PROCEEDINGS, 44th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 11-13, 2019
SGP-TR-214

# Reduced Models for Optimizing Well Placement and Scheduling

Rishi Adiga, John O'Sullivan and Andy Philpott

radi130@aucklanduni.ac.nz

**Keywords:** Optimization, well placement, scheduling, decision making, reduced model.

## ABSTRACT

Drilling geothermal wells has a very high capital cost. The location and operation of wells affects their production, so it is important to maximize value from wells by optimizing these decisions. The economic outcomes from particular well placement and operating policies can be estimated using reservoir simulations. A method has been developed to efficiently predict production outcomes from different combinations of possible wells and production starting times, using a relatively small number of reservoir simulations. This can then be used with optimization methods to select the best well locations and production starting times. A Mixed Integer Programming (MIP) model is presented to show this. Binary decision variables were used to select the combination of wells that would maximize total Net Present Value (NPV). Combined this approach provides an efficient method for finding optimal drilling plans, given a calibrated reservoir model.

## 1. INTRODUCTION

### 1.1 Motivation

The use of renewable forms of energy is growing globally. However, geothermal is lagging behind other forms, with a 2015 average growth rate of 2.4%, compared to an average across all renewable sources of 12% (Renewables 2016 Global status report 2016). One of the main reasons for this is that geothermal requires a much higher capital investment than the rest, a significant portion of which can be attributed to the cost of drilling wells. In Iceland, for example, the costs associated drilling and constructing wells comprise 34% of total capital expenditure (Gehringer and Loksha 2012). Also, Blankenship *et al.* estimated that drilling related expenses can exceed 50% of total plant costs (Blankenship, et al. 2005). Along with high upfront costs, geothermal ventures also involve high degrees of risk. Well drilling can be a hit-and-miss activity; a global study on the success of geothermal wells conducted by the International Finance Corporation (IFC) estimates a success rate of about 50% for the first well in a field (Success of Geothermal Wells: A Global Study 2013). The success rate improves as more wells are drilled in a field, but even over the first 30 wells the study's estimate for cumulative success rate is only about 70%.

Well-related costs can be a make-or-break factor in a geothermal project, and improving success rates for wells will bring large gains in reducing capital sunk into unproductive wells. The IFC report also found that while the success rates of exploration phase wells have been increasing notably over the years, those of development wells and operational wells have not. This suggests that methods for collecting information have been progressing, but methods for using that information to make decisions have not. As such, there is strong motivation to optimize the process of making well placement and scheduling decisions, with limited information available.

### 1.2 Outline

This paper presents a method, called the Production Scaling method, that has been developed to efficiently construct production predictions for any combination of wells and production starting times, from a small number of reservoir simulations. These predictions are in terms of the NPV contributions of each well, and NPV penalties representing the interference between them. This is important because it allows a whole solution space to be defined beforehand in a computationally cheap manner. Then, when running an optimization procedure, the entire solution space of different plans can easily be traversed without running a new simulation for each one. This method is shown to be accurate with respect to the reservoir model, for an example reservoir model based on the Kerinci geothermal system in Indonesia.

An accompanying (MIP) model is also presented for using these NPV predictions to optimize well placement and scheduling decisions. First, a background is given in Section 2 on how well placement and scheduling decisions are made in practice and on Integer Programming, along with a discussion of some previous work. Then, the Production Scaling method is presented in Section 3, including the optimization approach and MIP framework, and the actual calculation of the NPV and penalty predictions. Finally, the application of the method on the Kerinci system and the ensuing results are discussed in Section 4.

## 2. BACKGROUND

### 2.1 Current Practice

The use of numerical simulation as a tool for resource estimation and to inform drilling and production decisions has become increasingly common. Reservoir models are created and calibrated based on observations and field data such as topological measurements, magneto-telluric surveys and exploration well data, in a process known as natural state modeling. Forward simulations of natural state models are run until they converge a steady state representing the pre-production reservoir. A calibrated natural state model is then used as the initial state for simulation of production.

Using future scenario simulations to make decision recommendations is very time consuming and computationally intensive as it involves running many simulations for different scenarios, and recommendations are made by inspecting these simulation predictions, without rigorous optimization. This is especially true for large, high fidelity models that can take many weeks just to run a single simulation. This paper presents a framework for systematically using the future scenario simulation process to arrive at optimal well drilling and scheduling recommendations, given a calibrated natural state model and using as few future simulation runs as possible.

### 2.2 Previous Work

Previously, there have been many attempts to use mathematical techniques to inform geothermal well placement decisions. They have generally focused on using metaheuristics to find good solutions and fall broadly into two categories: gradient-based methods, and stochastic search algorithms. Stochastic here refers to the mechanism for searching the solution space. A common stochastic method used is Particle Swam Optimization (PSO). Ansari *et al.* used PSO to select locations for 4 production and 4 re-injection geothermal wells out of a set of 11 existing but abandoned wells in the US Gulf Coast (Ansari, Hughes and White 2014). Genetic Algorithms (GA) have also been widely used; Montes et al. developed and tested a GA on two example reservoirs (Montes and Bartolome 2001). Another stochastic method that has been used in this area is Simulated Annealing (SA). Beckner and Song used SA with a Travelling Salesperson formulation to optimize well placement and scheduling on an example petroleum field (Beckner and Song 1995).

Gradient based methods have also been used for the well placement problem. Sarma and Chen use an adjoint based gradient method on a continuous approximation of some example oil reservoirs (Sarma and Chen 2008). There have also been combinations of these methods; Bangerth *et al.* used a Simultaneous Perturbation Stochastic Approximation, which is a stochastic version of a steepest descent algorithm, and compared it to a Finite Difference gradient method and a SA method (Bangerth, et al. 2006). Though these approaches all have their advantages and disadvantages, none of them guarantee optimality (with respect to the numerical model). They all aim to find good solutions with as few simulation runs as possible. Helgason et al. ranked all blocks in an example reservoir by NPV to find an optimal location (Helgason, Valfells and Júlíusson 2017). This is essentially a grid search enumerating over the entire solution space and choosing the best one, but it is guaranteed to be optimal if only one well is being selected.

Research in the petroleum industry is more advanced, however, and strict form optimization methods have been used for decision making with hydrocarbon resources. Grossman and Goel use a stochastic MIP to optimize the planning of an offshore gas field project, accounting for decision-dependent uncertainty in their formulation (Grossman and Goel 2003). Rios *et al.* use a Partially Observable Markov Decision Process (POMDP) formulation to optimize well selection conditional on information that might accrue after each drill (Torrado, Rios and Tesauro 2017). In both these cases however, the optimization required running a new simulation or "function evaluation" at each node or potential plan being considered. No approach has been developed so far that guarantees optimality over a possible solution space while keeping the number of simulation runs low.

### 2.3 Overview of Integer Programming

A Mixed Integer Programming (MIP) model is used for the optimization, with binary decision variables modeling the selection of which wells to drill in each time period. Solution methodologies that deliver exact solutions for MIP models have made huge advances over the last twenty years, and now it is routine to solve such models with thousands of binary variables. The most popular solution approaches for MIP models use repeated application of well-known algorithms for Linear Programming (LP) models. In LPs, the objective function and constraints are linear, but variables can take fractional values. MIP solvers sequentially add extra constraints to LPs that preclude fractional optimal solutions, and enumerate different ways of fixing decision variables to binary values in a search procedure that eventually yields a provably optimal solution (Schrijver 1998). The state-of-the-art solver Gurobi was used for solving the MIP model (Gurobi optimizer reference manual 2017).

The coefficients of these variables in the objective function encode the information used to compare them (the NPVs and penalties). The optimization chooses "blindly", in the sense that it does not know the structural information of the reservoir, and is based solely on the effect of that structure in producing the NPVs and penalties. The simplest scenario of selecting only one well doesn't need a MIP, since the NPVs can just be compared and the highest selected, as was done by Helgason *et al*. The more complex scenario of selecting multiple wells and different times in this fashion warrants Integer Programming, but also requires information about the interactions between wells.

## 3. METHOD

### 3.1 MIP Model

This MIP model is presented mainly to introduce the framework with which the Production Scaling method was developed, and to demonstrate it's utility. Given a set of candidate feedzones that wells can be drilled to (blocks in a discretized reservoir model), and a pre-determined number of time periods and intervals, the optimal selection of wells to drill and start production at each time is desired. Consider $K$ potential candidates and $T$ time periods. A given well starting at a given time in isolation yields an NPV of $C_{ii}$, where $i$ is an index encoding both the well location and starting time period - for simplicity this will henceforth just be referred to as "well $i$". If we drill two wells $i$ and $j$ then we will receive an NPV of $C_{ii}$ plus $C_{jj}$. Well $j$ will also decrease the NPV from well $i$ by a (negative) penalty $C_{ij}$, and well $i$ will decrease that of well $j$ by a penalty $C_{ji}$. The decisions to make are which well to drill at the start of each time period. The MIP model is as follows overleaf.

Maximize:

$$\sum_{i=1}^{KT}\sum_{j=1}^{KT} C_{ij}\, x_{ij}$$

Subject to:

$$x_{ij} \geq x_{ii} + x_{jj} - 1, \qquad i,j = 1,\dots,KT; i \neq j \ (C1)$$

$$\sum_{m=0}^{T-1} x_{i+mK,i+mK} \leq 1, \qquad i = 1,\dots,K \ (C2)$$

$$\sum_{i=1}^{K} x_{i+mK,i+mK} \leq 1, \qquad m = 0,\dots,T-1 \ (C3)$$

Where $x$ and $C$ are defined as:

$$x_{ii} = \begin{cases} 1, & if\ well\ i\ is\ chosen \\ 0, & otherwise \end{cases}$$

$$x_{ij,i\neq j} = \begin{cases} 1, & if\ both\ wells\ i\ and\ j\ are\ chosen \\ 0, & otherwise \end{cases}$$

$$C_{ij} = \begin{cases} NPV\ of\ well\ i, & if\ i = j \\ NPV\ penalty\ of\ well\ j\ on\ well\ i, & otherwise \end{cases}$$

$$K = number\ of\ candidate\ feezones$$

$$T = number\ of\ time\ periods$$

Constraint $C1$ enforces the selection of the penalties of chosen wells on each other. If wells $i$ and $j$ are chosen $x_{ii}$ and $x_{jj}$ will both be 1, so this constraint becomes: $x_{ij} \geq 1 + 1 - 1$, which essentially makes $x_{ij} = 1$ as it is a binary variable. Constraint $C2$ says that a particular well can be chosen at most once, and can only start production at one particular time period. Constraint $C3$ limits the number of wells selected to one per time period. The objective coefficients can be visualized as a matrix, depicted below for two time periods. The blocks correspond to the different time periods, within which each entry corresponds to a particular well NPV or penalty. The main diagonal contains the NPVs for each well starting at each starting time ($C_{ii}$), shown in green, and the off diagonals are the penalty terms ($C_{ij,i\neq j}$), in blue. The diagonals of the off-diagonal blocks don't have any meaning and are zeroed out, in red.

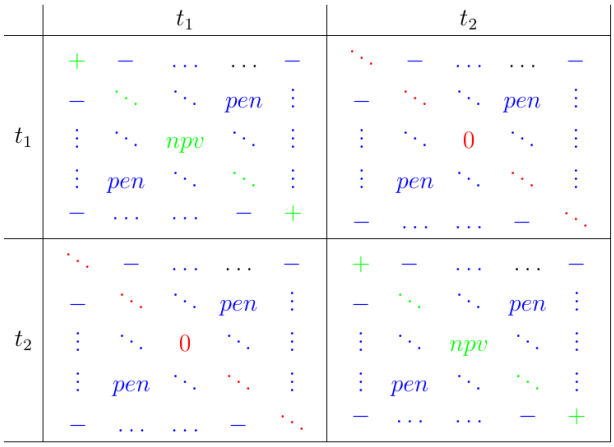


**Figure 1: Objective coefficient matrix with two time periods**

Any combination of well locations and production start times can be represented in this way, and their NPVs and penalties summed to get an estimate of the plan's monetary return. The MIP model itself is relatively simple. More sophisticated versions in a similar form have been made that include drilling costs, different kinds of constraints on well selection, and stochastic formulations considering uncertainty in reservoir calibration (Adiga, Philpott and O'Sullivan 2018). These all work with the same basic matrix structure of NPVs and penalties however, showing that very advanced and powerful analyses can be done when well production forecasts are calculated and represented in this form.

### 3.2 Production Scaling Method

Producing wells affect each other, as every well will change temperature and pressure distributions, and flow pathways in the reservoir. In general, simulating every possible combination of wells and start times is impractical, if not impossible. Therefore, a procedure was created that could make predictions for all possible combinations from a relatively small number of total simulation runs. NPV contributions from potential wells, and penalties for their effects on each other are calculated from simulation results for a subset of the possible production scenarios or plans being considered. Each simulation has wells placed at candidate locations (feedzones), and time history projections of well mass flows and enthalpies are recorded and multiplied to get heat flow predictions.

The fluid harvested from the wells should actually depend on the type of power plant installed. Dry steam plants require steam to directly turn the generator turbines, flash steam plants depressurize hot liquid to convert it to steam before driving the turbines, and binary cycle plants can use liquid at lower temperatures to heat a secondary working fluid with a lower boiling point, and use its steam to drive the turbines. There are also other considerations, such as heat loss during extraction, throttling wells to limit extraction, and possible re-injection of spent fluid back into the reservoir. These are all ignored to simplify the problem, and heat flow is used as the production quantity rather than steam flow or temperature regulated mass flow, assuming a direct conversion from heat to electrical energy via a fixed generation efficiency. A simple NPV calculation is used, multiplying the heat flows by fixed generator efficiency and electricity price to get cash flows, which are then discounted to time zero.

The candidate feedzones for placing wells in the simulation runs are selected based on simple physical cutoffs for temperature, depth and permeability. For a small enough reservoir model, all blocks can be considered, but sensible physical considerations allow many candidates to be culled, reducing the solutions space. These considerations also show that simple, programmable criteria can be used to define the set of candidates with minimal manual inspection. The wells in these simulations must use a deliverability model, with a fixed Productivity Index (PI).

The effect of extracting fluid from each feedzone on the potential resource available to every other candidate is considered individually, in a pair-wise manner. This is done by running simulations with wells producing from all the candidate feedzones, but only one with a normal PI (the *main* well) and the rest (*observer* wells) with reduced PIs, so they produce insignificant mass flows. The number of simulation runs in total is just the number of candidates considered, one with each of the candidates having the main well and the rest having the observer wells. Despite their very small mass flows, the decays in the observer wells' productions are indicative of the main well's effect on them and are scaled back up, converted to (negative) cash flows, and discounted to give the penalties. These penalties represent how much the main well's production takes away from the observer wells' potential production. Since all the observer wells have very small PIs and extract negligible amounts of resource, their effects on each other can be ignored and the main well's effects can be isolated.

Specifically, the observer wells' PIs are divided by a scale factor for the simulations. The mass flow curves for the observer wells are then multiplied by the same scale factor, and shifting by their baseline value (zeroed) to get mass flow *decay curves*. A decay curve from an example observer well is shown in Figure 2 below. The left axis shows the actual mass flow curve; its values are small and positive. The right axis shows the same curve after scaling and shifting, and its values are negative and larger in magnitude. Physically this represents the loss in potential mass flow production from the observer well due to extraction from the main well. These mass flow decay curves are then multiplied by the observer wells' enthalpies to get apparent heat flow decay curves, which are the loss of potential heat flow from observer wells due to the main well's production, and are converted to cash flows and discounted to get the penalties of the main wells on the observers.

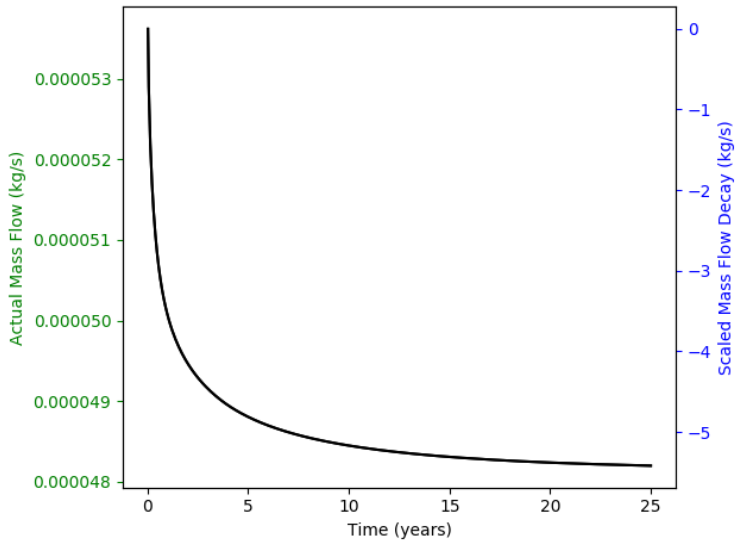


**Figure 2: Example observer well mass flow and mass decay**

The theory behind this is based on how reservoir simulations calculate production from wells in a deliverability model. Mass flow of a fluid phase $\beta$ from a grid block is defined as follows (Coats 1977):

$$q_\beta = \frac{k_{r\beta}}{\mu_\beta}\rho_\beta * PI * (P_\beta - P_{wb})$$

Here, $\mu_\beta$ and $\rho_\beta$ are the dynamic viscosity and density of the fluid respectively, $k_{r\beta}$ is the relative phase permeability of the fluid in that grid block, $P_{wb}$ is the flowing bottomhole pressure and $P_\beta$ is the fluid pressure in the grid block. To carry out an approximate analysis, the bottomhole pressure and relative phase permeability are assumed to be fairly constant parameters. Assuming compressibility effects are insignificant, this essentially says mass flow rate is proportional to the PI multiplied by the pressure difference at the grid block. The grid block pressure will also be affected by the mass extraction, however that is negligible for the observer wells because they produce so little. One way to think of this is that local pressure drawdown at a grid block due to a well producing there is the well's "effect on itself", which is accounted for by the simulation for the main well. For the observer wells however, the pressure decline at the grid block is solely due to effect of the main well, and as such, the PI can be treated as a linear scale factor between grid block pressure and mass flow production. Therefore, the observer wells' mass flows can be scaled up to give estimates of what the main well's effect on their mass productions would be, if they were also producing at the normal PI instead of the downscaled PI. Essentially, this assumes a time-constant linear relationship between an observer well's mass production and its PI, bypassing the other dynamics that would come into play if it was actually producing at the normal PI.

### 3.3 Time Shifting Production Curves

Thus far only well placement has been considered, assuming all wells are drilled and start producing at the same time. Modeling the effect of different production starting times is more complicated, but can still be done using the same set of simulation run outputs. This is accomplished by shifting the well production and decay curves in time and truncating them based on estimating the reduction in potential resource available to wells starting at different times, due to extraction from others. In general this is a very elaborate process. To start with, consider the case where there are only two wells: let $a$ and $b$ be specific wells starting at time periods $t_a$ and $t_b$ respectively, with $t_b > t_a$. Figures 3 through 6 are used to illustrate the processes described in the following explanation for two example wells, with $t_a$ = 5 years and $t_b$ = 15 years, showing production curves in green and decay curves in blue. First, consider effect of well $a$ on well $b$. The production curve from well $a$ as the main well is shifted in time to start at $t_a$ (shown in Figure 3 below), converted to a cash flow and discounted to time zero, giving the NPV of well $a$ if it starts producing at time $t_a$.

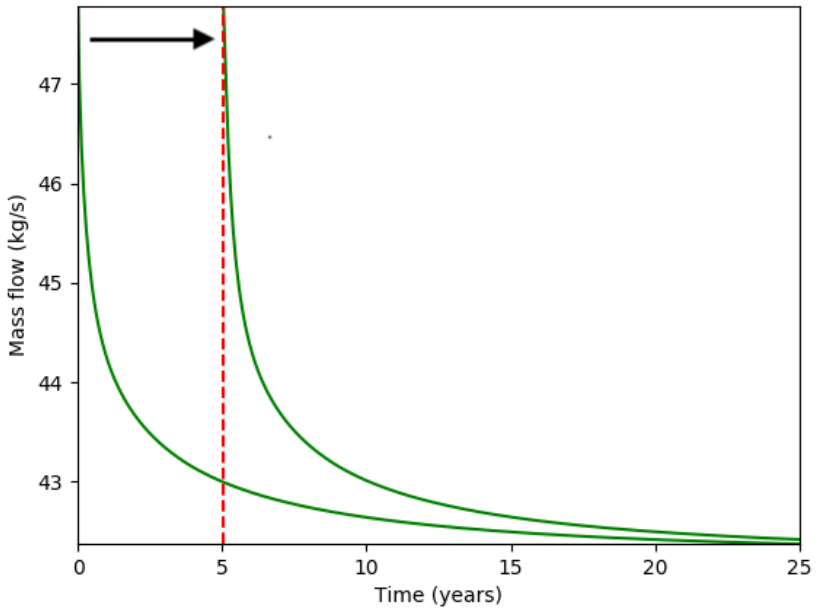


**Figure 3: Time-shifted production curve for well $a$ with $t_a = 5$ years**

Well $a$ operates from $t_a$ until $t_b$ before well $b$ starts, so the production decay curve from well $b$ as the observer well with well $a$ as the main well is shifted to start at $t_a$ and integrated from $t_a$ to $t_b$ (shown in Figure 4 overleaf). This integrated area is the estimated total loss of resource available to well $b$ due to well $a$'s operation in that time, called its *potential reduction*. Well $b$'s decay curve from $t_b$ onwards (after the integrated area) is converted to a cash flow and discounted to time zero to get the penalty of well $a$ starting at time $t_a$ on well $b$ starting at time $t_b$. The first part of the curve before $t_b$ is not included in calculating this penalty, as it is accounted for in well $b$'s potential reduction.

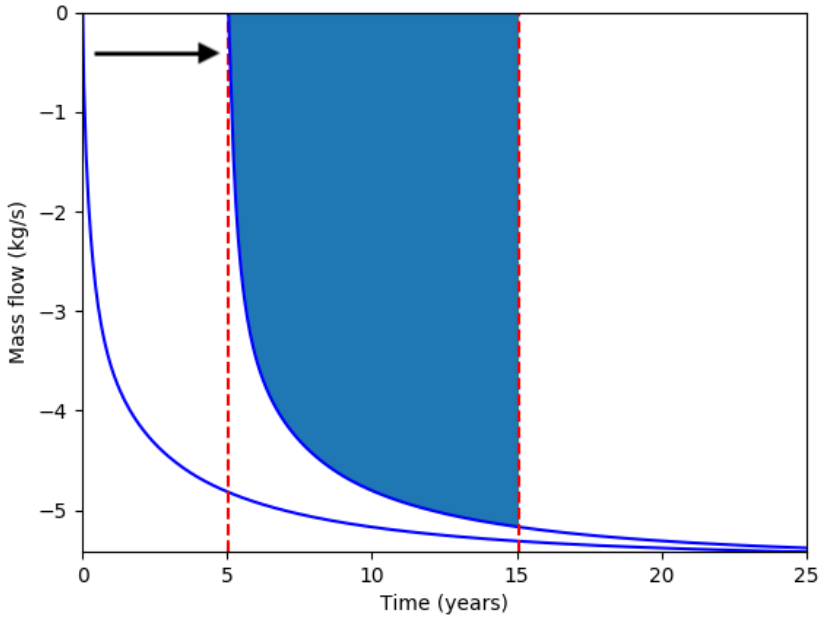


**Figure 4: Time-shifted decay curve for well $b$ with $t_a = 5$ years and $t_b = 15$ years**

Then, the production curve for well $b$ as the main well is integrated from time zero till the integral matches well $b$'s potential reduction due to well $a$'s operation (shown in Figure 5 below). The point along the production curve where this is the case is set as its new initial point. It is truncated to start from there, and now represents an estimate for what the production from well $b$ would be if it starts operating, given that well $a$ has already been operating for $t_b - t_a$ years. This truncated curve is then shifted forward in time to start at $t_b$, converted to a cash flow and discounted to time zero to get the NPV of well $b$ starting at time $t_b$.

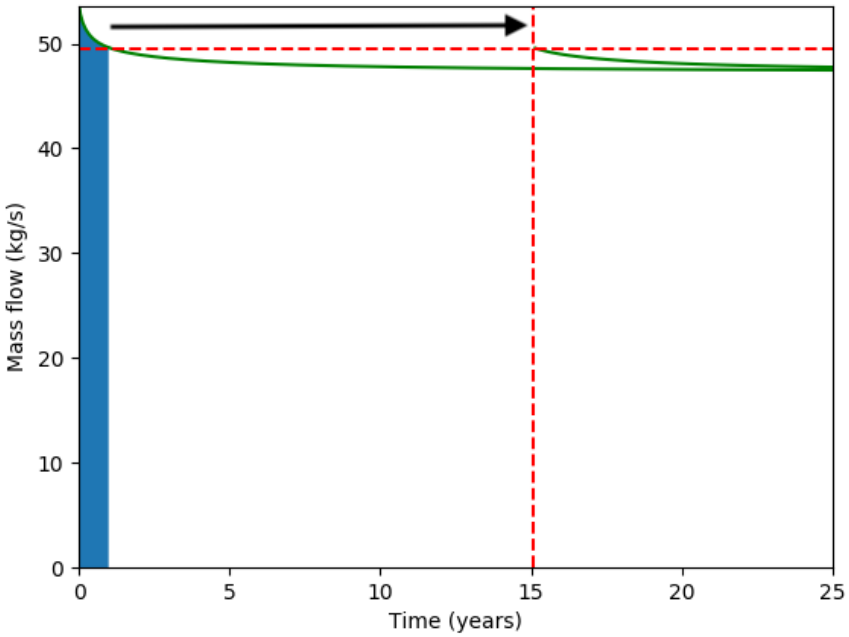


**Figure 5: Time-shifted production curve for well $b$ with $t_a = 5$ years and $t_b = 15$ years**

Finally, the decay curve for well $a$ as the observer well with well $b$ as the main well is also truncated to start from the same new initial point as the production curve for well $b$, (Figure 6 overleaf). It now represents an estimate for what the decay in well $a$'s production due to well $b$ operating would be, given that well $a$ has already been operating for $t_b - t_a$ years. This truncated curve is also shifted forward in time to start at $t_b$, converted to a cash flow and discounted to time zero to get the penalty of well $b$ starting at time $t_b$ on well $a$ starting at time $t_a$.

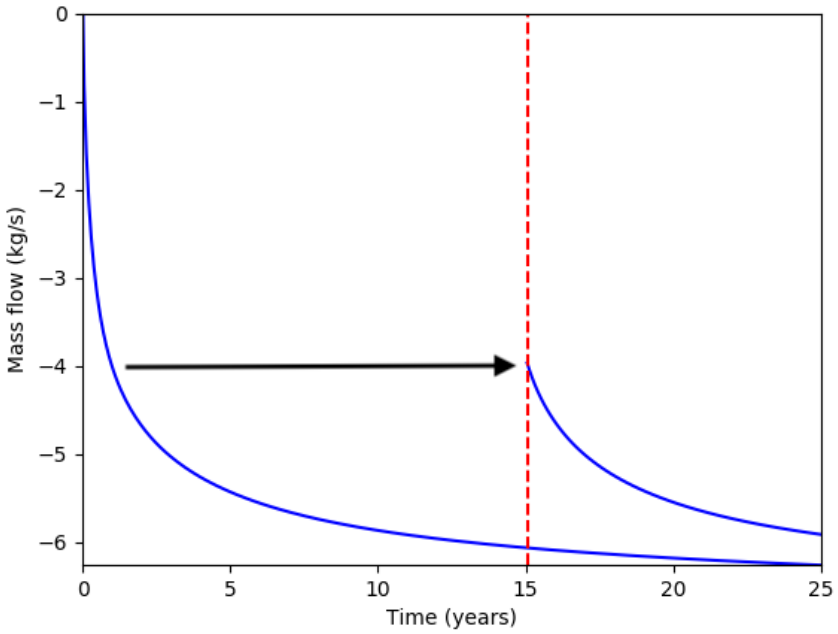


**Figure 6: Time-shifted decay curve for well $a$ with $t_a = 5$ years and $t_b = 15$ years**

However, this only works if those are the only two wells considered. This is because a well's potential reduction is dependent on the other well, but it would get included in its NPV (on the main diagonal of the objective coefficient matrix), whereas it should be included in the penalty term (in the off-diagonals). The NPVs of each well should only be a function of that well, and the penalties should be functions of pairs of wells. So instead, the NPVs are calculated for each well shifted to start at each time period individually, without any other well effects considered, as was done for just well $a$ in the above example. These are used to populate the main diagonal of the objective coefficient matrix (the $C_{ii}$s). Then, the above-described method is used in full to calculate NPVs and penalties in a two well setting, pairwise for every combination of wells and start times, $i = a + (m-1)\,K$ and $j = b + (n-1)\,K$, where $a$ and $b$ are well locations indices, and $m$ and $n$ are time period indices corresponding to the well start times $t_a$ and $t_b$. The NPV of each well $i$ and the penalty on it due to the other well $j$ are added to give its total NPVs in the two-well setting, denoted by $C'_{ij}$. The individual NPVs for both wells are then subtracted from their two-well setting NPVs to give the final penalty terms: $C_{ij} = C'_{ij} - C_{ii}$ and $C_{ji} = C'_{ji} - C_{jj}$, which populate the off-diagonals of the NPV matrix.

## 4. RESULTS

### 4.1 Reservoir Model & Parameters

The simulation model used to test the method is a relatively small one, and is based on the Kerinci geothermal system in Sumatra, Indonesia. It includes a recharge area that is 16km by 14km wide, and extends between 3.5km to 4km below the surface; the system is under the slopes of a volcano so the surface topology varies quite a bit in elevation. The reservoir is intersected by four faults that essentially bound it. Two main faults (high permeability) run in a near northeast-southwest direction, and two lesser (lower permeability) ones run northwest-southeast. The reservoir is covered by a low permeability clay cap. The numerical model was discretized into 8195 blocks and 528 nodes, in 483 columns and 19 rock layers, plus an atmospheric layer. Its natural state was calibrated with 82 defined rock types and 2 deep up-flows, to match synthetic down-hole temperature data generated for exploration wells. Future simulation runs were set up to consider a 25 year production lifespan, and took approximately one minute on to run to completion using AUTOUGH2 (Yeh, Croucher and O'Sullivan 2012) on a standard Windows desktop machine. A total of $T = 5$ time periods were considered across the project lifespan, in intervals of 5 years, for $K = 41$ blocks selected as candidate feedzones. AUTOUGH2 produces listing files to store the results of these simulations, from which production time histories were extracted and processed using PyTOUGH (Croucher 2015) modules in Python. A slice through the model showing the fault structure is given in Figure 7 overleaf.

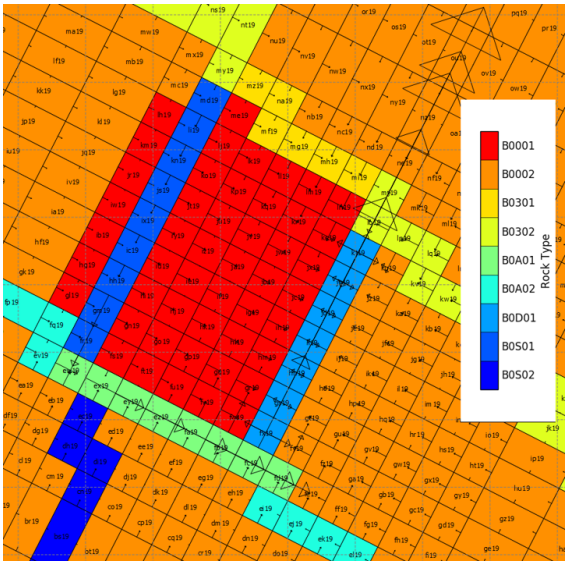


**Figure 7: Horizontal slice through the numerical model, showing the fault structure**

The PI was set at $5x10^{-13}$ for main wells, and $5x10^{-19}$ for the observer wells in each simulation. The scale factor used in the NPV penalty calculations was defined as the quotient of the two PIs, $1x10^{6}$. With these settings, the 41 simulations were run and the NPV matrix was constructed from their outputs as described previously. The Productivity Indices are actually parameters that should be calibrated to available data or defined based on some physical considerations, so they cannot be arbitrarily set in a real predictive reservoir model. However, since this reservoir model was calibrated to synthetic data and the aim is to demonstrate the application of the Production Scaling method, they were fixed. The PIs were experimented with before settling at these values. The PIs of the observer well have to be sufficiently low so they don't significatly affect each other, otherwise the decay curves obtained won't be representative of the main well's effect on them.

The other consideration is that the PIs of the main wells might not be the limiting factor influencing production, for example if the reservoir is very tight (low permeability), the local pressure drop caused by production is greater because fluid can't easily flow in from other parts of the reservoir. In this case, increasing PI won't increase mass production anymore after a certain point, as it is limited by the pressure decline. Then, scaling up observer wells' production curves can grossly over-predict the magnitude of the decay curves, as the linear scaling assumption breaks down. In general, mass production is not linearly related to PI, but this linear assumption works sufficiently well with these parameters for this model.

### 4.2 Single Start Time Accuracy

Initially, the accuracy of the method was checked for the case of multiple wells all starting production at time zero. That is, the sum of the predicted NPVs and penalties for each plan was compared to the NPV calculated from the total well output of directly simulating that combination of wells together in AUTOUGH2. If the method isn't accurate over the whole solution space (all combinations), then it is possible that the true optimal selection for the simulation model can get overlooked. Checking this required simulating every combination of wells, calculating the resulting NPVs and comparing with the method's predictions. This was done selecting every combination of four feedzones out of a reduced set of 20 candidates, which was a subset of the original 41 candidates. This amounted to 4845 simulation runs in total. The total NPVs for all combinations were calculated both using the Production Scaling method and from directly simulating the four wells together, then ranked and compared. The top 20 combinations from direct simulation are given below in Table 1.

**Table 1: Total NPVs and rankings from prediction and direct simulation for the top 20 well combinations.**

| Direct Simulation | | Production Scaling Method | | NPV % Diff |
|---|---|---|---|---|
| **Rank** | **Value** | **Rank** | **Value** | |
| **1** | 1.34E+08 | 1 | 1.30E+08 | 3.0 |
| **2** | 1.33E+08 | 5 | 1.29E+08 | 3.4 |
| **3** | 1.33E+08 | 4 | 1.29E+08 | 3.2 |
| **4** | 1.33E+08 | 3 | 1.29E+08 | 2.9 |
| **5** | 1.33E+08 | 2 | 1.30E+08 | 2.5 |
| **6** | 1.33E+08 | 6 | 1.29E+08 | 3.0 |
| **7** | 1.30E+08 | 10 | 1.26E+08 | 3.1 |
| **8** | 1.30E+08 | 11 | 1.26E+08 | 3.1 |

| | | | | |
|---|---|---|---|---|
| **9** | 1.30E+08 | 8 | 1.26E+08 | 3.0 |
| **10** | 1.30E+08 | 9 | 1.26E+08 | 3.0 |
| **11** | 1.30E+08 | 13 | 1.26E+08 | 3.2 |
| **12** | 1.30E+08 | 17 | 1.26E+08 | 3.3 |
| **13** | 1.30E+08 | 7 | 1.26E+08 | 2.7 |
| **14** | 1.30E+08 | 18 | 1.26E+08 | 3.3 |
| **15** | 1.30E+08 | 16 | 1.26E+08 | 3.2 |
| **16** | 1.30E+08 | 14 | 1.26E+08 | 3.1 |
| **17** | 1.30E+08 | 19 | 1.26E+08 | 3.3 |
| **18** | 1.30E+08 | 20 | 1.25E+08 | 3.3 |
| **19** | 1.30E+08 | 15 | 1.26E+08 | 3.1 |
| **20** | 1.30E+08 | 12 | 1.26E+08 | 2.8 |

It can be seen from this table that the optimal well combination for the Production Scaling method is also optimal with respect to direct simulation. The NPV errors are also consistently small, with all of them being less than 4% across all 4845 combinations. Despite being very accurate, it doesn't give the exact same rankings for the solutions as the direct simulation. For example, the second best solution as per the method is actually the fifth best for direct simulation, and vice versa. Plotting the total NPVs from direct simulation against those predicted by the method for all combinations, as in Figure 8 below, shows an almost linear trend. The correlation between the method NPVs and direct simulation NPVs is above 99%.

Also, there are bunches of local scattering. This is a clustering effect, with solutions grouping together in bands that can be clearly ordered. Within these groups however, similar solutions get "swapped", in the sense that one is slightly better the direct simulation, but the Production Scaling method predicts the order the other way around. They are also shown in Table 1 as well, separated by thick red boundaries between them. The top five solutions are common to both the method and direct simulation, despite not being in quite the same order. The next eight solutions and the seven after them (in simulation rank) form two more bands respectively, with almost all the solutions in them being common to both direct simulation and the Production Scaling method.

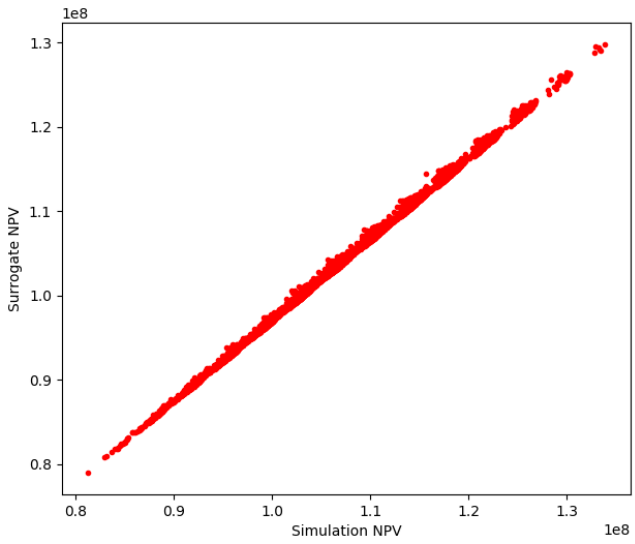


**Figure 8: Comparison of Surrogate Model to Direct Simulation**

### 4.3 Multiple Start Time Accuracy

The method's accuracy was then tested for the case of multiple wells starting production at different times along the project lifecycle, for a reduced set of drilling plans. Each simulation had four wells producing, just as before, but with only one starting at time zero and another starting at each subsequent time interval (every five years). This now required simulating every permutation of four wells, not combination as the order in which the wells start production matters, giving a much greater number of simulation runs. As such, the set of candidates was further reduced to 10, selected randomly out of the original set of 41. Now each plan was a permutation of four wells out of a set of 10, amounting to 5040 different plans. All of these were simulated directly with the four wells starting at their respective times, and the total NPV was calculated from the production results and compared to the NPVs predicted by the Production Scaling method with time shifting, as before.

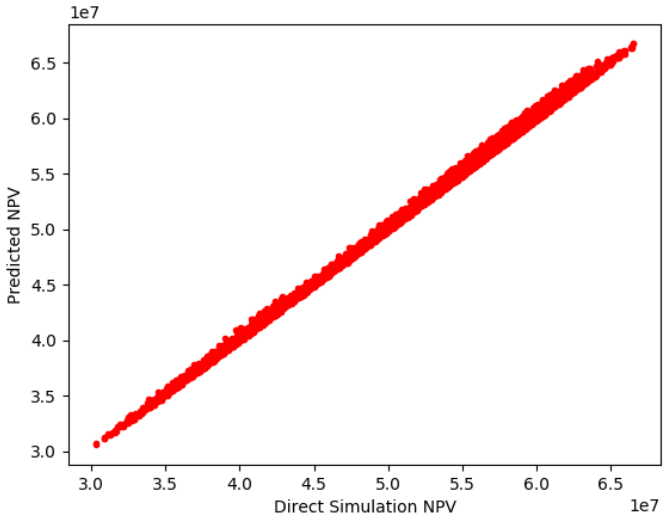


**Figure 9: Comparison of Surrogate Model to Direct Simulation**

The two sets of NPVs are shown plotted against each other in Figure 9 above. It can be seen that the predicted NPV is very close to that obtained from direct simulation. The trend is still almost linear and the correlation between the method's NPVs and the direct simulation NPVs is still over 99%. The time shifting of production curves works well and the method is very accurate when the wells start at different times. The value of such a method is that its predictions from just 41 simulation runs are comparable in accuracy to the alternative which requires 5040 simulation runs – a massive computation saving, as well as the possibility of using with an optimization framework like the MIP model presented in Section 3.1 for decision making. The optimal plan chosen by the MIP corresponds to the highest valued data point in Figure 9.

### 4.4 More Wells

While the Production Scaling method has been shown to be quite accurate for the four-well case, it is important to see how this holds as the number of wells chosen is increased. Instead of checking all possible plans for different numbers of wells, only the optimal plans were checked. A simpler version of the MIP model given earlier was used, without considering different time periods and assuming all wells start producing at time zero, shown below. Here constraint $C2$ limits the number of wells selected to $k$.

Maximize:

$$\sum_{i=1}^{K}\sum_{j=1}^{K} C_{ij}\, x_{ij}$$

Subject to:

$$x_{ij} \geq x_{ii} + x_{jj} - 1, \qquad (C1)$$

$$\sum_{i=1}^{K} x_{ii} \leq k, \qquad (C2)$$

This optimization was run multiple times using the Production Scaling method's calculated NPVs, with the limit on the maximum number of selected wells in constraint $C2$ gradually increased from four up to 15. The optimization only ever chose 11 wells at maximum, even when the well limit was 12 or more. For these cases, it didn't select as many wells as it could have, because at that point the penalties from additional wells began to outweigh their own NPV contributions, so the optimization would choose not to add them. This demonstrates that such an approach can be used to determine not only where to drill wells, but also how many to drill. It is also worth mentioning here that Gurobi took longer to solve the MIP when the well limit was increased. The optimization would run in under a second for the four-well case, but took several minutes for the 10-well case. This is because there are far more possible combinations of 10 wells than there are of four wells, so the solution space covered during the solve was much larger.

The optimal wells selected in the four well limit scenario remained in the optimal selection as the well limit was increased, with other wells being added to the selection. The Production Scaling method's predictions were checked for all the limit scenarios by comparing against direct simulation of the selected wells together. The percentage error of NPV between the method's prediction and direct simulation was plotted against the number of wells, both for the four wells that remained optimal in all scenarios, and for the total NPV over all the wells, shown overleaf in Figure 10.

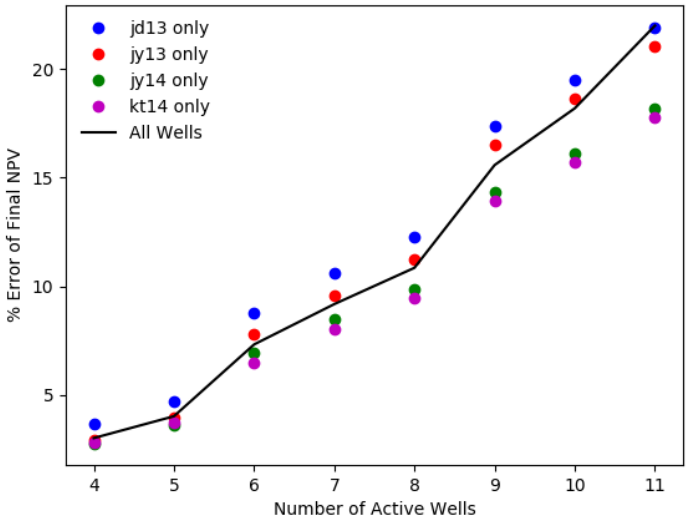


**Figure 10: NPV errors for optimal wells vs. number of optimal wells**

While the model is very accurate for small numbers of well, the discrepancy from the direct simulation values grows quite large as the number of wells increases. It reaches about 10% for eight wells, and about 20% for 10 wells. The errors also relate to well location. Wells tapping feedzones *jy14* and *kt14* have lower errors than the other two because they are deeper in the reservoir; the wells added as the well limit was increased went to shallower regions and therefore were further away and had less effect on these two wells than on the other two. The relationship between the accuracy of the predictions and the number of wells selected will depend on the reservoir. It is possible that for a larger reservoir the method will remain accurate for larger numbers of wells selected.

## 5. DISCUSSION

### 5.1 Conclusion

This paper presented an approach to facilitate the optimal selection of multiple production wells at different starting times, using as few simulation runs as possible. The Production Scaling method was developed to efficiently make production predictions for different combinations of well location and start times without simulating each one individually. This defines an entire solution space cheaply, allowing for the use of powerful optimization techniques to make well placement and scheduling decisions. An example MIP model is presented along with the method to show this, which assumes fixed time periods and optimizes where to drill wells to and in what order. The method is based on scaling simulation well production forecasts and shifting production curves in time. It simulates wells at each location in a candidate set and calculates NPVs for each well, and NPV penalties for the effect of each well on every other possible well.

It was tested on a reservoir model of the Kerinci field in Indonesia and is shown to be very accurate when compared to the direct simulation of wells together for small numbers of well, both when all wells start producing at the same time, and when they start production at different times. It was compared against direct simulation for all possible solutions selecting four wells in a reduced solution space, and its predicted solution NPVs were very strongly correlated with those from direct simulation. As such a solution found to be optimal using its predictions will at least be near optimal, with respect to this simulation model. However, as the number of wells chosen increased from four to 11, the NPV error of the model went up from less than 4% to over 20%. This is probably related to reservoir size, applied to a larger reservoir the method would likely be accurate for larger numbers of wells chosen.

### 5.2 Future Work

The application of this approach to the Kerinci reservoir model has been a proof-of-concept demonstration. Currently, it is being implemented on a large and well-developed model of the Ohaaki reservoir system in the North Island of New Zealand. There is an existing power station at Ohaaki that has been operating for decades, and as such, there is an abundance of real data and experience with this model. It is also a complex and realistic model, with multiple productive zones as well as boiling in certain areas, making it a great test case. Along with further testing the accuracy of the Production Scaling method's predictions, the relationship between the number of wells chosen and accuracy of the method will be further investigated. The issue of PI and scaling is also being worked on, to develop a more robust method to incorporate a non-linear relationship between a well's PI and the pressure decline caused by its production. This will make the method more realistic, accurate and better applicable to real reservoir systems. Once done, this will be a very powerful tool for increasing the scope of production forecasting, and to bridge the gap with optimizing decision-making.

**REFERENCES**

Adiga, R, A Philpott, and J O'Sullivan. "Stochastic MIP Models for Geothermal Well Selection and Scheduling." *ORSNZ.* Palmerston North, 2018.

Ansari, Esmail, Richard Hughes, and Christopher D White. "Well Placement Optimization for Maximum Energy Recovery from Hot Saline Aquifers." *Thirty-Ninth Workshop on Geothermal Reservoir Engineering.* Stanford, 2014.

Bangerth, W, H Klie, M F Wheeler, P L Stoffa, and M K Sen. "On optimization algorithms for the reservoir oil well placement problem." *Computational Geosciences* 10, no. 3 (2006): 303-319.

Beckner, B L, and X Song. "Field Development Planning Using Simulated Annealing - Optimal Economic Well Scheduling and Placement." *Field Development Planning Using Simulated Annealing - Optimal Economic Well.* Dallas, 1995.

Blankenship, D.A., et al. "Research Efforts to Reduce the Cost of Well Development for Geothermal Power Generation." *40th U.S. Symposium on Rock Mechanics.* Anchorage, 2005.

Coats, K., H. "Geothermal Reservoir Modeling." *52nd Annual Fall Technical Conference and Exhibition of the SPE.* Colorado: Society of Petroleum Engineers, 1977.

Croucher, Adrian. *PyTOUGH user's guide.* Auckland: University of Auckland, 2015.

Gehringer, Markus, and Victor Loksha. *Geothermal Handbook: Planning and financing power generation.* Washington, DC: Energy Sector Management Assistance Program, World Bank, 2012.

Grossman, I.,E., and V. Goel. "A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves." *Computers and Chemical Engineering*, 2003.

*Gurobi optimizer reference manual.* Gurobi Optimization, Inc., 2017.

Helgason, Dagur, Ágúst Valfells, and Egill Júlíusson. "Algorithm for Optimal Well Placement in Geothermal Systems Based on TOUGH2 models." *42nd Workshop on Geothermal Reservoir Engineering.* Stanford, 2017.

Montes, Guillermo, and B P Bartolome. "The Use of Genetic Algorithms in Well Placement Optimization." *SPE Latin American and Caribbean Petroleum Engineering Conference.* Buenos Aires, 2001.

Moon, Hyungsul, and Sadiq J Zarrouk. "Efficiency of geothermal powerplants: a worldwide review." *New Zealand Geothermal Workshop 2012.* Auckland, 2012.

*New Zealand's Energy Outlook.* Ministry of Business, Innovation & Employment, 2013.

Onwunalu, Jerome E, and Louis Durlofsky. "A New Well-Pattern-Optimization Procedure for Large-Scale Field Development." *SPE Journal* 16, no. 03 (2011): 594-607.

Ozdogan, Umut, Akshay Sahni, Burak Yeten, Baris Guyaguler, and Wen H Chen. "Efficient Assessment and Optimization of A Deepwater Asset Using Fixed Pattern Approach." *SPE Annual Technical Conference and Exhibition.* Dallas, 2005.

*Renewables 2016 Global status report.* Paris: Renewable Energy Policy Network for the 21st Century, 2016.

Sarma, P., and W. H. Chen. "Efficient Well Placement Optimization with Gradient-Based Algorithms and Adjoint Models." *SPE Intelligent Energy Conference and Exhibition.* Amsterdam: Society of Petroleum Engineers, 2008.

Schrijver, Alexander. *Theory of Linear and Integer Programming.* New York: John Wiley & Sons, Inc., 1998.

*Success of Geothermal Wells: A Global Study.* Washington, DC: International Finance Corporation, World Bank, 2013.

Torrado, R., R., J. Rios, and G. Tesauro. "Optimal Sequential Drilling for Hydrocarbon Field Development Planning." *AAAI Conference on Innovative Applications.* San Francisco, 2017.

Yeh, Angus, Adrian Croucher, and M.J O'Sullivan. "Recent developments in the AUTOUGH2 simulator." *TOUGH Symposium.* Berkeley, 2012.