Greenfield resource assessment: maximising early stage data to constrain uncertainty

Tia Jones[1], Claudia Saunders[1], Ken Dekkers[1], Michael Gravatt[1], Ruanui Nicholson[1], Oliver Maclaren[1], Theo Renaud[1], Michael O'Sullivan[1] and John O'Sullivan[1]

[1]Department of Engineering Science and Biomedical Engineering, The University of Auckland, New Zealand

michael.gravatt@auckland.ac.nz

Keywords: *Resource assessment, reservoir simulation, uncertainty quantification, approximate Bayesian computation.*

ABSTRACT

Geothermal fields are a valuable resource in the energy sector. However, there is significant cost and risk involved in determining the viability of a field for production. We have, therefore, improved a method to assess the potential of a geothermal resource using numerical modelling and uncertainty quantification.

We have further conducted a study on the impact of different data types for resource assessment of a geothermal field. The research focuses on understanding the significance of each data type and its contribution to the accuracy of the resource estimations. A systematic methodology was developed to analyse and compare the effectiveness of various data types in assessing key parameters, including reservoir temperature, surface features, mass outflows, and clay cap formations. This was conducted by filtering models on the above parameters individually and in succession. These filtered models were then used in production simulation to gauge a representation of the predicted geothermal resource output.

1. INTRODUCTION

Modes of sustainable energy such as geothermal are becoming increasingly important as population growth and climate change threaten the longevity of traditional non-renewable energy sources such as oil and gas. In the current era, with an increased awareness of kaitiakitanga, alternatives to fossil fuels are being increasingly implemented. Unlike solar or wind power, geothermal is a consistent energy source provided the reservoir is maintained correctly.

Assessing the potential of a geothermal resource in the early stages of exploration and development is essential for obtaining the investment needed to continue. Existing methods, such as stored heat or power density (Grant, 2015; Zarrouk and Simiyu, 2013) make significant and often incorrect assumptions about the reservoir. Both methods usually assume the reservoir will be similar to nearby developed reservoirs or will only use partial data sets. For example, the area of the reservoir will be determined by the extent of an MT survey or the DC resistivity boundary. However, the complete 3D structure of the alteration will be ignored.

On the other hand, reservoir modelling for developed fields is an invaluable tool for assessing the sustainability of the field provided it has been adequately calibrated to available natural state and transient production data (O'Sullivan et al., 2001). This is done by changing permeability and porosity in the rock types in the model and varying the quantity and energy in hot fluid entering the bottom of the model (upflows). In the early stages of exploration or development, these data sets often are sparse or not available. This means that there are often multiple combinations of parameters that could equally explain the early data.

Recently, Dekkers et al., 2022a, proposed a new resource estimation technique where thousands of possible representations of the geothermal reservoir are created by randomly varying parameters of a reservoir model, which is tightly coupled to the best current conceptual understanding of the field (O'Sullivan, et. al, 2023). The thousands of models are filtered based on which models match the available data. The filtering process is done using Approximate Bayesian Computation (ABC). In this paper, ABC is applied to a geothermal context. However, it is applicable for a wide range of problems (Garthwaite et al., 2005; Kaipio and Somersalo, 2006; Beamont, 2019). The selected models are then algorithmically run through a production scenario where the maximum steam is extracted from each model while applying appropriate reinjection strategies. This provides a distribution of steam flow rates across the forecast time period, and in turn, power output from the field can be calculated. Dekkers et al., 2022b expanded on this concept, applying it to a representative African rift geothermal system and de Beer et al., 2023 and Power et al., 2023 both improved the initial prior distributions used and the way that various data types were incorporated.

This paper focuses on two important aspects for getting robust resource estimates. The first is the inclusion of a broad range of exploration data types. We assess the value of different data types by looking at how the uncertainty of power estimates changes when data is considered. We also assess resampling. This generates more representative models by sampling from a distribution generated from the filtered models.

2. METHODOLOGY

In this section we introduce the model and simulator used in this research, uncertainty quantification, the types of data used.

2.1 Synthetic model and Waiwera

The methodology in this paper is applied to a synthetic geothermal field representative of a volcanic New Zealand geothermal system (Renaud et al., 2021). As shown in Figure 1, this model has seven lithologies, three fault structures and an inferred alteration model. This digital conceptual model is transferred to a reservoir model grid, which has 11,764 blocks. To allow for heterogeneity in the permeability distribution, we consider each combination of factors separately in terms of rock properties such as permeability. For example, where a fault passes through a lithology, a rock type is assigned to allow a difference in flow properties to the surrounding rock. For this model, it results in 104 rock-types, which have permeability in three directions and porosity, and the resource evaluation technique can vary.

The reservoir simulator used in this research was Waiwera (Croucher et al., 2020). This is a parallelised open-source geothermal reservoir simulator developed at the University of

Proceedings 45th New Zealand Geothermal Workshop
15-17 November, 2023
Auckland, New Zealand
ISSN 2703-4275

Auckland in conjunction with GNS. It was used as the performance makes it possible to run thousands of geothermal reservoir models in a timely manner (O'Sullivan et al., 2021). Each sample model took approximately 20 seconds to converge to a steady natural state solution using 40 cores.

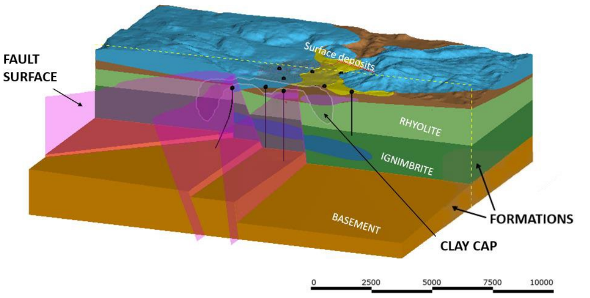

Figure 1: Synthetic geothermal field which has 7 lithologies, 3 faults and an inferred alteration. This translates to a reservoir model that has 11,764 blocks (Renaud et. al, 2021)

Once these samples have been created, they undergo a filtering process, whereby the models are filtered based on existing real-world data in the geothermal field. These filters help to narrow the sample set to a group that more accurately represents the specific field being researched, and thus provide higher accuracy of model predictions.

2.2 Uncertainty Quantification

Geothermal reservoir modelling comes with a level of uncertainty for predictions due to noisy data and model limitations. To correctly portray the accuracy of these models, uncertainty quantification should be performed. A popular method of uncertainty quantification for complex parameter spaces such as geothermal fields is Approximate Bayesian Computation (ABC). In this framework, unknown parameters are initially modelled as random variables, thus forming a relationship expressed by:

$$d = g(m) + \epsilon$$

Where d represents the field data, m the parameters, g the forward model and ϵ reflects additional sources of noise and variance (Cui et al., 2011). Once new data becomes available these distributions will be updated using Bayes formula:

$$\pi(m|d) \propto \pi(d|m)\pi(m)$$

where $\pi(m|d)$ is the 'likelihood', $\pi(m)$ the prior distribution, and $\pi(d|m)$ the posterior distribution (Nicholson, et al., 2020). Given the assumption that both the prior and noise distributions are Gaussian, the posterior can be given as

$$\pi(m|d) \propto \exp\left(-\frac{1}{2}\Big((g(m)-d)^T C_d^{-1}(g(m)-d) + \left(m - m_{prior}\right)^T C_m^{-1}\left(m - m_{prior}\right)\Big)\right)$$

Where C_d and C_m are the covariance matrices of the data (noise) and prior, respectively. As $g(m)$ is non-linear this is not in fact a simple Gaussian distribution.

For prior samples a covariance matrix was formed using the correlations between rock-types described by de Beer 2023 enforcing sensible geological constraints. In resampling a covariance matrix, first a matrix is formed with rows for each sample and columns for each parameter:

$$\begin{bmatrix} x_{1,1} & \cdots & x_{1,J} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,J} \end{bmatrix}$$

Where N is the number of samples and J the number of parameters. A column mean is found for each parameter: $\bar{x}[j] = \frac{1}{N}\sum_{i=1}^{N} x_{i,j}$. Subtracting this mean from each column gives the centred matrix:

$$X_c = \begin{bmatrix} x_1[1]-\bar{x}[1] & \cdots & x_1[j]-\bar{x}[j] & \cdots & x_1[J]-\bar{x}[J] \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_i[1]-\bar{x}[1] & \cdots & x_i[j]-\bar{x}[j] & \cdots & x_i[J]-\bar{x}[J] \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_N[1]-\bar{x}[1] & \cdots & x_N[j]-\bar{x}[j] & \cdots & x_N[J]-\bar{x}[J] \end{bmatrix}$$

The covariance matrix can be computed as $K_{zz} = \frac{1}{N-1}X_c^T X_c$.

2.3 Quantifying Data Types

The resource assessment model as much data is available, this paper specifically focuses on these four data types to filter models:

1. **Downhole Temperatures:** if the field contains wells, temperatures can be measured at specific depths down the wells and used to filter the models.
2. **Clay Cap:** The clay cap is an area of clay which seals over the geothermal field. Filtering can be done by specifying higher temperatures under the clay cap where geothermal activity is likely (Dekkers et al., 2022a).
3. **Surface Features:** This involves filtering on the temperature of certain blocks in the model where there are notable surface features such as geysers or bubbling mud. These are indicators that geothermal activity is happening below the surface.
4. **Surface Outflows:** Similarly, to surface features, the mass flow rate at certain blocks can be used to filter models.

Filtering on these data types can be toggled on and off to produce different groups of filtered models. Data types can be looked at in isolation or in combination with other data types through successive filtering.

3. RESULTS

In this section first look the impact of filtering on different data sets on production results and the how the combination of different data sets affects predictions of power output. Secondly we used resampling to obtain models that are a better representation of the data that the original prior models.

3.1 Filtering on different datasets

Successive filtering was a method used to quantify the impact of filtering on different data types. This was done by successively adding data types in order.

As more filters were added, it was ensured that they were equally weighted to each other. At each filter scenario (e.g., surface features, surface features, surface outflows, etc.) samples were run through a production model, which estimates the power output of the synthetic model when filtering on each data type. In all filtering scenarios, the best 10% of models were taken. An example of this is shown in

Figure 2 where the 10% best models that match the data for Well-3 are shown in red. The power output data was used to quantify the effect of the different filter types by assessing changes in power output as filters were added.

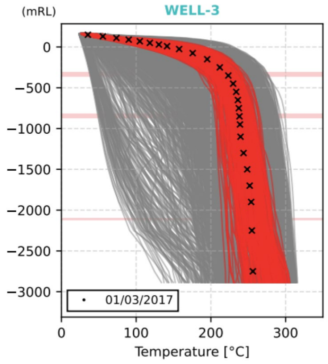

Figure 2: Example of filtering on downhole temperature. Prior distribution shown in grey, data shown by crosses and filtered samples shown in red.

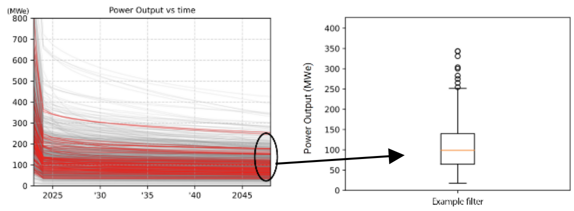

Figure 3: Example of power production forecast in time (left), distribution of power output at last time step quantified as box and whisker plot (right). Red lines are filtered models similar to Figure 2.

The estimated production output is created by simulating 25 years of future production using realistic extraction strategies. The model uses an iterative algorithm to select production and reinjection targets for the filtered models (Dekkers et al., 2022a). Since the main objective of the resource assessment method is to determine the potential power output of a field it was decided that the effects of filtering are assessed based final field power output. Since we have a distribution of power output this is displayed as a box and whisker plot, the derivation of which is shown in Figure 3.

By considering each data type individually we show the effect on predicted power output in Figure 4. For this model we see that if models are selected to match the temperature of a cold conductive well, the uncertainty (or standard deviation) is reduced by the most and the mean power prediction is the lowest (54 MWe). If models are selected based on match to a hot (convective) well measured, then the mean power predicted is significantly higher (95 MWe) however the standard deviation of predictions remains large. This difference in temperature distributions for filter on a hot well verse a cold well can be seen in Figure 5 and Figure 6 respectively. Favouring models that have surface expressions either through filtering on surface temperature or flow at the surface result in prediction of large systems. This requires more investigation as coarse reservoir models may over predict heat flow when a temperature comparison is done. For example, requiring a model that has 60°C in a surface block which is 100m x 100m would have higher heat flow than an isolated hot pool with that temperature and hence may be over representing how hot the system needs to be to match that temperature.

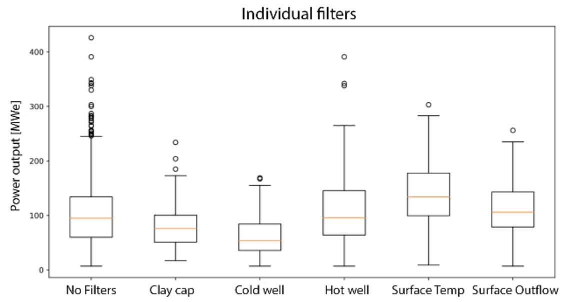

Figure 4: Predicted power production by filtering on individual data, 10% best models accepted. Hot well in this case refers to Well-3 with filtered models shown in Figure 5. Cold well refers to filtering on Well-4 shown in Figure 6.

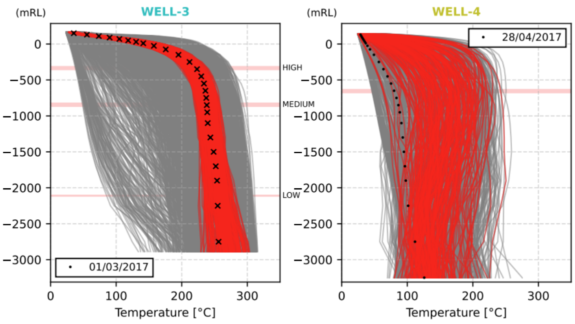

Figure 5: Effect of filtering on Well-3. Crosses show data used for filtering while dots show "unseen data" to the filtering process.

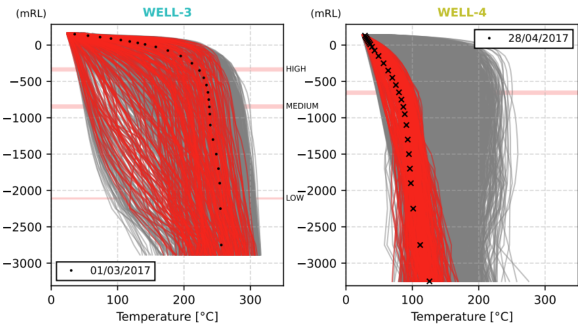

Figure 6: Effect of filtering on Well-4. Crosses show data used for filtering while dots show "unseen data" to the filtering process.

Next successive filtering was used to see to evaluate the effectiveness of filtering on multiple data types on the uncertainty prediction. Particularly we found the order of filtering was important. In all cases we are keeping 10% of prior models in the end. For successive filtering, we keep 50% of models based on the first filter and then reduce to 10% of models based on the second filter. Figure 7 shows both filtering on Well-3 first then Well-4 and vice versa. We see that by filtering on a cold well second, there is significant reduction in uncertainty by applying the second filter. While filtering on the hot well second, slightly increases the uncertainty. Firstly this shows that order of filtering will affect the results, but more importantly it shows that more models are required to sufficiently match each data type to the level of accuracy that the data is trusted. By taking the best 50% of models of cold well the middle box and whisker plot in (B) in Figure 7 has significantly higher uncertainty than taking the 10% of best models as shown for the cold well in Figure 4. However generating enough models to accurately capture all data available is difficult. In the next section we discuss resampling which increases the number of samples that fit our data by drawing from a distribution created by samples that match the data.

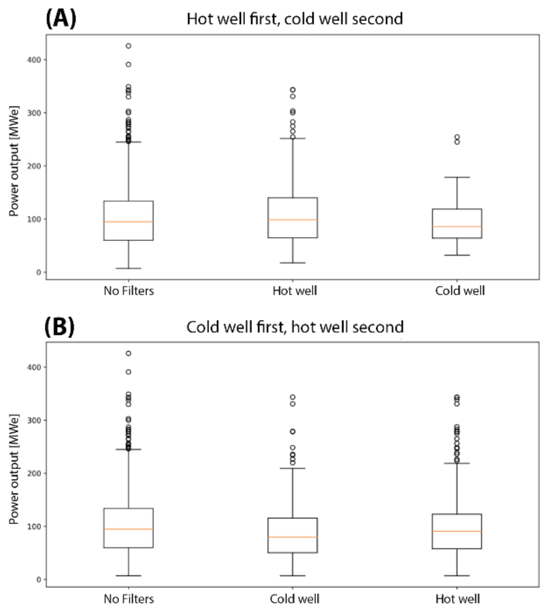

Figure 8: Successive filtering, (A) shows filtering the first 50% of samples on data from a hot well, then reducing to 10% of samples from data from a cold well. (B) shows cold well filtered first then hot well.

3.2 Resampling

One way to improve the filtering process was by resampling from a new distribution based on the filtered samples. This was achieved by creating a new covariance matrix from the 10% filtered set of sample parameters using numpy.cov. The original samples had a log gaussian distribution, so to utilise the multivariate normal function in the numpy package the log of the parameter values was found, and the covariance matrix was created from these logged parameters. Note also that for this particular function, the input matrix must have parameters as rows and samples as columns, which is the transpose of the covariance matrix described in Section 2.2. Hence, this matrix was transposed before being used in the numpy.cov function.

Random samples with a multivariate normal distribution were then extracted from this covariance matrix using numpy.random.mulitvariate_normal. 2000 samples were generated, as this is the same number of prior samples in the synthetic model. Inputs were the previously calculated covariance matrix, and the mean for each parameter.

This process generated a new set of samples with new parameter values for porosity, permeabilities and upflow. These samples with new parameters were then reprocessed to compare their distributions to that of the prior and posterior,

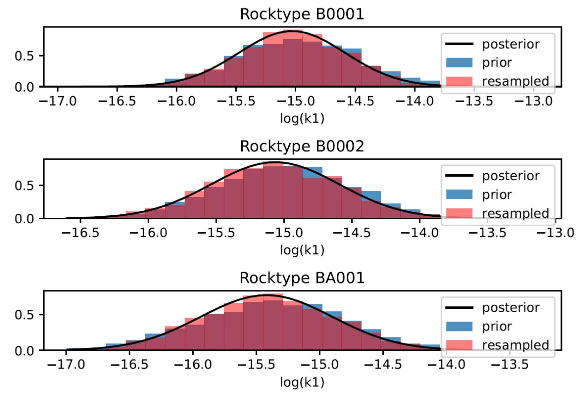

Figure 7: Sample distributions for the prior, posterior and resampled sets for K1 permeability in example rock-types.

to assess the ability of this method to create reliable samples that match available data.

Figure 8 shows the prior, posterior and resampled distributions for the K1 (x-axis permeability) of several rock types. Because the samples are log gaussian distributions, the x-axis is on the log scale. From these distributions it appears that resampling does not have a significant impact on the uncertainty of the model, as the distributions standard deviation is only slightly reduced. However, since the co-variance matrix is constructed based on filtered models, relationships between rock-types which were considered independent in the prior distributions were encoded.

Next models were run in Waiwera to get a distribution of temperatures to compare to the prior and the posterior. This was done for various filter percentages (the percent of samples we accept). Figure 9 shows the downhole temperature for well 3 for 1%, 10% and 50% filter acceptance. By calculating an Root Mean Squared (RMS) error between the filter data and each sample we can rank each sample by how well it matches the data.

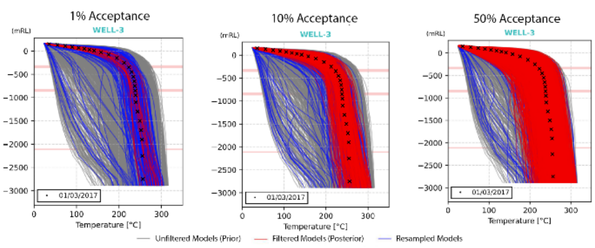

Figure 9: Unfiltered models (prior), filtered models (posterior) and resampled models based on various filter acceptance rates, 1% (left), 10% (middle) and 50% (right).

Figure 10 shows the original samples ranked by RMS error in black. The blue line represents the 10% acceptance rate used in the middle plot of Figure 9 and the red line shows the RMS error of the resampled models. We see for a 10% acceptance rate of the original prior, 89.9% of the resampled models are below the cut off error threshold. The acceptance rate is artbitrary for resampling, ideally maximisation of models that match available data is the goal. However, the optimal threshold for resampling will be field dependent and data dependent. We do show that a large number of the resampled

Proceedings 45th New Zealand Geothermal Workshop
15-17 November, 2023
Auckland, New Zealand
ISSN 2703-4275

models respect our data at least as well as the filtered models and therefore the number of models that can be used for forecasting has been more efficiently generated.

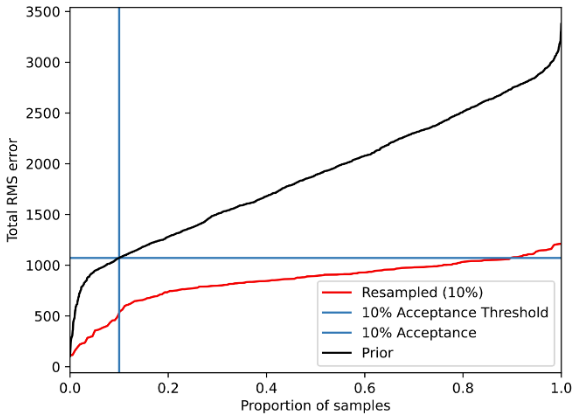

Figure 10: Root Means Squared error from the prior models (black) and the resampled models (red). The filtering threshold used for resampling is shown in blue and 89.9% of resampled models are below this threshold.

One limitation of the resampling process was the assumption of normal distributions for the parameters. This may not in fact be true when looking into the physics of geothermal fields. For example, the x and y permeabilities of rock are often more directly correlated than in the z-direction. Additionally, permeability of the rock can be affected by the presence of a fault, and a highly dense fault could block flow through it altogether. These are boundary conditions which can later be added to the resampling methods.

4. CONCLUSION

This paper demonstrates the effect of different data types such as well data and surface feature data on power predictions of a new resource assessment method from Dekkers et. al, 2020a. It also shows how resampling can be used to increase the number of samples that respect the available data. New data collected can systematically be included by further constraining the filtered models.

By more accurately including the data types available in the early stage in a robust resource assessment framework, we can better predict what geothermal fields are expected to produce. This gives more certainty to investors and decision makers to develop geothermal fields in a sustainable way and the technology shown here is already applied in a commercial setting.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. URL https://www.nesi.org.nz.

REFERENCES

Beamont, M.A.: Approximate Bayesian Computation. Annual Review of Statistics and Its Application, 6, (2019), 379-403.

Croucher, A., O'Sullivan, M.J., O'Sullivan, J.P., Yeh, A., Burnell, J., Kissling, W.: Waiwera: A parallel open-source geothermal flow simulator, Computers & Geosciences, 141, (2020), 104529, ISSN 0098-3004, https://doi.org/10.1016/j.cageo.2020.104529.

Cui, T., Fox, C., & O'Sullivan, M.J.: Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis-Hastings algorithm. Water Resources Research, 47(10), (2011), W10521.

de Beer, A., Gravatt, M.J., Renaud, T., Nicholson, R., Maclaren, O.J., Dekkers, K., O'Sullivan, J. P., Power, A., Popineau, J., & O'Sullivan, M.J.: Geologically consistent prior parameter distributions for uncertainty quantification of geothermal reservoirs. Proceedings, 48th Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California, February 11-13, (2023).

Dekkers, K., Gravatt, M., Maclaren, O.J., Nicholson, R., Nugraha, R., O'Sullivan, M., Popineau, J., Riffault, J., & O'Sullivan, J.: Resource Assessment: Estimating the Potential of a Geothermal Reservoir. Proceedings 47th Workshop on Geothermal Reservoir Engineering, Stanford, CA (2022a).

Dekkers, K., Gravatt, M., Renaud, T., de Beer, A., Power, A., Maclaren, O., Nicholson, R., Riffault, J., O'Sullivan, J. (2022b). Resource Assessment: Estimating the Potential of an African Rift Geothermal Reservoir. 9th African Rift Geothermal Conference. Djibouti, Djibouti.

Garthwaite, P.H., Kadane, J.B., & O'Hagan, A.: Statistical Methods for Eliciting Probability Distributions. Journal of the American Statistical Association, 100(470), (2005), 680-701.

Grant, M.A. (2015) Resource assessment, a review, with reference to the Australian Code, Proc., WGC 2015, Melbourne, Australia.

Kaipio, J., & Somersalo, E.: Statistical and Computational Inverse Problems. Springer Science and Business Media (2006).

Nicholson, R., Alferink, H., Patton-Simpson, E., Gravatt, M., Guzman, S., Popineau, J., O'Sullivan, J.P., O'Sullivan, M.J. and Maclaren, O. J. (2020). An Introduction to Optimal Data Collection for Geophysical Model Calibration Problems. 42nd New Zealand Geothermal Workshop. Waitangi: University of Auckland.

O'Sullivan, J., Renaud, T., Croucher, A., Yeh, A., Renaud, T., Gravatt, M.J., Riffault, J., Popineau, J., & O'Sullivan, M.J.: Comparisons of AUTOUGH2 and Waiwera on geothermal fields. Proceedings 43rd New Zealand Geothermal Workshop 23-25 November 2021 Wellington, New Zealand (2021).

O'Sullivan, J., Popineau, J., Gravatt, M., Renaud, T., Riffault, J., Croucher, A., Yeh, A., & O'Sullivan, M. : An integrated, mesh-independent geothermal modelling framework. Environmental Modelling & Software, 163, (2023) 105666.

O'Sullivan, M.J., Pruess, K., & Lippmann, M.J.: State of the art of geothermal reservoir simulation. Geothermics, 30(40, (2001), 395- 429.

Renaud, T., Popineau, J., Riffault, J., O'Sullivan, J., Gravatt, M., Yeh, A., Croucher, A., & O'Sullivan, M.: Practical Workflow for Training in Geothermal Reservoir

Modelling. Proceedings 43rd New Zealand Geothermal Workshop, Wellington, New Zealand (2021).

Power, A., Dekkers, K., Gravatt, M., Maclaren, O., Nicholson, R., O'Sullivan, J., de Beer, & A., O'Sullivan, M.: (2023) Improved Filtering for a new Resource Assessment Method. Proceedings 48th Workshop on Geothermal Reservoir Engineering, Stanford, CA.

Zarrouk, S.J. & Simiyu, F. (2013) A review of geothermal resource estimation methodology, Proc., 35th New Zealand Geothermal Workshop, Rotorua, NZ.

Proceedings 45th New Zealand Geothermal Workshop
15-17 November, 2023
Auckland, New Zealand
ISSN 2703-4275