Eindhoven University of Technology

MASTER

# Inverse Modelling and Uncertainty Quantification of a Computational Model of Rotorua

Dekkers, K.

*Award date:*
2021

*Awarding institution:*
University of Auckland

Link to publication

TU/e Department of Mechanical Engineering, Energy Technology
The Geothermal Institute, The University of Auckland

# Inverse Modelling and Uncertainty Quantification of a Computational Model of Rotorua

Graduation Project

**April-February – 2020-2021**

| Name: | ID-number: |
|---|---|
| K. Dekkers | 0863554 |

| Responsible supervisors: | |
|---|---|
| prof.dr.ir. O'Sullivan, M.J. | (UoA) |
| dr.ir. Speetjens, M.F.M. | (TU/e) |
| dr.ir. Gravatt, M.J. | (UoA) |

Auckland, New Zealand

# Abstract

The Rotorua geothermal field (RGF) is a geothermal reservoir located on the North Island of New Zealand. The reservoir has been extensively exploited since the 1950s. Overexploitation caused adverse effects on its geothermal activity. Therefore, the 1986 Bore Closure Programme was implemented to regulate production to recover the geothermal features of the reservoir.

Numerical models of geothermal reservoirs can predict future behaviour and help to prevent adverse effects caused by exploitation. The RGF has been successfully modelled in the past, however, a recently developed new model provides an improved fault structure and an improved numerical accuracy compared to previous models. Moreover, the new model is compatible with the geothermal simulator Waiwera, which reduces computation time for forward simulations and has easy access to the model sensitivity. Reduced computation times are advantageous for manual calibration because this requires many forward simulations.

The Rotorua model has been manually calibrated by matching the model results with the available field data, which improves model accuracy. However, the model predictions are uncertain because of model uncertainties and noise in the measurements. Therefore, quantifying the uncertainty of the model predictions is vital.

This study used the Bayesian framework, which combines the field data and the model sensitivity to reduce the model parameter uncertainty. The reduced parameter uncertainty caused a reduction in the model results and predictions, particularly in the Kuirau Park area.

Furthermore, the production and reinjection of the individual wells at Rotorua are highly uncertain due to a lack of records. Nevertheless, the uncertainty quantification implies that uncertain production and reinjection have negligible effects on the uncertainty of the model predictions.

In conclusion, quantifying the uncertainty of the model predictions provides helpful insight for sustainable reservoir management. The model predictions of the RGF, including their uncertainties, indicate that the current production and reinjection at the RGF is sustainable for the next 20 years. Further calibration will improve model accuracy. By simulating different future scenarios, we can, for example, predict if net production can be increased without causing adverse effects for the reservoir.

# Contents

# 1 Introduction

## 1.1 Background

Nowadays, more renewable energy sources are used all over the world. These energy sources will replace fossil fuels, which are getting scarcer and harm the environment. Therefore, researchers have been looking at geothermal reservoirs as it is both renewable and more eco friendly. These reservoirs are usually located at tectonically and volcanically active zones. The basic idea of a geothermal reservoir is to extract energy, in the form of heat, from the subsurface of the earth. This energy is extracted by producing hot water in its liquid form or steam which flows through a rock matrix, creating a hydrological system. Generally, the energy from hot water is used directly, whereas steam is converted to electricity.

Liquid-dominated geothermal systems were initially thought to be useless and therefore not exploited. However, in the early 1950s, New Zealand started to gain experience in the separation of steam from liquid-dominated mixtures. New Zealand is world-leading in the field of geothermal reservoir engineering, which is useful for the development of their many potentially exploitable geothermal systems, like the Rotorua geothermal field (RGF). Geothermal models can predict future reservoir behaviour, which helps to manage the reservoir in an efficient, sustainable and renewable way. (Grant, 2013)

There are various methods for harnessing geothermal energy, all of them require a form of heat exchange. A few of these methods are (BTES) Borehole Thermal Energy Storage, (ATES) Aquifer Thermal Energy Storage, geothermal heat pumps and (EGS) Enhanced Geothermal Systems. BTES and ATES are mostly used in European countries because these systems work at relatively low temperatures. The BTES and ATES are both used as seasonal thermal storage. These systems extract warm water from the earth in winter, used for heating, and in summer, cold water is extracted, used for cooling (Paksoy & Beyhan, 2014; Reuss, 2015). The difference between these systems is that ATES is an open system and, therefore, requires an open rock structure so water can flow easily, see Figure 1, while BTES, on the other hand, is a closed system and uses bores to store the warm and cold water, see Figure 2.

A geothermal heat pump indirectly uses geothermal energy. In a closed system, water is heated below the ground, and by means of a heat exchanger, the energy is then transferred to a domestic heat distribution system as can be seen in Figure 3 (Davidsdottir, 2012). The last method to harness geothermal energy is an EGS. These systems enhance the permeability structure through hydraulic, thermal or chemical stimulation (Huenges, 2016). Therefore, an EGS can be used in areas with low permeability and/or low water presence. Nevertheless, there are no successful EGS projects yet, thus further research is required.

The RGF has production bores, injection bores and downhole heat exchangers. These bores and heat exchanger can be categorised as a combination of the ATES and geothermal heat pumps. However, in contrast to the ATES system, the RGF has a constant flow through the aquifer and is not used as thermal storage. The hot water extracted from production wells is used directly in spa pools, and indirectly for industrial and domestic heating using heat exchangers. Additionally, geothermal heat pumps are used for domestic heating.

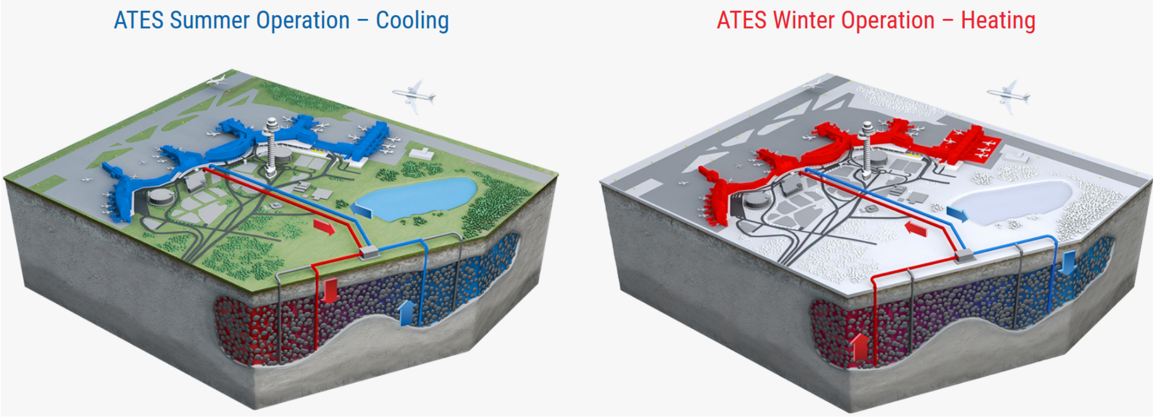


Figure 1: A schematic picture of an ATES system (?).

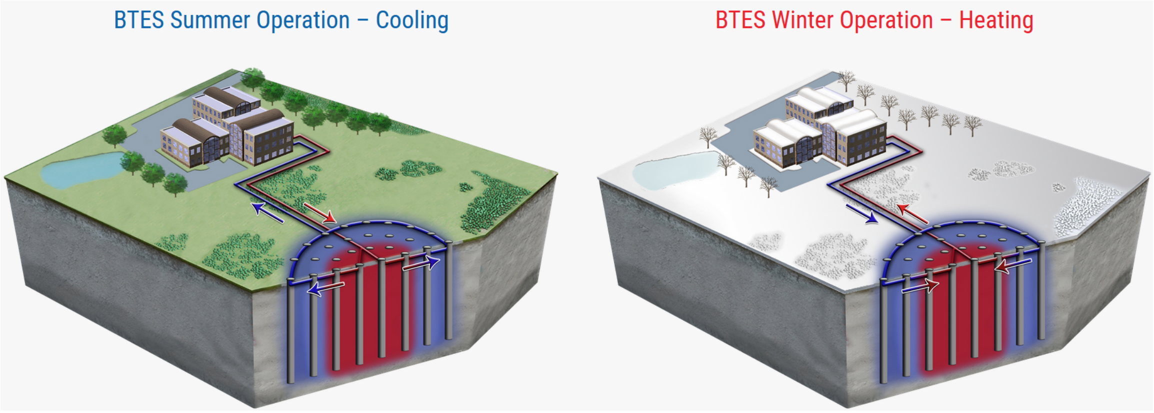


Figure 2: A schematic picture of a BTES system (?).

## 1.2 Rotorua geothermal field RGF

The RGF field is located in the North Island of New Zealand, in the urban area of Rotorua. European settlers have exploited the geothermal field since the 1800s and by the indigenous Māori people for many years before that (Ratouis et al., 2016b). Apart from an energy source, the RGF is a tourist attraction (Butcher et al., 2000), moreover, according to Neilson et al. (2010) it is of great importance for the Māori culture, and Acland & Molloy (2006) state the significance of the extensive biodiversity. Since the 1950s the energy demand has increased enormously, resulting in more than 900 wells being drilled by 1985, many of which were for personal use (Ministry of Energy, 1985). Back at that time, there was little knowledge and experience on the use of geothermal reservoirs. Therefore, without

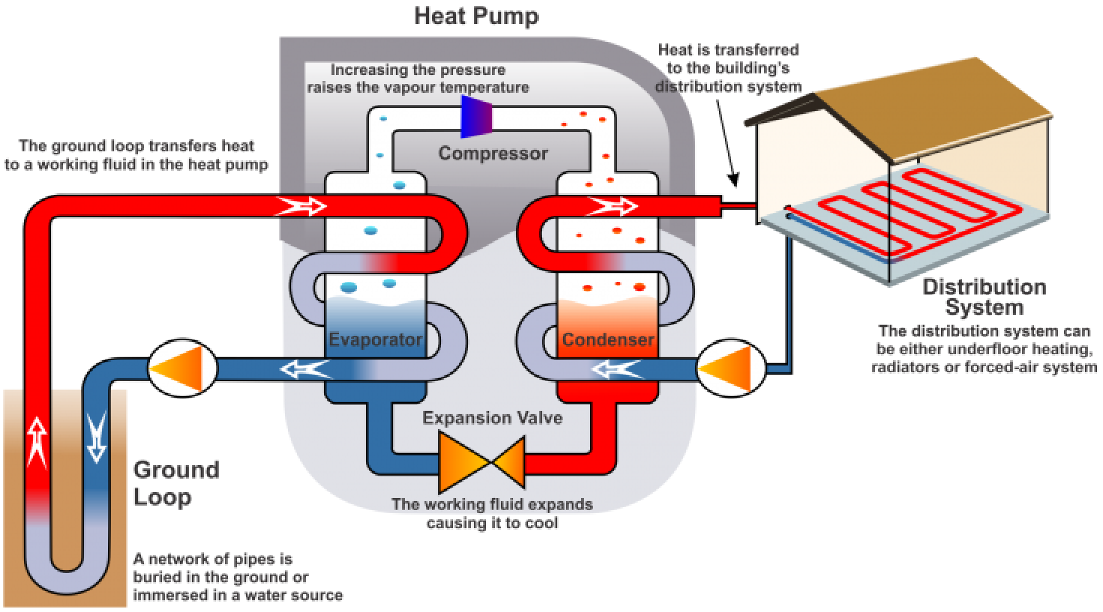


Figure 3: A schematic picture of a geothermal heat pump system (New Zealand Geothermal Association, 2021).

any regulation, the hydrothermal activity at the surface of Rotorua came to a dangerously low level (Lloyd, 1979). Scientists, who were involved in monitoring the Rotorua reservoir, advised the New Zealand government to regulate hot water extraction from bores and wells. The government acknowledged the deterioration of the geothermal field, and in 1986 the Bore Closure Programme was implemented, which meant that all bores within a radius of 1.5 km of Pohutu Geyser at Whakarewarewa had to be closed, see Figure 5. In addition, Gordon et al. (2005) encouraged the reinjection of water into the reservoir to maintain pressure and provide recharge, thus supporting the field's recovery. Figure 4 shows a graph of the production history, indicating a clear drop in production and an increase in reinjection after implementing the 1986 Bore Closure Programme.

## 1.3 Developing a model

Geothermal reservoirs contain high enthalpy, two-phase flow mixtures and a complex geological structure. Furthermore, there are limited available measurements, which means that not every detail of the reservoir's 3D structure is known. The complexity of the reservoir and the limited available data make it challenging to develop an accurate model (Ratouis et al., 2016b).

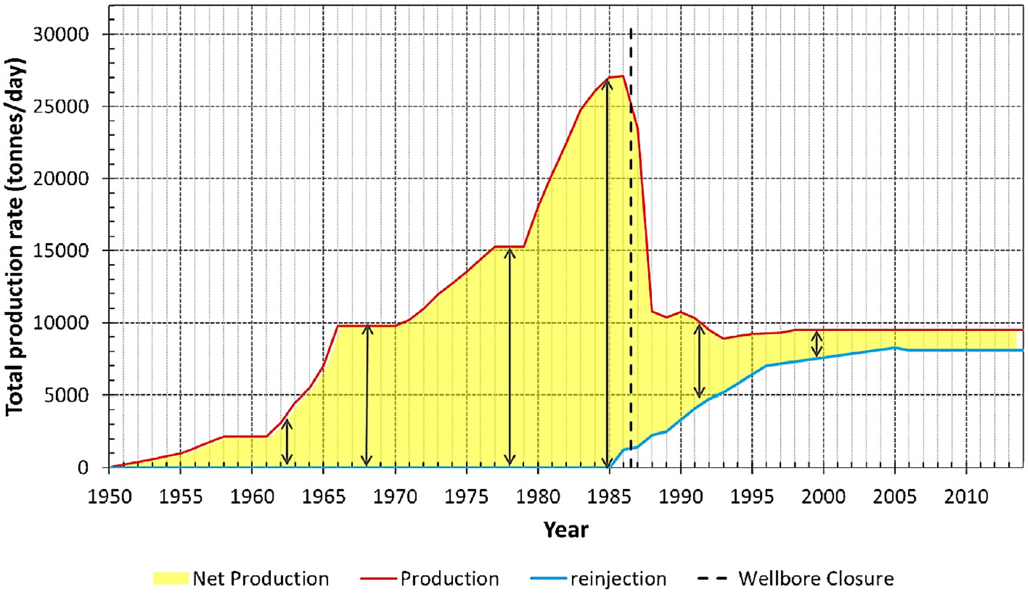


Figure 4: The cumulative estimated production and reinjection rates at the Rotorua geothermal field from 1950-2014 (Ratouis et al., 2016b).

Past studies have shown that temperature measurements can create a better understanding of geological structures. Anderson (2005) and Saar (2011) used temperature measurements to improve the understanding of the geological structures from groundwater systems. Other researchers (e.g., Woodbury & Smith (1988); Bravo et al. (2002); Jiang & Woodbury (2006); Rath et al. (2006); Vogt et al. (2014); Zhang et al. (2014)) applied inverse modelling using water and temperature distribution data. However, all these past studies were only applied to single-phase systems. Whereas geothermal reservoirs often deal with high-enthalpy systems and possibly a two-phase flow of water and steam, which creates a highly nonlinear forward problem making it considerably more complex for inverse modelling.

## 1.4 Modelling aims

The main goal of geothermal reservoir modelling is to make predictions of future behaviour. Therefore, we need an accurate model that represents the actual reservoir. The model can be divided into a natural state (which is the steady state of the reservoir pre-production), the production history (the reservoir since the start of production until now) and future scenarios (predictions on future behaviour).

### 1.4.1 Natural state

In a geothermal reservoir, a convective heat plume originates from the heat source deep in the earth. The heat plume rises above this heat source. The development and shape of this

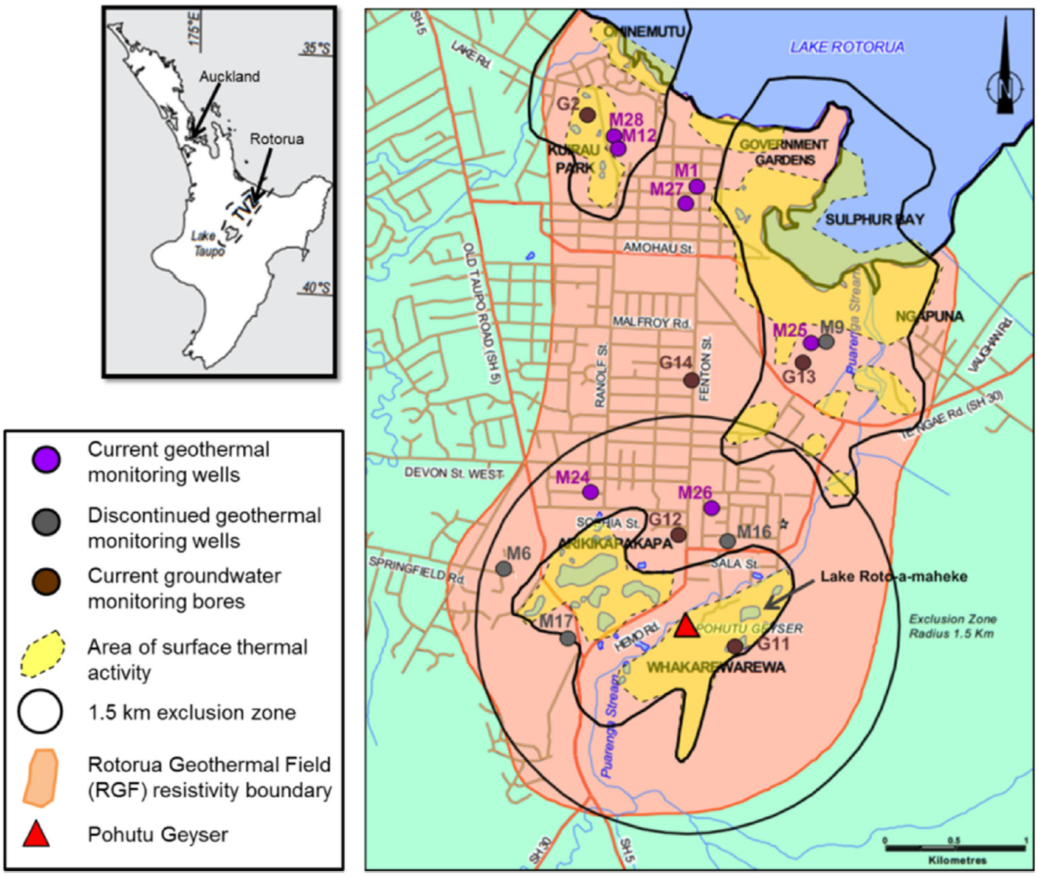


Figure 5: A map of the geothermal field at Rotorua (Ratouis et al., 2016b).

convective heat plume, seen in Figure 7, depends on the geological structure of the reservoir (e.g., faults, formation boundaries, caprock).

The natural state of a geothermal reservoir is typically referred to as the pre-production state of the reservoir. A natural state model with arbitrary initial conditions will be simulated until steady state conditions are reached. A transient approach is used, and state variables are updated with every time-step. The time-step is increased every time the change in the model results is low. When a large enough time-step, say about 30 million years, has been reached, one can assume that the result is a good approximation of the steady state conditions (Bjarkason et al., 2019).

### 1.4.2 Production history

The final state of the natural state simulation is used as the initial state of the production history model. The production history model is a transient model, which means the model changes over time due to production and reinjection. The RGF has many wells used either

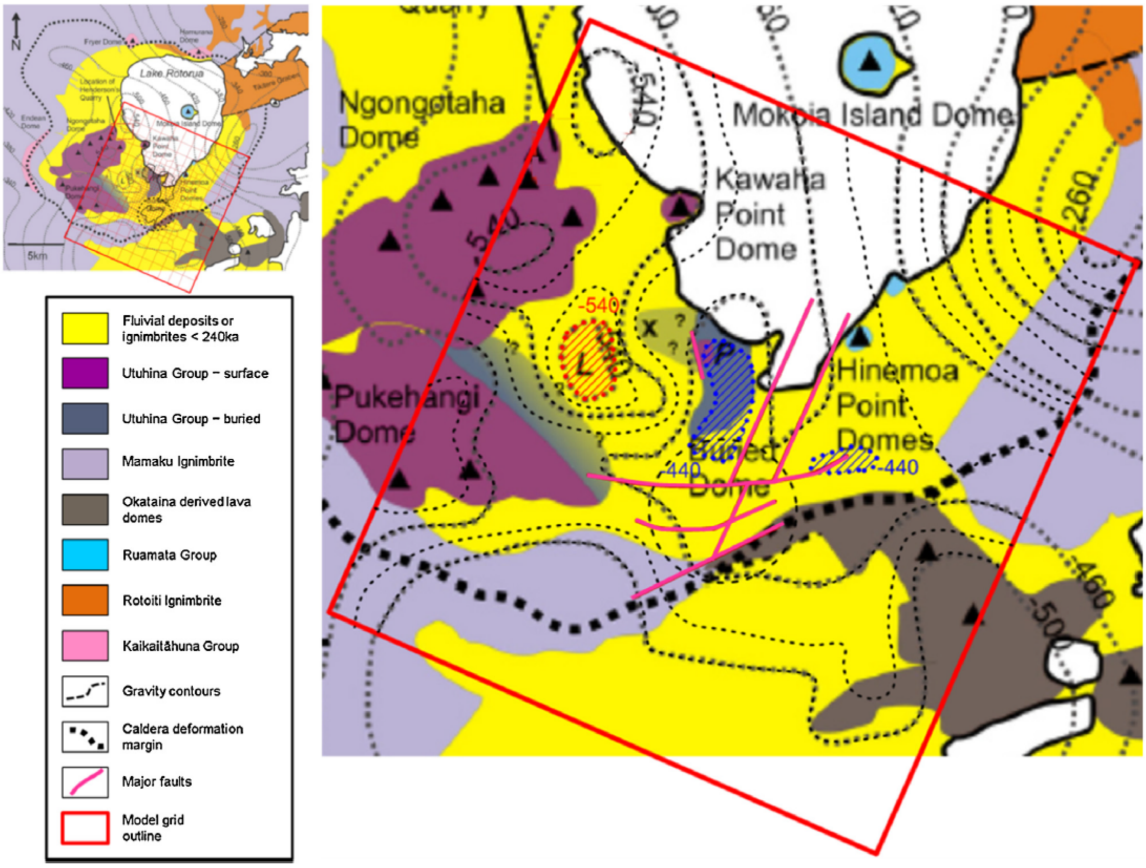


Figure 6: A geological map of the geothermal field at Rotorua used to create a numerical model (Ratouis et al., 2016b).

domestically or commercially. Typically for geothermal reservoirs, the production and reinjection flow rates of the wells are well known. However, in the case of Rotorua, only the total production and reinjection rates of individual wells were reliably estimated, and the flow rates of the individual wells are only rough estimates.

### 1.4.3 Future scenarios

The basis for accurate model predictions is an accurate model (natural state and production history). Inverse modelling or model calibration is the process of creating an accurate model, explained in Section 1.5 and 3.3. Based on the calibrated model, future scenarios can be built, simulated and analysed. However, due to many sources of uncertainty in the model and measured field data, the predictions will be uncertain. Therefore, quantifying model and prediction uncertainty is vital to improve informed decision making on reservoir management.

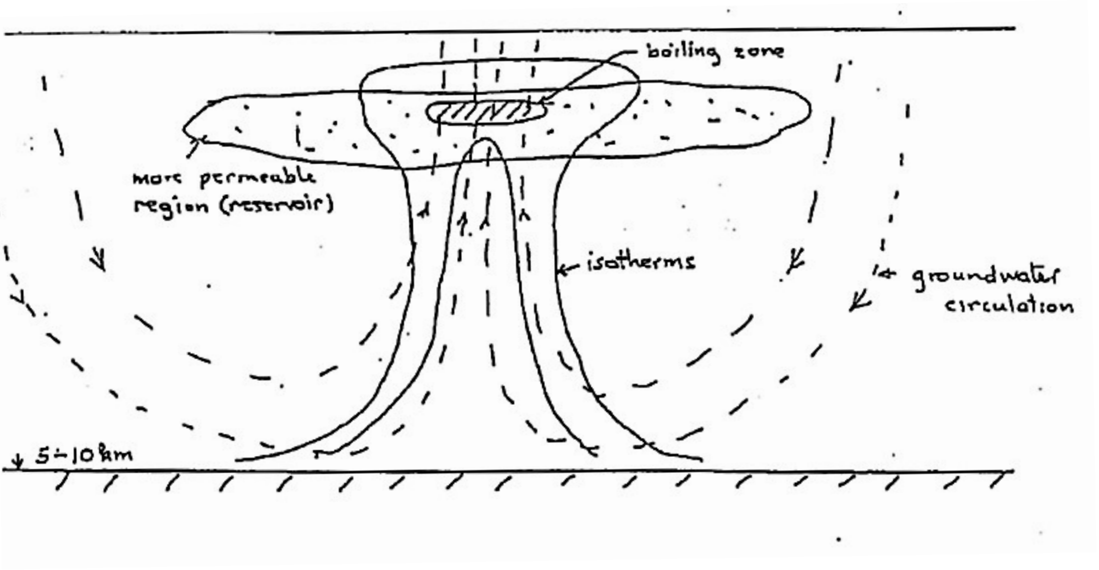


Figure 7: A schematic drawing of a convective geothermal system at natural state, showing the flow of water (O'Sullivan, 2009)

## 1.5 Inverse modelling

The idea of inverse modelling, or calibration, is to match the model results with the measured field data. The first step of the calibration process is manual calibration and typically followed by automatic calibration. Well-known automatic calibration software packages are iTOUGH2, AUiTOUGH2 (based on iTOUGH2) and PEST. Recently, a promising new software, iWaiwera, has been in development at the Geothermal Institute of the University of Auckland. The main advantage of this calibration software is the reduction of computation time (Gonzalez-Gutierrez et al., 2020). Currently, however, iWaiwera only works for specific geothermal models and, in the time-scope of this study, application on the Rotorua model was not accomplished.

## 1.6 Uncertainty quantification

As mentioned before, model predictions improve informed decision making on reservoir management. However, these predictions contain many sources of uncertainty. Errors made in the model conceptualisation has the most significant impact on model accuracy and predictions. Furthermore, model parameters (e.g. permeabilities and deep hot upflows) are uncertain due to the limited availability of field data, and the field data in itself contain measurement errors. Lastly, finite numerical accuracy is another source of uncertainty (Finsterle, 2007). Therefore, quantifying the uncertainty is crucial, giving more comprehensive results which provide further guidance for reservoir management. In earlier studies, the uncertainty has been analysed of inverse problems in general (S. E. Kaipio J.P., 2006; Stuart,

2010) and specifically on geothermal inverse problems (Maclaren et al., 2016, 2020; Vogt et al., 2010). The uncertainty quantification of the Rotorua model is unique due to its high dimensionality (94,701 blocks) and high uncertainty, particularly in the production history data. The methods used to quantify uncertainty will be explained in Section 3.5-3.7 and the uncertainty results will be analysed in Chapter 5.

The next chapter outlines how the geothermal model was developed. Then in Chapter 3 the methods for model simulation, calibration and uncertainty quantification are explained. Followed by, the manual calibration results and the uncertainty analysis in Chapter 4 and 5, respectively.

# 2 Modelling

The previous chapter concludes that model predictions, including uncertainty quantification, is vital for sustainable management of a geothermal energy source. Therefore, ongoing research is required, and the computer model needs continuous updating to improve model accuracy. The RGF has been modelled successfully before (Ratouis et al., 2014, 2016a,b, 2017). In these earlier studies predictions were based only on the calibrated model. In the present study, a new rectangular grid was implemented, consisting of double the number of blocks. The new rectangular grid is compatible with the relatively new software Waiwera, which has the advantage of shorter computation times and easy access to the model sensitivity. Reduced computation times make it more feasible to run numerous sample models. Easy access to the model sensitivity allows reducing parameter uncertainty. Both are beneficial for the uncertainty analysis, further explained in Chapter 5. The sections below describe the steps proceeding from geothermal field observations to a numerical model.

## 2.1 Conceptual model

The development of a conceptual model starts with observations from the geological structure, geophysical processes, geochemistry and reservoir monitoring (Ratouis et al., 2016b). These observations are combined to provide an understanding of the behaviour of the reservoir. The geological structure is derived from the geological setting and ground measurements.

The RGF is located in the Rotorua rhyolitic centre, part of a volcano-tectonic depression called the Taupo Volcanic Zone (TVZ) (Ratouis et al., 2016b). The volcanic activity is the energy source, and the geological setting contains information about the rock structure. Geochemistry data indicate flow patterns and can be used to monitor the response of production and reinjection strategies. Furthermore, geochemistry data give insights into the quality and sustainability of the reservoir. A network of monitoring bores (M-wells and G-bores) measures temperatures and water levels (water level deviations are related to reservoir pressure changes). Temperature and water level observations imply specific heat flow patterns and the thermodynamic state of the reservoir. Monitor wells are used to observe the effect of changes in production and reinjection rates (Ratouis et al., 2016b). Scientists and engineers worked together to develop a conceptual model. In Figure 8 a graphical representation of this conceptual model can be observed, including the geological structure, geochemistry and geophysical processes.

## 2.2 Numerical model

The conceptual model is transferred to a 3D Leapfrog model (more information about Leapfrog can be found in Appendix D.1). This 3D model contains information on the topography and rock structure and can be converted onto a grid to create a numerical model. Numerical models are mainly used to predict future reservoir behaviour, however, numerical models can also predict state variables at locations that are difficult or expensive

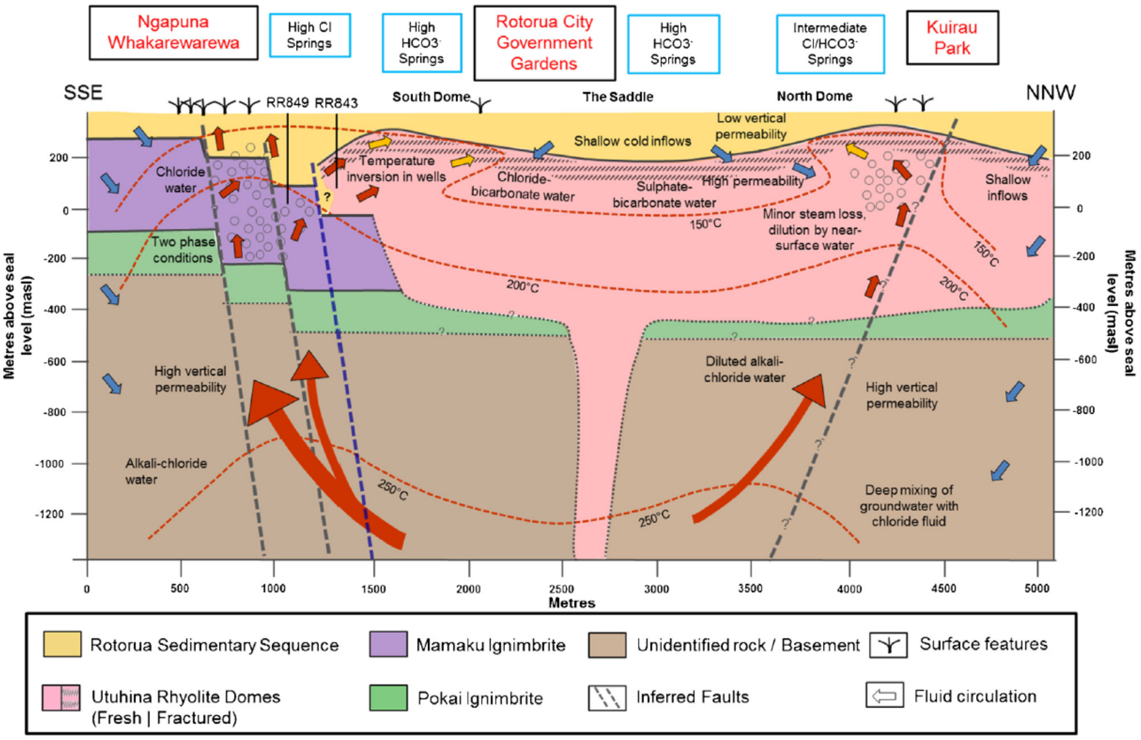


Figure 8: A schematic image of the RGF (vertical slice from SSE-NNW), illustrating the main geological structures, processes and temperatures (Ratouis et al., 2016b).

to measure. Software like TOUGH2 or Waiwera is used to simulate fluid flows and heat transfer to make these predictions.

Figure 9 shows a combination of the numerical grid and the 3D Leapfrog model of the RGF. There are some substantial improvements in the new model compared to the old model, namely;

- The new rectangular grid, making it compatible with Waiwera. Waiwera can simulate the model on multiple cores in parallel, speeding up computation time, and easy access to the model sensitivity, improving the uncertainty quantification.
- The number of blocks has roughly doubled (94,701 vs 48,034), improving numerical accuracy.
- The fault structure has been improved and is a much better representation of the geology, see Figure 12.
- The clay cap has been implemented in the model as a new rock type, making the model more realistic. The clay cap stretches from -700 m to 290 m and covers most of the geothermally active area, see Figure 10 and 11. Nonetheless, due to high

surface activity observations, the clay cap is assumed to be ”leaky”, meaning that the permeability of this clay cap is reasonably high $\mathcal{O}(10^{-15})$ to $\mathcal{O}(10^{-12})$. Typical permeability values for the Rotorua model vary between $10^{-16} - 10^{-10}\text{m}^2$, where a low permeability is correlated with a high flow resistivity and a high permeability is correlated with a low flow resistivity

Van Vlijmen (2020) transferred model parameters like rock properties and deep upflow sources from the old grid to the new grid, and did some manual calibration. Nevertheless, due to the short time-scope of his internship, he did not achieve a reasonable temperature match between the model and the data. The AUTOUGH2 model has been converted with Python to a Waiwera compatible format. The use of Waiwera speeds up the manual calibration process and improves the uncertainty quantification process. The latest natural state version of Van Vlijmen (2020) is referred to in this report as the initial model.

Table 1: Specifications of the numerical model.

| | |
|---|---|
| Area | 18.3 km x 12.4 km (l x w) |
| Blocks | 94,701 |
| Rock types | 340 |
| Smallest block size | 125 m x 125 m x 5 m (l x w x h) |
| Largest block size | 500 m x 500 m x 200 m (l x w x h) |
| Number of layers | 45 |
| Top layer elevation | 750 m |
| Bottom layer elevation | -1500 m |

The numerical model has to comply with governing equations like conservation of mass, conservation of energy and Darcy’s law (fluid flow through porous media). The governing equations will be discussed in Section 3.1. In the next section, the model constraining boundary conditions are described.

## 2.3 Boundary conditions

The boundary conditions of the numerical model can be divided into the side, bottom and top boundary conditions:

- The side boundaries are set far from the active system and can therefore be assumed to be closed (meaning no heat or mass in/out-flow).
- The bottom boundary has a constant conductive heat flow of 80 mW/m$^2$. Furthermore, high enthalpy water and $CO_2$ flow upwards through the faults at the bottom layer. Figure 12b shows the bottom layer of the numerical grid including fault structure. The deep upflow sources are located at the faults and have a total hot water inflow of 64,200 tonnes per day by Scott & Cody (1997). The inflow is divided over three main areas, see Table 2.

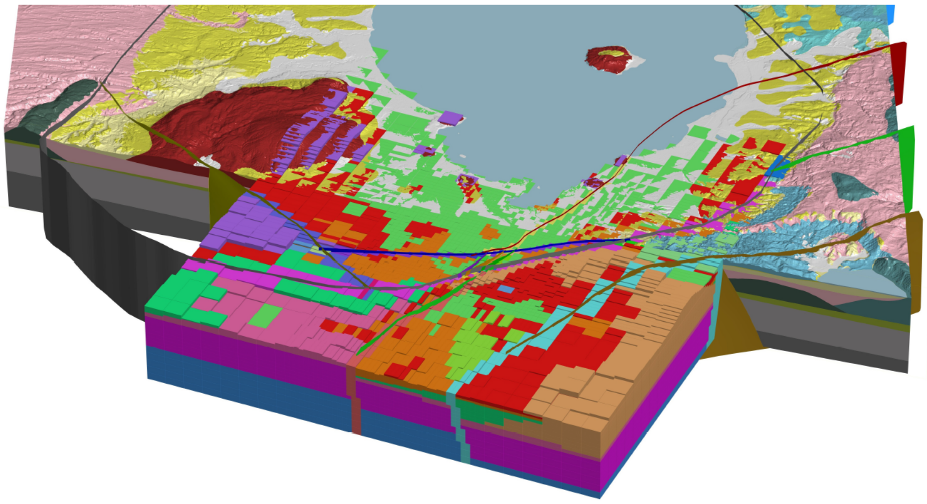


Figure 9: A slice of the the Rotorua geothermal reservoir. Both the numerical grid and the 3D geological model are included (Van Vlijmen, 2020).

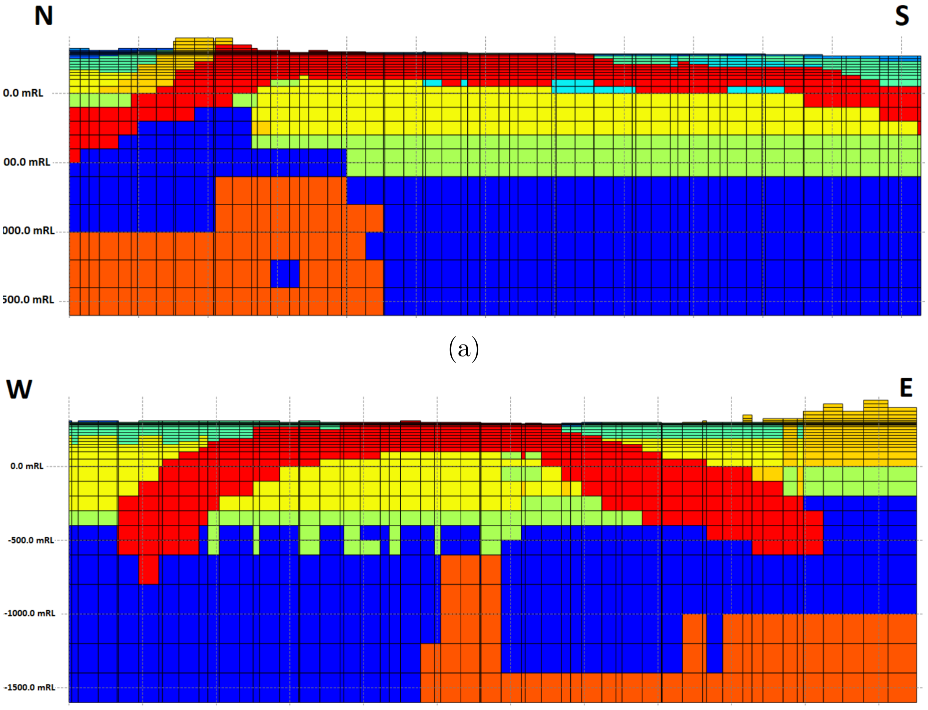
- The top surface is the atmosphere, and therefore a Dirichlet boundary condition is applied (1 bar, 15 °C). This condition also allows for fluid to flow in and out of the system (Ratouis et al., 2016a).

Constant rainfall is assumed at the top surface to reduce model complexity. Moreover, hydrostatic pressure is applied at the lake surface, and a lake temperature of 10 °C is assumed. The lake follows the bathymetry inferred by the International Lake Environment Committee Foundation (ILEC).

The yearly production and reinjection rates from 1950 to 2020 have been reliably estimated, with reasonable estimates for specific areas. Nevertheless, there are no good records for the individual wells, which adds to the uncertainty of the model. Rough estimates of the individual wells have been implemented in the production history model as time-dependent production and reinjection rates (Ratouis et al., 2016b).

Table 2: Hot water upflow at the three main faults at the bottom layer (-1500 m).

| Area | Mass [tonnes/day] | Temp [°C] |
|---|---|---|
| Kuirau Park | 7,100 | 225 |
| Ngapuna Stream | 28,400 | 165 |
| Whakarewarewa | 28,700 | 190 |
| Total | 64,200 | |



(a)

(b)

Figure 10: Two vertical slices, (a) from north-south and (b) east-west, of the numerical grid with the main rock types in different colours. The clay cap is indicated in red and stretches across a large part of the geothermal field.

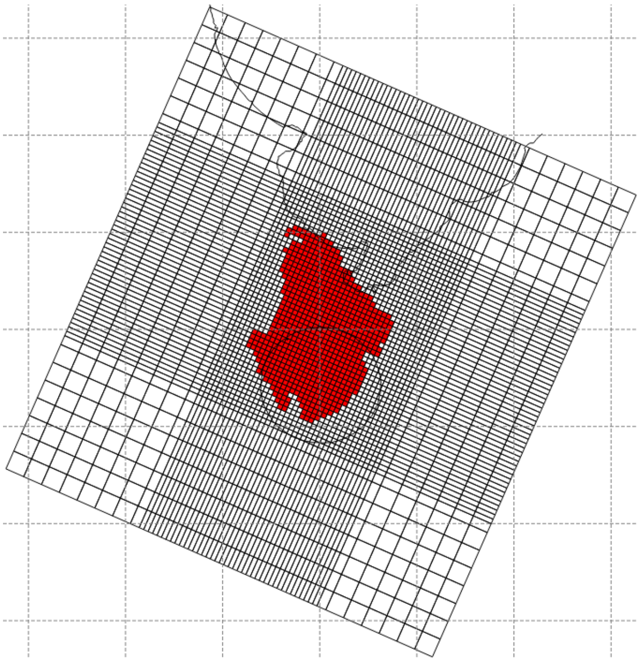
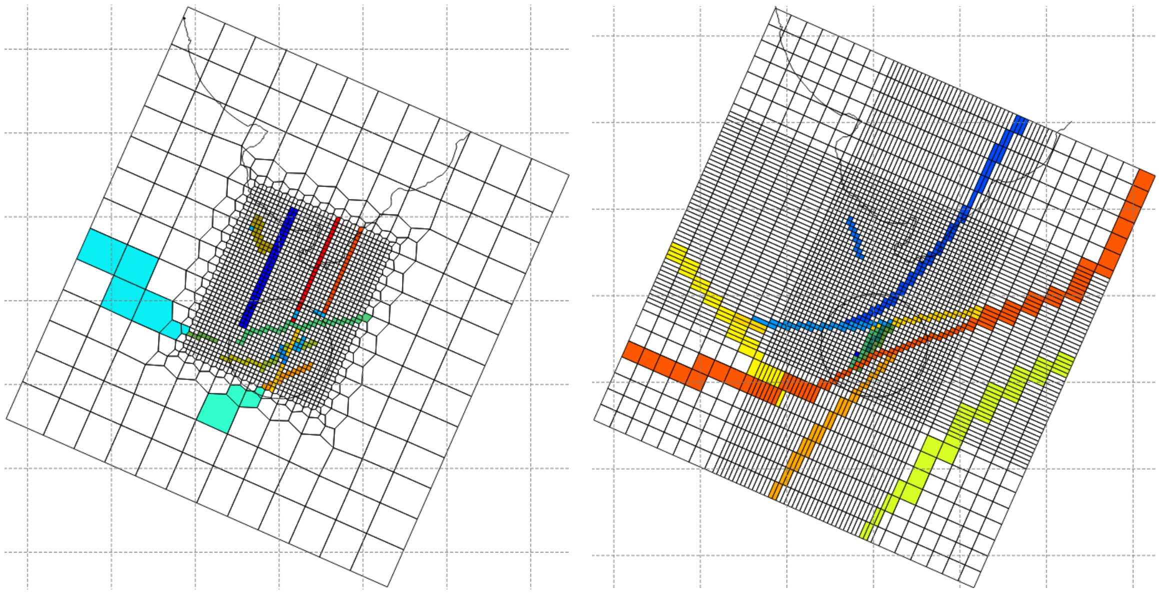


Figure 11: The numerical grid at -180 m, the clay cap is indicated in red, and the exclusion zone and lake are drawn in black lines for reference.



(a) (b)

Figure 12: A comparison between the old (a) and new grid (b) including fault structures. The different fault lines are indicated in colour and the exclusion zone and lake are marked for reference.

# 3 Methodology

In this study, the relatively new software Waiwera (Croucher, 2020) is used to simulate the forward model. Waiwera is a geothermal simulation package developed at the Geothermal Institute of the University of Auckland, based on TOUGH2 principles (Pruess et al., 1999; Pruess, 2004). Although TOUGH2 is more widely used for reservoir modelling, Waiwera has the advantage of solving the forward model in parallel on multiple cores, decreasing computation time. Another major advantage of using Waiwera is easy access to the model sensitivity, which will be used in the parameter uncertainty reduction in Section 3.5. The files required for Waiwera simulations and more on how Waiwera works is explained in Appendix A.1 and B, and the Waiwera manual (Croucher, 2020).

## 3.1 Conservation laws

Waiwera and TOUGH2 simulate the forward model by solving the conservation of mass, energy and momentum equations over time using the finite volume method (Pruess et al., 1999). The finite volume method is used to discretise the system over time and space. The mass, energy and momentum balance equation can be presented by:

$$\frac{d}{dt}\int_{V_i} M_\kappa dV = \int_{\Omega_i} \mathbf{F}_\kappa \cdot \mathbf{n} d\Omega + \int_{V_i} q_\kappa dV \tag{1}$$

Here on the left hand side of equation (1) $t$ represents time, $M_{\kappa=1,\ldots,N}$ is the accumulation term for the mass of all components (e.g. water and $CO_2$) and $M_{\kappa=N+1}$ is the energy accumulation term, where $\kappa$ represents the mass or heat component. The amount of mass or energy is contained in a specific volume $V_i$. On the right hand side of the equation, $\mathbf{F}_{\kappa=1,\ldots,N}$ is the mass flux component and $\mathbf{F}_{\kappa=N+1}$ is the heat flux component through a specific surface $\Omega_i$, $\mathbf{n}$ is the normal unit vector pointing inwards on surface $d\Omega_i$, and $q_\kappa$ represents the source/sink term within volume $V_i$ (i.e. production/injection wells).

The mass accumulation term can be described by equation:

$$M_{\kappa=1,\ldots,N} = \phi(\rho_l S_l X_{\kappa,l} + \rho_g S_g X_{\kappa,g}) \tag{2}$$

Here $\phi$ is the porosity, $\rho$ is density, $S$ is the saturation, $X_\kappa$ the mass fraction of mass component $\kappa$ and the subscripts $l$ and $g$ represent the liquid and gas phase respectively.

The energy accumulation term can be described in a similar way:

$$M_{N+1} = (1-\phi)\rho_r C_r T + \phi(\rho_l u_l S_l + \rho_g u_g S_g) \tag{3}$$

Where the subscript $r$ denotes the rock type, $C$ the heat capacity, $T$ the temperature and $u$ the specific internal energy.

The conservation of momentum is represented in the form of a multiphase version of Darcy's law. Using Darcy's law the mass flux of component $(\kappa = 1, ..., N)$ can be described by the following equations:

$$\mathbf{F}_{\kappa} = \mathbf{F}_{\kappa,l} + \mathbf{F}_{\kappa,g} \tag{4}$$

$$\mathbf{F}_{\kappa,l} = -\mathbf{k}\frac{k_{r,l}}{\nu_l}(\nabla P - \rho_l \mathbf{g})X_{\kappa,l} \tag{5}$$

$$\mathbf{F}_{\kappa,g} = -\mathbf{k}\frac{k_{r,g}}{\nu_g}(\nabla P - \rho_g \mathbf{g})X_{\kappa,g} \tag{6}$$

where $\mathbf{k}$ is the permeability tensor, $k_r$ is the relative permeability, $\nu$ the kinematic viscosity, $P$ is pressure and $\mathbf{g}$ is the gravitation vector.

The following equation is used to describe the energy flux:

$$\mathbf{F}_{\kappa=N+1} = -K\nabla T + (h_l \mathbf{F}_{\kappa,l} + h_g \mathbf{F}_{\kappa,g}) \tag{7}$$

Here $K$ is the thermal conductivity and the second term on the right-hand side is the convective component with $h$ the enthalpy for the specific liquid or gas phase.

## 3.2 Finite volume discretisation and time incrementing

In order to numerically solve the conservation of mass, energy and momentum equations, the reservoir domain is discretised using a finite volume mesh (Pruess et al., 1999). The conservation equations have to be solved for every block $V_i$ in the discretised mesh. The following two equations are used to calculate the mass and energy average $M_{\kappa,i}$, and the source/sink average $q_{\kappa,i}$ average in block $V_i$.

$$\int_{V_i} M_{\kappa} dV = V_i M_{\kappa,i} \tag{8}$$

$$\int_{V_i} q_{\kappa} dV = V_i q_{\kappa,i} \tag{9}$$

The momentum term in equation (1) is also known as the accumulation for all fluxes over a certain surface $A_{ij}$ between connecting volume blocks $i$ and $j$. Shown in the equation below:

$$\int_{\Omega_i} \mathbf{F}_\kappa \cdot \mathbf{n} d\Omega = \sum_j A_{ij} F_{\kappa,ij} \tag{10}$$

Combining equations (8-10) results in a discretised conservation equation for each volume block $V_i$:

$$\frac{d}{dt} M_{\kappa,i} = \frac{1}{V_i} \sum_j A_{ij} F_{\kappa,ij} + q_{\kappa,i} \tag{11}$$

For every block in the mesh, equation 11 is solved. The thermodynamic state of each block can be represented by a set of primary variables (Croucher, 2020). For a two-phase flow system, as in the case of the RGF, Waiwera uses pressure and vapour saturation as primary variables. Once these primary variables are known, all other fluid properties can be derived.

The next step for numerically solving a thermodynamic model like geothermal reservoirs is to discretise the system in time. The conservation equations of mass, energy and momentum can be represented by the function:

$$\boldsymbol{g}(\boldsymbol{y}, \boldsymbol{m}) = 0 \tag{12}$$

Where $\boldsymbol{g}$ is the forward model, $\boldsymbol{y}$ the variables like pressure and vapour saturation and $\boldsymbol{m}$ the model parameters like permeabilities and deep hot upflows. Numerical solvers like AUTOUGH2 and Waiwera take small time-steps $\Delta t^n$ and solve the system to conserve the mass, energy and momentum balance equations. This process is done iteratively with the Newton-Raphson method by calculating the residual $R$ and altering the primary variables $\boldsymbol{y}$ until a predefined tolerance is met. The residuals can be derived from equation (11):

$$R^n_{\kappa,i} = \frac{M^n_{\kappa,i} - M^{n-1}_{\kappa,i}}{\Delta t^n} - \frac{1}{V_i} \sum_j A_{ij} F^n_{\kappa,ij} + q^n_{\kappa,i} \tag{13}$$

Here $R^n_{\kappa,i}$ is the residual at the $n^{th}$ time-step with $\Delta t^n$ being the time-step. The resulting residuals are nonlinear and a function of the known primary variables at the previous time-step $n-1$, the unknown primary variables at the next time-step $n$ and the model parameters. The primary variables here are pressure $P$ and vapour saturation $S_g$. For complex systems like geothermal reservoirs, solving the residuals for all mass and energy components in every volume block is computationally expensive and time-consuming. (Croucher, 2020)

### 3.3 Inverse modelling

A geothermal reservoir simulation starts with a set of model parameters and solves the forward problem, $\boldsymbol{g}(\boldsymbol{y}, \boldsymbol{m}) = 0$, in order to compute model results like temperature and pressure. The process of a forward simulation is visualised in Figure 13. The final state of the natural state simulation is used as the initial state of the production history model.

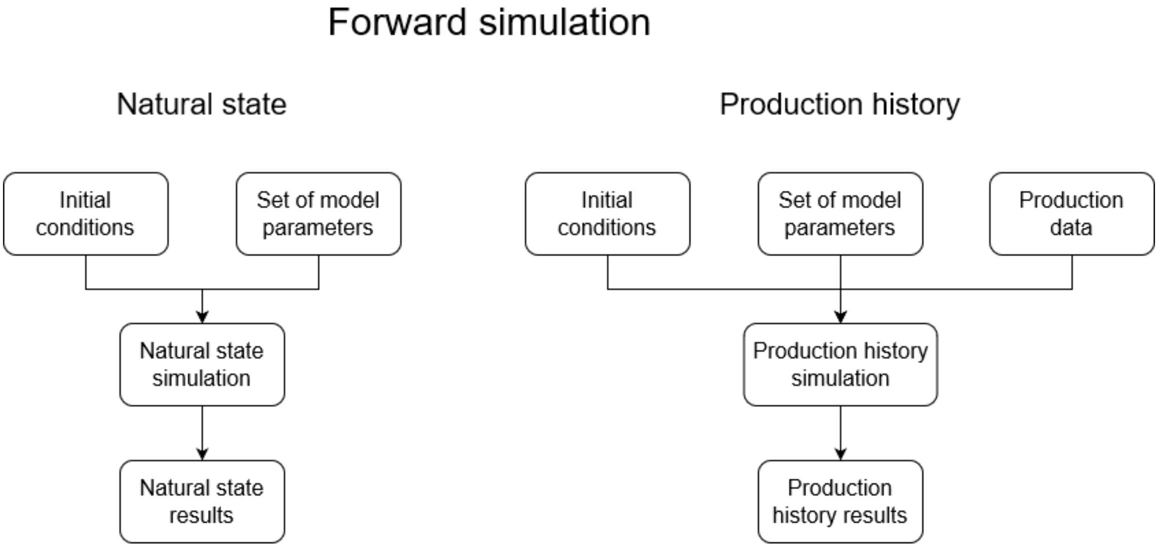


Figure 13: The process of a natural state and production history forward simulation.

In an ideal world, the model parameters are well-known, making the model a near-perfect representation of the actual reservoir. However, in reality, the model parameters have to be estimated, also known as model calibration or inverse modelling. Geothermal models are highly nonlinear and complex, and there is only limited available field data, making it an ill-posed inverse problem.

Magnetotellurics and geochemical measurements of the geothermal field give a first indication of the geological structure. An initial model can be developed using these data. After an initial forward simulation, the model results can be compared with other field measurements (e.g. temperature and pressure). Reservoir engineers aim to match model results with the available field data to create an accurate geothermal model (Bjarkason et al., 2019).

Inverse modelling can mathematically be seen as an optimisation problem. The objective is to minimise the difference between model outcomes and field data. The optimisation problem can be described by the following set of equations:

$$\begin{aligned} \min \quad & f(\boldsymbol{y}, \boldsymbol{m}) \\ \text{such that} \quad & \boldsymbol{g}(\boldsymbol{y}, \boldsymbol{m}) = 0 \end{aligned} \tag{14}$$

Here $f\ (\boldsymbol{y}, \boldsymbol{m})$ is the to be minimised objective function and, as mentioned before, the system has to comply with the conservation equations $\boldsymbol{g}(\boldsymbol{y}, \boldsymbol{m}) = 0$. The minimisation of the objective function is difficult due to the nonlinearity and complexity of the model, which causes multiple local minimums. Therefore, to make sure the geothermal model is physically correct and prevent run failures, the first part of the calibration process is done manually. Typically, manual calibration is followed by automatic calibration using software like iTOUGH2, PEST or iWaiwera.

## 3.4 Manual calibration

As mentioned in the previous section, manual calibration is crucial to develop an accurate geothermal model. Reservoir engineers combine all available field data and their knowledge on geothermal reservoirs to calibrate the geothermal model. The main model parameters for calibration are the rock permeabilities, and the strength of deep hot upflows. The model results are compared with measured field data like temperature, transient pressure and surface outflows. In contrast to automatic calibration, manual calibration allows the modeller to add new rock types and change the spatial location of existing rock types. The key steps of manual calibration, for both natural state and production history, are illustrated in Figure 14.

Visualising model results using graphs and graphical tools like Tim Isn't Mulgraph (TIM) (Yeh et al., 2013), see also Appendix D.2 help to analyse and calibrate the model. TIM is used as a pre- and post-processing graphical tool, this means the model parameters and the model results are visualised. The model parameters and results are shown in 2-D for horizontal layers and vertical slices. Figure 12 in Chapter 2 shows an example of the fault structure at the bottom layer captured with TIM.

The calibration process takes many iterations and is very time consuming since one full forward simulation takes 15 hours. The RGF model consists of 94,701 blocks and 341 rock types, meaning that some rock types have high frequencies in the model, some even up to 14,000 blocks. Consequently, adjusting the permeability of a (frequent) rock type, for example, can have a massive impact on the model results, or can even cause run failures. Again TIM offers the modeller support by visualising geological structures, temperature distributions, pressure distributions, flow directions and surface outflows. The manual calibration would typically be followed by automatic calibration, however, due to technical issues, this was not possible for the RGF model. Nevertheless, a short description of iWaiwera is given in Appendix C.

## 3.5 Uncertainty quantification

A numerical geothermal model is always an approximation of a real geothermal reservoir due to many sources of uncertainty and the complexity of the reservoir. The applicability of any model is strongly dependent on the degree of (un)certainty, thus, it is essential to quantify and analyse the uncertainty. Moreover, quantifying the uncertainty allows for

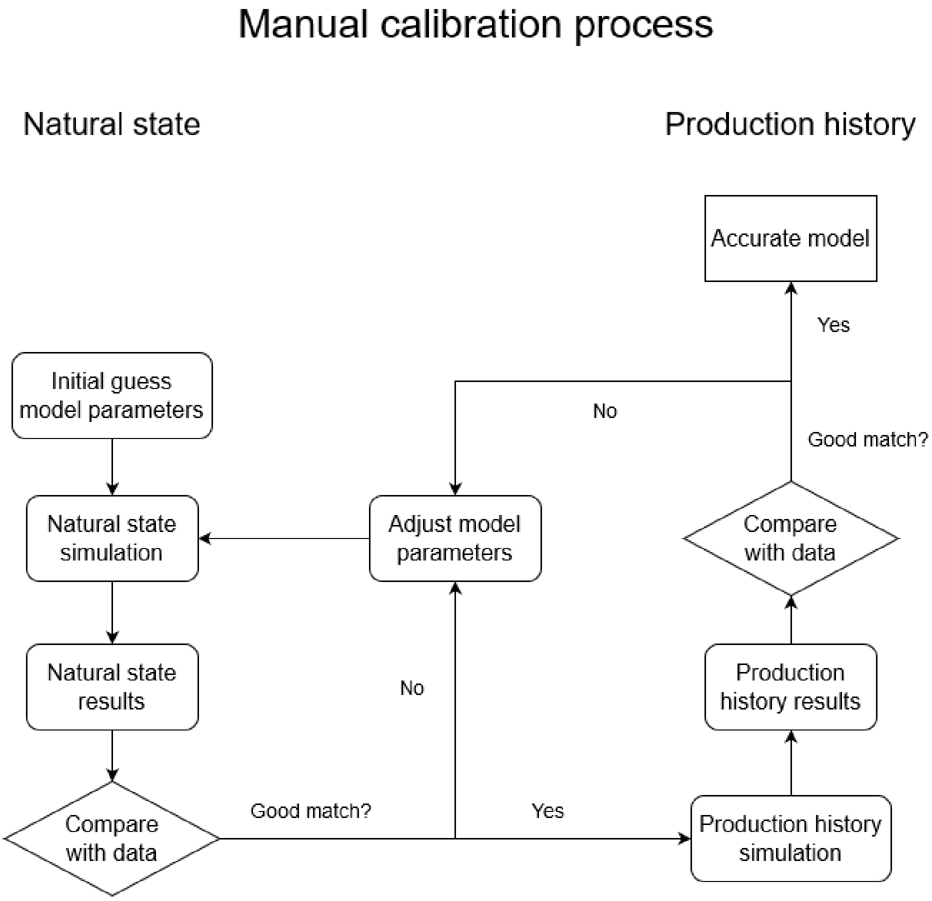


Figure 14: The process of the manual calibration of the natural state and production history models.

more comprehensive model results (compared to a single best estimate), providing better information for reservoir management decision making.

As mentioned previously, the numerical model is highly nonlinear and complex. Therefore, to estimate the uncertainty in any predictions of interest requires generating sample models and running forward simulations. This process is typically computationally (very) expensive and is generally only feasible using a supercomputer. For this study, the New Zealand eScience Infrastructure (NeSI) (New Zealand Geothermal Association, 2021) was available to run the sample simulations simultaneously. For every forward simulation, 40 CPUs, with 2GB RAM for every CPU, were allocated. Using NeSI, a single forward simulation takes four hours.

The uncertainty quantification process is graphically represented in Figure 15. This diagram shows the requirement of a calibrated model, a sensitivity matrix and a prior probability distribution of the parameters. The calibrated model will be further discussed in

Chapter 4, while the sensitivity matrix and prior probability distribution of the parameters will be further explained in Chapter 5. The next step is to carry out Bayesian inference, the theory of which is explained in the next section (while the application to the Rotorua model will be outlined in Section 3.7). Application of Bayesian inference results in a posterior probability distribution for the model parameters, i.e., the conditional probability distribution of the model parameters given the measured data. From the posterior, sample models are generated and then simulated using NeSI. The results of the simulations are used to analyse and quantify the uncertainty of model results and predictions.

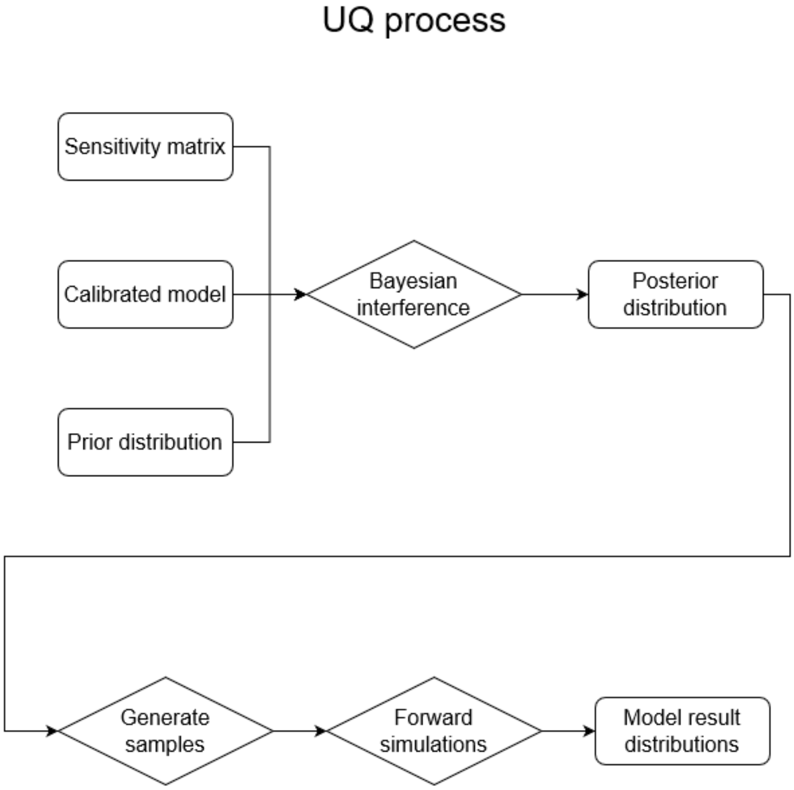


Figure 15: The process used to quantify the uncertainty in the model results of the geothermal model of Rotorua.

The Bayesian framework has been widely applied to inverse problems. For example, J. P. Kaipio & Somersalo (2007) apply Bayesian statistics to analyse the effects of discretisation errors of a general linear inverse problem, and Stuart (2010) shows examples of the Bayesian framework applied to various inverse problems originating from boundary and initial-boundary value problems. The Bayesian framework has also specifically been applied to geothermal inverse problems. For example, Maclaren et al. (2020) presents a method to incorporate additional approximation errors, caused by using a coarse model rather than a finer model, into a Bayesian framework. This method corrects for the approximation error while speeding up the Markov chain Monte Carlo (MCMC) sampling process by using the coarser model. In this study, we include the field data and model sensitivity to reduce parameter uncertainty using a Bayesian framework. Furthermore, we will study the impact of reduced parameter uncertainty, and uncertain production and reinjection on the model results and predictions.

### 3.6 The Bayesian framework

In the Bayesian framework all unknowns are taken as random variables and are described by probability distributions (Bernardo & Adrian F. M. Smith, 2009; Gelman et al., 2013; Tarantola, 2005; S. E. Kaipio J.P., 2006). The framework is based on Bayes' theorem, generally formulated as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{15}$$

where $P(A)$ is the prior probability, or simply the prior, which describes all prior beliefs of $A$, $P(B|A)$ is the likelihood function, which can be seen as the probability $B$ given $A$, while $P(A|B)$ is the posterior probability, or simply the posterior, of $A$ given $B$. The denominator $P(B)$ is the probability of evidence $B$, also known as the marginal likelihood, which is typically ignored as it is constant. Consequently, the posterior is proportional to the product of the prior and the likelihood function:

$$P(A|B) \propto P(B|A)P(A) \tag{16}$$

### 3.7 Bayes' theorem applied on Rotorua

Geothermal models contain various sources of uncertainty, such as uncertain parameters and noisy measurements. In order to update and reduce model uncertainties based on measured data, the Bayesian framework can be applied. First, equation (16) is expressed in terms of model parameters $\boldsymbol{m}$ and measured data $\boldsymbol{d}$:

$$P(\boldsymbol{m}|\boldsymbol{d}) \propto P(\boldsymbol{d}|\boldsymbol{m})P(\boldsymbol{m}) \tag{17}$$

where the data can be observed field data and new insights acquired during the calibration process. $P(\boldsymbol{m})$ represents the prior probability distribution of the model parameters, $P(\boldsymbol{d}|\boldsymbol{m})$ contains the new information, or evidence, and $P(\boldsymbol{m}|\boldsymbol{d})$ is the posterior probability distribution of the model parameters. The prior contains all our prior beliefs of the parameters like bounds, smoothness and correlation, and the likelihood incorporates any new knowledge (typically measured field data). Combining the prior and likelihood gives us the posterior, which is used to generate sample models for the uncertainty analysis.

The likelihood incorporates the uncertainty of the field data, which is assumed to be independent Gaussian noise with a mean of zero, and independent of the model parameters. Furthermore, we assume the prior is a Gaussian distribution, and the prior and likelihood are independent of each other. As a result, the posterior parameter probability distribution can be characterised by the equation below:

$$P(\boldsymbol{m}|\boldsymbol{d}) \propto exp(-\frac{1}{2}((\boldsymbol{g(m)}-\boldsymbol{d}_{\text{obs}})^T \boldsymbol{C}_d^{-1}(\boldsymbol{g(m)}-\boldsymbol{d}_{\text{obs}})+(\boldsymbol{m}-\boldsymbol{m}_{\text{prior}})^T \boldsymbol{C}_{\text{prior}}^{-1}(\boldsymbol{m}-\boldsymbol{m}_{\text{prior}}))) \tag{18}$$

where $\boldsymbol{g(m)}$ is the forward model, $\boldsymbol{m}_{\text{prior}}$ are the prior model parameters, $\boldsymbol{C}_d$ is the covariance matrix of the data and $\boldsymbol{C}_{\text{prior}}$ the prior covariance matrix of the parameters. The form of these covariance matrices will be discussed in the next section. Equation (18) fully characterises the posterior, however, generating the full posterior is computationally expensive and time consuming. In order to save time and make it easier to generate sample models on demand, we assume a Gaussian approximation of the posterior with mean $\boldsymbol{m}_{\text{MAP}}$, which is the maximum a posteriori, i.e., the point in parameter space maximising the posterior, i.e.:

$$\boldsymbol{m}_{\text{MAP}} := \underset{\boldsymbol{m}\epsilon\boldsymbol{M}}{\text{argmax}}\ P(\boldsymbol{m}|\boldsymbol{d}) \tag{19}$$

For the Gaussian approximation we first need to assume a local linearisation of the forward model:

$$\boldsymbol{d} \approx \boldsymbol{g}(\boldsymbol{m}_{\text{MAP}}) + \boldsymbol{S}(\boldsymbol{m} - \boldsymbol{m}_{\text{MAP}}) + \boldsymbol{e} \tag{20}$$

where $\boldsymbol{S}$ is the sensitivity of the model outcomes $\boldsymbol{g}(\boldsymbol{m})$ with respect to the model parameters $\boldsymbol{m}$ and $\boldsymbol{e}$ contains any form of error like modelling errors and measurement noise. The local linearisation of the forward model is equivalent to taking a local Gaussian approximation to the posterior distribution, i.e.:

$$P(\boldsymbol{m}|\boldsymbol{d}) \approx \mathcal{N}(\boldsymbol{m}_{\text{MAP}}, \boldsymbol{C}_{\text{post}}) \tag{21}$$

where $\boldsymbol{C}_{\text{post}}$ is the posterior covariance matrix, which takes prior and new knowledge into account. The mean $\boldsymbol{m}_{\text{MAP}}$ and covariance matrix $\boldsymbol{C}_{\text{post}}$ are combined to generate samples using the multivariate normal distribution sampling function from Python (Van Rossum & Drake, 2009). In order to generate sample models we need the posterior covariance matrix, which can be approximated (Tarantola, 2005; S. E. Kaipio J.P., 2006) as:

$$\boldsymbol{C}_{\text{post}} \approx (\boldsymbol{S}^T \boldsymbol{C}_{\text{d}}^{-1} \boldsymbol{S} + \boldsymbol{C}_{\text{prior}}^{-1})^{-1} \tag{22}$$

where $\boldsymbol{C}_{\mathrm{d}}$ is the data covariance matrix, which contains independent Gaussian noise on the measurements, and $\boldsymbol{C}_{\mathrm{prior}}$ is the prior covariance matrix, which incorporates variance, correlation and smoothness of the parameters. The sensitivity matrix $\boldsymbol{S}$, see equation (23), can be easily computed from Waiwera. Furthermore, the likelihood is represented in the first part of the right hand side $\boldsymbol{S}^T \boldsymbol{C}_{\mathrm{d}}^{-1} \boldsymbol{S}$, and the prior in the second part $\boldsymbol{C}_{\mathrm{prior}}^{-1}$. The sensitivity matrix is formulated as:

$$\boldsymbol{S} = \frac{d\boldsymbol{g(m)}}{d\boldsymbol{m}} = \begin{bmatrix} \frac{dg_1}{dm_1} & \cdots & \frac{dg_1}{dm_{\mathrm{N_m}}} \\ \vdots & \ddots & \vdots \\ \frac{dg_{\mathrm{N_d}}}{dm_1} & \cdots & \frac{dg_{\mathrm{N_d}}}{dm_{\mathrm{N_m}}} \end{bmatrix} \tag{23}$$

where $N_d$ is the number of measured data points and $N_m$ the number of model parameters. In the next section, the construction of the data and the prior covariance matrix will be explained in detail.

# 4 Model Calibration

The main goal of reservoir modelling is to be able to predict future behaviour. Model predictions are used as a guidance for efficient and sustainable reservoir management. The first step is to create a model that is an accurate representation of the real reservoir. The process of creating an accurate model is known as inverse modelling or model calibration. However, a calibrated model provides only a single best estimate of the model results and predictions. The calibration process and results will be discussed in this chapter.

As a consequence of modelling errors and noise measurements, the calibrated model results and predictions are uncertain. In order to provide more extensive guidance for reservoir management, it is vital to analyse and quantify the uncertainty of the model results and predictions. The uncertainty quantification of the Rotorua model was carried out from a Bayesian perspective, as introduced in sections 3.5 to 3.7. The results and analysis of the uncertainty quantification will be discussed in Chapter 5.

## 4.1 Model setup

In Section 2.2 the new numerical model of Rotorua is introduced. The new model includes a finer grid and consists of a different (rectangular) geometry than the old model. Because the new numerical grid is introduced, the new model requires calibration to improve the accuracy of the model and its predictions. First, the geological structure was transferred from the 3D Leapfrog model onto the numerical grid by Van Vlijmen (2020). Thereafter, Van Vlijmen (2020) attempted to manually calibrate the natural state model, however, due to time limitations of his internship, he could only do a handful of calibration iterations. His natural state model was build in AUTOUGH2, hence, the model was first converted to a Waiwera compatible format. The latest version of his model is used in this study as the initial natural state model. After further manual calibration, of which the key steps are explained in Section 3.3, the calibrated model results will be compared with the initial model results in the following sections.

Furthermore, there was no production history model yet using the new numerical grid. The production history model requires a calibrated natural state model (to prevent run failures), and production history data (like production and reinjection flow rates). The available information of the production and reinjection wells at Rotorua has been outlined by Ministry of Energy (1985). The well locations before and after the 1986 Wellbore Closure Programme are shown in Figure 16. The figure shows a clear reduction of production wells and an increase in reinjection wells as was required by the Rotorua Geothermal Regional Plan. The total production and reinjection flow rates were reliably estimated (Ministry of Energy, 1985; Environment Bay of Plenty (EBOP), 1999), illustrated in Figure 17. The plot shows a large production decline after the 1986 Wellbore Closure Programme and a steady increase in reinjection. Since 2006 there is a steady production and reinjection with a total reinjection rate of 90% of the total production rate, as agreed in the Rotorua Geothermal Regional Plan. Unfortunately, there is little information about the flow rates of the individual wells. Therefore, the production and reinjection rates of the individual wells have been estimated, and cause additional uncertainty in the production history model. The production uncertainty is unique for Rotorua because, generally for geothermal reservoirs, there are good production records.

Besides production data (required to build the production history model), there are measured data, used to calibrate the model, like temperature measurements. All data came from various sources and were in various formats (like spreadsheets, configuration (CFG) files and graphical images). Thus, they needed to be converted to a standard format compatible with various geothermal simulators. The data for every well is converted and combined using Python scripts to JavaScript Object Notation (JSON) files, adding up to a library of 600 JSON data files. The JSON format is compatible with AUTOUGH2 and Waiwera, and uses dictionaries to create a clear structure, containing information such as the well name, well coordinates, temperature measurements and production rates. An example JSON file of well RR724 is included in subsection A.2. It is essential to assess the quality and reliability of the measured data because calibrating to faulty measurements causes inaccuracies in the model.

## 4.2 Manual calibration of natural state and production history

The natural state and production history models require calibration to predict future behaviour accurately, the calibration process is illustrated in Figure 14 of Section 3.4. Model calibration is the process of matching the model results with the observed field data. The calibration process generally consists of a manual and automatic part. Manual calibration is time-consuming and complex. This process consists of running many forward simulations and adjusting model parameters after each simulation to better fit the model results with the field data, as explained in Section 3.4. Nevertheless, manual calibration is a necessary step in the calibration process because manual calibration allows the geothermal modeller to incorporate expert knowledge while making adjustments, i.e., add new rock types or change the spatial locations of a specific rock type. These adjustments are not possible in the automatic calibration part.

First, we look at the natural state calibration because a sufficiently calibrated natural state model will prevent run failures or unrealistic production history results. A forward natural state simulation requires initial conditions and a natural state model (including all parameters), see also Figure 13. If model results are improved with an adjusted set of model parameters, the natural state results are used as the new initial conditions. Model parameters have to be adjusted carefully because too many or too radical adjustments will cause run failures.

The geothermal field's temperature distribution is strongly linked to the geological structure and the deep hot upflow sources of the reservoir (O'Sullivan et al., 2001; O'Sullivan & O'Sullivan, 2016). Hence, the main model parameters used to calibrate the natural state are the rock permeabilities and the deep hot upflow rates. There are 2500 temperature data points available for Rotorua, distributed over 184 wells. These temperature measurements form downhole temperature profiles that can be compared with the natural state model results, see for example Figure 22. Although 2500 temperature data points seem like a sig-

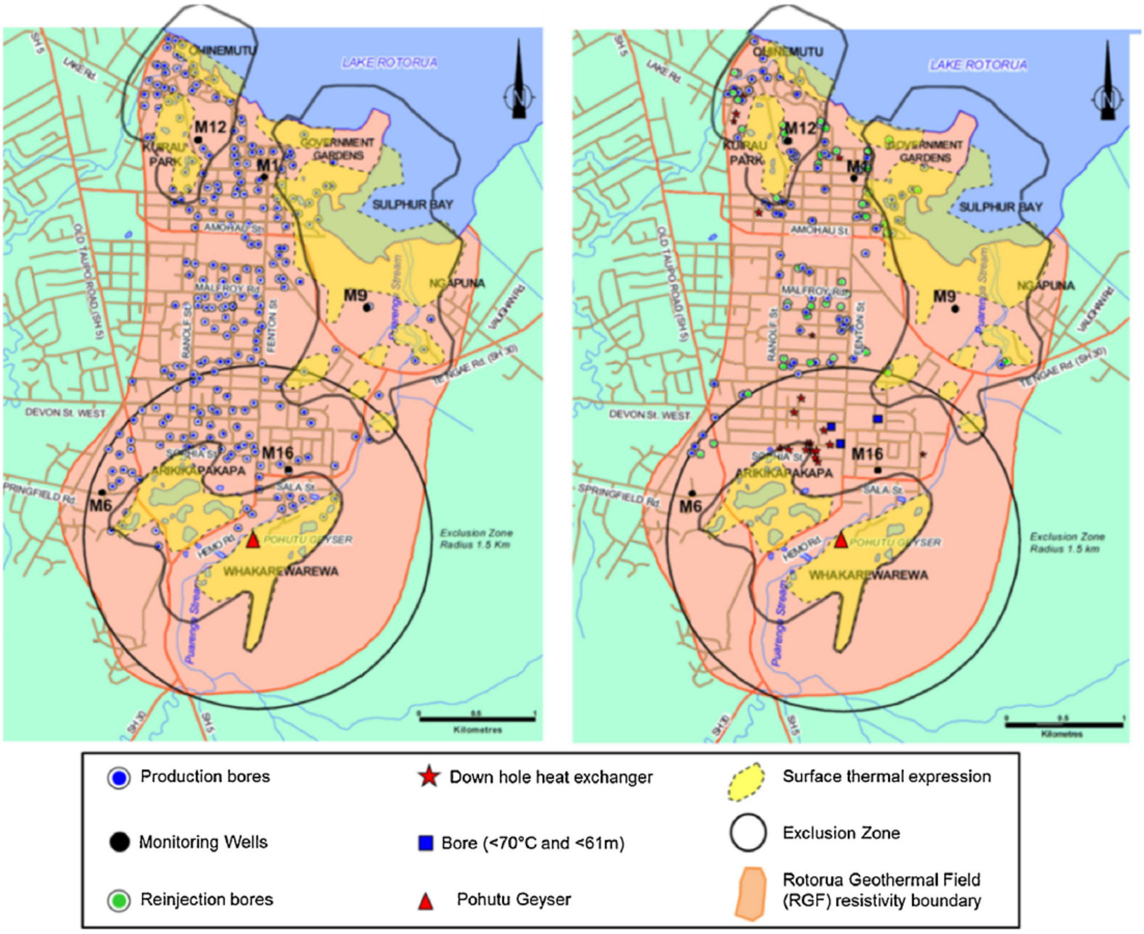


Figure 16: Production and reinjection well distribution map before (left) and after (right) wellbore closure program in 1986.

nificant amount of data for model calibration, the temperature measurements are shallow, 0 to 300 mRL (meters relative level or meters above sea level). The shallow measurements make it challenging to recreate the geological structure, particularly at great depths, -1500 to 0 mRL. The geothermally active areas of Rotorua are known and provide validation by comparing these areas with areas of high surface temperatures and mass outflows of the natural state model, shown in Figure 24. The natural state temperature and surface activity results will be discussed in the next section.

The production history model covers the period of 1950 to 2020. There is limited data available for the production history calibration. A few monitoring wells recorded the relative change in water level, which can be approximated to a change in pressure by $\Delta p = \rho g \Delta h$. Here $\Delta p$ is the pressure change, $\rho$ the water density, g the gravitational acceleration and $\Delta h$ the relative change in water level. In order to validate if the production history model

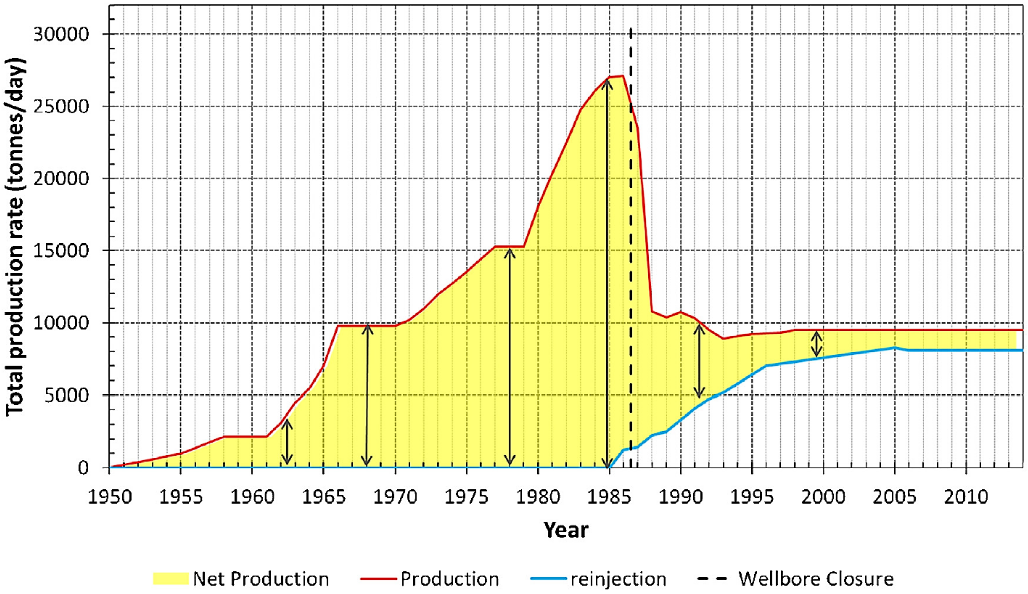


Figure 17: The cumulative estimated production and reinjection rates at the Rotorua geothermal field from 1950-2014 (Ratouis et al., 2016b) (repeated figure from Chapter 1).

is running correctly, we can check the production rates of the model. The production and reinjection are set to deliverability, which means the wells only produce and inject what is physically possible. Figure 18 shows an old model that cannot produce the amount given as input data, which indicates a high flow resistivity. A tight permeability structure causes high flow resistivity. Consequently, the natural state model needs to be re-calibrated by increasing the permeability, particularly in the areas causing the production issues.

The initial natural state model, build by Van Vlijmen (2020), required further calibration because the temperature of the model did not match the temperature data adequately. Additionally, the initial natural state model caused run failures for the production history simulations. In this study, 80 calibration iterations have been carried out.

The calibrated natural state model matches the temperature data better, and the production history model runs successfully (reaching the estimated input production rates) and matches reasonably with pressure data. The main differences between the calibrated and the initial model are:

- The permeability of every rock type (excluding the clay cap) has been increased by a factor of 10-100, resulting in lower flow resistivity.
- The permeability of every clay cap rock type has been increased by a factor of 10-400, also creating lower flow resistivity. Observations of high surface activity at Rotorua indicate a "leaky" clay cap, supporting the increase of clay cap permeabilities.

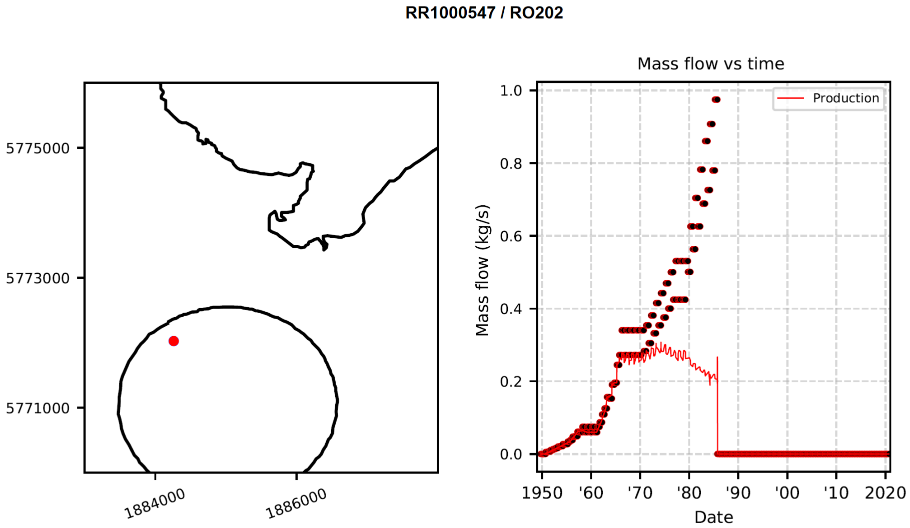


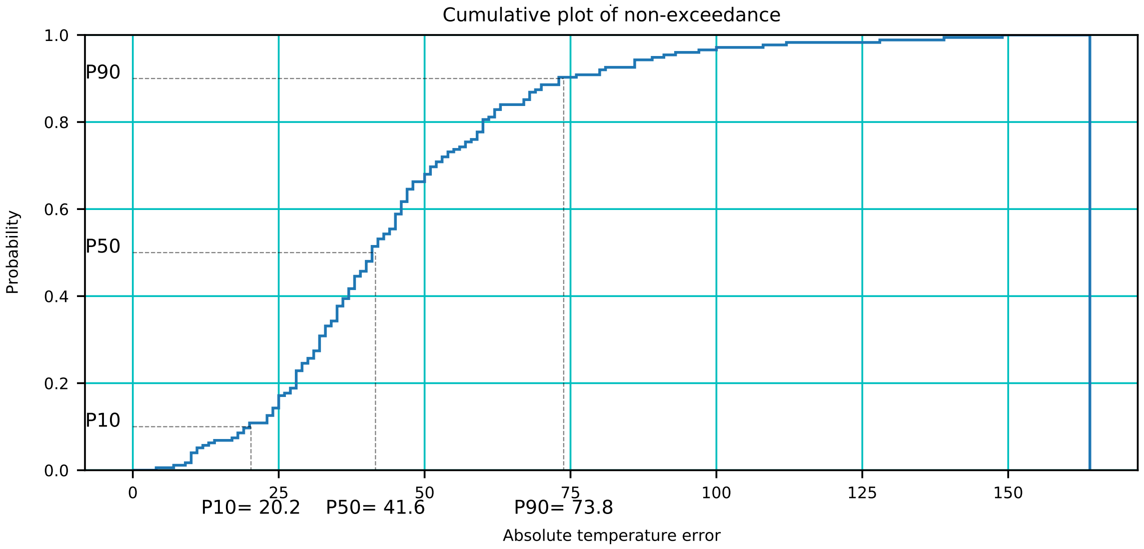
Figure 18: An example production well, from an old model, (RR1000547) not reaching the required production rates given as input. The well location is indicated on the map on the left. The graph on the right shows the the production model input rates as dots and the model production output rates as a solid red line. The mismatch between model input and output indicate that the model requires (further) calibration.

- Two rock types have been added to allow a smoother transition between rock structures.
- The spatial location of specific rock types has been changed to match specific temperature profiles.
- The total hot upflow rates have been increased by 11%.

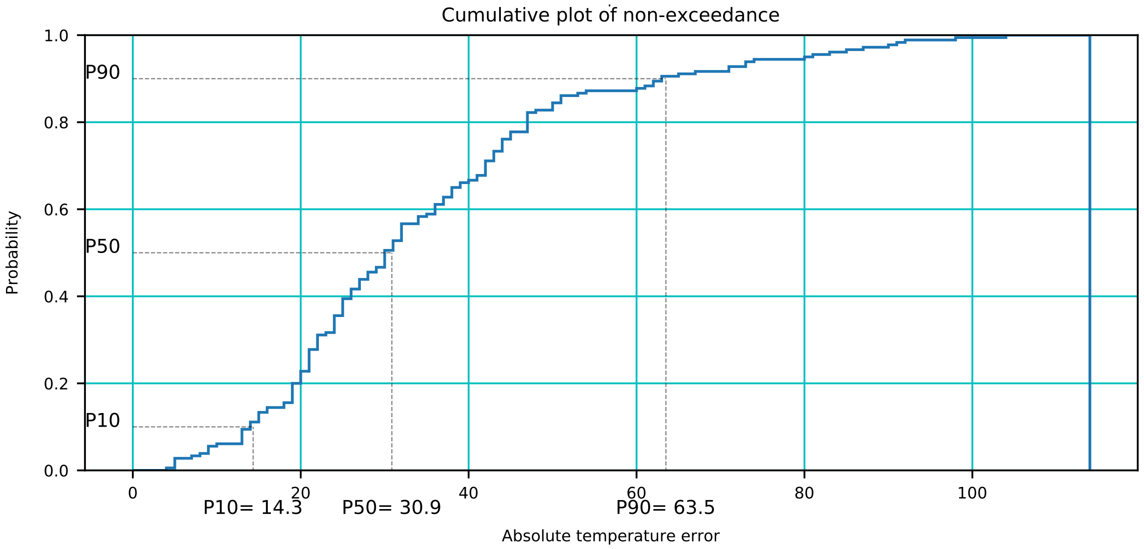
## 4.3 Calibrated model results and predictions

The calibrated model results show an improved match with field data compared with the initial model. Figure 19-21 show an overview of the overall model improvements. Figure 19 compares the cumulative plots of non-exceedance of the initial (a) and the calibrated model (b). On the horizontal axis is the absolute temperature Root Mean Square Deviation (RMSD) of the wells, and on the vertical axis are the probabilities of the wells not exceeding these RMSD temperature values. The mismatch between temperature model results and data has been significantly by calibration. The temperature RMSD of 90% of the wells are now below 63.5 °C, compared to the 73.8 °C of the initial model, which is an overall temperature reduction of 14%.

The histograms in Figure 20 show the temperature mismatch distribution of the wells. The horizontal axis presents the temperature RMSD of the wells, where a positive temper-



(a) Initial model results.



(b) Calibrated model results.

Figure 19: The cumulative plot of non-exceedance absolute temperature error, the initial model at the top and the calibrated model at the bottom. The absolute temperature RMSD of the wells is on the horizontal axis and the probability of not exceeding the temperature RMSD value is on the vertical axis.

ature RMSD means the well's temperature is too hot in the model compared to the field data, and a negative temperature RMSD means the well is too cold. On the vertical axis is the frequency of the RMSD. The histogram of the calibrated model shows a reduced width, supporting the conclusion of an overall temperature reduction.

Furthermore, the percentage of wells with a positive temperature mismatch has been reduced from 90% to 82%. In contrast, this means that the percentage of wells with a negative temperature mismatch has increased from 10% to 18%, however, these cold wells have only a small temperature deviation (RMSD $< -50°\text{C}$).

From the statistics of Figure 19 and 20 we concluded that the temperature match between the calibrated model and field data has improved. Figure 21 presents a RMSD temperature map with the well locations. The map of the calibrated model shows an increased number of wells with good temperature matches indicated by green crosses ($-10°\text{C} <$ RMSD $< 10°\text{C}$). Moreover, most of the dots have reduced in size (overall improvement of temperature match). However, there are still areas with a high density of temperature mismatches (RMSD $> |40°\text{C}|$). These areas require attention if the model is further calibrated.

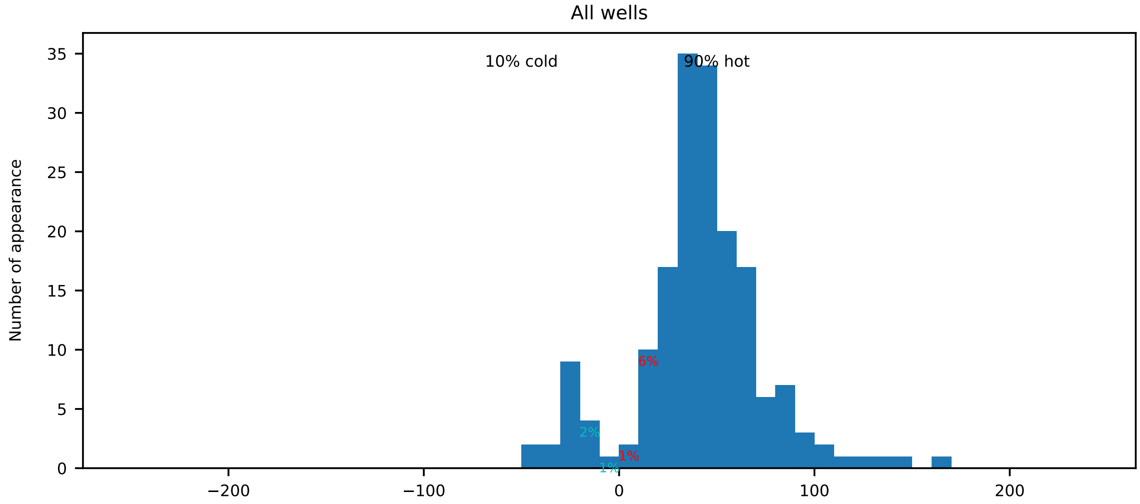
Model adjustments can improve the temperature match for some wells, but worsen the temperature match for other wells. Therefore, it is vital to analyse the overall temperature results (like in Figure 19-21) and the temperature results for each well for every iteration of the manual calibration process. There are 184 downhole temperature profiles used for the natural state calibration. Figure 22 shows a selection of three downhole temperature profiles from wells located in the areas Arikikipakapa, Devon Street and Government Gardens. The calibrated model results of these wells match the temperature measurements better than the initial model results.

Although the new model has better numerical accuracy with a finer grid (smallest block size is 125x125m), occasionally, two or three wells are located in the block of the grid. While these wells have moderate differences in temperature profiles, the model temperatures are the same, making it impossible to match both temperature profiles. An example of two wells located in the same block can be seen in Figure 23. Furthermore, the figure shows potential for model improvements by calibration, because in most parts of the well the temperature data deviates 50°C from the model.

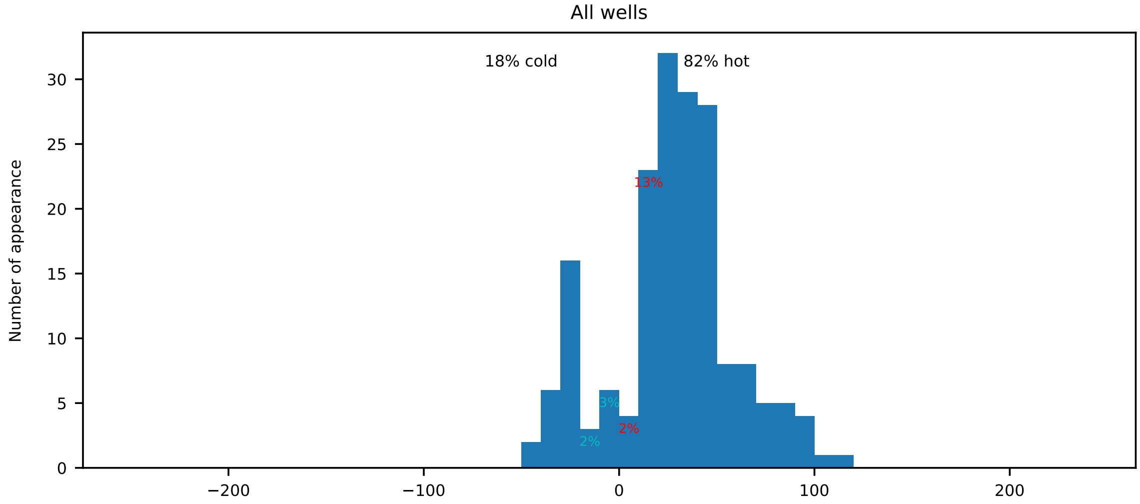
In general, the temperature match has been improved. However, it is challenging (or even impossible) to match the temperature in every well, particularly with manual calibration due to finite numerical precision, and the nonlinearity and complexity of the model.

Rotorua has many geothermally active areas, indicated on the map in Figure 24a. The location of these areas can be used to validate the calibrated model. Model results like surface temperature and mass outflow indicate geothermally active areas, as can be seen in Figure 24b and 24c. Areas with high surface model temperatures correspond reasonably well with the observed geothermally active areas. Areas in the model with high mass outflow show some resemblance with the observed geothermally active areas at Kuirau Park, Pohutu Geyser (centre of exclusion zone) and south Ngapuna. The calibrated model shows geothermal activity in the right locations, however, the model requires further calibration to increase the mass outflow at mostly the Arikikapakapa and Whakarewarewa.

Typically, manual calibration is followed by automatic calibration, which uses algorithms



(a) Initial model results.



(b) Calibrated model results.

Figure 20: The histograms show the percentage of the wells that are too hot and too cold with the initial model results at the top and the calibrated model results at the bottom.

to improve the temperature match further. In this study, it was not possible to manually calibrate the model further due to time limitations, and automatic calibration was not possible due to technical development issues of iWaiwera.

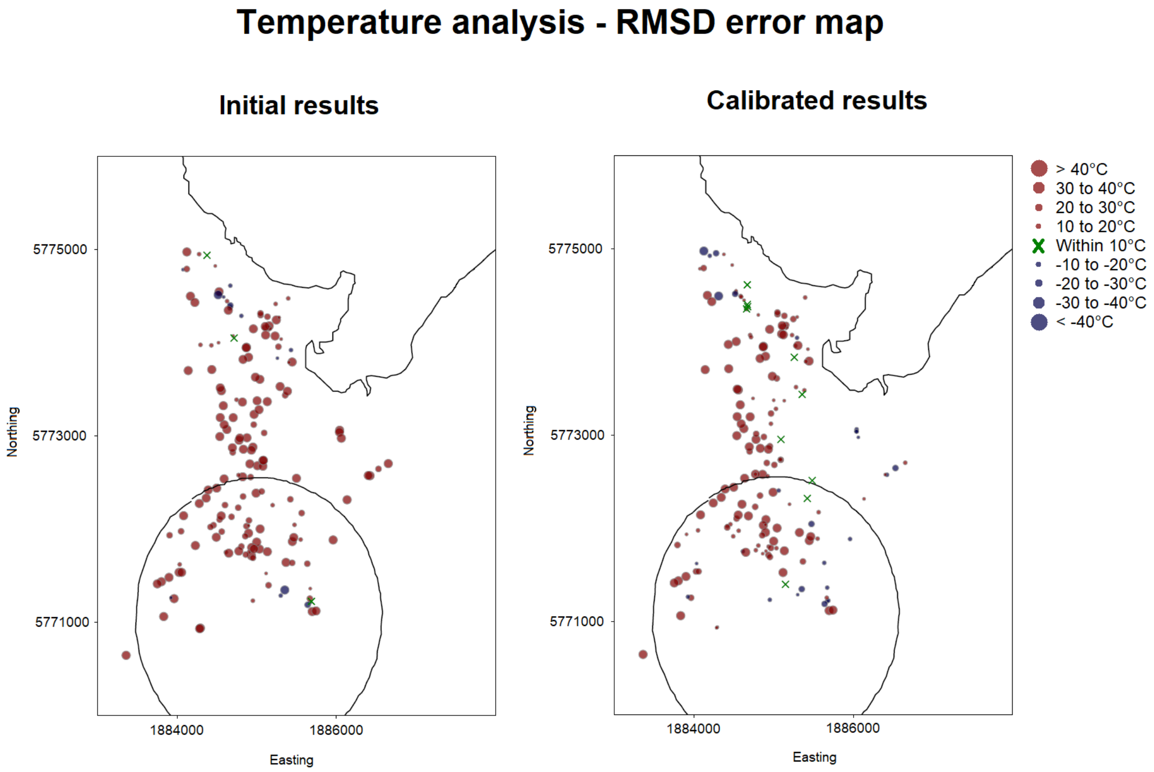


Figure 21: The temperature root mean square deviation (RMSD) error map of the initial and the manually calibrated NS model. A green cross is a good fit ($-10°C < \text{RMSD} < 10°C$), a red dot means the well is too hot ($\text{RMSD} > 10°C$) and a blue dot means the well is too cold ($\text{RMSD} < -10°C$). A bigger dot is equal to a larger RMSD.

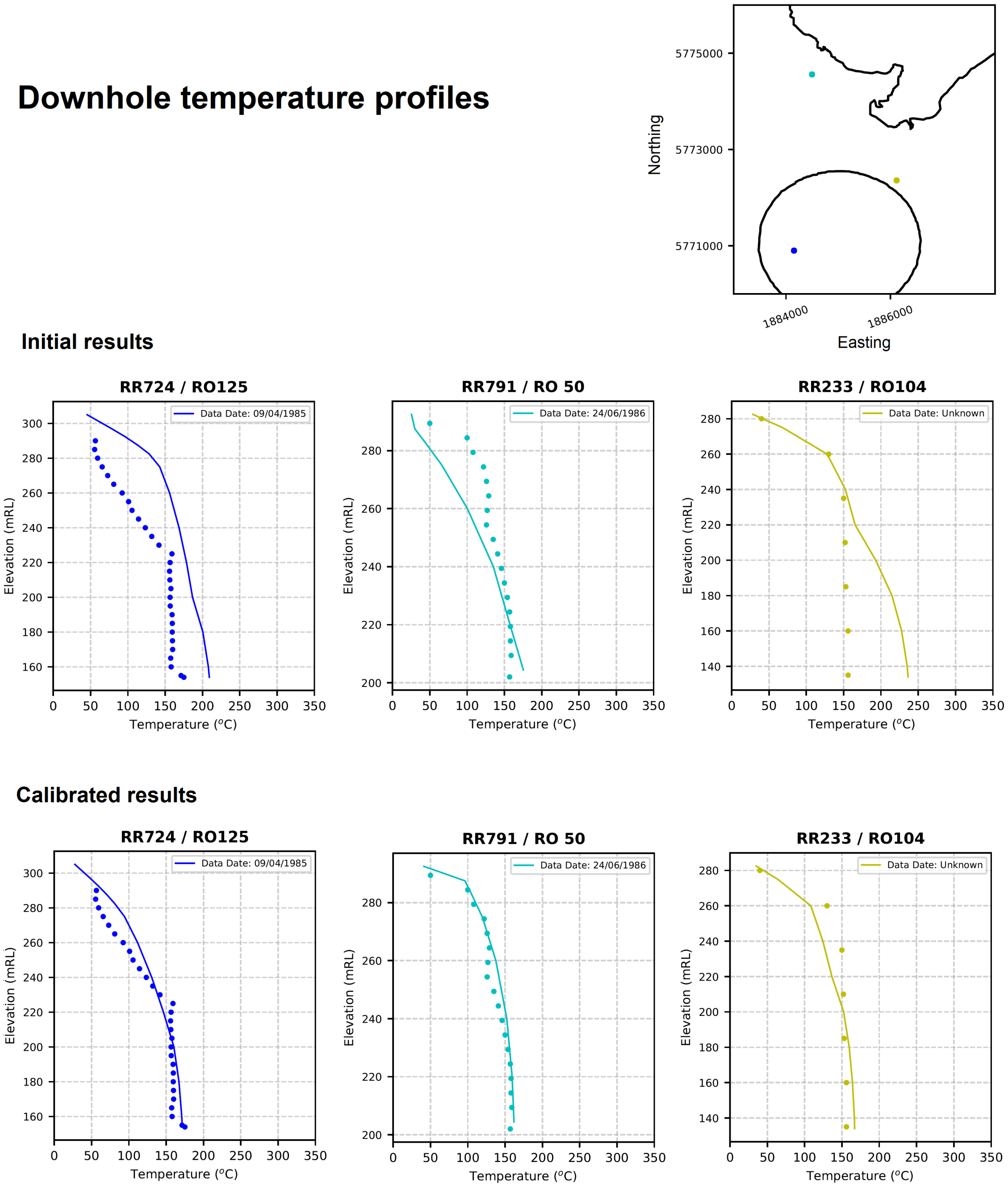


Figure 22: A selection of three downhole temperature profiles is shown to compare initial and calibrated model results. The wells RR724 (dark blue), RR791 (turquoise) and RR233 (yellow) located in the respective areas Arikikipakapa, Kuirau Park and Ngapuna. The well locations are indicated in colour on the map at the top right, the exclusion zone and the lake are marked for reference.

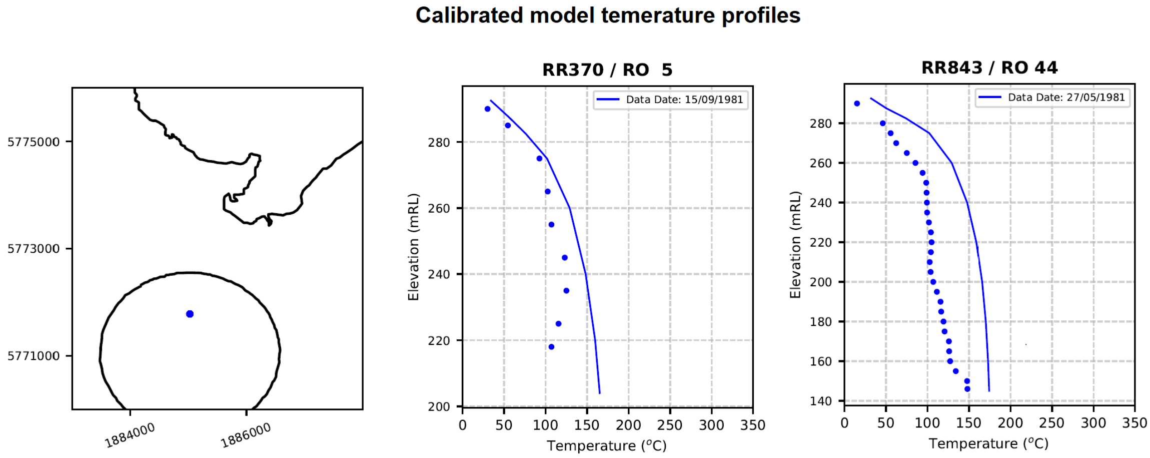
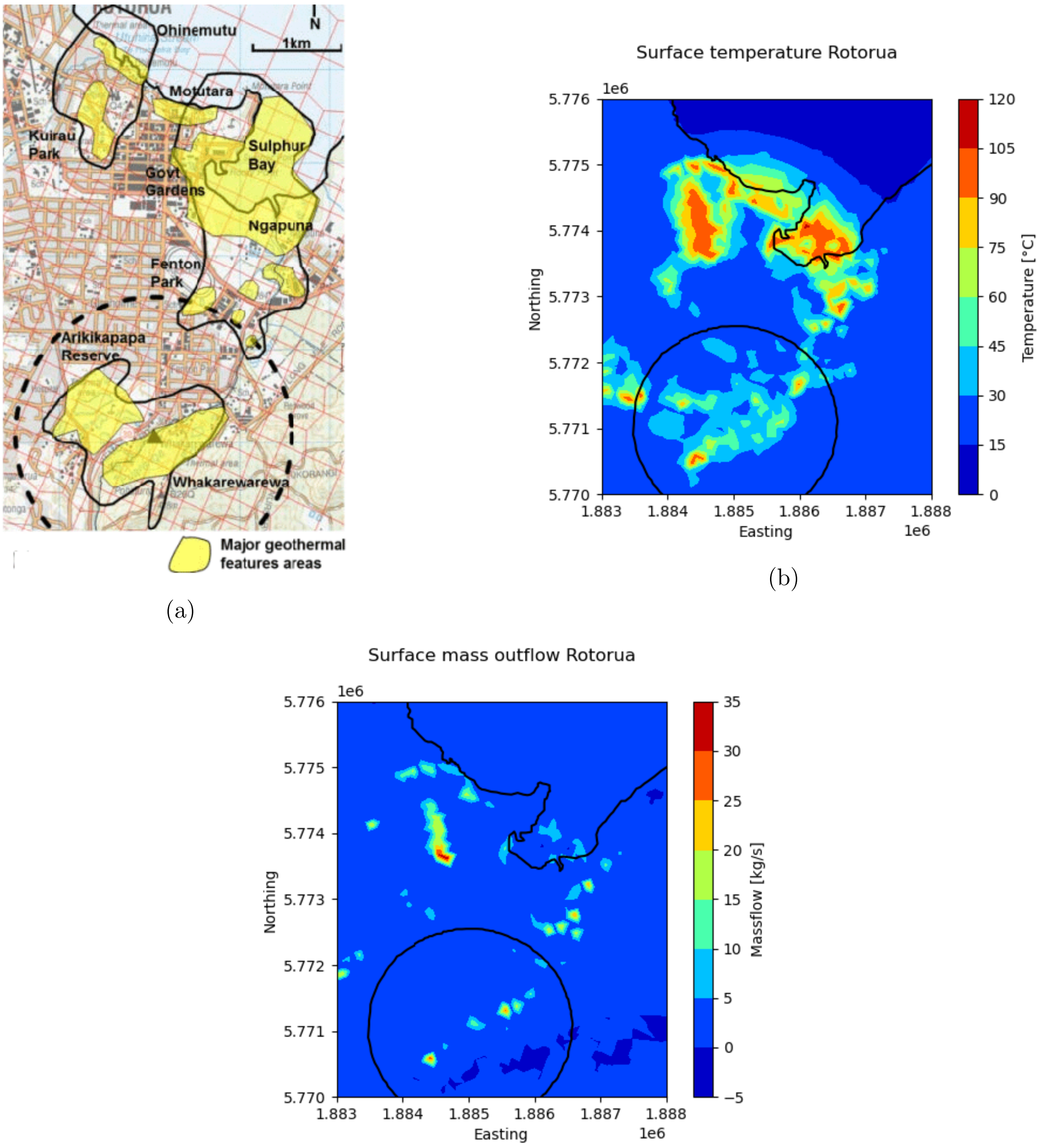


Figure 23: A selection of two wells located at the same block with different temperature data.

Like the natural state model, the production history model also requires manual calibration, which is done by matching the transient pressure model results with the available transient pressure data. While calibrating the production history model, the modeller has to be cautious for natural state result changes. A production history model improvement may worsen the natural state results. Permeabilities and hot upflows are used again to calibrate the production history model. The transient pressure results of a representative selection of three monitoring wells can be seen in Figure 25. The wells are located in the areas of Kuirau Park, Government Gardens and Whakarewarewa. The model results show a reasonable match with the field data, especially considering that the mass flow rates are very uncertain for Rotorua. However, the pressure results of well RR777 shows room for improvement. No further calibration was carried out, due to technical development issues of iWaiwera and time limitations.

The wellbore closure program was implemented in 1986 due to observations of decreasing subsurface pressure, surface temperature and mass outflow. These observations coincide with the production history model results shown in Figure 25 and 26. The pressure decreases until 1986, after which a significant increase can be observed due to production reduction and reinjection implementation (Figure 25). Furthermore, the surface features like temperature and mass outflow in large areas of the geothermal field show a decline between 1950 and 1986 (Figure 26). Furthermore, in large parts of the geothermal field, surface activity increased between 1986 and 2020 because of the wellbore closure (Figure 27). The surface feature decline in 1950-1986 and the increase in 1986-2020 of the model correspond with the field observations.

One of the main goals of geothermal modelling is to be able to make predictions about future behaviour of the reservoir. As such, the production history model was run through



(a)

(b)

(c)

Figure 24: The thermally active areas indicated on the map of Rotorua (a) (Ratouis et al., 2014), the NS surface temperature (b) and mass outflow (c). The lake and exclusion zone are marked for reference.

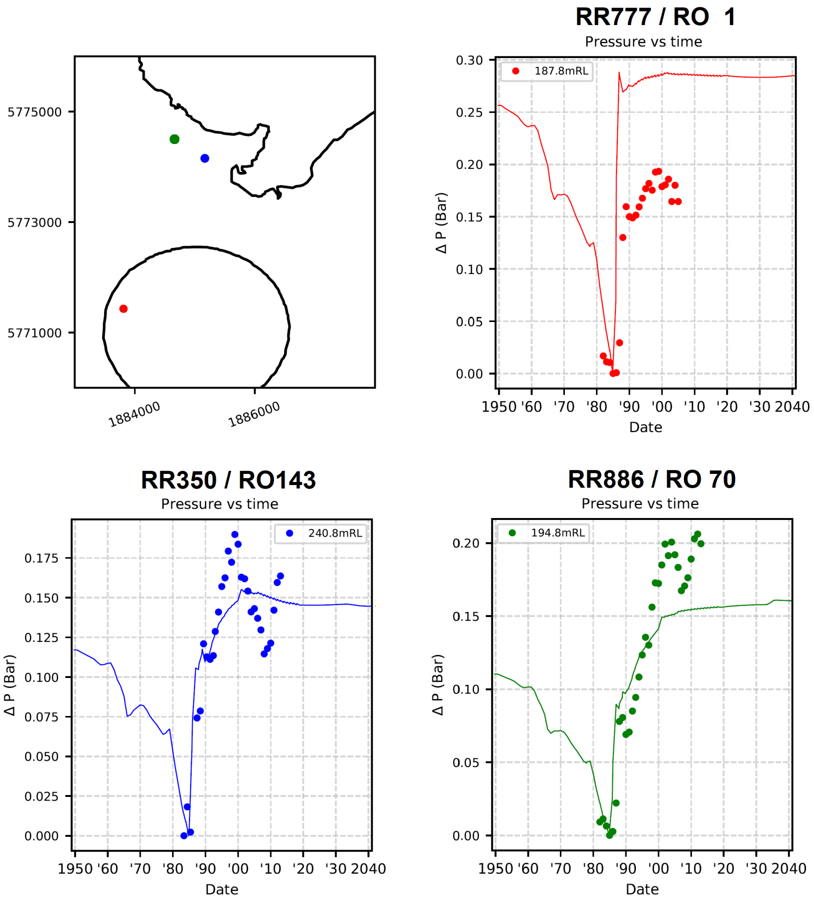
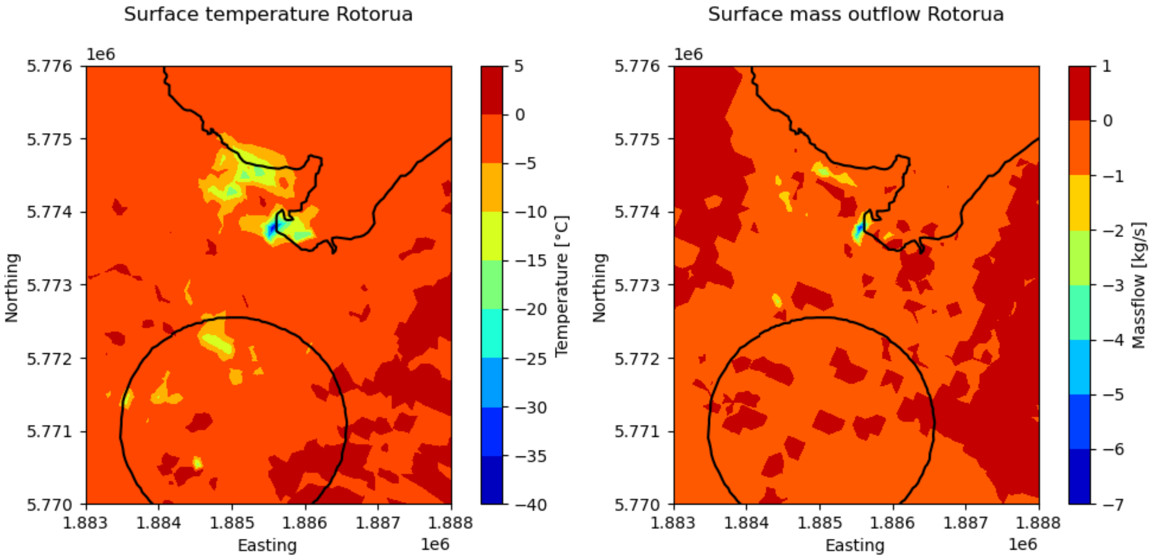


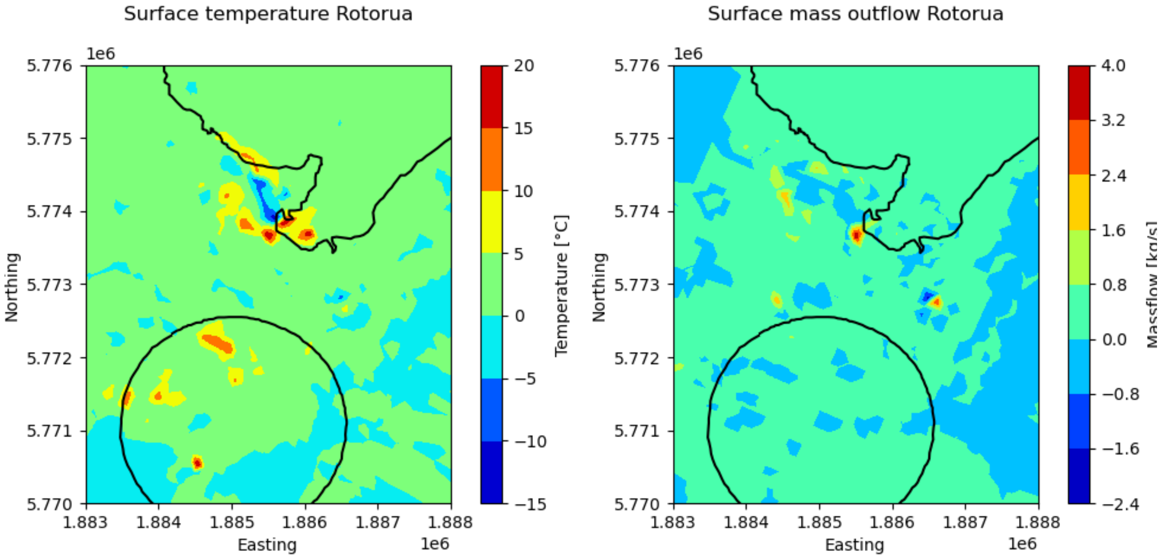
Figure 25: Transient pressure plots of three different wells at different well locations. The solid line is the calibrated model result and the dots are the field data.

to 2040 by extending the current seasonal production and reinjection mass flow rates. The results of Figure 25 and 28 indicate that if we keep using the geothermal reservoir for the coming 20 years with the current rates, we will be able to sustain the reservoir. The pressure stays relatively constant between 2020 and 2040 (Figure 25), and the surface temperature and mass outflow changes are negligible (Figure 28). However, due to large uncertainties in the model parameters, the predictions become more valuable by performing an uncertainty quantification, as discussed in the next chapter. Furthermore, future work could explore additional future scenarios with, for example, increased production rates.



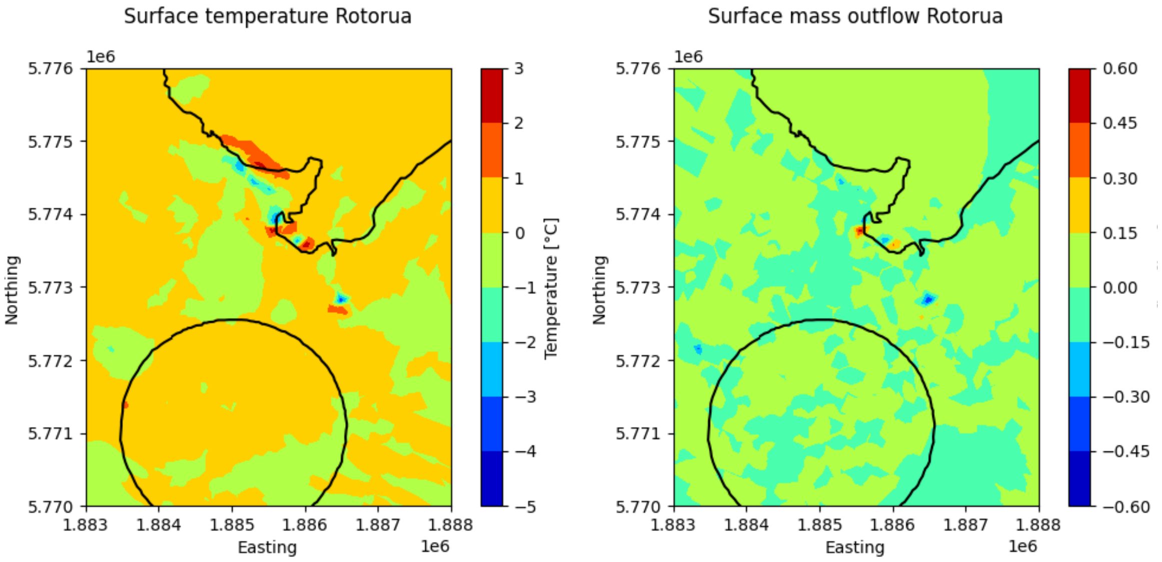
(a) (b)

Figure 26: The surface temperature (a) and mass outflow (b) change between 1950-1986.



(a) (b)

Figure 27: The surface temperature (a) and mass outflow (b) change between 1986-2020.



(a)

(b)

Figure 28: Predictions of the surface temperature (a) and mass outflow (b) change between 2020-2040.

# 5 Uncertainty Quantification

Quantifying and analysing the uncertainty of the model results and predictions provides helpful insights for reservoir management. We can use the Bayesian method described in Section 3.5-3.7 to reduce the uncertainty of the results and predictions of the calibrated model. As a recap, to quantify the uncertainty of the nonlinear geothermal model, we need to generate sample models and run numerous forward simulations. Every sample model was simulated using NeSI (the supercomputer at the University of Auckland). We generate the sample models using a posterior probability distribution of the model parameters $P(\boldsymbol{m}|\boldsymbol{d}) \approx \mathcal{N}(\boldsymbol{m}_{\text{MAP}}, \boldsymbol{C}_{\text{post}})$, incorporating our prior beliefs of the parameters and the measured data. The Rotorua model has been manually calibrated, which provides the $\boldsymbol{m}_{\text{MAP}}$. The advantage of using the Bayesian approach is that we can reduce model parameter uncertainty by including the measured data and the model sensitivity in the posterior covariance matrix $\boldsymbol{C}_{\text{post}} \approx (\boldsymbol{S}^T \boldsymbol{C}_{\text{d}}^{-1} \boldsymbol{S} + \boldsymbol{C}_{\text{prior}}^{-1})^{-1}$. Reducing the parameter uncertainty will reduce model result and prediction uncertainties. The advantage of using Waiwera, besides faster computation, is the easy access to the model sensitivity. The next section will outline the data covariance matrix and the prior covariance matrix (of the model parameters).

Every natural state and production history sample model has been simulated on NeSI using 40 CPUs with 2GB RAM for every CPU. Natural state sample models that took over 6 hours to converge were rejected. Out of the 226 natural state sample simulations, 75 have successfully converged, the rest were rejected, which is equal to success ratio of 33%. The successful natural state sample models were used as initial states for the production history sample models. The rejected natural state sample models were either physically impossible or deviated too much from the calibrated model.

## 5.1 Uncertainty in model parameters and field data

At the moment, it is only possible to compute the model sensitivity of the natural state with Waiwera. Since we only have the natural state model sensitivity, we can only include the temperature data in the data covariance matrix. In order to speed up the uncertainty quantification process, the number of temperature data points has been reduced from 2500 to 1400. The reduction of data points was achieved by excluding the data points within a 10m vertical distance from each other. Since most grid blocks have a height of 10m or more, the data reduction will cause no adverse effects. As described in Section 3.7, the data covariance matrix contains independent Gaussian noise with mean zero. We estimated the temperature measurement error to be $10 - 20\%$ of the measured data, resulting in an average error of $\sigma_{\text{d}} = 15°\text{C}$. Since the data is uncorrelated, we can formulate the data covariance matrix as $\boldsymbol{C}_{\text{d}} = \text{diag}(\boldsymbol{\sigma}_{\text{d}}^2)$.

The natural state's prior covariance matrix includes rock permeabilities and hot upflow rates because these parameters influence the geothermal reservoir's temperature distribution the most. In total, there are 1181 uncertain natural state model parameters. In order to generate production history samples, we take the uncertainty of the production and injection rates of 500 wells into account. Since it is not possible to compute the model sensitivity of the production history, the production history sample models are generated with the prior covariance matrix of the production and injection rates. In general, the prior covariance matrix can be formulated as:

$$\boldsymbol{C}_{\text{prior}} = \boldsymbol{\Sigma}_{\text{prior}} \cdot \boldsymbol{R} \cdot \boldsymbol{\Sigma}_{\text{prior}}^{\text{T}} \tag{24}$$

where $\boldsymbol{\Sigma}_{\text{prior}}$ is the prior standard deviation matrix of the parameters, i.e. $\boldsymbol{\Sigma}_{\text{prior}} = \text{diag}(\boldsymbol{\sigma}_{\text{prior}})$, and $\boldsymbol{R}$ the correlation matrix.

First, we consider the parameter uncertainty of the natural state model. The rock permeabilities of the RGF are inferred with magnetotellurics with values ranging from $\mathcal{O}(10^{-16} - 10^{-10}\ \text{m}^2)$. Since the permeabilities $k$ are on a logarithmic scale the standard deviation of the permeabilities is also logarithmic. As an example we look at the the prior permeability covariance matrix of a single rock type. In this study, the log-normal standard deviation of all permeabilities is taken as $\sigma_{\log(\text{k})} = 0.125$, the reason will be explained later in this section. The prior standard deviation of the permeability directions ($\text{k}_1$, $\text{k}_2$ and $\text{k}_3$) of every rock type is formulated as:

$$\boldsymbol{\Sigma}_{\text{k}} = \text{diag}(\boldsymbol{\sigma}_{\log(\text{k})}) = \begin{bmatrix} \sigma_{\log(\text{k}_1)} & 0 & 0 \\ 0 & \sigma_{\log(\text{k}_2)} & 0 \\ 0 & 0 & \sigma_{\log(\text{k}_3)} \end{bmatrix} \tag{25}$$

Typically, for the rock types defined in the model, the permeability directions are correlated. Correlation factors vary between 0-1, where 0 is no correlation and 1 is total correlation. The horizontal permeability directions ($\text{k}_1$ and $\text{k}_2$) have a high correlation (0.8), while the correlation between horizontal directions ($\text{k}_1$ and $\text{k}_2$) and vertical direction ($\text{k}_3$) is typically lower (0.5). The correlation matrix of the permeability directions of every rock type is described by:

$$\boldsymbol{R}_{\text{k}} = \begin{bmatrix} R_{\text{k}_{11}} & R_{\text{k}_{12}} & R_{\text{k}_{13}} \\ R_{\text{k}_{21}} & R_{\text{k}_{22}} & R_{\text{k}_{23}} \\ R_{\text{k}_{31}} & R_{\text{k}_{32}} & R_{\text{k}_{33}} \end{bmatrix} = \begin{bmatrix} 1.0 & 0.8 & 0.5 \\ 0.8 & 1.0 & 0.5 \\ 0.5 & 0.5 & 1.0 \end{bmatrix} \tag{26}$$

The deep hot upflow is distributed over multiple source blocks at the bottom layer of the model. These sources are taken as uncorrelated and with a standard variation of 10% of their manually calibrated value, explained later in this section. Since the upflow rates are uncorrelated, the prior covariance of the upflow rates can be defined by $\boldsymbol{C}_s = \text{diag}(\boldsymbol{\sigma}_s^2)$. With $\boldsymbol{\sigma}_s$ the standard deviations i.e., $\boldsymbol{\sigma}_{\mathbf{s}} = 0.10 \cdot \boldsymbol{\mu}_{\text{s}} = 0.10 \cdot [\mu_{\text{s},1}, \mu_{\text{s},2}, \ldots, \mu_{\text{s},\text{N}_\text{s}}]^T$. Where $\mu_{\text{s,i}}$ is the calibrated upflow rate of source block $i$ and $N_s$ is the total number of deep upflow source blocks.

Combining the prior covariance matrices of the permeabilities and upflow rates results in a total prior covariance matrix for the natural state $\boldsymbol{C}_{\text{prior}} = \text{diag}(\boldsymbol{C}_{\text{k}}, \boldsymbol{C}_{\text{s}})$. We can compute the posterior covariance by using the prior covariance, the data covariance and the model sensitivity, i.e., $\boldsymbol{C}_{\text{post}} \approx (\boldsymbol{S}^T \boldsymbol{C}_{\text{d}}^{-1} \boldsymbol{S} + \boldsymbol{C}_{\text{prior}}^{-1})^{-1}$. We expect that the posterior covariance will reduce the uncertainty of the model parameters. As explained previously, we use the posterior to generate sample model, i.e., $P(\boldsymbol{m}|\boldsymbol{d}) \approx \mathcal{N}(\boldsymbol{m}_{\text{MAP}}, \boldsymbol{C}_{\text{post}})$. In order to see if the parameter uncertainty has been reduced, the prior and posterior probability distribution of the parameters are compared.

As an example, we compare the prior and posterior of the permeability of rock type Q0001, as can be seen in Figure 29. On the diagonal of this figure is shown the marginal probability distribution of the permeability directions. The posterior marginal probability distributions show a significant width reduction compared to their prior, implying a reduced uncertainty. The blue scatter plots of the prior on the off-diagonal of the corner plot visualise the prior correlation of the permeability directions, as described in equation (26). The red scatter plots of the posterior on the off-diagonal show a reduction in size, however, the correlation has also decreased.

Nevertheless, this is still in line with our prior beliefs. For example, our prior belief is that the horizontal permeabilities are of the same order of magnitude, hence, a correlation factor of $\text{R}_{12} = 0.8$ (scatter plot left-centre). The application of the Bayesian framework reduced the correlation factor of the posterior to $\text{R}_{12} = 0.27$ (scatter plot top-centre), implying lower horizontal correlation. However, the posterior scatter plot (top-center) has reduced in size and still indicates that the horizontal permeabilities are of the same order of magnitude.

By applying the Bayesian framework, we achieved a significant uncertainty reduction for over 25% of the model parameters. The uncertainty reduction implies that we can, as expected, reduce the natural state model uncertainty by including the temperature data and the model sensitivity. Furthermore, we can conclude that the parameter uncertainty reductions are caused by a high density of temperature data points and/or a high parameter sensitivity. Parameter sensitivity is caused by the frequency and spatial location of the parameter.

For example, due to the high frequency and spatial location of rock type Q0001 (this rock type was assigned to 14,000 blocks out of a 94,000 total), the permeability uncertainty was reduced, as can be seen in Figure 29. The hot upflow source blocks are located at the bottom layer of the model at -1500m, while the temperature measurements are shallow at 0-300m. Furthermore, the source blocks only occur one time in the model. As a consequence, the upflow rates show no uncertainty reduction. Furthermore, rock types that are located far from the temperature measurements and/or have a low occurrence in the model show little to no uncertainty reduction of the permeability.

Large model alterations lead to run failures. Therefore, before numerous natural state models are simulated, the variance of the model parameters should be tested. Our first estimate for the standard deviation of the permeability and upflow rate was $\sigma_{\log(\text{k})} = 0.25$ and $\sigma_s = 0.20\mu_{\text{s}}$, respectively. These variations lead to no successful natural state simulations. Therefore, some tests were carried out to get an acceptable convergence ratio of the natural state simulations, while maintaining reasonable uncertainty bounds. A summary of the tested standard deviations is shown in Table 3. The standard deviation of the clay

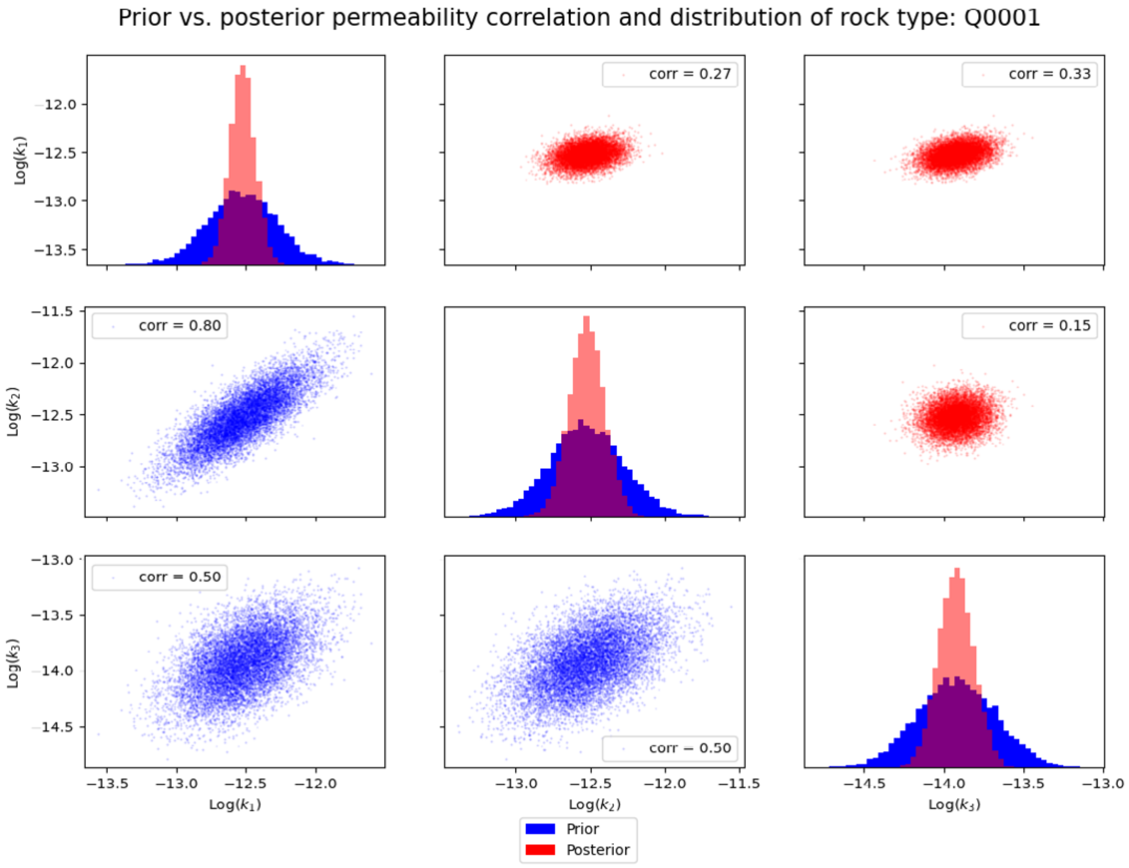


Figure 29: Corner plot of prior (blue) and posterior (red) probability distributions of the rock permeability of Q0001. On the diagonal is shown the marginal probability distributions for the permeability in each direction ($k_1$, $k_2$, and $k_3$), while on the off-diagonal are shown the two-dimension joint marginal probability distributions (which show the correlation).

cap's permeability is adjusted separately from the other rock types because the clay cap has low permeability and high sensitivity to the temperature results at the well locations. The various standard deviations were tested on 10 natural state sample models. A success ratio of 2 out of 10 simulations was acceptable (shown in the last column as "converged = yes"). Finally, the standard deviation of all permeabilities and the upflow rates was set to $\sigma_{\mathrm{log(k)}} = 0.125$ and $\sigma_s = 0.10\mu_{\mathrm{s}}$. respectively, as mentioned previously in this section. These standard deviations were used to compute the posterior probability distributions of the parameters. The posterior probability distributions were used to generate 226 sample models from the calibrated natural state model. The uncertainty of the natural state results will be discussed in the next section.

In order to generate production history sample models, the successful natural state

| Settings | $\sigma_{\log(\mathrm{k})}$ (except clay cap) | $\sigma_{\log(\mathrm{k})}$ (only clay cap) | $\sigma_s$ (upflow rates) | Converged (yes/no) |
|---|---|---|---|---|
| v1 | 0.5 | 0.5 | $0.2{\cdot}\mu_{\mathrm{s}}$ | no |
| v2 | 0.25 | 0.25 | $0.2{\cdot}\mu_{\mathrm{s}}$ | no |
| v3 | 0.25 | 0.25 | $0.1{\cdot}\mu_{\mathrm{s}}$ | no |
| v4 | 0.125 | 0.0 | $0.1{\cdot}\mu_{\mathrm{s}}$ | no |
| v5 | 0.125 | 0.0 | $0.05{\cdot}\mu_{\mathrm{s}}$ | yes |
| v6 | 0.125 | 0.0 | $0.0{\cdot}\mu_{\mathrm{s}}$ | yes |
| v7 | 0.125 | 0.05 | $0.05{\cdot}\mu_{\mathrm{s}}$ | yes |
| v8 | 0.125 | 0.125 | $0.05{\cdot}\mu_{\mathrm{s}}$ | yes |
| v9 | 0.125 | 0.125 | $0.1{\cdot}\mu_{\mathrm{s}}$ | yes |

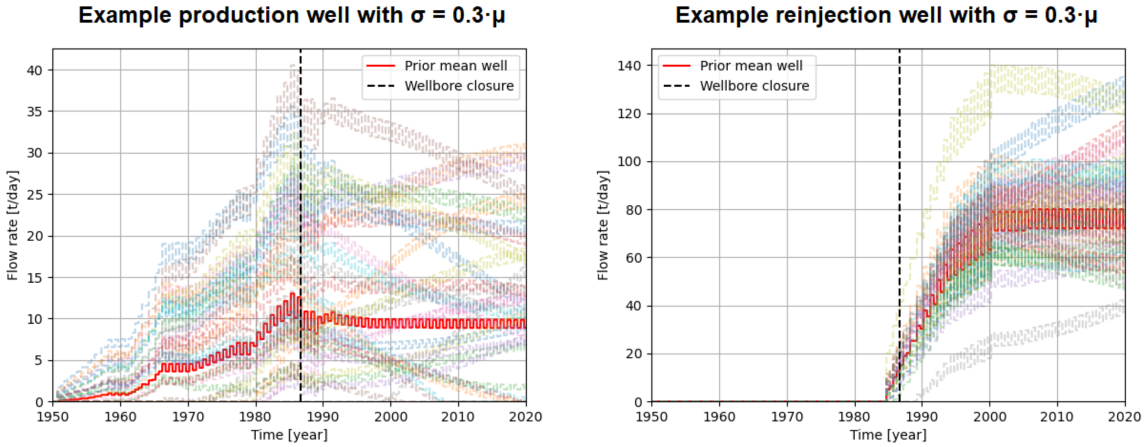
Table 3: Sample settings with various standard deviations for the model parameters

sample models are used, and the prior of the production and injection rates $P(\boldsymbol{\mu}) \approx \mathcal{N}(\boldsymbol{\mu}_{\mathrm{prod}}, \boldsymbol{C}_{\mathrm{prior}})$. Due to the absence of the production history sensitivity, we were not able to compute the posterior. The known requirements are that the production and reinjection flow rates have to stay positive and negative, respectively, and from the year 2005, the total reinjection had to be 90% of the total production (Gordon et al., 2005). The prior covariance matrix of the production and injection wells is constructed with the general equation (24).

In order to construct the prior covariance, we require the standard deviation of the standard deviation of the production and injection rates $\boldsymbol{\Sigma_{\mathrm{prod}}} = \mathrm{diag}(\boldsymbol{\sigma_{\mathrm{prod}}})$, and the correlation matrix $\boldsymbol{R}_{\mathrm{prod}}$ for every well. Since the production and injection rates of the individual wells were roughly estimated, the standard deviation of the flow rates is taken as 30% of the estimated flow rates, i.e., $\boldsymbol{\sigma_{\mathrm{prod}}} = 0.30 \cdot \boldsymbol{\mu}_{\mathrm{prod}} = 0.30 \cdot [\mu_{\mathrm{prod,t_1}}, \mu_{\mathrm{prod,t_2}}, \ldots, \mu_{\mathrm{prod,t_{N_t}}}]^T$, where $\mu_{\mathrm{prod,t_i}}$ is the flow rate at time-step $n$ and $N_t$ is the number of time-steps. The correlation matrix $\boldsymbol{R}_{\mathrm{prod}}$ encodes the smoothness of the production and injection rates by using the following formula:

$$\{\boldsymbol{R}_{\mathrm{prod}}\}_{\mathrm{ij}} = e^{-\frac{(t_{\mathrm{i}}-t_{\mathrm{j}})^2}{2l^2}} \tag{27}$$

where $l$ is the characteristic timescale. With the standard deviation and the correlation, we can now compute sample production and injection rates. Figure 30 shows the samples of a production and reinjection well. Here the red line is the estimated flow rate, the dashed black line the wellbore closure date and the semitransparent other lines are the generated sample flow rates. Furthermore, Figure 31 shows the cumulative production and reinjection rates of all the wells in Rotorua. The cumulative sample flow rates deviate only slightly from their estimated flow rates, which is caused by the summation of the sample flow rates that averages out the deviations of the individual wells. The total flow rates were estimated with more certainty than the flow rates of the individual wells, which corresponds with the results from Figure 31.



(a) (b)

Figure 30: An example production (left) and reinjection well (right) with the estimated flow rate in solid red line and the sample flow rates in semitransparent dashed lines. The vertical black dashed line is the wellbore closure date.

## 5.2 Uncertainty of natural state results

In the previous section, we reduced natural state parameter uncertainty by combining the temperature data and the natural state model sensitivity in a Bayesian approach. The results of 75 successful natural state samples are used to analyse the impact of reduced parameter uncertainty.

Figure 32 shows the uncertainty results of various temperature profiles at different geothermally active areas. The top left shows the well locations, the solid lines in the graphs are the calibrated model results, the grey lines are the results of the natural state sample models and the dots are the measured data points with 15°C error bars.

The well RR532 (orange) and RR335 (dark blue) show nearly perfect matches between temperature data and calibrated model results. However, the difference is that the model uncertainty of well RR532 is much lower than of well RR335. We can conclude that the uncertainty of the model results can be reduced by reducing model parameter uncertainty using a Bayesian framework. The model uncertainty reduction of well RR532 implies a high density of temperature measurements in the Kuirau Park area near the well and/or sensitive parameters are located near the well. In contrast, the model results of well RR335 show a higher uncertainty, which implies a lower density of temperature measurements and/or a low parameter sensitivity. Furthermore, we can conclude that while calibration improves model accuracy, it has little to no effect on the uncertainty of the natural state temperature results.

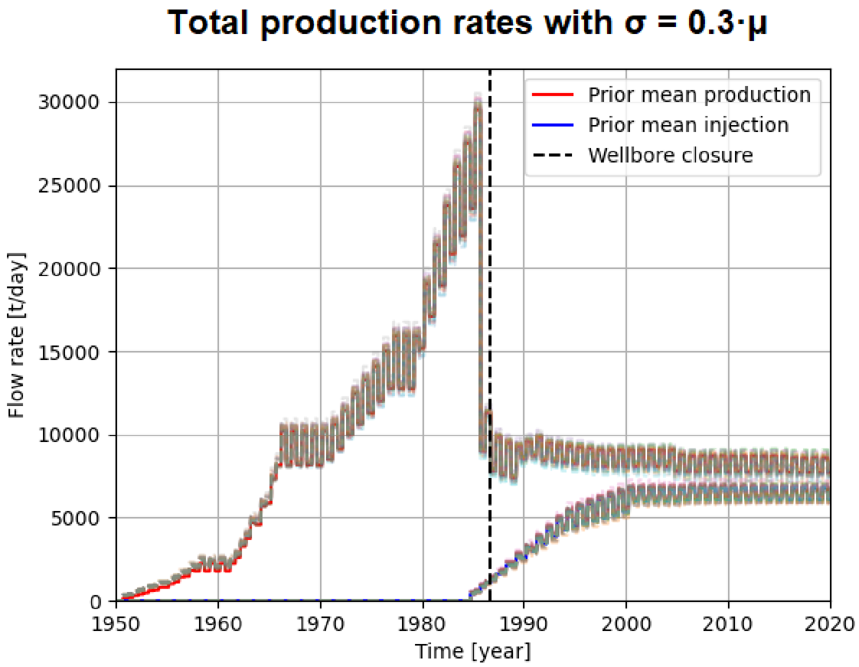


Figure 31: The cumulative sample production and reinjection flow rates of the wells in Rotorua.

The uncertainty results of the other wells indicate that further (manual or automatic) calibration and/or larger standard deviations of the model parameters is required to capture every temperature data point in the uncertainty bounds of the model results.

The uncertainty results of the surface temperature and mass outflow of the natural state are visualised in the map of Rotorua, as can be seen in Figure 33. The areas with high standard deviations of surface temperature (Figure 33b) generally correspond with the areas with high surface temperature from the calibrated model (Figure 24b from Chapter 4). However, mainly the Kuirau Park area shows a significantly low standard deviation (Figure 33b) compared to the high surface temperatures from the calibrated model (Figure 24b), indicating a low uncertainty. The low uncertainty of the surface temperature at Kuirau Park supports our previous conclusion that the model uncertainty is reduced by reducing parameter uncertainty. Furthermore, the standard deviation of the mass outflow rate (Figure 33c) generally shows a correlation with the magnitude of the mass outflow rate of the calibrated model (Figure 24c). No mass outflow uncertainty reduction can be observed, which can be explained by the unavailability of mass outflow data.

The main goal of geothermal modelling is to be able to predict future behaviour. For Rotorua, it is relevant to make predictions of surface features, since the surface features are of cultural and touristic importance. Therefore, we will compare the change in surface features and its uncertainty from 1950 to the predictions in 2040.

The geothermal surface features at four areas will be analysed and compared, namely, the surface temperature probability distribution at Whakarewarewa, the mass outflow probability distribution at Government Gardens, and the pressure probability distribution under Kuirau Park and the Pohutu Geyser. The standard deviations of the surface temperature (Figure 34a) and the mass outflow (Figure 34b) are 8.8% and 7.3% of their mean value, respectively. The standard deviation of the pressures under Kuirau Park and the Pohutu Geyser are 0.9% and 1.2% of their mean pressure, respectively. The pressure uncertainties are particularly low, implying that the natural state model pressure results are insensitive to model parameter uncertainties. The uncertainty percentages will be compared in the next section with the uncertainty results in 2020 (the present) and 2040 (the predictions).

### 5.3 Uncertainty of production history results and predictions

The 75 successful natural state sample model have been converted to production history sample models. The production history sample models have additional parameter uncertainty from the 500 production and injection wells. The sample models have been simulated from 1950 to 2040, where the results from 2020 to 2040 represent the model predictions.

The uncertainty results of Figure 33 and 35 indicate that the uncertainty of the surface features remains the same from 1950 to 2040. Therefore, we can conclude that the uncertain production and reinjection of the individual wells have a negligible effect on the uncertainty results and predictions of the surface features. This negligible effect can be explained by the feedzone location of the uncertain production and reinjection, being 100-300m below the surface, and the cumulative production and reinjection are relatively certain.

The surface temperature and mass outflow uncertainty results at Whakarewarewa and Government Gardens, respectively, are shown in Figure 35 for the years 2020 and 2040. We compare these results with the results from Figure 33, which shows a mean temperature increase of 0.2°C between 1950 and 2020, and a mean mass outflow reduction of 0.2 kg/s between 1950 and 2020, while the standard deviations remain constant. Furthermore, the mean and standard deviation of the temperature and mass outflow remain constant between 2020 and 2040.

The subsurface pressure probability distributions under Kuirau Park and the Pohutu Geyser show a similar negligible change in mean value and standard deviation. This can be observed by comparing Figure 34 and 37.

The standard deviation of the surface temperature, mass outflow and subsurface pressure remains nearly constant between 1950-2040, supporting the previous conclusion that the individual production and reinjection uncertainties do not affect the uncertainty of the surface features.

Figure 38 shows the transient pressure probability distribution plots of two monitoring wells, RR350 and RR886. The probability distribution is relatively wide, which indicates that the pressure results are relatively sensitive to the uncertain production and reinjection rates. The pressure results are taken relatively close to the feedzone, causing the relatively high sensitivity. Moreover, almost all pressure data points lie within the sample pressure distribution, implying that the production and reinjection uncertainties have been estimated correctly.

In conclusion, the uncertainty of the permeabilities and hot upflow rates cause the most uncertainty in the model results and predictions. The model and prediction uncertainty can be reduced by including the data points and the model sensitivity in a Bayesian approach. Furthermore, while the uncertain production and injection rates have a negligible effect on the surface features, they have a moderate effect on the pressures near the feedzones.

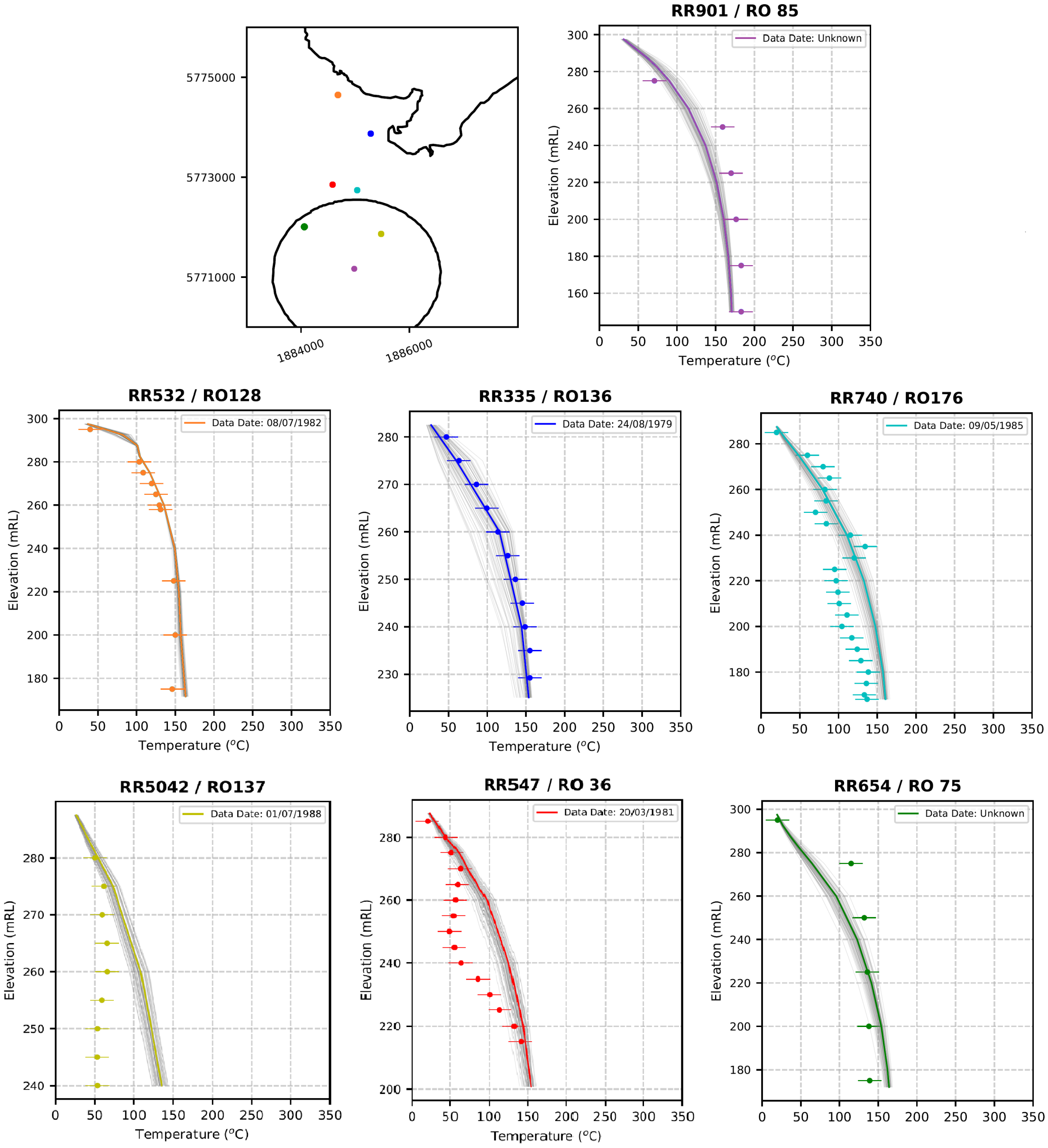
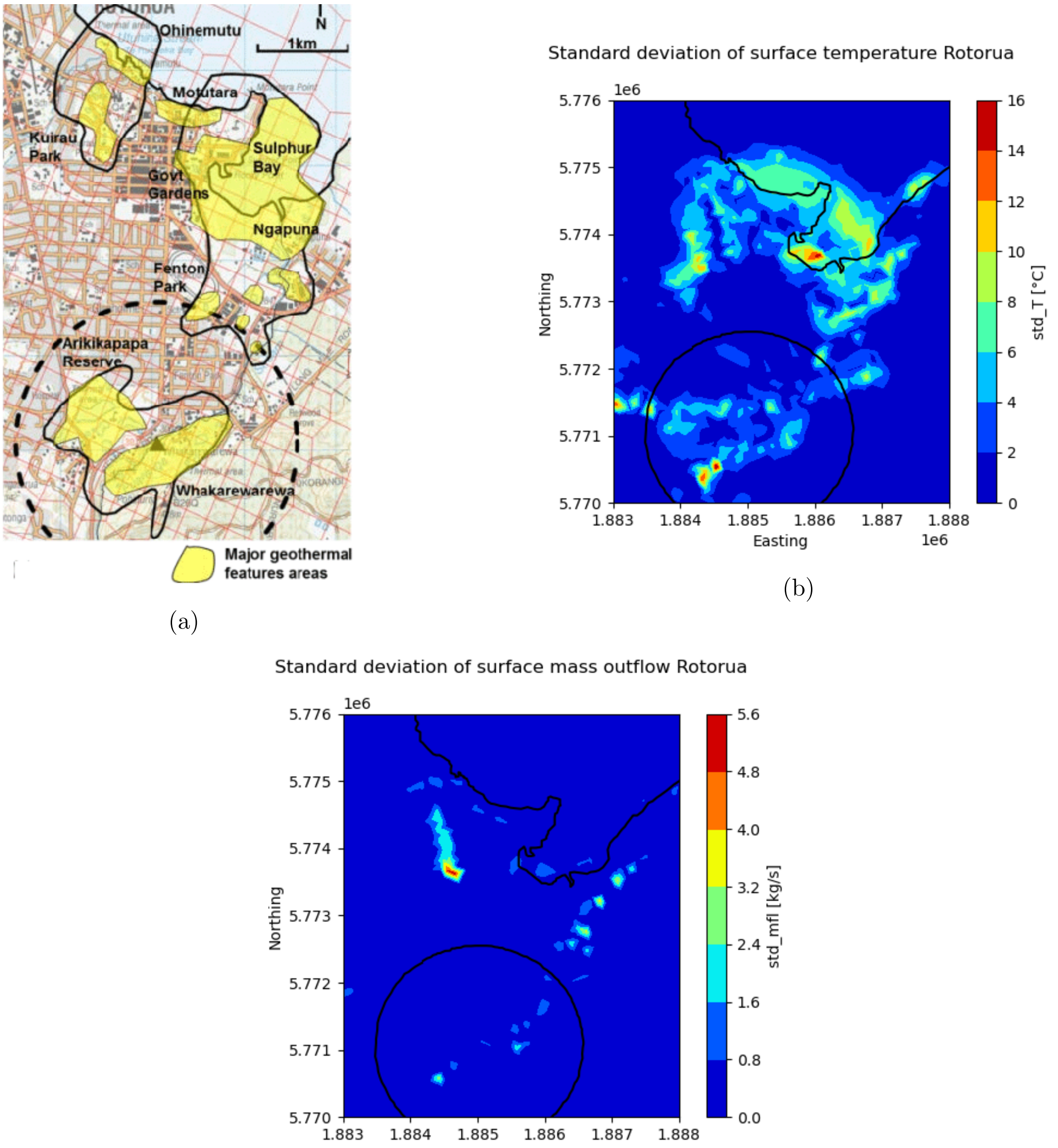


Figure 32: Uncertainty results of various NS temperature profiles from wells at different locations of the RGF. Top left shows the location of the wells: Arakikipakapa in green, Devon Street in turquoise, Government Gardens in dark blue, Kuirau Park in orange, Malfroy Road in red, Ngapuna in yellow and Whakarewarewa in purple. The measured temperatures are the dots with 15°C error bars, the solid lines are the calibrated model results and the lines in light grey are the sample model results.

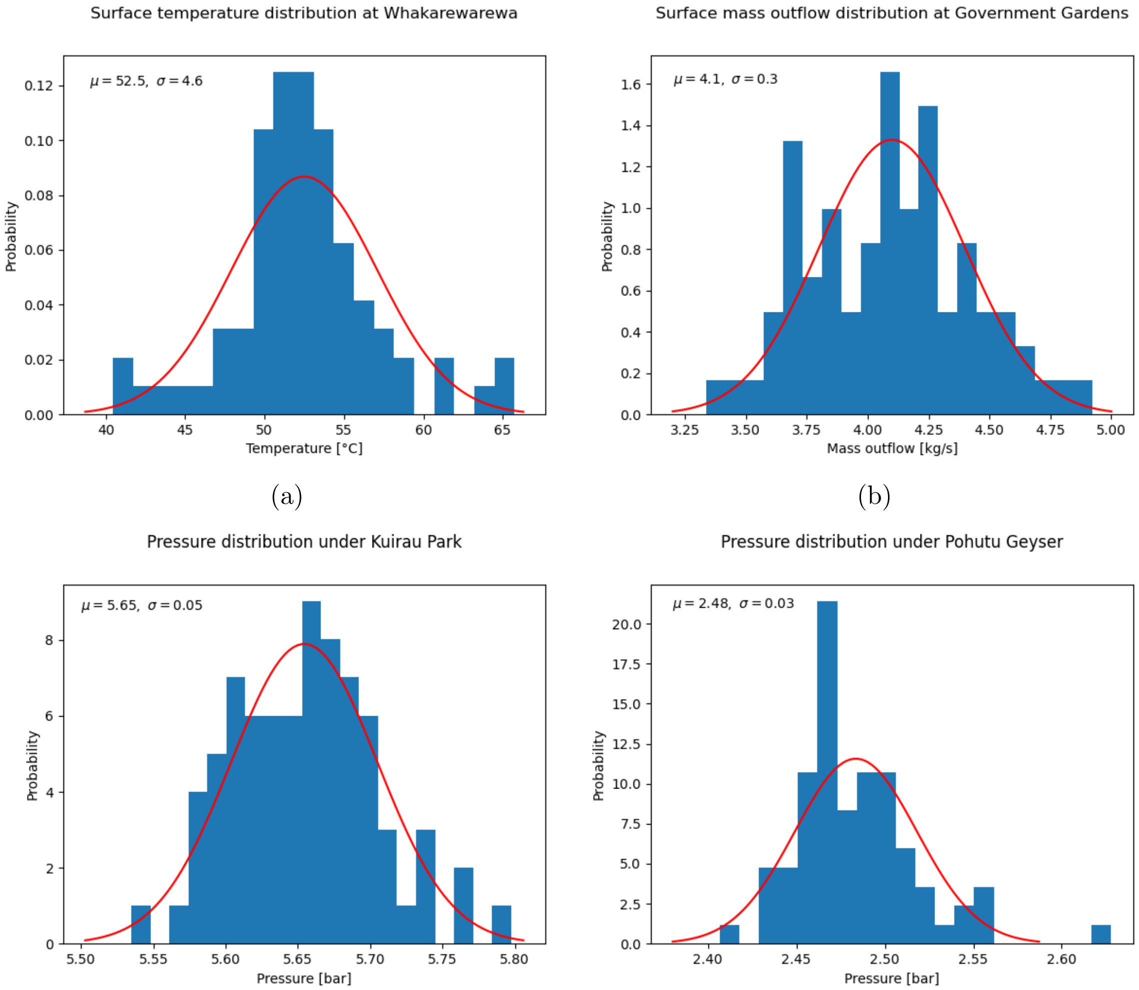


(a)

(b)

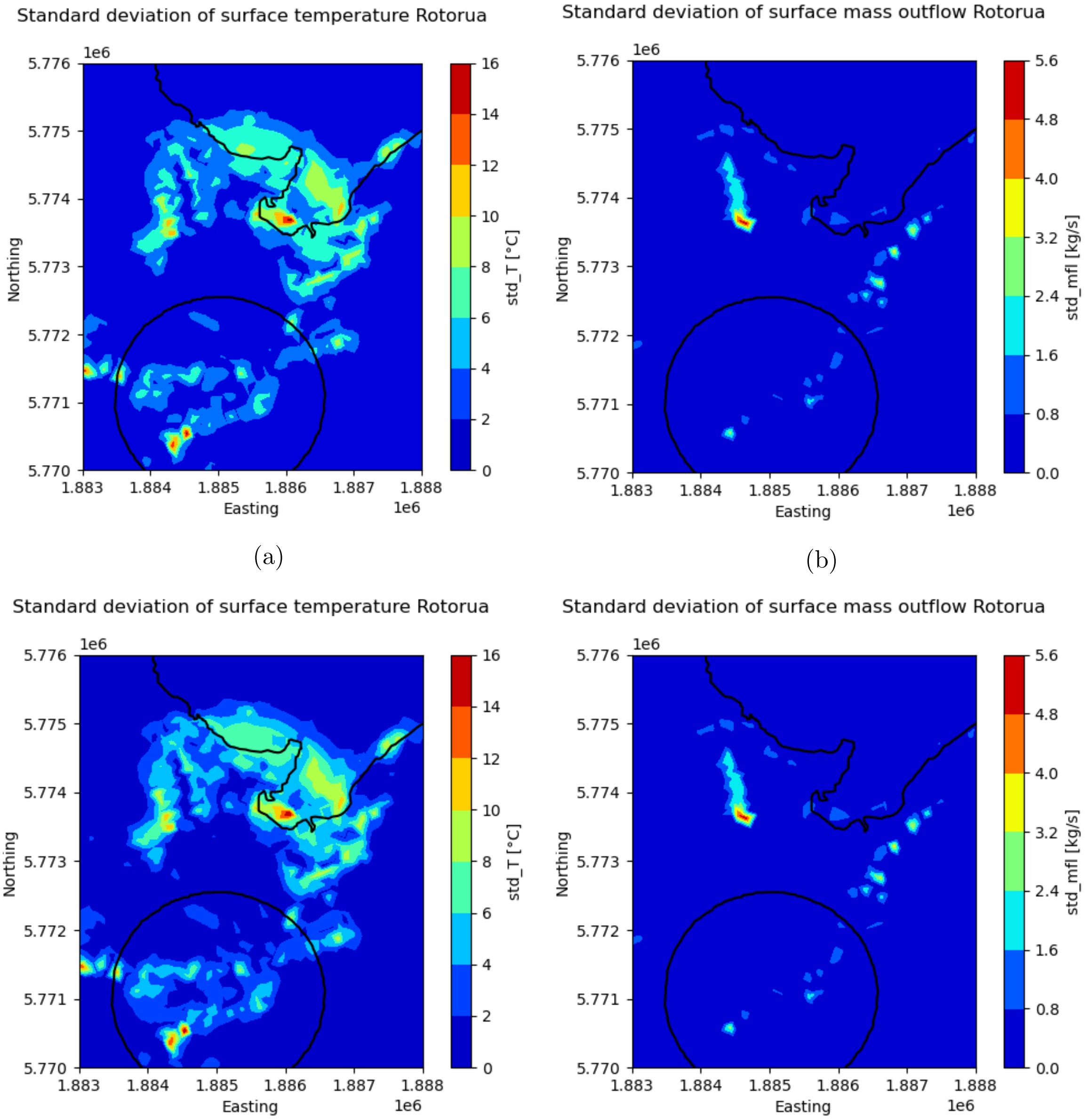
(c)

Figure 33: The thermally active areas indicated on the map of Rotorua (a) (Ratouis et al., 2014), the standard deviation of the surface temperature (b) and the mass outflow (c) of the (natural state) model at 1950.



(a) (b) (c) (d)

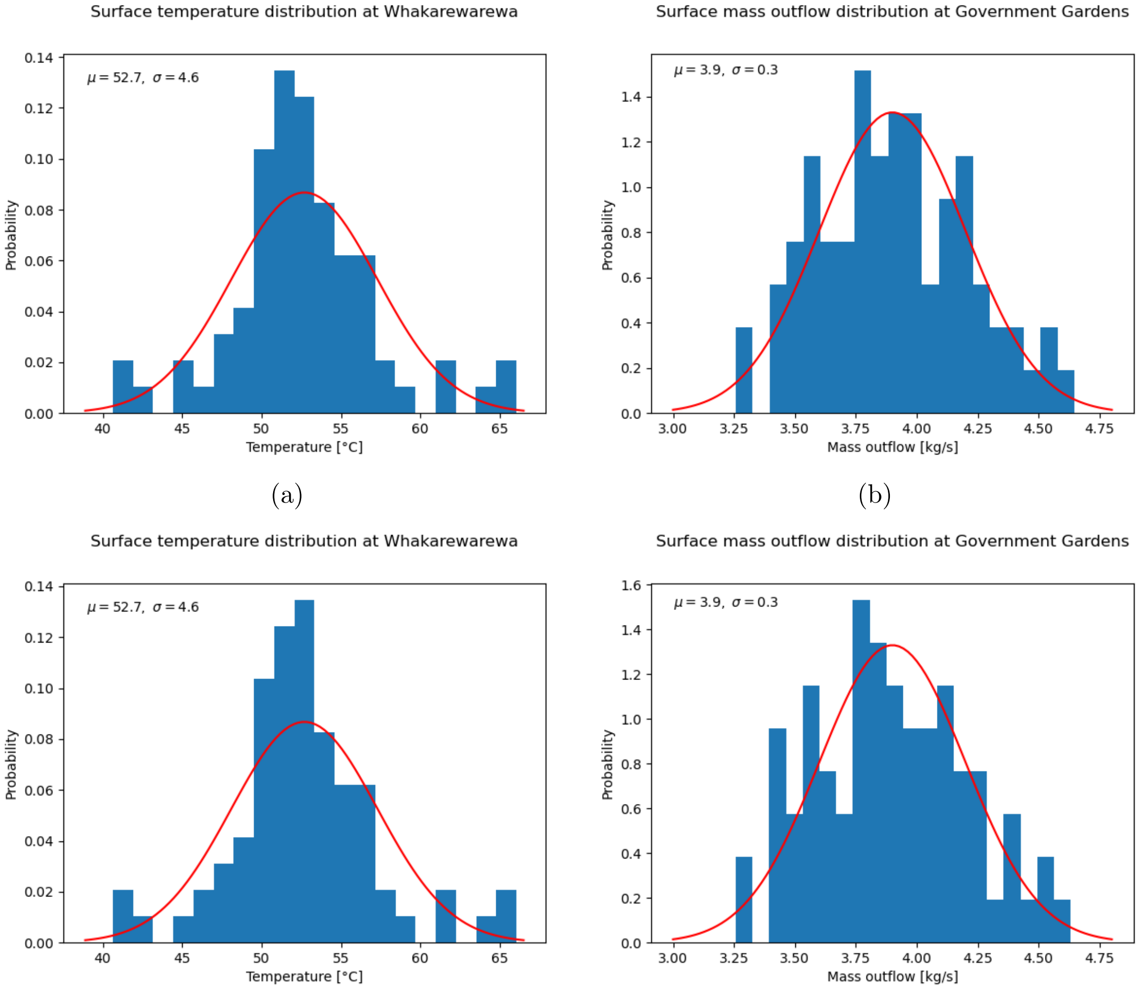
Figure 34: The surface temperature probability distribution at Whakarewarewa (a), the mass outflow probability distribution at Government Gardens (b), and the subsurface pressure probability distribution under Kuirau Park (c) and the Pohutu Geyser (d) of the natural state model (year 1950).



(a) (b)

(c) (d)

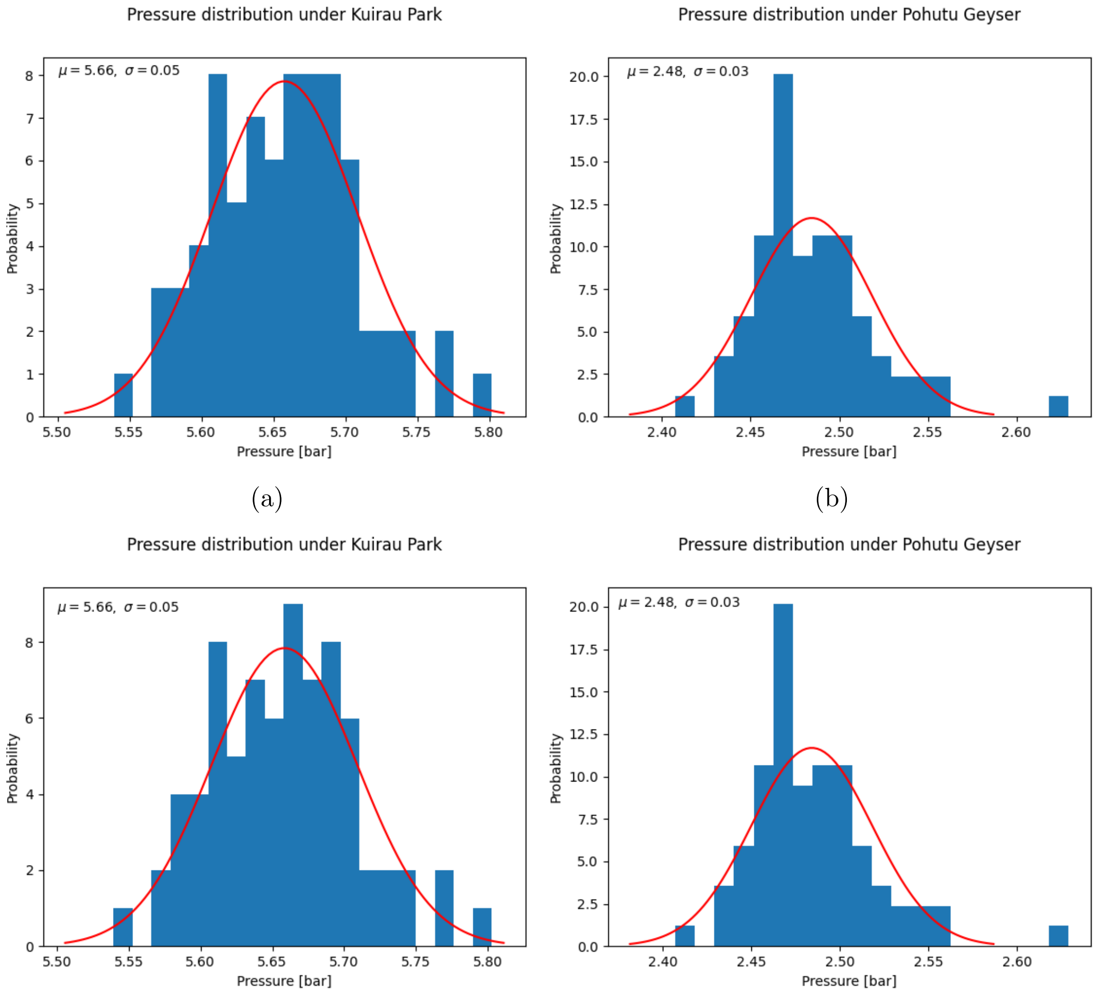
Figure 35: The standard deviation map of the surface temperature in 2020 (a) and 2040 (c), and the mass outflow in 2020 (b) and 2040 (d).



(a) (b)

(c) (d)

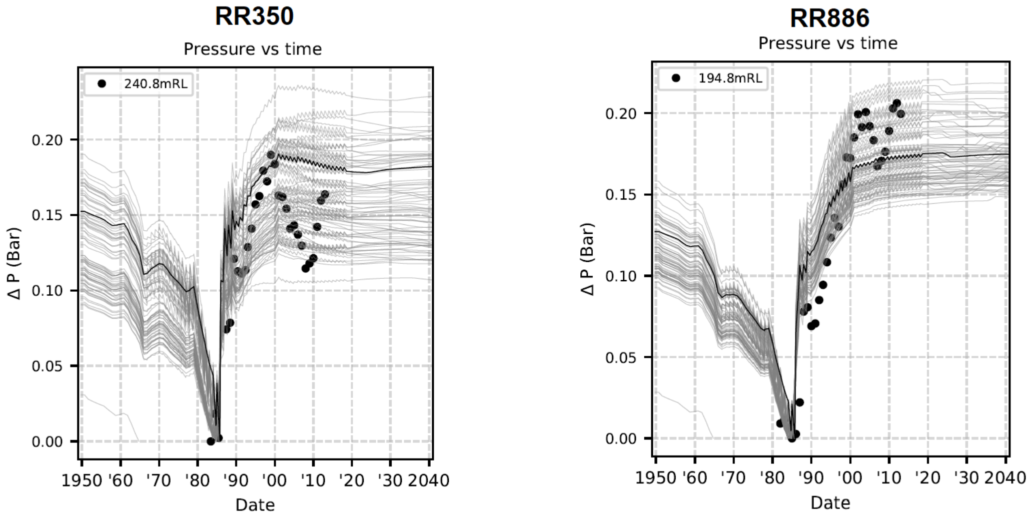
Figure 36: The surface temperature probability distribution at Whakarewarewa in 2020 (a) and 2040 (c), and the mass outflow probability distribution at Government Gardens in 2020 (b) and 2040 (d).



(a) (b)

(c) (d)

Figure 37: The pressure probability distributions under Kuirau Park in 2020 (a) and 2040 (c), and under the Pohutu Geyser in 2020 (b) and 2040 (d).



(a) (b)

Figure 38: Transient pressure sample distribution plots from well RR350 and RR886. The solid line is the calibrated model result, the dots are the data points and the light grey lines the sample results.

# 6 Conclusion and Recommendations

## 6.1 Conclusion

A geothermal reservoir is a promising renewable energy source. However, without careful management this energy source may lose its geothermal activity, with the results being irreversible. Geothermal models can predict future reservoir behaviour, providing helpful insights for sustainable reservoir management.

The Rotorua geothermal field has successfully been modelled in the past. However, a new Rotorua model has been developed, providing an improved fault structure and improved numerical accuracy due to a finer numerical grid (consisting of 94,701 blocks compared to 48,034). Furthermore, the new grid is made up of rectangular blocks, making it compatible with Waiwera (a recently developed geothermal simulator by Croucher (2020)). In this study, Waiwera is used to simulate the Rotorua model because Waiwera reduces computation times and has easy access to the model sensitivity. Shorter computation times speed up the manual calibration and uncertainty quantification process. The model sensitivity combined with the field data is used in a Bayesian framework to reduce the uncertainty of the model results and predictions.

Firstly, to develop an accurate model that makes accurate predictions, the new model needed to be calibrated with the available field data. The model was manually calibrated and showed a significantly improved match between the calibrated model results and the field data compared with the match from the initial model build by Van Vlijmen (2020). Furthermore, by calibrating the natural state model, the production history model was able to run successfully, while the initial model failed because the permeability structure was too tight.

The calibrated model predictions indicate that the current production and reinjection will be sustainable for the coming 20 years (until 2040). However, due to limited available field data, and uncertainties in the model and the field data, particularly the production data, the model predictions are also uncertain. Therefore, it is crucial to quantify and analyse the uncertainty of the model and the predictions.

Although further calibration will improve model accuracy, the model was not further manually calibrated due to time limitations. Moreover, while typically manual calibration is followed by automatic calibration to improve the model accuracy further, this was not yet possible with iWaiwera (the automatic calibration software package from Waiwera) due to technical development issues.

The uncertainty of the model results and predictions depends on the model parameter uncertainty and the noise in the measurements. The natural state model contains 1181 uncertain parameters, and the production history contains 500 uncertain production and reinjection wells. Typically, for geothermal reservoirs, the production and injection rates are well-known, however, in the case of Rotorua, there are only reasonable estimates of the total production and injection rates.

The 1181 natural state parameters and the production and injection rates of the 500 wells were varied to generate 226 sample models. These sample models were simulated simultaneously using the supercomputer NeSI (*New Zealand eScience Infrastructure*, 2021), and 75 models succeeded (a success rate of 33%). By using the model sensitivity (obtained from Waiwera) and the field data in a Bayesian framework, we were able to reduce the uncertainty of roughly 20% of the parameters. The decreased parameter uncertainty caused a reduction of the uncertainty of the model results, particularly of the temperature results in the Kuirau Park area. The reduced uncertainty in Kuirau Park is caused by a high density of temperature measurements and/or a high sensitivity of the model parameters in that area. Nevertheless, a significant number of temperature measurements are not within the uncertainty bounds of the natural state temperature results. Further calibration can improve the match between model temperature results and measurements, and larger variances of the model parameters can provide wider uncertainty bounds of the model temperature results.

The uncertain production and injection rates have a negligible impact on the uncertainty of the surface features. The impact is negligible due to the considerable distance between the feedzones and the surface (100-200m) and/or the large number of wells that average out the uncertain production and injection rates of the individual wells. If we look at a greater depth near the feedzones of the wells, the uncertain production and injection rates show a moderate impact on the uncertainty of the transient pressure results and predictions. The moderate impact is expected since the pressure results and predictions are taken closer to the source of uncertainty, the feedzone.

The model predictions imply that with the current production and reinjection at the RGF, the surface features will be maintained for the next 20 years (until 2040). The uncertainty of the pressure and mass outflow predictions is relatively low, while the uncertainty of the temperature predictions is moderately high in areas with high temperatures. In conclusion, prediction uncertainty can be reduced with the Bayesian framework, and uncertainty quantification provides helpful insights for sustainable reservoir management. Furthermore, the current use of the reservoir is sustainable, however, particularly surface temperatures should remain closely monitored.

## 6.2 Recommendations and future work

In this study, we successfully predicted future behaviour of the RGF with reasonable certainty. However, the uncertainty bounds of the model results do not match with a considerable number of data points, and only one future scenario is predicted.

In future work, it is vital to include the majority of the data points within the uncertainty bounds of the model results because this will increase the credibility of the model predictions. Including more data points can be accomplished by further calibration or increasing the width of the uncertainty bounds. Further calibration includes manual and automatic calibration. In this study, further manual calibration was not achieved due to time limitations, and automatic calibration was not possible due to technical development issues of iWaiwera. If the iWaiwera issues remain unsolved, we should consider to convert the Waiwera model back to an AUTOUGH2 model and explore the feasibility of automatic calibration with AUiTOUGH2.

Additionally, if further calibration does not achieve the desired results, the variance of the model parameters can be increased to create wider uncertainty bounds of the model results. The parameter variance should be increased in small increments until wide enough uncertainty bounds of the model results are accomplished. The small increments are necessary to prevent run failures.

Additionally, the current parameter variances could be done in small increments to achieve a higher success ratio of the sample simulations. However, this process will significantly increase computation times.

Finally, in this study, only one future scenario was predicted by extending the current production and reinjection for 20 years. Different future scenarios will provide valuable insights. For example, we can explore the possibility of increasing net production while sustaining the reservoir's geothermal activity.

# 7 Acknowledgements

I would like to thank the Geothermal Institute of the University of Auckland for giving me the opportunity to do this research using their resources and facilities. Furthermore, I would like to thank Michael O'Sullivan, Michel Speetjens and Michael Gravatt for their supervision and Ru Nicholson, Oliver Maclaren, John O'Sullivan and Susana Guzman for their advice and support.

I wish to acknowledge the contribution of NeSI to the results of this research. New Zealand's national compute and analytics services and team are supported by the New Zealand eScience Infrastructure (NeSI) and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation and Employment. (*New Zealand eScience Infrastructure*, 2021)

Thanks to Seequent for the license of Leapfrog Geothermal (*Leapfrog Geothermal - Seequent*, n.d.).

# References

Acland, J., & Molloy, L. (2006). Our world heritage: a tentative list of New Zealand cultural and natural heritage sites. *A Report to the Department of Conservation by the Cultural and Natural Heritage Advisory Groups, Department of Conservation.*

Anderson, M. P. (2005, 11). Heat as a Ground Water Tracer. *Groundwater*, *43*(6), 951–968. doi: 10.1111/j.1745-6584.2005.00052.x

ARANZ Geo Limited. (2015). *User Manual for Leapfrog Geothermal version 2.8* (Tech. Rep.).

Bernardo, J. M., & Adrian F. M. Smith. (2009). Bayesian Theory . *John Wiley & Sons*, *405*.

Bjarkason, E. K., O'Sullivan, J. P., Yeh, A., & O'Sullivan, M. J. (2019, 3). Inverse modeling of the natural state of geothermal reservoirs using adjoint and direct methods. *Geothermics*, *78*, 85–100. doi: 10.1016/j.geothermics.2018.10.001

Bravo, H. R., Jiang, F., & Hunt, R. J. (2002, 8). Using groundwater temperature data to constrain parameter estimation in a groundwater flow model of a wetland system. *Water Resources Research*, *38*(8), 28–1. doi: 10.1029/2000WR000172

Butcher, G., Fairweather, J. R., & Simmons, D. G. (2000). The economic impact of tourism on Rotorua.

Croucher, A. (2020). *Waiwera user guide Release 1.1.0* (Tech. Rep.).

Davidsdottir, B. (2012, 1). Sustainable energy development: The role of geothermal power. In *Comprehensive renewable energy* (Vol. 7, pp. 273–297). Elsevier Ltd. doi: 10.1016/B978-0-08-087872-0.00715-0

Environment Bay of Plenty (EBOP). (1999). Rotorua Geothermal Regional Plan. *Resource Planning Publication 99/02*(ISSN 1170 9022).

Finsterle, S. (2007). *iTOUGH2 User's Guide* (Tech. Rep.).

Gelman, A., Carlin, J., Stern, H., Dunson D.B., Vehtari, A., & Rubin, D. (2013). *Bayesian Data Analysis* (Third Edition ed.; Chapman & Hall/CRC Texts in Statistical Science, Ed.). Taylor & Francis.

Gonzalez-Gutierrez, B., Nicholson, R., O'sullivan, J. P., Yeh, A., Bjarkason, E., O'sullivan, M. J., & Maclaren, O. J. (2020). Accelerating the Solution of Geothermal Inverse Problems Using Adjoint Methods in Waiwera: a Case Studies on Kerinci and Wairakei.

Gordon, D. A., Scott, B. J., & Mroczek, E. K. (2005). *Rotorua geothermal field management monitoring update.* Environmental Publication.

Grant, M. (2013). *Geothermal reservoir engineering.* Elsevier.

Huenges, E. (2016). Enhanced geothermal systems: Review and status of research and development. In R. DiPippo (Ed.), *Geothermal power generation* (pp. 743–761). Woodhead Publishing. Retrieved from `https://www.sciencedirect.com/science/article/pii/B9780081003374000255` doi: https://doi.org/10.1016/B978-0-08-100337-4.00025-5

Jiang, Y., & Woodbury, A. D. (2006, 12). A full-Bayesian approach to the inverse problem for steady-state groundwater flow and heat transport. *Geophysical Journal International*, *167*(3), 1501–1512. doi: 10.1111/j.1365-246X.2006.03145.x

Kaipio, J. P., & Somersalo, E. (2007). Statistical inverse problems: discretization, model reduction and inverse crimes. *Journal of computational and applied mathematics*, *198*(2), 493–504.

Kaipio, S. E., J.P. (2006). Statistical and computational inverse problems. *Springer Science & Business Media*, *160*.

*Leapfrog Geothermal - Seequent.* (n.d.). Retrieved from `https://www.seequent.com/products-solutions/leapfrog-geothermal/`

Lloyd, E. (1979). Report on Probable Effects of Exploitation of Rotorua Geothermal Field on the Natural Activity. *Geothermal Circular No. EFL-6, NZ geological Survey, unpublished report, 14 pp.*

Maclaren, O. J., Bjarkason, E., O'sullivan, J., & O'sullivan, M. J. (2016). *INVERSE MODELLING OF GEOTHERMAL RESERVOIRS-A HIERARCHICAL BAYESIAN APPROACH* (Tech. Rep.).

Maclaren, O. J., Nicholson, R., Bjarkason, E. K., O'Sullivan, J. P., & O'Sullivan, M. J. (2020, 10). Incorporating Posterior-Informed Approximation Errors into a Hierarchical Framework to Facilitate Out-of-the-Box MCMC Sampling for Geothermal Inverse Problems and Uncertainty Quantification. Retrieved from `http://arxiv.org/abs/1810.04350`

Ministry of Energy. (1985). The Rotorua geothermal field. *In: A report of the Geothermal Monitoring Programme and Task Force 1982-1985. Ministry of Energy..*

Neilson, G., Bignall, G., & Bradshaw, D. (2010). Whakarewarewa a living thermal village–Rotorua, New Zealand. In *Proceedings world geothermal congress* (pp. 1–7).

New Zealand Geothermal Association. (2021). *New Zealand Geothermal Association — Geothermal Heat Pumps - New Zealand Geothermal Association.* Retrieved from `https://nzgeothermal.org.nz/heat-pumps-new-zealand-geothermal-association/geothermal-heat-pumps/`

*New Zealand eScience Infrastructure.* (2021). Retrieved from `https://www.nesi.org.nz/`

Nocedal, J., & Wright, S. (2006). *Numerical optimization.* Springer Science & Business Media.

O'Sullivan, M. (2009). *Types of Geothermal Systems [Powerpoint slides].* University of Auckland.

O'Sullivan, M., & O'Sullivan, J. (2016, 1). Reservoir modeling and simulation for geothermal resource characterization and evaluation. *Geothermal Power Generation*, 165–199. Retrieved from `https://www.sciencedirect.com/science/article/pii/B9780081003374000073?via%3Dihub` doi: 10.1016/B978-0-08-100337-4.00007-3

O'Sullivan, M., Pruess, K., & Lippmann, M. (2001, 11). State of the art geothermal reservoir simulation. *Geothermics*, *30*, 395–429. doi: 10.1016/S0375-6505(01)00005-0

Paksoy, H., & Beyhan, B. (2014, 10). Thermal energy storage (TES) systems for greenhouse technology. In *Advances in thermal energy storage systems: Methods and applications* (pp. 433–548). Elsevier Inc. doi: 10.1533/9781782420965.4.533

Pruess, K. (2004, 8). The TOUGH Codes—A Family of Simulation Tools for Multiphase Flow and Transport Processes in Permeable Media. *Vadose Zone Journal*, *3*(3), 738–746. Retrieved from `https://doi.org/10.2136/vzj2004.0738` doi: 10.2136/vzj2004.0738

Pruess, K., Oldenburg, C., & Moridis, G. (1999). *ERNEST ORLANDO LAWRENCE E3ERKE˜EY NATIONAL LABORATORY TOUGH2 User's Guide, Version2.0* (Tech. Rep.).

Rath, V., Wolf, A., & Bücker, H. M. (2006, 10). Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on automatic differentiation: sensitivity calculation, verification, and synthetic examples. *Geophysical Journal International*, *167*(1), 453–466. doi: 10.1111/j.1365-246X.2006.03074.x

Ratouis, T. M. P., O'Sullivan, M. J., Alcaraz, S. A., & O'Sullivan, J. P. (2017, 9). The effects of seasonal variations in rainfall and production on the aquifer and surface features of Rotorua geothermal field. *Geothermics*, *69*, 165–188. doi: 10.1016/j.geothermics.2017.05.003

Ratouis, T. M. P., O'Sullivan, M. J., & O'Sullivan, J. (2014). MODELLING OF THE ROTORUA GEOTHERMAL FIELD INCLUDING CHEMISTRY.

Ratouis, T. M. P., O'Sullivan, M. J., & O'Sullivan, J. P. (2016a). MODELLING GEOTHERMAL SURFACE FEATURES IN THE ROTORUA GEOTHERMAL FIELD.

Ratouis, T. M. P., O'Sullivan, M. J., & O'Sullivan, J. P. (2016b). A Numerical model of Rotorua Geothermal Field. *Geothermics*, *60*, 105–125. doi: 10.1016/j.geothermics.2015.12.004

Reuss, M. (2015, 1). The use of borehole thermal energy storage (BTES) systems. In *Advances in thermal energy storage systems: Methods and applications* (pp. 117–147). Elsevier Inc. doi: 10.1533/9781782420965.1.117

Saar, M. O. (2011). Review: Geothermal heat as a tracer of large-scale groundwater flow and as a means to determine permeability fields. *Hydrogeology Journal*, *19*(1), 31–52. doi: 10.1007/s10040-010-0657-2

Scott, B. J., & Cody, A. D. (1997). EFFECTS OF BORE CLOSURE AT ROTORUA, NEW ZEALAND.

Stuart, A. (2010). Inverse problems: a Bayesian perspective. *Acta numerica, 19, 451.*

Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation. *Society for Industrual and Applied Mathermatics.*

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual.* CreateSpace.

Van Vlijmen, M. J. M. (2020). *Implementing a new model for Rotorua in AUTOUGH2 and comparing results* (Tech. Rep.). Eindhoven University of Technology.

Vogt, C., Mottaghy, D., Rath, V., Marquart, G., Dijkshoorn, L., Wolf, A., & Clauser, C. (2014, 9). Vertical variation in heat flow on the Kola Peninsula: palaeoclimate or fluid flow? *Geophysical Journal International*, *199*(2), 829–843. doi: 10.1093/gji/ggu282

Vogt, C., Mottaghy, D., Rath, V., Wolf, A., Pechnig, R., & Clauser, C. (2010). *Quantifying Uncertainty in Geothermal Reservoir Modeling* (Tech. Rep.).

Woodbury, A. D., & Smith, L. (1988, 3). Simultaneous inversion of hydrogeologic and thermal data: 2. Incorporation of thermal data. *Water Resources Research*, *24*(3), 356–372. doi: 10.1029/WR024i003p00356

Yeh, A., Croucher, A. E., & O'sullivan, M. J. (2013). *TIM-YET ANOTHER GRAPHICAL TOOL FOR TOUGH2* (Tech. Rep.).

Zhang, Z., Jafarpour, B., & Li, L. (2014, 6). Inference of permeability heterogeneity from joint inversion of transient flow and temperature data. *Water Resources Research*, *50*(6), 4710–4725. doi: 10.1002/2013WR013801

# A Data Structure

## A.1 Waiwera input and output

In this study, Waiwera is used to simulate the geothermal reservoir of Rotorua. In order to simulate the natural state model, there are three input files required:

- The natural state model file (JSON format) contains information like simulation settings, model name, parameter values, boundaries etc.
- The initial condition file (H5 format), which contains the initial state of the model.
- The geometry file (MSH format), which contains the 3D grid of the model.

At the end of a natural state simulation, two files are created:

- The results file (H5 format) contains the natural results, e.g. temperature, pressure and mass flows, of the initial state and steady state (end time-step).
- The log file (YAML format) contains the log of the whole simulation, which can be valuable for run failures.

For the production history simulation Waiwera requires four inputs:

- The production history model file (JSON format) contains information like simulation settings, model name, parameter values, boundaries etc.
- The initial condition file (H5 format), which contains the initial state of the model, the final state of the natural state simulation.
- The geometry file (MSH format), which contains the 3D grid of the model.
- The well data library, containing a JSON file for each well (including the well location, production rates, injection enthalpy etc.). The structure of such a JSON file with well data is presented in an example in Appendix A.2.

The output of the production history simulation creates two files:

- The results file (H5 format) contains the production history results, e.g. temperature, pressure and mass flows, at every time step of the simulation.
- The log file (YAML format) contains the log of the whole simulation, which can be valuable for run failures.

## A.2 Well data

A JSON dictionary for well RR724 shows how a JSON data file is build up and what information it contains.

```
{
  "Well_Name": "RR724",
  "AUT2_Name": "RO125",
  "Well_Track": {
    "Units": {
      "x": "m",
      "y": "m",
      "z": "masl"
    },
    "Data": [
      "[1884281.9, 5770931.2, 310.0]",
      "[1884281.9, 5770931.2, 154.0]"
    ]
  },
  "Downhole_Temperature": {
    "09/04/1985": {
      "Units": {
        "Date": "dd/mm/yyyy",
        "Elevation": "mRL",
        "Temperature": "degC"
      },
      "Source": "RGF_Temperature_data_24_04_2017_deletebad.xlsx",
      "Comment": "",
      "Data": {
        "290.0": 56.5,
        "285.0": 55.5,
        "280.0": 59.5,
        "275.0": 65.6,
        "270.0": 73.0,
        "265.0": 81.1,
        "260.0": 92.4,
        "255.0": 101.0,
        "250.0": 105.6,
        "245.0": 114.3,
        "240.0": 123.4,
        "235.0": 132.0,
        "230.0": 141.7,
        "225.0": 159.1,
```

```
        "220.0": 156.6,
        "215.0": 155.9,
        "210.0": 156.3,
        "205.0": 157.7,
        "200.0": 156.6,
        "195.0": 156.8,
        "190.0": 159.3,
        "185.0": 159.6,
        "180.0": 159.5,
        "175.0": 159.8,
        "170.0": 160.0,
        "165.0": 157.4,
        "160.0": 158.0,
        "155.0": 171.3,
        "154.0": 175.1
      },
      "StableFlag": true
    }
  },
  "Comments": {
    "cfg_Names": [
      "SP019"
    ],
    "Monitor_Name": [
      "M17"
    ],
    "Consent_Numbers": [
      "NA"
    ]
  },
  "Feed_Zones": {
    "Units": {
      "Top": "m",
      "Bottom": "m",
      "Quality": "Text",
      "Source": "Text"
    },
    "Data": [
      "[226.5, 158.2, '', 'SP019_seasonal.cfg']"
    ]
  },
  "Production_Generator_Name": "P 125",
```

```
"Production_Mass_Flow": {
  "Units": {
    "Date": "dd/mm/yyyy",
    "Massflow": "t/hr"
  },
  "Source": "SP019_seasonal.cfg",
  "Comment": "",
  "Data": {
    "05/10/1949": 0.0,
    "01/01/1950": 0.0,
    "29/03/1950": 0.0,
    "05/04/1950": 0.0,
    "02/07/1950": 0.0,
    "28/09/1950": 0.0,
    "05/10/1950": 0.0196632,
    "01/01/1951": 0.0196632,
    "29/03/1951": 0.0196632,
    "05/04/1951": 0.0245772,
    "02/07/1951": 0.0245772,
    "28/09/1951": 0.0245772,
    "05/10/1951": 0.0393156,
    "01/01/1952": 0.0393156,
    "28/03/1952": 0.0393156,
    "05/04/1952": 0.0491436,
    "02/07/1952": 0.0491436,
    "27/09/1952": 0.0491436,
    "05/10/1952": 0.058975200000000005,
    "01/01/1953": 0.058975200000000005,
    "29/03/1953": 0.058975200000000005,
    "05/04/1953": 0.0737208,
    "02/07/1953": 0.0737208,
    "28/09/1953": 0.0737208,
    "05/10/1953": 0.0786384,
    "01/01/1954": 0.0786384,
    "29/03/1954": 0.0786384,
    "05/04/1954": 0.098298,
    "02/07/1954": 0.098298,
    "28/09/1954": 0.098298,
    "05/10/1954": 0.09829080000000001,
    "01/01/1955": 0.09829080000000001,
    "29/03/1955": 0.09829080000000001,
    etc.
```

```
    }
  },
  "Transient_Pressure": {
    "309.8": {
      "Units": {
        "Elevation": "mRL",
        "Date": "dd/mm/yyyy",
        "Pressure": "bar a"
      },
      "Source": "MWells_GBores.xlsx",
      "Comment": "relative pressure change",
      "Data": {
        "01/01/2006": 0.15038078797975893,
        "01/01/2007": 0.0,
        "01/01/2008": 0.19900066641984712,
        "01/01/2009": 0.3206671647411372
      }
  }
}
```

# B Waiwera

Waiwera is a new numerical simulator for geothermal systems. This free open-source software package was developed by the Geothermal Institute of the University of Auckland in collaboration with GNS Science (Croucher, 2020). Waiwera is based on simulation software like TOUGH2 and also uses the PyTOUGH library from Python for simulations. Furthermore, Waiwera incorporates some major advantages over TOUGH2. The simulations are executed in parallel on shared- or distributed-memory computers and clusters, saving computation time. PETSc (Portable Extensible Toolkit for Scientific Computation) is included for parallel data structures, linear and nonlinear solvers (Croucher, 2020).

Additionally, an automatic calibration software iWaiwera, as part of Waiwera, is under development. iWaiwera uses the adjoint method for inverse modelling, making the computation time considerably shorter. Waiwera provides easy access to the model Jacobian that is essential for the adjoint method. As for convergence of the Newton-Raphson iterations, Waiwera creates dimensionless residuals, so convergence is reached faster.

# C iWaiwera

Geothermal models are highly nonlinear and complex. Furthermore, there is only limited field data available, creating an ill-posed inverse problem. Therefore, software like iTOUGH2 and PEST is used to perform automatic calibration. iTOUGH2 and PEST use the finite differencing method to calibrate a geothermal reservoir automatically. This method requires the model to run forward simulations over and over again, changing the model parameters $N_m$ after every simulation (Bjarkason et al., 2019). This is a very time-consuming process and can take up days or even weeks for fine numerical models to find a best-fit solution.

Although iWaiwera encountered problems when attempting to automatically calibrate the Rotorua model, we discuss the methodology here because iWaiwera has a great automatic calibration potential. iWaiwera uses the adjoint method and has an easy access to the model Jacobian. By combining the adjoint method and the model Jacobian, iWaiwera chooses smart parameter values, avoiding numerous computationally expensive simulations and saving a significant amount of computation time (Gonzalez-Gutierrez et al., 2020). In general the inverse problem can be described as an optimisation problem by the following equations:

$$\begin{aligned} \text{min} \quad & f(\boldsymbol{y}, \boldsymbol{m}) \\ \text{such that} \quad & \boldsymbol{g}(\boldsymbol{y}, \boldsymbol{m}) = 0 \end{aligned} \tag{28}$$

The objective function $f$ is the function that has to be optimised, in this case minimised. Where $\boldsymbol{y}$ is the output results from simulation and $\boldsymbol{m}$ the model parameters that have to be estimated for optimisation. These output results and model parameters have to comply with governing equations, like Darcy's law and the mass and energy balance equations, represented by the constraint equation of $\boldsymbol{g}(\boldsymbol{y}, \boldsymbol{m}) = 0$.

Automatic calibration software often uses the sum of the regularised least squares to calculate the minimum of the objective function (Gonzalez-Gutierrez et al., 2020) written as:

$$f(\boldsymbol{y}, \boldsymbol{m}) = \frac{1}{2}\left|\boldsymbol{Z}_1(\boldsymbol{g}(\boldsymbol{m}) - \boldsymbol{d}_{\text{obs}})\right|^2 + \frac{\beta}{2}\left|\boldsymbol{Z}_2(\boldsymbol{m} - \boldsymbol{m}_{\text{prior}})\right|^2 \tag{29}$$

Here the $\boldsymbol{Z}$ matrices define the weight factors, taking the importance of the model parameters into account and our beliefs of the prior estimate. Furthermore, $\boldsymbol{m}_{\text{prior}}$ are the prior estimates of the model parameters. The regularisation factor $\beta$ represents how much confidence there is in the prior estimated model parameters. The forward model $\boldsymbol{g}(\boldsymbol{m})$ and the observed data $\boldsymbol{d}_{\text{obs}}$ can be defined as the observation residual $\boldsymbol{r(m)}$:

$$\boldsymbol{r}(\boldsymbol{m}) = \boldsymbol{g}(\boldsymbol{m}) - \mathbf{d}_{obs} \tag{30}$$

The adjoint method, used in iWaiwera, uses Lagrange multipliers for the differentiation of the objective function (Gonzalez-Gutierrez et al., 2020). Implementing the Lagrangian $\mathcal{L}$, while taking the constraint function $\boldsymbol{g}$ into account, results in an unconstrained objective function:

$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{m}, \boldsymbol{\lambda}) = \boldsymbol{f}\,(\boldsymbol{y}, \boldsymbol{m}) + \boldsymbol{\lambda}^T \boldsymbol{g}\,(\boldsymbol{y}, \boldsymbol{m}) \tag{31}$$

Here $\boldsymbol{\lambda}$ is introduced as a vector of adjoint variables. In order to solve the optimisation problem, or get to a minimum for $\boldsymbol{f}$, one has to set all partial derivatives of the Lagrangian to 0 (Nocedal & Wright, 2006). Resulting in the following set of equations representing the Jacobians of the Lagrangian:

$$\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}} = \boldsymbol{g}\,(\boldsymbol{y}, \boldsymbol{m}) = 0\\
&\frac{\partial \mathcal{L}}{\partial \boldsymbol{y}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}} + \left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}\right)^T \boldsymbol{\lambda} = 0\\
&\frac{\partial \mathcal{L}}{\partial \boldsymbol{m}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{m}} + \left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}}\right)^T \boldsymbol{\lambda} = 0
\end{aligned} \tag{32}$$

The first equation is similar to the governing equation in (28) satisfying Darcy's law and the mass and energy balance equations for the forward problem. The second equation is solved to compute the adjoint variables, which can be seen as the forward problem in the linearised form. This reduces the computational cost a lot compared to solving the full nonlinear forward problem solved in, for example, iTOUGH2. In order to fully minimise the objective function $\boldsymbol{f}$ the partial derivative of the Lagrangian with respect to the model parameters has to be equal to 0, giving the last equation in equation set (32).

iWaiwera can use either the Gauss-Newton (GN) or Levenberg-Marquandt (LM) algorithm to solve the optimisation problem. This requires the Hessian matrix, which is computed by a Gauss-Newton approximation. This approximation is chosen because it only has first-order Jacobian terms, which is usually easier to compute than the full Hessian matrix. The GN Hessian matrix approximation $H_{GN}$ for the objective function (28) is denoted by this equation:

$$\begin{aligned}
H_{GN} &= \left(\frac{\partial[\boldsymbol{Z}_1(\boldsymbol{Q}(\boldsymbol{y}) - \boldsymbol{d}_{\text{obs}})]}{\partial \boldsymbol{m}}\right)^T \left(\frac{\partial[\boldsymbol{Z}_1(\boldsymbol{Q}(\boldsymbol{y}) - \boldsymbol{d}_{\text{obs}})]}{\partial \boldsymbol{m}}\right)\\
&= \left(\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{m}}\right)^T \boldsymbol{Q}^T \boldsymbol{Z}_1^T \boldsymbol{Z}_1 \boldsymbol{Q} \left(\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{m}}\right)
\end{aligned} \tag{33}$$

The GN Hessian matrix can be written in this form because $\boldsymbol{Q}$ is taken as an interpolation operator in iWaiwera and, therefore not dependent on $\boldsymbol{y}$ or $\boldsymbol{m}$. In addition, the observed data is not dependent on $\boldsymbol{m}$ either.

The sensitivity matrix $\boldsymbol{S}$ can be calculated using the partial derivatives of the governing equation with respect to $\boldsymbol{y}$ and $\boldsymbol{m}$ shown in the equation below.

$$J = \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{m}} = \left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}\right)^{-1} \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}} \tag{34}$$

Solving the nonlinear forward problem is done with the model Jacobian, or partial derivative $\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}$, and is easily accessible in Waiwera. The second partial derivative $\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{m}}$ can be solved analytically, as described in the work of Bjarkason et al. (2019).

Two additional partial derivatives are required to solve the Lagrangian for inverse modelling. These are the partial derivatives of the objective function with respect to the simulation outcome and model parameters, shown in equation (35) and (36), respectively.

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}} = \boldsymbol{Q}^T(\boldsymbol{y})\boldsymbol{Z}_1^T\boldsymbol{Z}_1(\boldsymbol{Q}(\boldsymbol{y}) - \boldsymbol{d}_{\text{obs}}) \tag{35}$$

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{m}} = \beta \boldsymbol{Z}_2^T\boldsymbol{Z}_2(\boldsymbol{m} - \boldsymbol{m}_{\text{prior}}) \tag{36}$$

It has already been shown that Waiwera can save a considerable amount of computation time for solving the nonlinear forward problem. Recently, Gonzalez-Gutierrez et al. (2020) also showed that using the adjoint approach with iWaiwera for automatic calibration is very promising. Two existing geothermal systems, Kerinci in Indonesia and Wairakei in New Zealand have been tested with iWaiwera for inverse modelling of the natural state. Although the results were promising, still some improvements is required, like fixing run failures and incorporating uncertainty quantification (Gonzalez-Gutierrez et al., 2020).

# D Visualisation Software

## D.1 Leapfrog

Leapfrog, a 3D geological software package, is used to automatically map the geological structure of the reservoir onto the new grid. Leapfrog assigns the rock type with the highest presence in a block to that specific block in the grid. However, this automatic assigning of rock types is not flawless, therefore, the numerical model has to be checked and adjusted manually. Figure 9 shows a combination of the numerical grid and the new 3D geological model. Fault lines can be observed, and the different colours of the blocks in the grid correspond to a different rock type. (ARANZ Geo Limited, 2015)

## D.2 TIM

The graphical tool Tim Isn't Mulgraph (TIM) is used to visualise models and results. TIM is based on Mulgraph and also makes use of the PyTOUGH library. Compared to Mulgraph, TIM has improved usability and is more powerful. The GUI, TIM, allows the user to access models easily and modify them. Furthermore, within TIM it is possible to run AUTOUGH2 simulations. (Yeh et al., 2013)

# Declaration concerning the TU/e Code of Scientific Conduct for the Master’s thesis

I have read the TU/e Code of Scientific Conduct[i].

I hereby declare that my Master’s thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Conduct

Date

20-02-2021……………………………………………………..…………..

Name

Ken Dekkers……………………………………………………..…………..

ID-number

0863554……………………………………………………..…………..

Signature

……………………………………………………..…………..

*Submit the signed declaration to the student administration of your department.*

---

[i] See: https://www.tue.nl/en/our-university/about-the-university/organization/integrity/scientific-integrity/
The Netherlands Code of Conduct for Scientific Integrity, endorsed by 6 umbrella organizations, including the VSNU, can be found here also. More information about scientific integrity is published on the websites of TU/e and VSNU