Exploration of a Digital Twin Concept for Income Maximisation of the Waikaremoana Power Scheme

A thesis

submitted in fulfilment

of the requirements for the degree

of

Master of Engineering

at

The University of Waikato

by

Seng Theng (Marcus) Ly

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2022

I. Abstract

Digital twins are digital emulations of real-world objects or systems which mirrors the asset in terms of both behaviour and likeness. Genesis Energy Ltd, an electricity generation company in New Zealand, is interested in developing a digital twin for one of its hydropower assets, the Waikaremoana Power Scheme (WPS). The scheme is a multi-lake cascading system in the North Island with a generation capacity of 138 MW using seven turbines. The ultimate goal for digital twin development with Genesis is to create a tool capable of providing decision making support for traders managing the utilisation of the WPS in the NZ electricity market to maximise income and efficiency while minimising losses. This project is a tentative exploration into how an early digital twin concept could be built for the WPS with the end objective of maximising utilisation through optimising unit commitment and scheduling.

Plant data accuracy and reliability was examined as it is a foundational element to any digital twin. It was found that the WPS possessed accurate instruments for parameters like power output and water levels but relied on correlations for many flow readings around the scheme. Data sampling methods were also examined, and it was found that averaged data was better at short sampling intervals due to reduction of noise.

A flow model was built in Microsoft Excel using a first principles-based approach, assembled using mass and energy balances along with characteristic equations for the scheme. The accuracy of the model was tested against net flow values via lake sensor readings. It was found that on average, there was a difference of around 2 m3/s for all three lakes but increased in proportion to the model net flow rate. As part of the flow model, the efficiency characteristic functions were found using a regression refinement process, starting from linear regression refining eventually to a multivariable linear (polynomial) model. The model was validated using test data from each of the generation units with excellent or good fits found for all units. .

A profit-based optimisation formulation was developed based on literature and the flow model developed. The formulation was applied to a simplified case comprising of a single unit, lake, and time slice problem and testing the relationship between spot price and water value. The problem was solved using Excel Solver and a GRG nonlinear method. Depending on whether the spot price was greater or smaller than water value, the optimiser chose to generate at maximum efficiency before increasing to maximum output as spot price rose. The optimisation encountered difficulties when extending the problem to include multiple units. The Excel Solver was unable to find the global optimum without increasing compute times to unacceptable levels. To proceed further with this optimisation problem, it would need to be moved to a platform with more complex solvers.

II. Acknowledgements

I would like to express my deepest gratitude to Martin Atkins, who as my supervisor helped to guide me throughout my time researching. He provided a lot of insight about the research process and was always reassuring with a lot of patience and understanding. I would not have been able to finish this project without his help.

I want to give a major thanks to Genesis Energy, who were kind enough to host me for this project and gave me the opportunity to get some experience working in a great corporate environment. A special thanks to Michael Eschenbruch, my manager at Genesis. Without his guidance over the course of the project, his knowledge and help sourcing data from various places at Genesis, I could not have completed this project. A big thanks also to Jane Bydder and Pradeep Birudala for helping with proofreading and programming. Thanks also to the innovation and reliability team for being so welcoming.

I would like to express my gratitude to Callaghan Innovation for helping to fund this project with their research grant and also to the University of Waikato for helping with the funding of the Master's with a scholarship.

A big thanks to the Ahuora team, who were an extremely kind and fun group of people to work with. I am grateful that I was able to take part in some of the group events.

Finally, I would also like to thank my family, Che, Vic, Mum, and Dad, for all the support they've given over the course of all my studies.

III. Table of Contents

IV. Table of Figures

V. Table of Tables

VI. Nomenclature

Label	**Parameter**
A	Lake Surface Area
c	Start-up Maintenance Cost
h	Lake Level
$\overline{h}$	Maximum Lake Level
$\underline{h}$	Minimum Lake Level
H	Gross Head
h_{fb}	Forebay level
h_{tail}	Tailrace level
K	Production Function
P	Power Generated
q	Total Flow In
q_{Dis}	Discharge Flow Rate
q_{Gi}	Turbine Flow Rate
q_{L}	Leakage Flow Rate
q_{Sp}	Spill Flow Rate
S	Electricity Market Spot Price
V	Available Water in Reservoir
W	Water Value
y	Start-up Binary Variable (0,1)
η	Turbine Overall Efficiency

Chapter 1
Introduction

1.1 Context

Hydropower has been used as a reliable source of renewable energy for thousands of years and is essential to New Zealand's power generation system. It harnesses the potential energy of water in rivers or lakes at higher elevations, converting this energy to rotational kinetic energy. There has been evidence to suggest that hydropower in the form of waterwheels have been used as early as 400 BC in Ancient Greece and China (Munoz-Hernandez et al., 2013). The technology has evolved immensely over the years, developing into today's efficient hydro turbines, able to transfer up 90+ % potential energy into the rotation of a shaft. (Munoz-Hernandez et al., 2013)

Today's hydro turbines convert hydraulic energy into electricity by rotating a shaft with an iron core known as a rotor, this induces an electrical current in a series of coil windings surrounding the rotor, known as a stator. Modern hydropower schemes can be classified into 2 main types; impoundment and run-of-river (Harding & Mills, 2020). Run-of-river schemes typically involve diverting a portion of a river's flow to a small reservoir while impoundment involves damming a lake or river, often at higher elevations. In both cases, water from the reservoirs is transported through conduits such as tunnels and penstocks. The hydraulic pressure of water in the penstocks drive hydro-turbines which rotate coupled electrical generator shafts. The water is then expelled further downstream or into a lake at a lower elevation. Depending on the head and flow conditions, different types of turbines are used; reaction turbines (e.g., Kaplan: high flow, low head. Francis: moderate flow, moderate head) and impulse turbines which use the velocity force of the water (e.g. Pelton: low flow, high head)(Global Change Data Lab, 2022).

New Zealand is a country that relies heavily on hydropower for its renewable generation mix. Owing to the development of many hydropower sites in the 20th century, New Zealand has been able to boast a relatively high percentage of renewable electricity (a large part of which is hydropower) generation for many years (fluctuating between 60 and 90 % of annual generation since the 1970s)(Ministry of Business, Innovation and Employment, 2021). In 2020, hydropower makes up around 56 % of the NZ electricity generation mix, this is relatively high in comparison to other countries but modest compared to world leaders Norway and Brazil at 92 % and 64 % respectively (Global Change Data Lab, 2022).

Although hydropower is a reliable and renewable energy source with low emissions and running costs, it currently only makes up around 15% of the total electricity generated in 2021 worldwide (Global Change Data Lab, 2022). This modest value can be attributed to the geographic and water resource requirements, in addition to the environmental impacts of dam building and the large capital investment requirements. Despite these disadvantages, hydropower will play an essential role in many electrical grids due to its generation flexibility and quick starting capabilities.

The above factors become increasingly important as countries attempt to decrease their carbon emissions and move towards net zero emissions. To match similar pledges by other nations, the New Zealand government has made commitments towards net zero emissions by 2050 and has set a goal of attaining 100% renewable electricity generation by 2035 for a normal hydrological year. Despite already having a high level of renewable electricity generation, New Zealand faces several challenges when attempting to reach its 100% renewable electricity target such as diminishing returns (to compensate for intermittent availability, wind or solar capacity must be overbuilt), and energy security concerns when it comes to grid reliability and suitable dry year cover.

Wind power is projected to play a major role in the push towards renewables; Over the last 15 years New Zealand has increased its wind generation capacity from 54 MW in 2002 to 689 MW in 2016 and has added a further 350 MW in 2021. The Interim Climate Change Committee (2019) expects that by 2035 wind generation will increase to 3400 MW and make up 19% of the country's electrical generation mix (up from 5.5% in 2020)(Poletti & Staffell, 2021). In 2021, the advice given by the NZ Wind Energy Association to the Climate Change Commission for their report suggested that to meet the 13.3 TWh wind generation goal by 2035, 2800 MW of new generation would be needed on top of the 2021 installed capacity of 690 MW.

Although wind generation commissioned in New Zealand has been able to achieve a respectable 40 % capacity factor annually, wind power can be quite inconsistent on a short-term basis; daily generation for a farm can vary between 10 % and 70 % capacity factors with little warning (Poletti & Staffell, 2021). Moreover, unlike hydro, wind generation cannot be held, generators generally aim to operate units whenever wind is available. Because of this inconsistency and lack of storage capability, wind power pairs well with hydropower; hydropower can provide quick response reserve capabilities if an unexpected decrease in wind occurs while wind generation allows hydro to store more water for dryer periods.

Compared to other countries, New Zealand's hydroelectric power schemes have relatively small reservoirs that rely on consistent rainfalls to maintain annual generation capacity (Pritchard, 2021). New Zealand's hydropower schemes are run-of-river configurations and lack the storage flexibility of pumped hydro schemes. Water is generally held during the summer period for winter use when electricity demand is high and precipitation inflows decrease due to snow formation.

This has meant that New Zealand's electricity generation system is rather susceptible to negative impacts caused by low rainfall (dry years). Consequently, the country still requires the services of fossil fuel powered generation sites such as the coal and gas fired Huntly Power Station to provide a reliable energy security pillar particularly during a dry year. Due to the factors discussed above, efficient management of hydropower reservoirs will be increasingly important as more wind generation is brought online and traditionally reliable fossil-fuel generation is wound down.

Genesis Energy Ltd is a New Zealand generator-retailer company that owns several hydropower and thermal generation assets around the country, and it is exploring ways of evaluating the overall performance in terms of fuel use and maximising its generating potential. A significant portion of its generation assets are represented by hydropower schemes, ordered from largest to smallest generators are Tongariro, Tekapo, and Waikaremoana . Reservoir level management is also heavily intertwined with the power market; generators must decide how much generation to offer, what spot prices to offer the generation at and how long of duration to produce the greatest returns. This will consider inflow and storage considerations (spilling water will be in effect wasting fuel and thus revenue), tangible costs such as repair and maintenance costs and intangible costs such as opportunity cost. Opportunity cost is defined as the difference between the returns using the most profitable option and the returns from the actual decision made. This drives much of the complexity of the unit commitment problem and will be important in evaluating the overall performance of a generation site. Electricity spot market traders at Genesis currently use a mix of experience and intuition to optimise the revenue generated by each of its generation assets. However, the company is looking for innovative ways to create a 'benchmark' of sorts to evaluate the general performance of its traders and pinpoint areas where improvements in asset use can be implemented.

Recently, the concept of 'digital twins' has become an increasingly hot topic for companies and research groups looking to optimise their assets. The definition for a digital twin (DT) can vary wildly depending on the area of study, ranging from one-to-one working digital models to organisation data flow models. Some research groups have divided DTs into types based on level of fidelity and connectivity, coining the terms: Digital Models, Shadows, and Managers. According to IBM, "A DT is a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning and reasoning to help decision-making" (Armstrong, 2020). A DT aims to reproduce a system in the digital space that can replicate its likeness and behaviour while also having a connection between the real and digital space for the twin to continually adapt to changes in the real world. Genesis would like to explore the development of a DT for the Waikaremoana Power Scheme (WPS) to use as a support tool for its wholesale market traders. It is currently difficult for Genesis to evaluate the performance of the hydro scheme outside of financial comparisons. The ideal tool would be able to model the scheme hydraulic relationships to optimise the short-term and long-term hydro unit commitment considering design and operational parameters, weather forecasting, electricity market forecasting and water storage.

1.2 Objectives

This research thesis aims to explore the possible development of a DT for the WPS, focusing on the preliminary steps and early-stage development required to lay the groundwork for a functional DT model in the future.

The scope of this work primarily involves understanding the environment being modelled in terms of the characteristics of the hydropower scheme itself and its relationship with the power market. The project is to outline all the key components of the power scheme and identify the hydraulic relationships linking these components including efficiency functions and flow balances. To obtain the data to utilise these relationships, the relevant sensor and correlative data need to be sourced and their reliability assessed. As an important part of the flow balance model, efficiency functions specific to each turbine must be found. The project is to use correlative methods like regression and historical operations data to determine the best model to describe the efficiency function.

From here the project will look to create a flow model of the WPS using flow and energy balances through a first principles-based approach. The flow model will be made up of an amalgamation of all the balances, input data, hydropower, and efficiency characteristic functions to form one continuous and connected balance. The final piece of the project will address the long-term objective set by Genesis Energy; decision making support. The DT must be able to provide some insight into how to best manage the scheme at a given point in time to achieve the optimal amount of profit. A section of work carried out will be to do with the unit commitment and scheduling optimisation problem, learning the problem formulation for typical hydropower schemes and applying it to the WPS.

- What data sources are available for a Waikaremoana Power Scheme DT and are they sufficient for building an accurate model?
- What are the biggest challenges when it comes to building a DT model for a hydropower system like the Waikaremoana Power Scheme?
- What are the challenges to do with a mathematical optimisation of this problem? Is the optimisation process best done using traditional methods and what newer methods could be explored?

1.3 Thesis Outline

The contents of each thesis chapter are outlined below, each chapter covers a key area of the research topic and/or model development.

Chapter 2 details the literature review undertaken for this project. This chapter presents relevant theory and findings from published research relating to the main areas of interest for this thesis. It includes a look at digital twins, turbine efficiency, and optimisation of the hydropower scheduling problem, which includes formulation approaches for this problem and a look at the many methods for solving this problem.

Chapter 3 covers background material about the system of interest which includes the major areas, the WPS and New Zealand's power sector. It provides contextual information about the problem environment, detailing the main features of both the power scheme and the power sector which it participates in.

Chapter 4 examines some of the groundwork involved with developing a digital twin. It discusses how the digital twin concept can be applied to the Waikaremoana system and describes the sources of data for this project. This includes outlining the sampling methods, types of sensors used for data gathering and the limitations if any of these readings.

Chapter 5 details the first step in the behaviour modelling course, a curve fitting of the hydropower characteristic equations for the hydropower units at WPS using an incremental regression modelling process. It takes a step-by-step approach to finding the best fitting model for the turbine operational data before presenting a visual representation of this model, discussing the non-ideal fits and their possible causes.

Chapter 6 presents the next step for developing the behaviour aspect of the WPS digital twin in terms of water and energy flow. It explores a first principles-based approach using mass and energy balances to account for all known inputs and outputs of the system and derives all the necessary equations linking the main features of the WPS. It further uses these equations and the efficiency characteristic equations to assemble a 'first pass' flow model extending from the top of scheme to the bottom.

Chapter 7 tackles the optimisation or management aspect of the digital twin and attempts to apply the problem formulation previously covered in Chapter 2. It outlines the development process of a simplified version of popular profit-based optimisations including the objective function, equations, and constraints. It further details the solving of this optimisation problem using a nonlinear solver while incrementally increasing the complexity of the formulation.

Chapter 8 summarises the conclusions made in this project and the future research recommendations.

Chapter 2
Literature Review

2.1 Introduction

The thesis has thus far introduced the high-level energy problem, highlighting the importance of hydropower and its increasingly important role in generation as more renewable sources like wind power are added to the New Zealand generation mix. The overall objective of the project was defined as the exploration of the digital twin concept for the WPS and setting of the groundwork for a digital model capable of providing decision making support to traders managing the utilisation of WPS in the New Zealand power market.

This chapter presents the findings of a literature review covering several key project aspects. These key aspects include digital twins, turbine efficiencies, and problem formulation regarding unit commitment and hydro scheduling. These topics are not all directly linked but each plays a big role in the WPS model building process.

The DT concept is often open to interpretation, this section helps to ground the definition by examining how researchers have refined definitions and classifications. Turbine efficiency function development is a cornerstone to both energy and mass balances, without which the model would be unable to function. Lastly, as a decision-making support tool, the model must be able to optimise to some degree. Understanding the optimisation terms commonly used by others in this field will be helpful to recognize which will be suitable for this project. Each of the three main sections in this chapter cover general theory found in literature followed by the presentation and discussion of past studies by researchers in these subjects.

2.2 Digital Twins

The digital twins (DT) idea was originally sown by NASA in 1970, when they used detailed virtual simulations to test repair scenarios for Apollo 13 (W. Yu et al., 2022). The term ‘Digital Twins’ would not be established for another 30 years when it was coined by Michael Grieves and John Vickers of NASA in 2003 (Jones et al., 2020). Grieves described the early concept and model of a DT in a lecture about Product Lifecycle Management. He described the DT as having three components; a physical product, a virtual representation, and a two-way directional connection that could both feed data from the physical space to the virtual domain and transfer processed information and processes back to the physical space. Digital representations during this time were still quite rudimentary and data collection from physical objects was limited and often a manual process (Grieves, 2014). The relationship between the three components can be seen in the Figure 2-1.

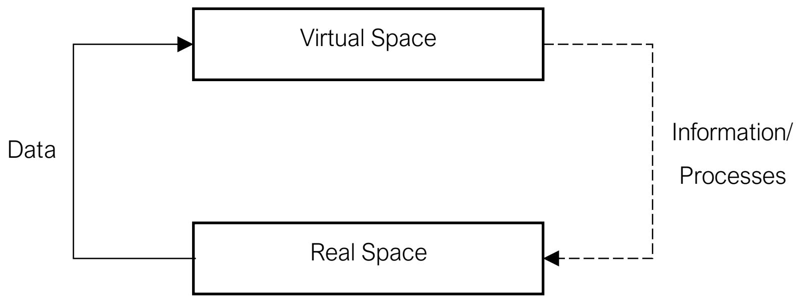

Figure 2-1: Digital Twin Block Flow Diagram (Jones et al., 2020)

DTs have become an increasingly popular research field amongst researchers and industry innovators alike. The required technology to realise the potential of DTs has taken many leaps forward since Grieves penned the concept. Technology and computing power have advanced to a point where widespread implementation of replicant digital models is feasible. Digital models now have a far higher level of fidelity and can often mirror their physical counterpart to an impressive degree. Data collection is more sophisticated through automated sensors, datalogging and centralised control systems.

Energy efficiency is a hot topic for all industries as the pressure to address climate change and reduce emissions heats up. Companies in the manufacturing and energy sector are increasingly seeking innovative ways to increase production efficiency by utilising newer tools in the digital space. Many different companies and research groups have started to invest into DT development for a wide variety of applications. As such, many variations of the DT concept have emerged along with many different definitions.

IBM defines a digital twin as:

"*A virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning and reasoning to help decision-making*." (IBM, 2022).

Although definitions differ slightly from one another, Grieves' main structure and concept have continued to remain at the core. DTs remain in essence a digital representation of a real object or system that can model their behaviour in a digital domain while having the ability to update in real time using information exchanged between real and digital domains. A review of DT research by Yu et al. (2022) from the Ahuora Energy Research Group at the University of Waikato has provided a refined definition that can be applied to a broad range of disciplines.

"A digital (or virtual) representation that looks-like, behaves-like, and connects-to a physical part or system with the goal of improving or optimising decision making for any time horizon."

2.2.1 Classifications

DTs can be applied to a wide variety of industries with varying degrees of detail. Yu et al. (2022) have formalised a framework for the classification of DTs according to the levels of fidelity and sophistication in the three main DT components: likeness, behaviour, and connection (Jones et al., 2020).

Likeness refers to the visual aspects of the DT, where the appearance of the model attempts to visualise or mirror the physical asset. The framework separates the level of likeness into three distinct categories:

- 1-D representation – e.g., a process flow diagram
- 2-D representation – e.g., a process flow diagram with dimensions and co-ordinates
- 3-D representation – e.g., virtual 3D model

For process modelling, the behaviour attribute is often perceived as the most important part of a DT. Behavioural replication of a system attempts to process a set of inputs and provide outputs that are true to the physical object or system. Yu et al. (2022)'s framework categorises the level of behavioural fidelity into the following:

- Single-state, static information – e.g., average process system state.
- Discrete, event-driven, multiple stead-states model – e.g., multiple equilibrium process system states
- Dynamic, time-driven, transient model, e.g., model predictive control.

The last attribute in the framework is the connection component of the model. It is often what separates a traditional simulation model with a DT. The transfer of information between the physical to the digital domains allows the model to continually update and improve based on the live condition of the real/physical object or system. The level of connection can be categorised as follows (Kritzinger et al., 2018):

- Digital Model - All indirect (e.g., non-automated) data flow between physical part or system and the DT – e.g., CAD drawing based on physical systems.
- Digital Shadow - Direct data flow (e.g., automated) from the physical part or system and indirect data flow from the DT in return e.g., a plant operator's process data display.
- Digital Manager - Two-way direct data flow between the physical part or system and the DT e.g., a plant control system.

Yu et al. (2022) offers a naming convention for these classifications in the form of DT_{ikj} where i, k, and j (i,k,j=1,2,3) denotes the level of likeness, behaviour, and connection.

In addition to the three attributes, DT classification can also include specifications of physical-scale and timescale. These scales are helpful in further differentiating types of DT research. For physical-scale classification, the framework has defined four levels.

- Nano, e.g., molecular level
- Micro, e.g., single operation or part
- Meso, e.g., collection of operations, including a factory or site
- Macro, e.g., a community, local area, or region

The timescale defines the time horizon which the DT will resolve and operate in. DT time horizons can range from seconds for real time process control to weeks and in some cases even decades for process design DTs.

2.2.2 Examples

Most DT research in the last 10 or so years have been focused on the manufacturing and building sectors, having a combined Scopus publication count of over 1400 since 2010 (W. Yu et al., 2022).

A research area showing great potential for DT development but with relatively light research coverage is the process and energy industries. DTs in this area fall under the Energy Digital Twins (EDTs) term. The main goal for EDTs research is to develop a tool capable of managing energy streams in a manufacturing plant and improving energy and fuel use efficiency. EDTs will be essential for managing energy consumption and improving heat and energy use efficiencies for manufacturing plants, particularly at a time where industries are exploring ways to reduce carbon emissions (W. Yu et al., 2022). From Yu et al.'s classification framework, it will be useful to explore some EDT examples with different classifications.

The first example is a 'simpler' EDT with 1-D likeness, discrete multi-state replication and indirect data flow, DT_{121}. Xu et al. (2019) built a numerical DT model of a 320 MW coal-fired thermal power plant using the Thermoflow software suite. The model consists of numerical physics models of the boiler island, steam-turbine island, and emission control equipment. It could simulate the thermodynamic performance of the plant, calculate the heat and mass balance of the overall system, sub-systems, and major plant components for any given operating condition. Xu et al. calibrated the model using recent operating performance data from the plant and validated it by simulating various part load operating scenarios to compare against actual plant performance.

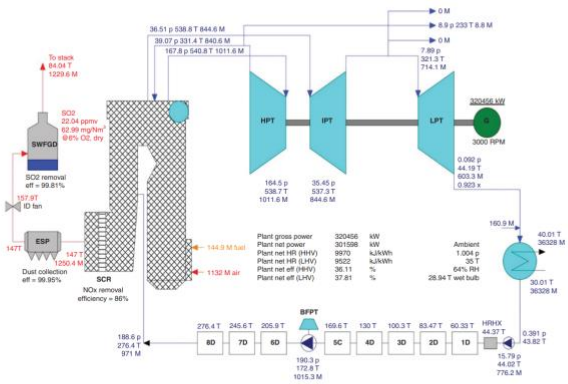

Figure 2-2: Xu et al.(2019)'s Coal Power Plant DT

Visuals-wise the DT used a simple process flow diagram interface which displayed the system parameters as seen in Figure 2-2. Connectivity was not specifically mentioned and looks to have followed the digital model class.

This DT was used to identify and investigate various optimization problems, including a problem with excessively high air-heater flue outlet temperature. A solution was found which reduced outlet temperature by up to 25 °C without modifying the existing air heater, with a ~86 kJ/kWh improvement in heating rate. Although not classified as a 'highest fidelity' DT, it still proved to be a valuable tool for steady state optimisation.

J. Yu et al. (2020) has created a DT[133] model to replicate the flow behaviour of high-pressure (HP) control valves for steam turbines using a hybrid modelling method. Hybrid refers to using a combination of first principles physical mechanisms and control systems operational data for the model. The model calculates the flow rate of HP control valves, derives the control stage flow and efficiency characteristics, and obtains the characteristic functions.

As shown in Figure 2-3, the DT model uses these functions to calculate performance indicators, exit pressure and exit temperature. These simulated indicators are compared against their actual values from operations data, any differences indicate a component performance change and raises a maintenance or troubleshooting flag.

In terms of visual representations, the model uses simple flow diagrams and time series representations. Operational data is continually fed to the model which calculates the characteristic functions and parameters for the DT. In return, the DT provides performance monitoring data to the real system, thus meeting the digital manager classification by using a two-way flow of information.

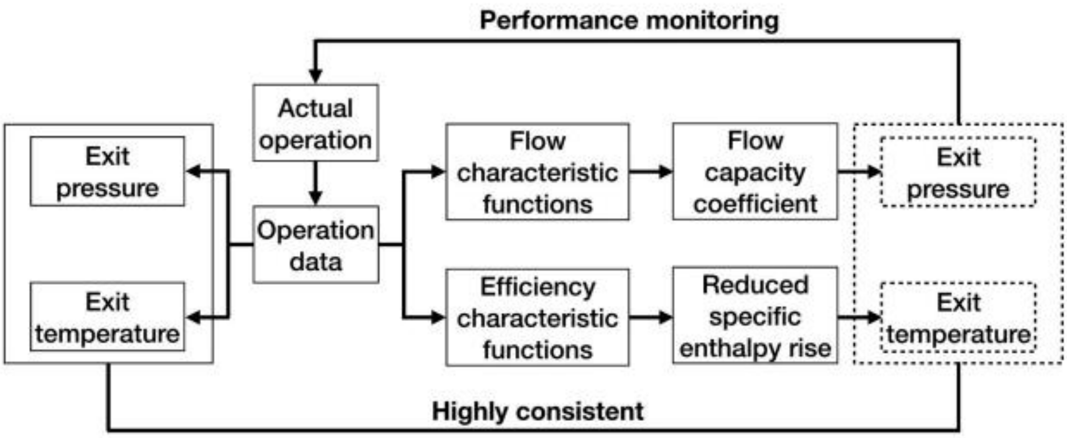

Figure 2-3: J. Yu et al.(2020)'s Digital Twin schematic diagram

When applied to a 330 MW subcritical steam turbine and a 1000 MW ultra-supercritical steam turbine, the model was able to accurately simulate exit pressure and temperature over the full unit working range. With relative errors of 0.5 % and 0.8 % for pressure and 0.2 % and 0.3 % for temperature (subcritical, ultra-supercritical).

The DT showed it was able to provide high precision real time simulation of control stages under dramatic load changes, at times fluctuations of greater than 30 % of rated power. The results prove the model to be effective for both subcritical and ultra-supercritical steam turbines using high-pressure control valves with different operating curves.

2.2.3 Types of Hydropower Digital Twin Research

As referenced in section 2.2, Digital Twins, the definition of a DT has a variety of different interpretations. This statement remains true for DT concepts in the hydropower research space. Although the number of research papers specifically referencing 'digital twins' for hydropower schemes are presently quite sparse, several papers have applied a couple different interpretations.

H.Yu et al. (2021) has used 3D laser scanning to build a one-to-one visual 'digital twin', a virtualized 3D model of the internal environment and equipment for a hydropower plant in Fuchunjiang, China. H.Yu et al.'s model consisted of scans stitched together using co-ordinate splicing technology to produce an accurate model. This research may not address this project's main aims in terms of modelling hydropower scheme behaviour but does offer a glimpse at one method of building a surface 'likeness' model for hydropower plants. This could be especially useful for older plants with incomplete drawings, missing drawings, or built without the assistance of CAD modelling. Interestingly, this research also references the use of this technology to scan full sections of concrete tunnels to detect surface defects, this could prove to be useful for mapping future repairs for Waikaremoana's concrete lined intake tunnels.

Dreyer et al. (2021) examines the use of a DT 'Hydro-Clone' commercial system to provide real time data about penstock integrity for the La Batiaz hydropower plant, Switzerland. The plant is often used for flexible generation and ancillary services, the increase in load variation and on/off cycling has made the penstocks more vulnerable to premature fatigue wear.

Hydro-Clone is a 'ahead of time simulation monitoring' software, designed to continuously model dynamic behaviour of the plant particularly hydraulic transients. It interfaces with the plant SCADA and PLCs to monitor for unusual events and simulates hydraulic and electrical component behaviour. It is capable of replicating in real time the pressure variations caused by hydraulic flow changes in the intake conduit. From these pressures it can calculate the stresses in the penstock using either analytical formulae or finite element modelling. The stresses are used as indicators of overall penstock wear and health. Dreyer was able to apply and validate this system by comparing with strain gauge results. Although an interesting and useful 'maintenance digital twin' concept, Hydro-Clone does not particularly deal with water management and high-level dispatch of a hydropower plant.

Wang et al. (2021) details an early project proposal, plans to produce a DT open platform framework specifically designed for hydropower systems. The research is based out of Oak Ridge National Lab in the United States (US) and is a long-term project lasting five years. The project is primarily meant to help address the issues with aging hydro infrastructure in the US and aims to build a platform that can collect hydropower data and continuously update dynamic models of various station components. The DT will run in parallel to real hydropower systems and allow users to optimise actual plant operation, test optimisation plans before implementation, perform fault diagnostics, and monitor the condition and health of components in the system. Plans also include the creation of visualisations with augment reality capabilities and the ability for the DT to receive real-time data and continuously update the dynamic models.

This project by Wang et al. looks to address all three aspects of the DT, replicating the system in terms of likeness, behaviour, and connection with the real system. Information from this group would have been invaluable for this thesis, however the group has only started to lay the groundwork for the project, with their first published piece detailing their progress expected in August 2022.

These research pieces on hydropower 'digital twins' demonstrate that development of true digital twins for hydropower scheme management have not yet begun in earnest in the academic field.

2.3 Turbine Efficiency

Maximisation of unit efficiency is a recurring objective in the hydro scheduling and unit commitment problem. The definition of efficiency can vary greatly depending on the area of study and the framing of the problem being solved. At Genesis, unit efficiency is evaluated in terms of the effective fuel (water) use per unit electricity produced, known as a k-value. Although this is a useful relative indicator of unit efficiency, it does not consider the differing potential energy values of water stored at different elevations. The following equation is the basis for estimating the theoretical energy potential of a water reservoir:

$$E_{Potential} = \eta V \rho g h_{Geo\ Cent} \tag{2-1}$$

Where $h_{Geo\ Cent}$ is the geometric centre height of the body of water.
Converting to calculating power generation using flowrate:

$$P(W) = \eta \rho q g H_{net} \tag{2-2}$$

Where P is the output power, η is the overall unit efficiency, ρ is the density of water, g is the acceleration of gravity, q is the water flow rate through the turbine, and H is the net head. Net head using the following:

$$H_{Gross} = h_{forebay} - h_{tailrace} \tag{2-3}$$

$$H_{Net} = h_{forebay} - h_{tailrace} - h_{penstock\ losses} \tag{2-4}$$

Penstock head losses may prove to be difficult to calculate due to penstock age and testing requirements to calculate penstock friction coefficient. Using gross head will instead incorporate variation due to penstock losses into the overall efficiency value instead, the difference is denoted due to this change using η^*. It should also be noted that density, ρ, varies with temperature. However, the density of water varies very little across the temperature range applicable for a temperate lake at atmospheric pressure. Water temperatures at the WPS lakes are not actively monitored but has been measured in the past.

Lake Waikaremoana's water temperature gradient was measured by Main (1976) in September 1972. Waikaremoana temperatures measured by de L. Main found that at a 9.3 °C average surface temperature, there was a 0.5 °C temperature gradient to a depth of 165 m. Below this depth isothermal conditions were observed. Figure 2-4 below shows the density variation of water between 0 and 30 °C using information from USGS.

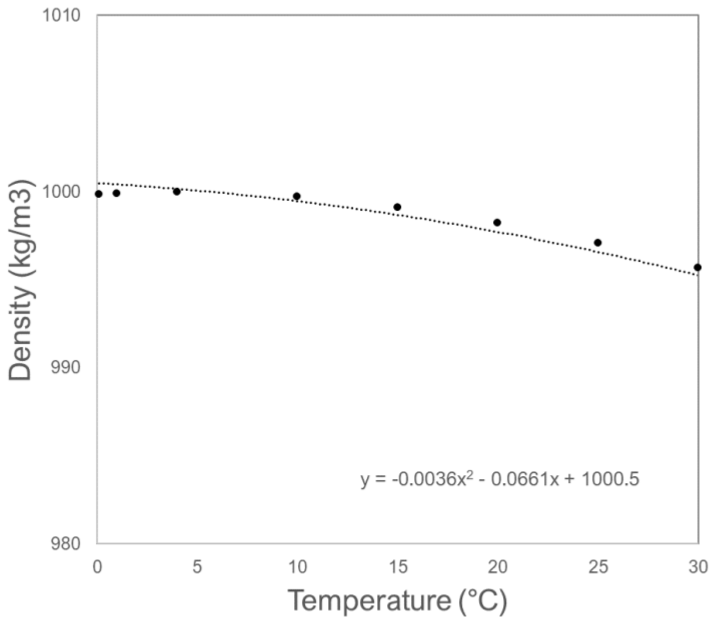

Figure 2-4: Water Density-Temperature Correlation (Engineering Toolbox, 2022)

Figure 2-4 shows that water density has an inverse parabolic relationship with temperature. The density variation in the lake's normal operating water temperature range between 8 and 18 °C (Hawke's Bay Regional Council, 2022) is only around 1.6 kg/m^3 or 0.16 %. This variation is negligible and thus density is assumed to be constant for calculations.

The efficiency equation is simplified with: $\mathrm{K} = \frac{\rho g}{10^6}$

$$P(MW) = \eta^* qH\mathrm{K} \tag{2-5}$$

Rearranging for η:

$$\eta^* = \frac{P}{qH\mathrm{K}} \tag{2-6}$$

Where η* denotes an overall efficiency that uses gross head. K (not to be confused with k-value) is known as the production function (Dal' Santo & Simões Costa, 2016) and is used to simplify this function. K = 0.00981 MJ/m^4, assuming gravitational acceleration = 9.81 m/s^2 and density of water = 999 kg/m^3.

The efficiency calculated in the above equation expresses the turbine's effectiveness in converting gross head potential energy into mechanical rotational energy. This equation shows that in the context of converting potential energy, k-values can only be used to compare hydro units with similar if not identical head values. In the case of Waikaremoana, k-values can be misleading as the units at Tuai Station are able to utilise over 200 m head whilst Kaitawa Station and Piripaua are only able to utilise 110-130 m head. Efficiency comparisons between stations using k-values may not be indicative of controllable machine performance but also uncontrollable head differences.

A water turbine in essence is a water pump that operates in reverse, although the impeller/runner designs will differ to optimise for pumping or extracting hydraulic energy. This fact allows pumped hydro stations to operate without separate pumping and turbine units. Similarly, to a water pump, turbine efficiency is a function of three operating parameters: flow rate, power, and water head. As a general trend, for a given net head, the efficiency of a hydro turbine is often expressed as a concave function. Increasing as discharge flow rate increases until maximum efficiency is reached; up to 95% for modern turbines. Increasing flow rate beyond this point will lead to a drop off in efficiency (Kong et al., 2020).

Unit efficiency in the equation (2-1) generally indicates the turbine efficiency of the unit but depending on where power is measured, can be expressed as the overall or global unit efficiency using a product of the turbine and generator efficiency (Kong et al., 2020) and used by Diniz & Maceira (2008).

$$\eta_{Overall} = \eta_{Turbine} \times \eta_{Generator} \times \eta_{Transmission} \tag{2-7}$$

Generator efficiency is defined as the energy transfer effectiveness of the generator converting rotational mechanical energy into electrical energy, this is typically greater than 95% and increases as a function of power output (Kong et al., 2020).

Turbine efficiency can be expressed as a function of power, flow rate and/or head. This can be in several different forms depending on the complexity required for the application. Increasing the complexity of the efficiency function improves the accuracy of the model but can introduce non-linearity or other computational complexity that increases the difficulty of optimisation methods. Some research papers simply assign a fixed efficiency value, using the average generation efficiency such Lima et al (Lima et al., 2013).

Others such as Daadaa et al. (2021) use dual efficiency values (a maximum efficiency, and a maximum power output efficiency) and expresses efficiency as the ratio of MW/m^3 like Genesis.

2.3.1 Hill diagram

Turbine efficiency is typically described by a hill diagram provided by the turbine manufacturer. The diagram presents the effective efficiencies in a topographical format as seen in Figure 2-5, considering all three turbine parameters: power generated, water discharge flow rate and net head.

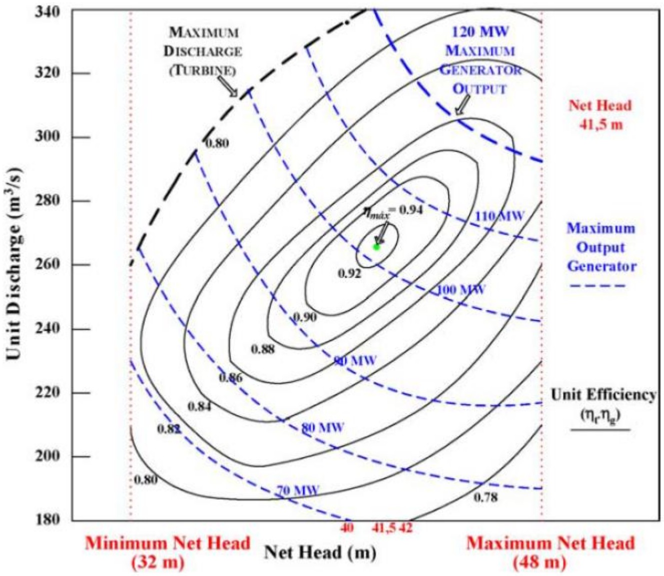

Figure 2-5: Example of a Turbine Hill Diagram (Finardi & da Silva, 2005)

The efficiency points of the hill diagram can be obtained directly from the manufacturer or adapted from interpolating hill diagrams in literature. Diaz (2011) derived an analytical expression from a hill diagram by sampling 378 points from the curves and passing these points through an algorithm on the statistical package, R, to estimate the coefficients of an efficiency function. This study however was based on the performance of a general Francis turbine hill diagram rather than a unit specific adaptation.

For older turbines, the original hill diagrams can become lost as manufacturing and operating companies restructure and change filing systems. In addition, as turbines age, the efficiency profile of the unit may change due to part replacement and wear/degradation of the runner, penstock, seals etc. Adapted hill diagrams may become too unreliable to provide accurate efficiency data.

2.3.2 Using operational data

Another option to map hydro unit efficiency profile is through regression analysis of operational data. Historical data for the three main turbine parameters are used to calculate a theoretical efficiency value using equation 2-6. This overall efficiency value will depend on all losses in the system, including turbine, generator, and penstock losses.

Regression analysis tests the relationship between one or more independent predictor variables (power, flow, or head) and the dependent variable (efficiency). The nature of the relationship between each set of predictor variables and efficiency is tentatively tested using a scatter plot of predictor variable against efficiency. The plot will show the general trends for the efficiency and may suggest a suitable regression model.

Simple linear regression forms the basis of regression analysis and is also used as a base for more complex models. Linear regression assumes that the relationship between predictor and efficiency can be approximated by a straight line. The coefficients of the line can be found using least squares methods and the fit of the model evaluated using statistics software such as Microsoft Excel's Analysis ToolPak package.

Multivariable and polynomial regressions can be considered variations of the linear regression. Multivariable linear regression considers multiple independent predictor variables and determines the effect each predictor has on the dependent variable. An important aspect of multivariable and polynomial regression is the independence of predictor variables or multicollinearity. The influence of each predictor cannot be accurately determined if the predictor variables are correlated. Multicollinearity can occur to different degrees and can be tested by calculating variance inflation factors for each predictor. Minor to moderate multicollinearity can be adjusted but strong multicollinearity may be difficult to fix without removing predictors.

Polynomial models introduce higher order terms for predictor variables and can be useful to model nonlinear behaviour for efficiency. The higher order terms are added as separate variables and analysed in a multivariable linear regression fashion. Zulkifli et al. (2015) modelled efficiency as a function of only one of the predictor parameters, power generated. The formula used in this paper was a 2nd order polynomial where a_1 and a_2 indicate the parabola shift parameter and rate of curvature:

$$\eta = a_1 P^2 + a_2 P + c \tag{2-8}$$

This formula does not consider the effect of water head on efficiency but could be useful in modelling the influence of power predictor on efficiency.

Dal Santo and Costa (2016), Finardi and Da Silva. (2005), and Diniz et al. (2007) all suggest that efficiency is a quadratic multivariate function of net head and discharge flowrate with the form:

$$\eta(H,P) = \gamma_0 + \gamma_1 H + \gamma_2 q + \gamma_3 Hq + \gamma_4 H^2 + \gamma_5 q^2 \tag{2-9}$$

Diniz et al. also suggests an alternative 4th degree polynomial function that builds on the above function:

$$\eta(H,P) = \gamma_0 + \gamma_1 H + \gamma_2 q + \gamma_3 Hq + \gamma_4 H^2 + \gamma_5 q^2 + \gamma_6 H^2 q + \gamma_7 q^2 H + \gamma_8 H^2 q^2 \tag{2-10}$$

This function could be an alternative model to investigate if simpler models are unable to adequately fit testing data.

Both power generated and water flow rate are predictor parameters, but they are not included together in models as they are strongly correlated with each other and inclusion of both creates strong multicollinearity. Logically, as water flow rate through a turbine increases, the moment on the turbine runner and generator shaft will increase. The rotational force of the shaft has a direct relationship with the amount of power generated and amount of load carried by the generator. The strong correlation between power and water flow rate has led to papers like Diniz et al. to only use flow rate or power generated with operating head.

Although papers such as Finardi et al. suggest using discharge flow rate and net head is preferred, Dal Santo and Costa noted that discharge flow rate could be substituted with power output.

$$\eta(H,P) = \gamma_0 + \gamma_1 H + \gamma_2 P + \gamma_3 HP + \gamma_4 H^2 + \gamma_5 P^2 \tag{2-11}$$

2.4 The Unit Commitment & Scheduling Problem

The hydropower management problem has been framed in several different ways in literature. The most prominent approach has been the scheduling, unit commitment and unit load dispatch problem. Unit commitment refers to decision making in when to utilise a unit to generate for a certain period, unit load dispatch deals with the running points of the generating units. Unit commitment only considers short-term generating decision making, considering only the next trading period or day ahead system state. Scheduling considers longer term scheme strategy and often involves utilising a function for longer term water management.

This problem has been studied extensively over the years and most pieces of literature examined for this thesis have taken this angle. Optimisation methods for this problem have often used a mathematical method to simplify and solve the problem numerically. Other methods such as object-oriented programming and reinforcement learning have been utilised by some papers but remain relatively uncommon for this problem.

2.4.1 Mathematical Approach

The mathematical approach to solve this problem is to define an optimisation problem in terms of an objective function and applicable constraints. The problem is solved using numerical methods such as linear programming (LP), nonlinear programming (NLP), mixed integer linear/nonlinear programming (MILP, MINLP), and dynamic programming (DP).

2.4.2 Objective Function

The objective function can be defined in a variety of different forms depending on the priorities of the hydropower plant operator. The most common objective function type focuses on the monetary aspect of running hydropower generators. This often develops into a plant/scheme profit maximisation objective, maximising the difference between revenue and losses. The function can include different terms depending on the complexity and factors included.

Shawwash's (2000) paper on short term schedule optimisation for a hydro scheme in Canada uses an objective function that maximises revenue for a system with both hydro and thermal generation. The revenue term is calculated as the product of the units of energy exported to grid and the per unit price of electricity. This revenue term is replicated in other papers with monetary based objective function such as Catalao et al. (2011) and Lima et al. (2013).

$$R_t = P_t \times S_t \tag{2-12}$$

Where R_t is revenue, P_t is power generated, and S_t is the electricity spot price for period t.

Papers by Diaz et al. (2011) and Aasgaard et al. (2018) apply objective functions that represent profit and use two terms to represent revenue. The immediate revenue received from generating and the theoretical future value of water held in the reservoirs. Different companies and groups can have different methods to determine this water value e.g., a water value calculation method can use the futures electricity market to determine the highest expected electricity price in a specific period. This price can be used as a general benchmark to determine if generation in the present is worthwhile over generation in the future.

$$R_F = W \times V_t \tag{2-13}$$

$$R_F = S_F \times k \times V_t \tag{2-14}$$

Where R_F is future revenue from generating, S_F is the futures spot price, k is the turbine conversion factor and V_t is the reservoir volumes in the current period.

The above-mentioned profit based objective papers also include a start-up cost term in their objective function. Starting up generators from a stationary state can place additional stress on turbines and generators, leading to faster wearing of parts and more frequent maintenance requirements. This term represents this cost by penalising start-ups using an approximated maintenance cost per start-up.

$$C_s = c \times y \tag{2-15}$$

Where C_s is the total start-up maintenance cost, c is the approximated maintenance cost per start-up, and y is a binary term that indicates whether the unit requires a start-up for the period. Bardsley and Chaudery (2000) uses an interesting approach to work around excessively starting and stopping generation units. Instead of optimising generation for smaller time steps (such as 30 min trading period), a time step of 8 hours is used. This sets a minimum generation length of 8 hours. However, this would only be applicable during times of steady spot pricing, a large change in spot price during the 8-hour period could result in a suboptimal solution.

Chang et al. (2001) also uses a similar profit-based objective, including terms for the value of available water and start-up costs. However, this paper does not include generation revenue for its application and instead includes a term for penalising missed reservoir target limits. However, it is stated that for deregulated applications, current period generation revenue should be included.

For all objective functions, the function total is calculated across several sets. The total studied time span is set, *T*, each time step within T is denoted by *t*. Each station or stage of the power scheme is labelled by the variable, *l*, and makes up the set, L. To differentiate each unit at each station, the set, I, is used with each unit represented by *i*.

The objective functions by Catalao et al. and Lima et al. do not include the reservoir future revenue value and are better applied to short term commitment problems. They include terms for present generation revenue and start-up cost:

$$Max \sum_{TLI} P_{tli}.S_t + \sum_{TLI} c_i.y_{tli}$$

(2-16)

For longer term optimisations, the full objective function by Diaz et al. and modified from Aasgaard et al. include the above terms and the term for total future generation revenue.

$$Max \quad \underbrace{\sum_{TLI} P_{tli}.S_t}_{Generation\,Value} + \underbrace{\sum_{LI} W_t.V_{Tl}}_{Stored\,Value} + \underbrace{\sum_{TLI} c_i.y_{tli}}_{Start-up\,Cost}$$

(2-17)

Papers such as Arce et al. (2002), Dal Santo et al. (2016), and Amani-Alizadeh et al. (2021) have taken a different approach and based their research on power stations with demand constraints; units are required to meet a set demand without regard to a market price. This could be deemed as the better strategy for regulated power markets. These papers have used objective functions that focus on minimising a loss of performance value (LoP). Factors that affect LoP include operating turbines at non-ideal performance points, reduced net water head due to flow dependent forebay and tailrace levels, and head loss in penstocks due to friction. These factors can be expressed in terms of power loss, LoP, and determine this loss by comparing against a reference value.

The three factors that contribute to the LoP term are losses associated with operating at suboptimal points on the turbine efficiency curve, head reduction due to tailrace elevation, and head loss from friction in the penstocks. Each have been represented by the below equations specified in Dal Santo et al.

$$LoP_\eta = K(\eta_{max} - \eta)qH$$

(2-18)

$$LoP_H = K\eta\left(H^{ref} - H\right)q$$

(2-19)

Where H^{ref} is a reference net head value. It has been previously defined by others, as the net head corresponding to the tailwater elevation at minimum turbine discharge.

$$h^{loss} = \alpha.q^2$$

(2-20)

$$LoP_F = K\eta\left(h^{loss} - h^{loss,ref}\right)q$$

(2-21)

Where h^{Loss} is the net head loss due to friction and α is the pipe loss factor. The LoP due to friction term, is also calculated by comparing against a reference head loss when operating at minimum discharge. These terms could be incorporated into an unregulated system with a profit based objective function by calculating the potential revenue loss due to LoP and subtracting from generation revenue R.

Another objective term to consider could be the reserves market, free or idle generator capacity can be bid into the reserves market for another revenue source.

2.4.3 Constraints

This section will outline the common constraints used in literature. Hydropower scheme constraints can range from simple numerical limits to complex functions. Kong et al. (2020) has a compiled a list of common problem constraints with their associated equations.

1. Reservoir water balance/Flow Continuity – a standard constraint for any system with a reservoir and water management requirements. The ΔT and 3600 term is a conversion for flow rate into volume.

$$V_t = V_{t-1} + \Delta T.3600(q_{Inflow} - q_{Outflow})$$

(2-22)

2. Storage Limits – restricts the capacity of the reservoirs to operate within the environmental consents and consists of upper and lower limits. Constraint is consistent across all income-based papers reviewed. Limits denoted by over and underbars. Chang et al. also added penalty variables and costs associated with spillage or falling below minimum levels. Can be converted to be in terms of volume.

$$\underline{h} \le h \le \overline{h}$$

(2-23)

3. Head Variation and Flow related head loss – significance as LoP was discussed above; they can also be considered in terms of constraints.
    a. Depending on the reservoir size, lake levels can vary with discharge flow rates. Catalao et al.'s study on a Portuguese hydropower system stated that this constraint was most applicable to small and medium sized reservoirs. Variables $h_{Forebay}$ and $h_{Tailrace}$ in these cases become functions of discharge flow rates. The relationship between discharge and tailrace level can be determined using operational data correlations (Dal Santo et al.), a significant relationship would indicate the need for this constraint. Skejlbred et al. mentions that spillway flows could also contribute to lake level variation and should be considered.
    b. Penstock losses are often defined as a quadratic function discharge flow such as Dal Santo et al. and Arce et al.

$$h^{loss} = \alpha . q^2$$

(2-24)

Loss factor, α, is specific to the penstock, depending on length, diameter, curvature, and roughness. An intake is often made up of different sections (tunnels, splits, and penstocks), making effective loss factor more complex and difficult to determine (Kong et al., 2020). Papers like Li et al. (2014) simplify the problem by assuming a constant penstock loss.

$$H = h_{Forebay} - h_{Tailrace} - h_{Penstock\ Losses} \tag{2-25}$$

4. Hydropower Production – another constraint consistent across all research papers examining the hydropower problem. This constraint uses the hydropower equation previously covered in the Turbine Efficiency section which defines the power, flow rate, and head relationship known as the unit performance curve.

$$P(MW) = \eta q H K \tag{2-26}$$

Where η has been defined as the global unit efficiency by papers such as Dal Santo et al. and Diaz et al. This efficiency as stated in the Efficiency section has been defined as a fixed value by Lima et al. or dual efficiency values by Daadaa et al. The most prevalent form is the concave polynomial function shown below, used by Dal Santo (2016), Finardi et al. (2005) and Diniz et al. (2007).

$$\eta(H, P) = \gamma_0 + \gamma_1 H + \gamma_2 q + \gamma_3 Hq + \gamma_4 H^2 + \gamma_5 q^2 \tag{2-27}$$

5. Discharge and Output Power Limits – sets the physical limits of the units when generating power. Units have a rated minimum and maximum acceptable value for discharge flow rate and power output. These limits will depend on the acceptable operating zones and rough running ranges on the unit performance curve. The operating zones are essential to avoid cavitation, excessive vibration, or inefficient running. The limits are often defined as fixed values, as in Dal Santo et al. and Catalao et al. or can be head dependent functions. The common definition of the constraint is shown below.

$$\underline{q}_{Dis}.\omega \leq q_{Dis} \leq \overline{q}_{Dis}.\omega \tag{2-28}$$

$$\underline{P}.\omega \leq P \leq \overline{P}.\omega \tag{2-29}$$

Variable, ω, is a binary variable that indicates the operational status of the unit for the current period (0: Off, 1: On).

6. Startup/Shutdown Status – constraints to define the start-up status of the unit given the unit operational status of the current and previous time periods. Determines the start-up condition for the unit based on the operational status for the current and previous periods. The first equation below is used by papers that include the start-up cost including Aasgaard et al. and Lima et al.

$$y_t \geq \omega_t - \omega_{t-1}$$

(2-30)

Finardi and Scuzziato (2013) uses a slightly different form below:

$$y_t = \omega_t(1 - \omega_{t-1})$$

(2-31)

Where y is a binary variable, mentioned in the objective function section, represents the start-up condition for this period. Subscripts t and t-1 denote the time step. A value of 1 indicates a start-up is required for the period and 0 indicates otherwise.

7. Coupling Mid/Long Term Strategy – Constraining the optimisation problem to a mid- or long-term scheduling strategy that can use load obligations, water level targets or price coupling.

2.4.4 Model Formulation Types

Depending on the complexity of the objective function and constraints selected, different exact optimisation formulations have been used to obtain the optimal unit commitment and scheduling solutions.

2.4.4.a Linear Programming

Earlier optimisation formulations for hydropower used linear solving methods due to the lack of computing power and strong optimisation algorithms.

Pekutowski et al. (1994) conducted a study to solve the short-term hydro scheduling (STHS) problem using linear programming for a hydro system in Tasmania, Australia. This model sought to maximise the amount of energy stored at the end of a study period. It featured many of the constraints previously outlined: demand obligations and level, discharge, and spill limits. Interestingly it used a piecewise hydro unit generation equation to approximate the unit performance and noted that in practice units would either run at the point of highest efficiency or highest power output. The piecewise linearisation of the turbine performance curves is a useful way to simplify the model but only produces accurate values at maximum efficiency and power output as seen in Figure 2-6 below.

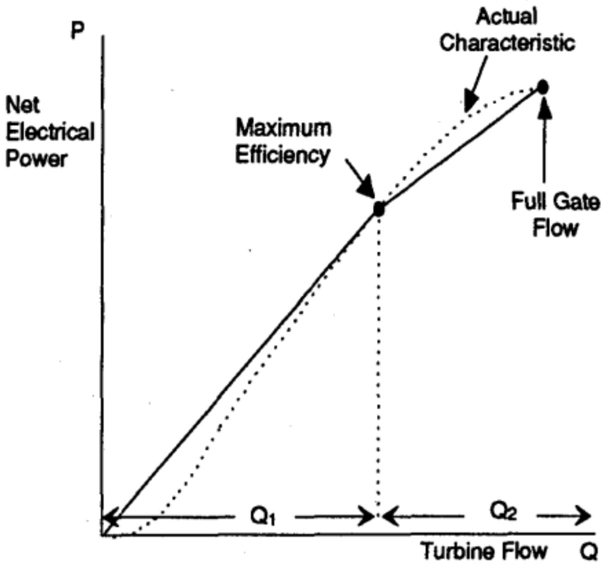

Figure 2-6: Unit Characteristic Performance Curve and Piecewise Characteristic by Piekutowski and Liwinowicz (1994)

Further constraints included coupling with longer-term strategy through target level constraints that penalised deviation from the scheduled target level band.

The model was solved using the MINOS linear programming package using a simplex method, it was able to resolve the model in an acceptable timeframe for a 24-hour study period but not for a 168-hour period. Schedules produced by this model was compared against historical operations records and showed annual energy savings of 0.3 to 0.4 % and up to 5 % stored potential energy savings on constrained days.

Shawwash et al. (2000) similarly developed a short-term optimisation model for a hydro scheme in British Columbia also using linear programming and solved using a simplex algorithm. It was used to determine optimal hourly generation schedules in a competitive power market.

The overall methodology included separate stages, the first was to assume a unit commitment combination using a Static Plant Unit Commitment program. The SPUC tabulated the optimal plant discharge for each increment of plant loading, forebay and for each unit availability combination. The objective for this stage was to minimise the plant discharges and used an AMPL (A Mathematical Programming Language) pre-solver to reduce the number of variables and constraints. In the main modelling stage, it used a revenue based objective function that included terms for hydro generation, thermal generation, and a future value of water for any volume differences from the target water storage. Similarly, to Pekutowski et al., the constraints for the system included the common limit constraints, flow continuity equation and a demand matching constraint. The model was solved using a CPLEX solver via the simplex method.

The model generated 24-hour generation schedules for seven major plants in the British Columbia Hydro system. These were compared against generation schedules prepared through traditional methods. Evaluation of results covering 185,000 MWh of generation showed that this model could be credited with between 0.25% and 1.0% of the optimised schedule generation value.

The model did have limitations, it did not consider the non-linear nature of power generation performance curves and flow patterns. It did not consider the cost of starting up units. Although it took into account the value of excess stored water, the long-term generation strategy was left assigned to water storage targets. Furthermore, due to the nature of linear programming, variables could only appear in up to two constraints to avoid over specification.

Shawwash et al. mentions that the model development process first started with a Microsoft Excel 5.0 linear solver model. Which "demonstrated the potential applicability of the technique and the benefits of modelling short-term operations".

Linear programming appears to be a good starting point for the short-term hydro scheduling problem, but its limitations may become evident when dealing with non-convex behaviour like that seen in hydro performance curves. LP models are easily solved using well-known efficient algorithms but do not fully represent the physical characteristics of hydro generation.

2.4.4.b Nonlinear Programming and Dynamic Programming

Nonlinear Programming (NLP) and Dynamic Programming (DP) offer methods to solve unit commitment models that use hydropower functions with nonlinear behaviour. Although a turbine performance curve may have areas of linearity, accurate performance models as functions of discharge flow rate and net head will often display nonlinear and non-concave behaviour.

Nonlinear solvers use gradient reducing methods to find the maximum or minimum points of a nonlinear function. Convergence to a global optimal point is not guaranteed and may depend on starting points. In the hydro commitment scheduling space, some studies have used nonlinear programming problem formulations but have solved them using dynamic programming instead of using nonlinear solvers. NLPs are commonly incorporated with a mixed integer component for mixed integer nonlinear programming method, discussed in the next section.

Catalao et al. (2006) has solved the scheduling problem using a purely NLP solving approach. The model is solved mathematically using quadratic programming (a type of NLP). The optimisation formulation itself uses a revenue maximising objective function, calculated with terms for the future value of stored water and revenue generated from power generation. It uses standard constraints: water continuity, hydropower function, head variation, and limits (storage, level, discharge).

The model was applied to a system of three cascaded reservoirs over a time horizon of 168 h. The system of reservoirs had net head values that were highly dependent on flows, with net head able to vary between 0 and 100 % net head. The study looked at four different storage management strategies to maximise generation profit.

The results were compared against a LP model, each of the four cases showed a small increase in profit. The largest at a 7.86 % increase was from Case 4 where Res. 3 storage was kept high and resulted in a higher cumulative net head. The paper outlined the importance of head dependency but was not yet able to define clear rules about the ideal storage volumes for the optimum hydropower scheduling.

It showed the next step from a LP model with nonlinear components but does not yet consider the cumulative cost of starting and stopping units. Furthermore, the system studied has significant head variations based on flow, indicating the reservoirs are on the smaller side.

DP has been a popular approach for solving models with nonlinear and nonconcave functions. DP breaks the problem into smaller subproblems. Each subproblem is presented as a stage and solved as an ordinary optimization problem. For each stage there is the concept of a state which reflects the information required to assess impact the current state will have on future actions. A decision is taken for the stage that produces a return or cost; the overall goal is to maximise this return (or minimise the cost). DP uses a recursive optimisation procedure to map out the optimal decision-making policy using the 'principle of optimality' (Bradley et al., 1977). For discrete systems, DP suffers from the curse of dimensionality; In hydro scheduling, an increase in the number reservoirs increases the number of variables and calculations exponentially. DP is usually limited to systems with no more than three reservoirs (Pérez-Díaz et al., 2010).

Perez-Diaz et al. (2010) utilises a novel DP model to solve the hydro scheduling problem for a single plant with multiple units fed by a small reservoir. Perez-Diaz et al. has decomposed the STHS problem into unit commitment (which units will generate) and unit load dispatch (how much each unit committed generates) subproblems but has solved them simultaneously using DP.

As a first step of the model, DP pre-processing is used to obtain a 'cloud' of plant operating points (generation characteristics) shown in Figure 2-7, representing optimal instantaneous unit commitment and dispatch for a given reservoir volume and water discharge.

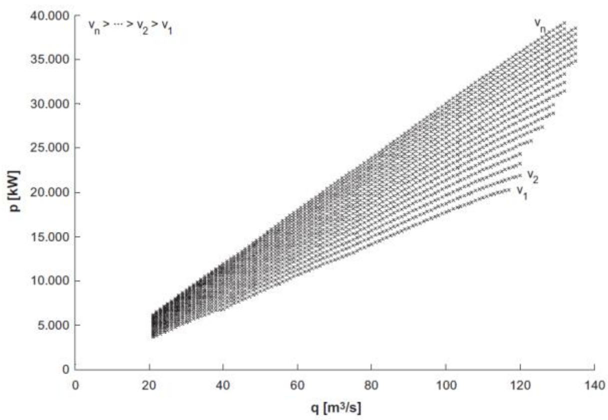

Figure 2-7: Plant Generation Characteristic 'Cloud' (Pérez-Díaz et al., 2010)

This was followed by formulating the optimisation problem for short term hydro commitment. The model uses a revenue maximisation objective function and has an NLP formulation as it considers head effects for a nonlinear hydropower characteristic constraint. Also included in the model are considerations for unit start-ups and shutdowns, head variations as a function of level change, and changing operation limits based on actual head. Unit start-ups and shutdowns are considered implicitly as part of the hydro generation characteristic functions instead of using generation state variables, the inclusion of which significantly increases problem complexity.

The optimisation model was solved using DP. Scheduling was broken down into hourly stages. The maximum revenue for each stage was calculated backwards sequentially from the last stage to the first using a recursive equation. The model optimisation results were compared against an NLP model for 15 different daily cases and showed an increase of between 0.9 and 3.3 % increase in revenue.

It should be noted that this study looks at a single hydro plant with a small reservoir where constrained flow plays an important role, optimal operating points are affected by how much volume and flow will be available. Although start-ups and shutdowns are addressed as part of the generation characteristic, previous stage unit states are not taken in account.

2.4.4.c Mixed Integer Programming

Improvements and developments in the computational solving algorithm space have allowed studies to address details of the scheduling problem that involve integer variables. Integer variables are a type of variable that is only able to take on integer values. A subset of integer variables are binary integer variables which can only take on the values of 1 or 0. Binary variables are useful to represent on and off states in a system or the occurrence of a state such as start-ups and shutdowns.

In hydropower scheduling, integer programming has been used to formulate the occurrence of start-ups/shutdowns and their total associated cost. Binary variables can also be used to represent simple forbidden generation zones for hydropower functions. Integer programming has been used in conjunction with linear or nonlinear models to formulate mixed integer linear programming (MILP) or mixed integer nonlinear programming (MINLP) models.

Catalao has produced several papers that use MINLP to solve the hydro unit commitment and scheduling problem. The most notable was a 'Profit-based evaluation' (Catalão et al., 2011). The formulation uses nonlinear functions and binary variables which places this study into the MINLP territory.

The objective function was profit based consisting of terms evaluating short-term generation revenue less unit start-up costs. Notable constraints: flow dependent tailrace elevation – linear or nonlinear, head dependent hydropower function. Binary constraint – start-up and shutdown status. The other constraints were standard: water balance/continuity, limits (storage, discharge, spillway). These limits can also be linked to binary variables.

The MINLP model was applied to a cascading hydropower scheme in Portugal, consisting of 7 different hydro plants. The plants are positioned and linked in a mix of series and parallel connections. The reservoirs for each plant vary in size, with only 3 and 6 considered storage-plants while the rest are considered run-of-river. Run-of-river in this paper refers to plants without significant water storage rather than a general term for plants pulling water from upstream and discharging downstream. The optimisation uses an hourly time interval over a one-day time horizon. The model was implemented using MATLAB and solved using an optimisation solver package, Xpress-MP. Results from the model were compared against NLP and MILP results although the start-up terms were removed for comparisons against NLP.

MINLP results were comparable to NLP in terms of profit, average discharge, and storage, however the compute time for MINLP was over twice the length of the NLP model. Against MILP, results for MINLP showed small increases in storage and profit, approximately a 1 % increase for a time horizon of one day. Compute time for MINLP was again much longer than the simpler model, at 9.26 s against 1.75 s for MILP. Compute times for MINLP were relatively longer than the simpler models, however, the absolute compute lengths were still within the acceptable time lengths. Catalao et al. concluded that while the compute times were longer, the results from the proposed MINLP model were more realistic due to the use of the nonlinear head relationships.

MILP formulations are like MINLP models but follow a method similar to Pekutowski et al.'s steps in linearising the nonlinear functions using piecewise linear approximations.

Chang et al. (2001) uses a MILP formulation to determine optimal generating schedules for the short-term hydro scheduling problem. The model includes variables specific to pumped hydro systems, but the general model can be applied to both conventional hydro and pumped storage systems.

The structure of the model follows the general structure of models already covered in this section but includes several different additions. A profit based objective function is used that considers total reservoir water value, variation from target water levels and start-up costs. Several uncommon constraints have been added in addition to the standard hydraulic constraints. The most notable of these are:

- Spinning reserve requirements are specified in a similar fashion to demand constraints seen previously, the possibility of adding a spinning reserve revenue term is also discussed.
- Binary constraints for pumped hydro, the system cannot simultaneously generate and pump.
- Minimum up-time and down-time limits, although start-up costs may address this issue, generators may require more clearly defined run duration limits.

The model was tested on two real systems and solved using a modelling/optimisation package: AMPL as the modelling language and CPLEX as the solver. The first system is referred to as the Southern Generating Group in the South Island of New Zealand. The hydro system consists of 9 hydro plants and 39 generating units, with most being a part of the Waitaki Hydro Scheme. The time horizon was a 24-hour period with 48 half-hour scheduling (spot market trading) intervals. The second system was the Swiss Rail hydro system, consisting of two separate river systems with a total of 7 hydro plants and 32 generating units. The scheduling is resolved over a 24-hour horizon at hourly intervals.

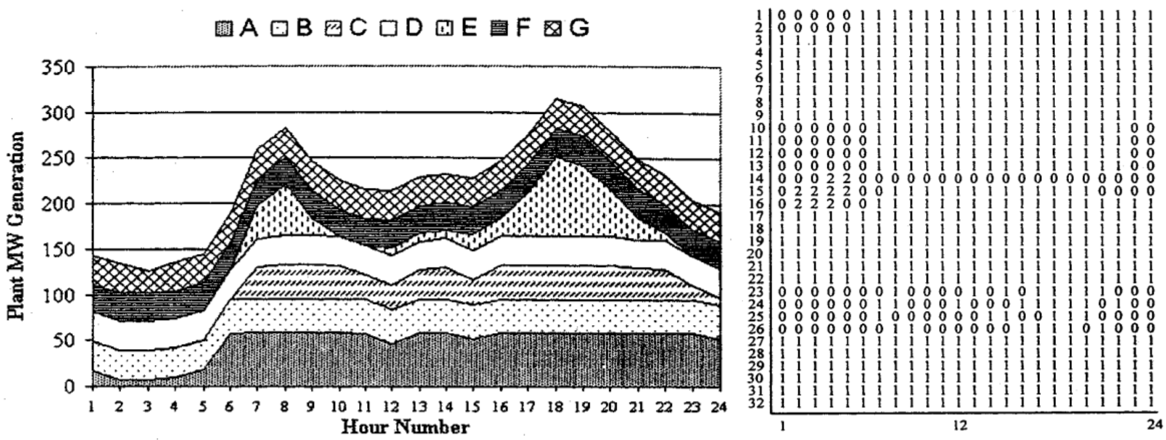

Figure 2-8: Scheduling (Left) and Hourly Hydro Unit Commitment Status (Right) for System 2 (Chang et al., 2001).

AMPL as a modelling language uses a pre-solver to simplify the problem before solving using a solver, CPLEX, via interior point, and branch/bound methods. Results from the model provide MW amount schedules and unit commitment statuses as shown in Figure 2-8. The model was able to solve a relatively large, simplified problem size in an acceptable amount of time, the first test system had around 4000 rows and 6400 columns of data and took under five minutes to resolve a day study using an old 233 MHz Pentium IBM workstation, archaic compute power by today's standards.

Results showed sensible optimal solutions for both systems, but no mention was made of comparisons against conventional scheduling methods or schedules found using other numerical models. Chang et al. concluded that MILP provided a simple, efficient, and suitable method for decision support. MILP could be a viable alternative to MINLP if complexity and compute efficiency become notable issues.

2.4.5 Object Oriented Programming Approach

Objected Oriented Programming (OOP) is a method of software development using an OOP coding language. 'Object Oriented' refers to the modelling of each system component as a data object, each with defined properties and attributes. These objects are created using Classes, an object blueprint or template that specifies 'Properties/Attributes' (information fields held by the object) and 'Methods' (class specific commands which objects created from the class can call).

OOP is useful for large and complex applications as the structure allows code to be easily reused through instantiation and inheritance. Instantiation is the creation of instances or objects with defined data structures using classes. Inheritance is a feature when creating a subclass using an existing class, the 'child' class inherits the attributes and methods of the 'parent' class without the need to rewrite any of its code. The child class can introduce new or changed, attributes and methods without affecting the parent class.

Use of OOP in hydropower optimisation related research works has been rare, with only a single group directly using OOP.

Garrido et al. (2009) has used an OOP based language, EcosimPro, to build a library of hydropower components and a simulation tool for small run-of-river hydropower schemes. Each major component in the system such as reservoirs, generators, turbines, gates, control systems etc. were defined as a class or subclass. Each class defined the relevant physical principles equations for their corresponding component. This approach provided the ability to easily build systems of different configurations by connecting modular components until a top-level model is complete. The group also developed a VisualBasic simulation tool to configure the model parameters, simulate the model and test its performance.

The library was used to build models based on real run-of-river stations, including the Villafranca, El Carpio, and Marmolejo hydropower stations in Spain. The model for Villafranca tested several inflow rates scenarios to predict the expected unit performance. The results were then validated against real operational data.

The paper details the equations used to build components in the library and shows some blocks of EcosimPro code. Although the code logic is similar to other OOP languages, more knowledge about the EcosimPro syntax would be needed to fully understand the code blocks. For the simulator, the paper outlines the logic flow diagram of the simulator but does not go into depth on how the simulator is programmed and how the parameters are updated. The modularity of components and the ease of adapting the library for different configurations could make OOP a worthwhile approach.

2.4.6 Reinforcement Learning Approach

Reinforcement learning (RL) is an exciting subset of machine learning that has seen a quick rise in recent years. It attempts to model the real-life learning behaviour of animals, where an individual can continually learn from past experiences to maximise decision making in a particular system. An RL system consists of an ‘agent’ and an ‘environment’, the agent is a learning algorithm that can interact with the environment through ‘actions’. The relationship is seen in Figure 2-9. Actions on the environment causes changes to the environment and provides the agent with an environment ‘state’ and a ‘reward’; a measure of how good the action was, trains the agent toward the desired behaviour (Riemer-Sorensen & Rosenlund, 2020). The agent aims to maximise the reward received over the planning horizon.

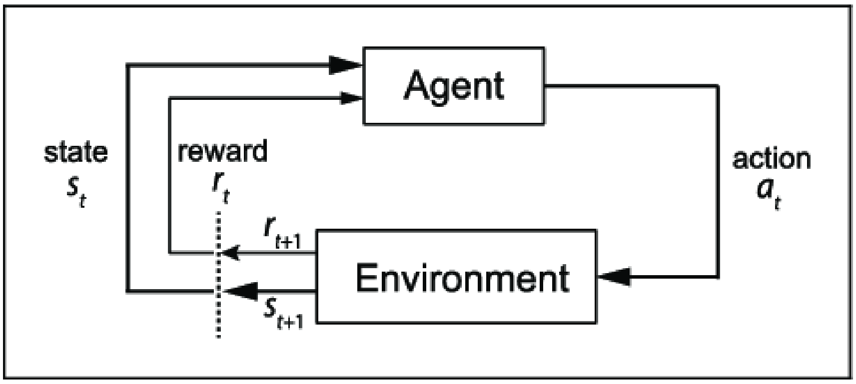

Figure 2-9: Reinforcement Learning Diagram

The agent is trained on test data until it can recognise the optimal decision for many possible states; the known optimal decisions for all states is called a 'policy'. The policy is formed by evaluating the value function at each state, the expected total reward for a given state. The value function helps the agent to measure the long-term outcome of a state. Q-Learning is an algorithm commonly used in RL applications to teach an agent and generate a Q-table or policy by slowly updating the table depending on a learning rate and a discount rate. Deep learning (DL) is often integrated into Q-Learning to create a Deep Q-network (DQN) and develop the regular RL agent into a deep reinforcement learning (DRL) agent. The DQN replaces the simple table with a 'black box' artificial neural network which accepts input states and outputs optimal decision values and corresponding actions. Once the agent is trained, the model can be applied to test data to evaluate the performance of the agent and model.

RL uses a similar principle to dynamic programming by using a recursive function to produce a state and an iterative optimal policy. The distinction between the two is that DP assumes that the model has a complete mathematical picture of the problem, each optimal decision for the policy is found through mathematical evaluation. RL has an artificial agent figure out these optimal decisions through a trial process, trying each decision until an approximate optimal decision is reached.

As the popularity of RL has taken off, so have studies exploring the application of RL agents to solve the hydropower scheduling and commitment problem.

Riemer-Sorensen (2020) aimed to maximise the yearly revenue of a hydro system given week by week inflows by training a soft actor-critic (SAC) algorithm and using Deep Reinforcement Learning (DRL). SAC algorithms are an expansion of Q-Learning algorithms and are suited to deal with stochastic variations and continuous action spaces. SAC models separate the actor's decisions from the value function table used in Q-Learning and introduces a variable, entropy, that measures the randomness of the policy. The SAC attempts to maximise both long-term reward and entropy, high entropy ensures that the agent makes sufficient exploratory actions to avoid converging towards a suboptimal local optimum. The environment state is defined by week number, storage, weekly spot price, weekly inflow, and number of weeks to empty reservoir if running at full capacity. Action space is defined by the water discharge. A step function determines the environment reaction to the action, the next state and the amount of reward received.

The agent was trained over 300,000 episodes (an episode is each training instance where the agent takes actions until a terminal state is reach while attempting to maximise total reward and entropy) first using artificial data with obvious price and inflow patterns followed by training using 60 years of historic data for the NO3 region from the Norwegian Water Resources and Energy Directorate. The agent only needs to be trained once on each training data set but took a day to complete on a relatively modern processor. The trained model could provide a complete 52-week generation schedule in less than a minute.

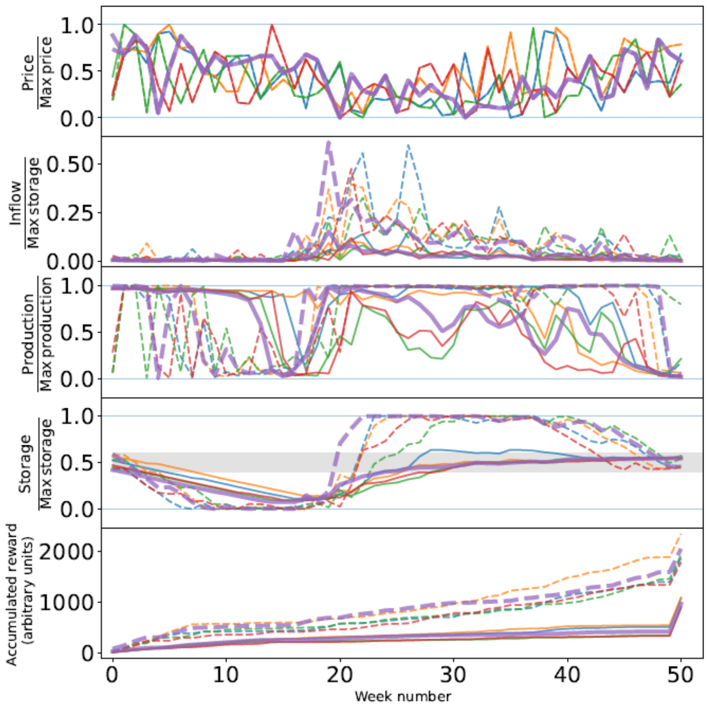

Figure 2-10: Schedule results using historic data scenarios, Higher Production Capacity (Dashed), Lower Production Capacity (Solid)(Riemer-Sorensen & Rosenlund, 2020)

The graphs from the study in Figure 2-10 shows the schedule parameters in 10 scenarios for a 52-week period using historic data. Result schedules showed the agent generally responded to increases in spot prices with higher production. During periods of lower spot prices, the agent attempts to save water according to predicted inflows and how quickly the reservoir could be depleted. Although the study tests the agent using historic data, it does not yet compare the performance of the agent against traditional trader methods for scheduling.

It should be noted that this agent was only trained and applied to a system with a single stage and linear relationships, the authors only aimed to demonstrate the minimal viable demonstration of this algorithm. This was very much an early exploration into applying DRL to hydropower.

Studies using DRL, show a promising approach to the hydro scheduling problem, it provides a robust automated process for solving this problem. Traditional methods require considerable computational power and will often require a human to decompose more complex systems. Building a DRL model would be a viable long term and comprehensive model goal but development was beyond the scope of this project.

2.5 Summary

- Researchers have created a classification method for DTs based on the level of fidelity and connection. EDTs are a sub-category of DT, coined for DT systems focused on managing energy streams in processing plants with the goal of improving energy and fuel efficiency. A DT for WPS could fall into this category as its overall goal relates to efficiently using its fuel (water).
- Selecting the correct level of fidelity is important to ensure computing resources are not overly taxing and resolution times are within an acceptable time horizon. The level of fidelity will be dependent on the application and its requirements.
- Turbine efficiency can be calculated in different ways, k-value is often used as an easy way to measure water use but does not consider the efficiency-head dependency. A traditional method for calculating the efficiency of a hydro unit is using the hydropower function.
- Turbine efficiency functions known as the characteristic function can be found by conducting regression analyses on unit historical operations data. Hydropower unit commitment researchers have proposed several different forms, the most promising being a multivariable polynomial regression with two predictor variables.
- The optimisation of a hydropower scheme participating in an electricity market is termed as the unit commitment and scheduling problem.
- Mathematical approaches are the most common method to optimise the problem among research papers. A profit-based function is the most widespread type of objective function, balancing the revenue from generation with potential revenue through storage. MINLP appears to be the most comprehensive formulation for the problem but requires more complex solving algorithms and may have more problems with non-convergence of the optimal point.
- Another promising method for solving the hydropower commitment problem is through a machine learning method, reinforcement learning. A successfully trained agent could act in a similar manner to a real trader, able to assess the current system state and makes decisions based on its hidden algorithmic knowledge.

Chapter 3
Background

3.1 Introduction

Chapter 2 presented the theory and past research behind three important aspects of building a support tool DT. It provided information regarding the definition of a DT and its current applications in both the hydropower and non-hydropower related settings. Subsequently, it discussed the function to evaluate turbine efficiency, along with methods of obtaining an efficiency model based on operational data by way of regression. Finally, the topic of optimisation was covered in detail, including traditional ways to solve the problem along with newer promising methods.

This chapter will provide contextual information about the different aspects of the project environment. The project environment consists of two main areas, the WPS, and the New Zealand power sector it participates in. Understanding the lake system, the hydrological factors, and the power generation assets available to it will be important for not only creating an accurate mass and energy flow model but also for understanding the WPS' capabilities and the many constraints it must operate within. The power sector is the other real-world system that must be considered and understood for hydropower management. WPS is a business asset, its operation is highly dependent on the electricity market and by extension so too are optimisations of the scheme operation. The sections below provide a well-rounded background information before discussing the role WPS plays in the market and the considerations that must be made by Genesis Energy when bidding WPS onto the market.

3.2 Waikaremoana

Waikaremoana, translated as the 'sea of rippling waves' from Maori to English, is a large naturally dammed lake located in New Zealand's northern Hawkes' Bay region. Nestled amongst rugged rainforests in the remote Te Urewera Park, Lake Waikaremoana in Figure 3-1, is among the North Island's largest and deepest lakes. Construction of the Waikaremoana Power Scheme (WPS) started in 1920's and was developed over the course of the 20th Century. It was built to utilise the hydropower potential of the lake owing to the convenience of the natural dam and the lake's altitude at 583 m above sea level. It is one of the world's only cases of a hydropower facility on a naturally dammed lake. The scheme consists of two other smaller lakes, Kaitawa and Whakamarino, each located at approximately 450 metres above sea level (masl) and 250 masl respectively. There are three power stations located below each respective lake (Kaitawa Station after Waikaremoana, Tuai after Lake Kaitawa and Piripaua after Whakamarino).

Figure 3-1: Waikaremoana Location (Left) & Lake Terrain Map (Right)(LINZ Data Service, 2022)

WPS is a cascading type hydro scheme where water discharged from each station is stored in the next lake before being utilised by the next station further down the scheme (excluding the final station, Piripaua). The widely accepted generation capacity for the scheme is 138 MW, utilising approximately 450m of head for 7 different turbine-generator units distributed between the stations. It is Genesis Energy's smallest hydroelectric power scheme by generation capacity; Tongariro (361.8 MW), Tekapo (190MW)(Genesis, 2022). Information regarding the origin of Lake Waikaremaona and the history behind the development of the WPS can be found in Appendix A.

3.2.1 Description

The WPS is made up of three power stations, Kaitawa, Tuai, and Piripaua, that draw water from three lakes of varying size seen in Figure 3-2, Waikaremoana, Kaitawa, and Whakamarino. The scheme is centred around harnessing the gravitational potential of water stored at Waikaremoana through several different stages or developments.

Waikaremoana has an area of 54 km^2 and an operating level range of 3 m. It receives inflows through a catchment area of around 350 km^2 (Taylor, 2019), averaging around 2440 mm on an annual basis. Total monthly rainfalls averages between 152 mm for the driest month and 270 mm for wettest.

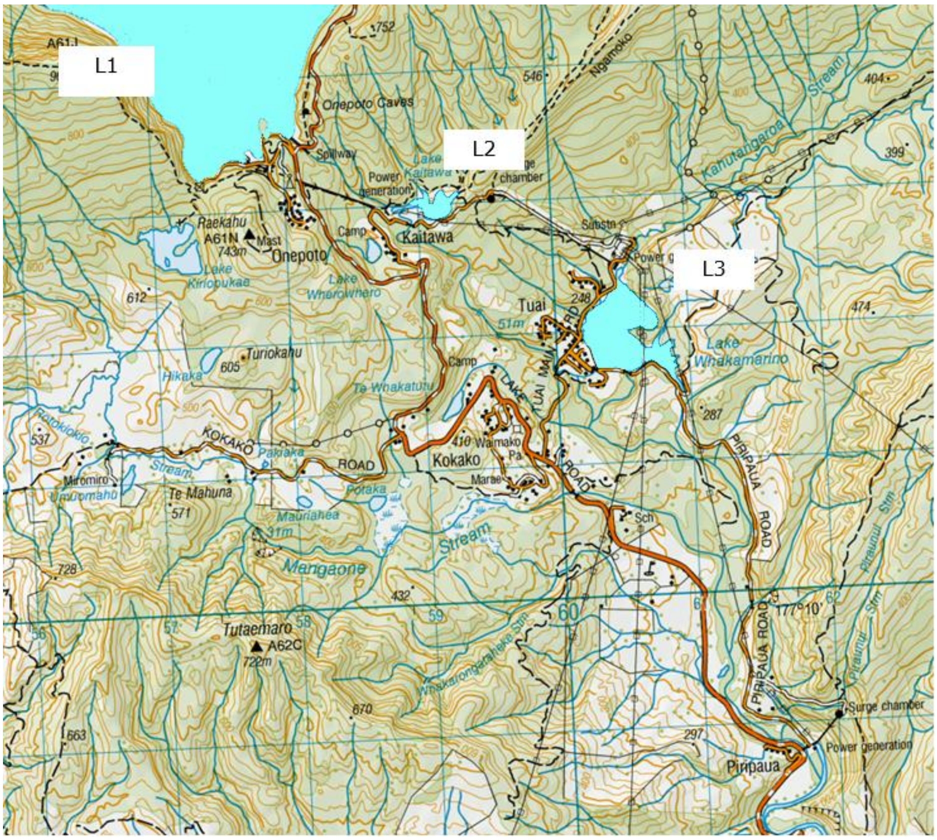

Figure 3-2: Waikaremoana Power Scheme Topographical Map (LINZ Data Service, 2022)

Waikaremoana's level is commonly at its peak during late spring and early summer and at its lowest in late autumn, early winter period. At maximum operating level, the volume of stored water available for use is approximately 159 million m^3, which translates to roughly 47 GWh when generating through Kaitawa Station at maximum efficiency. Even more energy can be extracted from this water by running it through the other stations in the scheme. Assuming no water is lost from the system, the above volume could theoretically generate a total of approximately 160 GWh.

Water can flow out of Waikaremoana through several avenues:

- through the Kaitawa Station intake at Onepoto, using a 220 m long, 3 m diameter tunnel (Engineering NZ, 2022), bifurcating into twin 270 m long, 2.4 m diameter tunnels (Taylor, 2019), then through two 670 m, 2.1 m diameter penstocks that reduce to 1.6 m (Taylor, 2019) as it reaches Kaitawa Station.
- through leaks in the lakebed (averaging around 5 m^3/s and increasing with lake level)
- through siphons that discharges water into a spillway leading to Lake Kaitawa.

Kaitawa Station, commissioned in 1947, is positioned on the banks of Lake Kaitawa and receives water through the Onepoto intake. The station has two hydro generation units, Units 6 and 7, fitted with vertical Francis turbines and each with a capacity of 18 MW. The turbines utilise a gross head of approximately 110 m and the turbine draft tube discharges water into Lake Kaitawa.

Lake Kaitawa is the smallest of the three lakes with an area of 0.061 km^2 and an operating range of 2.9 m. At full storage capacity, it holds enough for around 85 MWh generation at Tuai or 110 MWh if put through Piripaua as well. Aside from discharge flows from Kaitawa Station, Kaitawa also receives water through the siphon and spillway system, taking water from Waikaremoana via automatically primed siphons, passing through the spillway and into Lake Kaitawa. A sizeable portion of inflows into Kaitawa are from leaks through the slip material from Waikaremoana. The leaks form springs on the southern face of the slip material, flowing into the spillway and into Lake Kaitawa; on average the leakage flow measures approximately 5 m^3/s.

Water flows out of Kaitawa by way of the Tuai intake, 270 m long and 3.65 m diameter tunnel, leading into a 31.7 m diameter and 9 m deep surge chamber, then into three 1.1 km long and 2 m diameter reducing to 1.7 m penstocks once it reaches Tuai Station. Tip gates also allow water to be spilled into the upper Waikaretaheke River and to the Waikaretaheke Diversion. Water can continue along the Waikaretaheke River by passing through an automated gate or it can be diverted from the spillway to Lake Whakamarino using a gated canal which passes under the Waikaretaheke river using inverted siphons.

Tuai Station, opening in in 1929, is the main power station for the Waikaremoana scheme, housing the main control room and maintenance equipment and spares. It has the largest generation capacity of the three plants with three generating horizontal Francis turbine units: Units 1, 2, and 3, each capable of producing 20 MW of power. The turbines utilise a gross head of around 205 m before discharging flow into Lake Whakamarino.

Lake Whakamarino is the last reservoir in the scheme, covering an area of around 0.264 km^2 with a 1.4 m operating range. The operating volume of Whakamarino can store approximately 0.417 million m^3 of water which translates to around 111 MWh of generation based on the most efficient generation at Piripaua. Whakamarino receives water several paths: through Tuai’s discharge flow, by way of the Waikaretaheke River spillway from Lake Kaitawa and from the Kahutangaroa stream.

Whakamarino has similar outflow avenues to the other lakes, water can be passed to Piripaua via an intake on the eastern side of Whakamarino where a 2.6 km long, 4.5 m diameter tunnel takes it to the Piripaua surge chamber 32 m in diameter, 9 m deep. Two 2.6 m diameter around 548 m long penstocks connect the surge chamber to the Piripaua powerhouse.

Piripaua Station, commissioned in 1943, is positioned alongside the Waikaretaheke River with a nameplate generation capacity of 42 MW with both units running. The two units at Piripaua are Units 4 and 5, each utilise a gross head of around 110 m with a generation capacity of 23.6 MW each when running alone; the difference in total capacity and the lower sum of the individual capacities is attributed to Piripaua’s intake structure which consists of a single tunnel stretching over a far distance and involving inverted siphons. The long length of the single tunnel may have a considerable friction factor which results in more head loss at higher flows. Flow through the station discharges into the Waikaretaheke River.

Details about the three lakes are summarised below in Table 3.1.

Table 3.1: Summary of the Lake and Station

Lake	**Waikaremoana**		**Kaitawa**			**Whakamarino**	
Area	54 km^2		0.061 km^2			0.264 km2	
Station	**Kaitawa**		**Tuai**			**Piripaua**	
Average Head	120 m		205 m			110 m	
Turbine Type	Vertical Francis		Horizontal Francis			Vertical Francis	
Unit Number	**Unit 6**	**Unit 7**	**Unit 1**	**Unit 2**	**Unit 3**	**Unit 4**	**Unit 5**
Unit Capacity	18 MW	18 MW	20 MW	20 MW	20 MW	23.6 MW	23.6 MW
Stored Operating Water	159 million m^3		0.176 million m^3			0.417 million m^3	
Single Stage Stored Operating Generation	47,000 MWh		85 MWh			111 MWh	

A simplified PFD of scheme has also been drawn next in Figure 3-3.

3.2.2 Process Flow Diagram

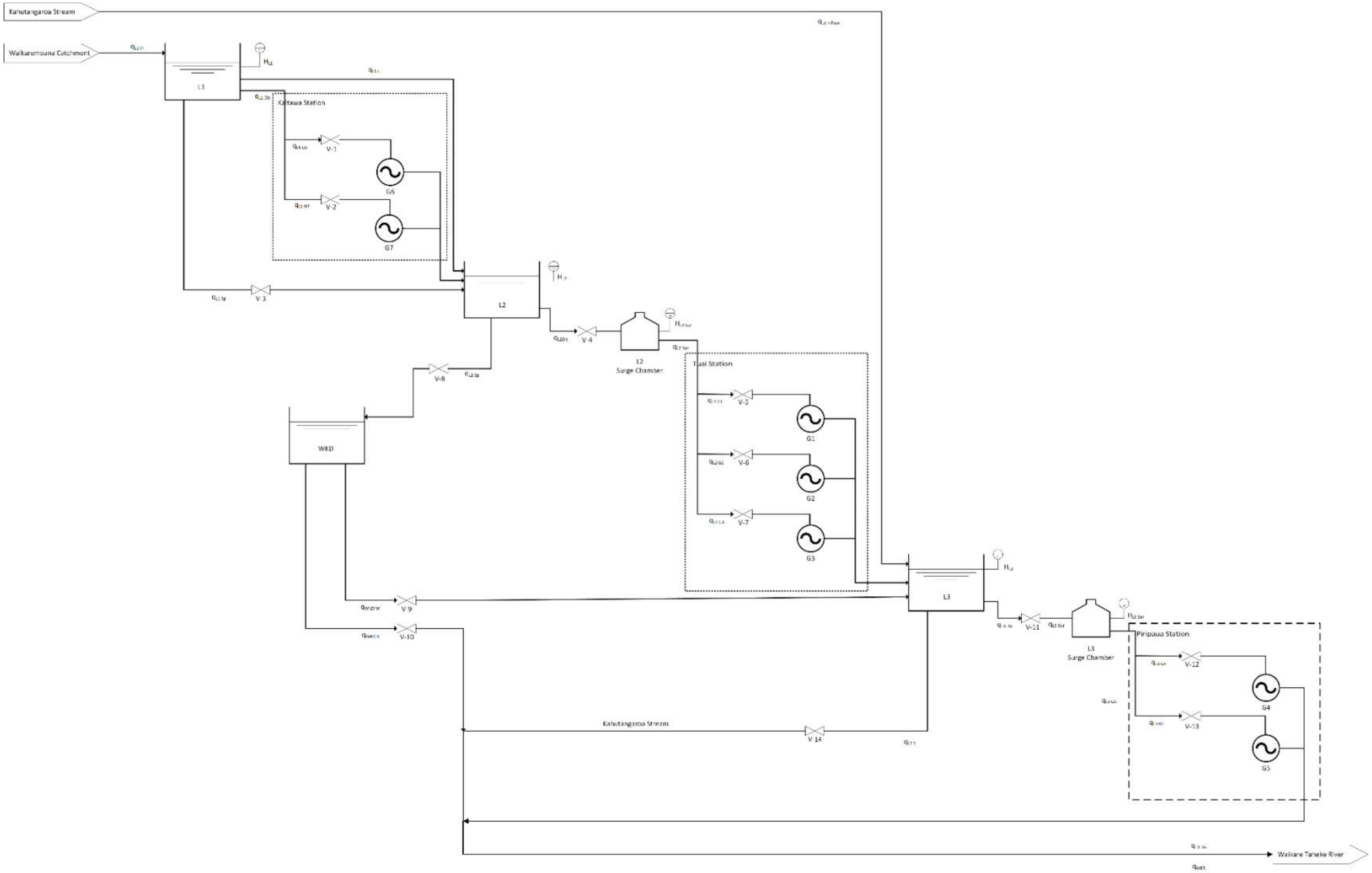

Figure 3-3: An Overview of the WPS in a Simplified Process Flow Diagram Format

3.2.3 Hydrology

Waikaremoana lies within the Wairoa basin with a catchment area of approximately 250 km^2. It is the North Island's fourth largest lake and its deepest, with a surface area of approximately 54 km^2, an average depth of 93 m, and a maximum depth of 248 m (Hopkirk, 2011). The current consented operating level is between 583.3 and 580.3 metres above sea level (masl); at its upper limit, it is estimated to have a useable water capacity of 157 million cubic metres.

Lakes Kaitawa and Whakamarino were flooded or enlarged specifically due to hydropower development but are still much smaller in comparison to Waikaremoana. Lake Kaitawa has a surface area of 0.061 km^2 and an operating level between 453 and 450.1 masl, and Lake Whakamarino 0.298 km^2 and level between 247.6 and 246.2 masl. Other characteristics are less well studied compared to the main lake.

3.2.3.a Inflows - Rainfall and Seasonal Variations

Waikaremoana has many streams flowing into the lake, around 114 of various size. These flows are difficult to measure due to their large numbers and their location within difficult to reach terrain. The largest streams in the catchment have been measured, the Aniwaniwa and Makau catchments, at 0.8 L/s/km^2 and 13.8 L/s/km^2 (Hopkirk, 2011)

The area surrounding Waikaremoana receives an average annual rainfall of over 2000 mm spread fairly evenly throughout the year. Mean annual rainfall data from different sites around the Waikaremoana area collected by the Hawkes Bay Regional Council (HBRC), is summarised by Hopkirk (2011) in Figure 3-4.

Catchment	Record begins	Mean annual rainfall (mm)	Min annual rainfall (mm)	Max annual rainfall (mm)
Erepeti Met	1928	1825.1	1165.3	2619.7
Aniwaniwa	1977	2232.4	1750.5	2892.7
Nga Tuhoe	1985	1683.6	1263.5	2255.0
Upper Waiau	1985	|1224.4	745.5	1547.5
Bushy Knoll	1986	1447.4	785.0	2352.5
Rocky Pad	1989	2144.2	1575.2	2754.0
Mt Manuoha	1989	2879.7	2164.0	2352.5
Waimaha	2000	1215.9	948.5	1466.3

Figure 3-4: Annual Rainfall at Sites Around Waikaremoana Provide by HBRC & Summarised by Hopkirk (2011)

Sites Aniwaniwa, Rocky Pad and Mt Manuoha are the locations closest to Waikaremoana, each with mean annual rainfall values above 2000 mm. Cant et al. (2004) has suggested a similar figure using data from 1921 to 1945. Rainfall tends to increase during the colder months, accounting for up to 40 percent of annual rainfall during the May-August period (Cant et al., 2004).

Rainfall measurements in Table 3.2 were also available from Genesis from a weather station at Onepoto, Waikaremoana. The rain fall data is consistent with the data found by Cant and the HRBC data organised by Hopkirk. The hotter months between Nov and Feb see the least rain while the highest rainfall occurs between June and September. Interestingly, August has on average noticeably less rainfall than months on either side which may be due to snow and ice formation during winter.

Table 3.2: Average Monthly Rainfall Data 2006-2021 from Genesis Onepoto Measurement (Total Rainfall in mm)

	Jan	***Feb***	***Mar***	***Apr***	***May***	***Jun***	***Jul***	***Aug***	***Sep***	***Oct***	***Nov***	***Dec***	***Total***
Min	31	50	63	82	38	94	62	88	53	47	59	57	*1535*
Mean	159	156	209	199	183	252	270	198	267	182	152	163	*2440*
Max	460	427	431	312	496	654	652	340	783	465	413	330	*4349*

3.2.3.b Outflows

Prior to the development of the Waikaremoana hydro upper development, the main natural outlet and only natural surface outlet for the lake was via overflowing of the natural dam. This flow was estimated to have occurred 50% of the time before hydro development at a rate of around 12 m^3/s (Anderson, 1948).

The other major outlet is subsurface flow via leakage through cavities in the natural dam, these flows emerge on the south face of the dam as small springs and streams. Tracer studies conducted by McPike (, as cited in Hopkirk (2011) have shown the passages through the debris to be complex with multiple and intersecting paths resulting in long tails and lag times for tracers. Water flowing from these springs and streams collect into the upper Waikaretaheke stream below Lake Waikaremoana and flows into Lake Kaitawa. Measurement of the upper Waikaretaheke stream provides a rough estimate for leakage flow, measuring roughly 12 m^3/s before sealing (Cant et al., 2004).

After the development of Kaitawa Station, work was undertaken between 1948 and 1955 to reduce leakage flow which bypassed the station. Partial sealing of the natural dam using a filter blanket of rocky debris, clay, and gravel was able to reduce leakage flow down to 4-6 m^3/s (Cant et al., 2004). Further sealing was planned but ultimately abandoned due to concerns related to altering the character of the environment. Additional sealing would likely destroy the linked natural springs.

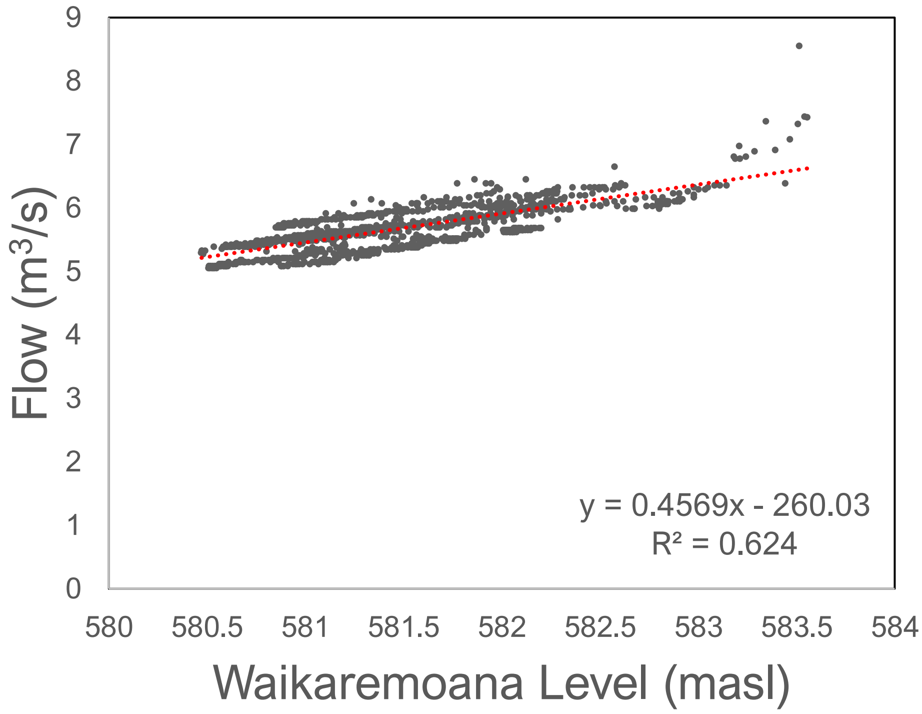

Figure 3-5: Waikaremoana Leakage Flow and Lake Level Relationship

Leakage flow is measured using the flowrate of the upper Waikaretaheke stream. Leakage data from recent years in Figure 3-5 show a positive linear relationship between lake level and leakage flow, generally between 5 and 7 m^3/s. This small increase of Cant's value could be due to erosion of the filter blanket or natural leak passages.

Surface evaporation in the Waikaremoana area was measured by Finklestein (1973) as part of a New Zealand wide study into evaporation. The study specifically measured open water evaporation using a tank evaporimeter at Onepoto, Waikaremoana. The measurement was used to calculate evaporation (mm) using a modified Penman equation. The modified Penman equation considers radiation, saturation, wind speed and temperature. Results showed that there was an average monthly evaporation of 50.42 mm over the year, peaking in January at 91 mm and dipping in June and July, both at 22 mm. Cumulative evaporation for the year totalled 595 mm.

3.2.3.c Water Storage

Due to its size, the majority of the power scheme's water storage is held in Lake Waikaremoana. Prior to hydropower development, lake level would follow natural rainfall patterns; reaching a peak after high rainfalls in winter-spring and falling to a minimum after the drier summer.

Following the development of the upper hydropower development for Kaitawa Station and implementation of lake level management, lake level is dictated by both rainfall and generation demand.

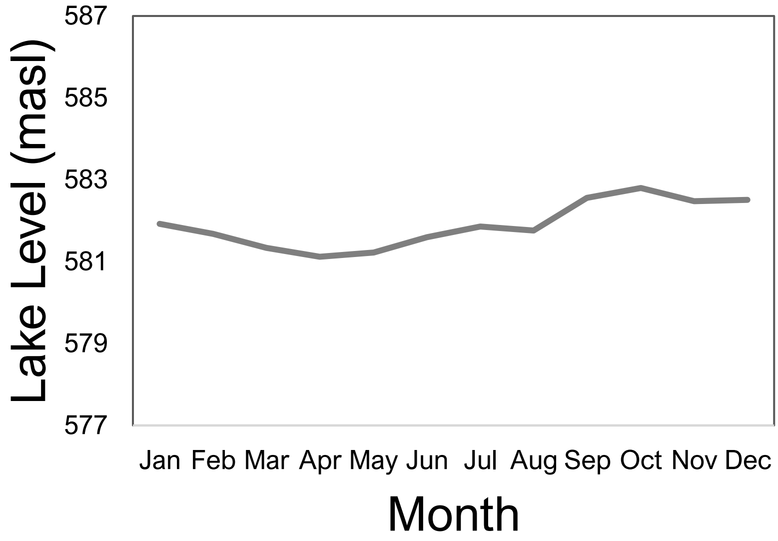

Figure 3-6: Average Monthly Waikaremoana Lake Levels 1997-2020

As shown in Figure 3-6, a monthly level average across the last couple of decades show that Waikaremoana's lake level is still highest in mid spring where high rainfall has accumulated, and peak generation demand has lessened as winter electricity demand decreases. Accumulation in winter is slowed by high electricity demands resulting in storage peaking mid spring. Water is generally saved and stored during the summer months and drawn down through generation and leakage in the lead up to winter.

3.3 New Zealand Power Sector

New Zealand has a liberalised power market and one of the least regulated in the world. It is a largely hydro-dominated system, though has comparatively limited hydro storage capacity. For much of its history, the electricity generation and transmission sector in New Zealand was completely Government controlled. It was not until the mid-1980s that it underwent a transformation through privatisation and corporatisation (Poletti, 2021).

In 1987, the New Zealand Electricity Department was corporatized into a state-owned New Zealand Electricity Corporation or ElectroCorp NZ (ECNZ). Several years later in 1993, transmissions services were split from ECNZ to form a new transmission operations company, Transpower - responsible for running, growing, and maintaining the national grid in its dual roles as Grid Owner and System Operator, as outlined below. To promote competition amongst generators, Contact Energy was started as a state-owned enterprise in 1996, using assets transferred from ECNZ (Shen & Yang, 2012).

To prevent distribution companies from hoarding savings in the wholesale market, the Government introduced the Electricity Industry Reform Act 1998. The Act prevented the dual ownership of distribution and electricity retail or generation businesses. Generation and retail business ownership was permitted and still exists today, with several 'gentailers' in the market which have a natural hedge between their generation and retail arms. In 1999 ECNZ was finally dissolved into three other state-owned generation companies: Genesis Energy, Meridian

Energy, Mighty River Power (renamed to Mercury in 2016)(Shen & Yang, 2012). These companies remained state-owned until 2013-2014 when they were partially privatised (remaining 51% Government-owned).

The electricity system today is made up of four sectors: generation, transmission, distribution, and retail.

Generation covers the sources of electricity to the grid. It comprises of generators that produce electricity through various sources such as thermal (mainly coal and natural gas), hydro, geothermal, wind and solar. New Zealand has five main generation companies (Meridian Energy, Contact Energy, Genesis Energy, Mercury and Trustpower) that own or control roughly 95% of New Zealand's electricity generation. These companies own 98 power stations and operate another 81 stations on behalf of other owners (Electricity Authority, 2018). Generators greater than 10MW are required to sell into the wholesale market.

Transmission refers to the national grid that connects generators to distribution networks or large consumers using high voltage power lines. The national grid spans thousands of kilometres across New Zealand and is owned and operated by a single state-owned enterprise, Transpower. Transpower is responsible for building and maintaining the grid including all transmission development processes such as obtaining consents, access rights and construction (Electricity Authority, 2018). Transpower's assets typically run from 50 kV up to 400 kV lines, including the 350 kV High Voltage Direct Current (HVDC) line which connects the North and South Islands. This includes a subsea cable system which runs under Cook Strait.

Distribution consists of the smaller lower voltage local power networks that distributes power from the transmission grid to individual homes or businesses. These networks are made up of substations and local level overhead wires and underground cables. New Zealand has 39 networks with most being owned and maintained by 29 distribution companies. (Electricity Authority, 2018)

Retailers sell electricity to individual customers and are required to purchase from the wholesale market (some large industrial consumers purchase direct from the wholesale market). The Electricity Authority has promoted an increase in the number of retail companies available for consumers to choose from, encouraging competition and innovation. In 2017, New Zealand had up to 48 different retailers across the country (Electricity Authority, 2018).

Transpower is also contracted as the system operator, who oversees the running of the national power system. Transpower is responsible for security of supply, quality, and ensuring adequate generation is available throughout the year by forecasting supply levels in New Zealand (Electricity Authority, 2018). Supply level is based heavily on the amount of water available for all hydro schemes in the country. In everyday operations, Transpower is in control of co-ordinating and matching supply with New Zealand's real time electricity demand for the lowest cost.

Lastly, the entire New Zealand Electricity Market is overseen by the electricity regulator, the Electricity Authority (EA). They aim to improve competition by increasing consumer choice and reducing market entry barriers instead of exercising direct regulatory control. The EA acts in an oversight capacity, monitoring the different classes of market participants to ensure they meet the requirements of the Electricity Industry Participation Code (EIPC)(Electricity Authority, 2018).

The wholesale electricity market consists of the spot market, hedged or futures market, and ancillary services market as outlined below.

3.3.1 Spot Market

The spot market is one of two main markets for market participants to purchase and sell wholesale electricity. Days are split into 48 half-hour trading periods, where generators with a capacity larger than 10 MW or generators connected to national grid can bid for the right to generate electricity to satisfy demand. Bids to buy and sell are made through the wholesale information and trading system (WITS) operated by New Zealand's Exchange (NZX). Up to five 'tranches' of price and quantity can be placed for each generating unit. Each bid specifies a quantity, price, and node location for the trading period. All offers for a trading period form an 'offer stack'. Some 'must run' generation can win the right to be allowed to bid in at $0.00/MWh to ensure it gets dispatched, including typically intermittent generation such as wind.

For each trading period the system operator is responsible for scheduling and dispatch, they must decide which bids in the offer stack to accept with the help of power flow analysis and security constrained economic dispatch programs. At a base level, offers are ranked by price. The system operator selects the combination of units able to reliably meet the required demand for the lowest cost. The problem becomes more complex when transmission constraints, transmission losses and security of supply is considered. The final spot price is determined by the bid placed by the marginal generator, the lowest cost generating unit required to meet the trading period demand. All generators dispatched are paid at this price for all their electricity generated in this period with variations depending on the node location. Generators can inject power into the grid at one of 52 grid injection points (GIPs), buyers are able to buy electricity from 196 grid exit points (GXPs).

The spot price varies according to supply and demand throughout the day, with typical peaks in the morning and early evening. The price can be quite volatile, spiking during times of low supply or high demand and dipping during an oversupply or low demand situation. Although over a long period of time, the spot price fluctuations tend to average out, the spot market and spot contracts are considered riskier ventures than hedged contracts. The hedged market is described in the next section.

In addition to operating the WITS, NZX is contracted to act as the Pricing, Reconciliation and Clearing Managers. The Pricing Manager, calculates and publishes the final clearing price for each node, these prices are used to settle transactions in the trading period. The Reconciliation Manager is responsible for reconciling the amount of power consumed against the amount of power injected into the grid according to meters.

The Clearing Manager uses information from the Pricing and Reconciliation Managers to ensure sellers are paid the correct amount for what they generated, and buyers pay the correct amount for what they consume.

3.3.2 Hedged Market

The other major submarket for wholesale electricity is the hedged market. To negate the volatility of the spot market, the hedged market allows consumers to buy contracts that guarantee the supply of electricity for a hedged or fixed price and vice versa for sellers. Although highs and lows in the spot market are generally expected to negate each other over a long period, hedging allows consumers and generators to manage risks associated with spot market price fluctuations, providing price certainty for a contracted period. (Electricity Authority, 2018) As an example, a generator may opt to cover shortages from planned plant outages with hedges to cover its retail position. Generally made up of financial instruments, this market is further divided into:

Over the Counter (OTC) - which involve direct buying and selling between market participants. Types of OTC can be contracts for difference, fixed price fixed volume, fixed price variable volume, and options (Electricity Authority, 2018).

Futures (ASX) – Futures contacts are traded through the Australian Securities Exchange and allows market participants to trade 0.1 MWh NZ power contracts for monthly, quarterly, or annual periods. The smaller size and standardisation of contracts allows smaller retailers to participate more easily. The Futures markets also lets participants observe the forward price of electricity - the market expectations for future spot price levels. Price expectations are updated daily for periods up to 3 to 4 years in the future. These prices will also try to reflect future changes to the market such as incoming generation investment (Electricity Authority, 2018).

Financial transmission rights (FTR) markets – allows participants to manage risks associated with the node-node price differences. These node prices can fluctuate significantly and can depend on the current generation offers, reserves, transmission constraints and losses in the grid. FTR contracts allow retailers to fix the spot price between two nodes, providing them certainty that their hedged contracts at a particular node can be used when retailing at other nodes without the need to account for locational price differences (Electricity Authority, 2017).

3.3.3 Ancillary Market

The ancillary market provides a place where grid support services can be offered and bought by generators. Each service helps to ensure grid integrity is maintained if an unexpected event occurs. The services are:

Frequency keeping – Frequency measures how well supply and demand are matched and must be kept within small margins for grid integrity reasons. Frequency keeping requires the provider to increase or decrease supply to ensure frequency is kept at 50 Hz. (Electricity Authority, 2018)

Instantaneous reserve – providing immediate back up generation in the event a supply failure event occurs such as another generator failing or a transmission line suddenly losing connection. This prevents overloading of generators still connected to the grid, frequency decay and subsequent grid collapse. Instantaneous reserve is made up of fast instantaneous reserve (FIR) and sustained instantaneous reserve (SIR). FIR is available within six seconds and must be able to operate for one minute. SIR is available within sixty seconds and must be able to operate for fifteen minutes. Instantaneous reserve is procured based on the size of the single largest risk that could occur during a particular trading period (typically the largest generating unit)(Electricity Authority, 2018).

Over frequency reserve – providing a quick decrease in generation in the event a significant load is unexpectedly disconnected from the system (e.g., an earthquake cutting a city's connection). This ensures the grid frequency is kept within the normal band and prevents damage to online generators from over speeding (Electricity Authority, 2018).

Voltage support – a voltage regulating service that provides or absorbs reactive power in the grid. Reactive power is 'additional' power that is required to generate electrical fields in motors, transformers, and other electrical components with coils but does not translate to meaningful work (Electricity Authority, 2018).

Black Start – providing generation from a full shutdown state in the event of a full grid outage. The provider must be a generator that can start without any external power from the grid. They will generally have an auxiliary power system that assists with generator start up. This will allow the system to return to service in a managed way, incrementally matching load with generation (Electricity Authority, 2018).

3.4 Waikaremoana Power Scheme Role

As discussed previously, the spot market clearing price is determined by the lowest cost generator required to meet the trading period demand. This generator is known as the 'marginal generator'. In a mixed thermal-renewable energy system the marginal generator is commonly set by the thermal generator due to the higher fuel costs.

3.4.1 General Bidding Strategy

The marginal generation cost of electricity using renewable energy is often minimal. Wind and rainfall in generation areas are generally abundant and without any direct costs. Although the 'fuel' for Waikaremoana's power stations is for all intents free, it is a finite resource and limited during some periods of the year. New Zealand has some hydro schemes which are effectively 'run-of-river', however most have dam systems and generating without taking into consideration storage levels and inflows will inevitably lead to empty reservoirs and high spot prices.

When water is scarce, it is important to hold back generation and save water until a satisfactory spot price is met by the market, this is commonly known as the water value. The water value is not an associated cost of water but is instead derived from the potential earnings of the scheme from generating, taking timing of expected inflows into account. Different generation companies may have different internal methods of calculating the water value as they consider their entire portfolio (for example generators which have hydro schemes in both the North and South Islands can expect different inflow patterns and timing). They consider a variety of different factors including available stored water, forecasted inflows, forecasted spot price (which in turn is affected by forecasted demand and supply) and other portfolio considerations such as plant outages.

Holding back generation until the spot price reaches the water value not only ensures that generation profit is optimised but also ensures that the scheme saves water for generation during times of higher demand and peaking. Generators must be careful however, holding back generation during times of high demand to further increase spot prices is unethical, prohibited by the EIPC and may lead to an Undesirable Trading Situation (UTS) and regulatory action.

Hydro schemes including Waikaremoana have storage limits to their reservoirs. When lakes are close to reaching their maximum limits or during times of abundant inflows, traders are expected to bid in at very minimal prices to ensure that the generator is selected to be dispatched by the system operator. This is to ensure that water is used to generate electricity, even for a comparatively low spot price, instead of being wasted as spillage. Unnecessarily spilling water which could otherwise be used for generation is frowned upon and has led to regulatory action by the Electricity Authority in the past.

3.4.2 Location Considerations

Waikaremoana and the Tuai node (TUI) is rather isolated which can be reflected in the nodal price compared to more central nodes like Otahuhu (OTA), Haywards (HAY) and Benmore (BEN). It is connected to the national grid through two 110 kV lines going from Tuai to Napier. It is however in the North Island where the bulk of New Zealand's electricity demand is located which provides it the advantage of not being constrained by the interisland HVDC line.

Waikaremoana is also the closest, largest source of generation for Gisborne and the East Cape. The only transmission line to the region is through the TUI node by means of a 110 kV line. Due to its proximity to Gisborne, transmission losses to the region are lower which reduces overall generation requirements. In the event of islanding (an isolation of a part of the grid) of the East Cape, Waikaremoana is well positioned to provide generation to the region while the connection is restored (Genesis Energy Ltd, 2015).

3.4.3 Other Considerations

The units at Waikaremoana can be bid in together, several times at different price points. If the full generation capacity is not expected to be dispatched, idle or free capacity can also be bid into the reserves market.

Spot traders at a generator-retailer such as Genesis must also consider their contractual obligations when bidding in units to the spot market. Genesis must generate enough to supply their retail business contracts or purchase power from the spot market themselves to make up any deficits in their supply.

In addition to standard Operations and Maintenance (O&M) costs, each unit in the Waikaremoana scheme has an associated start-up maintenance cost that needs to be considered. Starting up a turbine and generator from a dead stop places additional wear on the components which may result in more frequent maintenance events and associated costs. As such, it is important to distinguish between spot price spikes from true trends to ensure the units generate for longer periods of time and are not needlessly cycled.

Due to the connections between the schemes and requirements such as minimum flows, the output through each station must be carefully matched across the scheme.

3.5 Summary

The chapter has covered a considerable amount of background information related to the asset to be modelled and market environment it will interact with.

- WPS is a cascaded hydropower scheme that consists of three stations in series along three lakes. The scheme has a maximum generating capacity of 138 MW using 7 generating units distributed between the stations. The lakes and stations are linked in series where stations downstream are dependent on upstream sources to receive adequate water for generation.
- Waikaremoana is the largest lake and can hold enough water for 47 GWh of generation or 160 GWh of generation if passed through all three stations (cascade). The other two lakes, Kaitawa and Whakamarino, are much smaller and can hold another 85 MWh (130 MWh cascade) and 110 MWh. Due to their much smaller sizes, storage of water in the smaller lakes is quite limited, so flow management is an important factor in their operation.
- The New Zealand power sector uses a liberal electricity market where generators can bid for the right to generate in half hour blocks, by offering their generating units onto a spot market. All successful bids are paid at the same price, the lowest cost bid required to meet generation demand.
- WPS operates in this market and bases its bids on several factors; including the amount of water it currently holds, forecasted inflows and outages, and the spot price size compared to an opportunity cost value; theoretical future value of water. Management of the scheme is a game of bidding into the market at the right times to ensure that water is utilised in a way that minimises spilling and maximise generation revenue.

Chapter 4
Digital Twin Application and Data

4.1 Introduction

The previous chapters have covered a wide range of background information and theoretical relationships required to understand the state of the WPS and market environment it operates in. The thesis has discussed DTs and the types of implementation. In terms of the modelling aspect, information regarding the WPS layout, hydrology and turbine efficiency have been covered.

This chapter examines the first steps when considering the application of a DT for field without vast amounts of previous research papers. Chapter 2 presented a framework for classifying DTs, this section discusses the realistic and necessary levels of fidelity for a DT of WPS in terms of behaviour, visuals, and connectivity. Another topic of interest for the beginnings of DT development is data sources. The importance of data reliability cannot be overstated, all behaviour modelling cannot produce accurate results if compromised by bad data. The sections in this chapter examine the available data, the sampling methods, their method of measurement and their overall reliability.

4.2 Project Application

When exploring the development of a DT for a new or underdeveloped application, it is important to understand the requirements of the problem and the resources available to the DT. Depending on the needs of the application, different levels of detail for likeness, behaviour, and connectivity will be required to reach a suitable effective solution.

The required computing power to render and resolve a DT model naturally increases as the complexity of likeness and behaviour increases. Depending on the ultimate purpose of the DT, likeness and behaviour fidelity are often balanced to reduce model resolution times. An example of this balance are process simulation programs by companies like Aspen (2022) and Siemens (2022). These programs are also becoming known as part of the EDT field. EDTs focus on providing complex behavioural fidelity and balance compute requirements by using low likeness detail in the form of process flow diagrams.

Grieves (2014) mentions the usefulness of 'lightweight' virtual models, recommending that only the required geometry, characteristics, and attributes be selected to avoid carrying around unnecessary details. Yu et al. (2022) in the same vein states that an EDT with a higher level of fidelity does not necessarily equate to a better EDT but instead runs the risk of being over-engineered. Lightweight models dramatically reduce the size of the model and allows for faster processing and thus lower computational requirements.

This principle should also be applied to this project. To identify the project requirements and suitable level of fidelity, it is important to look back at the original objective set by Genesis. The objective is to develop a support tool to assist traders with decision making for the Waikaremoana scheme in the power market. The main objective of this project is higher level, management, scheme performance focused rather than micro, process control and equipment performance focused.

Likeness fidelity requirements are considered low for this application. While 3D representations of the scheme would be a desired feature for future comprehensive models, for this project, only a simple flow diagram is needed to visualise the main components of the scheme.

DT examples discussed above would suggest a low-level likeness model is accompanied by a more detailed behaviour model. However, the problem only requires resolution of a steady state model for every time step. The time step would likely match that of the real-world electricity market trading periods, set for every half hour. The inclusion of multiple steady states pushes the model into the mid class behaviour category. Accurate behaviour in terms of water flow, power generation and lake level projection would be considered the important attribute for the DT.

Although not of immediate importance, automated flow of operational data to the model would be a milestone feature for a complete Waikaremoana DT. The model could simply import data using one of Genesis Energy's data storage platforms. In terms of classification, the project is realistically looking at a 'digital model' due to the low level of likeness and manual extraction of data. In future, a 'digital shadow' with an automatic one-way flow of data would be the bigger goal.

4.3 Sampling Methods

Genesis records and stores data gathered from operations at Waikaremoana in multiple information databases, with most data for this thesis accessed using programs such as Hilltop Manager and PI-Datalink. The databases capture timeseries data, with PI coming from the control systems while Hilltop captures continuous and logged data relating to hydrology and weather. For the purposes of this project, data was retrieved using the PI-Datalink package on Microsoft Excel.

PI-Datalink provides several options for retrieving data including sampling method, start and end date, and interval size. Sampling method and interval size selection can have a dramatic impact on how data is displayed. The two main methods of sampling using PI is instantaneous sampling and calculated averaged values. Instantaneous sampling provides raw data taken from each sensor or correlation. It is inherently 'noisier' than averaged data but is 'truer' to the real-time behaviour of the parameter being measured. Instantaneous data is also able to reveal useful specific details such as data peaks and peak timings.

Averaged data is created using PI's calculated data function and as the name suggests, averages all instantaneous data values over the period of the interval. The interval for these instantaneous values is set at the minimum 1 second. Instantaneous data is tied to a single time value whereas averaged data is tied to a time interval. Due to the way the database is accessed via PI-Datalink, it is convenient to compare instant readings with averaged data using the start timestamp because of the way data is retrieved. When comparing the two using a single sample, averaged data will technically use values ahead in time of the instant data.

Table 4.1: Instant & Averaged Data Simple Dummy Data Example

Start Time	Instant Value		Start Time	End Time	Averaged Value
0	10		**0**	**1**	15
1	20		**1**	**2**	25
2	30		**2**	**3**	35
3	40		**3**	**4**	45
4	50		**4**	**5**	55
5	60		**5**	**6**	65
6	70		**6**	**7**	75
7	80		**7**	**8**	85
8	90		**8**	**9**	95
9	100		**9**	**10**	105
10	110		**10**	**11**	115

Table 4.1 displays a simplified comparison of the sampling methods using a set of linear dummy data, it shows the difference between instant and averaged data over a 10 second window at 1 second intervals. Instantaneous readings are taken at every start time. Average readings have a start and end timestamp, it averages all values (recorded at 1 second intervals) in between these two times.

The difference may not seem very significant with the above example, as it simply shows the small difference between using one reading at each second or two readings. The significance becomes far more pronounced as the time interval between data samples becomes bigger, data values at every second are averaged, so increasing the interval increases the number of samples included in the average. Larger time intervals will compare an instant data sample with averaged data that uses data samples as far ahead as the interval size. Operating conditions are much more likely to change over the course of an hour than a second. This means averaged data is far more likely to average data from two operating states.

To determine the difference between simultaneous and averaged data for efficiency analyses, a small test was conducted by testing efficiency results from instant and averaged data for Unit 6 over a period of several months. For the purposes of this test, the effect of hydraulic head is not viewed. Efficiency was calculated for each data sample method using the hydropower equation and plotted against power generated. With a sample interval of 1 hour, it was clear that using averaged data for calculating efficiency would produce an unacceptable amount of noise for power generation values below 10 MW. Figure 4-1 shows this, instantaneous sampled data (Left) produces a ‘cleaner’ efficiency curve as opposed to averaged data (Right), particularly for lower power outputs.

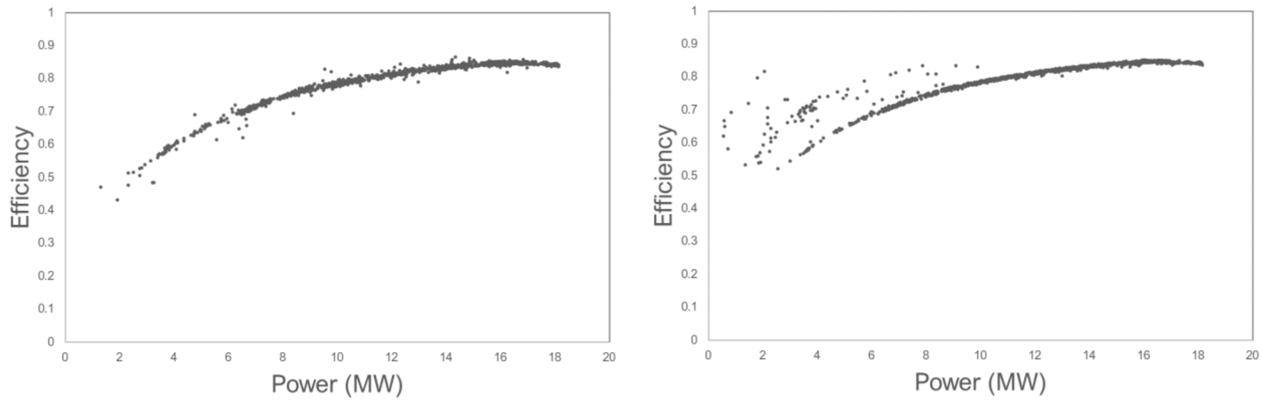

Figure 4-1: Power-Efficiency Plots, Instant Data (Left) & Averaged Data (Right) at 1 Hour Interval

The dispersion seen in Figure 4-1 Averaged Data (Right) is speculated to be caused by the averaging of data from different operating states. The longer time interval of one hour increases the chance that this occurs. An average of a series of low values can be skewed by a single high value.

Table 4.2: Averaged Data Skew Example

Time (Minutes)	Power	Flow Rate	Efficiency
0	17.071	18.124	0.838
20	0	0	-
40	0	0	-
60	0	0	-
Average	4.268	4.531	0.838

Table 4.2 uses simplified dummy data to demonstrate this problem. For an hour interval, thousands of samples will be averaged due to a sample every second. This example just uses 4 samples, but the principle will remain the same. The change in operating state skews the average but the ratio of power and flow rate of the averaged value will remain the same as the high value (Time = 0 in this case). Consequently, their efficiencies will be the same. This is not applicable for efficiency analyses as efficiency is not a constant value, it has a positive gradient relationship with both power and flow rate.

Interval size can also affect the precision of the efficiency curve. Averaging the sampling data for low interval sizes improves the grouping of the data trend while instantaneous data increases data point spread as shown in Figure 4-2. When reducing the sampling interval duration to 1 minute or even 30 seconds, the instant data curve becomes more dispersed due increased number of samples for a given period. The averaged data curve becomes more tightly grouped due to the removal of noise by averaging values.

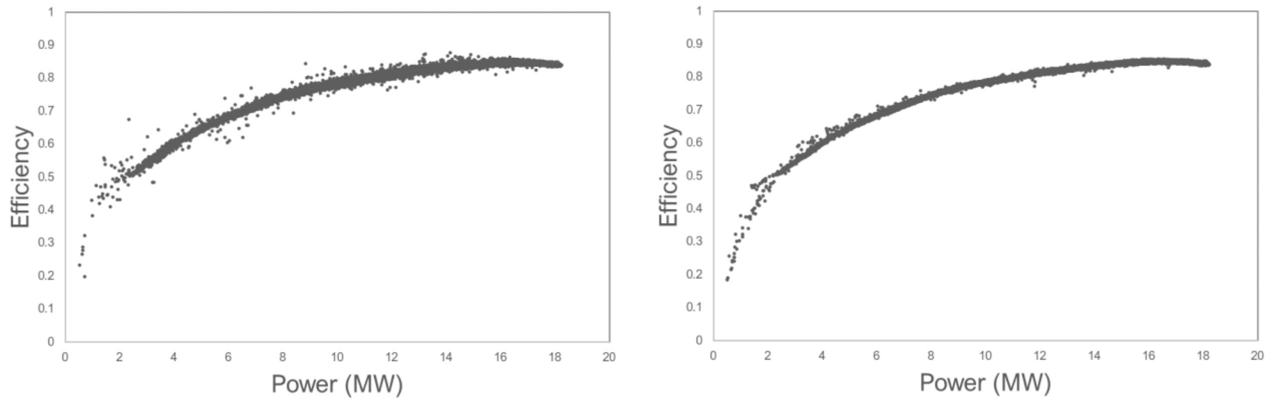

Figure 4-2: Power-Efficiency Plots, Instant Data (Left) & Averaged Data (Right) at 1 Minute Interval

To find the best possible fit for the turbines it is necessary to sample the turbine operational parameters at many different operating points to find corresponding efficiencies at as many different parameter combinations as possible. To avoid overly repeated data, a time interval of 1 hour would provide an acceptable difference in operating conditions between two data samples.

For the case of turbine efficiency regressions and other calculated relationships, an instantaneous sampling method for the parameters is superior, particularly at longer time intervals. However, for parameters that have noise, an averaged value would be the more suitable choice as it can remove those noise related peaks. Parameters that would find this useful include reservoirs levels and approximated flow rates like inflows. Averaged data is often used for plotting flow duration curves, water balance models and water quality modelling.

4.4 Input Data

Input data and/or historical data is crucial for modelling the scheme behaviour. The accuracy and reliability of source data is a central foundation for the usefulness of all analyses and modelling carried out in this project. Input data into the model can be split into two categories, data measured using sensors and data calculated through correlations or other models. The measured values being water levels, select flow rates, power output, and climate data. Measured values provide live feedback data, and ensures the scheme is operating within expected and consented specifications.

Hardware for measuring parameters remote from the power stations typically include a sensor, datalogger, and communication equipment to transmit data to a control circuit and database recorder.

Data loggers record and transmit sensor readings from the measurement site. Data loggers at WPS are commonly Unidata type loggers which record levels at a reduced level scale which is then later converted to actual masl by a PLC at the station. The other newer type of logger is a Campbell logger which measures level in terms of actual masl (Taylor, 2019).

Data is sent from the logger to two locations. To the distributed control system (DCS) using a Modbus Link, transmitted via a Kingfisher remote terminal unit which provides real time data to the scheme operators in the control room. Data is also sent via a GPRS radio link to Hilltop, Genesis Energy's hydrology analysis and compliance platform, which provides data to not only to Genesis' database but also other stakeholders that require the data e.g., for consent compliance etc. (Taylor, 2019).

4.4.1 Levels

Level readings are highly important in hydro scheme management as it directly affects the effective operation of the scheme (levels are used for direct calculations of head level and available water volume), level related historical analyses carried out (flow rate correlation models), and compliance with resource consents (avoidance of penalties associated with exceeding consent range).

A couple different types of sensors/transducers are used to measure water level at Waikaremoana. The most common and preferred method of measuring water level is the shaft encoder (shown in Figure 4-3) which uses a system consisting of a float connected to a counterweight via a pulley (Taylor, 2019).

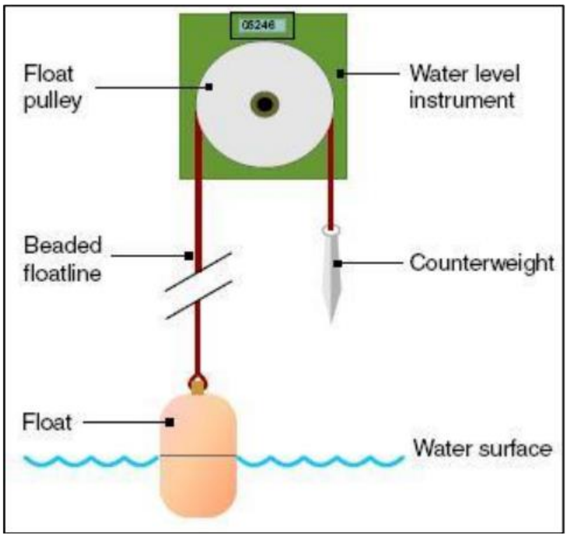

Figure 4-3: Shaft Encoder Type Level Transducer (Taylor, 2019)

The buoyancy of the float allows it to rise and fall with water level, turning the pulley attached to a counterweight. The pulley shaft is part of a water level instrument which measures the rotation of the shaft and converts this into a corresponding level reading.

Due to the size of the lakes and open nature of the reservoirs, waves and surface fluctuations are observed even at small time intervals of 1 second. Factors like dynamic flow conditions in and out of reservoirs and small measurement resolution levels (+/- 1 mm) also play a role in producing and revealing surface fluctuations. The fluctuations are somewhat mitigated using stilling wells and averaging sampled data.

If a stilling well cannot be used, dry pressure type transducers are used instead. The transducer produces gas bubbles down a gas tuber at a constant rate. The pressure transducer senses the back pressure from the bubbles which is proportional to the water level. Changes in back pressure indicate a change in level. This transducer is accurate to +/- 3 mm.

In addition to electronic level measuring instrumentation, each site also has separate alternative measurement gauges as references and in case electronic meters fail. These include manual staff gauge and electric plumb bobs.

Aside from the small issue with fluctuations, which can be reduced using stilling wells and averaging level readings, the electronic level measurements are relatively accurate and acceptable.

4.4.2 Flow Rates

Flow rates at Waikaremoana are not generally measured directly using traditional flow meters like orifice or venturi meters due to flow rate requirements and importance of reducing head loss. Many flow rates in the system are measured by inference and correlation. The flow rates for open channels such as the spillways uses stilling well level data to infer flow rates using a 'verified flow rating', a correlation between level and flow rate produced through testing and data analyses by Hilltop software (Taylor, 2019). In perfect operation, the spillways are less likely to be utilised outside of specific scenarios. As such, these measurements are of lower priority for this project.

For penstocks and turbines, discharge flow rates have traditionally been measured using pressure-flow rate correlations such as the Gibson method or the Winter-Kennedy method. These rely on theoretical relationships between pressure and flow rate, applying Newton's law of inertia and Bernoulli's fluid mechanic principle.

They are low cost and easily implemented methods for measuring flow and can provide acceptable measurements. However, they are an indirect method of measuring flow rate and may not always produce results accurate to the system. At Waikaremoana, pressure sensors on the turbine scroll casing have been used to produce flow rate readings but they are considered inconsistent by Genesis engineers due to possible fouling of the pressure sensors. They have instead produced a calculated flow rate using a correlation between flow rate, power output and head, which was developed from historic efficiency testing.

A newer, direct, and potentially more accurate method of measuring flow rate through penstocks is by using ultrasonic flow meters. More specifically, a transit time type ultrasonic meter which is more suitable clean water applications. A transit time ultrasonic meter, shown in Figure 4-4, has an upstream transducer produce a sound signal into the penstock and fluid, a second transducer slightly downstream detects this signal and sends it back. Fluid flow in the penstock changes the sound signal transit time. Fluid velocity is proportional to the difference between upstream and downstream transit times. Transit time ultrasonic sensors can be very accurate, with some manufacturers claiming an accuracy of +/- 0.1 %, however their effectiveness can largely depend on calibration and installation (Omega Engineering, 2018).

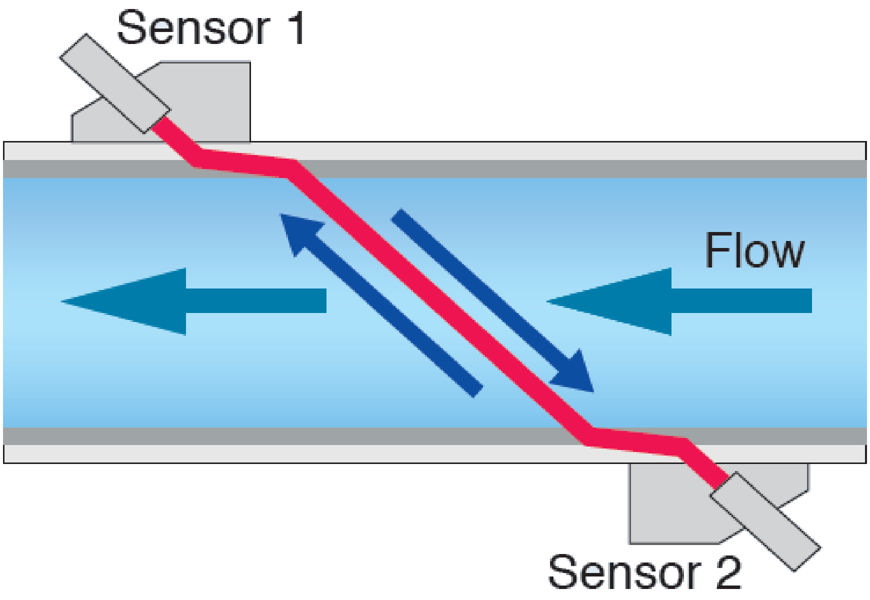

Figure 4-4: Example of a Transit Time Ultrasonic Flow Meter Layout (Fuji Electric, 2022)

Ultrasonic flow meters are considered relatively expensive, (in the realm of several hundred thousand per penstock) and their installation on all penstocks would require a great deal of capital and planning. For this reason, they have only been installed on the Piripaua penstocks for penstock integrity monitoring reasons. Ideally this reading would be used for the DT behaviour flow model but for the sake of consistency, analyses in this project will use the 'calculated' correlation flow rate values for all turbines.

4.4.3 Inflow

Water inflow measurements or calculated values for Waikaremoana is an important factor in creating an accurate water balance model for the power scheme. The water balance model is an essential piece to the overall model as it sets the expected amount of water available for hydro generation and the generation boundaries. The accuracy of this parameter will have a cascading effect on water volume calculations further down the scheme.

As previously discussed in Chapter 3, Lake Waikaremoana has a catchment area of approximately 250 km^2 with over 100 separate streams flowing into the lake. Flow measurements for these streams are generally infeasible due to their numerous numbers, inconsistent flow, and difficulty reaching locations owing to rugged steep terrain. The only consistent time series data available to examine for the streams come from the Hawkes Bay Regional Council; the two largest streams, the Aniwaniwa and the Te Kumi, flow from the north east and discharge into Waikaremoana's north eastern corner. However, these two rivers only make up a portion of the flow into Lake Waikaremoana. The other significant incoming source of water flowing into the scheme is the Kahutangaroa stream, which flows into Lake Whakamarino and has a field metering point.

Currently, for Lake Waikaremoana, Genesis has several indicators for inflows. The major measurement for inflow of Lake Waikaremoana is calculated daily on Hilltop using the difference between level at midnight and the previous or next day's reading while accounting for water extracted for hydropower and leakage. This calculated inflow value does not consider inflows mitigated by 'unmeasured' losses, these can include surface evaporation and leakage not captured by measurements of the Waikaretaheke stream flowing into Lake Kaitawa. However, Hopkirk (2011) in a study on Waikaremoana's inflows, found that these unaccounted losses were not big enough or significantly different from zero. Hence, they found that these were not big factors in producing errors. In this case, using this inflow in a flow balance to calculate net flow should produce a value consistent with net flow found using change in level.

4.4.4 Weather

Climate data for the wider Lake Waikaremoana area is fairly limited. The area around Waikaremoana used to host several National Institute of Water and Atmospheric Research (NIWA) weather stations to collect rainfall and atmospheric condition data. However, these have stopped collecting data since 1990 for most parameters and since 2010 for rainfall data. The Hawkes Bay Regional Council also monitored water quality and general lake health for Waikaremoana using live data via a lake buoy. Unfortunately, they have similarly ceased recording operations at this site.

Genesis still has a weather station located at Onepoto, Waikaremoana. This station provides operators with wind speed, wind direction, air temperature, pressure, humidity, and rainfall data. In the context of behaviour modelling, climate data would be useful for estimating the rate of evaporative losses and rainfall data could factor into inflow estimations.

4.4.5 Power

Power or generator output is measured using different sets of current transformers (CTs) connected to each generator. The basic components of a CT are seen in Figure 4-5. The CTs reduce the output current to a safe level for instruments to measure. Instruments effectively measure output power using either watt-meters, watt-hour meters or ammeters. The secondary coil seen in current produced by the current transformer is proportional to the generator supply current according to the ratio of turns on the transformer. The size of the reduced current is directly proportional to the power output of the generator. Measured power output from each unit can be measured using different current transformers.

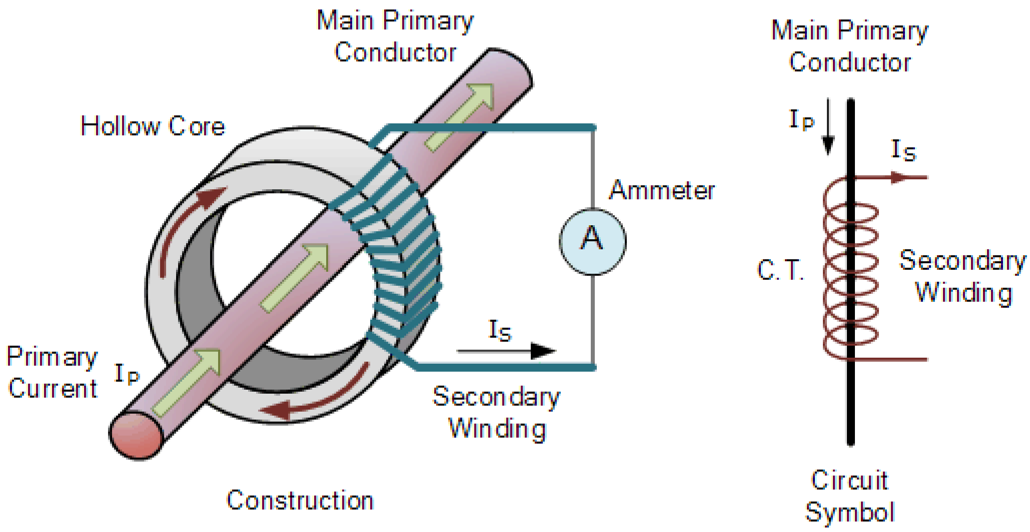

Figure 4-5: Components of a Current Transformer(Electronics Tutorials, 2022)

Each unit has a protection system current transformer which acts as part of a circuit breaker to trip the machine in case of abnormal running conditions. Power output measured using this transformer are lower in accuracy relative to other metering points but still offer readings accurate to within 0.5-1.0 %. Readings here offer the most flexibility due to its wide operating range from its rated current value. The power measured at this current CT or other CTs immediately at the generator output measures the full or gross generation coming from that generator. This data is the primary power output value used by Genesis' monitoring interfaces.

A quirk of this power output measurement occurs when the unit is in an off non-generating state, the meter will still output small fluctuating readings above 0 MW. Readings below 0.5 MW can be assumed to associated with an off, 0 MW state.

Measurements taken from the revenue metering CT produces the other major power output readings. This measurement is the most accurate, with an error of less than 0.3 % at the cost of a narrower range. The meter at this CT measures the power exported to the grid after accounting for all power used for the generating unit's ancillary systems. For hydropower generation units, only a small portion is usually used for ancillary purposes. The revenue metering reading as the name suggests determines the revenue paid to Genesis Energy.

Readings from the protection system CT meters will be the preferred data source for any efficiency related analyses of the generation units as the power function relates raw generator output. Although revenue metering CT meters are more accurate, the difference in accuracy is only marginal and minuscule when looking at absolute errors. In more elaborate models, revenue metering CT values could potentially be used to find the relationship between power output and ancillary load to further increase the accuracy of any revenue related calculations.

4.5 Conclusions

- Based on the long-term objective set by Genesis for a DT of the WPS, the problem is best served with a DT that is more focused on the behaviour side of the DT concept. The overarching goal of this project is higher level and management focused. Visual-likeness and connectivity are not likely to be high priority areas of the project.
- This chapter has covered some important considerations when retrieving data for use in any analyses or models intended for a DT project for WPS. Data availability and reliability is one of the biggest challenges when attempting to build a working behaviour model for a DT. It has a great deal of influence on the accuracy and relevance on this model and much of the behaviour modelling process is centred around identifying reliability or correcting for unreliable values. Two factors that affect reliability with measurement and representation of data are the sampling methods and measuring source.
- Regarding sampling, this section established that averaged sampling helps to reduce noise in the sensor data, particularly for low time intervals. It can have problems at higher time intervals as it skews data between operating states, the skewed data does not accurately represent non-static relationships in the data such as efficiency. Instantaneous sampling is more suitable for longer intervals and represents the true instant measurement of the parameter.
- Genesis Energy has a good range of sensor data available for use with good levels of accuracy. However, this chapter has highlighted some measurements calculated through correlations that could prove to be less reliable. Sensors for power output and water level provide reliable readings while some flow readings were dependent on correlations which may lead to issues depending on how well the correlations match up. The parameters most likely to present possible issues are inflow in Lake Waikaremoana and calculated flow rates particularly through the turbines.

Chapter 5
Turbine Efficiency Characteristic Functions

5.1 Introduction

An overview of the available data and the sampling methods in Chapter 4 has helped establish the considerations that must be made when sampling time series data from Genesis Energy's databases via Hilltop or PI-Datalink. The sections below will now focus on processing and analysing this data for different purposes in the project.

Characteristic hydropower functions that can produce accurate predictions for efficiency are essential for any hydropower DT model as part of the calculation chain. The hydro production parameters have proportional relationships according to the hydro power equation, percentage errors of efficiency will translate directly when calculating flow rate. Errors from this flow rate estimation will then have a carry-over effect on the overall scheme water flow model. Any investigations into the non-linear relationships using regression, whether it is modelling the efficiencies of the turbines or the volume of water stored in reservoirs, only provide estimations, and will has varying degrees of error from true values. Moreover, an important part of the optimisation process involves selecting the operating points with the best efficiency for each unit, making sure each unit utilises its available potential energy to its best extent.

This chapter describes an incremental refining process of regression and explains the reasoning behind each development process. This section begins with a simple linear regression and eventually moves toward the multivariable polynomial model suggested for efficiency functions by Dal Santo and Costa (2016).

5.2 Initial Data Processing

Although Genesis was able to provide many years of turbine operations data, this project will use Genesis' 'calculated' values for turbine flow rates due to reasons mentioned in Chapter 4, Data Source. The calculated values have only been introduced by the operations team in the past 24 months, starting in June 2020. The usable turbine analysis data period for Kaitawa and Piripaua Stations was narrowed to 18 months between July 2020 and February 2022. Tuai data was narrowed down a further three months to July 2020 to November 21 due to a surge tank level sensor failure. For each turbine the following sensor data is extracted from PI Datalink: forebay level, tailwater level, turbine flow rate, and power output. The data uses instantaneous sampling, at a 30 min time interval for reasons explained in Chapter 4.

The extracted data was cleaned by removing any invalid data entries including record errors with obvious faulty readings. Power output measurements below 0.5 MW have been considered as readings from units in an OFF state. These entries were filtered and removed from the dataset as they did not represent data from an ON operating state. It should be noted that due to the tunnel configuration of the units at Piripaua, unit characteristics may be different when both units are running compared to each unit operating on its own. As such, the results for those units have also been compared against a dataset that filters out entries where both units are running to check for significant differences.

The next step was to calculate the theoretical efficiency of the unit using the turbine power equation. As covered in the literature review, the overall efficiency of each generating unit can be calculated using the relationship.

$$\eta^* = \frac{P}{qHK} \tag{5-1}$$

Where K, the hydro unit production function, is calculated using $K = \frac{\rho g}{10^6}$. Density, ρ, was taken as 999.6 kg/m^3 and considered effectively constant.

Although this equation usually uses net head by subtracting losses due to conduit friction, friction losses were difficult to calculate as only friction factors for some conduits were available. Subsequently, friction losses are not explicitly removed from head and instead considered as part of the overall efficiency of the hydropower unit, efficiency is marked with * to note this difference.

From here, the analysis data set for each turbine consists of several thousand running cases/observations, these cases are randomised and split into analysis and validation groups in a 70:30 split. Randomisation ensures each group is representative of the complete data set. The analysis group is used to find the functional relationship between the dependent variable (effective efficiency) and the independent variables (power output, flow rate, and gross head). The validation group is used to test the fit and accuracy of the proposed model against an independent set of data.

5.3 Linear Regression Fit

For the purposes of this Chapter, the regression results of Unit 6 will be presented. Regression analysis was carried out using Python3 with the Sci-kit Learn library or using Excel's Analysis TookPak.

As a first step in the regression curve fitting process, each of the independent variables is plotted against efficiency and fitted with a simple linear regression.

The regression models are applied to our validation datasets and visually examined for any obvious relationships. Figure 5-1 shows that the power output and turbine discharge plots have strong similarities with each other; a strong 2nd order polynomial relationship with efficiency, and a relatively close grouping around the parabolic curve. Gross head displayed a very random spread without a clear relationship, only a negative gradient at the greatest efficiency values. The linear fits show plenty of room for improvement.

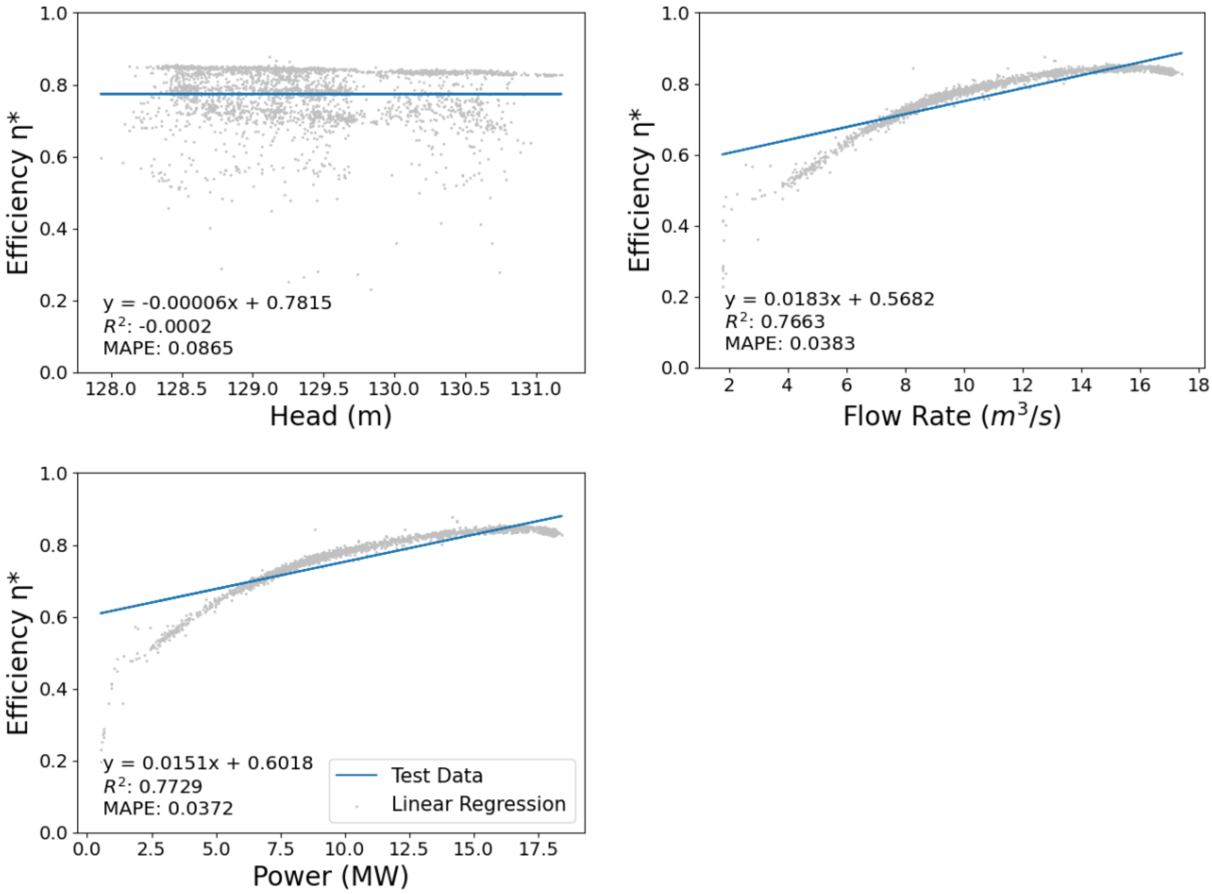

Figure 5-1: Unit 6 Linear Regression Plots Against Validation Data

Power and flow rate linear regressions coefficient of determination, R^2, ranged between 0.2 and 0.8, some plots showed that a linear regression for those variables had a good proportion of its error explained by a linear model. The linear model for head had an $R^2 \approx 0$ which suggested a linear model was inadequate to explain efficiency variation. The mean absolute percentage error (MAPE) is another useful metric to measure the predictive capabilities of a model. On average predictions made using the linear model were within 10 % of the actual validation data. It should be noted that MAPE can be skewed when the number of observations is high and concentrated in certain areas, which explains relatively low MAPE values despite considerable differences in the low flow rate/power regions of the plots.

From these results it confirmed that power, turbine discharge and head, while not having strong clear linear correlations, do have signs of varying relationships with efficiency. The next logical regressions to examine would involve models that follow the nonlinear quadratic pattern of the scatter plots and/or regressions with multiple independent variables. In the interest of time, for finding the most suitable model, only a single unit from each station is tested; Units 6, 1, and 4. Units at each station share head and power limits and should have turbines of similar specifications.

5.4 Multivariable Linear Regression

Literature has suggested the next option to obtain a better fitting regression would be a multivariable linear regression. This model attempts to establish a relationship between the dependent variable, efficiency and more than one independent variable: power, turbine discharge and head. The predictor variables for this type of regression must be independent to develop a dependable mode. Multi-collinearity becomes a problem when independent variables are related, small changes in data causes large changes in the model. Linear efficiency plots using power output or flow rate results in very similar looking plots. Logically power output and flowrate are physically related, as flow rate increases, more torque is applied to the turbine and shaft to produce a greater power output. Power and head should also have a direct relationship according to the power equation; however, this is not evident in the linear plot in Figure 5-2.

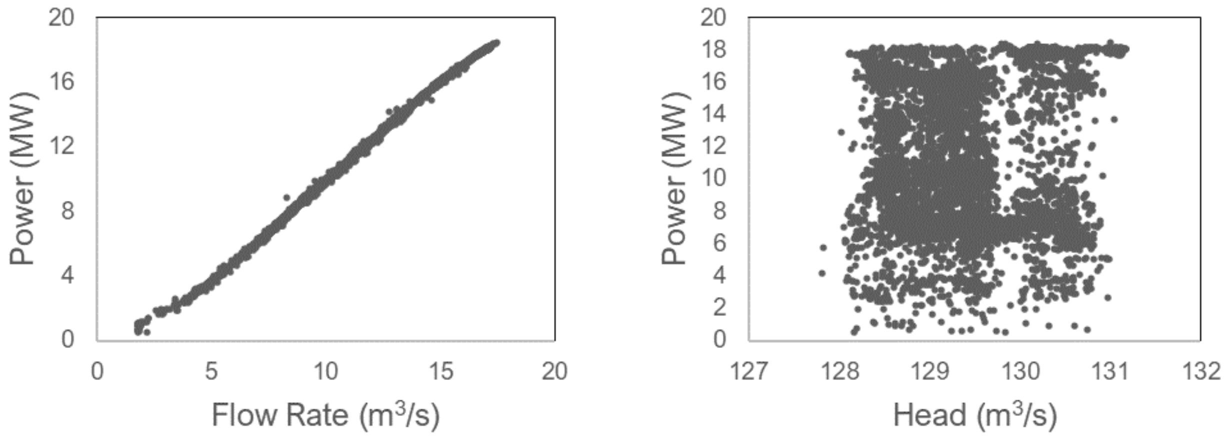

Figure 5-2: Predictor Variable Relationship, Unit 6 Flow Rate-Power (Left), Head-Power (Right)

Figure 5-2 demonstrates the largely linear relationship between the predictor variables power output, with small possible nonlinear regions at lower and higher outputs/flow rates. Therefore, it is sensible to conclude that power and turbine discharge cannot be used together for this regression. Traditionally in literature, discharge flow rate is favoured over power output in multivariable regression equations for turbine efficiency. However, the flow data Genesis records/data used for this analysis is only a calculated value (as opposed to a real sensor), the power output is a real measurement and therefore more reliable to use.

Multivariable linear regressions are a combination of their linear regressions. The two predictor variables selected for this regression will be power output and head. As only two variables have been selected as predictors, the model can be represented in a 3-Dimensional Plot, a combination of their respective 2D linear plots.

To check the validity of the model and the significance of each predictor variable, factors such as respective turbine unit's overall F-test significance and predictor p-values are examined. Overall F-test tests the null hypothesis, H_0: a model with only an intercept value, with independent variable coefficients = 0 fits the data as well as the tested model (Frost, 2019). The model is tested at a level of significance, $\alpha = 0.05$. This means if an

F-test results in a significance value below 0.05 the model is statistically significant enough to reject the null hypothesis, thus indicating that at least one of the predictors is significant (Bowerman et al., 2005).

Table 5.1: Multivariable Linear Regression Metrics Summary

Unit	R^2	Adjusted R^2	F-test	MAPE
6	0.807	0.807	0	3.268
1	0.838	0.838	0	2.255
4	0.650	0.650	0	5.492

As shown in Table 5.1 above, each of the units produced multi-variable regressions that were statistically significant over an intercept only model. P-value are tests against the H_0: the variable has no correlation with the dependent variable (Frost, 2019). Rejecting this null hypothesis indicates that there is a non-zero correlation for the variable. P-values for both the predictor variables were acceptable for all units.

The multivariable model yielded marginally improved regression metrics for most units. Using Unit 6 as an example, R^2 only improved 2% over the linear power and flow rate regressions. Usually, this marginal improvement alone would not determine this model to be an improvement. Note R^2 values should be taken with some scepticism when more predictors are introduced into the model, R^2 never decreases when additional predictor variables are introduced, even when the new variable is unrelated to the dependent variable. To ensure the model is not overfitted, the adjusted R^2 is examined. Adjusted R^2 corrects for the number of predictors and observations. An increase in adjusted R^2 may indicate that the addition of the head predictor variable improved the model more than would be expected by chance. However, at such a small percentage improvement, it is hard to say if there was a true improvement over a linear power regression. The MAPE value improvement of this model was also marginal, <1 % for most units. These results were also echoed across Units 1 and 4.

The multivariable linear regression model only shows weak signs of improvement over a singular linear model however improved versions of this model should still be explored. The linear plots initially observed indicated a quadratic relationship between power and efficiency.

5.5 Multivariable Linear (Polynomial Terms) Regression

The next progression in the regression model is a multivariable linear regression with polynomial terms for each predictor variable introduced. Each of the predictor variables has a 2nd degree polynomial term and lower order terms to model the parabolic behaviour observed. This form follows the suggested turbine efficiency power model suggested by Dal Santo (2016) and Finardi et al. (2005).

Table 5.2: Multivariable Polynomial Regression Metrics Summary

Unit	R^2	Adjusted R^2	F-test	MAPE
6	0.984	0.984	0	0.817
1	0.987	0.987	0	0.634
4	0.869	0.869	0	2.673

F-test significance was again ≈ 0, indicating that the model provides strong evidence that this model is statistically significant compared with an intercept only model. The p-values also support this, although Units 1 and 4 have p-values greater than the specified level of significance. This could be due to multi-collinearity between the predictor variables.

The change to a polynomial model by introducing polynomial terms via transforming the base variable data showed a considerable increase in regression evaluation metrics. Each of the tested units showed at least a 0.15 increase in adjusted R^2. Most notable are Units 6 and 1, reaching R^2 values around 0.98. MAPE values also saw a significant decrease; Unit 6 decreasing from 3.3 % for the multivariable linear model down to 0.8 % for this model.

5.6 Reducing Multi-Collinearity

The multi-collinearity (MC) problem was noted earlier in this chapter, it occurs when predictor variables in the model are correlated with one another. MC does not necessarily affect the predictive capabilities of the model, but it can make interpreting the variable coefficients and p-values of the model difficult. Coefficients become sensitive and may change wildly depending on the independent variables and regression dataset. The p-values of the affected variables become untrustworthy, and the true effect of each coefficient cannot be interpreted (Frost, 2019).

There are two types of MC: data and structural. Data MC occurs when independent variables have a real physical relationship that results in a correlation between the two sets of data, e.g., power output and flow rate outlined previously. Structural MC is present when a predictor variable data set is created using other predictor variables and caused by the model choice, e.g., x^2 variable data created from x variable data (Frost, 2019).

Data MC was previously addressed when creating the multivariable regression model, but structural MC has just been introduced when polynomial terms were added into the model. Polynomial terms are a transformed dataset from the base independent variable and will inherently depend on that variable.

The presence of MC can be verified by checking the variance inflation factor (VIF) values of each variable. These check for the presence of correlation between variables and the strength of these correlations. The VIF is defined as follows in equation 5-2:

$$VIF = \frac{1}{(1 - R^2)}$$

(5-2)

Where R^2 is from a linear regression of a tested variable using the other predictor variables as the independent variables.

VIF values start at 1 and do not have an upper limit. A perfect VIF is 1, where there is no correlation between the tested variable and the other independent variables. A VIF below 5 indicates that there is some correlation but within an acceptable range. VIF values > 5 indicates critical levels of multicollinearity where coefficients are unreliable (Frost, 2019). Table 5.3 outlines the VIF values for our latest regression model for Unit 6. The direct relationship between the base variables and their polynomial terms have resulted in highly large VIF values. Table 5.3 below shows this with the model's high VIF levels for each variable.

Table 5.3: Variable VIF Values Before Centring

Predictor Variable	**Variance Inflation Factor (VIF)**
Head, H	137737.6
Head Squared, H^2	138089.5
Power Output, P	30391.2
Power Output Squared, P^2	37.5
Mixed Poly Term, H·P	30873.7

Structural MC can be fixed by pre-centring the data, simply subtracting the mean from each observation. Interpretation of coefficients remains the same but when making predictions using the model, the input data must also be centred. By centring the base variables head and power output, the VIFs can be significantly reduced.

Table 5.4: Variable VIF values After Centring

Predictor Variable (Centred)	**Variance Inflation Factor (VIF)**
Head, H	1.15
Head Squared, H^2	1.21
Power Output, P	1.06
Power Output Squared, P^2	1.12
Mixed Poly Term, H·P	1.13

Table 5.4 shows that the VIFs have all decreased to low levels. The coefficients and p-values of the model have also shifted slightly. However, centring the data has not changed the predictive capabilities of the model and

the MAPE has remained the same. Other regression metrics of the model including R^2, adjusted-R^2 remains the same. Addressing MC is said to have made the coefficients and the p-values more trustworthy.

All unit F-test significance values were minimal and displayed as ≈ 0. This indicates there is very strong evidence to suggest that this regression model provides a better fit than an intercept-only model. The p-values were generally improved over the non-centred model, for most variables they were miniscule and below the specified significance level, indicating there was evidence that the model variables have a correlation with efficiency. There were some cases where p-values indicated there was insufficient evidence to conclude a correlation existed, these cases were for the variable, head squared: Units 1, 3, and 4. However, given that sufficient evidence was found for other units and literature supported this mode, the squared head variable was kept in the final regression model.

The equation of the final regression model is the form shown below in Equation 5-3, using Unit 6 as an example.

$$\eta^* = 0.8112 - 0.00408(H - \overline{h}) + 0.000571(H - \overline{h})^2 + 0.01551(P - \overline{P}) - 0.00182(P - \overline{P})^2 - 0.00013(H - \overline{h})(P - \overline{P}) \tag{5-3}$$

Table 5.5: Summary of Regression Coefficients, Centring Means, MAPE and Adj. R^2 Values

	Regression Coefficients						Centring Means		Error	
Unit	**Inter**	**H**	**H^2**	**P**	**P^2**	**H·P**	**H Mean**	**P Mean**	**MAPE %**	**Adjusted R2**
6	0.81120	-0.00408	0.00057	0.01551	-0.00182	-0.00013	129.44	11.43	0.817	0.984
7	0.80172	-0.00467	0.00041	0.01410	-0.00168	-0.00033	129.36	11.68	0.735	0.989
1	0.76753	-0.00221	-0.00009	0.01550	-0.00138	-0.00024	204.88	11.22	0.634	0.987
2	0.77611	-0.00088	-0.00120	0.01215	-0.00149	-0.00057	204.86	12.04	0.628	0.982
3	0.86063	-0.00360	-0.00144	0.00787	-0.00164	0.00149	204.41	15.06	2.351	0.847
4	0.82108	0.04057	-0.00173	0.01463	-0.00129	0.00106	113.82	12.25	2.673	0.869
5	0.79406	0.01937	0.00611	0.02019	-0.00164	0.00103	113.86	10.19	2.640	0.874

5.7 Contour Plots

The coefficients from Table 5.5 produced the characteristic efficiency equations for each unit. These equations were applied across the full range of maximum and minimum operating head and power to generate hill diagrams. The required flow rate at each point was then calculated by rearranging the power characteristic equation.

$$q = \frac{P}{\eta^* HK}$$

(5-4)

On a quick sanity check, the hill diagrams generally generated points with values that were logically sensible.

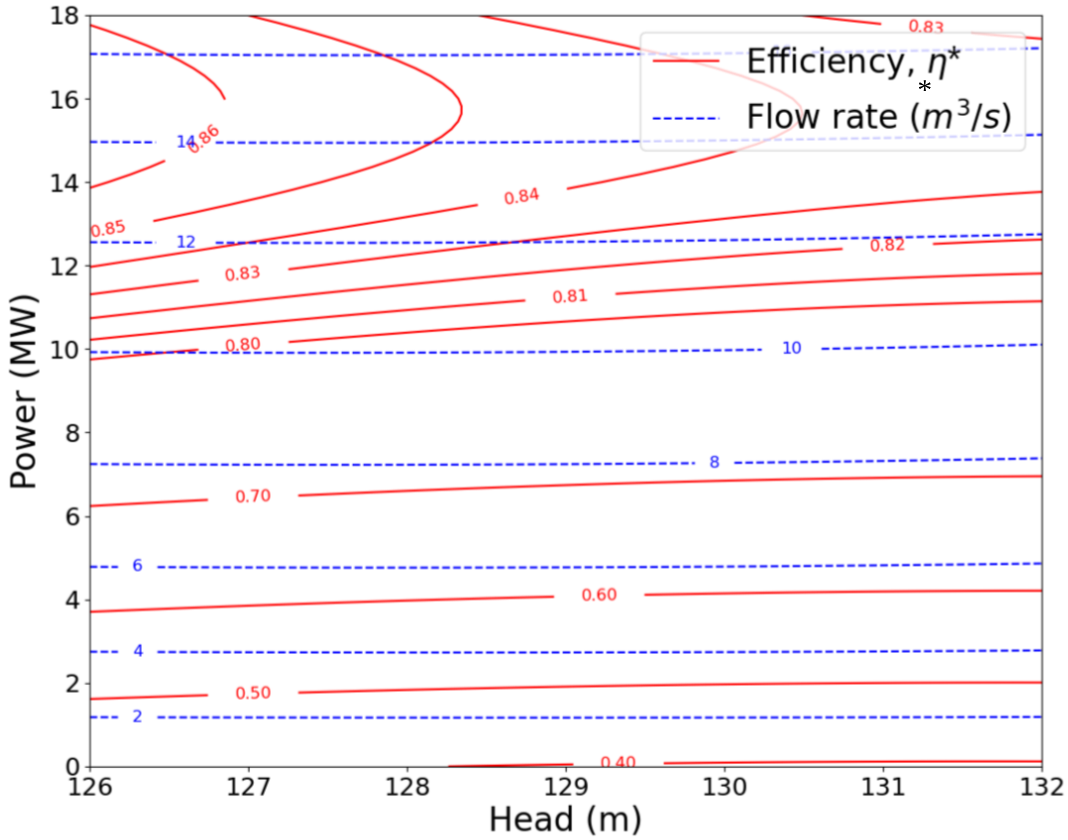

Figure 5-3: Hill diagram for Unit 6

Each turbine has an operating region where maximum efficiency is achieved, Figure 5-3 shows Unit 6's optimum operating region is in an area of lower head and higher power output. This represents the general trend for most of the turbines analysed. Hill diagrams for all units can be found in the appendices. The hill diagrams only emulated a portion of the hill diagram example given by literature as the operating range for Waikaremoana's turbines are relatively narrow compared to turbines specified in literature.

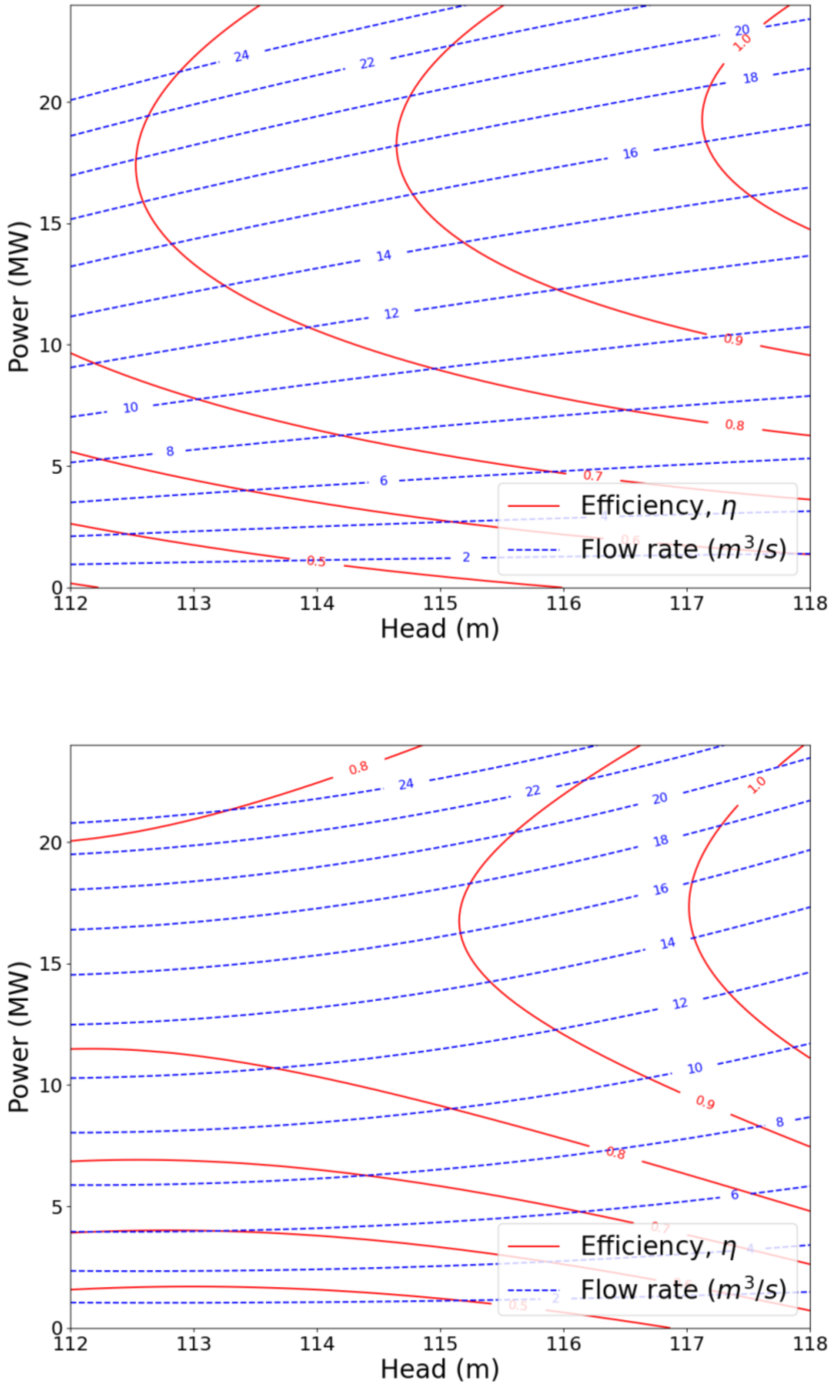

Figure 5-4: Hills diagrams for Unit 4 (Top) and Unit 5(Bottom)

Figure 5-4 shows the hill diagrams of Units 4 and 5 which differ from the general trend of the other units. The units at Piripaua according to these diagrams are more efficient at an area of higher head and higher power. Also, of note is the region of the graph with efficiencies larger than 1.0, logically this is invalid and will be a consequence of the input data where calculated flow rates are used. Furthermore, the equations used to generate these diagrams are only approximations using operational data regressions. The adjusted-R^2 values for Piripaua were among the weakest out of all units owing to the irregular scatter of the operating points. It could also simply be the case where the specific turbines are rated for use at head levels higher than the operational head at Piripaua while the units at Kaitawa Station and Tuai are rated for lower head levels then their current operating heads.

5.8 Variation From the Model

Although still having moderate regression metrics Units 3, 4, and 5 show a noticeably lower regression metrics, R^2 and MAPE, compared against the other units. Linear scatter plots for these units showed that they had irregular spikes or a wide scatter around the expected parabolic curve. The spikes may be caused by different running conditions that are not captured by a model that only considers head and power output.

Unit 3 at Tuai was considered a last on, first off unit during the period the regression data was gathered. Therefore, Unit 3 also had far fewer observations than the other units, often by a factor of 3. It was due for an overhaul and was only run when water constraints or opportune prices in the market necessitated it run.

The dispersion seen in the scatter plots for Units 4 and 5 could be attributed to a few factors. The regression plots were quite dispersed, with some samples reaching an impossible efficiency over 1. This clearly indicated there was some error in the analysed data sources. The other notable factor is the station intake situation; the units at Piripaua use a single long tunnel (2.6 km) for the intake. There are therefore higher relative friction losses from the tunnel when both units are running versus a single unit running. The data used for this regression was not separated between single and dual unit running. To check the effect of this factor, regressions of both running conditions will be compared with one another.

5.9 Single and Simultaneous Running Cases

Each station has either two or three units that can run individually or simultaneously with other units at the station. At Kaitawa and Tuai Stations, the units have short intake tunnels and individual penstocks that reduces the effect that units have on one another when running together. Piripaua is known to have some inter unit interaction when both units are running due its shared intake tunnel. This was previously suspected to have cause the wide regression dispersion and lower regression metrics.

To test the effect this interaction has on the regression results and validate this impact, regression data for Units 4 and 5 were further filtered into individual running data and simultaneous running data. For the Unit 4-only case, only data entries with Unit 4 power >0.5 and Unit 5 power <0.5 was selected. Vice versa for Unit 5. For the parallel running case, both Units 4 and 5 power output needed to be >0.5.

Table 5.6: Regression Metrics Comparing Mixed Case with Single & Simultaneous Running

Unit	R^2	Adjusted R^2	F-Test	MAPE
Unit 4 Any	0.869	0.869	0	2.67
Unit 4 Single	0.912	0.912	0	2.25
Unit 4 Simultaneous	0.792	0.791	0	2.91
Unit 5 Any	0.874	0.874	0	2.64
Unit 5 Single	0.879	0.879	0	1.55
Unit 5 Simultaneous	0.713	0.713	0	2.06

Above is Table 5.6 which compares, regression results using the initial filter criteria (any), and cases with single or simultaneous running turbines. It shows that further filtering data into single and simultaneous running cases only improved the regression results marginally. For the single running turbine case, Unit 4's adjusted R^2 only improved by around 4% while Unit 5 did not see much improvement at <1%. For MAPE values, Unit 4 did not see much improvement while Unit 5 saw overall ≈ 1% reduction.

Applying the selected regression on a simultaneous running case saw noticeable decreases in adjusted R^2 for both units. MAPE results diverged, Unit 4 case had an increased MAPE over the 'any' case while Unit 5 saw the opposite.

Improved regression metrics were expected if single and simultaneous running cases had large effects on the goodness of fit. Regression results from this test were mixed and improvements were marginal at best. These mixed results could neither confirm whether single and simultaneous running data are significant factors in producing better regression fits. Further investigation is required to validate the impact of a long single intake tunnel. Avenues to consider, choosing an alternative source of flow rate data such as the ultrasonic sensors and accounting for friction when calculating efficiency.

5.10 Summary of the models

Below in Table 5.7 is a summary of the statistics results for Unit 6.

Table 5.7: Unit 6 Different Model Summaries

Regression Model	R^2	Adjusted R^2	F-test	MAPE
Linear	0.773	0.773	0	3.830
Multi-variable Linear	0.807	0.807	0	3.268
Multivariable Polynomial	0.984	0.984	0	0.817
Multivariable Polynomial Centred	0.984	0.984	0	0.817

5.11 Future Improvements

In future, net head could be used for this regression model by finding the correct friction loss values for each conduit. The friction coefficients can be estimated or found by conducting pressure loss tests. However, correction of head may reveal a direct correlation between net head and power/flow rate due to the relationship between frictional head loss and flow rate. The centred model may not be able to reduce the data multi-collinearity in this case. As discussed, predictions made using the model would still produce comparable predictions, but the coefficients become volatile, and p-values are unable to be interpreted. To address this issue, a different regression method could be applied, like ridge regression which can deal with data multicollinearity.

5.12 Conclusions

The refining process for regressions of the turbine operational data yielded the following conclusions.

- Linear regressions produced plots with decent R^2 regression values but visually the plots showed that a linear plot was unsuitable particularly for lower or higher power outputs.
- Multivariable linear regression results showed only marginal improvement over regular linear regression. The improved R^2 values were also viewed with scepticism as R^2 always increases when more predictor variables are added. Adjusted R^2 showed very marginal changes between this and the single predictor linear regression.
- Multivariable linear regression with created polynomial terms saw a dramatic improvement in regression metrics. Each of the tested units showed improved adjusted R^2 by around 15%, with some units like Unit 6 and 1 reaching an adjusted R^2 of around 0.98. The introduction of polynomial terms did introduce structural MC into the data. This was reduced by centring, subtracting the mean values, this did not affect the regression statistics but decreased the MC metric, VIF dramatically.
- Units 3, 4 and 5 were the units that did not show excellent fits to the proposed model. Although their regression metrics were still considered ‘good’, they showed some odd behaviour in the results. Hill diagrams generated using Unit 4 and 5’s characteristic functions had areas of efficiency greater than 1. This oddity may be down to the quality of data for the two units at Piripaua. Unit 3 was due for a generator overhaul and as a result had substantially fewer data samples than the other units.
- Units 4 and 5 are known to affect each other’s performance when both are operating simultaneously. Both units showed scatter plots with the most amounts of dispersion around the proposed function. It was expected that further splitting the regression data into single and simultaneous running cases would see an improvement in both units’ regression value. Unit 4 saw a small increase, but the opposite occurred for Unit 5. This could be entirely due to the quality of data.

Chapter 6
Power Scheme Flow Balance

6.1 Introduction

The refinement process for discovering the best function to use to represent the efficiencies of the hydropower units was an important piece of work. The characteristic functions for most units should be able to predict efficiencies to within a few percent. These functions form part of the flow model chain which is the focus of this chapter. This chapter will report on the process to produce a working flow model that describes all the major hydraulic relationships in the WPS. It will present the visual representation of this model before testing the model using operational data.

In this section, a first principles-based approach is taken by first deriving the mass and energy balance forms for a generalized hydropower lake. These balances are applied to the lakes in the WPS to generate the relevant hydraulic equations for each lake. The accuracy of the flow model is tested by applying the flow model equations to a set of operational data. Net flow rates found using the flow model is reconciled against actual level changes measured by sensors.

6.2 First Principles Balance Equations

6.2.1 Mass Balance

To determine whether all water flows are accounted for, a mass/water balance can be applied across each reservoir. The hydropower system is for our intents and purposes, a pure water open system. The transfer of mass through a reservoir can be expressed as:

$$Accumulation = \ Mass\ In - Mass\ Out \tag{6-1}$$

$$m_{Acc} = m_{In} - \ m_{Out} \tag{6-2}$$

Accumulation will translate to the change in volume of the reservoir per period measured. ‘Mass In’ will include all sources of inflows into the reservoir, both natural inflows, discharge and spillway flows from power stations. ‘Mass Out’ will be the sum of all water drawn from the reservoir for generation and via spillway releases. Evaporation can also be factored into Mass Out.

Expressing in terms of volume, then flow rate:

$$\rho V_{Acc} = \rho V_{In} - \ \rho V_{Out} \tag{6-3}$$

$$V_{Acc} = V_{In} - V_{Out} \tag{6-4}$$

$$\frac{dV_{Acc}}{dt} = q_{In} - q_{Out} \tag{6-5}$$

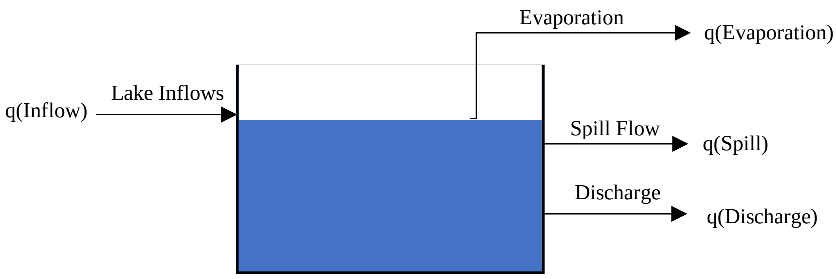

Figure 6-1: Mass Balance for a Standard Hydropower Reservoir

Calculating the volume in the current time instance using the above balance and previous time point's available water volume, V_{t-1}:

$$V_t = V_{t-1} + \Delta T(q_{In} - q_{Out}) \tag{6-6}$$

$$V_t = V_{t-1} + \Delta T(q_{Inflow} - q_{Discharge} - q_{Spill} - q_{Evap}) \tag{6-7}$$

Where ΔT is the time interval chosen and V_t is the available reservoir volume. V_t, can be calculated preferably using a level-volume correlation, otherwise using a simple multiplicative relationship between level and estimated lake surface area. Available volume only considers the water volume above the minimum operating level; hence the minimum level is subtracted for the current level.

$$V_t = A \times (L - \underline{L})$$

6.2.2 Energy Balance

An energy balance is important to measure how well the turbine harnesses the potential energy held by the water reservoirs. This directly relates to the turbine efficiency calculations. Energy can be lost due to many factors including friction along tunnel and penstock walls, turbine runner inefficiencies and generator losses.

As discussed in the literature review chapter, the potential energy of a water reservoir is calculated using the equation below. Only the useable volume should be considered, that is the volume above the minimum operating limit. Energy potential calculations will use the average elevation of the available volume, a simple average of the level and the level lower limit.

$$E_{Potential} = mgh_{Geo\ Cent} \tag{6-8}$$

$$E_{Potential} = V\rho gh_{Geo\ Cent} \tag{6-9}$$

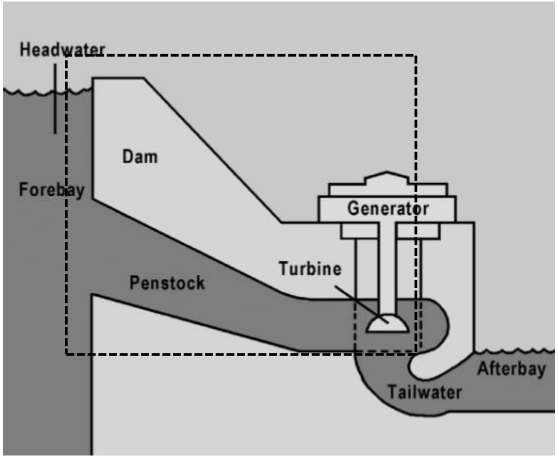

Figure 6-2: Hydro Dam with Energy Balance Boundary (U. S. Department of the Interior: Bureau of Reclamation, 2005)

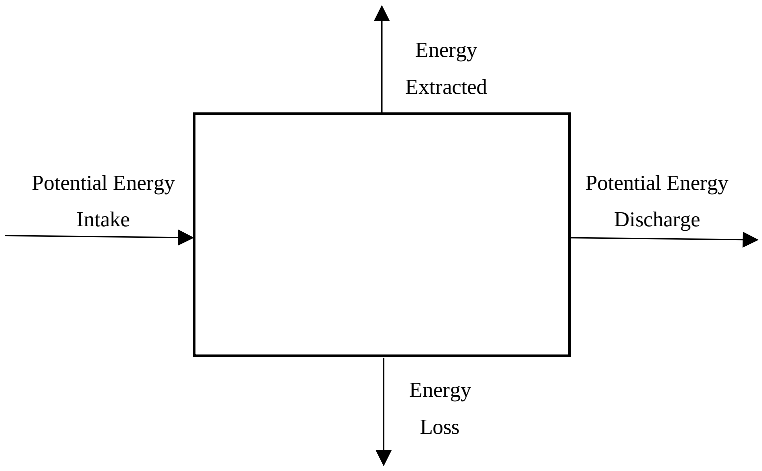

Figure 6-3: Power Station Water Energy Balance

$$E_{Intake} = E_{Extracted} + E_{Discharge} \tag{6-10}$$

$$V\rho g h_{Intake} = E_{Extracted} + V\rho g h_{Discharge} \tag{6-11}$$

$$V\rho g(h_{Intake} - h_{Discharge}) = E_{Extracted} + E_{Loss} \tag{6-12}$$

Converting the balance from volume (V) and energy (E) to flow rate (q) and power (P).

$$P_{Total} = P_{Generation} + P_{Loss} \tag{6-13}$$

$$P_{Generation} = \eta_{Overall} qKH \tag{6-14}$$

These balances can be applied across each of the WPS' lakes.

6.2.3 Applying to the Scheme

6.2.3.a Lake 1 - Lake Waikaremoana, Kaitawa Station

The balance can be applied across Lake Waikaremoana (L1), another outflow term for leakage flow is also introduced into the flow balance.

$$\Delta \dot{V}_{L1} = q_{L1\,In} - q_{L1\,Out} \tag{6-15}$$

$$\Delta \dot{V}_{L1} = q_{L1\,In} - (q_{L1\,L} + q_{L1\,Sp} + q_{L1\,Dis} + q_{L1\,Evp}) \tag{6-16}$$

$$V_{L1,t} = V_{L1,t-1} + \Delta T(q_{L1\,In} - q_{L1\,Dis} - q_{L1\,Sp} - q_{L1\,L} - q_{L1\,Evp}) \tag{6-17}$$

Volume available for Waikaremoana is calculated using Level-Volume correlation found from prior lake level testing. The plot showed a strong linear correlation with an almost perfectly linear fit and strong R^2 value of 0.99. It should be noted that volume is in the unit million m^3, volume measurements on the scale of thousands of m^3 will be very sensitive to small changes in level. Volume changes calculated for this smaller scale should viewed with some caution.

Level measurements, precise and accurate enough for volume calculations on the thousands scale may not be possible due to lake level fluctuation caused by waves.

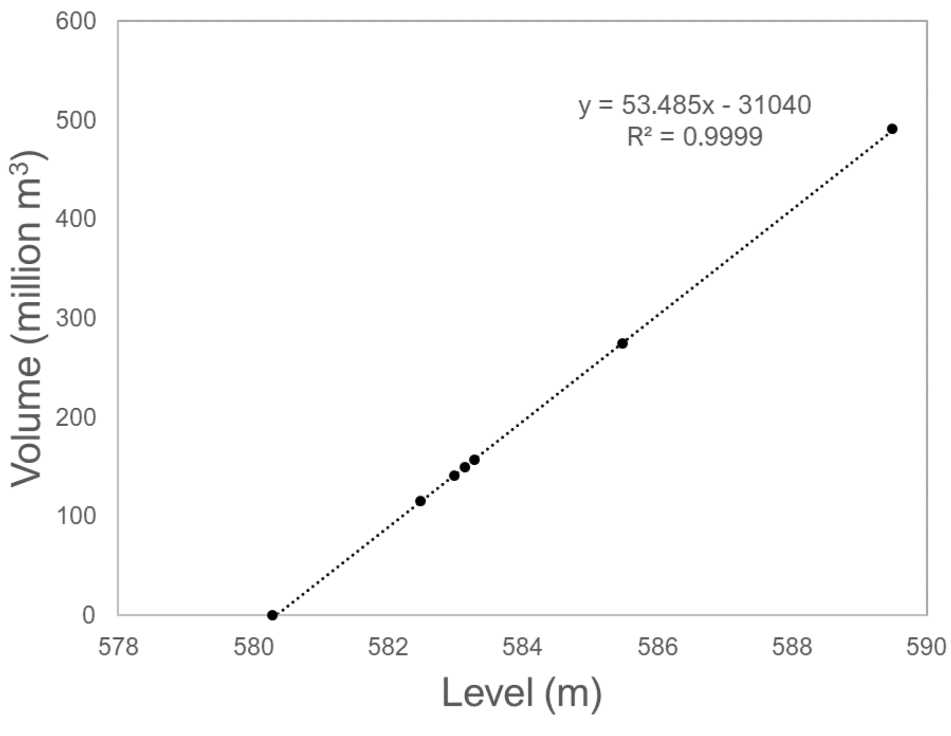

Figure 6-4: Waikaremoana Level-Volume Correlation

The approximate theoretical energy potential of available water at Waikaremoana using available volume and current head level.

$$E_{L1} = (V_{L1} - \underline{V_{L1}})\rho g H_{L1} \tag{6-18}$$

Balances for Kaitawa Station, including splitting of discharge flow into the individual penstocks. This equation assumes that discharge flow through the tunnel is evenly split into the two penstocks when both units operate.

$$q_{L1\,Dis} = q_{L1\,G6} + q_{L1\,G7} \tag{6-19}$$

Equation to find the gross head for the units at Kaitawa Station.

$$H_1 = h_{L1} - h_{L1\,Tail} \tag{6-20}$$

Hydropower efficiency was found for the units with the following the hydropower characteristic function suggested by several researchers such as Dal Santo and Costa (2016).

$$\eta^*{}_{G6} = \Upsilon_0 + \Upsilon_1 H + \Upsilon_2 P + \Upsilon_3 HP + \Upsilon_4 H^2 + \Upsilon_5 P^2 \tag{6-21}$$

The full form found from multi-variable regression analysis, with the coefficients and averages to correct for the centring of the parameters. The following equations for Unit 6 are the same in format for Unit 7 but uses the corresponding coefficients and average values.

$$\begin{aligned}\eta^*{}_{G6} = {} & 0.8112 - 0.00408(H_{L1} - 129.44) + 0.000571(H_{L1} - 129.44)^2 + 0.01551(P_{G6} - 11.43) \\ & - 0.00182(P_{G6} - 11.43)^2 - 0.00013(H_{L1} - 129.44)(P_{G6} - 11.43)\end{aligned} \tag{6-22}$$

As power is used in this equation, the hydropower equation will be used for calculating flow rate.

$$P_{G6} = K q_{L1\,G6} H_{L1} \times \eta^*{}_{G6} \tag{6-23}$$

$$q_{L1\,G6} = \frac{P_{G6}}{\eta^*{}_{G6} \cdot H_{L1} \cdot K} \tag{6-24}$$

6.2.3.b Lake 2 - Lake Kaitawa, Tuai Station

The water balance when applied across Lake Kaitawa (L2) is shown below. This balance has several more terms due to the multiple routes of travel for water from Waikaremoana and the option of spilling into the Waikare Tahake River via the diversion, denoted by the subscript SpD. Although the canal into Whakamarino is technically the 'diversion', for the sake of consistency across the lakes, spilling to the next lake remains denoted with Sp.

$$\Delta \dot{V}_{L2} = q_{L2\,In} - q_{L2\,Out} \tag{6-25}$$

$$\Delta \dot{V}_{L2} = (q_{L1\,Dis} + q_{L1\,Sp} + q_{L1\,L}) - (q_{L2\,Dis} + q_{L2\,Sp} + q_{L2\,SpD} + q_{L2\,Evp}) \tag{6-26}$$

$$V_{L2,t} = V_{L2,t-1} + \Delta T(q_{L1\,Dis} + q_{L1\,Sp} + q_{L1\,L} - q_{L2\,Dis} - q_{L2\,Sp} - q_{L2\,SpD} - q_{L2\,Evp})$$
(6-27)

The volume calculation of Lake Kaitawa does not have a specific level-volume correlation. Due to its relatively small size and more uniform shores, it simply a constant surface area of the lake and water level above minimum consent to calculate volume.

$$V_{L2} = A_2 \times (h_{L2} - \underline{h_{L2}})$$
(6-28)

$$V_{L2} = 6.1 \times 10^4 \times (h_{L2} - 450.1)$$
(6-29)

The other equations for stage two of the WPS is almost identical in format to lake one. The only differences being that the station, Tuai has a large surge chamber on its intake and the station has three units and penstocks. The approximate theoretical stored potential energy in Lake Kaitawa uses the same equation form as Lake Waikaremoana.

$$E_{L2} = (V_{L2} - \underline{V_{L2}})\rho g H_2$$
(6-30)

Although there should only be a small difference, the surge chamber level sensor data is set as the $h_{forebay}$ rather than the lake level. The hydraulic pressure from the lake tries to maintain a steady level in the chamber.

$$H_2 = h_{L2\,Srg} - h_{L3}$$
(6-31)

The flow through to the surge chamber is assumed to be consistent with flow out of the chamber to maintain the level. Discharge flow is split evenly into up to three penstocks.

$$q_{L2\,Dis} = q_{L2\,G1} + q_{L2\,G2} + q_{L2\,G3}$$
(6-32)

Again, each unit has its own hydropower characteristic function for efficiency calculation. The example function for Unit 1 is below.

$$\eta^*_{G1} = 0.76753 - 0.00221(H_{L2} - 204.88) - 0.00009(H_{L2} - 204.88)^2 + 0.01550(P_{G1} - 11.22) - 0.00138(P_{G1} - 11.22)^2 - 0.00024(H_{L2} - 204.88)(P_{G1} - 11.22)$$
(6-33)

Turbine flow rate is calculated through the same method as stage one.

$$q_{L2\,G1} = \frac{P_{G1}}{\eta^*_{G1} \cdot H_{L2} \cdot K}$$
(6-34)

6.2.3.c Lake 3 - Lake Whakamarino, Piripaua Station

The water balance for Lake Whakamarino (L3) follows the same format. Flow into the lake stems from Lake Kaitawa via the discharge and spillway canal.

$$\Delta\dot{V}_{L3} = q_{L3\ In} - q_{L3\ Out} \tag{6-35}$$

$$\Delta\dot{V}_{L3} = (q_{L2\ Dis} + q_{L2\ Sp} + q_{L2\ Ka}) - (q_{L3\ Dis} + q_{L3\ Sp} + q_{L3\ Evp}) \tag{6-36}$$

$$V_{L3,t} = V_{L3,t-1} + \Delta T(q_{L2\ Dis} + q_{L2\ Sp} + q_{L2\ Ka} - q_{L3\ Dis} - q_{L3\ Sp} - q_{L3\ Evp}) \tag{6-37}$$

Volume calculation for Lake Whakamarino uses the same form as for Lake Kaitawa. It is many times larger than Lake Kaitawa but still much smaller than Waikaremoana with better defined shorelines.

$$V_{L3} = A_3 \times (h_{L3} - \underline{h_{L3}}) \tag{6-38}$$

$$V_{L3} = 2.98 \times 10^5 \times (h_{L3} - 246.2) \tag{6-39}$$

Equations concerning the operation of Piripaua Station follow the same form as Kaitawa Station. Approximate theoretical stored potential energy in Whakamarino.

$$E_{L2} = (V_{L2} - \underline{V_{L2}})\rho g H_2 \tag{6-40}$$

Piripaua similarly to Tuai, has a surge chamber but only has two penstocks for two units.

$$H_3 = h_{L3\ Srg} - h_{L3\ Tail} \tag{6-41}$$

Flow intake into the chamber is similarly assumed to match flow out of the chamber to maintain level due to hydraulic pressure from the lake. Discharge flow is split evenly into up to two penstocks.

$$q_{L3\ Dis} = q_{L2\ G4} + q_{L2\ G5} \tag{6-42}$$

Example hydropower characteristic function for Unit 4, for efficiency calculation.

$$\begin{aligned}\eta^*{}_{G4} = {} & 0.82108 + 0.04057(H_{L3} - 113.82) - 0.00173(H_{L3} - 113.82)^2 + 0.01463(P_{G4} - 12.25) \\ & - 0.00129(P_{G4} - 12.25)^2 + 0.00106(H_{L3} - 113.82)(P_{G4} - 12.25)\end{aligned} \tag{6-43}$$

Turbine flow rate is calculated through the same method as stage one.

$$q_{L2\ G1} = \frac{P_{G1}}{\eta^*{}_{G1} \cdot H_{L2} \cdot K} \tag{6-44}$$

6.2.3.d Other

The last balance for system is for the flow out of the WPS via the natural river outlet, the Waikaretaheke River. Originally, overflows and leakage from Waikaremoana would use this path to flow toward the sea. After hydro development, the river remains used as spillways for each stage and the main outlet for water leaving the power scheme. The Upper Waikaretaheke Stream flows into Lake Kaitawa while the Lower Waikaretaheke acts as the scheme outlet. The Lower Waikaretaheke can receive spill flows from Lake Kaitawa and Lake Whakamarino in addition to discharge flows from Piripaua Station.

$$q_{WTR} = q_{L3\ Dis} + q_{L2\ SpD} + q_{L3\ Sp} \tag{6-45}$$

6.3 Assumptions and Simplifications

The hydro generator units are operated primarily using the wicket gates and main inlet valves. During shutdown, wicket gates are closed followed by the main inlet valve. The order is reversed for start-up. The headgate at the top of the water column remains open to avoid cavitation of the penstock in the event the penstock is unintentionally emptied.

A common concern among run-of-river hydroelectric plants is head variation due to discharge flow and subsequent loss of head. This is more of an issue for high flow and low head, river extracting systems. WPS has reservoirs with more storage than most of these systems. Levels at Whakamarino and particularly Waikaremoana are unlikely to be significantly affected by discharge flow rates due to their size. Lake Kaitawa is considerably smaller than the other two lakes and may have some flow dependent level variation. The level that could be most affected by flow variation is Piripaua's discharge outlet on the Waikaretaheke River which only has a tailrace pond. A scatter plot of this outlet's operating data, plotting tailwater level against Piripaua's discharge flow rate is shown below in Figure 6-5.

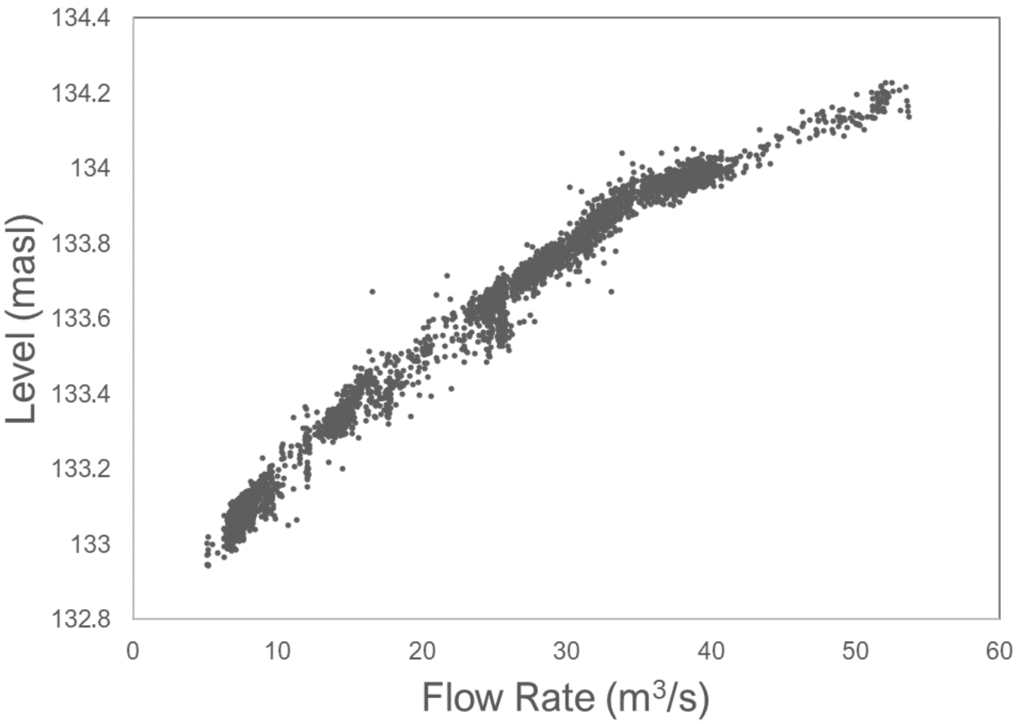

Figure 6-5: Piripaua Tailwater Head Variation, Flow Dependence

Figure 6-5 shows a clear linear or quadratic relationship between the two parameters. However, it should be noted that the level variation across the entire range of flow rates is only between 1-1.5 m. This is rather insignificant when head levels for all stations are in the range of 100 to 210 m. Variations of 1 to 2 m would only amount to affecting 1 or 2 % of total head. It would be a fair assumption to consider head variation to be negligible and a less important factor at this stage. Defining and factoring in head variation could be investigated later in future work.

Another type of loss often discussed is the friction loss when water travels through conduits such as penstocks and tunnels. Net head consists of forebay level less tailwater level and penstock head loss. These losses can be difficult to calculate as they require the characteristic friction coefficient of the conduit which would require discharge and pressure tests on the water column. Penstock losses are therefore incorporated into the turbine efficiency value as part of an overall unit efficiency, η^*.

The penstocks are normally kept filled with water unless maintenance on the penstocks is required. This means the water travel time from the forebay to the station discharge reservoir is negligible as water is shifted down the intake. Water spillways are open channels and not always utilised so will have some travel time. These could be found using flow tests in conjunction with the stilling well measuring points. However, in ideal operation, water should not be spilled to extract as much energy as possible. Spilling water only occurs to bypass a non-operational station, to avoid exceeding the upper limits of the reservoir consent or meet the minimum flow requirements for spillways.

The Waikaretaheke Diversion is used to bypass Lake Whakamarino when Whakamarino is in danger of exceeding its consented operating water level or when both Tuai and Piripaua cannot be operated. This option introduces additional complexity to the system because it creates asymmetry to the cascaded scheme. To obtain an easier replicable scheme in optimisation, the Diversion is simplified into two spillways, one from Lake Kaitawa into Whakamarino and the other into the Waikaretaheke River.

Calculating surface evaporation for lakes is a complex task. There have been a great number of formula put forth to model and predict water evaporation from open water sources. Each considers many different parameters, from irradiance to vapour pressure and wind speed. Other elaborate methods for calculating evaporation also require data not readily available such as net radiation. The simplest method for evaporation estimation would be using a blanket average value.

Finklestein (1973) measured evaporation and estimated the yearly total to be around 595 mm, one could simply assume an even distribution across the year for around 1.63 mm evaporation per day. This very rough estimate equates to approximately 1 m^3/s for Waikaremonana, 0.001 m^3/s at Lake Kaitawa and 0.006 m^3/s for Whakamarino. This is quite negligible in the grand scheme of things, so in this early-stage evaporation can be assumed to be negligible in the balance. This is supported by Hopkirk (2011), who carried out research regarding estimating leakage flow and evaporative losses not accounted for by available measured data. They concluded that these losses calculated in the study were not statistically significant enough from zero.

6.4 Visual Representation

Although of secondary importance to investigating the flow balance, a visual representation of system similar to a PFD would be a helpful addition. An Excel model of the WPS system was built using these balances and equations with the help of the PFD shown in the scheme background. The model can be used to calculate expected lake levels given a set of sensor data (inflows, levels) and decision variables (power output). The model diagram and flow are represented simply by Excel generated shapes. Data corresponding to each segment is annotated with all corresponding data, including constants, sensor, calculated, and decision variables. A random running state was taken from WPS's historical data and used as an example for a running state in Figure 6-6.

6.4.1 Excel Flow Diagram

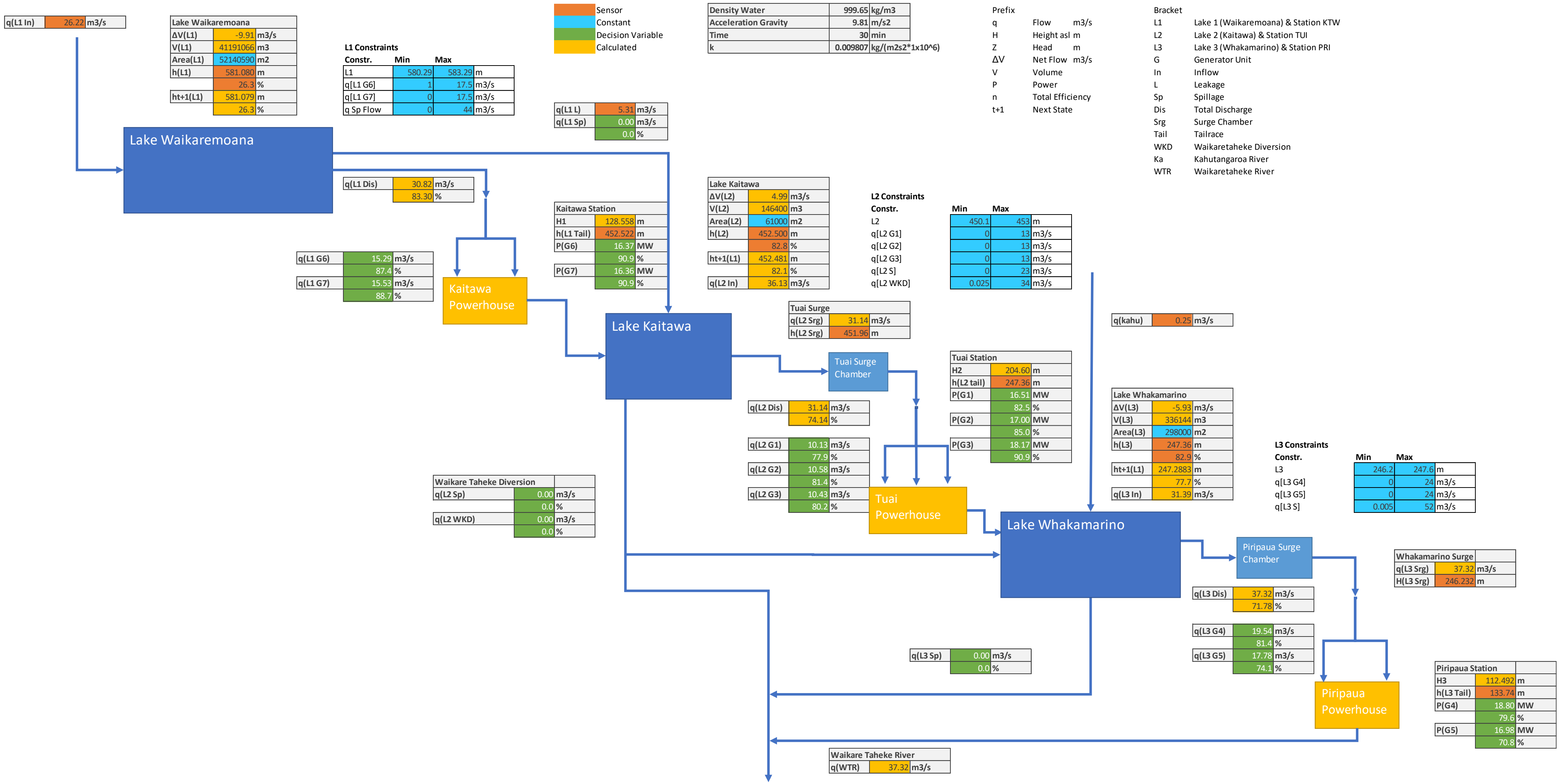

Figure 6-6: Overview of the Excel Flow Diagram

6.5 Net Flow Reconciliation

This section aims to examine the accuracy of the balance and flow model shown earlier. To test this, 6 months of daily data from Waikaremoana was used to compare net flow rates calculated through two methods. The first is calculated through the flow balance and the other is by examining daily level difference. Each data sample is taken at midnight, level data is a simultaneous sample while all other data was averaged over the day. An interval of a day was used to capture a wider range of operating states, an interval of minutes could limit the data to just an on or off state.

The process to get to the comparison data for a single hydropower stage was as follows in Figure 6-7:

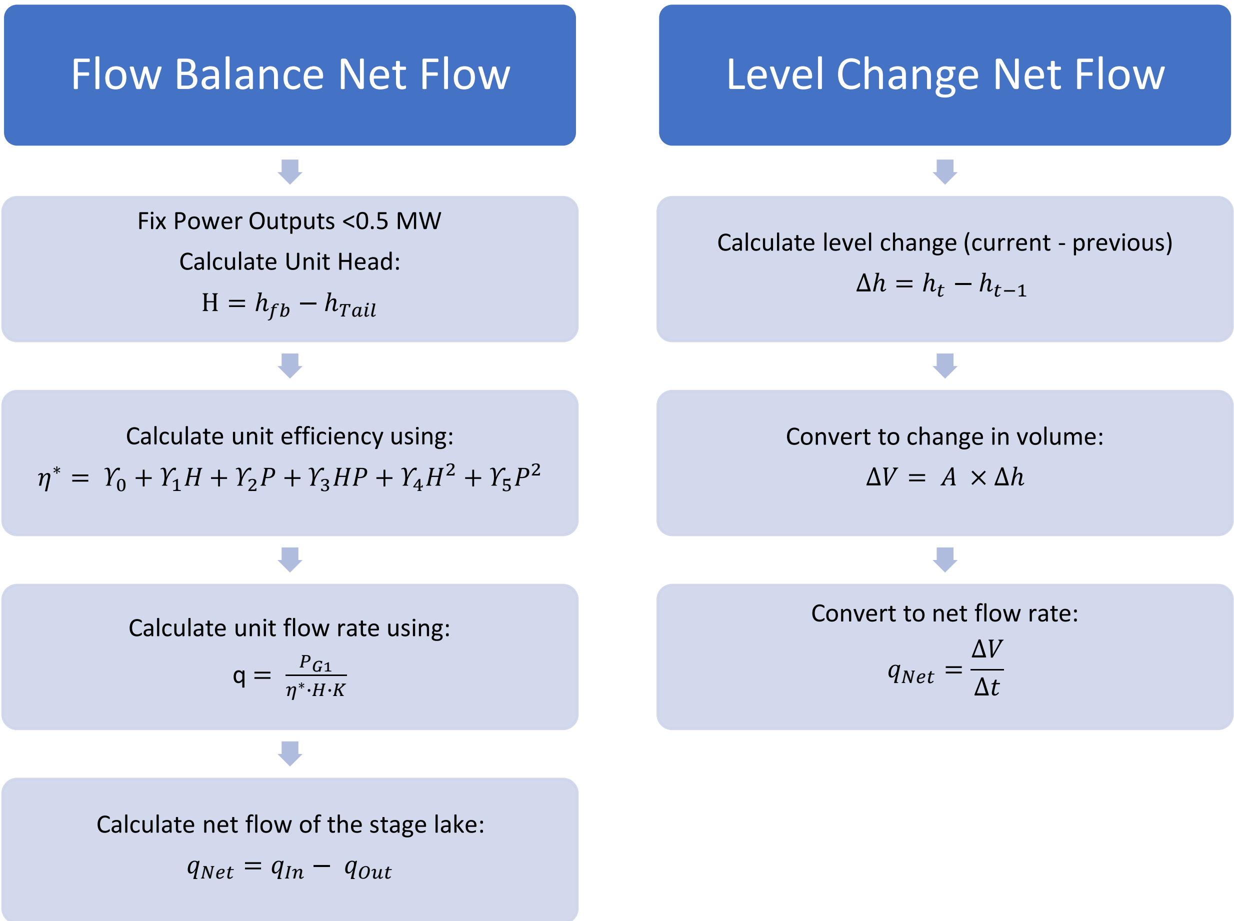

Figure 6-7: Calculation Process for Flow Balance and Level Reconciliation

6.5.1 Results

This comparison should show how closely the calculated value from a flow balance matches the actual measurable change in volume. With a perfect model, the two values would be expected to agree, however the results show that flow balance-net flow rate tends to dominate the difference between the two. For each lake, the difference in flow rate of the two methods was plotted against the flow balance net flow rate to show this domination.

The plots for each the lakes show a clear linear relationship between the reconciliation difference and the net flow calculated via flow balance. This suggests that flow difference could be largely driven by some factor used in the flow balance.

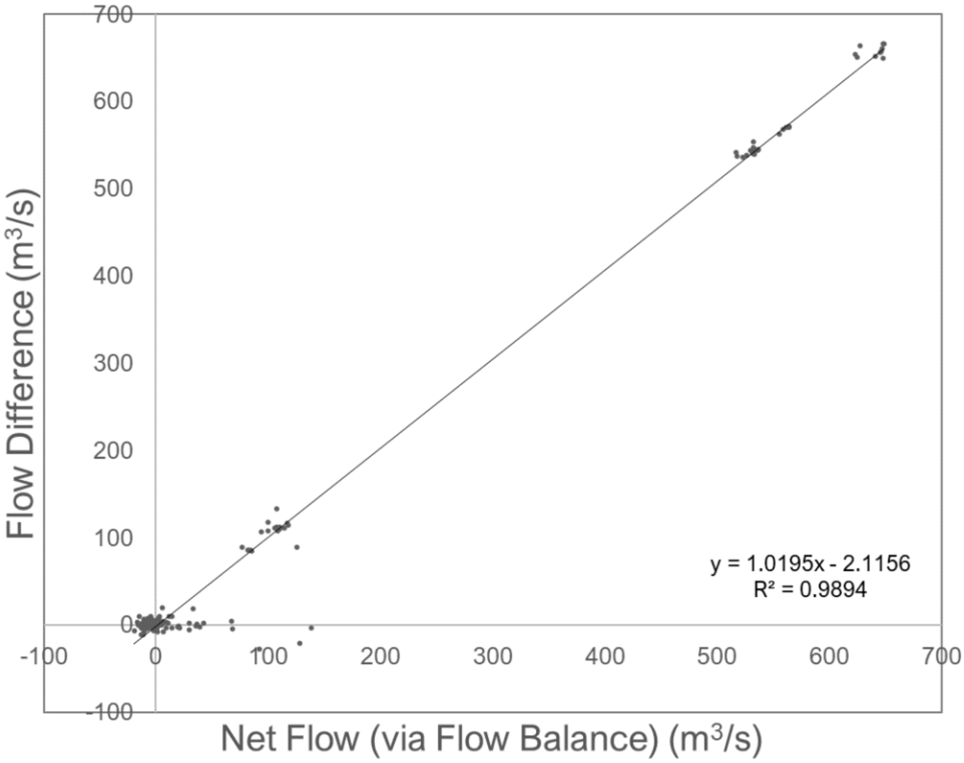

Figure 6-8: Lake Waikaremoana Balance Flow Difference Against Flow Balance Net Flow

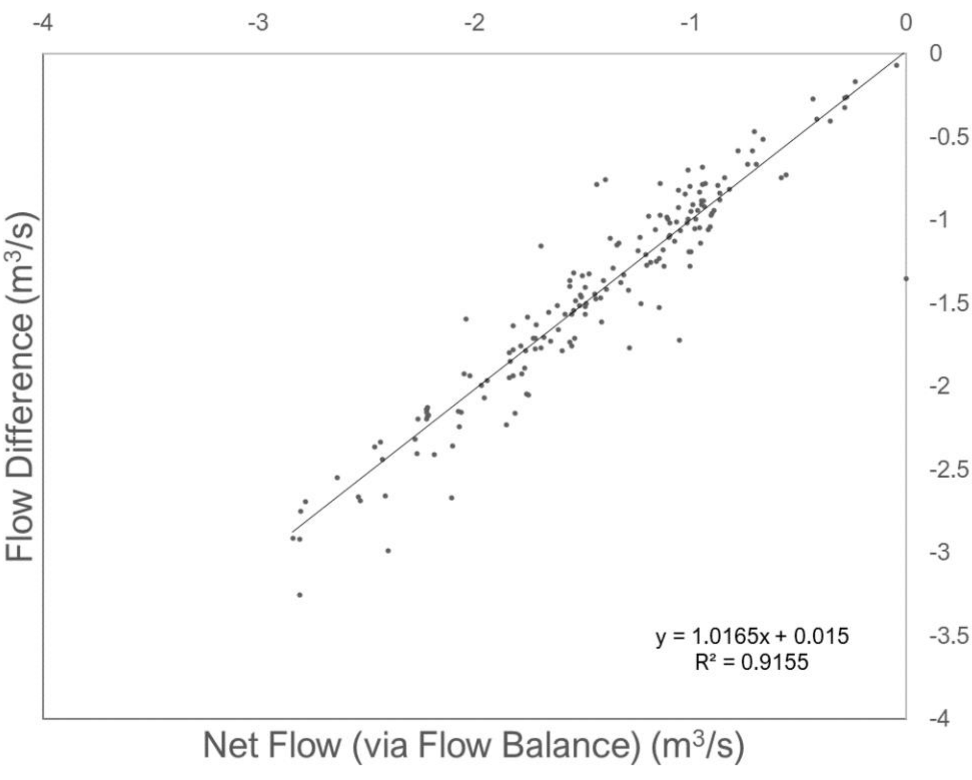

Figure 6-9: Lake Kaitawa Balance Flow Difference Against Flow Balance Net Flow

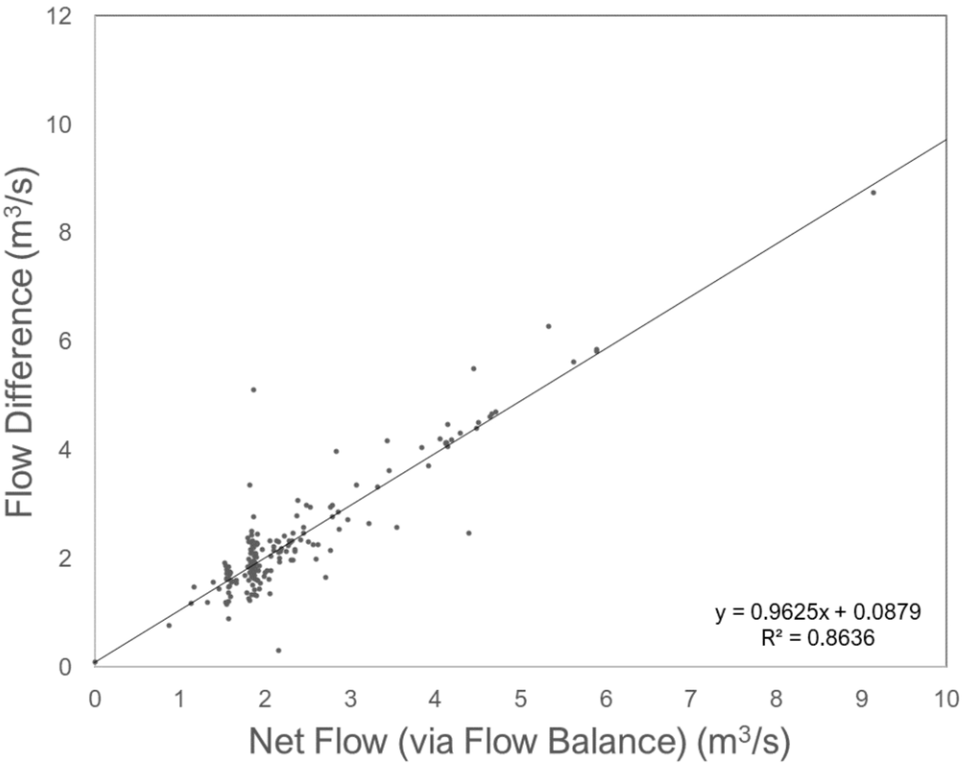

Figure 6-10: Lake Whakamarino Balance Flow Difference Against Flow Balance Net Flow

The biggest component for the flow balance at Waikaremoana is its inflow value. Figure 6-8 shows that for the Waikaremoana balance, most of the discrepancies between the two net flow rate values are scattered around the origin point. When it comes to larger flow balance net flows, caused by high daily inflows, the level change for the day does reflect these events and the difference matches the magnitude of these high inflows. Genesis Energy's inflow value for Waikaremoana is said to be derived from examining daily level changes but this data does not seem to support this.

Figure 6-9 shows that for Lake Kaitawa, there was a pattern of overestimating negative flow throughout the entire six months by the flow balance. This is relatively small difference of just over a cubic metre of water as shown in Table 6.1. The difference could be due to an overestimating of discharge flows or more likely an underestimation of the leakage flow from Waikaremoana; only leakage travelling through the spillway is measured.

Lake Whakamarino's results in Figure 6-10 shows a more scattered plot compared to the other two lakes. The additional natural inflow from the Kahutangaroa stream may play a part in this increased randomness around the linear trendline.

Table 6.1: Average Metrics for Flow Comparison

Lake	Mean Difference (m^3/s)	Median Difference (m^3/s)
Waikaremoana	109.65	2.09
Kaitawa	-1.36	-1.34
Whakamarino	2.37	2.01

The average metrics for this comparison show that in general, the absolute error between the two readings is in the region of only a few m^3/s. A median offers a better representation of the error for Waikaremoana as there are several very high outliers that skew the mean to a much higher level than a large portion of samples in the data.

Table 6.2: Lake Data and Net Flow Summation Table

Lake	Start Level (m)	End Level (m^3)	Sum Net Flow Level Based (m^3/s)	Sum Net Flow Flow Bal. Based (m^3/s)
Waikaremoana	581.60	581.59	-0.41	19732.37
Kaitawa	452.59	452.27	-0.03	-245.69
Whakamarino	247.23	247.28	0.29	431.49

An interesting observation to note, Table 6.2 shows that the level at the start of the six months of data for each of the lakes is nearly even with the level at the end of the data. To achieve this, the sum of the net flows should logically remain close to zero. This is the case with net flow calculated with level change but false for net flow via flow balance for all three lakes. This suggests there is an overestimation of a factor in each balance, or the use of averaged values has possibly skewed flow rates to higher than actual magnitudes.

6.6 Conclusion

- This chapter started from the basics and tried to apply a first principles approach to the DT behaviour aspect. Mass and energy balances was derived and applied to the scheme to develop a continuous flowsheet for the WPS.
- Several assumptions were outlined to help simplify the problem and remove factors in the balances and equations that provided only a small amount of influence for the amount of further work required.
- The balances and equations were adapted into an excel flow model form. Provided a visual representation of the scheme with annotations for relevant readings and calculated values.
- Finally, the net flow produced by the flow balance was compared against the net flow found by physical comparison of daily water levels. Discrepancies were generally smaller in magnitude with a strong relationship with flow balance net flow. A range of possible sources for these errors were discussed which included overestimation of inflows and underestimation of leakage flow.

This has provided a good starting point for the flow behaviour aspect of a DT.

Chapter 7
Early Optimisation Development

7.1 Introduction

The thesis has covered the behaviour replication aspect of the DT and has some visual representations of the system with process flow type diagrams. The previous chapters have worked on relationship functions needed for the flow model and presented the first attempt at a flow model.

The eventual goal of a DT for Genesis was a tool that could provide decision making support for wholesale electricity traders and give asset market performance evaluations. DTs not only try to emulate real systems, but they commonly provide data processing capabilities. This ‘data processing’ attempts to deliver useful information for the system often in the form of an optimisation, outputting optimal parameters to achieve the best outcome. This chapter will look to examine the formulation of an optimisation problem, including applying objective functions found in literature, the constraints, and accompanying equations. Such an optimisation would be highly complex, particularly when dealing with a cascading system and dynamic variables. This section tackles the problem by starting with a simplified form that does not involve any sets or binary variables as part of the manipulated variables.

7.2 Initial Formulation

As covered in 2.4.2, the objective function can be defined in a variety of different forms. The overarching objective for this problem is to optimise the operation of the power scheme. Waikaremoana participates in a deregulated power market and part of a greater power network, meaning generation by this scheme is not necessarily constrained to a single demand. Generation is adjusted according to the present spot market price and the calculated value of water. Due to role of the market and overall goal to optimise the scheme from a business point of view, a profit based objective function would be the best fit for this problem.

A common profit maximisation objective function found in literature, consisting of two components representing generation revenue and stored value for time set. Optimising for cascaded schemes and multi-unit stations requires the summation of further sets. For an initial formulation, these other sets are not yet considered.

$$Max \quad \underbrace{\sum_{T} \underbrace{P_t . S_t}}_{Generation\ Revenue} + \underbrace{W_T . V_T}_{Stored\ Value} \tag{7-1}$$

This function can be described as a complex system with dynamic relationships between variables in the terms. Two notable variables that contribute to complexity in this function is the spot price, S, and water value, W. Both spot price and water value are dynamic and interlinked. They are functions of not only the market but also physical factors (unit availability, water storage limits etc).

The first term is simply a product of the amount of electricity generated in the examined period and the per unit spot price. The spot price is the short-term market variable that primarily depends on immediate power system demand and available supply which can also be affected by long term supply security (reservoir levels, drought conditions, fuel shortage). Bidding decisions are made with only short-term and long-term projections of this price, spot price in the context of optimisation would be a predicted or approximated value. True spot prices are only determined after a trading period has occurred and calculated by the system pricing manager. Spot price is defined by the market, but the marginal value and bidding decisions will also depend on asset or portfolio specific requirements (overfilling reservoirs, contract obligations etc.).

The second term is a product of reservoir volume at the end of the optimisation time horizon and an estimated 'water value'. Water value represents the estimated long-term value of available water stored in the reservoirs. It can be calculated in different ways and can include many factors depending on the user. The method to calculate water value is often regarded as proprietary knowledge by generation companies and can be viewed as part of an overall trading and generation strategy.

One method of approximating water value is relying on the futures market, the ASX. The futures market as outlined in chapter 3.3.2 allows traders to buy and sell future units of electricity for a projected price. It sets a market estimate for the worth of future generation and thus the value of water. The ASX offers options and futures for the NZ electricity market for peak and base load contracts covering monthly, quarterly, or yearly contract lengths. In other parts of the world, in multipurpose reservoirs, water can be sold as a commodity to users for irrigation, municipal and industrial use. A commodity price would be a useful alternate price indicator but would also require a 'futures' price for stored value prediction (U. S. Department of the Interior: Bureau of Reclamation, 2005).

This value can then be modified depending on the time horizon, node location and predicted transmission losses. It should be noted that water value is also only an estimated future value, it does not technically guarantee a return and does not consider how stored water will be used to realise its value. The stored value term is calculated using the final stored volume of the reservoirs, it does not consider inflows that are removed from the system by means other than generation.

Although spot price and water value are complex and important as part of a general bidding strategy, they are only part of trader considerations. They are useful as an optimisation starting point and guidelines as part of an overall strategy.

7.3 Equations and Constraints

The following are equations and constraints used for a simple first optimisation.

1. Reservoir water balance/Flow Continuity

$$V_t = V_{t-1} + \Delta T.3600(q_{Inflow} - q_{Outflow})$$

(7-2)

2. Level and Spill flow limits

Unit	$\underline{h}$	$\overline{h}$	$\underline{q}_{sp}$	$\overline{q}_{sp}$
L1	580.29	583.29	0	44

3. Discharge and Output Power Limits

Unit	$\underline{q}_{dis}$	$\overline{q}_{dis}$	$\underline{P}$	$\overline{P}$
6	0	17.5	0	18.0

4. Turbine Efficiency

$$\eta^* = \gamma_0 - \gamma_1(H - H_{avg}) + \gamma_2(H - H_{avg})^2 + \gamma_3(P - P_{avg}) - \gamma_4(P - P_{avg})^2 - \gamma_5(H - H_{avg})(P - P_{avg})$$

(7-3)

	Regression Coefficients						**Centring Means**	
Unit	**γ0**	**γ1**	**γ2**	**γ3**	**γ4**	**γ5**	**H Mean**	**P Mean**
6	0.81120	-0.00408	0.00057	0.01551	-0.00182	-0.00013	129.44	11.43

5. Flow rate via power production equation.

$$q_{dis} = \frac{P}{\eta^*.H.K}$$

(7-4)

6. Lake Volume

Lake Waikaremoana volume can be calculated through the following:

$$V_1 = h_1 \times (5.349 \times 10^7) - 3.104 \times 10^{10}$$

(7-5)

7.4 Variables

Analysis Variables	**Design Variables**
Inflows Lake Levels Gross Head Start Volume Power Generated Turbine Efficiency Discharge Flow End Volume Spot Price Cost of Fuel/Water Value	Power Generated
Analysis Functions	**Design Functions**
Power Revenue Stored Value Total Value	Maximise Total Value

7.5 Initial Optimisation

7.5.1 Generation and Stored Value

As a starting point, the objective function is narrowed to a simple one-dimension form by examining just a single unit, for one lake over a single period. Head level is set to a constant level. This optimisation is to simply examine the behaviour of the objective function when spot price is varied against a constant water value.

$$Max \quad \underbrace{P_t . S_t}_{Generation\ Value} + \underbrace{W_T . V_T}_{Stored\ Value} \tag{7-6}$$

Where t denotes the start of the period, while T denotes the end of the studied period.

On a first inspection of this objective function, it would appear to be a simple linear programming problem balancing the generation value and stored value based on the relative magnitudes of spot price and water value. One would expect that if $S_t > W_t$, the maximum amount of water would be used for generation while no generation would be expected when $S_t < W_t$. When $S_t = W_t$, any balance of generation and non-generation should result in the same revenue.

In a single period, there is a limited amount of electricity that can be generated by any single unit, and therefore a limited amount of water that can be used for generation. The above objective function attempts to optimise the best use of this volume of water. An attempt was made to solve the above problem through Excel Solver using the GRG nonlinear solver with multiple starting points to ensure the result found was a global optimum. The manipulated variable by the solver was simply the power output of the unit. Unit 6 was used as part of the example, over a single period of one hour. This optimisation was to observe the system behaviour when spot price was varied against a constant water value. The water value is set at $200/MWh, an arbitrary value in this case but in practice would either reflect the internally calculated water value or simply use a futures market-based value such as the ASX. It assumed no inflows and a constant head (highest efficiency head). Total volume available was set equal to the volume required if the unit ran for one hour at maximum flow rate.

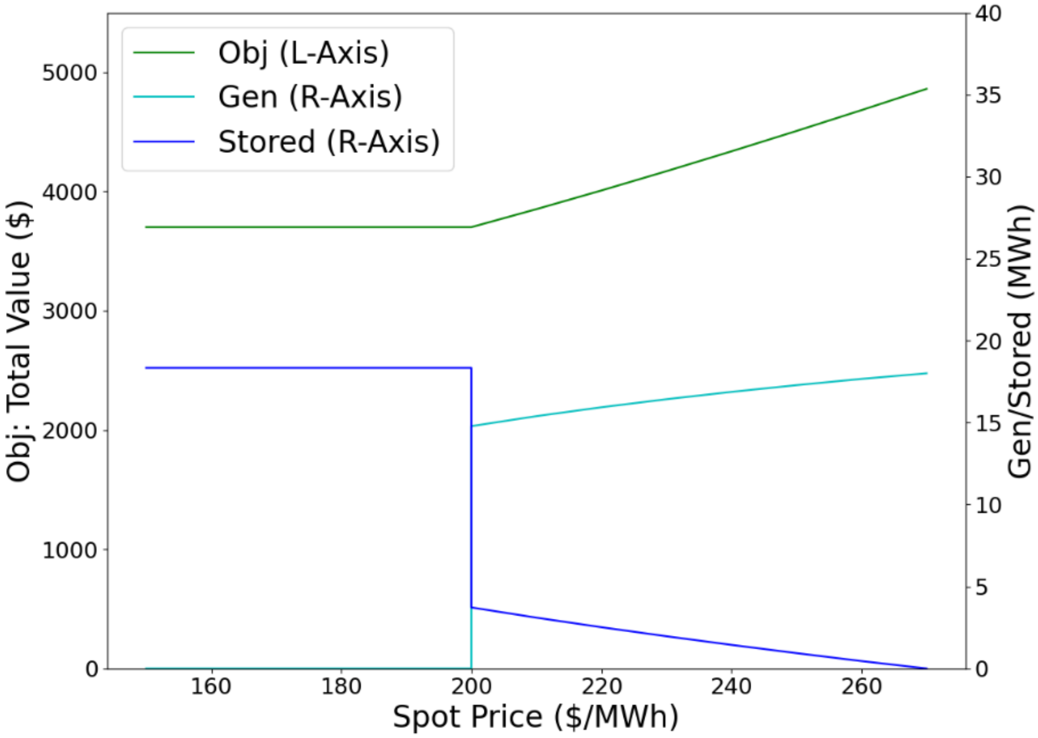

Figure 7-1: Generation-Stored Value Optimisation with Increasing Spot Price

The Excel Solver optimisation results are shown in Figure 7-1. The initial predictions proved to be accurate for when $S_t < W_t$, the optimum choice was to store water for this condition. When $S_t = W_t$, generating and storing water should provide the same per MWh value, but due to the nonlinear nature of the hydropower function, the unit must run at maximum efficiency to equal the value of storing all water. Running at maximum efficiency leaves a small portion of water for storage shown by the 'Stored' line in Figure 7-1, this provides additional value over simply running at full throttle. This fact gradually changes as S_t increases over W_t; it becomes more optimal to increase power output even through small decreases in unit efficiency.

A useful set of data from this test is the maximum efficiency points at the specified head, the average head at each station in this case. These points will also have a specific k-value (m^3/s per MW output) specific to these points, which will be useful to convert stored water into a stored generation.

Table 7.1: Efficiencies and k-values

Unit	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Max η*	0.844	0.831	0.811	0.801	0.870	0.863	0.856
Min k	0.933	0.948	0.614	0.622	0.573	1.039	1.046

This optimisation provides useful insight into the basic interaction between generation worth and stored value terms for a single generation unit, however, this objective function does not consider the costs associated with unit start-up.

7.5.2 Adding Start-up Costs

The additional stress and wear placed on hydro units during start-up will often increase maintenance requirements and/or the reduce the lifespan of the equipment. Start-up costs can be quite substantial, and can vary among different operators, units and stations depending on asset conditions and maintenance expenses. A common method for estimating this cost per start-up is by tallying all associated costs with maintaining and replacing the parts on the unit and dividing this by the number of cumulative start-ups (U. S. Department of the Interior: Bureau of Reclamation, 2005). An often-touted figure for start-up costs for hydrogenerators at Genesis is approximately $1000 per start-up.

As covered in the literature review, papers by Catalao et al. (2011) and Lima et al. (2013) among others introduced a binary term for operational status and start-up, indicating on/off states and the requirement for a start-up or not.

$$y_t = \omega_t(1 - \omega_{t-1})$$

(7-7)

The addition of these terms can be tested in the MS Excel Solver optimisation. The above equation was added as intermediary calculation and the start-up variable was defined as binary and added as a manipulated variable. The start-up cost term is added into the objective function as a product of the start-up variable and cost of each start-up, the function becomes the following:

$$Max \quad \underbrace{P_t.S_t}_{Generation\,Value} + \underbrace{W_T.V_T}_{Stored\,Value} - \underbrace{c.y_t}_{Startup\,Cost}$$

(7-8)

Spot price is again varied against a constant water value to observe the optimal generation pattern. Logically, for a system with a unit already running, start-up cost will be nullified. For a system that requires a start-up, the threshold at which generation outweighs start-up cost and stored value should increase compared with an objective without a start-up term.

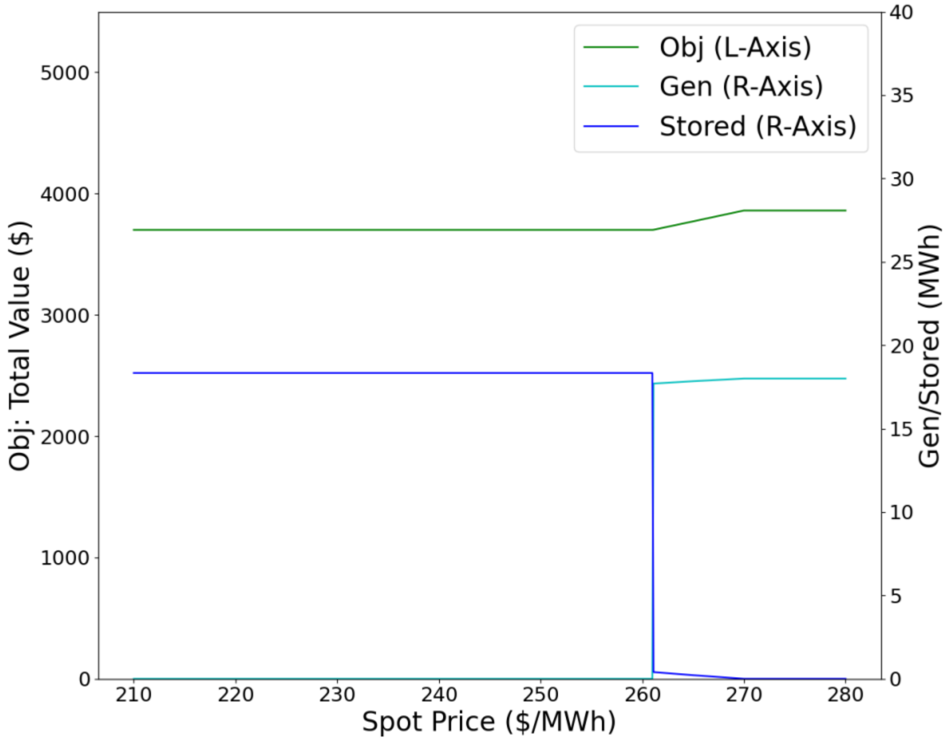

Figure 7-2: Generation-Stored-Startup Value Optimisation with Increasing Spot Price

Figure 7-2 above shows a similar result to Figure 7.1 but there is an evident shift up the spot price axis. Another observation is that the rate at which power output increases to its maximum is higher in this case, increasing from a high efficiency output at the threshold to its maximum at 270 $/MWh. Without a start-up term, the ramp up to maximum power output is approximately 70 $/MWh, between 200 and 270 $/MWh. The start-up cost pushes the optimisation to increase power output far more quickly to recoup its cost. The binary manipulated variable turned this into a mixed integer problem, where the Excel Solver package using a GRG Nonlinear solving method begins to run into difficulties finding a global optimum. Different results are reached depending on the initial values. A work around for this, is either by specifying multiple starting points or using an evolutionary method. For this problem, the minimum of 10 start points was adequate to reach the expected global optimum point without a big loss in compute time. The general behaviour behind the optimisation still revolves around the spot price and water value relationship but the fixed start-up cost changes two things. The spot price generation threshold increases to balance this cost. When choosing to generate, the rate at which power output increases (in relation to spot price) to maximum is steeper to make up for the new cost. It must be remembered though, the choice to generate will also be tied to the availability of water at the scheme. A possible implementation of this fact is a modification of the futures water value depending on the immediate water levels of the lakes.

The logic behind this optimisation becomes more complicated when examining more than a single time-period. For a multi-time set optimisation with unique spot prices for each time-period, the start-up cost would need to be weighed against a series of spot prices to optimise the best time to commit to a start-up.

7.5.3 Multiple Units and Lakes

Another step in the optimisation path is incorporating multiple units and multiple lakes. This means introducing a summation for each set in the objective function and constraints for each unit and lake. The hydropower unit commitment and scheduling problem becomes more complex when doing so. The objective function takes the form shown below:

$$Max \quad \underbrace{\sum_{I} P_{ti}.S_t}_{Generation\ Value} + \underbrace{\sum_{L} W_T.V_T}_{Stored\ Value} - \underbrace{\sum_{I} c_i.y_{ti}}_{Startup\ Cost} \tag{7-9}$$

Constraints for lake levels and spill flow, followed by constraints for unit discharge and power output. Each unit also needs their unique coefficients for their hydropower characteristic functions.

Table 7.2: Level and Spill Flow Limits Table

Unit	$\underline{h}$	$\overline{h}$	$\underline{q}_{sp}$	$\overline{q}_{sp}$
L1	580.29	583.29	0	44
L2	450.10	453.00	0	23
L3	246.20	247.60	0.005	52

Table 7.3: Discharge and Output Power Limits Table

Unit	$\underline{q}_{dis}$	$\overline{q}_{dis}$	$\underline{P}$	$\overline{P}$
6	0	17.5	0	18.0
7	0	17.5	0	18.0
1	0	13.0	0	20.0
2	0	13.0	0	20.0
3	0	13.0	0	20.0
4	0	24.0	0	23.6
5	0	24.0	0	23.6

Hydropower Characteristic Function:

$$\eta^* = 0.81120 - \gamma_1(H - H_{avg}) + \gamma_2(H - H_{avg})^2 + \gamma_3(P - P_{avg}) - \gamma_4(P - P_{avg})^2 - \gamma_5(H - H_{avg})(P - P_{avg}) \tag{7-10}$$

Table 7.4: Regression Coefficients Table

	Regression Coefficients						Centring Means	
Unit	**γ0**	**γ1**	**γ2**	**γ3**	**γ4**	**γ5**	**H Mean**	**P Mean**
6	0.81120	-0.00408	0.00057	0.01551	-0.00182	-0.00013	129.44	11.43
7	0.80172	-0.00467	0.00041	0.01410	-0.00168	-0.00033	129.36	11.68
1	0.76753	-0.00221	-0.00009	0.01550	-0.00138	-0.00024	204.88	11.22
2	0.77611	-0.00088	-0.00120	0.01215	-0.00149	-0.00057	204.86	12.04
3	0.86063	-0.00360	-0.00144	0.00787	-0.00164	0.00149	204.41	15.06
4	0.82108	0.04057	-0.00173	0.01463	-0.00129	0.00106	113.82	12.25
5	0.79406	0.01937	0.00611	0.02019	-0.00164	0.00103	113.86	10.19

For the sake of examining basic optimisation behaviour with these additional sets, the problem remains simplified by keeping the head level for each lake constant and setting initial levels to 50 % of the maximum level. In a more developed system, these head levels would vary across each time-period and may often reach levels where no more generation for the period is possible to remain within consent limits. Natural inflows into the lakes are also kept at a constant level. Lake volumes are calculated using lake area and available water height. Unit operating status were all kept as off to remove start-up cost from the problem for the time being, to better see the optimisation behaviour between spot price and water value without altering the profit threshold from start-ups.

Volume at the end of the time slice is calculated using the flow balance constraint. This volume is converted to a stored generation (MWh) value by using lowest k-value (m^3/s per MW) at each station which was previously found and documented in

Table 7.1.

It is assumed that generation stored at higher lakes will also be available for use at lakes further down the scheme, e.g., volume at Lake 1 consists of stored generation for all three stations.

$$Stored\ Gen.(MWh)_{L1} = \frac{V_1}{\max\left(k\ value_{L1}\right)} + \frac{V_1}{\max\left(k\ value_{L2}\right)} + \frac{V_1}{\max\left(k\ value_{L3}\right)} \tag{7-11}$$

This is only considered as an estimate as head levels (and k-values) may vary by the time the entirety of V_1 water makes its way down to the subsequent reservoirs. Stored generation for Lake 2 and 3 is similarly calculated.

$$Stored\ Gen._{L2} = \frac{V_2}{\max\left(k\ value_{L2}\right)} + \frac{V_2}{\max\left(k\ value_{L3}\right)} \tag{7-12}$$

$$Stored\ Gen._{L3} = \frac{V_3}{\max\left(k\ value_{L3}\right)}$$
(7-13)

A sum of each lake's stored generation gives the total stored generation.

$$Total\ Stored\ Generation\ (MWh) = Stored\ Gen._{L1} + Stored\ Gen._{L2} + Stored\ Gen._{L3}$$
(7-14)

This value is used as V_T in the optimisation objective function as part of the stored value term.

Another issue of note is the operation of the lake spillways, in the ideal optimal operation in terms of profit generation for the scheme, the spillways would not see little to no operation. Spilling water essentially wastes fuel that could otherwise be convert to revenue through generation. Realistically, there will be times when these spillways must be operated. When a station is offline, and water needs to bypass the station to reach other stations further down scheme. When an unexpected weather event has dumped more water than could be expected, which cannot be handled solely through powerhouse flow and threatens to increase lake levels over the consented levels. Lastly, when the scheme is required to provide spillway releases for recreational purposes. For this optimisation, spillway flows are kept to their minimum levels.

This optimisation builds on Section 7.5.1 using a summation of each set for their respective objective function terms. Excel Solver and its GRG nonlinear solving method are again used to try and solve this problem. The manipulated variables have been expanded to include all unit power output values. The limitations of Excel Solver become increasingly evident in this optimisation formulation. Attempts to solve the summed optimisation problem involving all units and lakes resulted in inconsistent results. The GRG nonlinear solver found different optimums depending on the initial values. The solver finds a local optimum and stops running after the objective value shows little change in around five iterations.

An example of this was when water value (W_T) is set at \$200 and spot price (S_T) at \$190. Logically when $W_T > S_T$, more objective value is gained in this scenario by not generating at all even before start-up cost is even considered. The solver however converges to a local optimum, if any of the starting values are close to the high efficiency running point the solver will move usually stay at this point instead of moving towards zero where the highest value is found. On occasions, some units are set to zero but where it does so is strongly dependent on the starting values.

Attempts to include an on/off binary variable as a manipulated variable leads to the solver being unable to find an optimum.

7.5.4 Limitations of Excel

The GRG nonlinear method stands for Generalized Reduced Gradient, the method looks at the gradient of the objective function as the input values change, reaching the optimum point when the gradient approaches zero

(P.E, 2016). When using only a single starting point, the solver may often find a local optimum depending on the initial values.

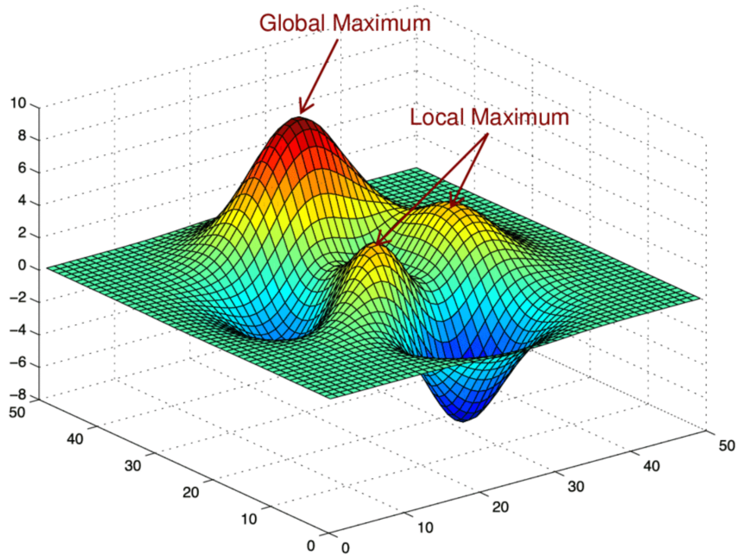

Figure 7-3: Illustration of local optimum and global optimum (Jin, 2015)

Figure 7-3 illustrates the difference between local and global maxima. A single GRG nonlinear optimisation will unlikely travel down a trough from a local maximum in order to reach the peak with the global maximum.

An alternate solving method available on Excel Solver is the Evolutionary method. Evolutionary methods attempt to emulate the process of natural selection in nature. The optimisation begins with many different random starting points or 'population'. The objective function is evaluated at each point. The points which result in the best objective value are selected, new points are made by modifying this point. This loop is continued until the objective value cannot be improved further. This method does not use derivatives and does not guarantee a global optimal point. Furthermore, the objective function is evaluated at each point individually, so the compute time is often longer than gradient based methods.

Another possible work-around of the local maximum issue while still using the GRG nonlinear method is the 'multi-start' option. This runs the solver multiple times using different random starting points, increasing the chance a global maximum is found. The minimum number of starting points for multi-start is 10. It is a useful compromise between GRG and evolutionary methods. Although it is quicker than the evolutionary method, the drawback to this option is that there is still a considerable increase in compute time compared with a single GRG nonlinear optimisation. For this formulation, solving took several minutes to resolve for a single time-period.

Excel solver also has other limitations. Although not a limitation for this optimisation problem, manipulated or decision variables are limited to 200. Solver also has trouble dealing with non-convex functions, like when binary variables are introduced as decision making variables. This has been true for this optimisation, adding the running (on/off) binary variable as a manipulated variable increased the solving time even for a problem with just a single unit.

7.5.5 Multi-Start Results

The multi-start option was used to attempt to solve the problem formulation in this section. The minimum number of start points was selected, which would increase the compute time by several times. In general, the optimisation is expected to behave very similarly to the single unit optimisation; the choice to generate should still be largely decided by the magnitude of spot price relative to water value. When spot price is below the water value and start-up cost threshold, the optimum configuration was no generation on all units. A little difference from the single unit case, was that the stored generation value at each lake is determined by the most efficient unit at the station/stage rather than individual units. e.g., at maximum efficiencies, Unit 6 is slightly more efficient than Unit 7 according to their regression equations, therefore the stored generation value is determined by Unit 6 rather than Unit 7. Revenue from Unit 7 must be greater than theoretically storing and generating later with Unit 6. This means the spot price thresholds in which each unit should start to generate will be slightly different from the single unit case due to the differences in efficiencies and use of the 'best' generator to value storing water.

Even with a multi-start optimisation, Excel Solver experienced problems with finding the global maximum depending on the initial values. Examples of the solver failing to find the global optimum can be observed at threshold points, e.g., $S_T = 199$, $W_T = 200$. Logic would dictate that no generation would result in the greatest objective value, but the GRG method will result in some units generating depending on the random initial values chosen. On the occasions the solver produced a logically sound result, the order in which the units were chosen to generate started from most to least efficient. Compute times increases from less than 30 seconds to over 250 seconds.

Although Excel Solver is a useful tool for simple optimisation problems, the limitations of such a platform have become evident in this scenario. To solve the most recent formulation and future formulations with the aim to optimise of a set of time slices, a more robust optimisation package needs to be used. Papers in literature have used commercial packages such as AMPL and CPLEX solver. Python3 offers a platform with various free optimisation algorithm packages that could potentially solve this problem efficiently. One possible package is the GEKKO optimisation library built by APMonitor. The Excel model would also need to be converted and organised from an Excel cell representation into an array format. Furthermore the visual representation for the flow model should be moved to a better platform like Python for easier development.

7.6 Future Work and Integration into the Digital Twin

Once the optimisation model has been translated into a state solvable by more complex mixed integer nonlinear algorithms, there many other factors and constraints that could be interesting to add into the optimisation model. The following factors could be investigated and implemented in future developments.

1. Head variations due to their dependence on flow rate.
2. Friction losses due to their dependence on flow rate.
3. Flow balancing across the entire scheme, ensuring that none of the smaller lakes run out of water.
4. Introducing a reserves market term into the objective function, units not generating can still make money as a reserve generator. Factoring this in could be an interesting change in dynamic.
5. Prediction models for spot price and water value, these values are often rough estimates, investigating more reliable methods to predict these values would be important undertaking.
6. Introduce rough running ranges, another constraint to improve turbine longevity and decrease maintenance costs.

DT platforms particularly for processing and energy focused industries have always been about optimising the system to increase efficiency and reduce fuel costs. The optimisation problem should function on top of the behaviour model. The behaviour model should provide the optimiser with the current state of the system and feed to it the current condition and constraints. From these the optimiser should determine a new best decision for every resolution period. The optimiser is only a piece to the DT tool. With the optimiser and flow model, there could also be integration of inflow models, spot price and water value prediction models that all feed into the model and optimiser.

7.7 Conclusions

- The objective function selected in the problem formulation was a profit-based function, with terms which are commonly seen in most published papers on hydropower scheduling optimisation. It was again selected here due to the big role the market plays in driving generation decisions and because the end goal is to maximise the amount of profit.
- Optimisation of a problem with only a single unit, lake and period was useful for viewing the relationship between spot price and water value. Spot price and water value have a complex interdependent relationship. Based on this optimisation, we can see that the objective function consists of a balance between generation and stored value. A unit will prefer to generate at max efficiency when the threshold is reached and slowly move towards max power output as the spot price raises over the water value further.
- The addition of start-up costs simply raised the threshold at which the best solution was to generate to recover the new introduced cost. An interesting behaviour to note was that the optimal generation point moved much more quickly toward maximum output as spot price was raised, likely due to the function moving to maximize revenue given the start-up cost.
- The limitations of Excel Solver GRG nonlinear began to show when multiple units were introduced into the optimisation problem. The solver had trouble finding a global maximum rather than a local maximum. This somewhat negated by using multiple start-points but this drastically increased the time to find a solution from several seconds to several minutes and still did not always guarantee a global optimum. Excel is not a suitable platform with a DT.
- To advance further in the optimisation pathway, the optimisation formulation must be moved to a platform that has access to more complex solvers that can deal with problems with multiple optimal points. A platform like Python using APMonitor’s GEKKO library has access to MINLP and MILP solvers for this type of problem but translating the problem into the format required for Python may prove to be challenging.

Chapter 8
Conclusions & Recommendations

8.1 Conclusions

This project has attempted a 'first step' in the development of a DT for the hydropower scheme at Waikaremoana. The following is a summary of the conclusions that have been made over all sections of this project.

- A DT model for the WPS should be more behaviour focused because the endgame use case for a WPS DT is decision making support, by means of optimisations or benchmarking for the traders. Visuals and connectivity have been considered as less important in early development processes; a simple flow diagram representation was sufficient. Regarding connectivity, data retrieval from PI Datalink may be too slow via Excel, a retrieval through Hilltop file export may provide a quicker alternative.
- For the behaviour modelling aspect of a DT, it was found to be a highly complex problem due to the hydropower scheme cascaded configuration, many non-linear relationships and interdependencies which has resulted in errors and uncertainties accumulating in attempts to reconcile between calculated and measured values.

Efficiency Regression

- Unit efficiency regression modelling showed that unit operational data showed good and excellent fits with the multi-variable polynomial model found in literature. The parameters, power output and head, were acceptable predictor variables for this efficiency model. Pre-centring the data was effective at reducing multicollinearity and provided more stable function coefficients and valid statistical p-values.
- The efficiency characteristic functions were able to generate hill diagrams consistent with those found in literature. The diagrams showed contour maps with visual regions of best efficiency. The results were not without odd regions of impossible efficiency, but this could largely be attributed to inconsistent input data rather than the process. This model still has plenty of room for improvement: improving the fit for a handful of units and accounting for friction loss to produce a traditional efficiency value rather than an overall efficiency.

Flow Balance, Hydraulic Relationships

- Attempts to perform a mass/water balance on the power scheme showed that a water flow balance of this system is highly reliant on input data. Comparisons of net flow levels calculated by flow balance against flow from water level change revealed small to considerable discrepancies. For most samples tested, differences were for most cases around the 2 m^3/s region indicated by the median difference. The difference graphed against net flow via flow balance showed a clear linear or quadratic relationship suggesting that the discrepancy is tied to a factor in the flow balance. The flow balance over each lake revealed that for most data points, the flow balance overshot (Waikaremoana and Whakamarino) or undershot (Lake Kaitawa) the expected net flow. This indicated that there is a systematic error in the balance.

- The biggest differences in reconciled lake levels were for Waikaremoana, specifically in scenarios where high inflow events occurred. Improved efficiency modelling and direct measurements of penstock flow rate is suspected to be the best areas to tackle to for behaviour model improvement.

Optimisation

- Optimisation of the hydropower scheduling and unit commitment problem is a complex non-linear balancing problem between generation and storing water for generating at a later period.
- Excel solver using a GRG nonlinear method was able to handle a simplified optimisation of a single unit, lake, and time interval. Under a constant water value and varying spot price, power output was manipulated to observe the basic behaviour of the proposed objective function. The choice to generate was confirmed to be dictated by spot price magnitude in relation to water value which acts as a threshold. Above the threshold, there is a transition from maximum efficiency to maximum output generation.
- The GRG nonlinear solver struggled to find a global optimum as the dimensions of the problem increased. The solver began finding local optimums when adding multiple units into the manipulated variable mix. The multi-start nonlinear option was tested but compute times increased to unreasonable levels. The same difficulties were encountered when adding binary variables as a manipulated variable.
- The optimisation formulation should be moved to a different platform with more efficient solving algorithms, that can deal with mixed integer and multidimensional problems. It is important to note that mathematical optimisations are top-down methods of optimisation, only providing the optimal decision

Overall, this project was able to present a first foray into the development of a DT for the WPS. It examined many of the considerations required for building a behaviour model and highlighted the areas of difficulty particularly regarding data reliability and hidden relationships with parameters.

8.2 Recommendations

The investigations into the many aspects of the project have revealed several areas where improvements can be made if a further, more developed DT is attempted in the future, spanning from input data quality to the translation of a hydropower optimisation to a Python format. Some of these improvements may have a bigger impact than others. The recommended areas to focus on for future work are as follows.

1. Work to improve the quality of input data for the DT. Data reliability is the corner stone of any DT project, particularly when it comes to modelling the system behaviour. The sections regarding efficiency and flow balance have raised questions about the accuracy of several calculated measurements.
    a. Consider ways to predict the true inflow value into Waikaremoana, particularly for high inflow events. Deep learning could be a useful tool to develop a black box model by feeding various Waikaremoana hydrology data. The model could help to predict true inflow based on hidden relationships within the hydrology data that would be unnoticed traditional regression methods.
    b. Examine the accuracy of calculated turbine flow rates used for turbine regressions, comparing against ultrasonic flow sensors installed at Piripaua. Use of true sensors for all units or more calibrations could help rule out turbine flow rates as a source of error.
    c. Gather more data from turbine operations
2. Improving the efficiency regressions. The regressions have shown a good fit with the model proposed in literature but there is still room for improvement and other methods to explore.
    a. Find all friction loss coefficients for each intake and penstock with a standard test across each of lake intakes. Use this to calculate head loss due to friction, net head, and subsequently traditional turbine efficiency.
    b. Explore other methods of regression or prediction, namely ridge regression which can deal with data collinearity and deep learning. DL could be a promising method for producing a black box tool for predicting efficiencies based on the production parameters.
3. Refine the flow model. The flow balance comparison against measured stored volume change showed a noticeable calculated net flow dependent gap between the two. Reducing this gap is essential for an accurate model due to the flow on effects as a cascaded water system. Investigate the causes for the differences between these net flows. Determine the factor that causes the difference to rise as the calculated net flow increases in magnitude and correct the model for this factor.
4. Although a 2D representation is sufficient and the visual aspect of the DT is less of a priority than the behaviour aspect, a dynamic 2D representation that could update based on the time interval would still be a useful tool to visualise the model status. Converting the excel flow diagram into a dynamic 2D diagram using a platform like Python would be an achievable and logical next step in terms of DT looks-like aspect.

5. Optimisation was the ultimate goal from the outset of the Waikaremoana DT idea. There is enormous potential and with many different options that could be explored to reach this goal.
    a. Continue following the mathematical optimisation pathway. Translate the optimisation problem to a programming-based language and use solvers capable of mixed integer and nonlinear solving such as GEKKO or CPLEX. Work to increase the number of sets or dimensions of the problem (solving over many time intervals before solving with multiple units then multiple lakes).
    b. Investigate the application and development of reinforcement learning for this optimisation. Convert the flow balance into a form that could be used as the basis for an ‘environment’. Translate the optimisation objective function terms and constraints into part of a ‘reward’ function.

Development of a working DT concept is a complex and difficult undertaking. Much more development is required to produce a working model with the desired capabilities. The recommendations are the next areas that could be worked on to reach the desired goal.

Chapter 9
References

Aasgård, E. K., Naversen, C. Ø., Fodstad, M., & Skjelbred, H. I. (2018). Optimizing day-ahead bid curves in hydropower production. *Energy Systems*, *9*(2), 257–275. https://doi.org/10.1007/s12667-017-0234-z

Amani, A., & Alizadeh, H. (2021). Solving Hydropower Unit Commitment Problem Using a Novel Sequential Mixed Integer Linear Programming Approach. *Water Resources Management*, *35*(6), 1711–1729. https://doi.org/10.1007/s11269-021-02806-6

Anderson, G. P. (1948). Waikaremoana: Problem of Lake Control. *Proceedings of the New Zealand Institution of Engineers, XXXIV*, p.508.

Arce, A., Ohishi, T., & Soares, S. (2002). Optimal dispatch of generating units of the Itaipu hydroelectric plant. *IEEE Transactions on Power Systems*, *17*(1), 154–158. https://doi.org/10.1109/59.982207

Armstrong, M. M. (2020, December 4). *Cheat Sheet: What is Digital Twin?* IBM Business Operations Blog. https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin/

Aspen. (2022). *Aspen HYSYS | Process Simulation Software | AspenTech*. https://www.aspentech.com/en/products/engineering/aspen-hysys

Bardsley, W. E., & Choudhry, S. (2000). An Approach to Estimating Hydro Power System Income Gain from Computerized Water Scheduling. *Natural Resources Research*, *9*(3), 215–222. https://doi.org/10.1023/A:1010143717898

Bowerman, B. L., O'Connell, R. T., Koehler, A. B., & Bowerman, B. L. (2005). *Forecasting, time series, and regression: An applied approach* (4th ed). Thomson Brooks/Cole.

Bradley, S. P., Hax, A. C., & Magnanti, T. L. (1977). *Applied mathematical programming*. Addison-Wesley Pub. Co.

Cant, G., Hodge, R., Wood, V., & Button, L. (2004). *The Impact of Environmental Changes on Lake Waikaremoana and Lake Waikareiti*. The Waitangi Tribunal.

https://forms.justice.govt.nz/search/Documents/WT/wt_DOC_93526217/Wai%20894%2C%20D001.pdf

Catalão, J. P. S., Mariano, S. J. P. S., Mendes, V. M. F., & Ferreira, L. A. F. M. (2006). Parameterisation effect on the behaviour of a head-dependent hydro chain using a nonlinear model. *Electric Power Systems Research*, *76*(6–7), 404–412. https://doi.org/10.1016/j.epsr.2005.09.002

Catalão, J. P. S., Pousinho, H. M. I., & Mendes, V. M. F. (2011). Hydro energy systems management in Portugal: Profit-based evaluation of a mixed-integer nonlinear approach. *Energy*, *36*(1), 500–507. https://doi.org/10.1016/j.energy.2010.10.014

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. https://doi.org/10.1109/59.962421

Cristian Finardi, E., & Reolon Scuzziato, M. (2013). Hydro unit commitment and loading problem for day-ahead operation planning problem. *International Journal of Electrical Power & Energy Systems*, *44*(1), 7–16. https://doi.org/10.1016/j.ijepes.2012.07.023

Daadaa, M., Séguin, S., Demeester, K., & Anjos, M. F. (2021). An optimization model to maximize energy generation in short-term hydropower unit commitment using efficiency points. *International Journal of Electrical Power & Energy Systems*, *125*, 106419. https://doi.org/10.1016/j.ijepes.2020.106419

Dal' Santo, T., & Simões Costa, A. (2016). Hydroelectric unit commitment for power plants composed of distinct groups of generating units. *Electric Power Systems Research*, *137*, 16–25. https://doi.org/10.1016/j.epsr.2016.03.037

Davies, T. R., McSaveney, M. J., & Beetham, R. D. (2006). Rapid block glides: Slide-surface fragmentation in New Zealand's Waikaremoana landslide. *Quarterly Journal of Engineering Geology and Hydrogeology*, *39*(2), 115–129. https://doi.org/10.1144/1470-9236/05-041

de L. Main, W. (1976). Morphology of lake Waikaremoana, New Zealand, and reconnaissance of its benthic fauna. *New Zealand Journal of Marine and Freshwater Research*, *10*(4), 597–611. https://doi.org/10.1080/00288330.1976.9515643

Diaz, F. J., Contreras, J., Munoz, J. I., & Pozo, D. (2011). Optimal Scheduling of a Price-Taker Cascaded Reservoir System in a Pool-Based Electricity Market. *IEEE Transactions on Power Systems*, *26*(2), 604–615. https://doi.org/10.1109/TPWRS.2010.2063042

Diniz, A. L., Esteves, P. P. I., & Sagastizabal, C. A. (2007). A Mathematical Model for the Efficiency Curves of Hydroelectric units. *2007 IEEE Power Engineering Society General Meeting*, 1–7. https://doi.org/10.1109/PES.2007.385632

Diniz, A. L., & Maceira, M. E. P. (2008). A Four-Dimensional Model of Hydro Generation for the Short-Term Hydrothermal Dispatch Problem Considering Head and Spillage Effects. *IEEE Transactions on Power Systems*, *23*(3), 1298–1308. https://doi.org/10.1109/TPWRS.2008.922253

Dreyer, M., Nicolet, C., Gaspoz, A., Gonçalves, N., Rey-Mermet, S., & Boulicaut, B. (2021). Monitoring 4.0 of penstocks: Digital twin for fatigue assessment. *IOP Conference Series: Earth and Environmental Science*, *774*(1), 012009. https://doi.org/10.1088/1755-1315/774/1/012009

Electricity Authority. (2017). *Overview of the FTR market*. https://www.ea.govt.nz/operations/wholesale/hedges/ftr-market/

Electricity Authority. (2018). *Electricity in New Zealand 2018*. https://www.ea.govt.nz/about-us/media-and-publications/electricity-new-zealand/

Electronics Tutorials. (2022). *The Current Transformer*. Electronics Tutorials. https://www.electronics-tutorials.ws/transformer/current-transformer.html

Engineering NZ. (2022, March). *Waikaremoana Power Stations | Engineering New Zealand*. https://www.engineeringnz.org/programmes/heritage/heritage-records/waikaremoana-power-stations/

Engineering Toolbox. (2022). *Water—Density, Specific Weight and Thermal Expansion Coefficients*. https://www.engineeringtoolbox.com/water-density-specific-weight-d_595.html

Finardi, E. C., & da Silva, E. L. (2005). Unit commitment of single hydroelectric plant. *Electric Power Systems Research*, *75*(2), 116–123. https://doi.org/10.1016/j.epsr.2005.01.008

Frost, J. (2019). *Regression analysis: An intuitive guide for using and interpreting linear models*.

Fuji Electric. (2022). Flow Meters for Fluid. *Fuji Electric Corp. of America*. https://americas.fujielectric.com/products/instrumentation/flow-meters/flow-meters-for-fluid/

Garrido, J., Zafra, Á., & Vázquez, F. (2009). Object oriented modelling and simulation of hydropower plants with run-of-river scheme: A new simulation tool. *Simulation Modelling Practice and Theory*, *17*(10), 1748–1767. https://doi.org/10.1016/j.simpat.2009.08.007

Genesis. (2022). *Our Generation Assets*. Genesis. https://www.genesisenergy.co.nz/assets

Genesis Energy Ltd. (2015). *Annual Environmental Report (AER) for the Waikaremoana Power Scheme (WPS) 2014/15* [Environmental Report].

Global Change Data Lab. (2022). *Our World In Data*. Our World in Data. https://ourworldindata.org/

Grieves, M. (2014). Digital twin: Manufacturing excellence through virtual factory replication. *White Paper*, *1*, 1–7.

Hardcastle, A. (1929, November). *Tuai*. Tuai | Items | National Library of New Zealand | National Library of New Zealand. https://natlib.govt.nz/records/22599247

Harding, S., & Mills, G. (2020). *Embedded Hydro Generation Opportunities in New Zealand*. Ministry of Business, Innovation and Employment. https://www.mbie.govt.nz/assets/embedded-hydro-generation-opportunities-in-new-zealand.pdf

Hawke's Bay Regional Council. (2022, June 16). *Hawke's Bay Lakes*. https://www.hbrc.govt.nz/hawkes-bay/rivers-and-lakes/hawkes-bay-lakes/

Hopkirk, C. L. (2011). *Forecasting Lake Waikaremoana Water Availability for Hydro Power* [Thesis, University of Waikato]. https://researchcommons.waikato.ac.nz/handle/10289/5297

IBM. (2022, April 5). *What is a digital twin? | IBM*. https://www.ibm.com/topics/what-is-a-digital-twin

J. Finkelstein. (1973). Survey of New Zealand Tank Evaporation. *Journal of Hydrology New Zealand Volume 12, 1973*, *12*(2), 119–132.

Jin, C. (2015). *A Sequential Process Monitoring Approach using Hidden Markov Model for Unobservable Process Drift*. https://doi.org/10.13140/RG.2.2.32537.83041

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, *29*, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

Kong, J., Skjelbred, H. I., & Fosso, O. B. (2020). An overview on formulations and optimization methods for the unit-based short-term hydro scheduling problem. *Electric Power Systems Research*, *178*, 106027. https://doi.org/10.1016/j.epsr.2019.106027

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, *51*(11), 1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474

Li, X., Li, T., Wei, J., Wang, G., & Yeh, W. W.-G. (2014). Hydro Unit Commitment via Mixed Integer Linear Programming: A Case Study of the Three Gorges Project, China. *IEEE Transactions on Power Systems*, *29*(3), 1232–1241. https://doi.org/10.1109/TPWRS.2013.2288933

Lima, R. M., Marcovecchio, M. G., Novais, A. Q., & Grossmann, I. E. (2013). On the Computational Studies of Deterministic Global Optimization of Head Dependent Short-Term Hydro Scheduling. *IEEE Transactions on Power Systems*, *28*(4), 4336–4347. https://doi.org/10.1109/TPWRS.2013.2274559

LINZ Data Service. (2022). *New Zealand Terrain Map* [Terrain]. LINZ Data Service. https://data.linz.govt.nz/data/

Miles, J. (2009). *Solving Operational Problems With Inverted Siphons*. North Shore City Council. https://www.waternz.org.nz/Attachment?Action=Download&Attachment_id=1343

Ministry of Business, Innovation and Employment. (2021). *Data tables for electricity* [Electricity statistics]. https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-statistics-and-modelling/energy-statistics/electricity-statistics/

Munoz-Hernandez, G. A., Mansoor, S. P., & Jones, D. I. (2013). *Modelling and controlling hydropower plants*. Springer.

Natusch, G. G. & Electricorp Production (N.Z.). (1992). *Power from Waikaremoana: A history of Waikaremoana hydro-electric power development*. Electricorp Production.

Omega Engineering. (2018, August 28). *What is an Ultrasonic Flow Meter?* Https://Www.Omega.Com/En-Us/. https://www.omega.com/en-us/resources/ultrasonic-flow-meters

P.E, C. Y. (2016, November 29). Excel Solver: Which Solving Method Should I Choose? *EngineerExcel*. https://engineerexcel.com/excel-solver-solving-method-choose/

Pérez-Díaz, J. I., Wilhelmi, J. R., & Arévalo, L. A. (2010). Optimal short-term operation schedule of a hydropower plant in a competitive electricity market. *Energy Conversion and Management*, *51*(12), 2955–2966. https://doi.org/10.1016/j.enconman.2010.06.038

Piekutowski, M., Litwinowicz, T., & Frowd, R. J. (1994). Optimal short-term scheduling for a large-scale cascaded hydro system. *IEEE Transactions on Power Systems*, *9*(2), 805–811. https://doi.org/10.1109/59.317636

Poletti, S. (2021). Market Power in the New Zealand electricity wholesale market 2010–2016. *Energy Economics*, *94*, 105078. https://doi.org/10.1016/j.eneco.2020.105078

Poletti, S., & Staffell, I. (2021). Understanding New Zealand's wind resources as a route to 100% renewable electricity. *Renewable Energy*, *170*, 449–461. https://doi.org/10.1016/j.renene.2021.01.053

Pritchard, S. (2021). *Pumped storage hydropower—A key to reliable renewables—NS Energy*. https://www.nsenergybusiness.com/features/pumped-storage-hydropower-a-key-to-reliable-renewables/

Read, S. A. L., & Riley, P. B. (1992). Lake Waikaremoana -Present day stability of landslide barrier. In *Landslides: Proceedings of the sixth international symposium*. A. A. Balkema.

Riemer-Sorensen, S., & Rosenlund, G. H. (2020). Deep Reinforcement Learning for Long Term Hydropower Production Scheduling. *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, 1–6. https://doi.org/10.1109/SEST48500.2020.9203208

Shawwash, Z. K., Siu, T. K., & Russell, S. O. D. (2000). The B.C. Hydro short term hydro scheduling optimization model. *IEEE Transactions on Power Systems*, *15*(3), 1125–1131. https://doi.org/10.1109/59.871743

Shen, D., & Yang, Q. (2012). Electricity Market Regulatory Reform and Competition – Case Study of the New Zealand Electricity Market. In *Energy Market Integration in East Asia: Theories, Electricity Sector and Subsidies* (pp. 103–140). Economic Research Institute for ASEAN and East Asia (ERIA). https://EconPapers.repec.org/RePEc:era:chaptr:2011-rpr-17-06

Siemens. (2022). *Plant Simulation and Throughput Optimization | Siemens Software*. Siemens Digital Industries Software. https://www.plm.automation.siemens.com/global/en/products/manufacturing-planning/plant-simulation-throughput-optimization.html

Taylor, N. (2019). *Waikaremoana Power Scheme Manual*. Genesis Energy Ltd.

U. S. Department of the Interior: Bureau of Reclamation. (2005). *Hydroelectric Power*. Reclamation: Managing Water in the West. https://www.usbr.gov/power/edu/pamphlet.pdf

Wang, H., Ahmed, O., & Smith, B. (2021, August 27). *Developing a digital twin for hydropower systems—An open platform framework—International Water Power*. https://www.waterpowermagazine.com/features/featuredeveloping-a-digital-twin-for-hydropower-systems---an-open-platform-framework-9037947/

Xu, B., Wang, J., Wang, X., Liang, Z., Cui, L., Liu, X., & Ku, A. Y. (2019). A case study of digital-twin-modelling analysis on power-plant-performance optimizations. *Clean Energy*, *3*(3), 227–234. https://doi.org/10.1093/ce/zkz025

Yu, H., Hou, J., & Xu, H. (2021). Research on the Digital Twin Model of Fuchunjiang Hydropower Plant Based on 3D Laser Scanning Technology. *IOP Conference Series: Earth and Environmental Science*, *784*(1), 012015. https://doi.org/10.1088/1755-1315/784/1/012015

Yu, J., Liu, P., & Li, Z. (2020). Hybrid modelling and digital twin development of a steam turbine control stage for online performance monitoring. *Renewable and Sustainable Energy Reviews*, *133*, 110077. https://doi.org/10.1016/j.rser.2020.110077

Yu, W., Patros, P., Young, B., Klinac, E., & Walmsley, T. G. (2022). Energy digital twin technology for industrial energy management: Classification, challenges and future. *Renewable and Sustainable Energy Reviews*, *161*, 112407. https://doi.org/10.1016/j.rser.2022.112407

Zulkifli, A. S., Muhamad Razali, N. M., Marsadek, M., Yahya, Z., & Juhana Tengku Hashim, T. (2015). Power Plant Performance Based on Power-Efficiency Curve of Hydroelectric Power Plant. *Applied Mechanics and Materials*, *785*, 516–520. https://doi.org/10.4028/www.scientific.net/AMM.785.516

Chapter 10

Appendix A

10.1 Historical Background

10.1.1 Lake Waikaremoana Origin

Lake Waikaremoana was first mapped geologically in 1876 and is believed to have been formed some 2200 years ago following a giant rock landslide (Davies et al., 2006). The landslide consists of a section of rock-avalanche debris 1 km^3 in volume and a relatively intact block 1.4 km^3. The entire landslide measures at a total of around 2.4 km^3 in bulked volume and covers an area of 18 km^2. The two distinct sections were once thought to be from two separate sliding events by Marshall in 1927 or three stages by Read et al. (1992), but scale model studies by Beetham et. al suggested that the rock avalanche and movement of the intact block were likely to have been involved in a simultaneous motion (Davies et al., 2006).

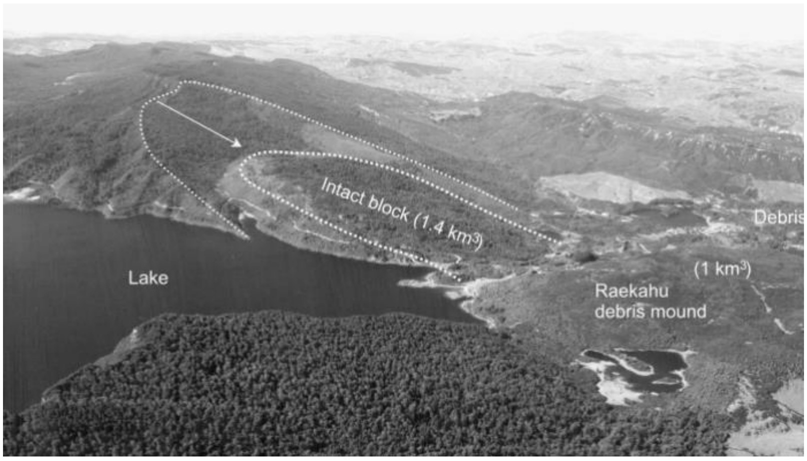

Figure 10-1: Aerial View of the Waikaremoana Landslide, photo by Lloyd Homer (Davies et al., 2006)

The 275 m thick intact block is theorised to have slid away from the escarpment with the rock avalanche, moving around 2 km and pushing rock-avalanche debris upwards of 150 m to form a debris mound now known as the Raekahu hill (Davies et al., 2006). Hopkirk (2011) has compiled a description of the landslide dam using Read et al.(1992). The full landslide dam spans 8 km along the Waikaretaheke valley and is on average 400 m high with an average surface slope of 6° towards the river and 16° towards the sea (Natusch & Electricorp Production (N.Z.), 1992). It consists of randomly oriented large sandstone and siltstone blocks. The block sizes can vary up to tens of meters in diameter and is supported by a matrix of fine-grained sand, silt, and pumice (Hopkirk, 2011).

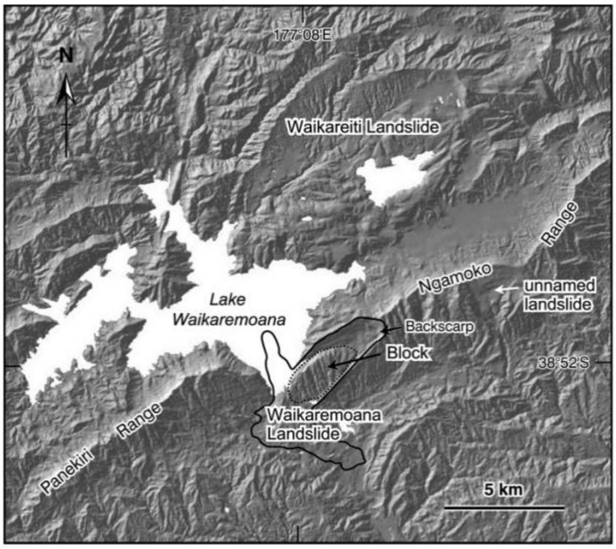

Figure 10-2: Satelite Map of Lake Waikaremoana and the Landslide (Davies 2006)

The cause of the landslide has been theorised to be a strong nearby earthquake due to the proximity of active fault lines and the frequency of seismic activity in the Napier region (Davies et al., 2006). Prior to the landslide the Waikaretaheke River flowed through a kilometre wide valley and between the Panekiri range to the SW and Ngamoko range to the east (Davies et al., 2006). Cavities in the broken slabs allowed some flow of the Waikaretaheke River through the debris but inflows from the valley slowly accumulated over ten years (based on a flow of 17 m^3/s) to form Lake Waikaremoana as it is now known (Natusch & Electricorp Production (N.Z.), 1992). The raising of the lake eventually allowed further flows to penetrate the slip and form water springs on the southern side of the debris which in turn, stabilised the lake level. Valley trees submerged by creation of the lake remained preserved for thousands of years before being removed for boat safety reasons, carbon dating of these trees provided the 2200-year ago estimate (Natusch & Electricorp Production (N.Z.), 1992).

Dams created from landslide events are said to be usually short-lived, with only 10% surviving beyond 1 year (Read & Riley, 1992). Riley and Read (1992) have theorised the reasons for its survival. The factors that determine landslide dam survival are overtopping, the size of the upstream impoundment, the inflows, and the landslide material. Erosion of the downstream face of the landslide is often the major cause of landslide dam failures. Waikaremoana's survival for over 2000 years can be attributed to the integrity of the intact block imbedded into the landslide debris and several factors that allowed the dam to regulate overtopping and minimise erosion of the downstream barrier face. These factors include: a hollow in the intact block that fills with water when lake overflow would otherwise occur, a large lake volume relative to catchment flow volume, the diversion of the Waikaretaheke River from the lower half of the landslide to an intact bed rock, and finally the porosity of the dam allowed leakage flows of up to 9-12 m^3/s without notable internal erosion of the landslide. These factors all contributed to regulating the discharge of the lake and minimising overtopping and spilling down the downstream face (Read & Riley, 1992).

10.1.2 Historical Development

10.1.2.a Initial Surveys

The first survey of Lake Waikaremoana is often cited as occurring in the mid to late 19th century during the 'Great Trigonometric Survey' in 1876. The hydropower potential of the lake was first commented on by J W Witty as early as 1866, however this date is highly questionable as electromagnetic generators were still in the development stage at this time. It is much more likely that his comments were written during a survey in 1876 or even as late as 1886 when the use of electricity in New Zealand was beginning to gain momentum.

In 1903, Minister of Works, William Hall-Jones ordered his department to examine as many potential hydropower sites as possible and commissioned an electrical engineer from the United States, L M Hancock for an independent assessment of potential hydropower sites in the North Island. Hancock was unable to see Waikaremoana but noted the potential of a partial scheme for the lake. P S Hay, the Superintending Engineer for the Public Works Department wrote a report in September 1904 that detailed estimates of a scheme at Waikaremoana using flow and barometric measurements from the lake. He estimated that an output of 50 MW would be possible utilising a head of 306 m. He noted the challenges with managing leaks through the lakebed, control of which at Onepoto could yield a doubling of the output. The demand for electricity in New Zealand during this time was approximately <40MW, as such Waikaremoana was classified as 'beyond present use' by Hay (Natusch & Electricorp Production (N.Z.), 1992).

It was not until 1916, when G P Anderson and a small team was tasked to examine Waikaremoana in accurate detail and produce a preliminary power station layout. A review of his data showed potential in utilising Lake Kaitawa, 120m below Waikaremoana as another reservoir by diverting to it, the natural Waikaremoana discharge and raising Lake Kaitawa's level by 3 m. A dam to create a third lake on the Whakamarino Flat with a powerhouse at Piripaua was also seen as a possibility.

Anderson's work was essential for all subsequent planning work for the development and was later used by Hay in 1918 in planning the layout for Tuai Station. In 1918 the Government produced a report with plans for a national grid with stations at Mangahao, Waikaremoana and on the Waikato River using 110kV lines to demand centers.

10.1.2.b The Temporary Station

In 1918, Vickerman and Lancaster, an engineering consultancy, co-founded by Hay, developed a proposal for a powerhouse on the edge of the Whakamarino flats fed using water taken from Lake Kaitawa via a pipeline. After some modification and discussion, it was agreed that a smaller initial temporary station would be built, providing 750 kW with a 33 kV transmission line to the town of Wairoa. It would provide enough power to Wairoa and for building the main stations (Natusch & Electricorp Production (N.Z.), 1992).

Work on the temporary station began in 1920, with transportation of materials 50km from Wairoa to the lake sites proved to be quite challenging due to road sections becoming muddy bogs. Work was undertaken to build metalled roads and several bridges to ease travel to the site in anticipation of further works to build the main Tuai Station. The turbines were supplied by Vickers in England and penstock (measuring 0.6m in diameter, splitting into two 0.3m penstocks on one end) from Glasgow. After delays, the station was officially opened in March 1923 with two 350 kW generating units. The station cost $236,000 and operated well until Tuai was brought online in 1929, several units were transferred to Tuai for use as auxiliary generators, the penstocks were salvaged and later used for irrigation lines in Otago.

10.1.2.c Tuai

Tuai is the main station of the Waikaremoana scheme and it's control room is the local control centre for the scheme. Kaitawa Station and Piripaua also have small control rooms that were initially manned, but it was planned from onset that the entire scheme would be controlled from Tuai using electrical signals. By 1924 it was becoming clear that the Mangahao scheme, only just coming into operation, could not meet growing electricity demand in the lower North Island, this prompted the start of preparatory work for the construction of Tuai Station in April 1926 (Engineering NZ, 2022)).

Due to the initial difficulty in delivering water safely and economically from Waikaremoana, the plan to build Tuai on the Whakamarino plain remained the most feasible. The Waikaretaheke River flowing from Waikaremoana would be diverted into Lake Kaitawa where a low dam would create the head pond for Tuai with a gross head of 205 m. Water would enter an intake at Lake Kaitawa, travel through a tunnel, surge chamber and penstocks before driving the turbines at Tuai and discharging onto the Whakamarino Flat. A small tailrace pond for Tuai was initially built with a weir to maintain a sufficient draught tube water level but ultimately a dam at the lower end of the Whakamarino Flat and flooding of the plain was planned. This reservoir would become known as Lake Whakamarino, storing water discharged from Tuai for use as the head pond for a possible lower development powerhouse.

Construction of Tuai proceeded quickly, with the official opening of the station occurring in November 1929, only two years and nine months after approval to start had been received. The station initially housed two 16 MW horizontal Francis turbine units supplied by Boving and Co in London.

Figure 10-3: Opening of Tuai Station, 1929 (Hardcastle, 1929)

Tuai operated satisfactorily after its opening and survived two earthquakes, the Napier earthquake in 1931 and 1934 earthquake, with minor damage to the water conveying equipment, moderate damage to concrete partition walls in the powerhouse and minor cracking to the Lake Kaitawa lower dam structure. The main powerhouse structure proved to be structurally firm.

In 1936, the need for additional power soon became apparent which prompted the start of works to add a third unit, 20 MW in size at Tuai to accommodate additional peaking capacity. The unit was sized while considering water availability at Kaitawa when an upper development for the scheme would be implemented. Work to add the additional penstock and generation equipment commenced in 1937 and the unit came into operation in late 1939 (Natusch & Electricorp Production (N.Z.), 1992).

10.1.2.d Piripaua

A lower development to take advantage of approximately 100 m of gross head available 5km downstream of Whakamarino had been viewed as a possibility since reports in 1918. With the suspension of works at Onepoto, intensive exploratory works was done in 1936 and 1937 for a station at Piripaua. In 1938, approval for the lower station arrived and commencement of work began. The planned plant was to utilise an average of 18 m^3/s flow and 110 metres gross head with two 20 MW generators.

A 2.6 km long tunnel (Taylor 2019), with a 4.5 m diameter was built to take water from Whakamarino to a surge chamber for Piripaua. The tunnel crosses tricky swampy terrain and a stream with the use of two inverted siphons, the first use of this technology in New Zealand. The siphons dropped the tunnel level below the undesirable terrain before passing under and climbing back up.

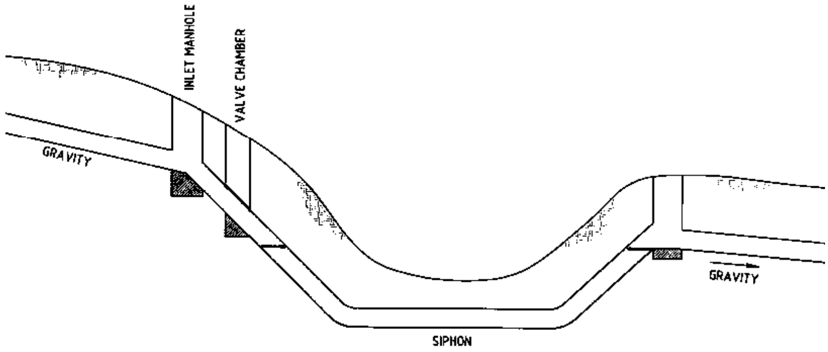

Figure 10-4: Example Diagram of an Inverted Water Siphon (Miles, 2009)

The first generator unit was commissioned in 1943 and the second in 1944. Plans for the scheme's final station, Kaitawa, were in the works even before Piripaua came into operation.

10.1.2.e Kaitawa

The final station to be built at Waikaremoana was on the edge if Lake Kaitawa. After plans for the upper development were shelved during the 1930s, in 1942 there was renewed interest for a station that would take advantage of the 110 m gross head from Waikaremoana and would increase the available water for the stations downstream. Tunnelling at Onepoto had been previously halted due to discovering considerable water flow once they reached the main body of the slip. Initial plans were for two independent 2.4 m diameter concrete lined tunnels that would feed into straight penstocks, but these plans were altered in 1941 to include a single 3 m diameter tunnel for the first 220 m section of the Onepoto intake tunnels.

The main obstacle to tunnelling was sudden water flows into the tunnel of up to 100 L/s, this was negated by improved pumping and grouting resources. As they tunnelled into the slip material, grouting material was continually applied from the surface and from within the tunnels. The tunnel was secured using continuous close timbering, followed by concreting material. Work on the tunnel was slow, progressing at a rate of 5 m per week.

Heavy grouting was required for the last 60 m of the twin tunnels, with one hole requiring 210 m^3 of grouting. At times the tunnel face had to be sealed using high pressure concreting to stem water flows before being drilled out to continue work. The last 60 m took seven months to complete. The entire tunnel took five years to complete in part due to tunnelling difficulties but also due to manpower shortages from the war.

The Kaitawa development did not include a true surge chamber, instead annular collars were built into the upper section of the gateshaft which allowed the machines to be started more quickly. Kaitawa Station could handle a rapid increase in load but not a sudden shutdown of the generators. Relief valves were built to discharge water in the event of a generator trip. Further water diversion structures were also built to divert water to Lake Kaitawa and on Lake Kaitawa to spill into the Waikaretaheke River.

The generators were commissioned in 1947 with a combined name plate capacity of 32 MW.

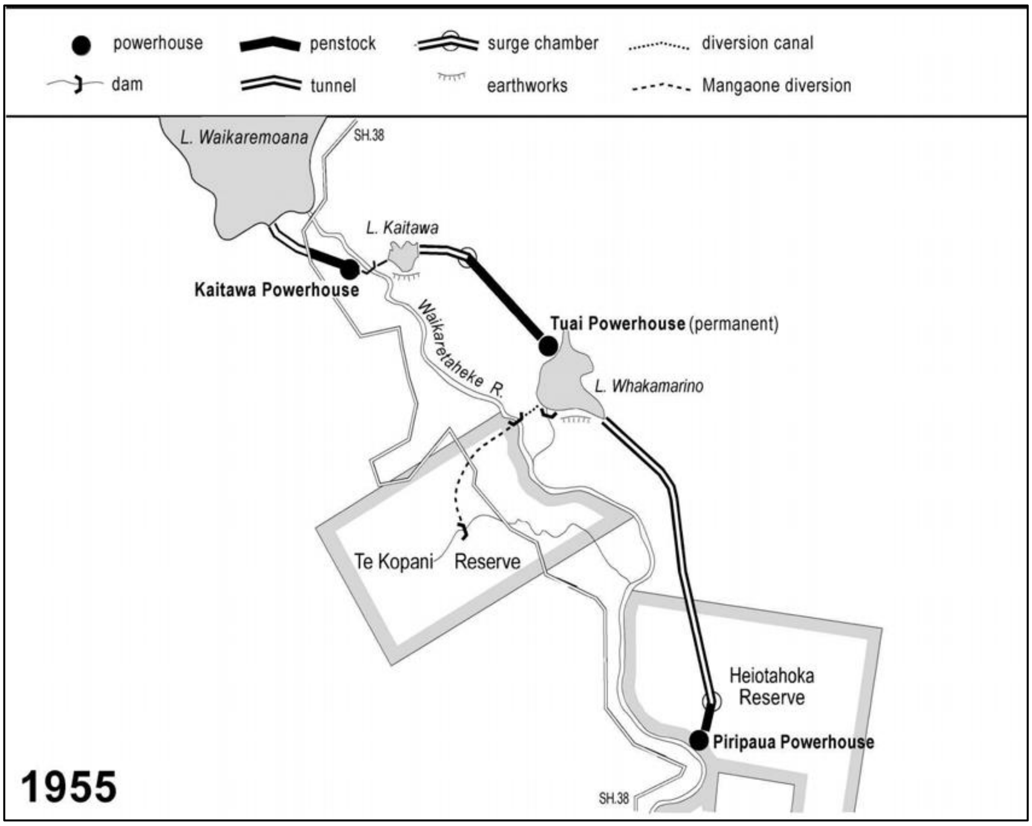

Figure 10-5: Hydropower Development 1955 (Cant et al., 2004)

10.1.2.f Further Developments and Upgrades

While Kaitawa Station was being built, New Zealand's power shortage was an ever-present problem. Power generation from Tuai and Piripaua was limited by natural inflows from Waikaremoana into Lake Kaitawa. To both alleviate this problem and assist with drawing down the lake for the intake development, several temporary water siphons were installed at Te Whara Whara in 1944 which drew water from Waikaremoana, down the upper Waikaretaheke stream (Waikaremoana spillway) and into Lake Kaitawa. Waikaremoana's lake level was reduced by up to 10 metres at times during the 1940s to 1960s. The temporary siphons were later replaced by two permanent siphons to ensure Kaitawa Station could be bypassed if an outage occurred.

Sealing the lake and ensuring leakage flow was instead put through the Kaitawa Station turbines could mean saving approximately 7 GWh per year. Saving this water would also raise the head level for Kaitawa Station, further increasing the power output.

A diver was used to examine the lakebed for leaks. While the Kaitawa Station intake was being prepared, the lake was drawn low which allowed higher leaks to be sealed with materials of similar size to the holes. Leaks below water surface level were closed by dropping material over the mapped holes. Leaks in deeper water needed a temporary jetty and other times a barge was used. Sealing was able to stem flow by approximately 50% from 10 m^3/s to 5 m^3/s. Stability of this sealing could not be guaranteed due to the wave action, sediment filtering and erosion.

Over the many years of operation, several of Waikaremoana's units have received various upgrades. At Tuai, all units have had their turbines overhauled and upgraded with uprated runners in 1986. Units 1 and 2 had their turbines overhauled again in 1994, with their scroll cases being replaced. In terms of their generators, each Tuai unit has had stator rewinds in the 1960s. Tuai Station is as of writing this thesis, is going through a generator rewind project with Unit 3 being the first to be completed in 2022 and the other two units to be completed in 2023 and 2024.

Units 4 and 5 at Piripaua received new Voight turbine runners in the 1996. Both units' generators received new stator cores and rewinds around the same time. Although an overhaul of the turbines may be necessary in the near future, the runners and generators are expected to last many more years.

Kaitawa Station's Units 6 and 7 received new turbine runners in 1991 and Unit 7 received a turbine overhaul in 1998. The current runners are expected to last for several more decades. The generators at Kaitawa Station have lasted for a considerable amount of time without major upgrades but have begun to approach their end of life with replacement due by the end of 2030.

10.2 Efficiency Regression Models

10.2.1 Linear Plots

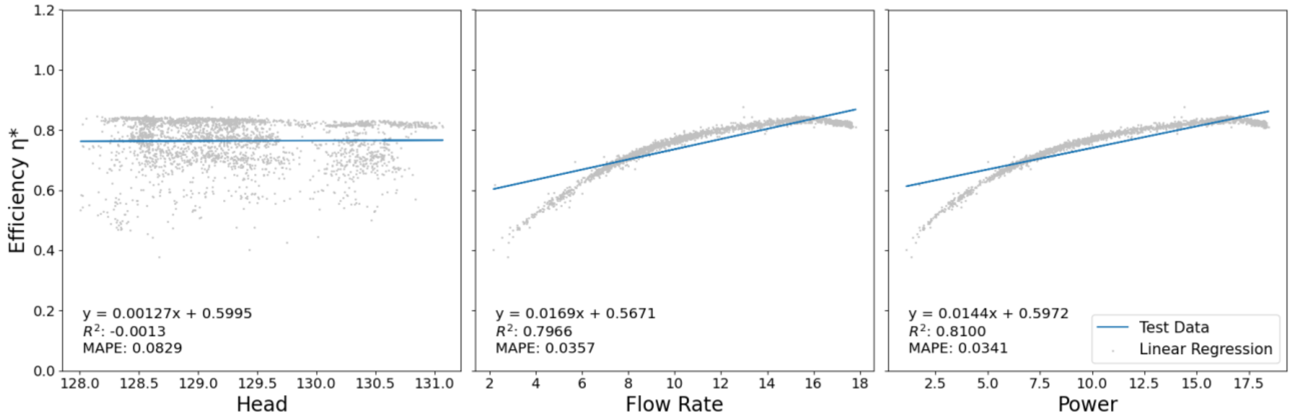

Figure 10-6: Unit 7 Linear Regression Plots Against Validation Data

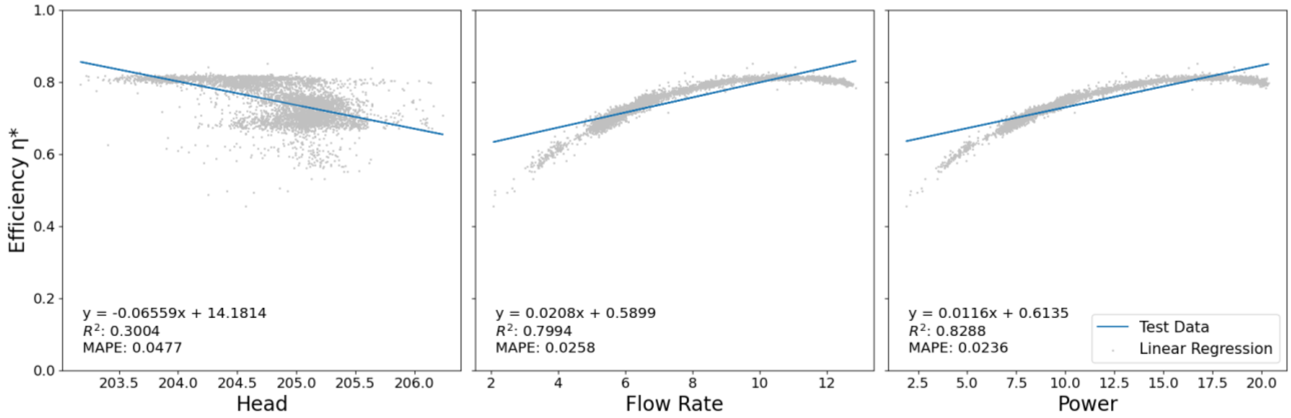

Figure 10-7: Unit 1 Linear Regression Plots Against Validation Data

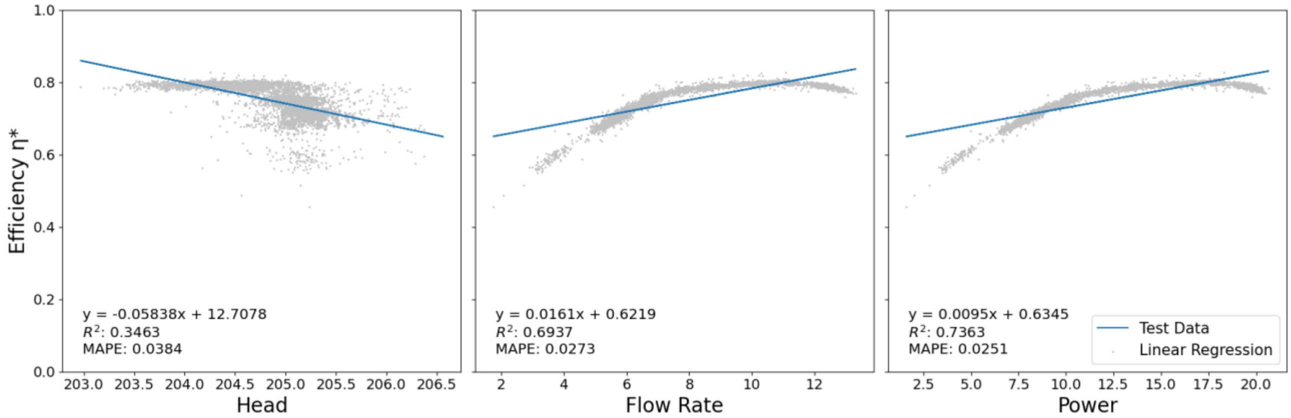

Figure 10-8: Unit 2 Linear Regression Plots Against Validation Data

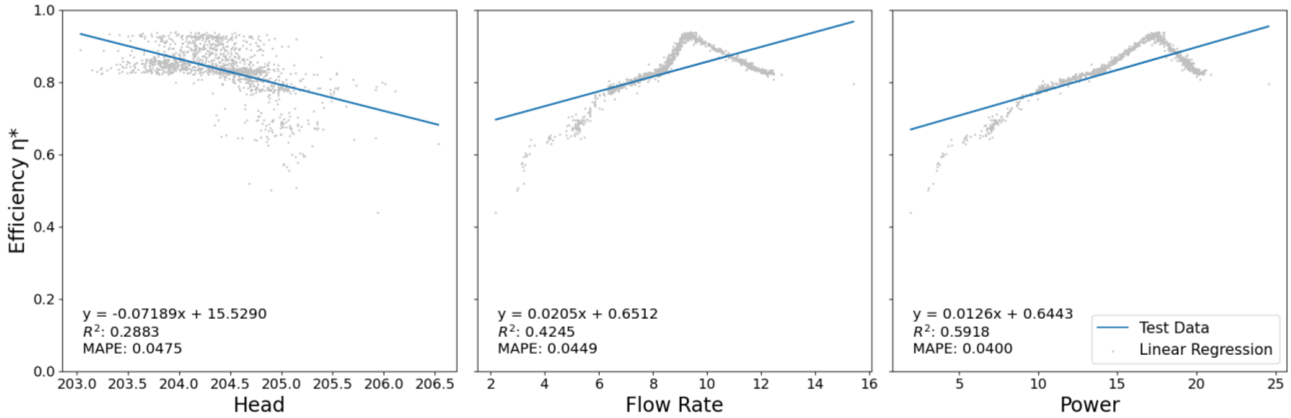

Figure 10-9: Unit 3 Linear Regression Plots Against Validation Data

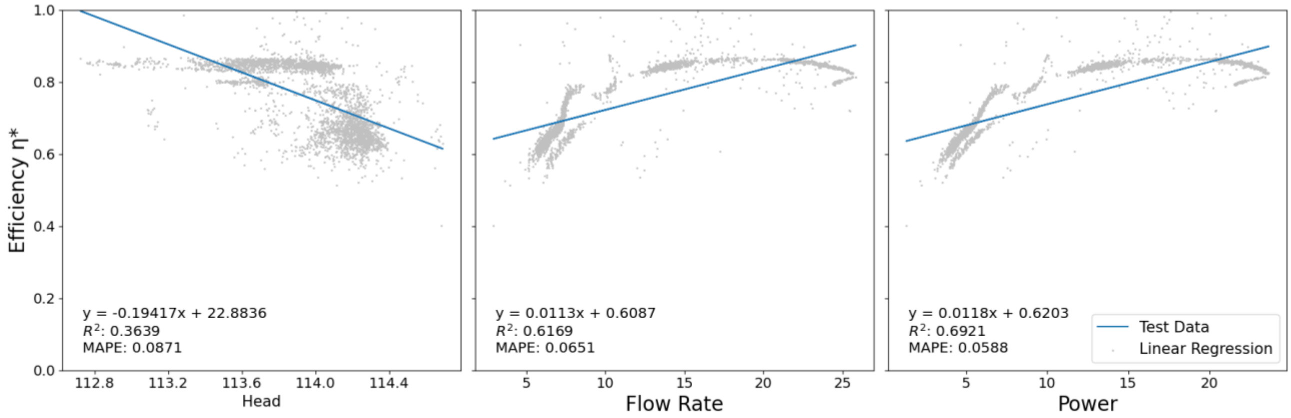

Figure 10-10: Unit 4 Only Linear Regression Plots Against Validation Data

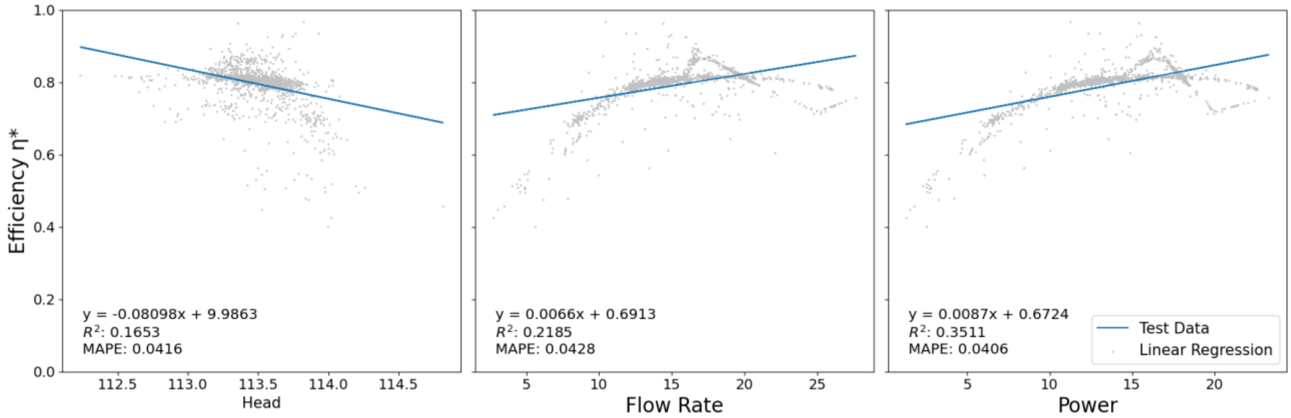

Figure 10-11: Unit 4 Tandem Linear Regression Plots Against Validation Data

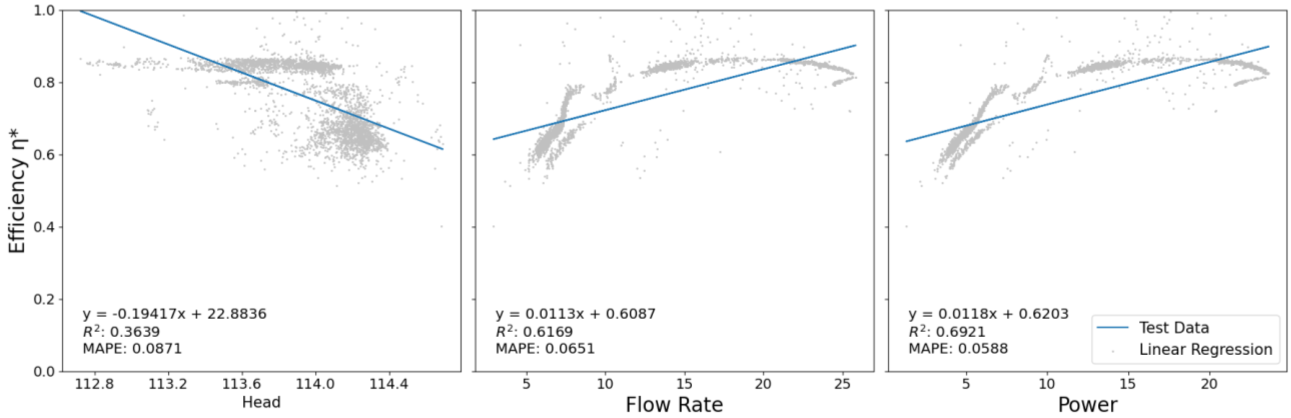

Figure 10-12: Unit 5 Only Linear Regression Plots Against Validation Data

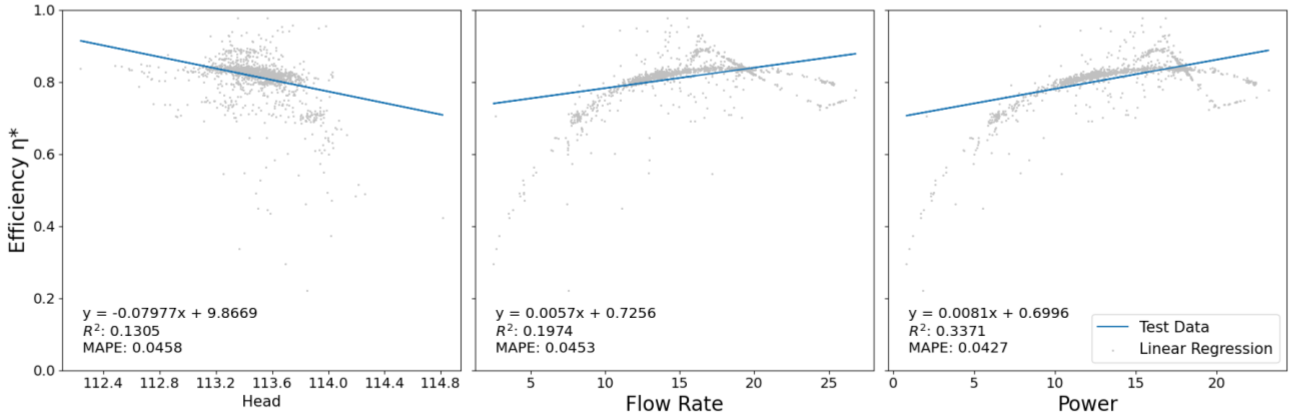

Figure 10-13: Unit 5 Tandem Linear Regression Plots Against Validation Data

10.3 Balance Equations

10.3.1 Head Variation

To check for head variation, flow rate was plotted against levels for each point where level is used in head calculations.

Kaitawa Station variations: Forebay level is measured against Waikaremoana net flow. Tailwater level is measured by discharge flow out of Kaitawa Station.

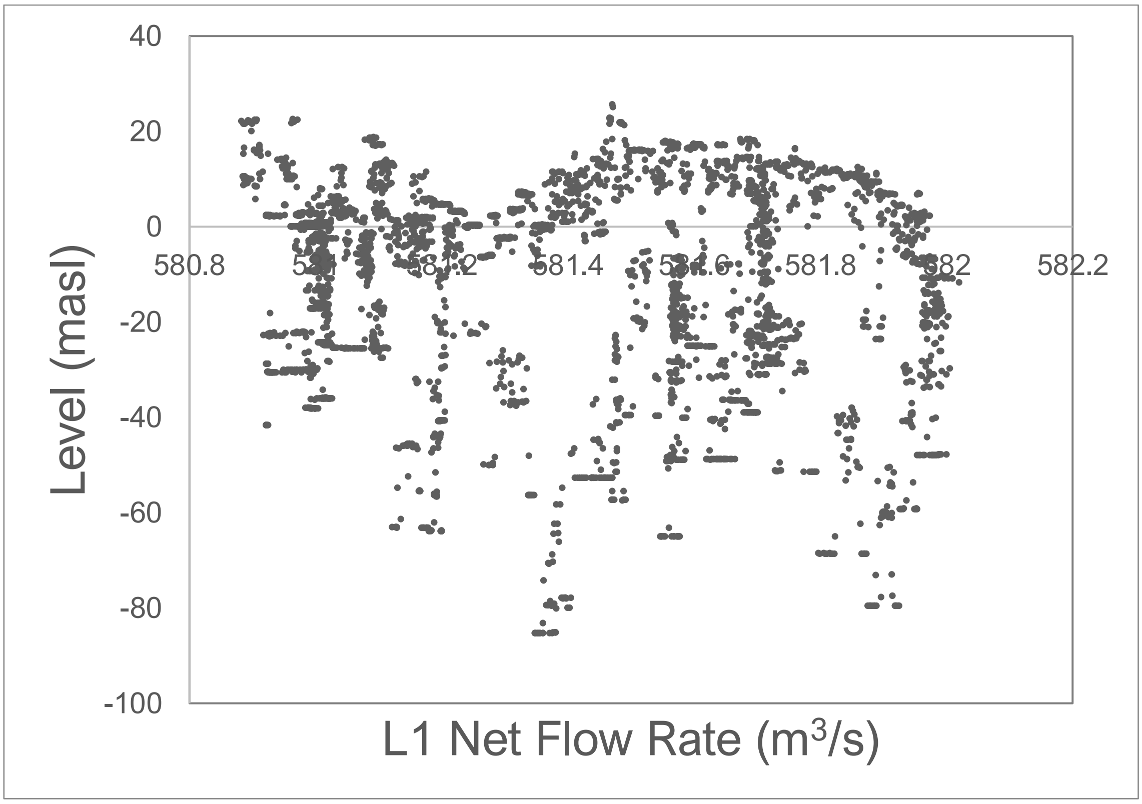

Figure 10-14: Waikaremoana Forebay Level Variation from Flow

Forebay level shows some hint of an abstract pattern but can also be interpreted as random and without a clear relationship.

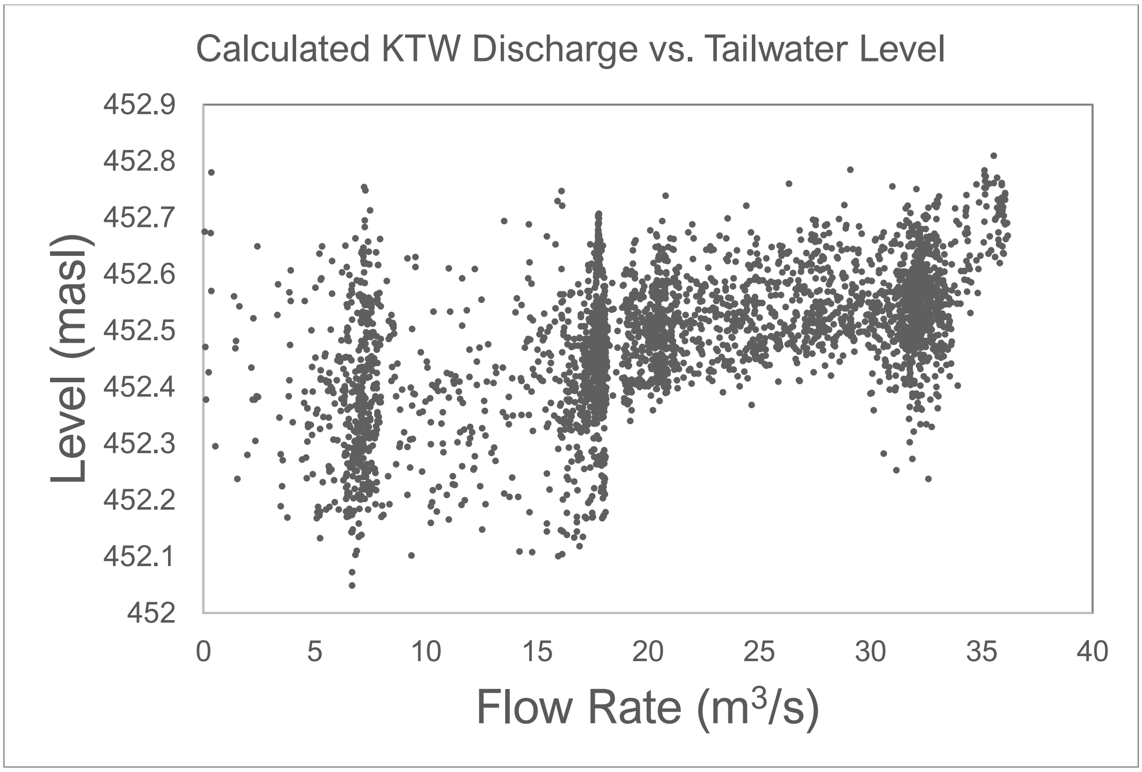

Figure 10-15: L1 Tailwater Variation from Flow

This plot shows a moderate positive gradient relationship between flow and tailwater level.

Tuai Station variations: Forebay level at the surge chamber is measured against Tuai discharge flow. Tailwater level for Lake Whakamarino measured compared against discharge flow out of Tuai Station.

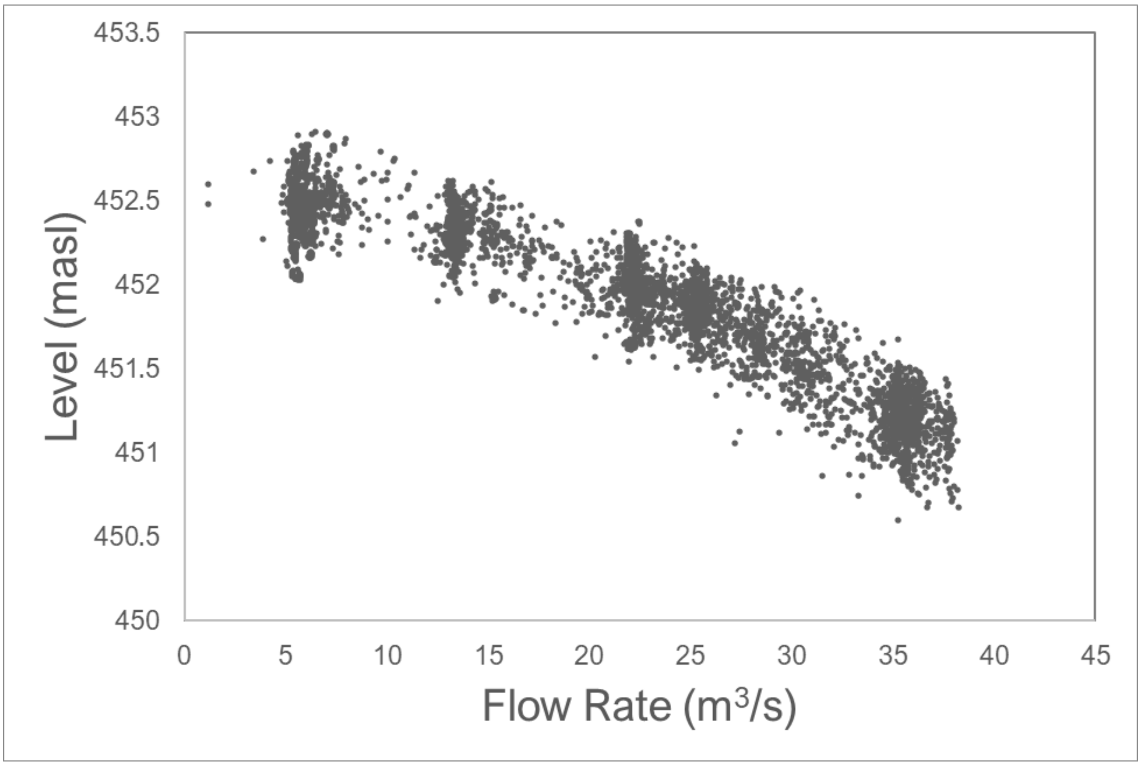

Figure 10-16: Kaitawa Forebay Level Variation with Flowrate

10.4 Hill Diagrams

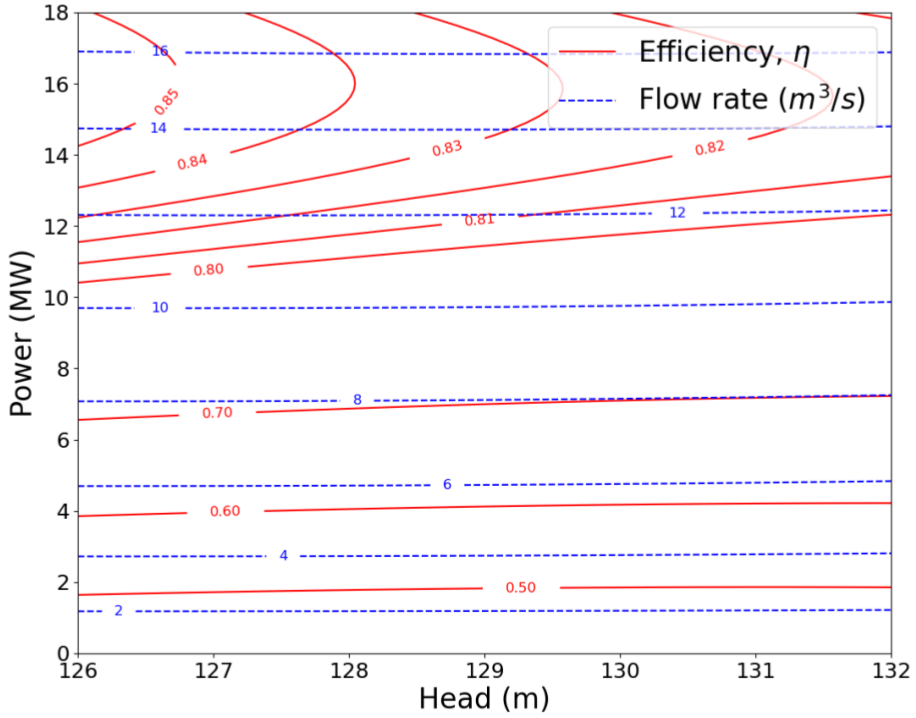

Figure 10-17: Hill Diagram Unit 7

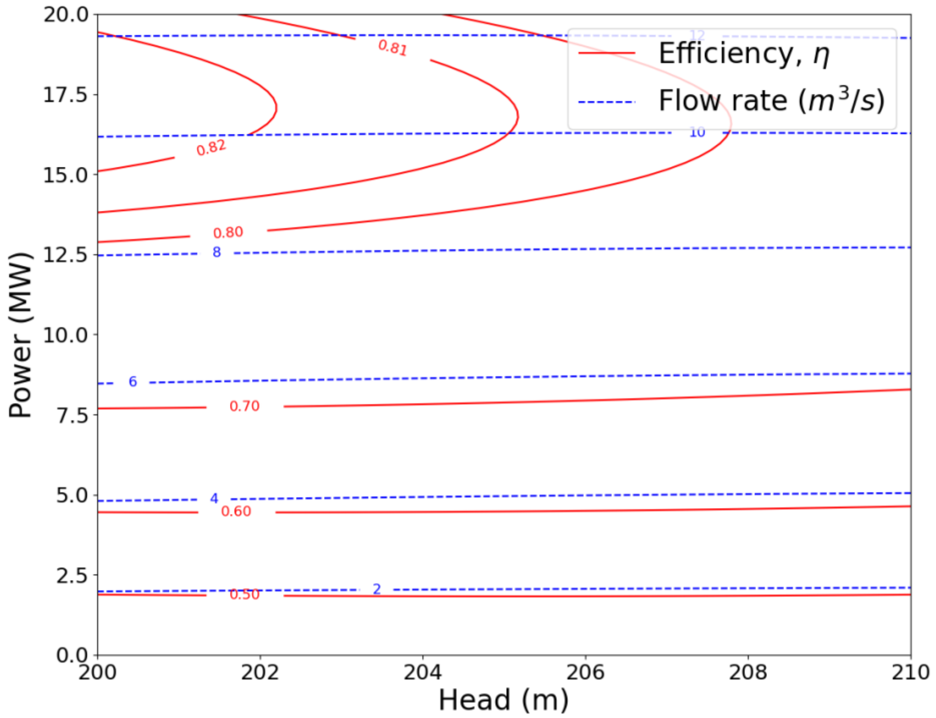

Figure 10-18: Hill Diagram Unit 1

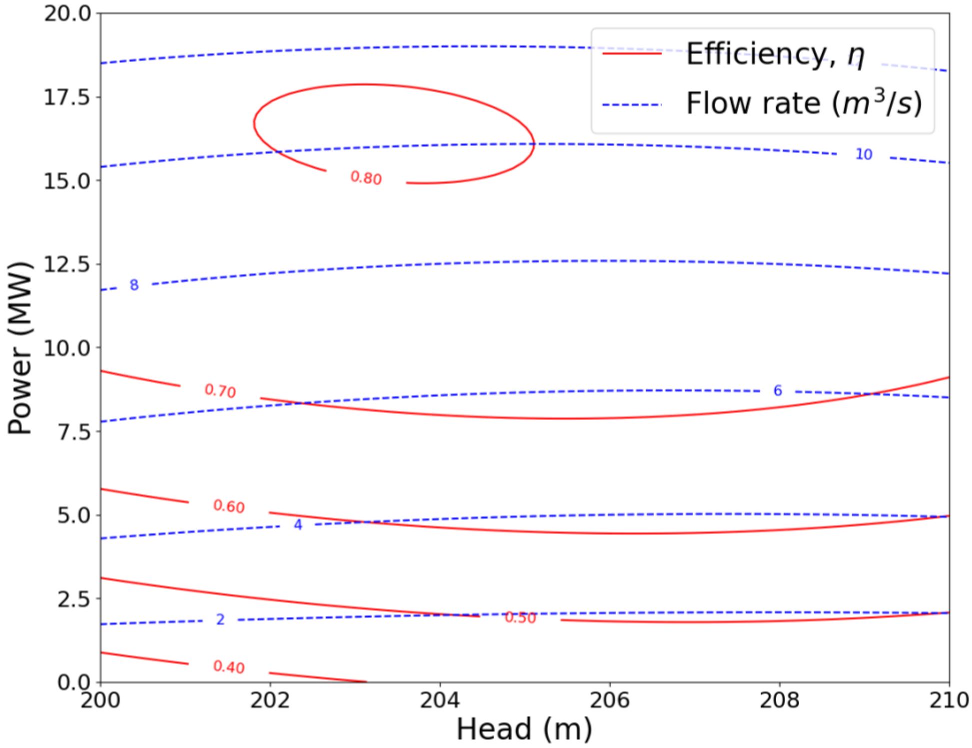

Figure 10-19: Hill Diagram Unit 2

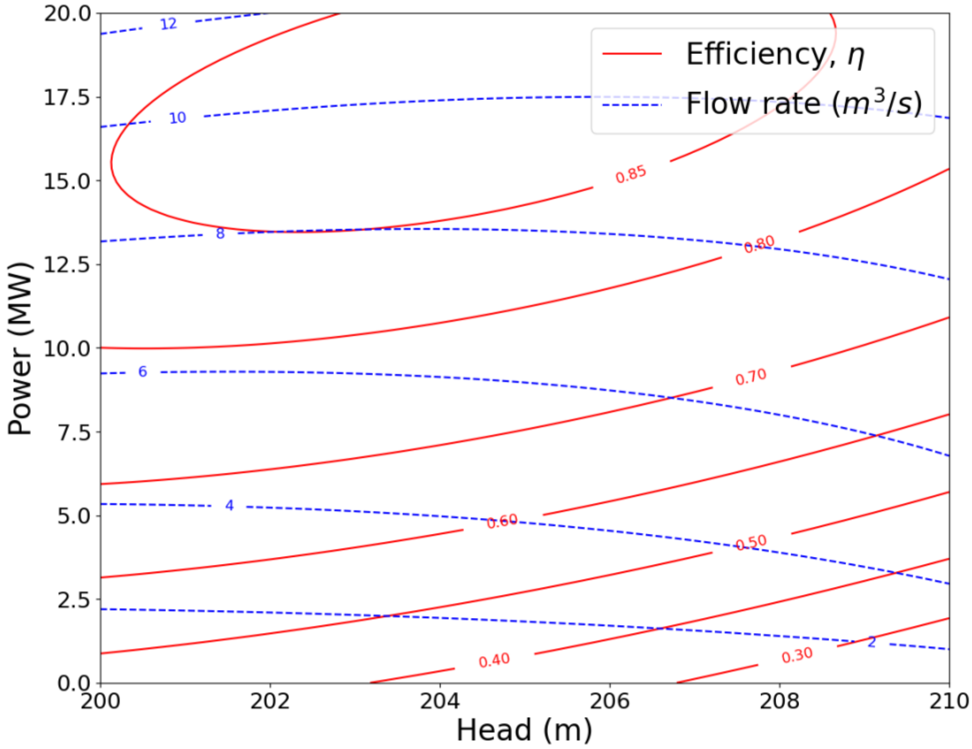

Figure 10-20: Hill Diagram Unit 3