hour period. The results indicate that during hours 1 to 2 as well as 5 to 6 a definite saving, equating to a cumulative value of R20900.00, could have been obtained if the model loading requirements were implemented. During hours 3 and 4, marginal savings were noted as a result of the units being required to operate at their highest attainable load.

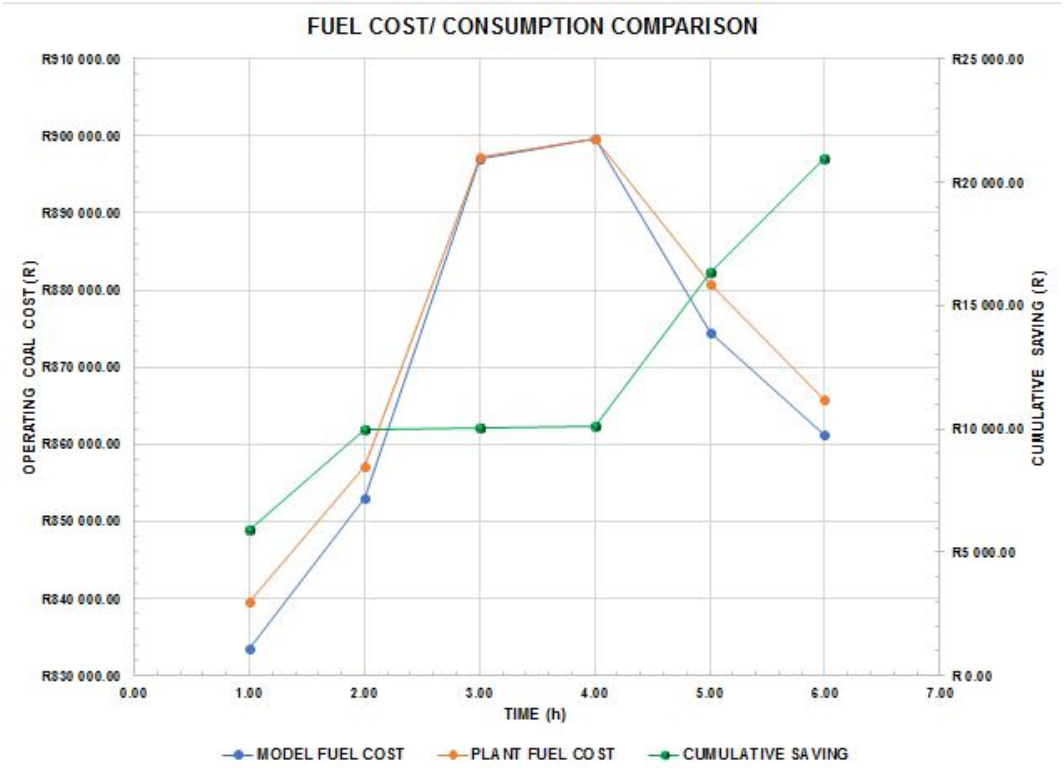

Figure 5.22: Model versus actual fuel consumption comparison

When extrapolating the possible cost saving the power station can receive by implementing the optimisation model in the production environment, assuming an R 20900 saving every 6 hours, an estimated saving of R30 572 844 is obtained for a period of 1 year. Note that the estimated cost saving is only applicable to 1 power station. If the preceding principle is expanded to a fleet of power stations such as discussed in Section 5.4.3, the expected cost savings will increase significantly.

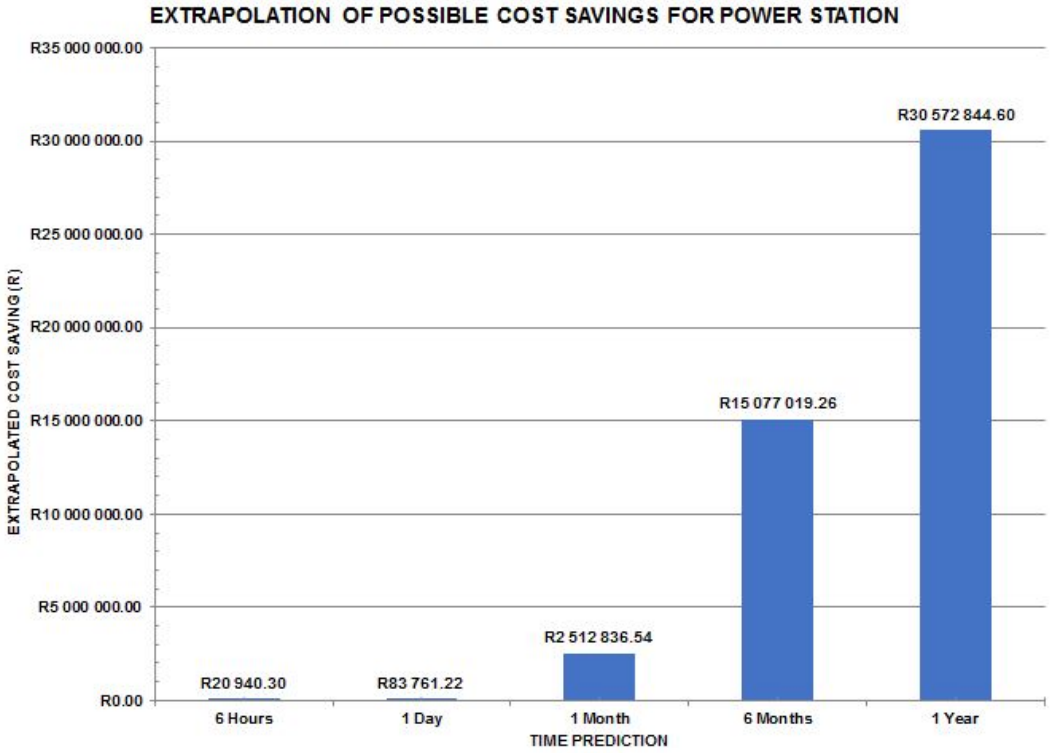

Figure 5.23: Extrapolation of possible cost savings for the power station

Another aspect to evaluate, in order to determine the effectiveness of implementing the optimisation model in the production environment, is by comparing the actual emissions production with the expected model emissions production. The comparative study results are depicted in Figure 5.24. When analyzing the results, it is clear that if the proposed model unit loading was implemented instead of the actual loading requirements as instructed by the system operator, a definite reduction in emissions production would have been noted. Similar to the fuel cost in Figure 5.22, a significant reduction in the emissions production would have been noted for periods 1 to 2 and 5 to 6. Note that the reason for the significant reduction in emissions is as a result of the alteration in the unit loading requirements. Emissions is a function of unit load as mentioned previously, and if a unit with poor emissions performance is deloaded, an immediate reduction in emissions will be noted. A marginal change in the emissions production is seen during hours 3 and 4 due to the units being loaded close to full load operation.

The results discussed in Section 5.4.1 serves as proof that by implementing the MILP UCEELD model in the power utility production environment, a definite reduction in capital expenditure, as well as an improvement in resource utilisation, will be expected. Now that the functionality of the MILP UCEELD has been proven in Section 5.4.1, it is imperative to ensure that the MILP model will be able to solve a comprehensive problem instance within a reasonable time. Although it is expected that the model will improve capital expenditure and resource utilisation as mentioned above, it will not be practical to implement if the model cannot solve a realistic sized problem within a reasonable time. For this reason, a detailed model scaling analysis and comprehensive model case study is performed in Sections 5.4.2 and 5.4.3 to determine the ability of the MILP UCEELD model in solving a realistically sized optimisation problem.

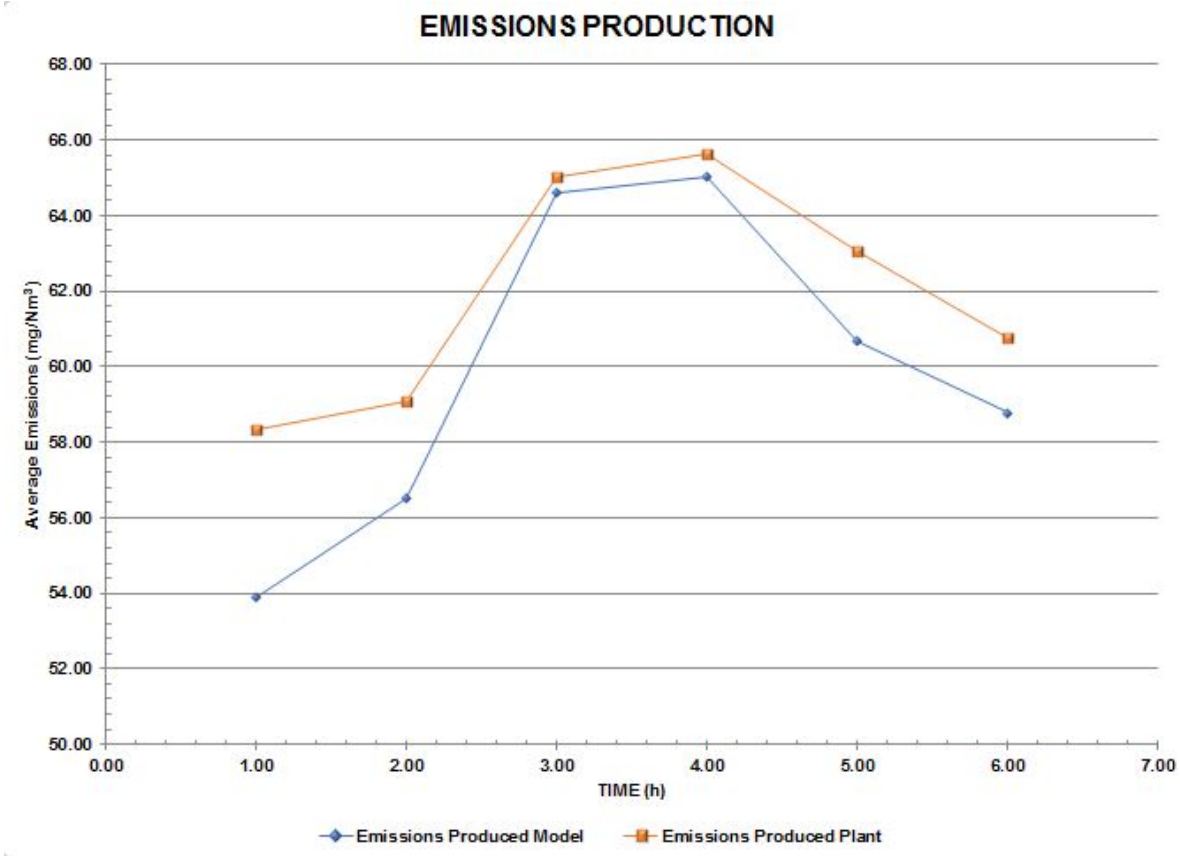

Figure 5.24: Model versus actual emissions production comparison

5.4.2 Influence of model scaling on solution time

The specifications of the computer used to solve the model scaling problem instances comprised of the following:

1. Acer laptop with an installed Intel Core i7-6500U processor
2. 4 GB DDR4 ram memory and an Intel HD 520 MB dynamic video memory graphics card

Note that when using a computer with increased computational capability, the solution time will definitely improve.

The initial problem instance that was solved comprised of 25 thermal generating units. The optimisation model was able to obtain an optimal solution within 39 seconds and calculated an objective function value of R6 533 440. After adding another 25 thermal units, the solution time increased to 1 minute 15 seconds with an expected rise in the objective function to a value of R 10 781 207. Note that an escalation of the objective function value is expected as a result of the increase in coal, water and fuel oil consumption associated with the addition of thermal units. The third problem instance entailed the solution of 75 thermal units with the model results obtained, indicating a solution time of 10 minutes 41 seconds and an objective function value of R31 499 704. The computational time required for a 98 thermal unit problem instance equated to a value of 12 minutes 21 seconds with an operational cost of R51 537 345 being calculated. After adding 7 hydro units to the 98 thermal unit model, the solution time increase to 15 minutes 37 seconds with the objective function value increasing marginally to a value of R 51 900 784. The increase in operational cost can be attributed to the increase in power demand induced by the author from problem instance 4 to 5. If the power demand remained constant, the objective function cost would have been seen to reduce as a result of the addition of the hydro units. Adding 4 pumped storage units to the previous problem instance does not significantly increase the model complexity and results in a solution time of 16 minutes 36 seconds. The objective function cost associated with the preceding model equates to a value of R 51 494 857. Comparing the operating cost of problem instance 6 with that of 5, it is apparent that there was a reduction in the overall cost. The decrease in cost can be attributed to the addition of the pumped storage units. The remaining two problem instances include the addition of multi-area power importation (with area 2 consisting of 1 hydro and 2 pumped storage units) and model stochasticity. Including the power importation functionality drastically increases the model complexity, leading to a solution time of 31 minutes 38 seconds. Similar is true when adding model stochasticity as it results in a solution time of 43 minutes 48 seconds. The objective function of the two mentioned instances are calculated to be R 50 446 031 and R 86 753 073 respectively. The drop in operational cost from R 51 494 857 to R 50 446 031 can be credited to the power importation being cheaper than the ability of the resident area to produce power. The increase in operating cost when considering the last problem instance can be as a result of the fuel cost stochasticity incorporated into the model. Refer to Figure 5.25 for a detailed summary of how the solution time is influenced by model scaling.

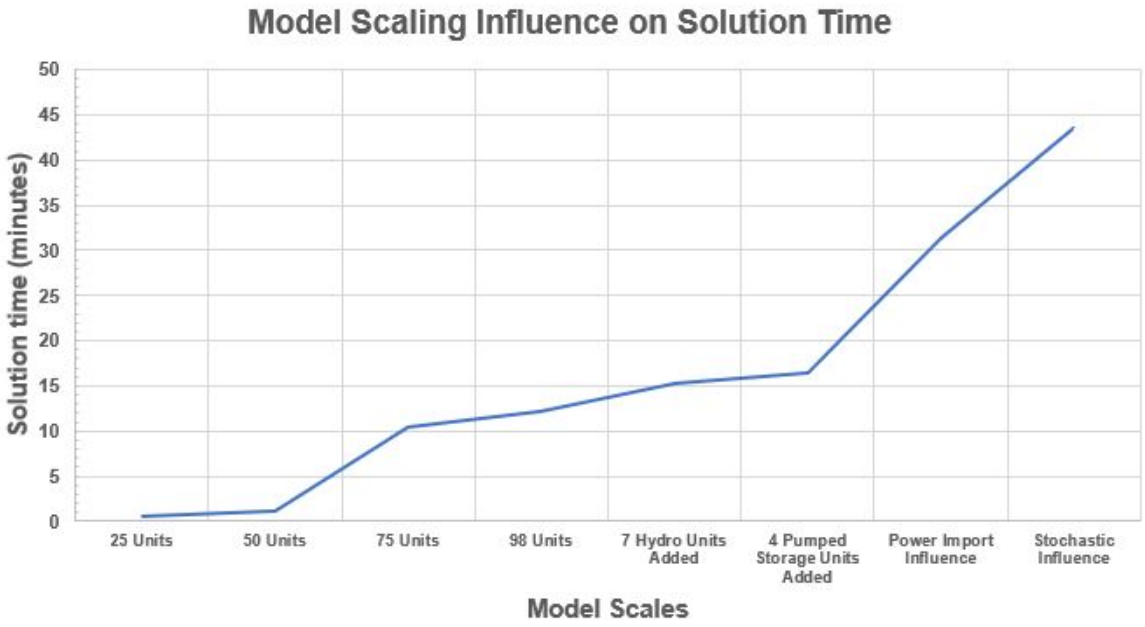

Figure 5.25: Solution time influenced by model scaling

During the execution of the different scaling problem instances, the error range as depicted in (4.5) was maintained in a range of 0 - 50 MW. It was only during the addition of model stochasticity where the range was increased to 0 - 11250 MW. The increase in error range was necessary in order to obtain a solution within a reasonable time. If not adjusted, the optimisation model will either run out of memory or will take a few hours to solve, depending on the computational power available on the computer used for solving the model. From the data depicted in Figure 5.25, it can be concluded that the MILP UCEELD problem solution time follows an exponential trend when scaling the problem from 25 thermal units straight through to the addition of the model stochasticity. The mentioned problem can be classified as an NP-hard problem as the computational complexity increases exponentially as the model is expanded. Considering the results obtained from scaling problem instances 1 - 8, it can be established that the model is capable of solving a real life realistic power utility problem within a reasonable time period. However, in order to implement the optimisation model in a production environment, it might be better to exclude the model stochasticity as it significantly increases the solution time. By Including model stochasticity the user is also required to adjust the power balance operating error range in order to obtain a model answer within a sensible time frame, with the determination of the error range also consuming additional time. In order to portray the capability of the MILP UCEELD model in solving realistic sized power utility optimisation problems, the unit commitment and load dispatch schedules obtained from the stochastic model in Figure 5.25 are portrayed in Section 5.4.3.

5.4.3 Comprehensive model

Figures 5.26 to 5.28 depict the results obtained for the comprehensive stochastic model consisting of 98 thermal, 8 hydro and 6 pumped storage units. For ease of data interpretation, the results of only thermal units 20, 37, 41, 80, 90 as well as hydro units 100, 104 and pumped storage units 1 - 4 are portrayed where indexes $n \in \mathcal{N}$ and $p \in \mathcal{P}$ are equal to 1. As mentioned in the introduction to Chapter 5, the power demand curve incorporated into the model corresponds to a realistic day ahead hourly demand curve. The initial peak in the demand curve resembles the morning peak demand whereas the second peak refers to the evening peak demand. The power flow from area 2 to 1 was set at 500 MWh to represent the maximal allowable amount of power that may be imported by the power utility in area 1. The power imported can only be supplied from either hydro or pumped storage units as no thermal units were considered for area 2.

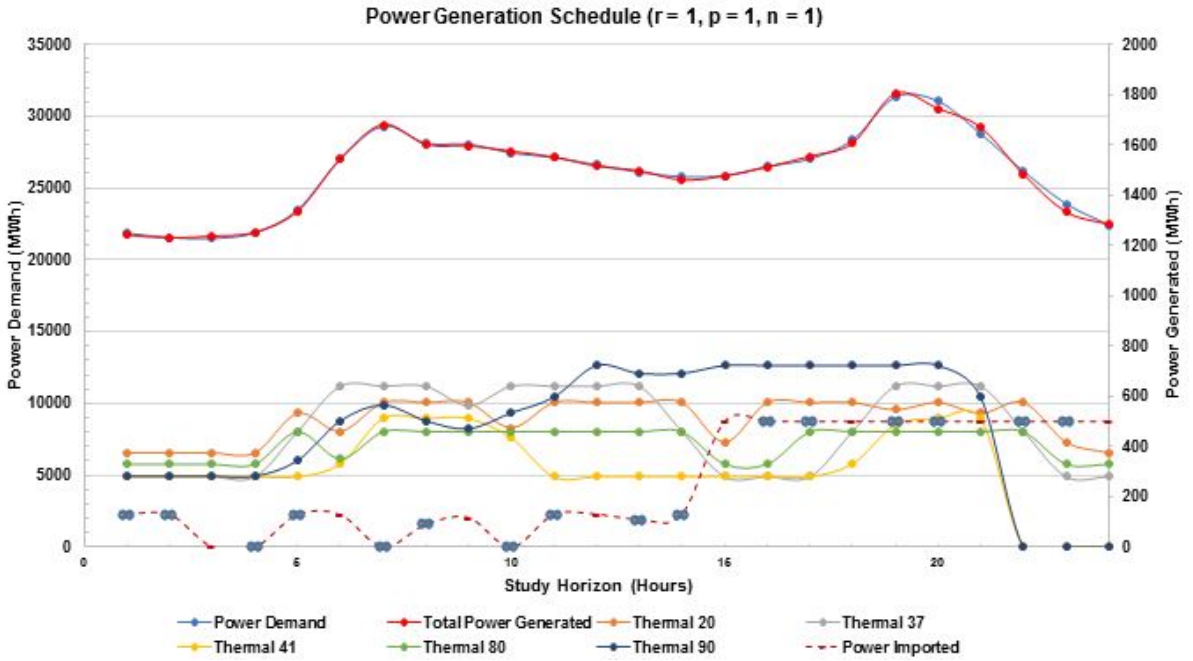

Figure 5.26: Comprehensive model results, thermal unit response

In evaluating the data depicted in Figure 5.26, it is apparent that the majority of the thermal units' loading mimic the power demand curve. During hour 1 to 5 the thermal units were loaded close to their minimum allowable loading. As the power demand started to escalate in hour 6, the loading of each thermal unit was increased to ensure that sufficient power was available to satisfy the demand. After hour 8, the power demand steadily reduced. In order to maintain grid stability, some of the thermal units were deloaded (units 20, 37, 41 and 90) whereas others were maintained at constant load (unit 80). The second peak is noted in hour 19 with the units' loading again being increased to meet demand. During hours 21 to 24, the demand reduced at a fast rate, resulting in unit 90 being decommitted from the grid. The remaining units were maintained on load, however they were deloaded close to their minimum allowable operation. In the above figure, it can also be concluded that the greatest amount of power was imported during hours 15 to 24. Prior to hour 15, power importation was limited as to ensure that the pumped storage units in area 2 satisfy their upper reservoir volume setpoint as specified by the author. From the above data, it can also be stated that thermal unit 41 is the most expensive unit when related to units 20, 37, 80 and 90 as its loading is less for majority of the time as compared to the other units.

The load profiles of both hydro and pumped storage units are portrayed in Figure 5.27. From the results depicted in the mentioned figure, it is apparent that the hydro units were only dispatched during periods 1 - 4 and 11 - 24. The preceding is as a result of the power generation schedule added to the model, by the author, to simulate the effect the commitment schedule obtained from the department of water affairs will have on the hydro unit loading. During hours 5 - 10 the hydro units were decommitted from the grid to prevent any negative influence on the downstream ecosystem as mentioned in Chapter 4. The pumped storage units were operated in pumping mode during hours 1 - 5. As the power demand increased, the pumped storage units were switched to generating mode in order to supply power to the grid during the morning peak demand (hours 6 - 10). After hour 10, the pumped storage units switched between shutdown and pumping modes in order to ensure the satisfaction of the upper reservoir volume setpoint prescribed by the author at time period 15. For the remaining hours of the study horizon, the pumped storage units operated in generating mode with the aim of supplying power to the grid during the evening peak demand period. Take note that no pumped storage units were operated in generating mode while neighboring units were switched to pumping operation. The units were either all operated simultaneously in pumping mode or a combination of pumping operation and single units being decommitted from the grid. When a pumped

storage unit was switched to generating mode, the remaining pumped storage units followed. The preceding substantiates the fact that the optimisation model correctly manages the commitment and loading of the hydro and pumped storage units.

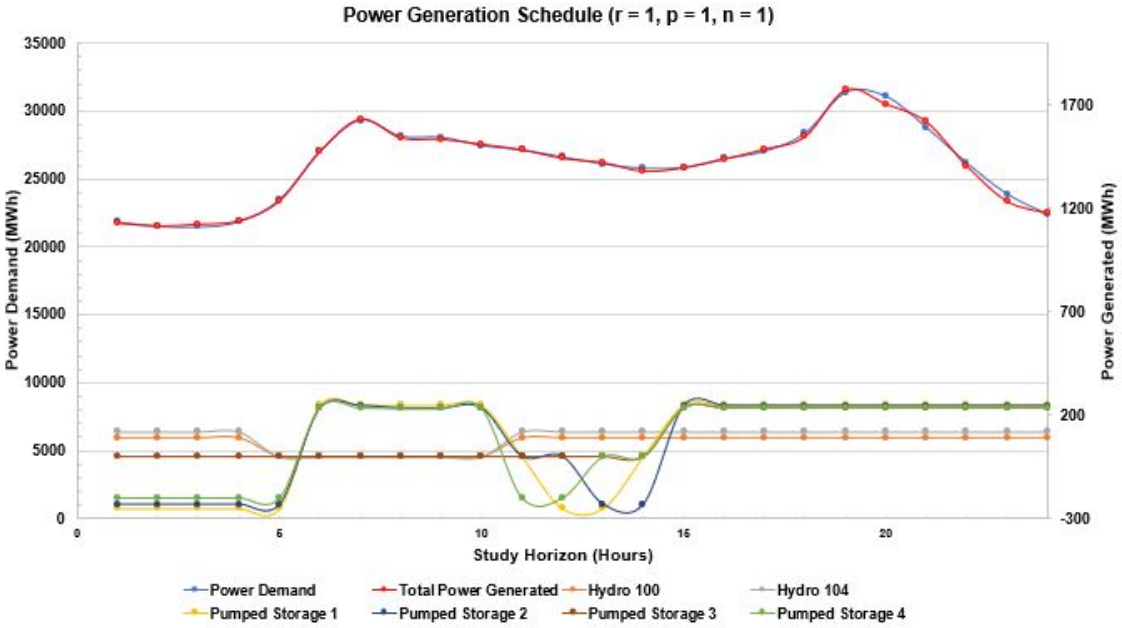

Figure 5.27: Comprehensive model results, hydro and pumped storage unit response

Figure 5.28 portrays the dynamics associated with the upper and lower pumped storage reservoir volumes, as induced by the pumped storage units' mode switching. During hours 1 - 4, the upper reservoir volume increased while the lower reservoir volume decreased. The reason being because the pumped storage units were operated in pumping mode. The same principle applied when analyzing the unit loading results depicted in hours 11 - 14 of the study horizon. Throughout hours 6 - 10 and 15 - 24, the upper reservoir volumes decreased as a result of the pumped storage units that were operated in generating mode. For the duration during which the pumped storage units were operated in generating mode, they were loaded to the maximum allowable output because of the negligible cost associated with pumped storage power generation.

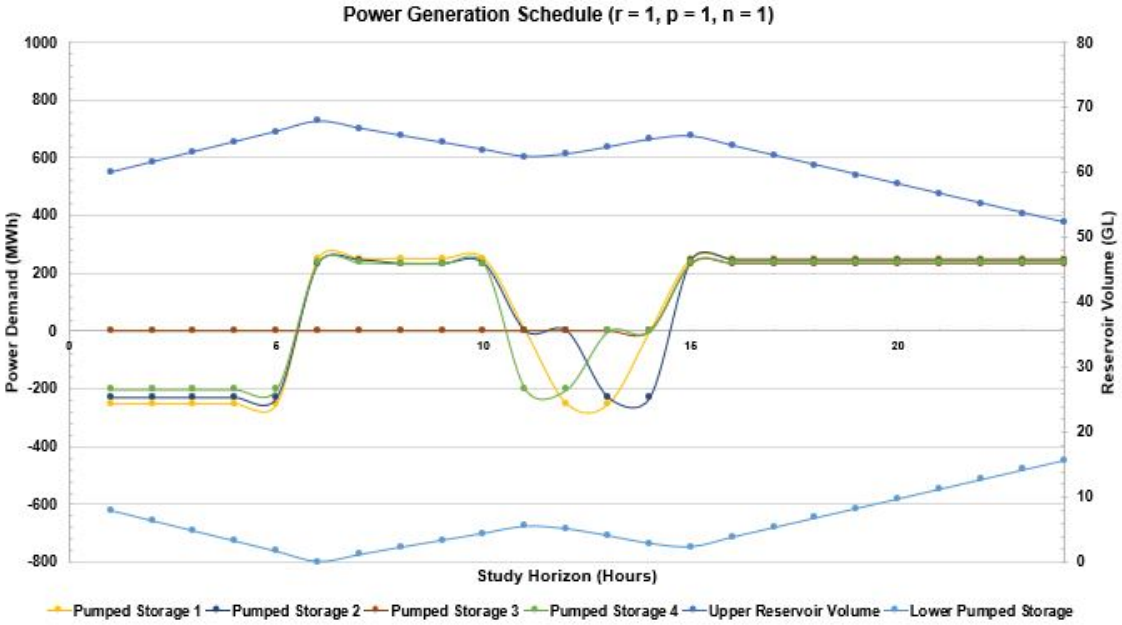

Figure 5.28: Comprehensive model results, pumped storage reservoir response

The results depicted in Section 5.4.3 serves as proof of the capability of the MILP UCEELD model to solve realistic sized power utility optimisation problems within a reasonable time. Therefore, it can be concluded that the MILP UCEELD model will be suitable for the implementation in a power utility production environment.

5.5 Visual basic GUI for data processing and acquisition

A data processing and acquisition tool was developed using Microsoft Excel Visual Basic (VBA), to reduce the time required by the user to prepare the input data for the Cplex optimisation model. In the data processing tool, graphical user interfaces (GUI's) were developed for different categories of the MILP UCEELD optimisation model. The GUI's comprise of thermal and hydro, pumped storage, day ahead forecast power demand, load demand stochasticity as well as fuel and emissions stochasticity user interfaces. Refer to Appendix B (Chapter 8) for a representation of the GUI's mentioned above. The GUI's are used to supply the VBA coding with input data from where it is transformed to the matrix format required by the commercial software, Cplex. The data processing tool does not only improve data acquisition time, but also prevent unnecessary data handling errors to ensure correct model outputs are obtained.

Chapter 6

Summary and Conclusion

6.1 Chapter summaries

Generating a unit commitment and load dispatch schedule with the aim of optimising resource utilisation and reducing operational cost, while satisfying the grid demand is a typical challenge for power utilities internationally. Constructing such a detailed schedule entails the consideration of various factors including aging infrastructures, stringent environmental legislation, multiple operational limitations and interconnected scheduling of different generating technologies as discussed throughout Chapter 1. This is a complex and monotonous exercise to perform using manual computations and therefore requires the intervention of mathematical optimisation modeling. This entails the use of computer programming to improve the feasibility and practicality of the process. It has therefore been an objective in this thesis to construct and develop a MILP UCEELD model which will be able to utilise realistically sized input data and provide the user with a comprehensive commitment and dispatch schedule within a reasonable time frame.

A technical background of different solution methods encompassed through literature with specific emphasis on linear and integer programming algorithms was provided in Chapter 2. Spesific attention was also paid to the standard formulations of both linear and integer programming problems together with a detailed elaboration on the theory behind the simplex algorithm as well as the branch-and-bound method. A brief overview of various heuristic solution algorithms was provided with a discussion of the core principle associated with each method. The aim of Chapter 2 was to ensure that the reader comprehend the fundamentals of the principles applied throughout literature to solve the unit commitment and load dispatch problems as to obtain a better understanding of the logic applied in this thesis.

The purpose of Chapter 3 was to provide the reader with a view of the fundamentals associated with different power generation technologies and the applicability of the UCEELD problem to the mentioned technologies. Specific information was provided regarding the terminology and technical aspects associated with coal fired, hydro and pumped storage units and the complexity of the problems faced on a daily basis. The aim of including the foregoing was to obtain insight and an understanding regarding the dynamics associated with the MILP UCEELD problem to be solved. A comprehensive literate review of the different solution algorithms applied by other researchers were also summarised in this chapter to obtain an idea of the algorithmic advances made by other academics.

The MILP model applied in this work for solving the UCEELD problem is presented in Chapter 4. The problem objective and constraints are formulated using mixed integer linear programming as baseline. In this chapter the different aspects associated with the model is defined with descriptive reasoning of why each was included into the model. Simplistic examples were provided in conjunction with a detailed discussion on the basic notations used, as to ensure ease of the model understanding. The model derivation is only focused on the inclusion of coal fired, hydro and pumped storage units

and did not consider other technologies such as nuclear, wind and photovoltaic power.

The computational results depicted in Chapter 5 provide a means of analyzing the capability and effectiveness of the proposed model in solving the UCEELD problem. The results depicted in this chapter were divided into verification and validation sections. Throughout model verification, focus was set on determining the model accuracy and establishing if the model's response was as expected for each of the constraints depicted in Chapter 4. The validation process was more concerned with the ability of the proposed MILP UCEELD model to solve realistically sized problems within a reasonable time frame. After evaluating the model results generated from the verification and validation processes, it could be concluded that the model's response was as expected and that the model was capable of solving realistically sized problem instances within a reasonable time. The proposed model was able to solve a full scale stochastic scenario within 43 minutes 48 seconds as denoted in Section 5.3.1. Considering the above mentioned, it is apparent that the model has the capability of being successfully implemented within a power utility production environment, and will be able to supply the user with an optimal unit commitment and load dispatch schedule.

Although the model proposed in this thesis is quite comprehensive as mentioned in the above sections, there still exists some future research that needs to be investigated in order to improve the ability of the model to simulate realistic real life conditions. A summary of the aspects that is recommended for inclusion in future work is provided in Section 6.2.

6.2 Future work

The MILP UCEELD model formulated in this thesis only takes into consideration thermal, hydro and pumped storage units when constructing the unit commitment and load dispatch schedule. Although the foregoing covers a great deal of the power generation technologies utilised by a power utility, not all existing technologies were incorporated into the model. An opportunity exists to integrate power generation technologies such as nuclear, wind and photovoltaic stations to the existing MILP UCEELD model. When considering renewable technologies such as wind and photovoltaic power, the stochasticity of these sources also needs to be taken into consideration. By including the mentioned technologies to the existing model, the complexity and solution time will most definitely increase. For this reason, an investigation will be required to determine the computational feasibility of using the MILP formulation approach in solving such a comprehensive UCEELD model. The computational feasibility must be evaluated by means of considering both solution time and the ability of the model to obtain global optimum results.

It might be beneficial to explore other means of formulating the MILP UCEELD model, after adding the additional power generation technologies, to ensure reasonable solution times are maintained. One option to scrutinize may be to supply the model with a predefined commitment schedule containing all base load units (thermal units), which are known to be on load and will remain committed for the period of the study horizon. By incorporating such a user defined input functionality, it will reduce the decision making computations required by the optimisation model significantly and may lead to improved solution times. The impact the different piecewise linear discretisation methods have on the MILP model solution time, as depicted in Section 2.2.2, also needs to be assessed during future work as to determine the optimal MILP UCEELD model formulation. An evaluation of the MILP UCEELD model solution time compared to other heuristic methods will also be beneficial when performing further research.

Further algorithmic work which can be added to the MILP model as defined in Chapter 4, include transmission line losses. These losses comprise of the amount of energy lost as a result of the transmission line lengths and capacity restrictions. Although research has shown that these losses are minute, by including these losses into the model, the capability of the model to simulate real life conditions will improve. However, in order to incorporate the foregoing losses, additional research

is required to determine an efficient method of linearising the nonlinear equations in order to fit it into the MILP framework. After inclusion of the transmission line losses, it might be beneficial to conduct a contingency analysis where different transmission lines are taken out of service to evaluate the effect these lines may have on the total operational cost. The contingency analysis may also be used to determine the criticality of each transmission line by determining which line outages may lead to infeasible solutions being obtained. The criticality of each transmission line can be utilised to determine the amount of maintenance resources to be assigned to each line as to ensure power availability and reliability is maintained. Additional aspects such as labor and maintenance cost associated with thermal, hydro and pumped storage generating units may also be added to the model objective function to improve the model's capability of simulating reality.

Incorporation of demand side resources is another aspect worthy of being investigated when considering future work on the MILP UCEELD model. These resources refer to energy consumers such as aluminium smelters with whom contracts are signed to manage the grid's energy consumption. The preceding sources are generally capable of either consuming energy or providing additional reserves to the grid in an event of grid instabilities. The demand side resources are usually capable of switching off production for a predetermined time period without suffering any losses. By utilising the characteristics of the demand side resources effectively, the optimisation model will be able to ensure grid stability by means of either scheduling the resources to, or decommitting them from the grid during power uncertainty. Being able to model the dynamics of the contracts agreed upon between a power utility and the demand side participants, will provide the system operator with additional resources to utilise with the aim of reducing total operational cost. A methodology that may be applied to simulate the demand side resources, is to model the different resources as one single virtual power plant and include it into the optimisation model accordingly. However, additional research is required to determine the feasibility of the foregoing proposal.

In the MILP UCEELD model specified throughout Chapter 4, spinning reserves were considered with no specific focus on the different categories. In order to improve the model's accuracy and its ability to simulate real life conditions, it might be beneficial to investigate the inclusion of the spinning reserves by differentiating between instantaneous, regulating and ten minute reserves. Instantaneous reserves entail the capability of certain units to supply additional power to the grid within 10 seconds with the ability to sustain the power supply for up to 10 minutes. Regulating reserves or also know as automatic generation control (AGC) are reserves provided by thermal units on a second by second basis to account for any power demand fluctuations, and is managed by the AGC control system. Lastly, ten minute reserves comprise of any units which can be committed to the grid within 10 minutes and is able to sustain the power supply for up to two hours.

In addition to the aforementioned, future work is required to expand the current day ahead MILP UCEELD model to a 7 day ahead model with the purpose of evaluating the dynamics of the pumped storage generating units. The purpose of including the pumped storage units to the day ahead model as depicted in Chapter 4, was to portray the characteristics of the constraints and its applicability to the commitment and dispatch process. However, in order to obtain a realistic and holistic view of the commitment and dispatch schedule associated to pumped storage units, a 7 day ahead schedule needs to be considered. Specific focus must be set on the solution time required to solve a 7 day ahead MILP UCEELD model as it is expected that the solution time will increase drastically. The reason being is because of the increase data quantity that needs to be processed. It might also be beneficial to evaluate other solution methodologies or constraint formulation methods which will be able to reduce the computational time of the 7 day ahead UCEELD model. Adding to the 7 day ahead model, research needs to be performed to incorporate dynamic generator maintenance outage schedule constraints. The purpose of the mentioned constraints will be to ensure that the model is able to account for variable unit outage schedules and not only consider static outage schedules as was done in this thesis. The researcher will need to evaluate the effect the inclusion of such constraints will have on the commitment and load dispatch schedule generated by the model.

Chapter 7

Appendix A: Model Parameters and Variables

Table 7.1: Model Parameters

Symbol	Definition	Units
$c^{(u)}{}_{\mathrm{h}}$	Financial fee associated with start-up fuel oil consumption for Thermal Units	R/h
$c^{(d)}{}_{\mathrm{h}}$	Financial fee associated with shutdown fuel oil consumption for Thermal Units	R/h
$\mathrm{r_p}$	The probability for a specific load scenario to occur	Fraction
$q^{(f)}{}_{\mathrm{ph}}$	Possible variable coal cost associated to each thermal unit	Fraction
$c^{(f)}{}_{\mathrm{hj}}$	Thermal unit coal cost due to load related coal consumption	R/h
$c^{(w)}{}_{\mathrm{h}}$	Thermal unit demineralised water consumption cost	R/h
$p^{(t)}{}_{\mathrm{hj}}$	Discretised generated load for thermal and hydro units	MWh
$o^{(t)}{}_{\mathrm{h}}$	Outage schedule for thermal and hydro units to prevent load scheduling of units scheduled for either IR/IN/GO	Binary
$t^{(u)}{}_{\mathrm{h}}$	The minimum up time required for thermal units before shutdown is permitted	Hours
$t^{(u1)}{}_{\mathrm{h}}$	The time a thermal unit was already operational at the start of the planning horizon	Hours
$t^{(d)}{}_{\mathrm{h}}$	The minimum down time required for thermal units before start-up is permitted	Hours
$t^{(d1)}{}_{\mathrm{h}}$	The time a thermal unit was already off at the start of the planning horizon	Hours
$e^{(f)}{}_{\mathrm{hj}}$	Particulate emissions produced by the thermal units due to load selection	mg/sm^3
$e^{(l)}{}_{\mathrm{hj}}$	Particulate emissions operating limit as stipulated by the Air Quality Act	mg/sm^3
$w^{(m)}{}_{\mathrm{n}}$	Maximum allowable water consumption permitted for thermal units	GL/h
$r^{(u)}{}_{\mathrm{h}}$	Thermal and hydro unit onload ramp-up capability	MW/h
$r^{(us)}{}_{\mathrm{h}}$	Thermal and hydro unit start-up ramp capability	MW/h
$r^{(d)}{}_{\mathrm{h}}$	Thermal and hydro unit onload ramp-down capability	MW/h
$r^{(ds)}{}_{\mathrm{h}}$	Thermal and hydro unit shutdown ramp capability	MW/h
$d^{(t)}{}_{\mathrm{h}}$	The maximum continuous rating both thermal and hydro units are capable of achieving under normal operating conditions (no limitations)	MWh
$p^{(d)}{}_{\mathrm{ptn}}$	Grid load demand governed by area power consumption	MWh

$g^{(d)}$	Total spinning reserves required for given planning horizon	MWh
$w^{(h)}{}_{\text{hj}}$	Hydro unit water consumption associated with load selection in planning horizon	m^3/s
$w^{(hm)}{}_{\text{h}}$	Maximum allowable water consumption permitted for hydro units	m^3/s
$v^{(ui)}{}_{\text{ts}}$	The initial top reservoir volume for a pumped storage station at the start of the planning horizon	GL
$v^{(ue)}{}_{\text{ts}}$	The required reservoir volume for a pumped storage station at time 6 in the planning horizon	GL
$o^{(p)}{}_{\text{su}}$	Outage schedule for pumped storage units to prevent load scheduling of units scheduled for either IR/IN/GO	Binary
$v^{(li)}{}_{\text{ts}}$	The initial bottom reservoir volume for a pumped storage station at the start of the planning horizon	GL
$l^{(p)}{}_{\text{suvj}}$	Discretised generated load for pumped storage units	MWh
$p^{(w)}{}_{\text{su}}$	Power consumption of a pumped storage unit associated with pumping operation	MWh
r_{sv}	Predefined volume ranges for the top pumped storage reservoir which is utilised for load selection	MWh
$r^{(u)}{}_{\text{su}}$	Pumped storage unit onload ramp-up capability in generating mode	MW/h
$r^{(us)}{}_{\text{su}}$	Pumped storage unit start-up ramp capability in generating mode	MW/h
$r^{(d)}{}_{\text{su}}$	Pumped storage unit onload ramp-down capability in generating mode	MW/h
$r^{(ds)}{}_{\text{su}}$	Pumped storage unit shutdown ramp capability in generating mode	MW/h
$v^{(f)}{}_{\text{suvj}}$	Pumped storage unit water consumption due to a specific load selection in generating mode	GL/h
$p^{(q)}{}_{\text{su}}$	Pumped storage unit water consumption due to operating in pumping mode	GL/h
$d^{(p)}{}_{\text{su}}$	The maximum continuous rating pumped storage units are capable of achieving under normal operating conditions (no limitations)	MWh

Table 7.2: Model Decision Varaibles

Symbol	Definition
z_{pthj}	A binary decision variable for thermal and hydro load selection to satisfy grid demand in the planning horizon with $z\varepsilon(0;1)$
$p^{(f)}{}_{\text{pt(i,j)}}$	A decision variable for load distribution between multiple geographical areas in the planning horizon with $p^{(f)}\varepsilon(0;b)$
y_{th}	A binary decision variable to track the operational hours for thermal and hydro generating units in the planning horizon with $y\varepsilon(0;1)$
w_{th}	A binary decision variable to track Unit start-up for thermal and hydro units in the planning horizon with $w\varepsilon(0;1)$
x_{th}	A binary decision variable to track Unit shutdown for thermal and hydro units in the planning horizon with $x\varepsilon(0;1)$
$s^{(t)}{}_{\text{pth}}$	A decision variable for spinning reserve allocation for thermal and hydro generating units in the planning horizon with $s^{(t)}\varepsilon(0;100000)$
$s^{(p)}{}_{\text{ptsu}}$	A decision variable for spinning reserve allocation for pumped storage generating units in the planning horizon with $s^{(p)}\varepsilon(0;100000)$

$v^{(u)}{}_{\text{pts}}$	A decision variable to track the upper reservoir volume (Mm3) of a pumped storage station throughout the given planning horizon with $v^{(u)}\varepsilon(0;100)$
$v^{(l)}{}_{\text{pth}}$	A decision variable to track the lower reservoir volume (Mm3) of a pumped storage station throughout the given planning horizon with $v^{(l)}\varepsilon(0;100)$
$v^{(f)}{}_{\text{ptsu}}$	A decision variable to track the water consumption into or out of both top and bottom pumped storage reservoirs for the given planning horizon with $v^{(f)}\varepsilon(-100;100)$
$p^{(r)}{}_{\text{tsu}}$	A binary decision variable to track the operational hours for a pumped storage unit in pumping mode throughout the given planning horizon with $p^{(r)}\varepsilon(0;1)$
$p^{(n)}{}_{\text{tsu}}$	A binary decision variable to track Unit start-up for pumped storage units in pumping mode throughout the planning horizon with $p^{(n)}\varepsilon(0;1)$
$p^{(f)}{}_{\text{tsu}}$	A binary decision variable to track Unit shutdown for pumped storage units in pumping mode throughout the planning horizon with $p^{(f)}\varepsilon(0;1)$
$t^{(r)}{}_{\text{tsu}}$	A binary decision variable to track the operational hours for a pumped storage unit in generating mode throughout the given planning horizon with $t^{(r)}\varepsilon(0;1)$
$t^{(n)}{}_{\text{tsu}}$	A binary decision variable to track Unit start-up for pumped storage units in generating mode throughout the planning horizon with $t^{(n)}\varepsilon(0;1)$
$t^{(f)}{}_{\text{tsu}}$	A binary decision variable to track Unit shutdown for pumped storage units in generating mode throughout the planning horizon with $t^{(f)}\varepsilon(0;1)$
f_{ptsuvj}	A binary decision variable for pumped storage load selection in generating mode to satisfy grid demand in the planning horizon with $f\varepsilon(0;1)$
$p^{(p)}{}_{\text{ptsu}}$	A decision variable to track the power generation (in generating mode) and or power consumption (in pumping mode) for pumped storage units throughout the planning horizon with $p^{(p)}\varepsilon(-1000;1000)$
$d^{(m)}{}_{\text{ptsuv}}$	A binary decision variable to track the upper bound of the top reservoir volume to determine the load range at which the Unit will be able to operate given the reservoir volume with $d^{(m)}\varepsilon(0;1)$
$d^{(n)}{}_{\text{ptsuv}}$	A binary decision variable to track the lower bound of the top reservoir volume to determine the load range at which the Unit will be able to operate given the reservoir volume with $d^{(n)}\varepsilon(0;1)$

Table 7.3: Model Indexes

Symbol	Definition
P	Index for the possible stochastic variability with $p\varepsilon P$
N	Index for the number of geographical areas considered with $n\varepsilon N$
T	Index for the number of time periods considered with $t\varepsilon T$
H(n)	Index for the number of thermal and hydro units considered with $h\varepsilon H(n)$
J	Index for the number of discretisation points used to linearise nonlinear functions with $j\varepsilon J$
V	Index for the number of reservoir volumes considered for each pumped storage unit with $v\varepsilon V$
S(n)	Index for the number of pumped storage reservoirs considered with $s\varepsilon S(n)$
U(s)	Index for the number of pumped storage units per reservoir considered with $u\varepsilon U(s)$

Chapter 8

Appendix B: Excel Model Graphical User Interface

In the subsequent section, the graphical user interfaces (GUI's) designed to improve the data processing and acquisition of the MILP UCEELD model inputs are provided. In Figure 8.1, the thermal and hydro GUI is portrayed. The succeeding GUI can be utilised to provide the VBA tool with the number of power generating areas, the quantity of units per area, the type of units (either thermal or hydro) and the design data for each of the mentioned units. The design data associated to the thermal units include parameters such as fuel usage and emissions production coefficients, make-up water consumption, minimum and maximum load ranges, prohibited operating regions, ramp rates as well as minimum up and down times. Other parameters also covered by the GUI include coal and fuel oil cost, outage schedules, spinning reserves, and legislative emissions limitations. Parameters associated with hydro units include majority of the above mentioned, with the fuel cost, fuel oil usage and emissions production being neglected. In order to properly account for the dynamics of the hydro stations, the curve coefficients representing the water flow rate through each turbine is added by means of the GUI as well as the maximum allowable volume flow to which the units are limited. The upload button on the GUI is utilised to submit the foregoing data to the VBA model in order to structure the information in the required matrix format.

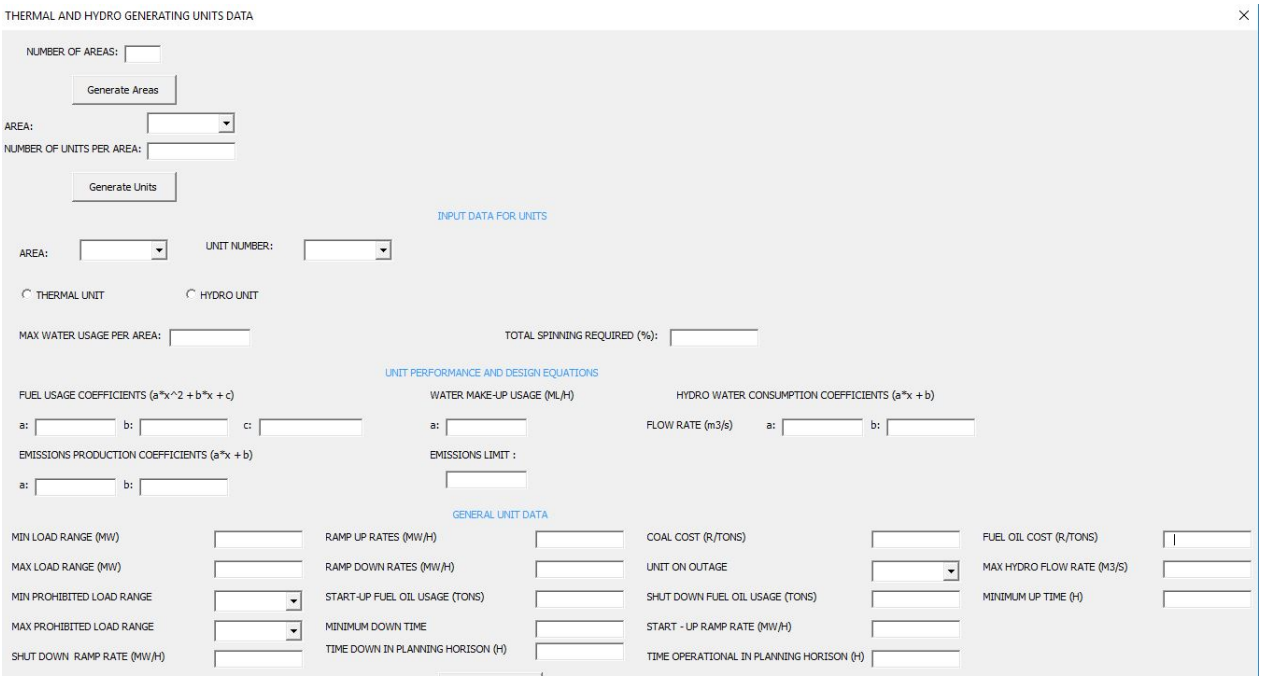

Figure 8.1: Acquisition Tool GUI for thermal and hydro unit data

Figure 8.2 represents the GUI designed for the data processing of pumped storage generating units.

The design of the GUI is similar to what was mentioned for Figure 8.1, with some additional functionalities incorporated to account for the dynamics of pumped storage units. Using the GUI, the user will be able to specify the number of areas, the quantity of reservoirs per area and the amount of pumped storage units per reservoir. Design data associated to the pumped storage units are also incorporated by means of the GUI. Majority of the parameters as mentioned in the above discussion are also applicable to the pumped storage units with the functionality of adding these parameters also included in the GUI portrayed in Figure 8.2. Note that the fuel, emissions and hydro parameters are not included in the GUI as it is irrelevant to the operation of pumped storage units. However, additional factors to consider include the initial upper and lower reservoir volumes, the upper reservoir volume setpoint, pumped storage turbine water consumption curve coefficients, the maximum and minimum allowable upper and lower reservoir volumes as well as the water and power consumption rates when operating a pumped storage unit in pumping mode. After submitting the data of each individual pumped storage unit, the information is processed by means of selecting the upload button.

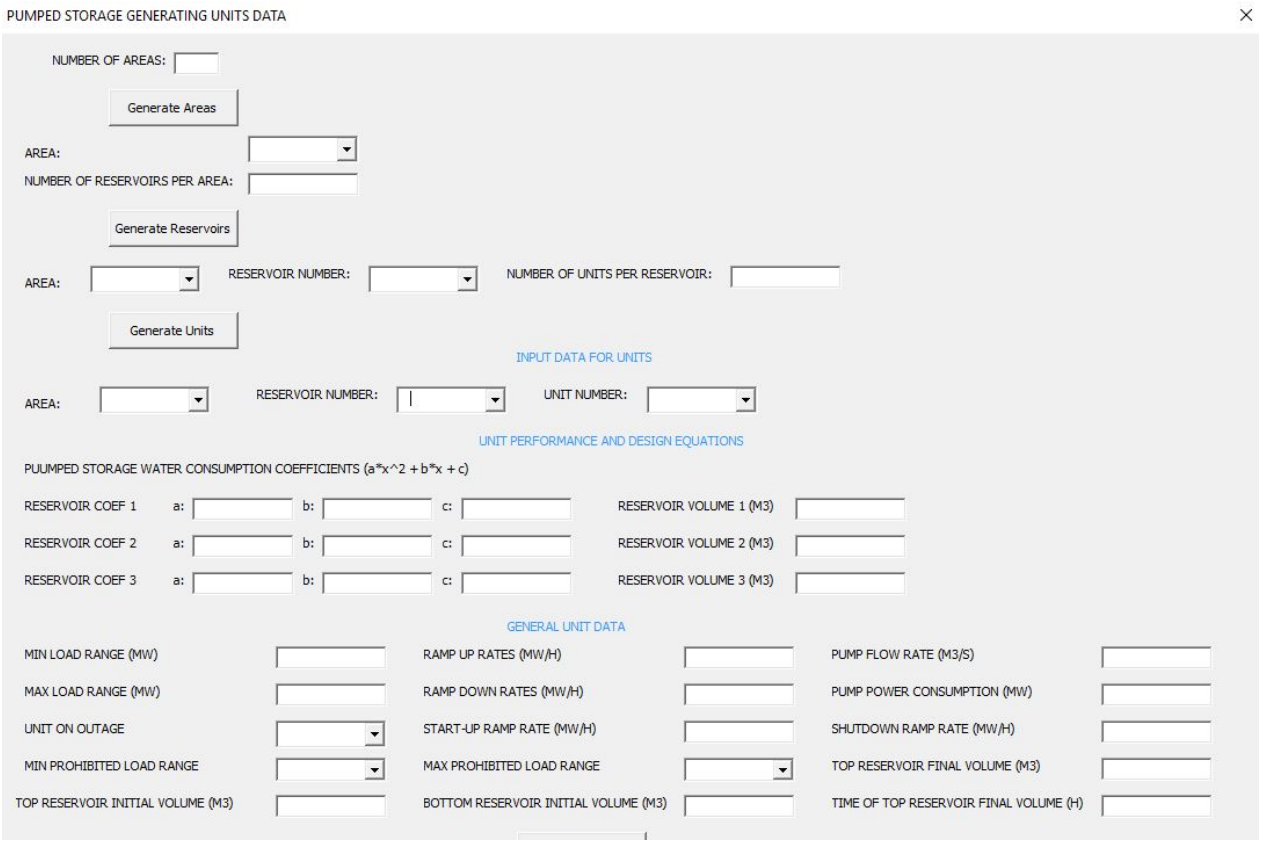

Figure 8.2: Acquisition Tool GUI for pumped storage unit data

In Figure 8.3, the GUI utilised to process the day ahead power demand forecast data, is presented. In the foregoing, the user is able to divide the power demand forecast data into hourly intervals. The GUI is structured in such a manner as to allow the user to select the area under consideration before submitting the power demand forecast. The area selection functionality is incorporated to ensure that the VBA coding assigns the correct power demand data to its corresponding area. If an incorrect power demand is assigned to an area, the mismatch could lead to erroneous results being obtained and incorrect decision making being made. For the purpose of this dissertation a GUI was designed to only process a 24 hour power demand data set, but can easily be expanded to a 7 day power demand data set interface. If however a 7 day power demand forecast is provided to the Cplex optimisation model, it is expected that the computational time required to obtain an optimal solution will increase. After submitting the demand forecast data, the information can be processed by selecting the upload button.

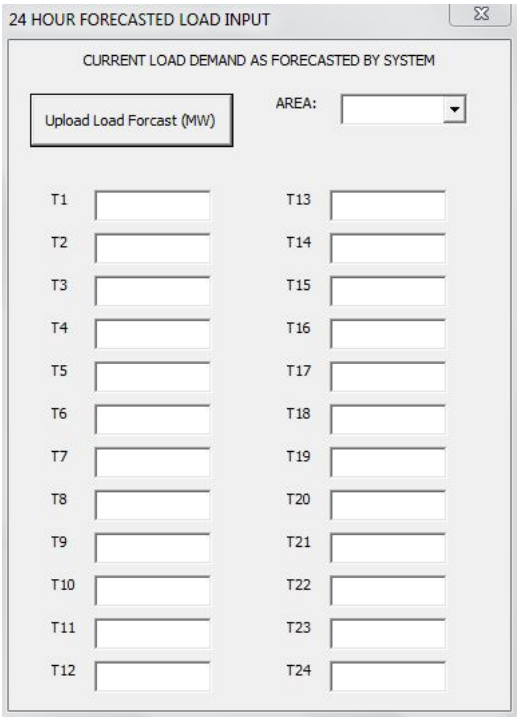

Figure 8.3: Acquisition Tool GUI for forecasted load demand

Another GUI included to the excel model to improve the data acquisition process, is the power demand stochasticity interface. The aforementioned is depicted in Figure 8.4. When using this GUI, the user is able to specify the number of stochastic scenarios with which the power demand can be altered during the study horizon. The mentioned scenarios are incorporated into the Cplex model to simulate power demand stochasticity. After selecting the submission button on the GUI, the number of stochastic scenarios is uploaded to the excel file. The drop down box enables the user to select a specific scenario number from the previously generated list. After selecting the scenario number, a fraction can be entered representing the weight to which the power demand needs to be adjusted for the specific scenario. A probability fraction is also specified representing the likelihood of the scenario from occurring. The data is uploaded to the excel file in matrix format by using the bottom submission buttons depicted in Figure 8.4.

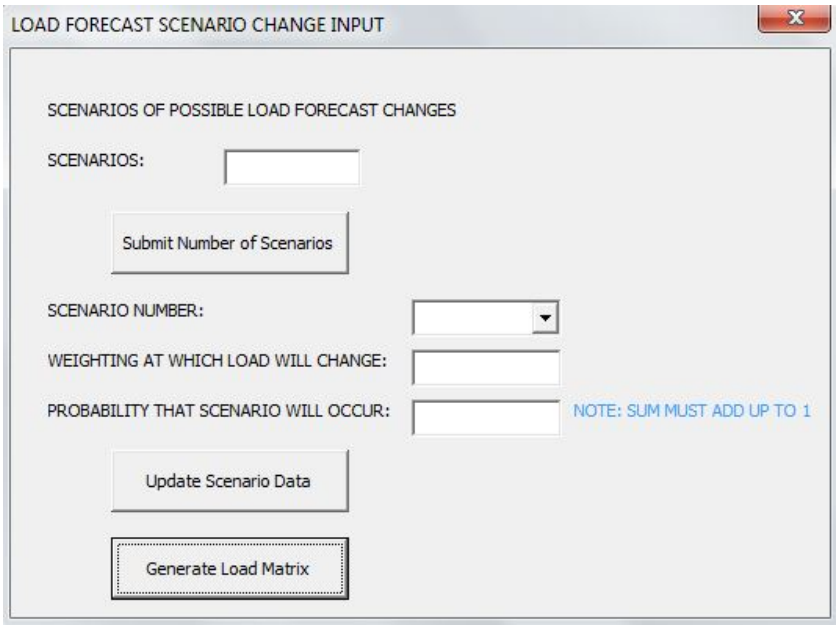

Figure 8.4: Acquisition Tool GUI for forecasted load demand stochasticity

The last GUI designed by the author is the fuel cost and emissions production stochastic interface. The purpose of this GUI is to allow the user to generate random weight fractions to which the fuel and emission variables need to be adjusted to simulate variable uncertainty. Utilising the GUI, the user is able to specify the number of stochastic scenarios that need to be considered. After selecting the number of scenarios, the data is uploaded to the excel file via the submission buttons. The matrix generation buttons call the normal distribution random number generator function in excel, and depending on the number of scenarios, the excel model will generate a certain amount of random numbers. The preceding is then implemented in the Cplex optimisation model in order to model fuel cost and emissions production stochasticity.

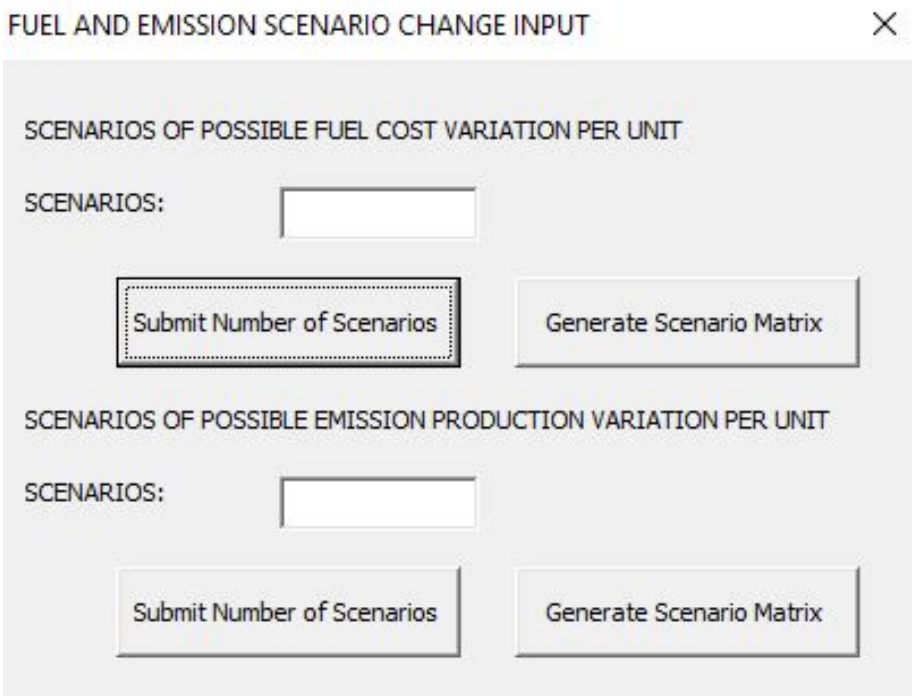

Figure 8.5: Acquisition Tool GUI for fuel cost and emissions production stochasticity

Chapter 9

Appendix C: Mathematical Model Summary

9.1 Model objective

$$\min \quad (F + W + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{h \in \mathcal{H}(n)} (c_h^{(u)} w_{th} + c_h^{(d)} x_{th})) \tag{9.1}$$

$$F = \sum_{p \in \mathcal{P}} \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{h \in \mathcal{H}(n)} \sum_{j \in \mathcal{J}} r_p q_{ph}^{(f)} c_{hj}^{(f)} z_{pthj} \tag{9.2}$$

$$\sum_{j \in \mathcal{J}} z_{pthj} = y_{th}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.3}$$

$$W = \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{h \in \mathcal{H}(n)} c_h^{(w)} y_{th} \tag{9.4}$$

9.2 Model constraints

$$0 \leq P^{(T)} + P^{(F)} + \sum_{s \in \mathcal{S}(n)} \sum_{u \in \mathcal{U}(s)} p_{ptsu}^{(p)} - p_{ptn}^{(d)} \leq 60.0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{9.5}$$

$$P^{(T)} = \sum_{h \in \mathcal{H}(n)} \sum_{j \in \mathcal{J}} p_{hj}^{(t)} z_{pthj} \tag{9.6}$$

$$P^{(F)} = \sum_{(i,j) \in \mathcal{A}} p_{pt(i,j)}^{(f)} - \sum_{(j,i) \in \mathcal{A}} p_{pt(j,i)}^{(f)} \tag{9.7}$$

$$p_{ptsu}^{(p)} - L^{(P)} + p_{su}^{(w)} p_{tsu}^{(r)} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.8}$$

$$L^{(P)} = \sum_{v \in \mathcal{V}} \sum_{j \in \mathcal{J}} l_{suvj}^{(p)} f_{ptsuvj} \tag{9.9}$$

$$s^{(d)} = p^{(d)}_{ptn}(g^{(d)}/100) \tag{9.10}$$

$$\sum_{h\in\mathcal{H}(n)} s^{(t)}_{pth} \quad + \sum_{s\in\mathcal{S}(n)} \sum_{u\in\mathcal{U}(s)} s^{(p)}_{ptsu} = s^{(d)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{9.11}$$

$$s^{(t)}_{pth} \leq (d^{(t)} y_{th} - \sum_{j\in\mathcal{J}} p^{(t)}_{hj} z_{pthj}), \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.12}$$

$$s^{(t)}_{pth} \leq r^{(u)}_{h} y_{th}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.13}$$

$$s^{(p)}_{ptsu} \leq (d^{(p)}_{su} t^{(r)}_{tsu} - p^{(p)}_{ptsu}), \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.14}$$

$$s^{(p)}_{ptsu} \leq r^{(u)}_{su} t^{(r)}_{tsu}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.15}$$

$$w_{th} - x_{th} = y_{th} - y_{(t-1)h} \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.16}$$

$$w_{th} + x_{th} \leq 1 \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.17}$$

$$\sum_{\substack{t\in\mathcal{T} \\ t\geq 1 \\ t\leq (t^{(u)}_h - t^{(u1)}_h)}} y_{th} \geq t^{(u)}_h - t^{(u1)}_h y_{1h}, \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.18}$$

$$\sum_{\substack{k\in\mathcal{T} \\ k\geq t \\ k\leq (t+t^{(u)}_h - 1)}} y_{kh} \geq t^{(u)}_h w_{th}, \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.19}$$

$$\sum_{\substack{t\in\mathcal{T} \\ t\geq 1 \\ t\leq (t^{(d)}_h - t^{(d1)}_h)}} 1 - y_{th} \geq t^{(d)}_h - t^{(d1)}_h (1 - y_{1h}), \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.20}$$

$$\sum_{\substack{k\in\mathcal{T} \\ k\geq t \\ k\leq (t+t^{(d)}_h - 1)}} 1 - y_{kh} \geq t^{(d)}_h x_{th}, \quad t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.21}$$

$$\sum_{h\in\mathcal{H}(n)} \sum_{j\in\mathcal{J}} z_{pthj} \geq 2 \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N} \tag{9.22}$$

$$p_{tsu}^{(r)} - p_{(t-1)su}^{(r)} - (p_{tsu}^{(n)} - p_{tsu}^{(f)}) = 0, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.23}$$

$$t_{tsu}^{(r)} - t_{(t-1)su}^{(r)} - (t_{tsu}^{(n)} - t_{tsu}^{(f)}) = 0, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.24}$$

$$p_{tsu}^{(n)} + p_{tsu}^{(f)} \leq 1, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.25}$$

$$t_{tsu}^{(n)} + t_{tsu}^{(f)} \leq 1, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.26}$$

$$t_{tsu}^{(n)} + p_{tsu}^{(n)} \leq o_{su}^{(p)}, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.27}$$

$$t_{tsu}^{(r)} + p_{tsu}^{(r)} \leq o_{su}^{(p)}, \quad t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.28}$$

$$\sum_{u \in \mathcal{U}(s)} L^{(P)} \leq M(1 - p_{tsu}^{(r)}), \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.29}$$

$$(\sum_{j \in \mathcal{J}} p_{hj}^{(t)} z_{pthj} - \sum_{j \in \mathcal{J}} p_{hj}^{(t)} z_{p(t-1)hj}) \leq R^{(U)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.30}$$

$$R^{(U)} = r_h^{(u)} y_{(t-1)h} + r_h^{(us)} w_{th} \tag{9.31}$$

$$(\sum_{j \in \mathcal{J}} p_{hj}^{(t)} z_{p(t-1)hj} - \sum_{j \in \mathcal{J}} p_{hj}^{(t)} z_{pthj}) \leq R^{(D)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, h \in \mathcal{H}(n) \tag{9.32}$$

$$R^{(D)} = r_h^{(d)} y_{th} + r_{ny}^{(ds)} x_{th} \tag{9.33}$$

$$(p_{ptsu}^{(p)} - p_{p(t-1)su}^{(p)}) \leq R^{(UP)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s), v \in \mathcal{V} \tag{9.34}$$

$$R^{(UP)} = r_{su}^{(u)} t_{(t-1)su}^{(r)} + r_{su}^{(us)} t_{tsu}^{(n)} \tag{9.35}$$

$$(p_{p(t-1)su}^{(p)} - p_{ptsu}^{p)}) \leq R^{(DP)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s), v \in \mathcal{V} \tag{9.36}$$

$$R^{(DP)} = r_{su}^{(d)} t_{(t-1)su}^{(r)} + r_{su}^{ds)} t_{tsu}^{(f)} \tag{9.37}$$

$$\sum_{\substack{j\in\mathcal{J}\\ j\geq pr_h^{(l)}\\ j\leq pr_h^{(h)}}} Z_{rptnyj} = 0, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, h\in\mathcal{H}(n) \tag{9.38}$$

$$\sum_{\substack{j\in\mathcal{J}\\ j\geq pr_{su}^{(l)}\\ j\leq pr_{su}^{(h)}}} f_{ptsuvj} = 0, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n), u\in\mathcal{U}(s), v\in\mathcal{V} \tag{9.39}$$

$$y_{th} \leq o_h^{(t)} \quad t\in\mathcal{T}, n\in\mathcal{N}, h\in\mathcal{H}(n) \tag{9.40}$$

$$\sum_{j\in\mathcal{J}} (q_{ph}^{(e)} e_{hj}^{(f)} z_{pthj}) \leq e^{(l)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, h\in\mathcal{H}(n) \tag{9.41}$$

$$\sum_{h\in\mathcal{H}(n)} ((q_h^{(w)} y_{th})/1505.49) \leq w_n^{(m)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N} \tag{9.42}$$

$$p_{pt(i,j)}^{(f)} \leq b, \quad p\in\mathcal{P}, t\in\mathcal{T}, (i,j)\in A \tag{9.43}$$

$$p_{pt(i,j)}^{(f)} \geq 0, \quad p\in\mathcal{P}, t\in\mathcal{T}, (i,j)\in A \tag{9.44}$$

$$\sum_{j\in\mathcal{J}} w_{hj}^{(h)} z_{pthj} \leq 0, \quad p\in\mathcal{P}, t\in\mathcal{T}: 5\leq t\leq 10, n\in\mathcal{N}, h\in\mathcal{H}(n): 99\leq h\leq 105 \tag{9.45}$$

$$\sum_{j\in\mathcal{J}} w_{hj}^{(h)} z_{pthj} \leq w_h^{(hm)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, h\in\mathcal{H}(n) \tag{9.46}$$

$$v_{pts}^{(u)} = v_{p(t-1)s}^{(u)} - \sum_{u\in\mathcal{U}(s)} v_{p(t-1)su}^{(fl)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n) \tag{9.47}$$

$$V_{pts}^{(u)} \geq V_{ts}^{(ue)} o_{su}^{(p)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n), u\in\mathcal{U}(s) \tag{9.48}$$

$$v_{pts}^{(u)} = v_{ts}^{(ui)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n) \tag{9.49}$$

$$v_{pts}^{(l)} = v_{p(t-1)s}^{(l)} + \sum_{u\in\mathcal{U}(s)} v_{p(t-1)su}^{(fl)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n) \tag{9.50}$$

$$v_{pts}^{(l)} = v_{ts}^{(li)}, \quad p\in\mathcal{P}, t\in\mathcal{T}, n\in\mathcal{N}, s\in\mathcal{S}(n) \tag{9.51}$$

$$v_{ptsu}^{(fl)} - \sum_{v\in\mathcal{V}}\sum_{j\in\mathcal{J}} v_{suvj}^{(fl)} f_{ptsuvj} + p_{su}^{(q)} p_{tsu}^{(r)} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.52}$$

$$\sum_{v\in\mathcal{V}} r_{sv}(d_{ptsuv}^{(m)} + (1 - t_{tsu}^{(r)})) \geq v_{pts}^{(u)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.53}$$

$$\sum_{v\in\mathcal{V}} r_{sv} d_{ptsuv}^{(n)} \leq v_{pts}^{(u)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.54}$$

$$d_{ptsu(v-1)}^{(m)} - d_{ptsuv}^{(n)} = 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.55}$$

$$\sum_{j\in\mathcal{J}} f_{ptsuvj} = d_{ptsuv}^{(n)}, \quad p \in \mathcal{P}, t \in T, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s), v \in \mathcal{V} \tag{9.56}$$

$$\sum_{v\in\mathcal{V}} d_{ptsuv}^{(m)} = t_{tsu}^{(r)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s) \tag{9.57}$$

$$\sum_{v\in\mathcal{V}} d_{ptsuv}^{(n)} = t_{tsu}^{(r)}, \quad p \in \mathcal{P}, t \in \mathcal{T}, n \in \mathcal{N}, s \in \mathcal{S}(n), u \in \mathcal{U}(s)c \tag{9.58}$$

Chapter 10

Bibliography

Abido, M. 2003. A novel multiobjective evolutionary algorithm for environmental/economic power dispatch. Electric Power Systems Research, 65(1): 71 - 81.

Abido, M. 2006. Multiobjective evolutionary algorithms for power dispatch problem. IEEE Transactions and Evolutionary Computation, 10(3): 315 - 329.

Afolabi, J.L. 2012. The performance of a static coal classifier and its controlling parameters. Leicester: University of Leicester. (Thesis - PhD).

Afroozi, N.M., Isapour, K., Hakimzadeh, M. & Davodi, A.2014. Solution economic power dispatch problems by an ant colony optimization approach. International Journal of Mathematical and Computational Sciences, 8(11): 1403 - 1407.

Afzal, M. & Madhav, M.K. 2017. Economic load dispatch and unit commitment solution using advanced optimization approach- a review. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 6(10): 7587 - 7593.

Agrawal, S., Panigrahi, B. & Tiwari, M. 2008. Multiobjective particle swarm algorithm with fuzzy clustering for electrical power dispatch. IEEE Transactions and Evolutionary Computation, 12(5): 529 - 541.

Ahsan, M.K., Pan, T. & Li, Z. 2018. The linear formulation of thermal unit commitment problem with uncertainties through a computational mixed integer. Journal of Power and Energy Engineering, 6: 1 - 15.

Amhamad, Y.N. & Shrivastava, J. 2016. Application of modified (PSO) and simulated annealing algorithm (SAA) in economic load dispatch problem of thermal generating unit. International Journal of Electrical Engineering & Technology (IJEET), 7(2): 69 - 78.

Anjali, T.H. & Kalivarathan, G. 2015. Analysis of efficiency at a thermal power plant. International Research Journal of Engineering and Technology: 2(5): 1112 - 1119.

Antal, B.A. 2004. Pumped storage hydropower: A technical review. Denver: University of Colorado. (Dissertation - Masters).

Ashfaq, A. & Khan, A.Z. 2014. Optimization of economic load dispatch problem by linear programming modified methodology. 2nd International Conference on Emerging Trends in Engineering and Technology, London (UK), 30 - 31 May. p. 76 - 79.

Azzam, M., Selvan, S.E., Lefevre, A. & Absit, P.A. 2014. Mixed integer programming to glob-

ally minimize the economic load dispatch problem with valve-point effect. *https* : *//arxiv.org/pdf/* 1407.4261.*pdf*. Date of access: 20 Aug 2018.

Basu, M. 2005. A simulated annealing-based goal-attainment method for economic emission load dispatch of fixed head hydrothermal power systems. Electrical Power and Energy Systems, 27: 147 - 153.

Basu, M. & Chowdhury, A. 2013. Cuckoo search algorithm for economic dispatch. Energy, 60(1): 99 - 108.

Basu, M. 2003. Hopfield neural networks for optimal scheduling of fixed head hydro-thermal power systems. Electric Power Systems Research, 64(1): 11 - 15.

Benhamida, F., Ramdani, Y. & Medles, K. 2018. A Hopfield neural network solution to economic dispatch problem including transmission losses. Journal of Electrical Engineering. *http* : *//www.jee.ro/covers/art.php?issue* = $WH1207053266W47f22bd202cc5$. Date of Access: 02 Sept 2018.

Benyahia, M., Benasla, L. & Rahli, M. 2008. Application of Hopfield neural networks to economic environmental dispatch (EED). Acta Electrotehnica, 49(3): 323 - 327.

Bhattacharya, A. & Chattopadhyay, P.K. 2010. Solving complex economic load dispatch problems using biogeography-based optimization. Expert Systems with Applications, 37: 3605 - 3615.

Borghetti, A., Ambrosio, C.D., Lodi, A. & Martello, S. 2008. An MIP approach to short-term hydro scheduling and unit commitment with head-dependent reservoir. IEEE Transactions on Power Systems, 23(3): 1115 - 1124.

Bradley, S.P., Hax, A.C. & Magnanti, T.L. 1977. Applied mathematical programming. MA: Addision-Wesley Publishing company. http://web.mit.edu/15.053/www/AMP-Chapter-09.pdf. Date of Access: 03 Sept 2018.

Bratko, D. & Doko, A. 2013. Water intake structures for hydropower.
http : *//dspace.epoka.edu.al/bitstream/handle/*1/1252/499 − 1387 − 1 − *PB.pdf?sequence* = 1&*is Allowed* = *y*. Date of access: 29 Aug. 2018.

Cai, J., Ma, C., Li, Q., Li, L. & Peng, H. 2010. A multi-objective chaotic ant swarm optimization for environmental/economic dispatch. Electrical Power and Energy Systems, 32: 337 - 344.

Chang, G.W., Aganagic, M., Waight, J.G., Medina, J., Burton, T., Reeves, S. & Christoforidis, M. 2001. Experiences with mixed integer linear programming based approaches on short term hydro scheduling. IEEE Transactions on Power Systems, 16(4): 743 - 749.

Chang, G.W. & Waight, J.G. 1999. A mixed integer linear programming based hydro unit commitment. IEEE Power Engineering Society. Summer Meeting Conference Proceedings, 18 - 22 July. P. 924 - 928.

Chaturvedi, K.T., Pandit, M. & Srivastava, L. 2008. Hybrid-neuro-fuzzy system for power generation control with environmental constraints. Energy Conversion and Management, 49(11): 2997 - 3005.

Chen, C.L. & Chen, N. 2001. Direct Search method for solving economic dispatch problem considering transmission capacity constraints. IEEE Transactions on Power Systems, 16(4): 764 - 769.

Chen, CL. 2007. Non-convex economic dispatch: a direct search approach. Energy conversion and management, 48(1): 219 - 225.

Chen, D., Batson, R.G. & Dang, Y. 2010. Applied integer programming modelling and solution. 1st. Hoboken, NJ: John Wiley & Sons.

Chen, P.H. 2008. Pumped-storage scheduling using evolutionary particle swarm optimization. IEEE Transactions on Energy Conversion, 23(1): 294 - 301.

Corneujols, G. & Tutuncu, R. 2007. Optimization methods in finance.1st ed. NY: Cambridge University Press.

Dantzig, G.B. 2002. Linear programming. Operations Research, 50(1): 42 - 47.

Dash, S.K. 2013. An artificial neural network method for optimal generation dispatch with multiple fuel options. ISSN, 2(1): 126 - 134.

Desell, A.L., McClelland, E.C., Tammar, K. & Horne, P.R.V. 1984. Transmission constrained production cost analysis in power system planning. IEEE Transactions on Power Apparatus and Systems, 103(8): 2192 - 2198.

Dillon, T., Edwin, K., Kochs, H. & Taud. R. 1978. Integer programming approach to the problem of optimal unit commitment with probabilistic reserve determination. IEEE Transactions on Power Apparatus and Systems, 97(6): 2154 - 2166.

Downes, G. 2012. Eskom's capacity and new build context. Presented at the 4th SA-EU Clean Coal Working Group, Johannesburg Emperor's Palace, SA, 5 November 2012. $http://www.energy.gov.za/files/4thEUSouthAfricaCleanCoalWorkingGroup/Eskom\%20SA_EU\%20bilateral\%20GD.pdf$. Date of access: 11 Jun. 2018.

Espana, G.M., Latorre, J.M. & Ramos, A. 2012. Tight and compact MILP formulation for the thermal unit commitment probleb. $https://www.iit.comillas.edu/aramos/papers/OnlineCompanion_Tight\&Compact_UC.pdf$. Date of Access: 30 Aug 2018.

Feige, U. 2011. The basics of linear programming. $http://www.wisdom.weizmann.ac.il/\ robi/teaching/2012a-AdvancedAlgorithms/Lecture1.pdf$. Date of Access: 02 Sept 2018.

Ferreres, X.R. & Font, A.R. 2010. Installation of a new hydropower plant in ockelbo. Sweden: University of Gavle. (Dissertation - Masters).

Fossati, J.P. 2012. Unit commitment and economic dispatch in micro grids. $http://www.um.edu.uy/docs7_unit_commitment_and_economic_dispatch_in_micro_grids.pdf$. Date of Access: 30 Aug 2018.

Gaing, Z. 2003. Particle swarm optimization to solving the economic dispatch considering the generator constraints. IEEE Transactions on Power Systems, 18(3): 1187 - 1195.

Galati, M. 2010. Decomposition methods for integer linear programming. Lehigh: Lehigh university. (Dissertation - PhD).

Gartner, B. 1995. Randomized optimization by simplex-type methods. Berlin, University of Berlin. (Dissertation - PhD).

Genova, K. & Guliashki, V. 2011. Linear integer programming methods and approaches - a survey. Cybernetics and Information Technology, 11(1): 3 - 25.

Ghasemi, A. 2013. A fuzzified multi objective interactive honey bee mating optimization for environmental/economic power dispatch with valve point effect. Electrical Power and Energy Systems, 49: 308 - 321.

Ghasemi, M., Aghaei, J., Akbari, E., Ghavidel, S. & Li, L. 2016. A differential evolution particle swarm optimizer for various types of multi-area economic dispatch problems. Energy, 107: 182 - 195.

Gil, E., Bustos, H. & Rudnick, H. 2003. Short term hydrothermal generation scheduling model using genetic algorithm. IEEE Transactions on Power Systems, 18(4): 1256 - 1264.

Govidsamy, R. 2013. Thermal performance evaluation of heat exchangers in pulverised coal boilers. Johannesburg: University of the Witwatersrand (Dissertation - Masters).

Gunda, J. & Acharjee, P. 2011. Multi objective economic dispatch using pareto frontier differential evolution. International Journal of Engineering Science and Technology, 3(10): 7389 - 7396.

Habibollahzadeh, H. & Bubenko, J.A. 1986. Application of decomposition techniques to short-term operation planning of hydrothermal power system. IEEE Transactions on Power Systems, 1(1): 41 - 47.

Hadji, B., Mahdad, B., Srairi, K. & Mancer, N. 2015. Multi-objective PSO-TVAC for environmental/economic dispatch problem. Environment and Sustainability, 74: 102 - 111.

Hedman, K.W., Ferris, M.C., OBeill, R.P., Fisher, E.B. & Oren, S.S. 2009. Co-optimization of generation unit commitment and transmission switching with N-1 reliability. *http : //www.pserc. cornell.edu/empire/*04_H*edman_GenUC*$_1$1022009*.pdf*. Date of Access: 30 Aug 2018.

Hua, W. 1990. Application of the revised simplex method to the farm planning model. Hangzhou, China. (Thesis - Masters).

Hunt, J.D., Freitas, M.A.V. & Junior, A.O.P. 2014. Enhanced-pumped-storage: combining pumped-storage in a yearly storage cycle with dams in cascade in Brazil. Energy: 78: 513 - 523.

IBM Corp. 2015. IBM ILOG Cplex optimization studio Cplex user's manual. *https : //www.ibm .com/support/knowledgecenter/SSSA*5$P_1$2.6.2*/ilog.odms.studio.help/pdf/usrcplex.pdf*. Date of Access: 11 Oct. 2018.

Ikura, Y. & Gross, G. 1984. Efficient large-scale hydro system scheduling with forced spill conditions. IEEE Transactions on Power Apparatus and Systems, 103(12): 3502 - 3520.

Jacobs, C. 2017. Hydro and Peaking station overview [e-mail]. 4 May. 2017.

Jadoun, K.V., Gupta, N., Niazi, K.R. & Swamkar, A. 2015. Modulated particle swarm optimization for economic emission dispatch. Electrical Power and Energy Systems, 73: 80 - 88.

Jayabarathi, T., Chalasani, S. & Shaik, Z.A. 2007. Hybrid differential evolution and particle swarm optimization based solutions to short term hydro thermal scheduling. WSEAS Transactions on Power Systems, 11(2): 245 - 254.

Jayabarathi, T., Jayaprakash, K., Jeyakumar, D. & Raghunathan, T. 2005. Evolutionary pro-

gramming techniques for different kinds of economic dispatch problems. Electric Power Systems Research, 73(2): 196 - 176.

Jayabarathi, T., Sadasivam, G. & Ramachandran, V. 2000. Evolutionary programming based multi-area economic dispatch with tie line constraints. Electric Machines and Power Systems, 28(12): 1165 - 1176.

Jeddi, B. & Vahidinasab, V. 2014. A modified harmony search method for environmental/economic load dispatch of real-world power systems. Energy Conversion and Management, 78: 661 - 675.

Jestin. L. 2017. Eskom specialization center in Energy Efficiency, Thermo-hydraulics applied to power plant design and operation monitoring. (Unpublished).

Jeyalalitha, K., Ramesh, K.N.S., Santhosh, K.R. & Sathish, K.N. 2008. Design of hydro power plant. Chennai: Anna Univeersity. (Report - Bachelor).

Joshi, K. 2018. Optimization of economic load dispatch problem by using tabu search algorithm. International Journal of Latest Trends in Engineering and Technology, 8(4): 182 - 187.

Kapooria, R.K., Kumar, S. & Kasana, K.S. 2008. An analysis of a thermal power plant working on a rankine cycle: A theoretical investigation. Journal of Energy in Southern Africa, 19(1) : 77 - 83.

Kapooria, S., Kumar, S. & Kasana, K.S. 2008. Technological investigations and efficiency analysis of a steam heat exchange condenser: conceptual design of a hybrid steam condenser. Journal of Energy in Southern Africa, 19(3): 35 - 45.

Khodja, F., Younes, M., Laouer, M., Kherfane, R.L. & Kherfane, N. 2014. A new approach ACO for solving the compromise economic and emission with the wind energy. Energy Procedia, 50: 893 - 906.

Koretsky, M.D. 2013. Engineering and Chemical Thermodynamics. 2nd Edition. River street Hoboken, NJ: John Wiley & Sons Inc.

Lewis, C. 2008. Linear programming: theory and applications. *https : //www.whitman.edu/ Documents/Academics/Mathematics/lewis.pdf*. Date of Access: 02 Sept 2018.

Liang, Y.C. & Juarez, J.R.C. 2014. A normalization method for solving the combined economic and emission dispatch problem with meta-heuristic algorithms. International Journal of Electrical Power and Energy Systems, 54(1): 163 - 186.

Li, L.D., Li, X. & Yu, X. 2008. Power generation loading optimization using a multi-objective constraint-handling method via PSO algorithm. Korea, Daejeon: IEEE International Conference on Industrial Informatics. P. 1632 - 1637.

Lombard, A.C.J. 2010. Design and implementation of a low cost grid-connected 10 kW hydro power system. Stellenbosch: Stellenbosch University. (Theses - Masters).

Lowery, P.G. 1996. Generating unit commitment by dynamic programming. IEEE Transactions on Power Apparatus and Systems, 85(5): 422 - 426.

Lu, Y., Zhou, J., Qin, H., Wang, Y. & Zhang, Y. Environmental/economic dispatch problem of power system by using an enhanced multi-objective differential evolution algorithm. Energy Conversion and Management, 52(2): 1175 - 1183.

Mahor, A., Prasad, V. & Rangnekar, S. 2009. Economic dispatch using particle swarm optimization: A review. Renewable and Sustainable Energy Reviews, 13: 2134 - 2141.

Mallach, S. 2015. Exact integer programming approaches to sequential instruction scheduling and offset assignment. Cologne: University of Cologne. (Dissertation - PhD).

Mandal, K.K., Mandal, S., Bhattacharya, B. & Chakraborty, N. 2015. Non-convex emission constrained economic dispatch using a new self-adaptive particle swarm optimization technique. Applied Soft Computing, 28: 188 - 195.

Manteaw, D.E. & Odero, N.A. 2012. Multi-objective environmental/economic dispatch solution using ABC_PSO hybrid algorithm. International Journal of Science and Research Publication, 2: 1 - 7.

Manivel, J., Manimaran, L., Thiyagarajan, M., Satheeshkumar, P. & Thiruppathi, R. 2017. Performance analysis of air preheater in 210mw thermal power station. International Journal of Advanced Research, Ideas and Innovations in Technology, 3(2): 619 - 630.

Myllyvirta, L. 2014. Health impacts and social costs of Eskom's proposed non-compliance with South African's air emission standards. http://*www.greenpeace.org/africa/Global/africa/publications/Health*%20*impacts*%20*of*%20*Eskom*%20*applications*%202014%20*final.pdf*. Date of access: 11 Jun. 2018.

Naama, B., Bouzeboudja, H. & Allali, A. 2013. Solving the economic dispatch problem by using tabu search algorithm. Energy Procedia, 36: 694 - 701.

Ni, E. & Guan, X. 1999. Scheduling hydrothermal power systems with cascaded and head-dependent reservoirs. IEEE Transactions on Power Systems, 14(3): 1127 - 1132.

Norang, I. 2015. Pumped storage hydropower for delivering balancing power and ancillary services. *https* : *//brage.bibsys.no/xmlui/bitstream/handle/*11250/2433581/13093$_F ULLTEXT$ *.pdf?sequence* = 1&*isAllowed* = *y*. Date of Access: 30 Aug 2018.

Norouzi, R.M., Ahmadi, A., Nezhad, A.E. & Ghaedi, A. 2014. Mixed integer programming of multi-objective security-constrained hydro/thermal unit commitment. Renewable and Sustainable Energy Reviews, 29:911 - 923.

Palanichamy, C., Naveen, P., Ing, W.K., Danquah, M.K. & Indumathi, J. 2015. Energy efficiency enhancement of fossil-fueled power systems. International Journal of Energy Economics and Policy: 5(3):765 - 771.

Pan, S., Jian, J. & Yang, L. 2017. Solution to dynamic economic dispatch with prohibited operating zones via MILP. *https* : *//arxiv.org/pdf/*1704.01801.*pdf*. Date of Access: 2 Sept 2018.

Pandit, M., Srivastava, L. & Sharma, M. 2015. Environmental economic dispatch in multi-area power system employing improved differential evolution with fuzzy selection. Applied Soft Computing, 28: 498 - 510.

Panigtahi, B.K., Pandi, V.R., Sharma, R., Das, S. & Das, S. 2011. Multiobjective bacterial foraging algorithm for electrical load dispatch problem. IEEE Transactions on Power Systems, 52(2): 1334 - 1342.

Papageorgious, L.G. & Fraga, E.S. 2007. A mixed integer quadratic programming formulation

for the economic dispatch of generators with prohibited operating zones. Electric Power Systems Research, 77(10): 1292 - 1296.

Papavasiliou, A. & Oren, S.S. 2013. A comparative study of stochastic unit commitment and security-constrained unit commitment using high performance computing. *https : //perso .uclouvain.be/anthony.papavasiliou/public*$_h$*tml/ECC.pdf*. Date of Access: 02 Sept 2018.

Park, J.H., Kim, Y.S. & Lee, K.Y. 1993. Economic load dispatch for piecewise quadratic cost function using hopefield neural network. IEEE Transactions of Power Systems, 8: 1030 - 1038.

Pieterse, J.C. 2016. High pressure feedwater heaters replacement optimization. Pretoria: University of Pretoria. (Thesis - Masters).

Prasad, A.D., Jain, K. & Gairola, A. 2013. Pumped storage hydropower plants environmental impacts using geometrics techniques: an overview. International Journal of Computer Applications: 81(14): 41 - 48.

Rall, D.P. 1974. Review of the health effects. Environmental health perspectives, 8:97-121.

Rao, S.S. 2009. Engineering optimization theory and practice. 4th ed. Hoboken, NJ: John Wiley & Sons.

Romano, R., Quintana, V.H., Lopez, R. & Valadez, V. 1981. Constrained economic dispatch of multi-area systems using the Dantzig-Wolfe decomposition principle. IEEE Transactions on Power Apparatus and Systems, 100(4): 2127 - 2137.

Salama, M.M., Elghazar, M.M., Abdelmaksoud, S.M. & Henry, H.A. 2013. Short term optimal generation scheduling of fixed head hydrothermal system using genetic algorithm and constriction factor based particle swarm optimization technique. International Journal of Scientific and Research Publications, 3(5): 2250 - 3153.

Salama, M.M., Elghazar, M.M., Abdelmaksoud, S.M. & Henry, H.A. 2014. Short term pumped storage scheduling using two proposed techniques. International Journal of Energy and Environment, 5(2): 219 - 238.

Sampath, R., Dhami, S.S. & Srivastava, S. 2015. A virtual model of steam turbine power generation unit. International Journal of Performability Engineering, 12(1): 33 - 43.

Santillan, L.K., Sene, K.B., Wamilda, R. & Ocampo, L. 2016. An integrated rule-based approach of unit commitment with economic load dispatch problem. DLSU research congress, Philippines (Manila), 7 - 9 March. p. 1 - 6.

Sasikala, J. & Ramaswamy, M. 2010. Optimal based economic emission dispatch using simulated annealing. International Journal of Computer Applications, 1(10): 55 - 63.

Sayah, S. & Zehar, K. 2006. Economic load dispatch with security constraints of the Algerian power system using successive linear programming method. Leonardo Journal of Science, 9: 73 - 86.

Secui, D.C. 2015. The chaotic global best artificial bee colony algorithm for the multi-area economic/emission dispatch. Energy, 93: 2518 - 2545.

Senthil, K. & Manikanda, K. 2010. Economic thermal power dispatch with emission constraint and valve point effect loading using improved tabu search algorithm. International Journal of Com-

puter Applications, 3(9): 6 - 11.

Shiina, T. 2004. Price-based unit commitment problem. *http : //www.kurims.kyoto − u.ac.jp/kyodo/kokyuroku/contents/pdf/*1373 − 25.*pdf*. Date of Access: 02 Sept 2018.

Sjelvgren, D., Andersson, S., Andersson, T., Nyberg, U. & Dillon, T.S. 1983. Optimal operations planning in a large hydro-thermal power system. IEEE Transactions on Power Apparatus and Systems, 102(11): 3644 - 3651.

South-Africa. 2010. National Environmental Management: Air quality Act 39 of 2004.

Still, D., Erskine, S., Walker, N. & Hazelton, D. 2008. The status and use of drinking water conservation and savings devices in the domestic and commercial environments in South Africa. *http : //www.wrc.org.za/Pages/Drought/*4.2/*Water*%20*demand*%20*management/status*%20*of*%20*DW*%20*devices*$_T$*T*%20358.*pdf*. Date of access: 11 Jun. 2018.

Suman, M., Rao, M.V.G., Hanumaiah, A. & Tajesh, K. 2016. Solution of economic load dispatch problem in power system using lambda iteration and back propagation neural network methods. International Journal of Electrical Engineering and Informatics, 8: 347 - 355.

Sun, C. & Lu, S. 2010. Short term combined emission hydro thermal scheduling using improved quantum-behaved particle swarm optimization. Expert Systems with Applications, 37(6): 4232 - 4241.

Swart, D.E. 2016. Performance assessment and mass energy balance for regenerative type air heaters. Potchefstroom: NWU. (Dissertation - Masters).

Teleman, Y. 2016. Rankine Cycles, modeling and control. Sweden: Lund University. (Thesis - Masters)

Treshow, M. 1980. Pollution effects on plant distribution. Environmental Conservation, 7:279-286.

Tseng, C.L., Li, C.A. & Oren, S.S. 2000. Solving the unit commitment problem by a unit decommitment method. Journal of optimization theory and applications, 105(3): 707 - 730.

Tuffah, M. & Gravdahl, J.T. 2013. Mixed-integer formulation of unit commitment problem for power systems: focus on start-up cost. *https : //core.ac.uk/download/pdf/*52115361.*pdf*. Date of access: 30 Aug 2018.

Van Rooyen, K.M. 2014. The effect of condenser backpressure on station thermal efficiency: Grootvlei power station as a case study. Potch: NWU. (Dissertation - Masters).

Vasanie, P. 2004. Eskom and Water. Proceedings of the 2004 water institute of Southern Africa (WISA) biennial conference, Cape Town, SA, 2 - 6 May 2004. http:*//docplayer.net* /38011728 − *Eskom − and − water − vasanie − pather.html*. Date of access: 19 Feb. 2017.

Walters, D.C. & Sheble, G.B. 1993. Genetic algorithm solution of economic dispatch with valve point loading. IEEE Transactions on Power Systems, 8(3): 1325 - 1332.

Wassung, N. 2010. Water scarcity and electricity generation in South Africa. Stellenbosch: Stellenbosch University. (Thesis - Masters).

Welty, J.R., Wicks, C.E., Wilson, R.E. & Rorrer, G.L. 2008. Fundamentals of momentum, heat, and mass transfer. 5th edition. River street Hoboken, NJ: John Wiley & Sons.

Williams, H.P. 2009. Logic and Integer programming. 1st ed. Heidelberg, NY: Springer.

Winner, E.W., Mooney, H.A. & Goldstein, R.A. 1985. Sulfur dioxide and vegetation physiology, ecology, and policy issues. California, Stanford: Stanford university press. Date of Access: 11 Jun. 2018.

Winston, W.L. & Goldberg, J.B. 2004. Operations research applications and algorithms. 4th ed. Belmont, CA: Curt Hinrichs.

Wong, D.P. & Wong, Y.W. Short term hydro thermal scheduling part. I. Simulated annealing approach. Generation, Transmission and Distribution, 141(5): 497 - 501.

Wu, L., Thimsen, D., Clements, B., Zheng, L. & Pomalis, R. 2014. A hybrid rankine cycle (HyRC) with ambient pressure combustion (APC). Applied Thermal Engineering, 73: 484 - 499.

Yahya, A.E.M., Shaban, M. & Yahya, Y. 2015. Applying unit commitment method in power station to minimize the fuel cost. Open Journal of Social Sciences, 3: 166 - 173.

Yang, J.S. & Chen, N. 1989. Short term hydrothermal coordination using multipass dynamic programming. IEEE Transactions on Power Systems, 4(3): 1050 - 1056.

Yang, X., Hosseini, S.S.S. & Gandomi, A.H. 2012. Firefly algorithm for solving non-convex economic dispatch problems with valve loading effect. Applied Soft Computing, 12: 1180 - 1186.

Zhou, W., Sun, H. & Peng, Y. 2010. Risk reserve constrained economic dispatch model with wind power penetration. Energies, 3: 1880 - 1894.

Zhu, J.Z. 2003. Multiarea power systems economic power dispatch using a nonlinear optimization neural network approach. Electric Power Components and Systems, 31(6): 553 - 563.

Zoi, V.A. 2013. Feasibility and environmental appraisal of a small hydro-power plant. Piraeus: University of Piraeus. (Thesis - Masters).