

Author's Accepted Manuscript

Optimisation of decision making under uncertainty throughout field lifetime: A fractured reservoir example

Dan Arnold, Vasily Demyanov, Mike Christie, Alexander Bakay, Konstantin Gopa



PII: S0098-3004(16)30190-X
DOI: <http://dx.doi.org/10.1016/j.cageo.2016.07.011>
Reference: CAGEO3800

To appear in: *Computers and Geosciences*

Received date: 6 August 2015
Revised date: 20 July 2016
Accepted date: 21 July 2016

Cite this article as: Dan Arnold, Vasily Demyanov, Mike Christie, Alexander Bakay and Konstantin Gopa, Optimisation of decision making under uncertainty throughout field lifetime: A fractured reservoir example, *Computers and Geosciences*, <http://dx.doi.org/10.1016/j.cageo.2016.07.011>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Optimisation of decision making under uncertainty throughout field lifetime: A fractured reservoir example.

Dan Arnold, Vasily Demyanov, Mike Christie, Alexander Bakay, Konstantin Gopa

Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, UK

dan.arnold@pet.hw.ac.uk

Abstract

Assessing the change in uncertainty in reservoir production forecasts over field lifetime is rarely undertaken because of the complexity of joining together the individual workflows. This becomes particularly important in complex fields such as naturally fractured reservoirs. The impact of this problem has been identified in previous and many solutions have been proposed but never implemented on complex reservoir problems due to the computational cost of quantifying uncertainty and optimising the reservoir development, specifically knowing how many and what kind of simulations to run.

This paper demonstrates a workflow that propagates uncertainty throughout field lifetime, and into the decision making process by a combination of a metric-based approach, multi-objective optimisation and Bayesian estimation of uncertainty. The workflow propagates uncertainty estimates from appraisal into initial development optimisation, then updates uncertainty through history matching and finally propagates it into late-life optimisation. The combination of techniques applied, namely the metric approach and multi-objective optimisation, help evaluate development options under uncertainty. This was achieved with a significantly reduced number of flow simulations, such that the combined workflow is computationally feasible to run for a real-field problem.

This workflow is applied to two synthetic naturally fractured reservoir (NFR) case studies in appraisal, field development, history matching and mid-life EOR stages. The first is a simple sector model, while

the second is a more complex full field example based on a real life analogue. This study infers geological uncertainty from an ensemble of models that are based on the carbonate Brazilian outcrop which are propagated through the field lifetime, before and after the start of production, with the inclusion of production data significantly collapsing the spread of P10-P90 in reservoir forecasts. The workflow links uncertainty estimation with the appropriate optimisation at appraisal, development and reservoir management stages to maximise oil recovery under uncertainty.

Keywords: uncertainty quantification, history matching, fractured reservoirs, model classification, optimisation, field development

1 Introduction

A key challenge in reservoir simulation is to know the appropriate number of models to run in order to make a good decision given estimates of uncertainty in reservoir description. Much work has been published on topics such as history matching, uncertainty quantification, sensitivity analysis and optimisation to tackle this challenge across 4 key stages for reservoir modelling & development:

1. Reservoir appraisal: Sensitivity analysis is (commonly) used to tell us the preproduction estimate of uncertainty.
2. Initial reservoir development planning: The best development options are explored to maximise value given the uncertainty.
3. Reservoir model history matching: Production data is integrated to improve the estimates of uncertainty
4. Mid-late life Reservoir management: Additional mid/late life development decisions are optimised given the uncertainty, which has been updated from history matching.

Stages 1 and 3 are concerned with reservoir forecasting given the uncertainty in the subsurface model. Stages 2 and 4 are concerned with identifying new development opportunities to maximise reservoir value given a range of different options, well locations and other engineering trade-offs (e.g. maximising oil vs minimising water). The stages represent a set of interconnected ill-posed inverse

problems, where many possible models/solutions may exist given the data yet provide a range of forecasts.

Stochastic optimisation methods are often employed to solve the inverse problem in reservoir simulation, and Bayesian formalism can be applied to the results of history matching to estimate the uncertainty. Alternatively, Park et al (2013) for instance estimated uncertainty using kernel probability density estimation, a different formalism to the Bayesian way we employ in this paper.

Uncertainty analysis is rarely however propagated from one stage to another in a rigorous yet practical way which is the aim of the workflow described in this paper. The aim is to create an improved workflow that would fully couple stages 1 – 4 and propagate the uncertainty through each stage of forecasting and optimisation. The reservoir development will be optimised across the full range of possible reservoir descriptions, given the considered uncertainty, to provide the best development option (or set of options to choose from). This could be described as optimisation under uncertainty.

One of the most statistically robust approaches to estimating the uncertainty in an inverse problem is to apply Monte Carlo (MC) techniques to infer unknown distribution of the uncertain model parameters. Bayesian approaches to uncertainty quantification often use Markov Chain Monte Carlo (MCMC) to quantify uncertainty. In all cases (particularly for MCMC as many samples are rejected in the process of running the algorithm) Monte Carlo methods are computationally expensive with many hundreds of thousands or millions of iterations being required for convergence in high dimensional problems.

Building a complete field lifetime workflow using these MCMC techniques aiming to propagate the uncertainty through the workflow would create an unmanageable computational cost, unless surrogate models are used, however these trade-off speed of solution against additional errors in uncertainty quantification. The resulting combinatorial explosion in, for instance, finding the optimal solution for drilling a new well given a large ensemble of possible reservoir descriptions could easily lead to millions of model runs – an unfeasibly large number for practical engineering problems.

This work also aims to reduce the computational workload of estimating and propagating uncertainties effectively to improve on the industry norm. The reality for many practicing engineers, given typical constraints on time and the computer power, is something equivalent to the following steps:

- Simulations are run to assess the uncertainty at the initial stage, when the uncertainty is theoretically at its greatest. These simulations are more commonly developed as Min-Max ranges rather than statistically significant confidence intervals.
- One single “most likely”/“base case” is carried forward for development planning and optimisation at the initial stage. Sometimes the min and max cases are tested against the optimal development plan for the “base case”.
- The base case model is compared to actual production data and history matched to the production data where possible. The history matching process can often involve adding parameters such as multipliers to the simulation model in order to add flexibility to the model and make matching easier. Where the base case cannot match, another model from the appraisal stage or a newly generated model may be used to improve the history match.
- The best history matched model is graduated to the next stage and used in subsequent reservoir development decisions, history matching and optimisation phases, being the “most likely” model.

The problem with the above approach is that at each stage, the uncertainty is not estimated with any reasonable level of statistical robustness (such as you would get from MCMC) and the uncertainty estimates are not propagated into the optimisation stages (Stages 2 and 4). Therefore, the “optimum” is only the optimal for the most likely reservoir description, rather than in respect to the reservoir uncertainty.

In this paper we propose a solution to the problem of propagating estimates of uncertainty throughout field lifetime with a minimum number of simulation runs to provide accurate estimates of uncertainty.

This is attempted by developing a full field lifetime workflow that:

- propagates the uncertainty from one stage to the next such that we optimise decisions given the level of uncertainty and no information is lost between each stage;
- calculates and updates statistical estimates of uncertainty at all stages;
- is computationally efficient in minimising the number of stochastic iterations required to assess uncertainty and optimise the development.

The objective is to create a complete, field lifetime workflow for uncertainty propagation that is efficient enough to be attempted on a real field scenario. Recent papers, such as Shirangi and Durlofsky (2015), demonstrate approaches for this type of (Closed-loop) workflow where uncertainty analysis and optimisation are coupled but required large numbers of iterations in to achieve a result (220,000 flow simulations) thus reducing the simulation model requirement would be an advantage for more complex field problems. Maucec et al (2011) used a Bayesian MCMC approach on a proxy model to estimate geological uncertainties then MDS to dynamically rank the realisations, however they did not carry these uncertainties into an optimisation step thus well covered steps 1 and 3.

This paper describes a novel combination of existing techniques to facilitate an optimisation under uncertainty process (carried out both before and after history matching) to maximise the value of the field throughout its lifetime.

The first section of this paper describes the set of techniques employed in this new workflow that enables accurate estimation of the uncertainty with a minimal cost in terms of simulation time. This is a combination of three key technologies: Multi-Objective optimisation, posterior NA-Bayes inference and model clustering/classification using Multi-dimensional scaling. These techniques are used in different combinations at each stage to best capture the uncertainty while minimising the computational cost. There is a brief description of each technique used.

The second section of the paper demonstrates variations of the new workflow on 2 synthetic case study examples, (1) a simple sector models and (2) a more complex fractured field example. The second case study is a relatively complex but realistic (70740 cells) synthetic fractured field example developed using fracture analogue data from van Eijk (2014), to create a set of Discrete Fracture Network (DFN)

models, which were then upscaled to create representative simulation grids. The field is 3 phase, with a gas cap and a relatively thin oil bearing section, making optimisation of well placement a challenge given the quantified range of uncertainty and model complexity.

To be consistent throughout the paper, a predefined set of nomenclature is used to explain important terminology around modelling. These terms are described as follows:

- Scenario – a high level geological idea that is to be tested. Different geological scenarios describe differences in the geological conceptual model.
- Realisation – this is the outcome from a stochastic process (which can be either stochastic simulation of the static model or the models produced from stochastic optimisation (history matching or optimisation)). We can produce many realisations of a scenario
- Model – the specific technique used to capture nature, for instance a DFN and a geocellular model capture nature in different ways thus contribute a different model.

2 Methodology: An approach for a full-lifetime workflow

2.1 A complete field lifetime workflow for uncertainty propagation and optimisation under uncertainty

Figure 1 describes the workflow used in this paper as applied to each stage of field lifetime (Appraisal, Development, History Matching, Management). Each stage has a key outcome/decision that is the basis of the modelling effort, which is plotted against the computational effort required to solve the problem using this workflow. The total computational cost of the 4 stages in this example is under 5000 simulations for the entire workflow.

The workflow techniques are described more fully in the next section, however the workflow can be generalised as follows:

1. An initial set of fracture configurations for the field are developed from appropriate analogue data to create a set of possible geological descriptions. The analogue data is used to condition a set of different DFN models that could represent the fracture patterns/orientations in the field. Which analogue data to use is a modellers choice, but the 2 examples in this paper use outcrop analogue data, which is described more fully below. The DFN model properties are then upscaled for use in a reservoir simulation model grid. An appropriate fracture upscaling method is used to create a set of permeability fields, which represent the ensemble of possible geological scenarios of the reservoir.
2. The upscaled DFN permeability fields are used as training images for Multi-Point statistics (MPS) to create a large number of realisations of the reservoir, covering the spread of uncertainty around each geological scenario. The MPS models are conditioned to any appraisal well data.
3. Models with similar property distributions (porosity, permeability, flow responses) are grouped together using a clustering technique to reduce the dimensionality without compromising the uncertainty range. The grouping is based on a similarity/dissimilarity matrix computed for a user selected metric (e.g. production rate) of the model.
4. A set of model clusters are identified and the median value model from the centre of each cluster is found. This reduces the number of models needed to estimate the uncertainty, down to the number of clusters identified. The cluster centre models are then used to recreate the P10-P90 range from Step 1 and the uncertainty represented by the selected cluster centre models can be compared to the exhaustive set of models to make sure the P10/P50/P90 estimates are close.
5. This is the start of the initial development (Stage 2). Optimisation is carried out under uncertainty (i.e. the optimal solution is obtained over the range of possible reservoir descriptions) by optimising across all of the cluster centre models. Optimisation is carried out using a Multi-Objective optimisation algorithm to find the best set of possible

development scenarios that balance different production target criteria (e.g. maximise oil and minimise water, NPV vs NPC), given the uncertainty.

6. A development option is chosen and the uncertainty range is then reassessed given the data from the new wells. The geological descriptions are reconditioned to the new well data and the range or uncertainty is compared to the original estimates. The number of geological descriptions could be exhaustive (the complete set generated in step 1 of this workflow) or only on the n-cluster centre models from step 4 (to save compute time).
7. The next stage (Stage 3) is to update the prior estimates of uncertainty using history matching. This is a 2 step process that first (1) automatically history matches the reservoir model using a stochastic optimisation algorithm, then (2) calculates the posterior Bayesian credible intervals by post-processing the ensemble of models using a Gibbs sampler. The Gibbs sampler is implemented in the NA-Bayes algorithm. To incorporate geological reservoir uncertainty from the previous steps, the cluster centre models are included as part of the history matching.
8. NA-Bayes selects a subset of models from the history match ensemble to be used in the estimation of uncertainty. A set of these resampled models will be used to represent uncertainty post-history matching. Some or all of the resampled models can be used in subsequent optimisation under uncertainty.
9. The final phase is to assess any EOR decisions under the updated estimate of uncertainty. Multi-objective optimisation under uncertainty is used to evaluate the mean response of candidate development options for the set reservoir descriptions selected from NA-Bayes. The outcome is the forecast for the new development decision under the estimated uncertainty from NA-Bayes.

Steps 1-4 above are summarised in Figure 2 which illustrates the process of analogue data collection to clustering models based on their similarity.

The key techniques applied in this workflow are clustering with Multi-dimensional scaling, Multi-Objective optimisation and the NA-Bayes Bayesian ensemble inference algorithm. These techniques are described in more detail in the following section.

2.2 The uncertainty quantification and optimisation techniques used in this paper

Extensive work was carried out on uncertainty quantification and optimisation in recent years and many techniques have been developed for uncertainty quantification and optimisation for various stages of reservoir development. The key challenge being researched is the reduction in CPU time required to achieve the best results for both tasks. This paper employs a set of techniques that reduce the required number of simulations throughout field lifetime. These are multi-objective optimisation, multi-dimensional scaling and NA-Bayes.

- Multi-Objective optimisation improves the performance of finding “good” models (e.g. good matches, good development plans). It improves performance of the optimisation and history matching phases by increasing the diversity of possible solutions for the same computational cost of Single Objective methods. Obtaining a more diverse set of solution is important to ensure robustness of the predictions and decisions made.
- Multi-dimensional scaling is employed with cluster analysis to identify the minimum set of models required to capture uncertainty. This subset can then be used to propagate the uncertainty into subsequent uncertainty estimation and optimisation steps.
- NA-Bayes allows the unbiased estimation of uncertainty from history matching from the biased sampling responses of stochastic optimisation, which preferentially samples from good fitting models.

Each approach is discussed in turn below.

2.2.1 Multi-Objective Optimisation

Reservoir optimisation problems such as history matching and development optimisation can be done efficiently using optimisation approaches such as stochastic algorithms, filter based methods such as EnKF (Oliver et al, 2011) or traditional gradient optimisation approaches reviewed in Gomez et al

(2001). The latter is used less widely as history matching is ill-posed and gradient methods suffer from entrapment in local minima.

Optimisation is carried out using an objective function that is to be maximised or minimised, typically a single objective function that is focused on a global optimum. In history matching, where many minima exist, it is preferential to maximise the number of diverse minima that are found and hence maximise the spread of forecasts to gain reliability.

Adaptive stochastic optimisation algorithms are often favoured in history matching as their stochastic nature can balance exploration of parameter space for local minima with minimising the number of simulation runs. Examples of modern stochastic search algorithms are Particle Swarm Optimisation (PSO) (Kennedy and Eberhart, 1995, Mohamed et al, 2010), Differential Evolution (DE) (Hajizadeh et al 2010) and Bayesian Optimisation Algorithms (BOA) (Abdollahzadeh et al, 2012).

In reservoir development optimisation there may be several competing objectives that need to be considered in any engineering decisions. For instance, there is a trade-off in many wells in maximising the oil production and minimising water. Multi-Objective versions of these algorithms (e.g. Hajizadeh et al 2011, Christie et al, 2013) allow trade-offs in match quality (for HM) or engineering criteria (for Opt) to maximise the diversity of models and in some cases improve algorithm performance. While single objective algorithms work by optimising one, predefined objective to find the best solution, multi-objective algorithms work by finding solutions that balance the trade-off between a number of conflicting objectives. As one objective's score is improved there is a corresponding worsening of the competing objectives. MOO finds solutions that balance this trade-off by finding those solutions that cannot improve any objective without worsening any of the other objectives – this is known as the non-dominated set or Pareto front (Teich, 2001).

In this paper we employ a Multi-Objective Particle Swarm Optimisation (MOPSO) algorithm like that demonstrated in Christie et al (2013), which showed fast convergence in 200-300 iterations on a complex real field problem.

2.2.2 NA-Bayes Bayesian inference algorithm (NAB)

Bayesian approaches to uncertainty estimation are common in post-production fields, where the history data is used to estimate the likelihood (Christie et al, 2006; Elsheikh et al, 2014). Models are history matched to data using an optimisation method and the model probability is updated based on match quality to calculate a posterior and narrow the estimate of uncertainty. The “gold standard” approach for Bayesian inference is to apply Markov Chain Monte Carlo (MCMC), which requires a large numbers of simulation runs to achieve convergence, hence MCMC is rarely applied to real-field problems.

Adaptive stochastic optimisation algorithms are used in history matching to reduce the number of required iterations to find good matches, unfortunately the posterior cannot be estimated directly from these algorithms as they inherently bias the sampling towards good fitting areas of parameter space (and hence bias the estimate of the posterior). This bias must be removed in order to make accurate estimates of the posterior from the outcomes of stochastic optimisation.

NA-Bayes (Sambridge, 1999) is applied to the misfit ensemble, generated from stochastic optimisation, to post-process the misfit results and approximate the posterior P10-P90 effectively (Christie et al, 2006). NA-Bayes (or NAB as it is sometimes abbreviated to) uses Voronoi cells to interpolate the misfit surface away from the sample points, assigning a constant value of misfit within the cell. It then employs a Gibbs sampler to take a random walks across the Voronoi tessellated parameter space where each proposed step of the walk is accepted/rejected based on the cell likelihood, thus the higher the likelihood the greater the chance of acceptance. While the probability of accepting a low misfit model is high, the probability of a random walk landing in the smaller volume Voronoi cells is also lower. Thus only a subset of the complete ensemble from stochastic history matching will be accepted during the resampling to approximate the posterior PDF.

We have employed NAB throughout this work as a way to approximate the posterior probability distribution (PPD) while maintaining the computational performance of stochastic optimisation in providing a range of different history matches. The combination with MOO and NAB has been shown to improve the overall estimate of the posterior as more diverse local minima are included in the inference (Hajizadeh et al, 2011).

All these approaches require the model uncertainty to be described by a parameterisation, which can be difficult to do for some cases where multiple interpretations of the reservoir exist (Bentley and Smith, 2008, La Pointe & Fox, 2011, Park et al, 2013, Jung et al, 2013, Arnold et al, 2013) or where parameter priors are highly non-uniform (Rojas et al, 2012).

2.2.3 Model selection using a Multi-Dimensional Scaling distance based approach

The application of distance- metric based methods in the petroleum industry was demonstrated previously in a number of papers, namely Scheidt & Caers (2007, 2008, 2009). The authors used a similar approach for uncertainty estimation as the one described in this paper on a pre-production turbidite reservoir. Jung et al. (2013) used Modified Hausdorff distance measure between dual-porosity patterns resulting from DFN upscaling and used these measures to obtain representative training images for faster Multi-Point statistical (MPS) modelling.

This work uses the Euclidean distance measure to indicate the similarity or dissimilarity of the model where similar models are closer together (Borg & Groenen, 2005). The realizations are compared and grouped together into clusters using spectral clustering extension of k-medoids in the multi-dimensional scaling (MDS) projection and are represented as points in a metric space of models. A kernel function is applied to the MDS space to provide a better spatial allocation of models for clustering. To improve the clustering process spectral clustering (after Dhillon et. al., 2005) was used to identify a single model to represent each of the clusters.

3 Case Study examples

The general workflow given in section 3.1 is tested on two field examples of naturally fractured reservoirs. In both cases the flow is predominantly from the fracture network, storage in the matrix (a fractured carbonate analogue). The examples represent a simple sector model with 1-3 wells and a more complex full field, three phase example.

In both case studies the permeability field is controlled by the fracture distribution. Fractures distributions are captured using a discrete fracture network (DFN), which is a stochastic method of fracture modelling to allow the generation of 3D, synthetic, probabilistically simulated fractures

(Dershowitz & Einstein, 1988; Dershowitz et al, 1991). DFN models are controlled by a range of parameters including those describing their physical characteristics, such as lengths, orientations, aperture (how open the fracture is), fracture intensity, statistical characteristics. Different geological scenarios can be set up by identifying the major and minor fracture sets and the ranges of their physical properties, from which realisations can be generated stochastically.

The DFN is upscaled using the Oda Gold method to reduce the computational resources and time required to evaluate the models. The Oda Gold approach (Elfeel, 2013) is an extension of the Oda method (Oda, 1985), a fast analytical (i.e. non flow-based) upscaling method for DFNs. Oda Gold improves on the accuracy of Oda by better accounting for the fracture connectivity in the upscaled model (to account for disconnected discrete fractures). Oda gold provides the full permeability tensor for use in reservoir simulation. In both case studies the flow is associated with the fractures only, therefore allowing a single porosity model to be used (the porosity homogeneous and kept constant between scenarios to allow for an easier comparison of flow response). In practice however the upscaled data may be used in more complex dual-porosity or dual-permeability simulations, depending on the problem.

The benefit of upscaling is to allow the consideration of more scenarios to better cover the range of reservoir uncertainties, however the corresponding flow response is subject to upscaling errors introduced by the choice of the upscaling approach (Elfeel et al, 2012, 2013).

The management of simulation time vs model validity is handled differently in the 2 examples, with the sector model using a production metric for clustering, where the heavy simulation work is carried out on coarse gridded models, given an acceptably low level of error from the coarse-fine upscaling process. The second example measures model similarity based on the permeability field to reduce the number of required flow simulations and to better account for spatial variability.

Case study 1: Sector model

The model represents 1000x1000x100 meter oil reservoir with a heterogeneous upscaled DFN fracture permeability and porosity taken as (homogeneous) matrix porosity to maintain OIIP between

realisations. In this simple case study 2 fracture orientations were identified, a parallel set (0°) and oblique (45°) which had uncertainties in the fracture length, fracture density and the relative proportion of each fracture set. Fracture density decreased linearly from the producer to injector side of the field described by an unknown linear function which was parameterised. In all 384 DFN scenarios were generated the values given in Table 1 with 4 realisations of the DFN being generated for each scenario.

A total of 1536 DFN realisations of the reservoir, in a fine 25x25x5m cell size grid, were created to cover a wide range of reservoir fracture distributions and properties. The 1536 models were then upscaled to a 50x50x10m cell size grid assuming that, given similar responses from the faster (40 seconds vs 6.5 mins) coarse grids, the coarse grid will have almost the same clustering response as the fine grid and will save time. To test the impact of switching to the coarse grid for clustering a full set of fine grid simulations were also run and the inference of P10-P90 was compared between the coarse and fine grids. In this example it showed no difference between the fine and coarse models, although the work of Josset et al (2013) demonstrates how to compensate for occasions where numerical errors introduce a bias in clustering of models of different scales in different metric.

The example was run through the proposed workflow across all 4 stages of field lifetime with the following outcomes.

Appraisal stage (workflow 1-4)

For the purposes of appraising the differences/similarities in reservoir response between the ensemble of models a single injector/producer pair were added to the model in opposite corners of the grid (i.e. the maximum distance apart) and run for 2500 days. This well configuration was used on all 1536 models of the coarse grid. These were simulated and the difference in their production responses was evaluated as a dissimilarity matrix to total oil production based on a Euclidean distance, which is then projected into a metric space of models using MDS. Total oil production data was used to estimate the model dissimilarity matrix however the choice of production metric will have a strong impact on the clustering. A comparison of the clusters from total oil and oil rate data is given in Table 2 which shows that even for correlated data the clustering can sometimes be very different (there are few shared models between clusters).

The locations of wells will impact on the resulting production response and therefore the model similarity estimates (from MDS) thus we might expect variations in the clustering outcomes given a different number and/or location of wells. This is one complexity of using production data if there is significant variations between different well placements due to spatial variations in the permeability field of each realisation.

Models were clustered based on production response using a spectral clustering extension of k-medoids. The number of clusters is chosen by trading off the minimum number of clusters and the minimising the error between the P10-P50-P90 quantiles forecasts from the complete ensemble of 1536 models and the forecast from N cluster centre models, where N is the number of clusters, in this case 6 (Gopa et al, 2013). Figure 3 shows (a) the clustering projections in metric space and (b) the spread of forecasts of the 6 cluster centre models with those of the complete ensemble. Figure 4 shows the P10-P90 predictions (dots) from the 6 cluster centre models for both with (a) and without (b) spectral clustering, where the spectral clustering improves the clustering performance.

Initial reservoir development (Workflow 5)

The optimisation task is the placement and operation of an injector/producer pair over a 10 year period. Figure 5 shows the constraints on possible locations for the injector and producer which are defined as priors to the optimisation algorithm; the red boxes define the 3 drill locations for the production well and the blue box is the possible location of the injection well. The production well rates are also optimised as part of the scheme, with a total of 6 optimisation parameters (see Table 3).

The MOPSO algorithm was used to optimise well placement by trading off 2 objectives: (a) maximising total oil production and (b) minimising water rates. The ideal proportion of oil and water production is not specified beforehand, rather MOPSO is able to explore the trade-offs and a best solution can be chosen from the ensemble of models produced. The optimisation runs were performed for each of the 6 cluster centre models to represent the uncertainty.

The mean model responses for the Pareto front models are used to identify the optimal well locations for the available drill zones. The mean FOPT responses for the producer well locations are given in Figure 6 and this information was used to locate the production well at location $i=31$ and the injection well at $i=20$ (see Figure 7(a)) to maximise the mean total oil production.

History matching (workflow 7-8)

History matching was carried out for oil and water rates from the truth case, with added noise to represent realistic measurement errors, using 500 iterations of the MOPSO algorithm with the parameterisation given in Table 4. The 6 cluster center models from the previous step are used to represent the range of reservoir uncertainty by including these scenarios into history matching as a discrete distribution.

The entire ensemble of 500 models are then resampled using NA-Bayes to calculate posterior probability of the HM model, and come up with the Bayesian credible intervals. The results of NA-Bayes are used in Stage 4 to assess the uncertainty in the forecasts from any subsequent development decision based on the updated P10-P90 envelope from Stage 1 given the history match results.

Figure 7(b) shows the performance of each grid in history matching, as the least squares norm for the sum of the objectives and shows good matches can be achieved in more than one model cluster. The results show clusters 3 – 5 provide the best overall matches (Figure 7 (c & d)).

The NA-Bayes resampled models are used to identify the ensemble of models to be used in forecasting and optimisation. NA-Bayes resampled models preferentially excluded clusters 1, 2 and 6 due to a lower likelihood/high misfit (Figure 7(b)).

Mid field-life Optimisation (workflow 9)

The objective here is to optimise the location of a new infill well added to the reservoir after 8.2 years of production (3000 days), given the updated reservoir uncertainty estimated from history matching and NA-Bayes. The production well location is optimised by changing 4 parameters: the i, j location of the well, the length of the production casing (k) and the maximum liquid production rate. The injector is optimized based on i, j location. The parameter ranges for the new well are given in Table 5.

To account for uncertainty in the optimization, a subset of the NA-Bayes resampled models were chosen to represent the uncertainty, in this case the most resampled models are from clusters 3-5. A total of 300 iterations of the MOPSO algorithm were run to identify the best strategies, trading off maximising total oil production against minimising the maximum water rate, for each of the 3 geological scenarios.

Heat maps of the possible well locations are shown in Figure 8 for the mean and max production responses. These results show the optimal well location is around $i=4, j=36$. This location was chosen and the production is forecast using the resampled models from NA-Bayes to predict the P10-P90 confidence intervals. The results of forecasting are shown in Figure 9 against the true reservoir response which falls within the confidence bounds before and after history matching (and the resulting reduction in variance of the forecasts).

3.1 Case study 2: PUNQ_FRAC field

A new synthetic NFR field example was created, based structurally on the PUNQ-S3 field (Floris et al, 2001) called the PUNQ_FRAC field. The new field is a fractured carbonate oil field, with a gas cap and weak aquifer support, thus any development requires the use of injection wells to maintain pressure but cope with a high degree of uncertainty in water breakthrough times. The field is located offshore (with high associated development and operating costs) and is to be developed as a tieback to an existing platform.

Flow in PUNQ_FRAC is controlled by the fracture network permeability in a low permeability matrix that contains the bulk of the STOIP. The PUNQ_FRAC truth case was generated from a DFN model

based on fracture statistics collected from the image analysis of 4 outcrops in the Jandaira formation, Brazil. An example of part of one outcrop is illustrated in Figure 10(a) (van Eijk (2014)) showing the interpreted fractures overlaying the high resolution aerial photographs used to elicit them. Fracture data was collected from all 4 outcrops and processed in a similar way to Bisdom et al (2014). High resolution digital imagery of the outcrop were taken and subsequently analysed using image recognition tools to identify the fractures visible from background karstic features. Fracture properties such as length, spacing and orientation were extracted from the digital dataset to build statistical models of the combined outcrop data that could be used in creating a range of possible DFN models. The data is incomplete due to the finite resolution and visibility of the outcrop, therefore the resulting statistical distributions are incomplete, and issue mentioned in Bisdom et al (2014) and resolved by assuming a power law distribution of fracture sizes. This demonstrates a need for even richer, more multi-scale dataset (thin section data, multiple outcrop analyses) to improve the prior model.

Data from the Jandaira formation demonstrated spatial variations in length, orientation and fracture spacing across the outcrop area. Three main fracture sets could be identified from the digital data based primarily on fracture orientation, associated lengths and intensity. Ten possible secondary fracture sets were also identified that are associated with the primary sets, both of which are given in Table 6. From unique combinations of these (primary and secondary) sets, a range of different geological scenarios were created based on the observed fracture statistics. Fracture height, aperture and permeability were all unknown and could be estimated using previously observed relationships with length. Fracture height is estimated using a length/height ratio of 8:1. The fracture apertures are estimated based on the measured relationships with fracture length/height described in Klimczak et al., (2010) which are also used to estimate permeability from pre-existing relationships between permeability and aperture (Dershovitz, 2012).

The range of geological scenarios is controlled by (1) the primary fracture set orientation, (2) the secondary fracture set orientation, (3) fracture set intensity parameter and (4) the relative proportion of primary vs secondary fractures. Three primary fracture sets were tested as well as 3 relative proportions and 2 fracture intensity values (full set is given in Table 6). In total, 168 distinct geological scenarios

were created as DFN models. More complex scenarios could be created from the data to capture the apparent non-stationary nature of the fracture set distributions, however this is beyond the scope of this work and would entail the development of more sophisticated prior definitions. The truth case is a different model, created using a unique combination of 6 of the observed fracture sets and intensities to generate a DFN that is not represented by any one of the 168 scenarios used to describe the prior but still based on the underlying data.

Each DFN was subsequently upscaled into a 70740 cell (54 by 87 by 15) simulation grid using the Oda Gold method (See Figure 11(c)). This is a higher resolution grid than in the original PUNQ model to allow for better preservation of the permeability heterogeneity but still significantly lower resolution than the DFN model and associated fracture data. Oda Gold is used to account for sub-grid connectivity present in the fractures to be maintained in the upscaled grid.

To account for the natural variability in the spatial distribution of fractures for a given geological scenario while also conditioning the realisations to static well data, the produced 168 geological scenarios were then used as training images in FilterSim algorithm (Zhang et al, 2006) to create a set of 1680 realisations of the reservoir (10 for each training image), conditioned to a single appraisal well drilled in the field. This is a more rapid alternative to the seeding of multiple DFN realisations and their subsequent upscaling, which would require significantly larger compute times to complete.

4 Stage 1: Appraisal (workflow steps 1-4)

4.1 Approach description and results

The differences between the permeability field (KXX tensor only) of the 1680 geological realisations (representing the range of reservoir uncertainties) are evaluated as a dissimilarity matrix of the Euclidean distance between models, which is then projected into metric space using MDS and clustered using k-medoids. The permeability matrices compared in this case rather than production rates to remove the need to simulate all 1680 models and to find models with similar spatial distributions of fractures.

The number of clusters chosen for the next step is a subjective modeller's choice and is impacted by a number of choices in the MDS and clustering process. For this field, 10 clusters were selected as an acceptable trade-off of low error and a manageable number of scenarios. Figure 12 demonstrates the permeability clustering effectiveness, showing the spread of production forecasts (for a random production setup) of the 10 cluster centre models and the original 1680 realisations and the P10-P50-P90 interval for the complete data sets.

Stage 2: Initial reservoir development (workflow steps 5-6)

The optimisation task is to optimise the number and location of the production wells in the initial development. This is achieved by exploring the trade-off in maximising NPV and minimising the development cost for the number, placement and completion configuration of initial injector and producer wells. This represents a useful economic trade-off to understand, particularly for smaller, more capially constrained companies.

The number, location and operation of up to 12 possible producers plus the option of up to 5 injectors are optimised at this stage. Figure 11(a) shows the range of locations (black boxes) for the producers (P1-P12) with the injectors being located at fixed locations to provide pressure support. The field includes at least 2 injection wells due to a lack of an active aquifer, with an additional 3 optional injection wells (I2, I3 and I5) that are included as part of the optimisation. The full list and range of the optimisation parameters are included in Figure 11(b) (Stage 1 parameters) which include X/Y location of each well, the completion length (6 options are the completion in layers (a) 1-5, (b) 5-10, (c) 10-15, (d) 1-10, (e) 5-15, (f) 1-15), if the well is drilled or not drilled and the maximum liquid rate (BPD). The dashed boxes show the locations of future infill wells (see Stage 4 below).

The workflow uses the 10 cluster centre models from the previous stage to represent the reservoir uncertainty. To optimise under uncertainty, each development scenario will be run on all 10 grid realisations and the optimiser will be driven by the mean response of the ensemble. The MOPSO algorithm was used to optimise well placements by trading off 2 objectives, (a) maximising mean

project NPV and (b) minimising Net Present Cost (NPC) to represent a constraint on the affordability of the project in capital outlay and operating costs for each development decision. NPV is calculated based on a fixed well and tie back cost over a period of up to 9000 days

The Pareto front results of MOPSO show the non-dominated set of the mean model objective scores (each point is better in at least one objective score) from which we can identify the best candidate development solution as some trade-off between NPV/NPC (Figure 13). The field development problem is complicated as the development is marginal with many well configurations leading to negative NPV values, however 3 cashflow positive candidate developments (within the red region in Figure 13) are selected based to examine the capital outlay vs NPV trade off.

Figure 14 shows the spread of NPV for 3 candidate developments options – as min, max, mean values – for the range of (10) geological uncertainties. The mean NPV for the 3 options are quite similar, however the variance in NPV is significantly higher for Option 1 than Option 3, with a considerable portion of the distribution in negative NPV. The greater capital outlay of Option 3 here helps to de-risk the project and is the chosen development scenario.

New permeability data from the production wells was used to recondition the MPS models. The 10 cluster centre models are now used as training images in Filtersim to generate a new set of 100 realisations of the reservoir, conditioned to both the appraisal and development well data (8 additional wells) and better represent uncertainty. The 100 models are clustered again using permeability to identify representative models for the reconditioned ensemble. A more brute force approach would be to recondition the 168 original DFN realisations to the complete well data but the subset of 10 models were chosen to reduce compute times.

The clustering workflow was repeated and identified 6 clusters covering the range of permeability field uncertainty (Figure 15). A silhouette plot (Rousseeuw, 1987) for the 6 clusters is given in Figure 16, and demonstrates the majority of the clusters are made up of similar models with silhouette values above 0.8. Plotting the number of clusters vs the % of models with a silhouette value above 0.8 indicates that 6-7 clusters is the ideal number for this problem (Figure 17).

The resulting spread of production responses for the chosen well configuration from MOPSO is given in Figure 18(a) which shows the spread of results and the P10- P90 inference for the 100 reconditioned models is compared with the P10-P90 estimates from 6 cluster centre models. Clustering on permeability rather than production responses has led to a noticeable discrepancy between the production responses for the 100 models and the 6 clusters which have a much narrower spread of forecasts. As a comparison the P10-P90 estimates from clustering on FOPT are shown in Figure 18(b) where the discrepancy between the P10-P90 estimates were be minimised by increasing the number of clusters to 15 if the additional simulation runs were carried out.

5 Stage 3: History matching and forecasting (workflow 7-8)

The objective of this stage is to update the model uncertainties based on the history match response. History matching was carried out for field oil and water rate using MOPSO algorithm with the parameterisation given in **Table 7**, and the standard least squares misfit objective function.

The 6 cluster centre models from the previous step are used in history matching to account for prior geological uncertainty, given the available well data. To incorporate this uncertainty in this stage, the grids are again included as a discrete variable that switches between the permeability fields along with more traditional history matching parameters such as skin factors, porosity and permeability multipliers to provide the models with some flexibility to match.

To estimate the uncertainty based on history match quality, the misfit ensemble is resampled using NA-Bayes to calculate the Bayesian credible intervals. The results of NA-Bayes are used in Stage 4 to assess the uncertainty in the forecasts from any subsequent development decision.

NA-Bayes provided an estimate of the P10-P90 based on the forecasts of 8 resampled models from the complete misfit ensemble of 300 models. The number of realisations could be further reduced to 4 by excluding models with a low number of resamples from MCMC. The 4 chosen models still accounted for more than 95% of the variability of the P10-P90 oil and water production estimates from the 8

models but reduced the number of scenarios to test by half. These 4 models were chosen to account for the geological uncertainty in a second stage of optimization - the placement of infill wells in the field.

The forecasts of P10-P90 for FOPT are shown vs the true production from the field in Figure 19.

6 Stage 4: Mid/Late field life optimisation (Workflow step 9)

The objective of this final stage is to optimise the location of up to 2 new infill wells in the reservoir, added after 3000 days (8.2 years) of production, given the updated reservoir uncertainty estimated from history matching and NA-Bayes. The aim is to maximize NPV but minimize the risk in adding the new well to the field, given an already profitable base production rate. Therefore the trade-off to be explored is to maximize incremental NPV while minimizing variance in NPV (i.e. maximize the mean, minimize the variance) using MOPSO.

The geological uncertainty remaining in the field, post history matching is propagated into the optimization step using the results of NA-Bayes (resampled models). The 2 new wells are optimized using the same parameterization as at the appraisal stage. This resulted in 10 parameters, the grid X and Y coordinates (within the blue dashed boxes in Figure 11(a)), the well completion length, well rate and well status (drilled or not drilled) for each well.

A total of 300 iterations of the MOPSO algorithm were run across the 4 geological realisations taken from NA-Bayes to identify the best development option as a mean incremental NPV (i.e. NPV above that of continued production with the current wells for the same time period).

Figure 20 shows the Pareto front for mean incremental NPV vs incremental NPV variance which (as for the development stage optimisation), show a number of options that can produce negative value, demonstrating that careful selection of development options is required. The red square in Figure 20 represents the chosen development option where the well adds significant value but with moderate estimated risk based on the geological uncertainty.

The new development option was added to the model and then forecasted under uncertainty based on the results of NA-Bayes. The resampled ensemble of models from NA-Bayes analysis from Stage 3 were run forward with the new well and the overall forecast uncertainty is shown in Figure 21. These

forecasts are compared with the future production of the synthetic “truth” case for with and without the new wells to demonstrate the increase in production. The forecasted P10-P90 confidence intervals for the new scenario are narrow, owing to the optimization step that focused on minimizing the variance of the model responses, however the true production increase from the new wells is significantly less (black dashed line) than that forecasted, resulting in a negative incremental NPV in reality.

A critical observation from these results is that the true level of uncertainty in the model is no longer captured in the ensemble of models used in characterization and decision making. The successive steps taken to reduce the number of models required to capture the uncertainty has also reduced the accuracy of the uncertainty estimate. The overly narrow uncertainty bounds represent too great a reduction in forecast spread from the initial prior uncertainty (168 geological scenarios). This reduction came from removing too many models as a result of clustering and through an overly narrow range of forecasts from history matching which combined to affect the posterior confidence intervals. As Figure 18 demonstrated, a choice of clusters based on permeability similarity is not necessarily tied to a similar diversity of flow response and the results of clustering was narrower than the P10-P90 estimated from clustering on production. While in the first example the impact of uncertainty narrowing was not seen to impact on the ability to forecast, Table 2 did demonstrate that even changing from total oil to oil rate can impact on the clustering response such that there was little similarity in the populations of models in each cluster from both methods (even though the curves are completely correlated).

7 Discussion

This paper demonstrates a workflow (Figure 1) that both assesses uncertainty and optimises reservoir performance under uncertainty which aims to increase efficiency by finding a minimum set of models needed. Figures 22 and 9 demonstrate a reduction in uncertainty over time (as production data is added) in a similar way to the idealized model of Dromgoole and Speers (1992), with varying degrees of success in the robustness of forecasting.

Clustering of models throughout the field lifetime in addition to conditioning the models to static and dynamic data provided a way to propagate estimates of uncertainty from appraisal through the initial development of the field to later stage history matching and EOR decisions.

Clustering can be seen as a method for managing uncertainty by reducing the number of models to a manageable level (in these cases 6 geological scenarios), while trying to maintain the full variability inherent in large numbers of scenarios. This will always mean a trade-off in terms of the robustness of any uncertainty estimates, particularly away from existing well locations (as is shown in the infill well drilling of the PUNC_FRAC example), thus is a disadvantage of clustering as it artificially reduces the amount of uncertainty used in decision making.

History matching and posterior inference also reduces the prior uncertainty but by estimating the posterior based on new data so, while also reducing uncertainty it does so in a useful way (eliminating models that don't match observations). Avoiding reduction in uncertainty due to clustering should be addressed by more advance clustering techniques – k-medoids is not necessarily the best in class, and by the choice of what properties are used in developing the clusters.

Another significant improvement to the history matching workflow would be to include a method for geological parameterization and history matching of this reservoir model, which would retain the geological variability but treat it as a more continuous space of geological parameters which would better inform our understanding of the subsurface uncertainty.

This workflow provides a way to reduce the required number of models needed to account for the uncertainty from an initial large set models at appraisal stage, to a subset of models from clustering after the initial reservoir development and history matching which were used to optimise the infill well positioning under uncertainty.

8 Conclusions

This paper demonstrated a way to minimise the number of flow simulation runs through the field development workflows without compromising the uncertainty estimation and update this through the field development stages. This allows us to manageably (a) optimise development decisions, while

taking account of the uncertainty in the reservoir; and (b) history match across different reservoir models. Figures 9 & 22 showed how the initial estimates of uncertainty collapsed to a much narrower estimate of uncertainty after history matching.

The generalised workflow is similar philosophically to the ideas of Gawith and Gutteridge (1999). Our workflow has achieved initial success by combining a number of key technologies, namely multi-objective algorithms, Multi-dimensional scaling and NA-Bayes. Each of these techniques has been shown to reduce the total number of simulations required to capture model diversity from competing reservoir descriptions/development options. This is a first attempt to combine them all into one piece of work, which resulted in a reduction of simulation effort to less than 5000 iterations to attain a robust set of reservoir decisions and forecasts for a life-of-field workflow.

Care must be taken in the clustering steps to maintain a sufficiently wide variation of models to account of uncertainties.

Acknowledgements

We would like to thank a number of contributors to this paper. Foremost our sponsors (EoN, JOGMEC, BG, RFD) whose funding of the Uncertainty JIP has made this work possible. The analogue data for the synthetic case study were collected by Giovanni Bertotti, Kevin Bisdorn and John Reijmer and the team from Delft University and made their finding available to us and we also appreciate their advice on the data interpretation. We would like to thank contributions from Dr Mohamed Ahmed and Prof Sebastian Geiger (Heriot-Watt) for their input into fracture modelling and upscaling. Thanks also to Prof Jef Caers (Stanford University) for his personal communications that helped shape the work and Pallav Sarma (Chervon) for most useful remarks and comments on the work presented at ECMOR. Finally we would like to thank Epistemy, Schlumberger, FracMan Golder Associates and Rock Flow Dynamics for access to their software to carry out this work.

References

1. Abdollahzadeh, A., Reynolds, A., Christie, M., Corne, D. W., Davies, B. J., & Williams, G. J. (2012). Bayesian optimization algorithm applied to uncertainty quantification. *SPE Journal*, 17(03), 865-873.
2. Arnold, D., Demyanov, V., Tatum, D., Christie, M., Rojas, T., Geiger, S., & Corbett, P. (2013). Hierarchical benchmark case study for history matching, uncertainty quantification and reservoir characterisation. *Computers & Geosciences*, 50, 4-15.
3. Bentley, M., & Smith, S. (2008). Scenario-based reservoir modelling: the need for more determinism and less anchoring. Geological Society, London, Special Publications, 309(1), 145-159. <http://doi.org/10.1144/SP309.11>
4. Bisdom, K., Gauthier, B.D.M., Bertotti, G. and Hardebol, N.J., 2014. Calibrating discrete fracture-network models with a carbonate three-dimensional outcrop fracture network: Implications for naturally fractured reservoir modeling. *AAPG Bulletin*, 98(7), pp.1351-1376.
5. Borg, & Groenen. (2005). *Modern Multidimensional Scaling Theory and Application*. Springer Series in Statistics.
6. Christie, M., Demyanov, V., Erbas, D. (2006). Uncertainty quantification for porous media flows *Journal of Computational Physics* 217 (1), 143-158
7. Christie, M., Eydinov, D., Demyanov, V. and Talbot, J. [2013] Use of multi-objective algorithms in history matching of a real field. *SPE Reservoir Simulation Symposium*.
8. Dershowitz, W., & Einstein, H. (1988). Characterizing Rock Joint Geometry with Joint System Model. *Rock Mechanics and Rock Engineering*, vol. 21, 21-55.
9. Dershowitz, W., Hurley, N., & Been, K. (1991). Stochastic Discrete Fracture Modelling of Heterogeneous and Fractured Reservoirs.
10. Dershowitz, B. (2012) Fracture aperture correlations. *Proceedings from 46th US Rock Mechanics Geomechanics Symposium*. Session Chair Interlude for Paper 469
11. Dhillon, I., Guan, Y., & Kulis, B. (2005). A Unified View of K-means, Spectral Clustering and Graph Cuts. University of Texas at Austin.
12. Dromgoole, P., Speers, R. (1992). Managing Uncertainty in Oilfield Reserves, *Middle East Well Evaluation Review*, 12,
13. Elfeel, M.A., & Geiger, S. (2012). Static and Dynamic Assessment of DFN Permeability Upscaling. *EAGE Annual Conference & Exhibition SPE Europec*. Copenhagen, Denmark .
14. Elfeel, M.A., Jamal, M., Enemanna, C., Arnold, D., & Geiger, S. (2013). Effect of DFN Upscaling on History Matching and Prediction of Natural Fractured Reservoirs. *75th EAGE Conference and Exhibition incorporating SPE EUROPEC 2013*. London.
15. Elsheikh, A. H., Demyanov, V., Tavakoli, R., Christie, M.A., Wheeler, M, F. (submitted 2014) Calibration of channelized subsurface flow models using nested sampling and soft probabilities" accepted for publication in *Advances in Water Resources*
16. Floris, F.J.T, Bush, M.D., Cuypers, M, Roggero, F., Syversveen, A-R. (2001). Methods for quantifying the uncertainty of production forecasts – a comparative study. *Petroleum Geoscience*, 7 Supp. 87.
17. Gawith, D. E., & Gutteridge, P. A. (1999). Decision-driven reservoir modelling: The next big thing. In *SPE symposium on reservoir simulation* (pp. 131-134).

18. Gomez, S., Gosselin, O., & Barker, J. W. (2001). Gradient-based history matching with a global optimization method. *SPE Journal*, 6(02), 200-208.
19. Gopa K., Demyanov V., Ahmed Elfeel M. and Arnold D. (2013) Uncertainty Quantification with Application of Distance Metric in Naturally Fractured Reservoirs Modelling, 6th Saint Petersburg International Conference and Exhibition
20. Hajizadeh, Y., Christie, M. A., & Demyanov, V. (2010, January). History matching with differential evolution approach; a look at new search strategies. In *SPE EUROPEC/EAGE Annual Conference and Exhibition*. Society of Petroleum Engineers.
21. Hajizadeh, Y., Christie, M. and Demyanov, V. [2011] Towards Multiobjective History Matching: Faster Convergence and Uncertainty Quantification. *SPE Reservoir Simulation Symposium*. The Woodlands, Texas, USA. doi:10.2118/141111-MS
22. Josset, L., Lunati, I. (2013). Local and Global Error Models to Improve Uncertainty Quantification, *Mathematical Geoscience*, 45:601–620
23. Jung, A., Fenwick, D., & Caers, J. (2013). Training Image-based Scenario Modeling of Fractures Reservoirs for Flow Uncertainty Quantification. *Computational Geoscience*.
24. Jung, A., Fenwick, D., & Caers, J. (2013). Updating Uncertainty in the Conceptual Geological Representation of Fractured Reservoirs Using Production Data. Presented at the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013, London, UK: EAGE. <http://doi.org/10.3997/2214-4609.20130924>
25. Kennedy, J. and Eberhart, R. [1995] Particle swarm optimisation. *Proceedings of the IEEE International conference on neural networks*. 4, 1942-1948. Piscataway, New Jersey: IEEE Service Center.
26. Klimczak, C., Schultz, R.A., Parashar, R., Reeves, D.M. (2010) Cubic law with aperture-length correlation: implications for network scale fluid flow. *Hydrogeology Journal*. June 2010, Volume 18, Issue 4, pp 851-862.
27. La Pointe, P. R., & Fox, A. (2011). Quantification of Conceptual and Parametric Uncertainties in Fractured Reservoir Models, DOI:10.1306/13301407M961409. In Y. Z. Ma & P. R. La Pointe (Eds.), *Uncertainty analysis and reservoir modeling* (Vol. 96, pp. 57-76). Tulsa, Oklahoma, USA: AAPG.
28. Maucec, M., Cullick, S., Shi, G. (2011). Geology-guided Quantification of Production-Forecast Uncertainty in Dynamic Model Inversion. *SPE Annual Technical conference and Exhibition*, Colorado. Society of Petroleum Engineers.
29. Mohamed, L., Christie, M. and Demyanov, V. (2010) Comparison Of Stochastic Sampling Algorithms For Uncertainty Quantification SPE-119139-PA. *SPE Journal*, 15(1), 31-38.
30. Oliver, D. S., & Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1), 185-221.
31. Park, H., Scheidt, C., Fenwick, D., Boucher, A., & Caers, J. (2013). History matching and uncertainty quantification of facies models with multiple geological interpretations. *Computational Geosciences*, 1-13.
32. Rojas, T., Demyanov, V., Christie, M., & Arnold, D. (2012). Reducing Uncertainty in Modelling Fluvial Reservoirs by using Intelligent Geological Priors. In *Geostatistical Congress*, Oslo.
33. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.
34. Sambridge, M. (1999). Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space. *Geophysical Journal International*, 138(2), 479-494.

35. Scheidt, & Caers. (2007). Using Distances and Kernels to Parameterize Spatial Uncertainty for Flow Applications. *Petroleum Geostatistics 2007*. Cascais, Portugal.
36. Scheidt, C., & Caers, J. (2008). *Representing Spatial Uncertainty Using Distances and Kernels*. International Association for Mathematical Geology.
37. Scheidt, C., & Caers, J. (2009 December). Uncertainty Quantification in Reservoir Performance Using Distances and Kernels Methods - Application to a West Africa Deepwater Turbidite Reservoir. *SPE Journal*.
38. Shirangi, M. G., & Durlofsky, L. J. (2015). Closed-Loop Field Development Optimization Under Uncertainty. Society of Petroleum Engineers. doi:10.2118/173219-MS
39. Teich, J. (2001). Pareto-front exploration with uncertain objectives. In *Evolutionary multi-criterion optimization* (pp. 314-328). Springer Berlin Heidelberg.
40. Van Eijk M. (2014), Analysis of the fracture network in carbonate rocks of the Jandaira formation in northeast Brazil, MSc thesis Delft University of Technology, unpublished.
41. Zhang: Zhang T, Switzer P, Journal AG (2006) Filter-based classification of training image patterns for spatial simulation. *Math Geol* 38(1):63–80

LIST OF FIGURES

Figure 1: Workflow overview from this paper showing all steps used throughout and the number of simulations required. Each step's outcome is referenced to a figure in this paper demonstrating the outcome/value. The deliverables from each step are given, ■ representing model forecasts for uncertainty and ■ representing development decisions based on simulation results.

Figure 2: Workflow for the extraction and use of prior data from fractured analogue outcrop for use in the creation of DFN models.

Figure 3: Projection of the 1536 models in MDS space shown in (a), colour coded for the 6 clusters. The flow responses for the ensemble of DFN realizations are shown in (b) and colour coded according to the cluster. The flow response of the 6 cluster centres representatives (dotted lines).

Figure 4: Robustness of P10, P50 and P90 quantiles obtained from 100 runs of the method for Cumulative Oil production (a) with spectral clustering and (b) without spectral clustering. The dots in both cases come from the prior uncertainty obtained from DOE.

Figure 5: The parameterisation of well placement along the i axis for producer and injector. The producer can be located in the red drill zones, while the injector can be located in the blue region.

Figure 6: Distribution expected (likely) cumulative production for all possible producer well locations, I , from all the model cluster centre DFN scenarios.

Figure 7: Results of history matching for truth case model (iteration 313) where (a) shows the well location, (b) and (d) shows the history data and match quality for (b) Field oil rate (FOPR) and (d) Field water rate (FWPR). The quality of match is influenced by the cluster choice as demonstrated by (c) which shows the misfit (match quality) vs cluster.

Figure 8: Heat maps of the combined model cluster samples showing, for both the i and j grid index, (a) the number of samples per cell, (b) the minimum achieved water production rate, (c) the average and (d) maximum total oil production.

Figure 9: Forecast of chosen new well location ($i=4, j=36$) for the truth case model for Total oil (FOPT (Red)) and water rate (FWPR (blue)). A comparison between the predicted average, predicted maximum and actual (truth case) FOPT and FWPR is given in the table. Initial estimates of P10-P90 ranges during appraisal are updated from the history matched results using NA Bayes algorithm (Sambridge, 1999) then forecast for the new well. P10-P90 reduces significantly upon the addition of production data and the forecast bounds the true production from the reservoir.

Figure 10: Aerial photo panel (after van Eijk, 2014) showing variations spatially in the fracture orientations, lengths and intensities for one of 5 analysed outcrops in the Jandiara study area. (a)

Fractures were identified from high resolution aerial photographs and are shown in yellow. The high resolution photo-panel is superimposed onto Google earth (2016) base image. (b) shows wider view of complete outcrop. Results from all 4 outcrops were used to create the fracture statistics for use in DFN model.

Figure 11: Overview of the PUNQ_FRAC field example and its parameterisation. (a) 2D top map of the PUNQ_FRAC field showing the oil leg (red) with a small aquifer (dark blue) and gas cap (light blue). Black squares define the parameter ranges for wells in initial development optimisation (stage 2). The production well names are given as P1-P12 and the location of the chosen development plan from Stage 2 are shown as hexagons. The blue dashed zone represents the region targeted for infill well drilling after history matching (Stage 4) for 2 wells P13, P14. (b) lists the parameter and their ranges used in the optimization stages (c) 2 different fracture network scenarios, created using the DFN and upscaled using Oda Gold.

Figure 12: Spread of oil production for all 1680 models (grey dots) vs the 10 cluster centre models (black dots) from initial model clustering. Blue, Green and Red lines represent P10-P90 interval for the 1680 models.

Figure 13: Mean NPV vs Mean NPC cross plot for sampling from MOPSO. Pareto front models (orange dots) show the trade off in increasing project investment vs NPV return. The Red box identifies 3 candidate development scenarios selected based on cost/value trade-off

Figure 14: Uncertainty in NPV of 3 development options. Options 1-3 are given with the central NPV value being the mean NPV of all 10 cluster models along with the min and max values of the ensemble. In this case the more expensive investment is less risky in terms of NPV, with the majority of the spread lying above NPV=0.

Figure 15: MDS projection of model similarity (Permeability) of 100 realisations where clustering has been applied to identify representative cluster centre models. Clusters can be identified by colour with cluster centre models shown using squares. Cluster centre models, with the exception of the green cluster, appear to be highly representative with similar models.

Figure 16: Silhouette plot for the 6 cluster centre models. Model similarity to other members of the cluster is measured as the silhouette value, with a value > 0.8 being chosen as a high similarity and therefore clustering robustness. Most models are above 0.8 in this case.

Figure 17: Number of model clusters vs % of the 100 conditioned models (post development) with a silhouette value > 0.8 in those clusters. The plot shows a peak around 6-7 clusters, with a drop and gradual increase in model similarity.

Figure 18: Cumulative Oil production (FOPT) for all 100 grid models (grey dotted lines) and the results of clustering for (a) Permeability field and (b) Cumulative oil production. The cluster models are shown (black dots) as are the P10-P90 forecasts for clustering (dashed lines) and the entire 100 models (solid lines). P10-P90 estimates from the permeability based clustering are offset from those of the entire ensemble due to the narrow spread of the 6 clusters. By comparison the clustering on FOPT uses spectral clustering to minimise the error in P10-P90, which enforces a good match but requires additional clusters to be selected based on the alternate metric. The 15 clusters are seen to cover the spread of oil production well in comparison to the 6 clusters selected based on permeability.

Figure 19: Spread of prior (dashed lines) and posterior (solid lines) P10-P50-P90 estimates shows the reduction of uncertainty based on the addition of production data, with the true field cumulative production lying within posterior confidence interval.


Figure 20: Mean NPV vs NPV variance crossplot for the results of MOPSO. The Pareto front (orange dots) show the trade-off in increasing the mean value and decreasing the variance in NPV (based on the NAB inference). The red square represents the chosen development scenario for the field.

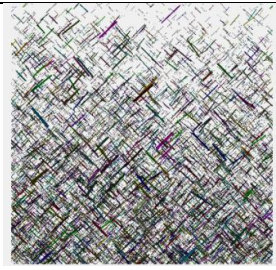
Figure 21: Cumulative oil production forecasts for with (dashed lines) and without (solid lines) infill wells. Posterior P10-P50-P90 estimates from NAB are plotted as blue, green and red lines vs the true values (black lines). The infill well forecasts diverge significantly from the true increase in production (truth case) in comparison with the forecasts without new wells (which performed well).

Figure 22: Reduction in uncertainty in the P10-P90 intervals over field lifetime. The min-max range of the entire P10-P90 ensemble is shown as a comparison. The initial P10-P90 estimates for the 10

clusters are adjusted and narrowed by the conditioning to the new well data and the identification of 6 new clusters. The true reservoir is located outside of the 6 cluster P10-P90 range until these are history matched to production data for the first 2850 days. The results show a reduction in uncertainty in the posterior inference from history matching + NA-Bayes (which is good) but also from the 2 stages of clustering (removing uncertainties).

Table 1: Geological scenarios tested in the DFN model. Scenarios are based on the proportion length and density of 2 fracture sets that vary according to all unique combinations of cases listed above. The 2 fracture sets feature 4 different mean fracture lengths, 6 fracture proportions from 0-1) and 4 min and 4 max density values values that define the variation in fracture density from a min to max value (top to bottom of DFN). All resulting models are upscaled into grid permeability values for the simulation grid.

Fracture set	Length variation (l) (power law)	Proportion of fracture set (prop())	Fracture density (d)
Parallel to main stress (NS)	Cases run	Cases run	Linear variation along j direction:
	10m 20m 30m 40m	0.0001 0.2 0.4 0.6 0.8 1	
Oblique to main stress	Cases run	Cases run	Cases Run Min cases = 0.03, 0.04, 0.05, 0.06
	10m 20m 30m	Prop(obliq) = 1-Prop(vert)	



40m

Max cases = **0.08, 0.09, 0.1, 0.11****Calculated Fracture properties**

Aperture (A)

$$A = 0.00002 * \frac{l}{2} \text{ (correlated with length)}$$

Permeability (Kfr)

$$k_{fr} = 10^{15} * \frac{A^2}{12} \text{ (Cubic relation to aperture)}$$

Table 2: Comparison of model cluster for MDS clustering based on Total oil production, FOPT (given by the columns) and by Field oil rate, FOPR (given by the rows). The number of common models between each of the 6 clusters for FOPT and FOPR is compared. No one cluster is maintained by changing from FOPT to FOPR.

	1	2	3	4	5	6	TOTAL
1	92	93	11	32	32	50	310
2	94	33	7	65	3	102	305
3	7	57	1	0	104	0	169
4	36	122	5	15	67	22	268
5	0	23	0	0	169	0	192
6	49	2	2	149	0	89	291
TOTAL	278	330	27	262	375	263	1535

Table 3: Optimisation parameter ranges for simple box case study. Ranges listed in brackets, fixed values listed and single numbers. The producer i ranges are given in 3 brackets defining the 3 drill zones defined in Figure 4.

	i location	j location	casing length (number of cells in k)	maximum liquid production rate (BPD)
Producer	(1-10, 15-25, 30-40)	(30-40)	(2-20)	(500 – 2500)
Injector	(1-40)	(1-15)	20	voidage replacement

Table 4: Table of history match parameters used by PSO algorithm

History match parameter	Parameter (prior) range
Vertical permeability multiplier	0.5 – 1
Horizontal permeability multipliers (X)	0.5 – 1
Horizontal permeability multipliers (Y)	0.5 – 1
Porosity	0.01 – 0.25
Simulation grid file (cluster)	1 – 6
Skin factor (producer P1)	-5 – 2
Skin factor (injector I1)	-5 – 2
Residual oil saturation (Sor)	0.05 – 0.35
Critical water saturation (Swc)	0.05 – 0.35
Corey oil exponent	1 – 5
Corey water exponent	1 – 5

Table 5: Parameterisation and ranges for optimisation run

Optimization parameter	Parameter range
i location	1 – 27
j location	30 – 40
k casing thickness	2 – 20
Maximum liquid rate (Prod)	500 – 2500

Table 6: Geological uncertainties in fracture set orientation, proportion and intensity based on the Jandaira formation outcrop data. A total of 168 unique combinations of the fracture orientation, intensity and relative fracture proportions. Each scenario was used to create a DFN based on fracture length statistics created by van Eijk (2014) based on drone aerial photography analysis.

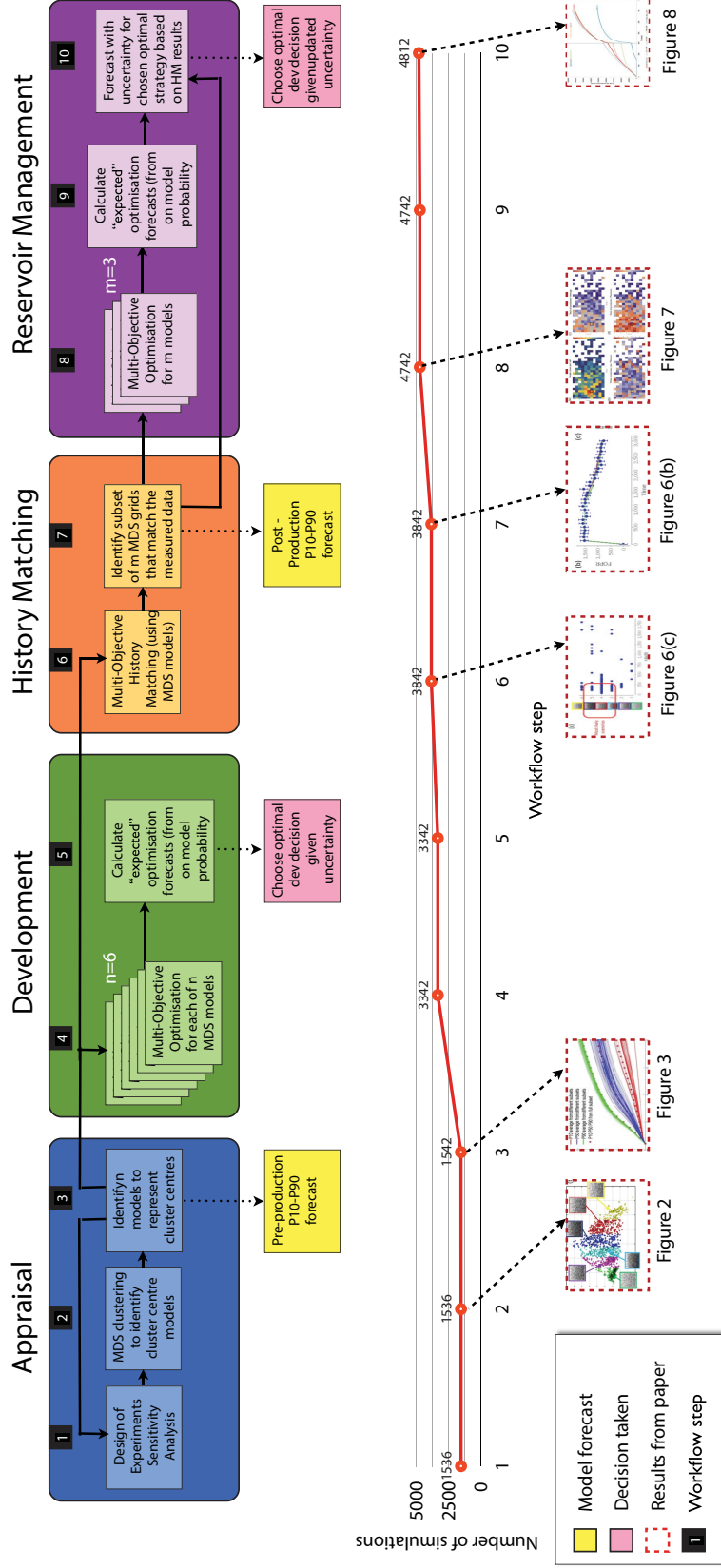
Main fracture set orientations	Secondary fracture set orientations	Fracture intensity (fraction)	Proportion of primary fracture set (%)
90-180	65	0.2	90
30-120	30	0.1	70
65-155	30-120		50
	30-65		
	65-120		
	120		
	155		
	155-65		
	30-155		
	120-155		

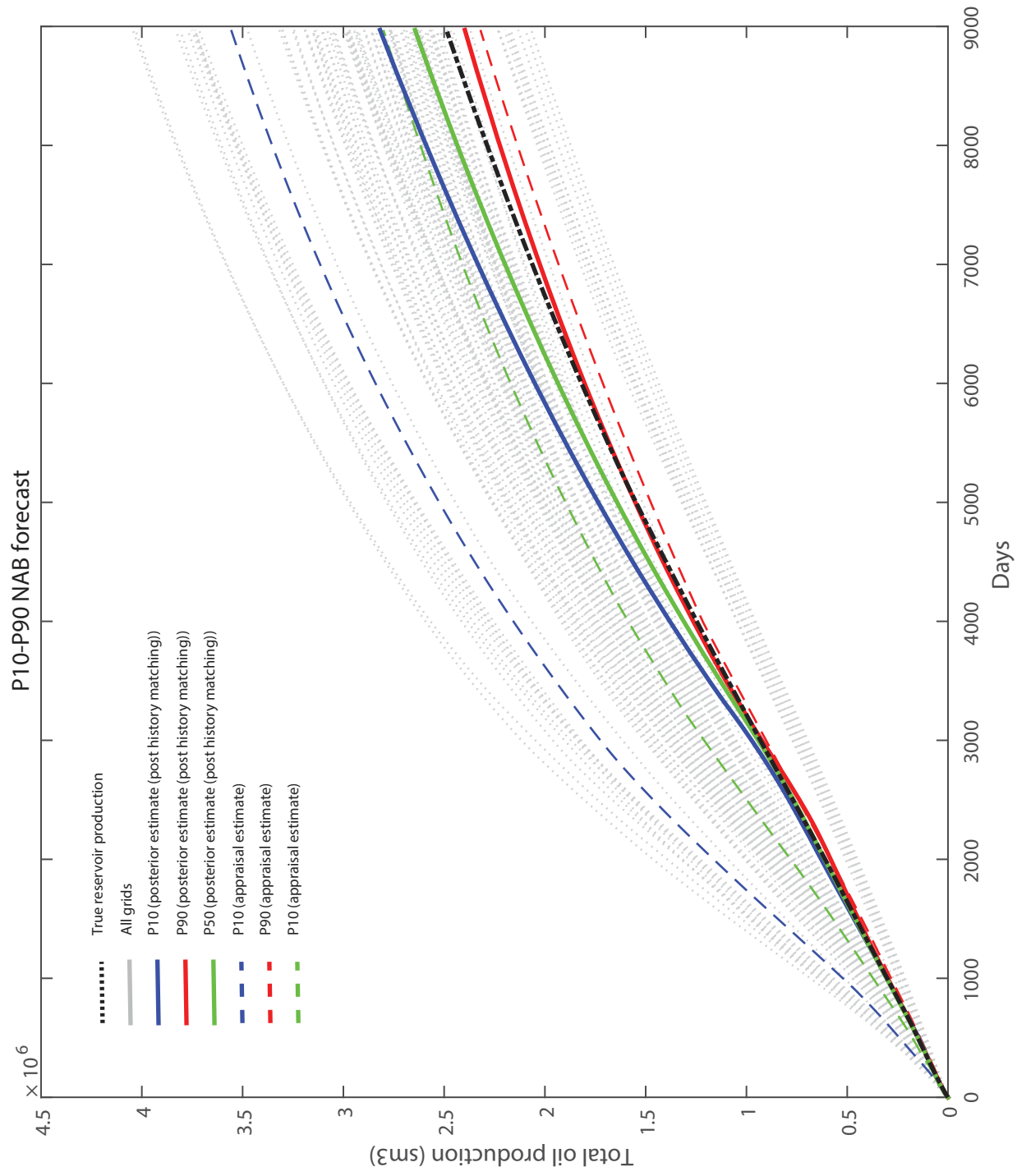
Table 7: Table of history match parameters used by PSO algorithm

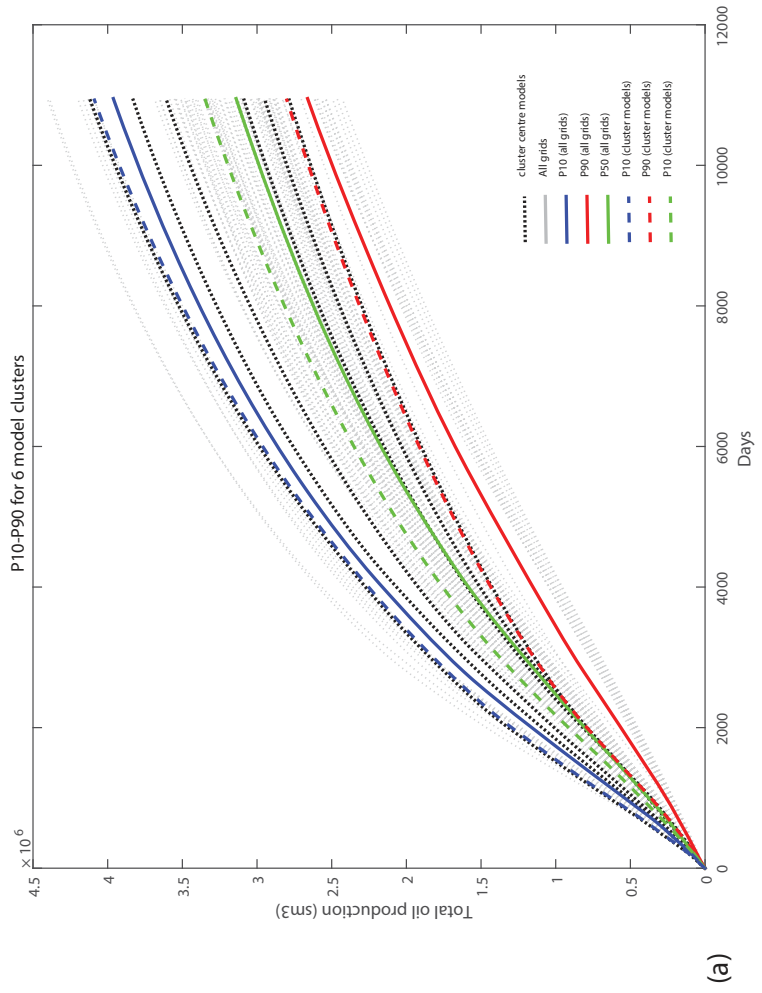
History match parameter	Parameter (prior) range
Vertical permeability multiplier	0.5 – 1
Horizontal permeability multipliers (X)	0.5 – 1
Horizontal permeability multipliers (Y)	0.5 – 1
Porosity	0.01 – 0.25
Simulation grid file (cluster)	1 – 6
Skin factor (producers)	-5 – 2
Residual oil saturation (Sor)	0.05 – 0.35
Critical water saturation (Swc)	0.05 – 0.35
Corey oil exponent	1 – 5
Corey water exponent	1 – 5

Highlights

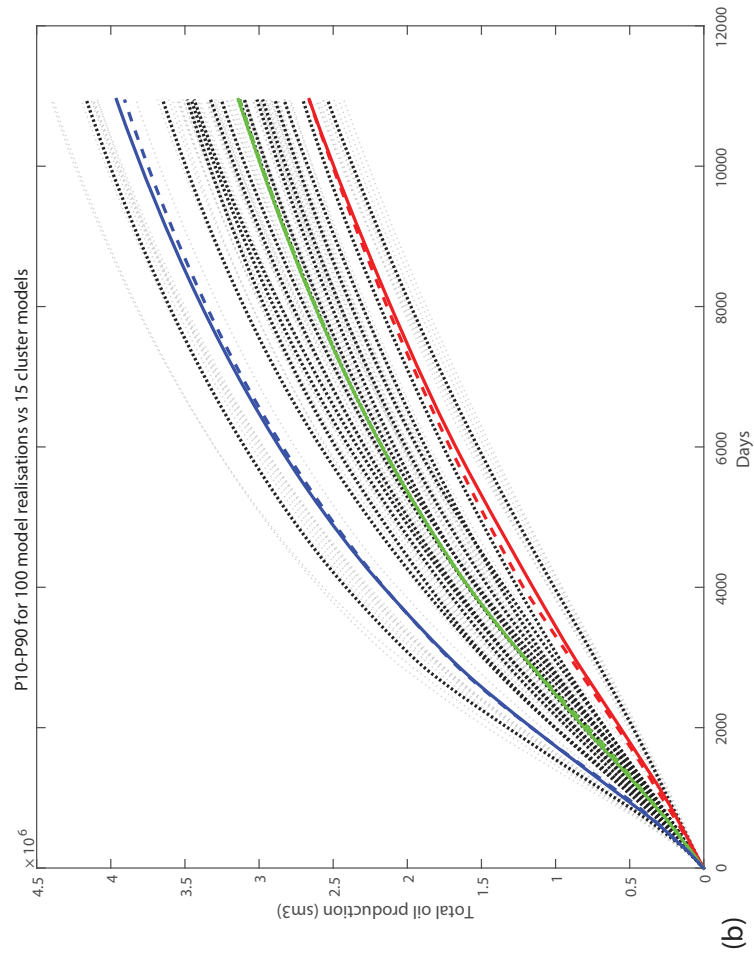
- Proposed workflow for propagating throughout field lifetime using a combination of techniques
- Key components are model classification using MDS, multi-objective history matching and optimisation using MOPSO and Bayesian inference
- Workflow aims to achieve results of MCMC approaches but with fewer simulations to fit the needs to oil company teams where CPU time is limited – under 5000 simulations
- MO history matching and Bayesian inference reduces compute time in post-production phase, MDS classification reduces workload in preproduction, MOO under uncertainty finds optimal across ensemble of models.
- Applied to 2 case studies, 1 a sector model, 2 a full field example



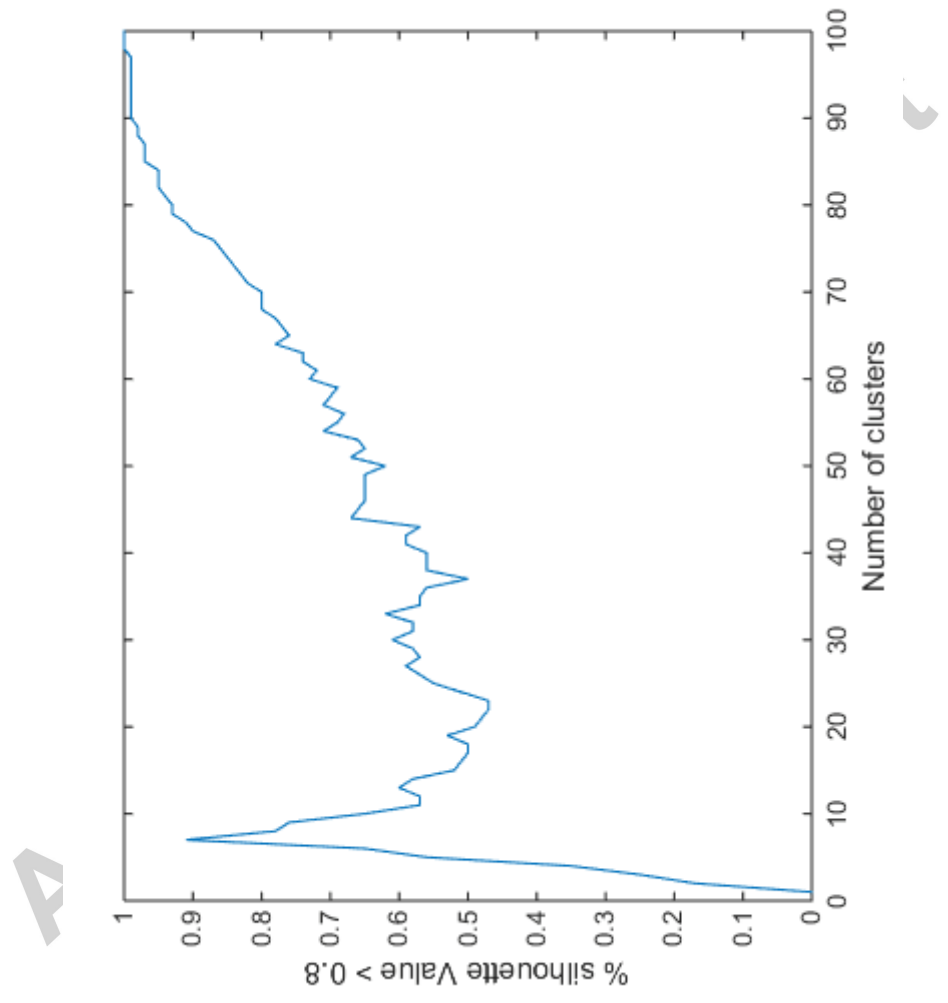




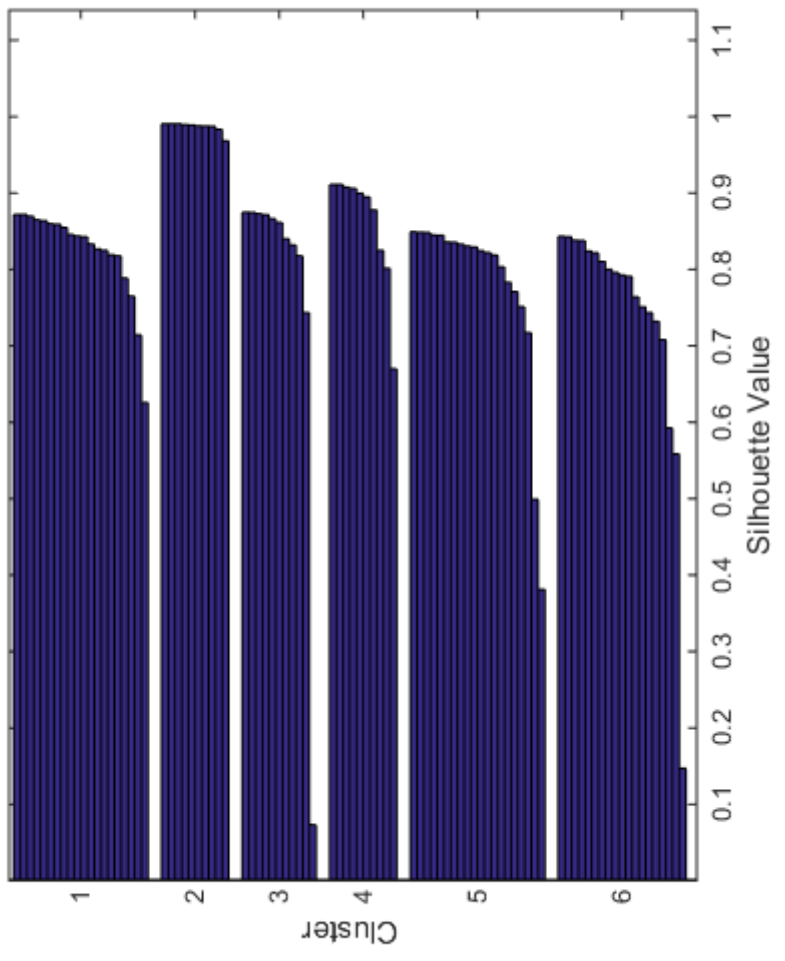
(a)



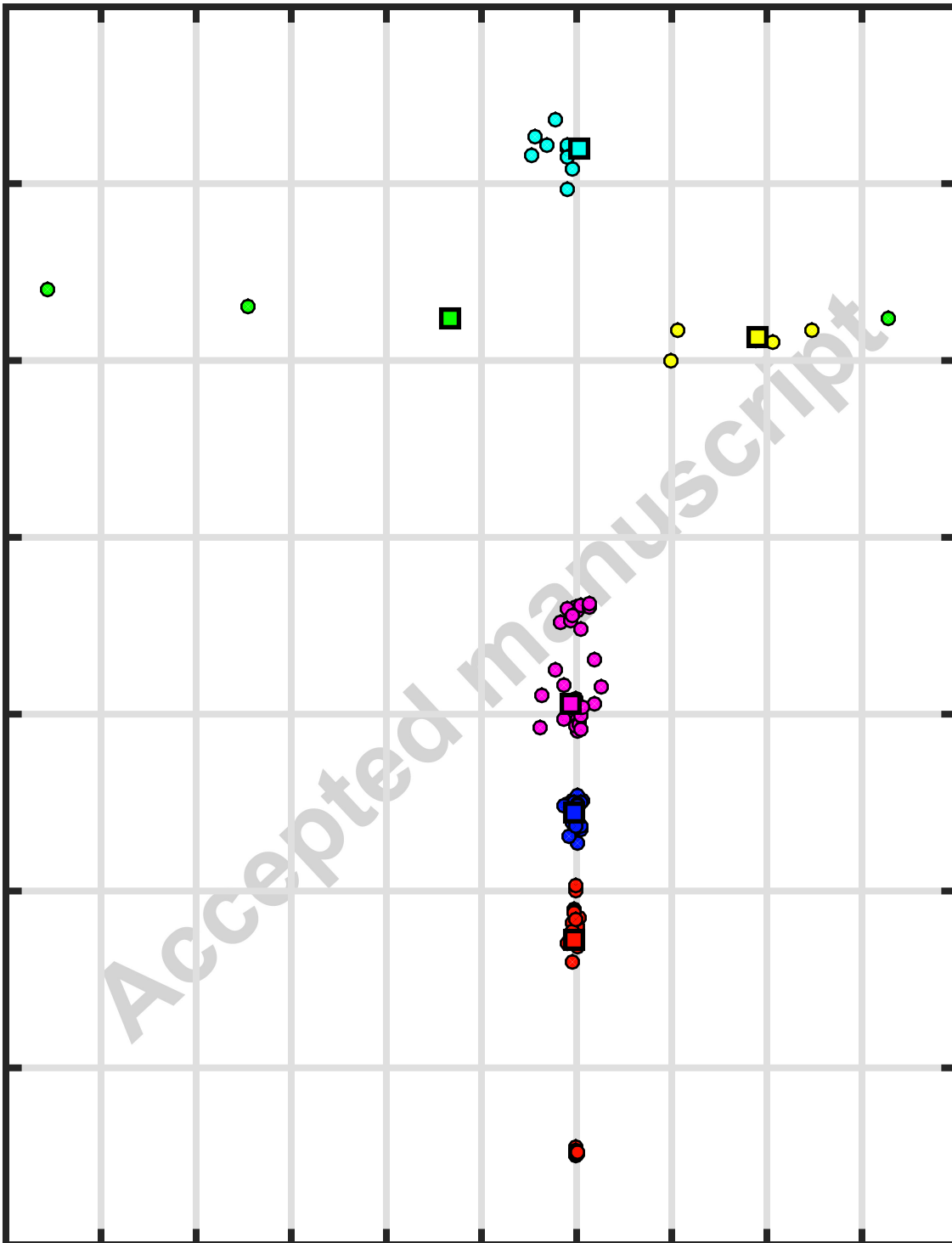
(b)



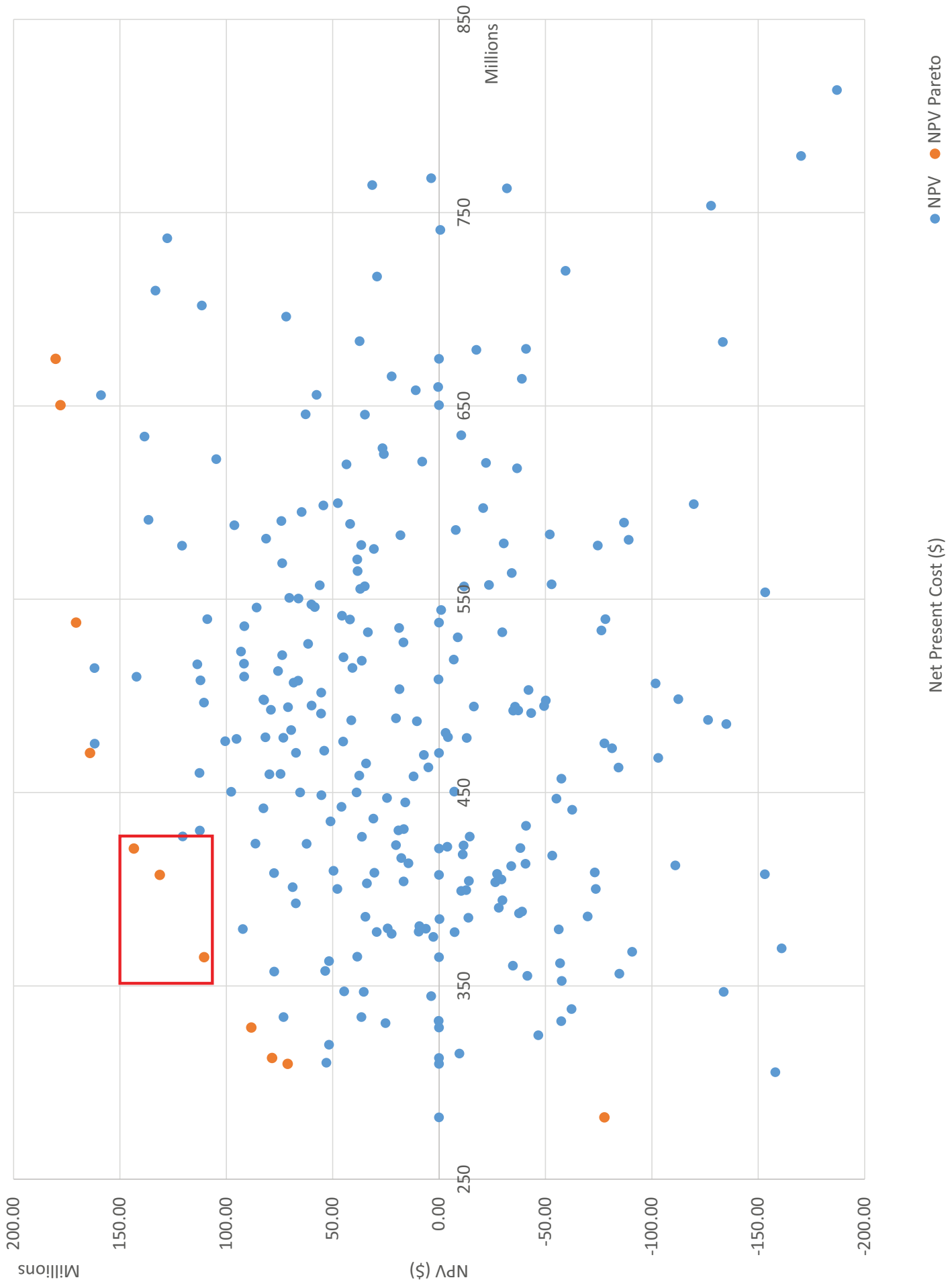
A

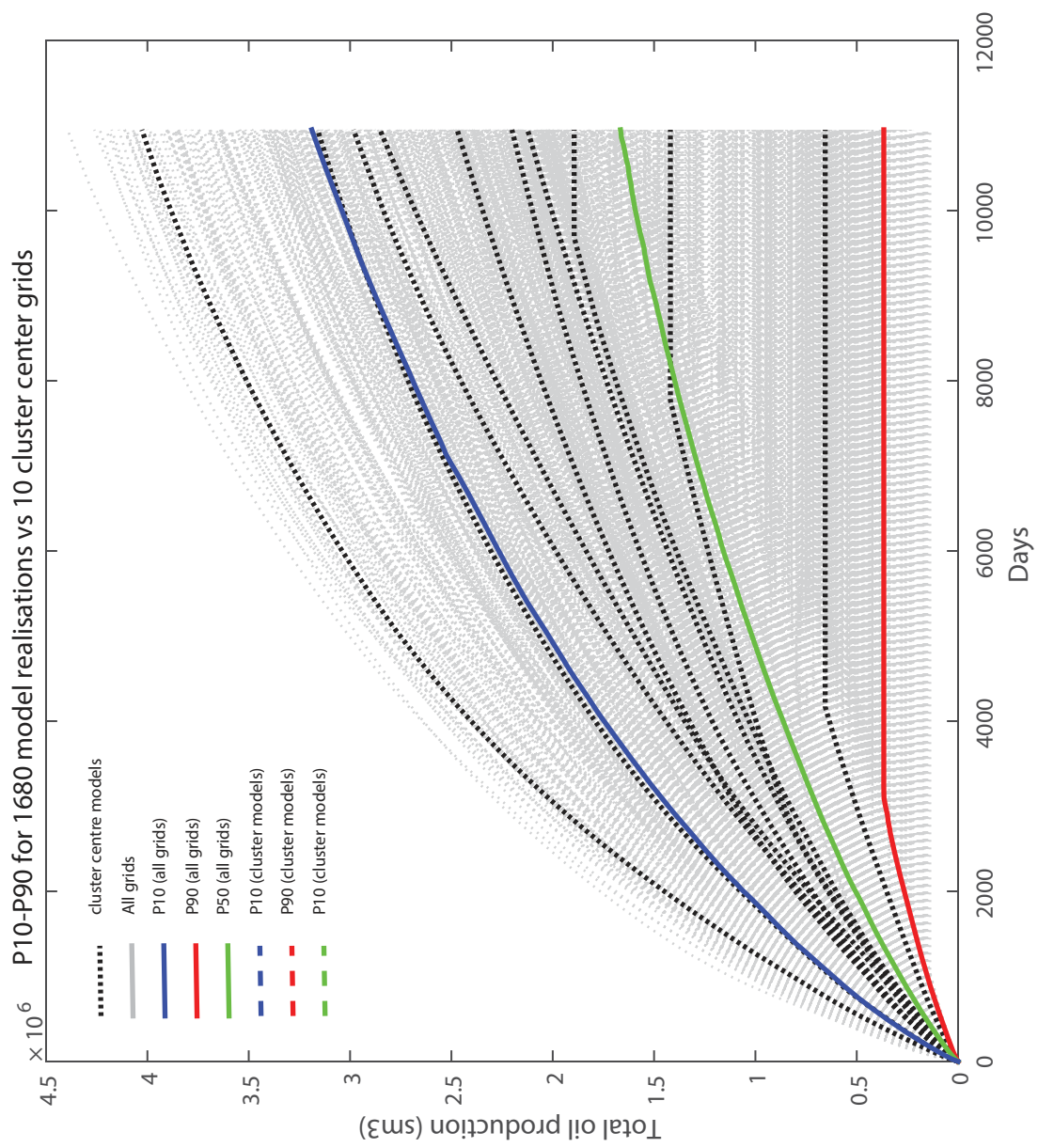


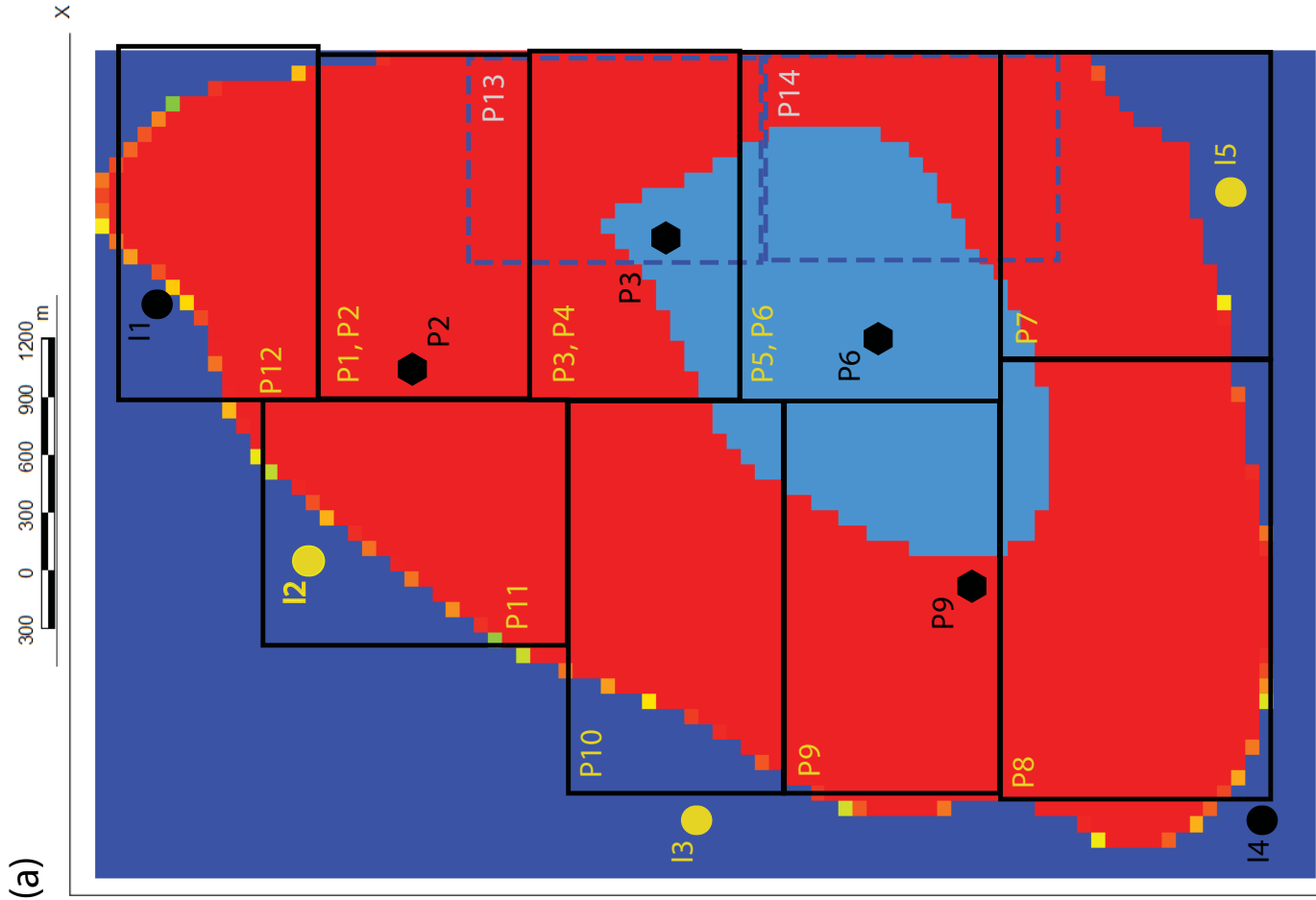
6











(b)

	X	Y	Z	OPEN or SHUT	Rate
Producers	P1 - P12 (black box)		(1 - 6)	(0 - 1)	(300 - 1500)
Injectors	-	-	-	(0 - 1)	-

Stage 1

Producers	P13 - P14 (dashed box)	(1 - 6)	(0 - 1)	(300 - 1500)
-----------	------------------------	---------	---------	--------------

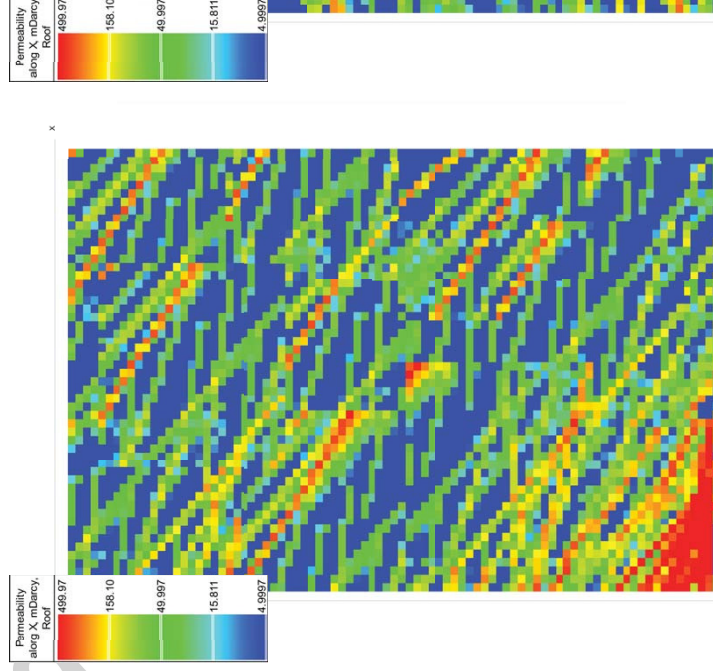
Stage 3

Z value	Z cell range
1	(1 - 5)
2	(5 - 10)
3	(10 - 15)
4	(1 - 10)
5	(5 - 15)
6	(1 - 15)

Completion length (Z) is parameterised as a discrete uniform distribution of 6 options, defining the upper and lower limits of the completion.

The table to the left lists the 6 parameter values and their equivalent cell ranges in the z direction. The model has a total of 15 cells vertically.

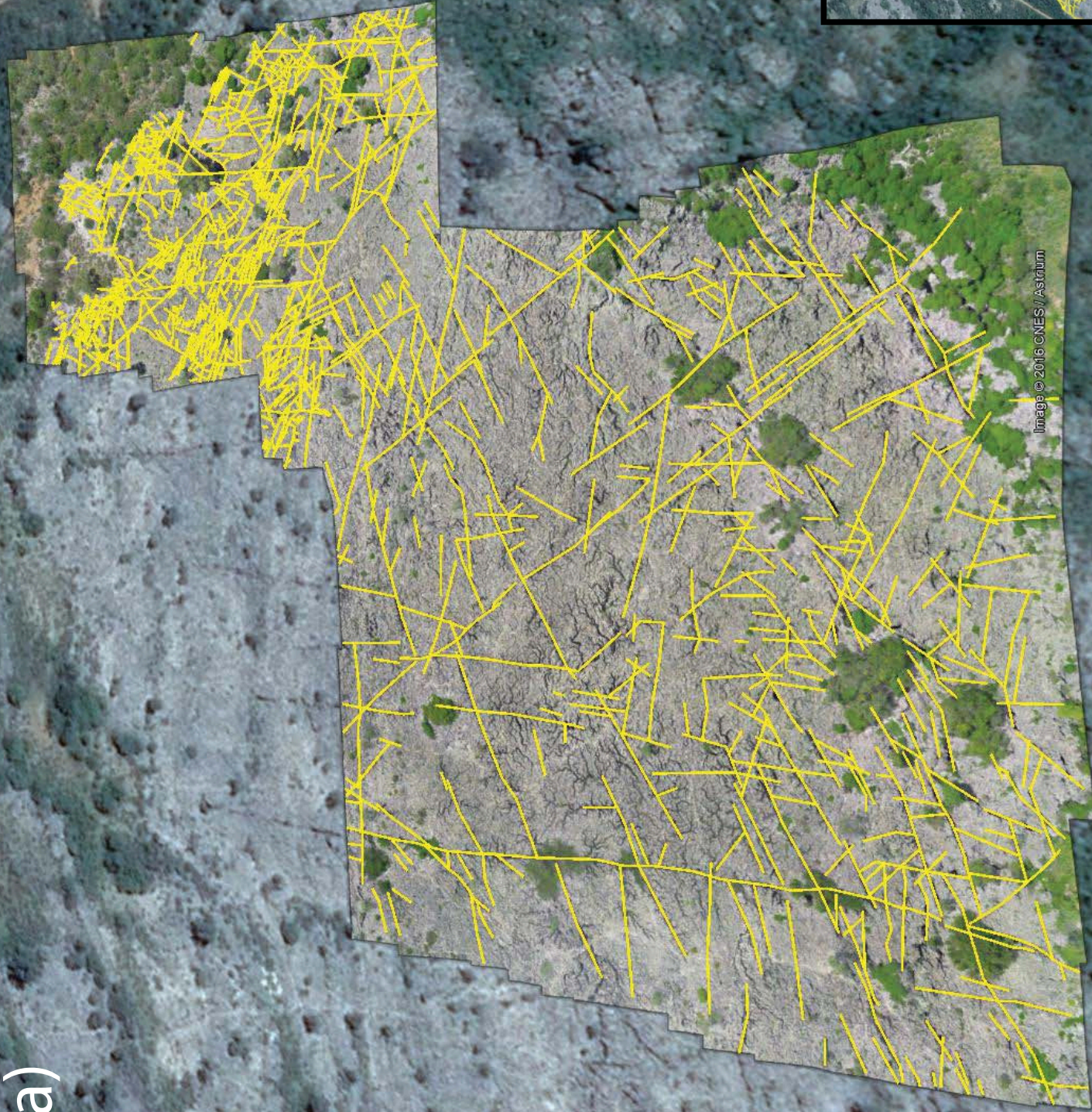
(c)



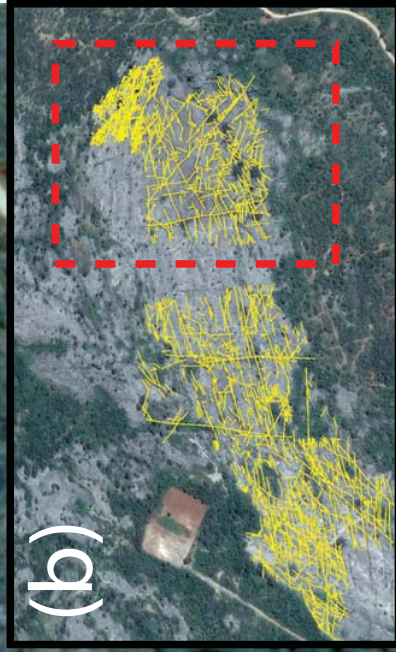
Scenario 1

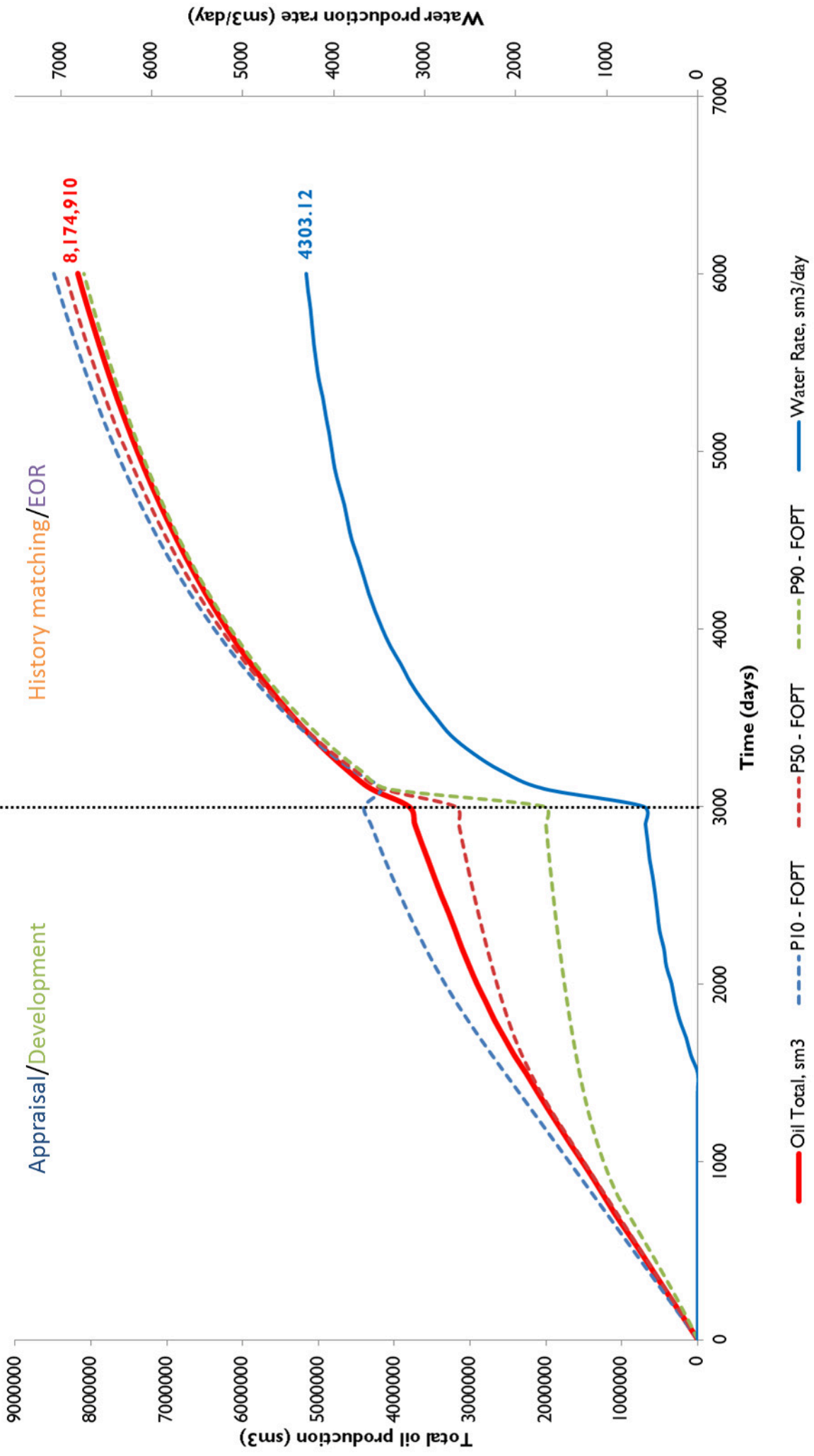
Scenario 2

(a)

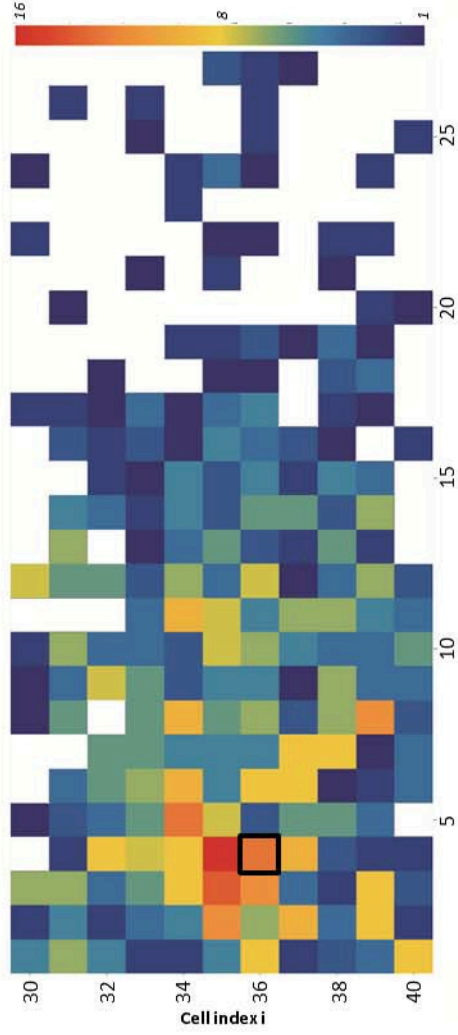


(b)



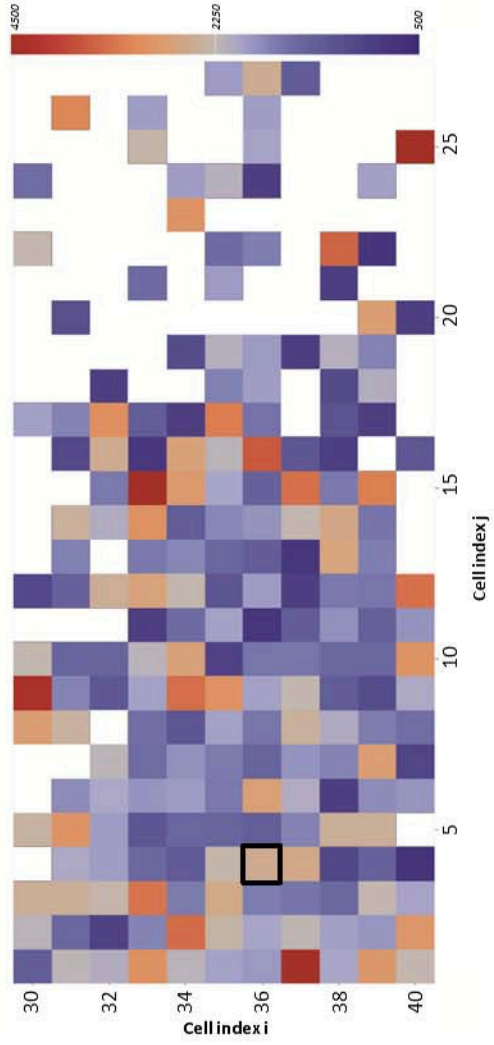


(a) Sample count per cell



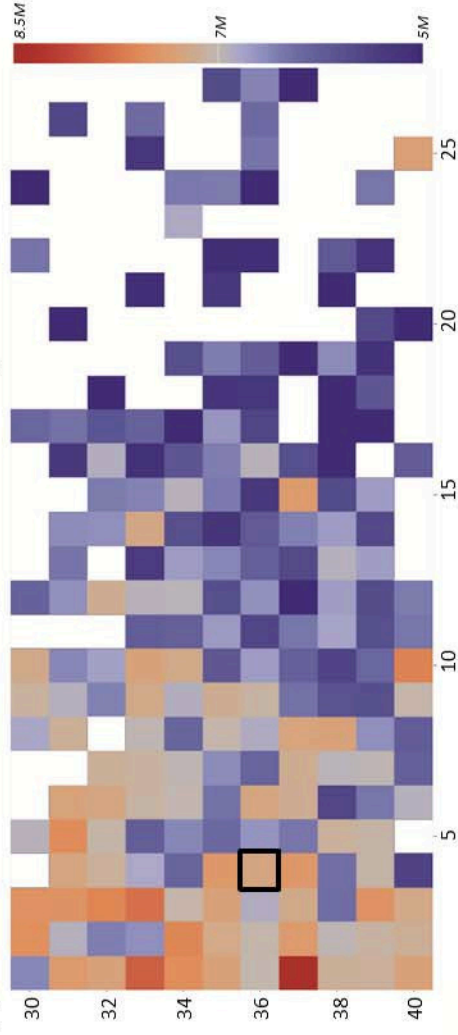
(b)

Water production rate: Minimum



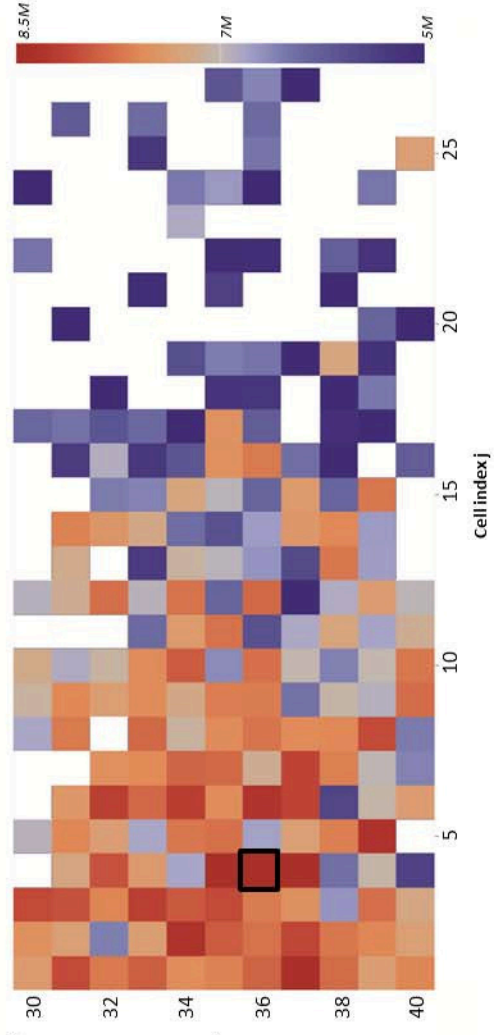
(c)

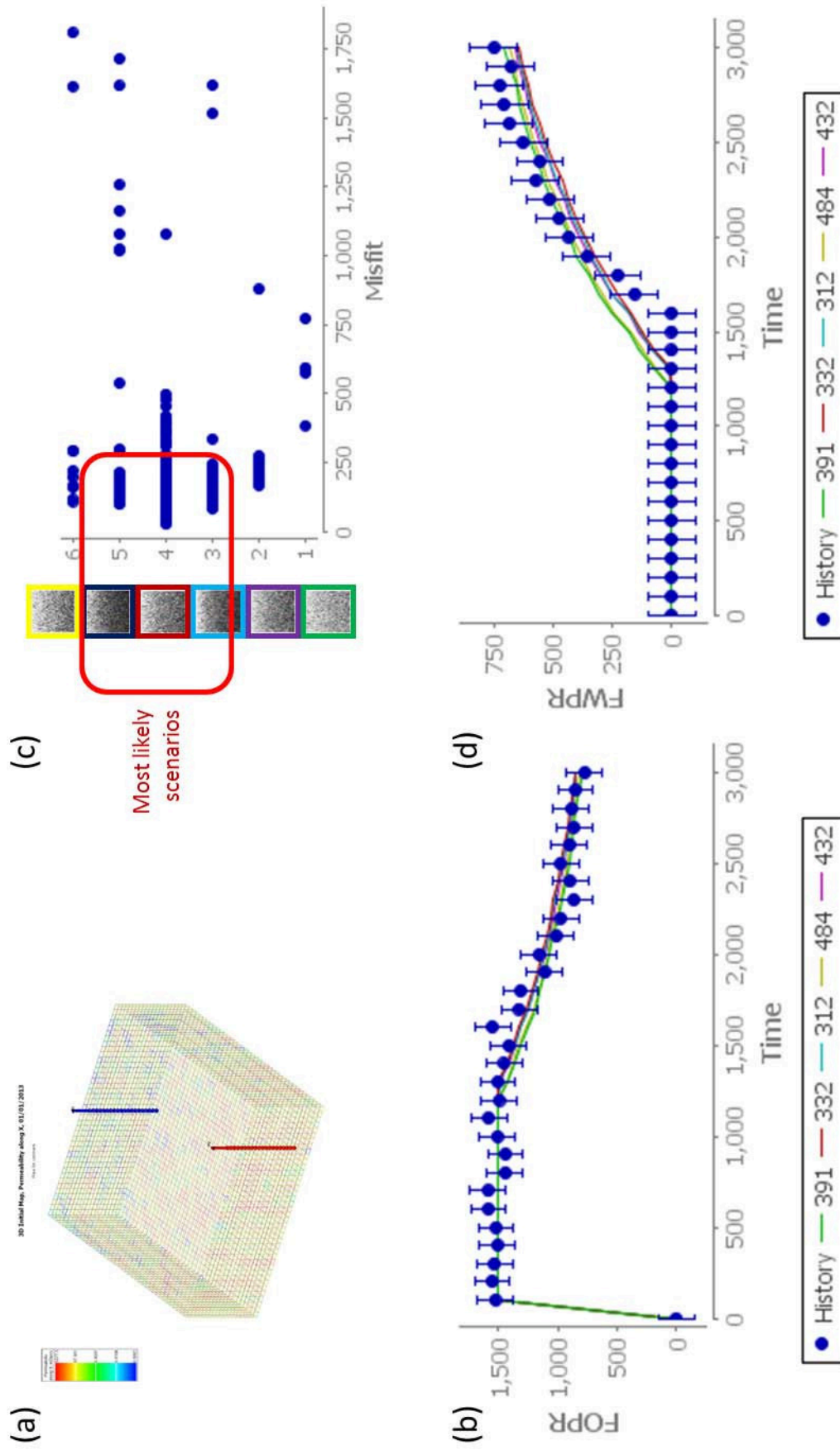
Total oil production: Average

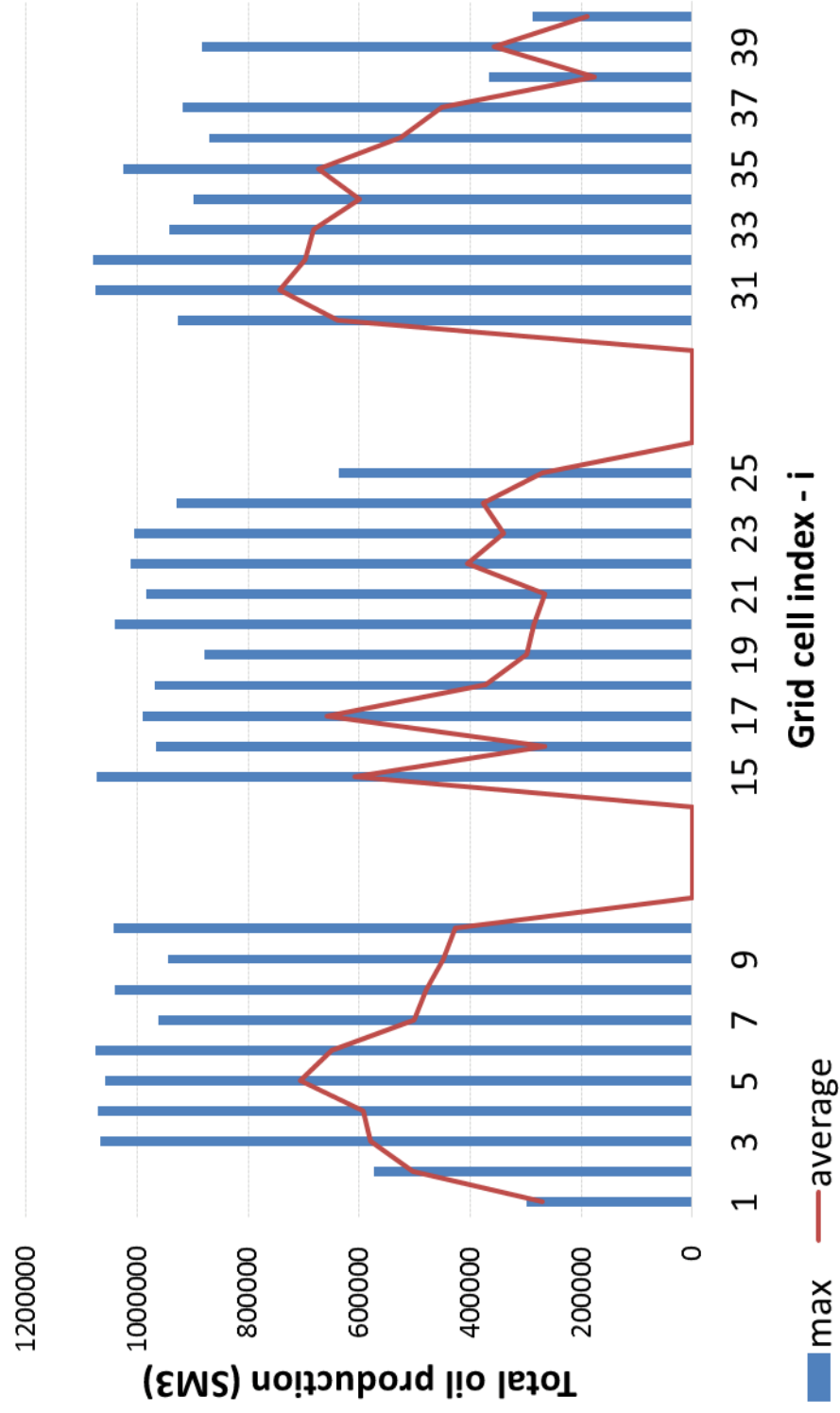


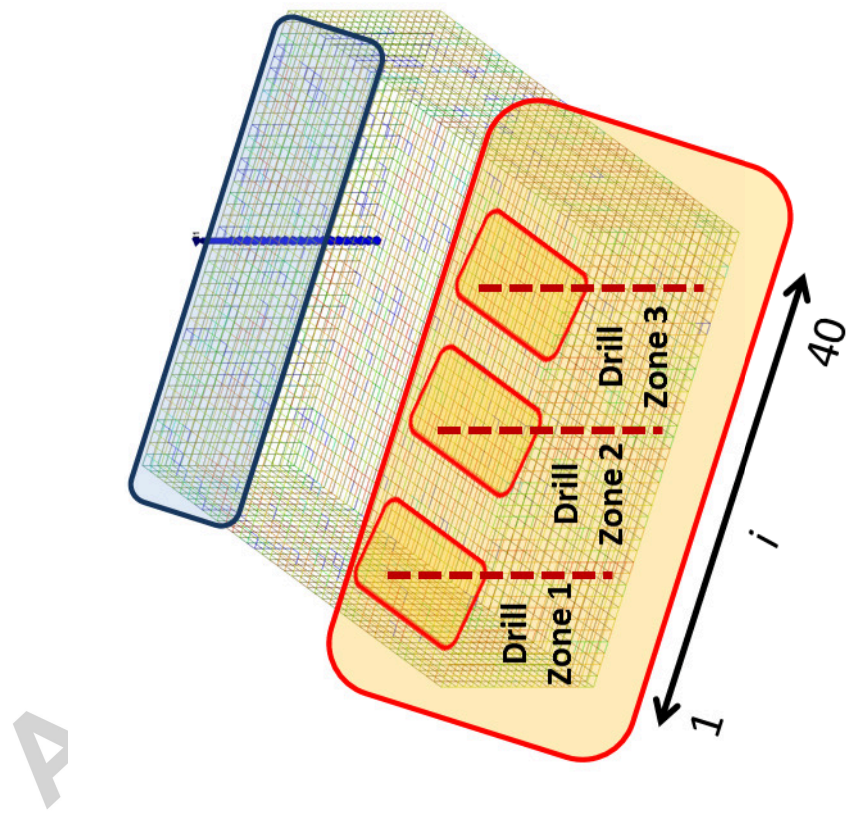
(d)

Total oil production: Maximum

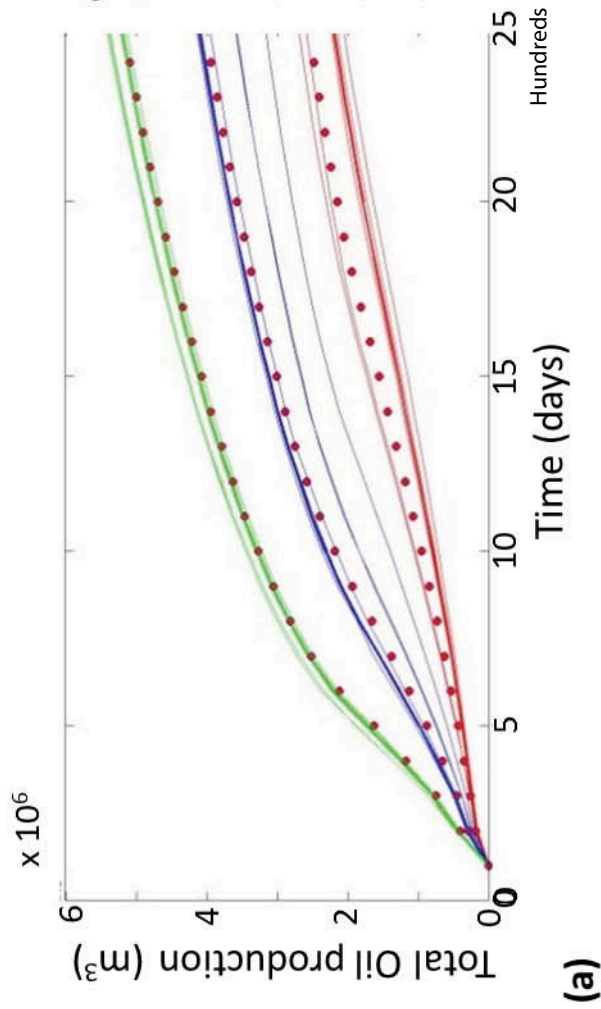




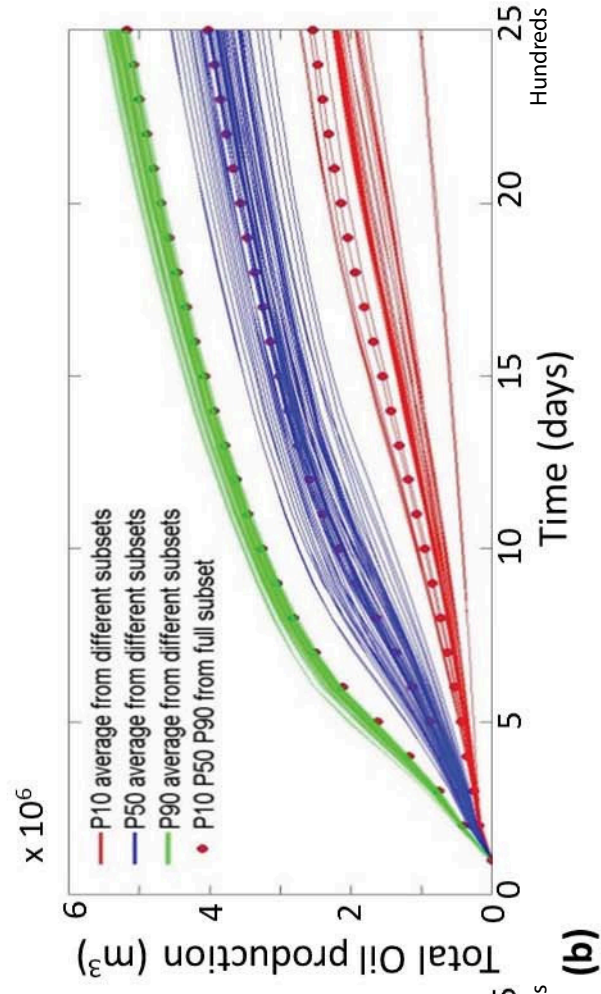




Ac

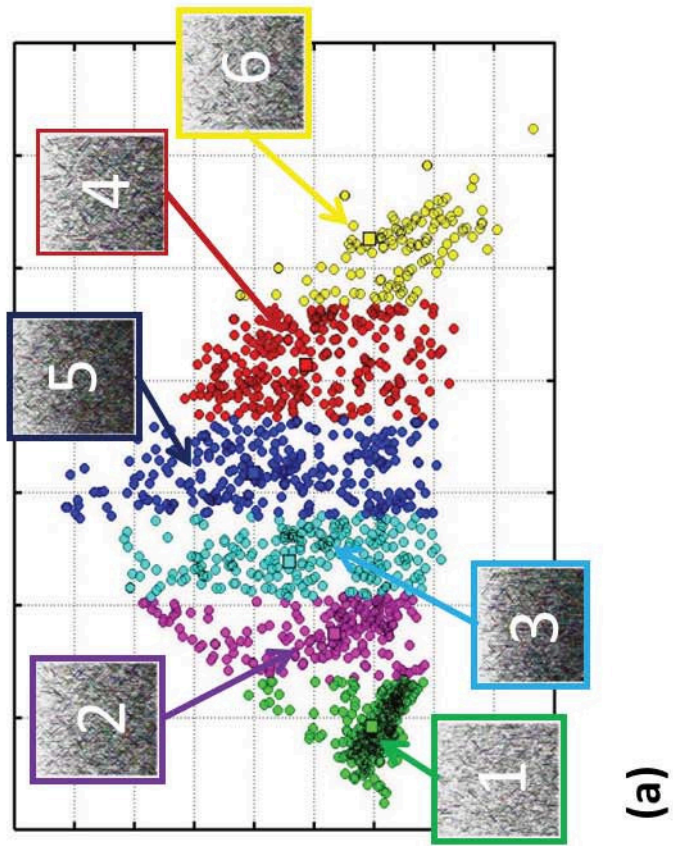
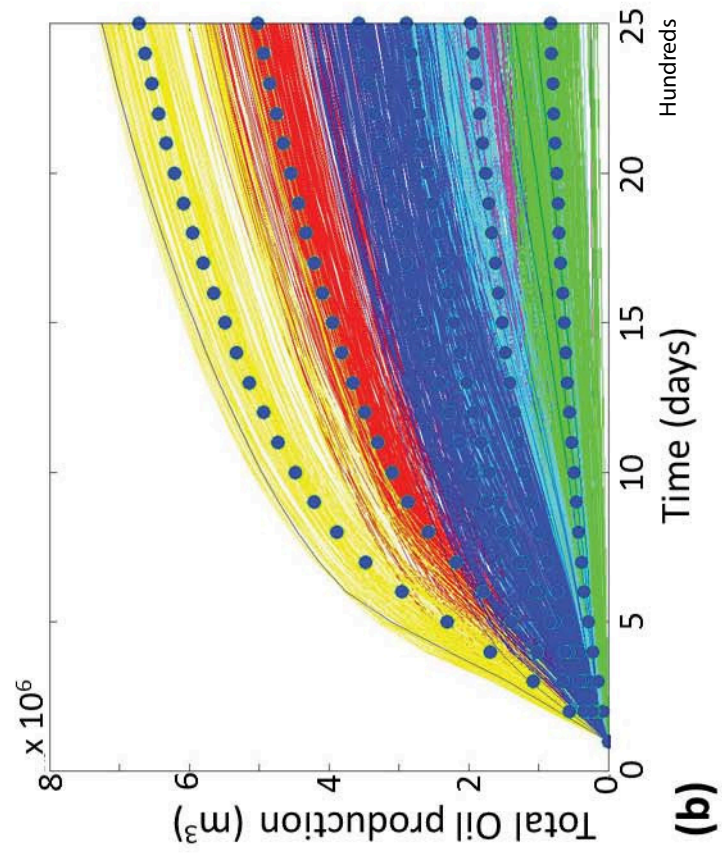


(a)



(b)

nt



(1)

Collect and analyse geological analogue data

e.g. in Case Study 2 this is achieved using Drones to take a high resolution aerial photographs of the outcrop, then extract the fractures from the background using image analysis tools (Bisdom et al, 2014)

(2)

Identify key fracture scenarios, elicit fracture properties

e.g. in Case study 2, fractures elicited from image analysis are used to build statistical models of the fractures across a set of nearby outcrops. General groupings are identified around the major set orientations

(3)

Build DFN models of fracture scenarios, conditioned to analogue data

e.g. in Case study 2, fractures properties and relationships for orientation, intensity (number of fractures per unit area), length, and relative proportion are elicited and used to create a range of scenarios of the fracture sets, observed in the outcrop

(4)

Upscale DFN (Oda Gold or equivalent) to gridded model

e.g. in Case study 2, this was done using MPS, using the upscaled DFN models as training images

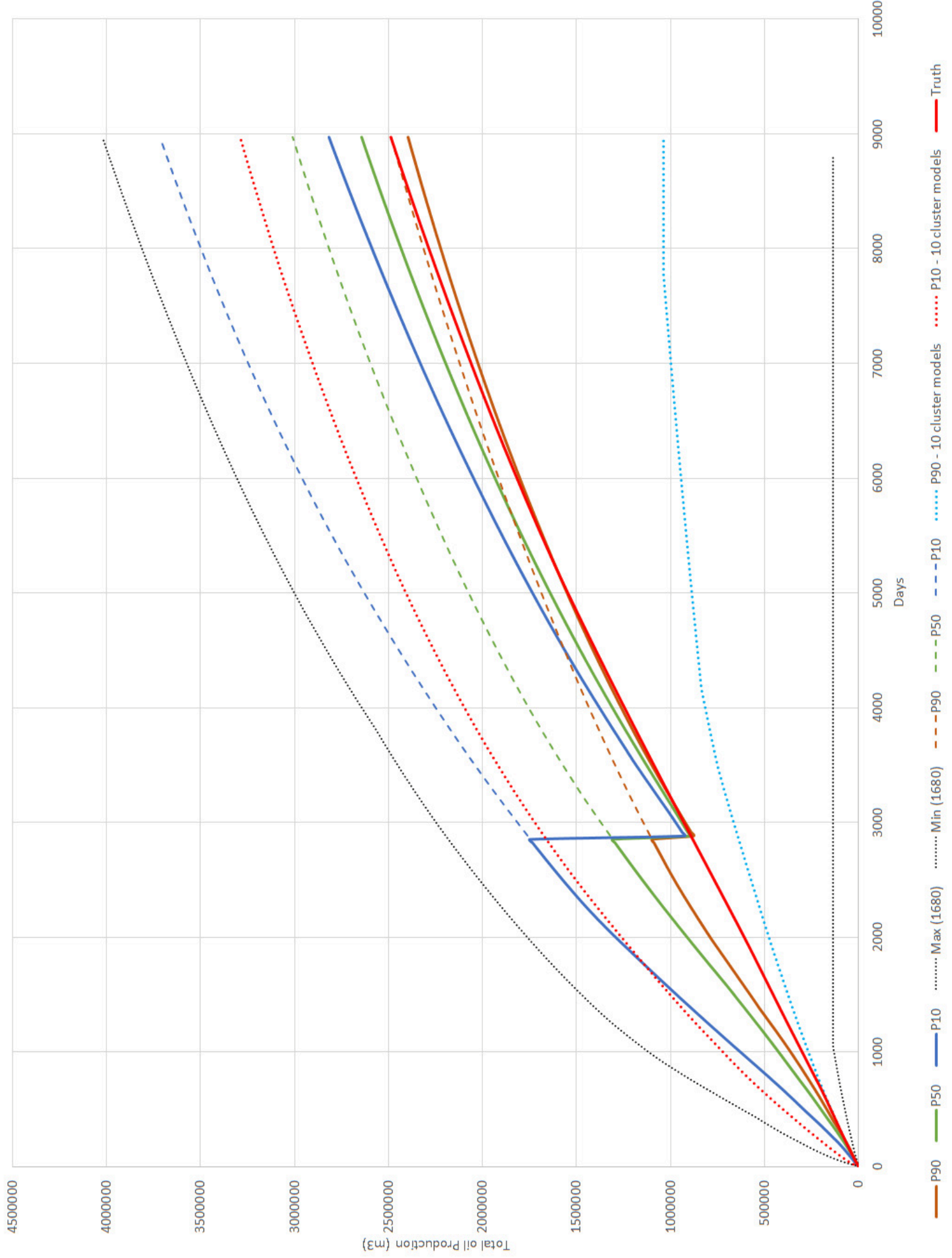
(5)

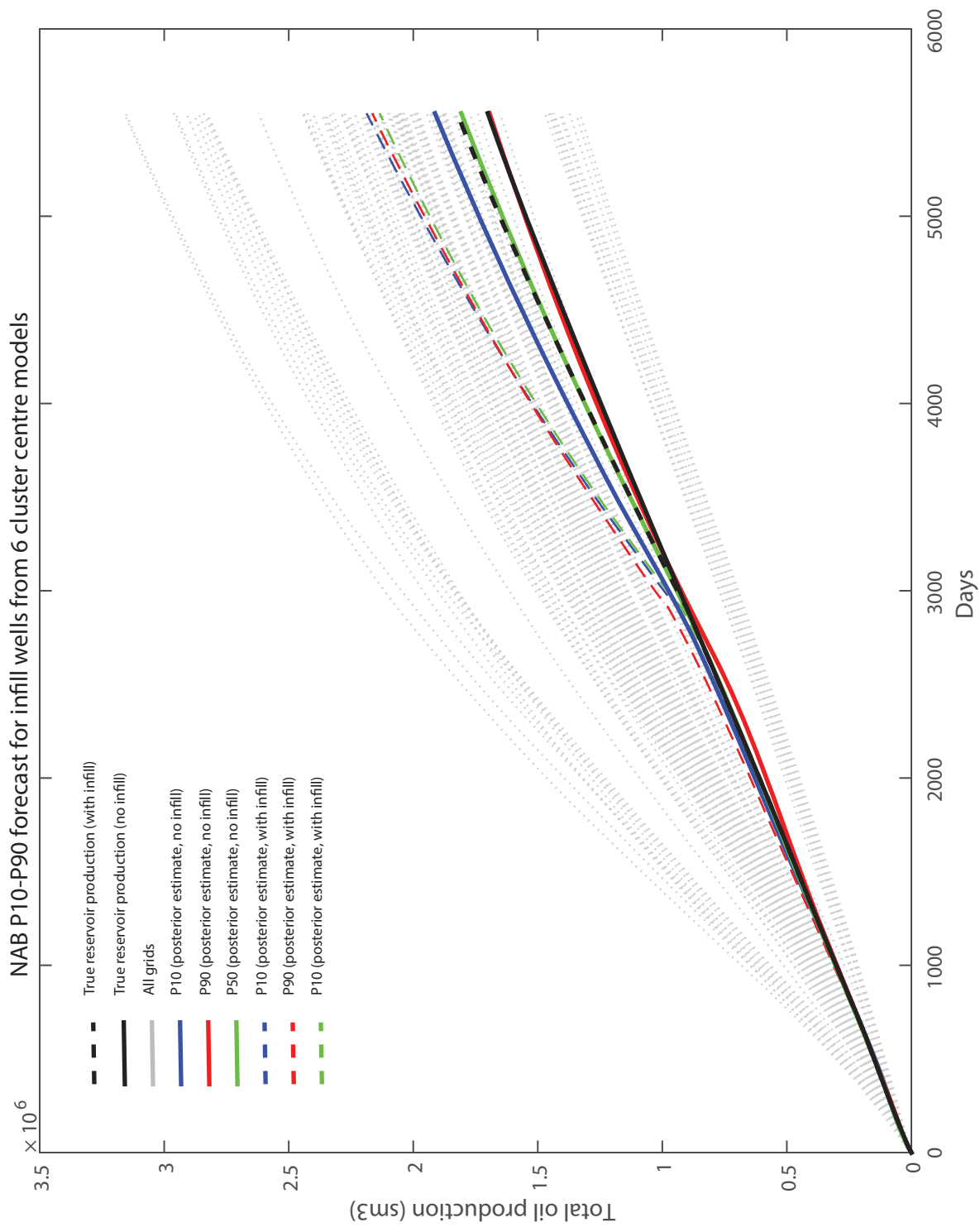
Generate multiple realisations of the gridded model for each scenario

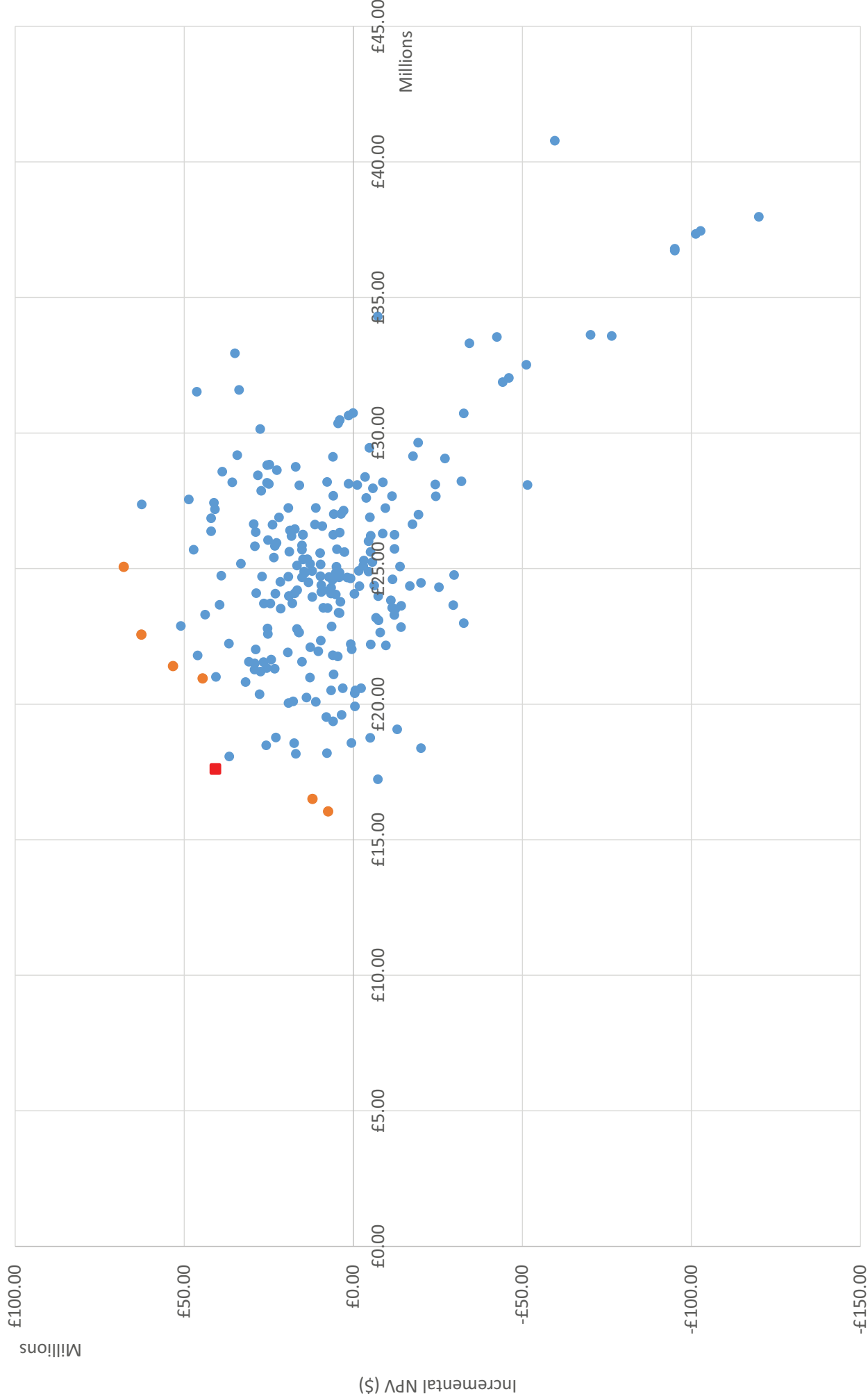
Find clusters of similar models based on distance metric of a model property

(6)

e.g. in Case study 2, this was done using the permeability field to identify models that have a similar distribution of fractures







Standard Deviation of Incremental NPV (\$)

- NPV
- NPV Pareto