공학석사 학위논문

Reliable Uncertainty Quantification of Channel Reservoirs Using Model Regeneration Scheme with a Deep Learning Algorithm

딥러닝 기반 모델 재생성 기법을 이용한 채널저류층의
불확실성 평가

2021 년 2 월

서울대학교 대학원

에너지시스템공학부

이 유 준

Reliable Uncertainty Quantification of Channel Reservoirs Using Model Regeneration Scheme with a Deep Learning Algorithm

딥러닝 기반 모델 재생성 기법을 이용한 채널저류층의
불확실성 평가

지도교수 최 종 근

이 논문을 공학석사 학위논문으로 제출함
2021 년 2 월

서울대학교 대학원
에너지시스템공학부
이 유 준

이유준의 공학석사 학위논문을 인준함
2021 년 2 월

위 원 장   박 형 동

부위원장   최 종 근

위   원   정 훈 영

Abstract

Production behaviors in channel reservoirs are highly affected by characteristics of channels. Hence, reservoir model calibration by matching its production data has often implemented. However, when this inversion method is applied to channel reservoirs, there are several critical issues, such as geological information loss and high dependency of results to initial models.

Therefore, in this research, a model regeneration scheme is introduced for reliable uncertainty quantification on channel reservoir models without a conventional model inversion method. It is composed of three parts: feature extraction, model selection, and model generation. In the feature extraction, drainage area localization and discrete cosine transform are utilized to extract channel features of near-wellbore area. In the model selection, K-means clustering and an ensemble ranking method are implemented to select models similar to a true reservoir. In the model generation, generative adversarial networks (GAN) and transfer learning are adopted to generate new models.

The proposed method is to, first, select good models from an initial ensemble by the feature extraction and comparing production data. New models are generated with the selected models by GAN. The feature extraction and model selection process are repeated to select final models from the selected and generated models. The final models are utilized to quantify its uncertainty by predicting future productions.

The proposed method is analyzed with 3 different 2D channel reservoir cases. The analysis shows that it allows to obtain reliable models for production forecasts with reduced uncertainty. This is by effectively characterizing connectivity of channels and permeability distribution of near-wellbore and by increasing a probability of existence of models similar to the true model.

Table of Contents

List of Tables

List of Figures

1. Introduction

Petroleum is a naturally occurring hydrocarbon mixture existing in a phase of gas, liquid, and semi-solid depending on temperature, pressure, and composition of hydrocarbons (Choe, 2017). Petroleum is typically contained in subsurface porous and permeable rock formations, called reservoirs. Characteristics of reservoirs have great impacts on production behaviors of petroleum fluids. Hence, characterizing a reservoir is crucial for petroleum field development.

Reservoir characterization is often implemented by reservoir modeling. It allows to obtain a reservoir model that reflects features of a reservoir of interest using geological information such as core sample data. When available geological information is limited, reservoir modeling becomes challenging. A model generated with limited data contains high uncertainty. Therefore, uncertainty quantification of a reservoir is necessary in order to obtain a reliable reservoir model under this circumstance. This can help estimating future production effectively, which leads to successful decision making.

However, having a reliable reservoir model is difficult to achieve in regard to very heterogeneous reservoirs such as channel reservoirs. Uncertainty significantly increases in channel reservoirs due to their unique characteristics. A channel reservoir consists of permeable sand formations in a channel shape and impermeable shale formations existing outside of channels in the reservoir.

Due to vivid separation of sand and shale formations, petroleum fluid mostly flows through channels. Hence, the connections and patterns of channels need to be properly characterized to understand production behaviors in

channel reservoirs. Under circumstances of limited geological information and high heterogeneity, it is desirable to perform uncertainty quantification using an ensemble of equi-probable model realizations for channel reservoirs (Choe, 2013).

Many researchers have performed channel reservoir characterization in a stochastic manner by utilizing an ensemble of model realizations to consider uncertainty (Jung and Choe, 2010; Kang and Choe, 2017; Lee et al., 2017; Jung et al., 2018). They conducted stochastic history matching, which is calibrating statistic properties of reservoir models, such as permeability, by matching their production data to observed data of a true model. This method is also known as model inversion. However, a possibility of geological information loss is an inevitable issue of history matching (Chang et al., 2016; Kim et al., 2016). When it is applied to channel reservoirs, history matching results in neglecting key channel characteristics such as connectivity and continuity (Jo et al., 2017).

To overcome this issue, many researchers have studied to effectively extract main characteristics of channel reservoirs. One approach is localizing drainage area (Jung and Choe, 2012; Yeo et al., 2014; Jung et al., 2017a). They conducted uncertainty quantification by utilizing production responses with corresponding effective zones. This scheme improved the prediction of future productions. Another approach is applying discrete cosine transform (DCT) to reservoir models. Jafarpour and McLaughlin (2007) first introduced DCT application for characterizing channel reservoirs. The ability of DCT capturing connectivity and continuity of channel reservoirs is verified in many studies (Kim et al., 2016; Jung et al., 2017b).

Another issue of the stochastic history matching is that its results are

highly affected by initial reservoir models. Therefore, numerous methods of selecting proper models from an initial ensemble have been studied (Lee et al., 2016; Jung et al., 2018; Lee et al., 2018). Kang et al. (2019) classified models into groups by K-means clustering and selected models similar to the true reservoir. They improved the history matching results by using the selected models as the initial models. Ranking reservoir models is another powerful method for organizing and selecting models. Jung et al. (2017a) ranked the models by calculating similarity of production data to the true response and selected the top ranked models to perform uncertainty quantification.

However, in case of channel reservoirs with high uncertainty, probability of proper models existing in an initial ensemble is very low. Under this circumstance, selecting models from the initial ensemble for the history matching does not improve its results. To increase the probability of existence of models similar to a true reservoir, Lee et al. (2017) suggested re-static modeling. They applied a distance-based clustering and selected a group of the models that have the most similar production responses to the true responses. The selected models were utilized to generate new reservoir models. With them, they performed a selection scheme to estimate uncertainty in production forecasts.

In addition, Kang et al. (2020) proposed a method of generating new reservoir models with deep convolutional generative adversarial networks (DCGAN). They selected models similar to the true model by applying principal component analysis (PCA) and K-means clustering. They trained DCGAN with the selected models as input and generated new models. They performed their selection scheme again to select final models for uncertain

quantification.

DCGAN was proposed by Radford et al. (2015), and it is now one of the successful network designs of generative adversarial networks (GAN). However, under limited data and resources, training DCGAN becomes very challenging due to its vast number of parameters (Wang et al., 2018). One profound technique that allows to overcome this issue is transfer learning (Yoshinski et al., 2014; Wang et al., 2019; Noguchi and Harada, 2019). It is to transfer knowledge obtained from a pre-trained network to a new network to be trained with a different dataset.

This study introduces a model regeneration scheme for reliable uncertainty quantification in channel reservoirs without a conventional model inversion method. This scheme consists of three parts: feature extraction, model selection, and model generation. In the feature extraction part, drainage area localization and DCT are applied to reservoir models. In the model selection part, the models are classified by k-means clustering with the extracted feature information. They are ranked based on the classification results by an ensemble ranking method. In the model generation part, DCGAN is trained with selected top-ranked ensemble models as input data. To successfully train DCGAN, transfer learning is implemented.

Theoretical backgrounds of the techniques adopted for this research are explained in Chapter 2. In Chapter 3, the proposed method with the model regeneration scheme is described. Chapter 4 presents results of analyzing the proposed method with 3 channel reservoir cases. In the last chapter, Chapter 5, summarizations and conclusions of the study are discussed.

2. Theoretical backgrounds

2.1 Feature extraction

2.1.1 Drainage area localization

Drainage area localization is assigning near-wellbore data of every reservoir model realization to each production well. The main goal in reservoir characterization is to obtain models that have similar characteristics to a true reservoir model. In order to meet this goal without a model inversion method, proper model selection needs to be performed from an ensemble generated to cover geological uncertainty of a reservoir in interest. Model selection requires response data of reservoir models and a true observed data. Models that have similar response data to the true observed data are more probable to have similar reservoir characteristics.

During the selection process, utilizing field total production data is not desirable. Figure 2.1 demonstrates a possibility of two different reservoir models having the same field total production data because they are perfectly symmetrical. To prevent this issue, production data of each well need to be used for proper model selection. It is more effective to use near-wellbore features when comparing production data of each well due to their significant impacts on production data of a well. Therefore, assigning near-wellbore data of every reservoir model to individual well is necessary.

For the drainage area localization process, models are, first, divided into smaller sections as many as the number of wells. The sections are identical in

size for every model. The divided sections are then assigned to corresponding wells. This process can be shown in Figure 2.2 and Figure 2.3.
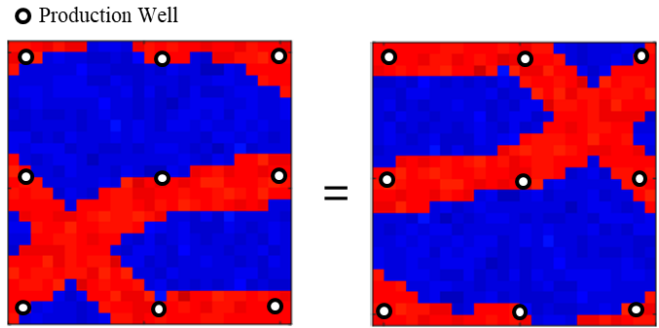

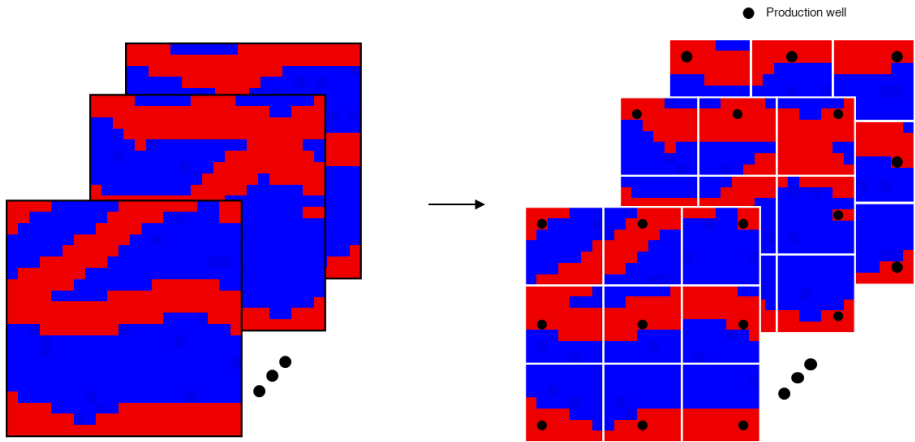Figure 2.1 Two different models that have the same field total production behaviors

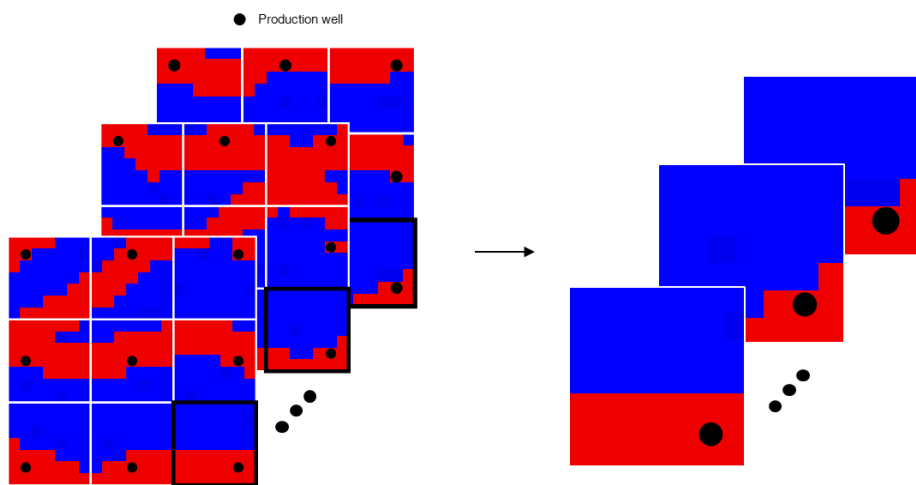Figure 2.2 An example of dividing models into smaller sections

Figure 2.3 An example of assigning near-wellbore data to a production well

2.1.2 Discrete cosine transform

DCT is transforming data in terms of a sum of cosine functions oscillating at different frequencies. It is a widely used transformation technique in signal processing and data compression. DCT for one dimension (1D) is defined by the following equations:

$$v(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos\left[\frac{\pi(2n+1)k}{2N}\right], 0 \le k \le N-1 \qquad (2.1)$$

$$\alpha(k) = \begin{cases} \sqrt{2/N}, & k = 0 \\ \sqrt{1/N}, & 1 \le k \le N-1 \end{cases} \qquad (2.2)$$

where, $u(n)$ is an input data sequence, and $v(k)$ is kth DCT coefficients. N is the size of data. In Equation (2.1), data are converted to coefficients of cosine functions.

DCT for 2D data can be performed by applying 1D DCT in two perpendicular directions. 2D DCT is often utilized for image compression. Because low frequency components of DCT contain general information of an image, main characteristics of an image can be gained by only utilizing the components of low frequency. As shown in Figure 2.4, low frequency components exist on the upper-left region. A frequency of cosine functions decreases towards the up and left direction.

DCT is applicable for characterization of channel reservoirs due to distinct patterns of channels. With low frequency components of DCT, channel reservoir models can be reduced in dimensionality without losing information

of their main channel trends. In this study, they are utilized to project reservoir models and to differentiate them in a reduced dimension space. Figure 2.5 shows transforming near-wellbore data, obtained by drainage area localization, to DCT coefficients.
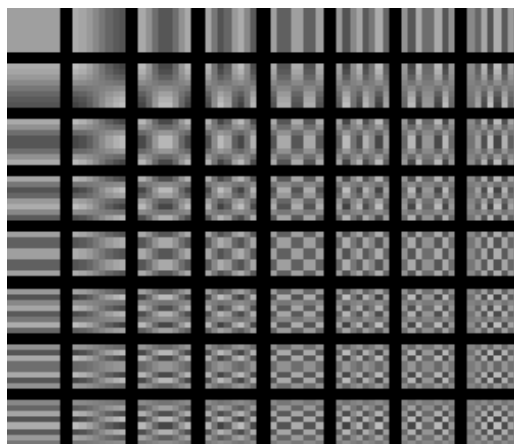
Figure 2.4 Discrete cosine transform based for 8 by 8 image representation
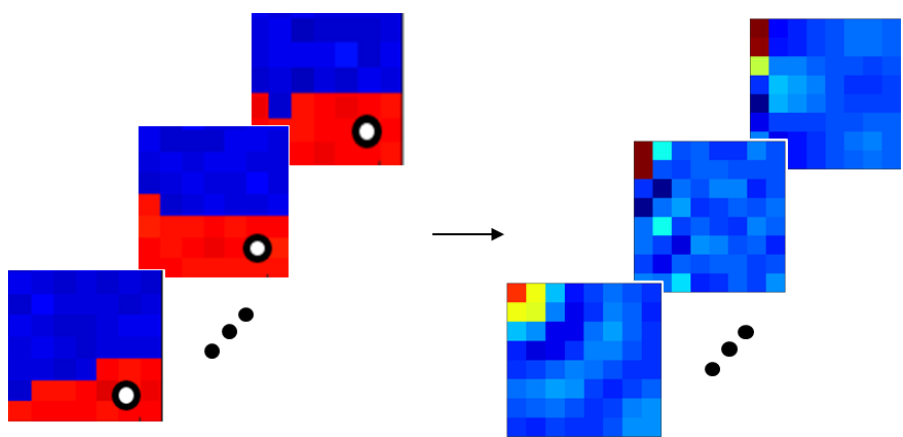
Figure 2.5 An example of applying DCT to assigned near-wellbore data

2.2 Model selection

2.2.1 K-means clustering

K-means clustering is one effective technique for classifying a dataset into groups. It is well-known for its simplicity and fast calculation. It groups a dataset into a pre-determined amount of clusters by minimizing the objective function J as Equation (2.3).

$$J = \sum_{k=1}^{K} \sum_{n=1}^{N} \left\| x_n^{(k)} - c_k \right\|^2 \tag{2.3}$$

where, $x_n^{(k)}$ is data in cluster k, and c_k is a centroid of cluster k. When grouping N data points into K clusters, k-means clustering finds a location of centroids, which has minimum distance between centroids and their data.

The process of k-means clustering is as follows. First, centroids are assigned at random locations in data space. These locations are assumed to be a center of clusters, and data close to these centers are included into corresponding clusters. Then, with the classified data, a new center of the clusters is calculated. The data are classified again depending on their locations relative to the new centers. These steps are repeated until the objective function converges. Figure 2.6 shows an example of 400 data points classified into 10 groups.

In this study, reservoir models are classified into groups by using DCT coefficients of near-wellbore data for every section assigned to a production well. During this classification, Silhouette method is implemented. It is to

optimize the number of clusters by calculating silhouette values which are defined as Equation (2.4).

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

(2.4)

where, $a(i)$ represents data cohesion in a cluster, and $b(i)$ represents separation between clusters. For a data point, i, $a(i)$ is calculated by averaging distance between i and other data points within a cluster. $b(i)$ is calculated by averaging distance between i and data points in a neighbor cluster closest to i.

(a) Dataset of 400 points   (b) Dataset classified into 10 clusters
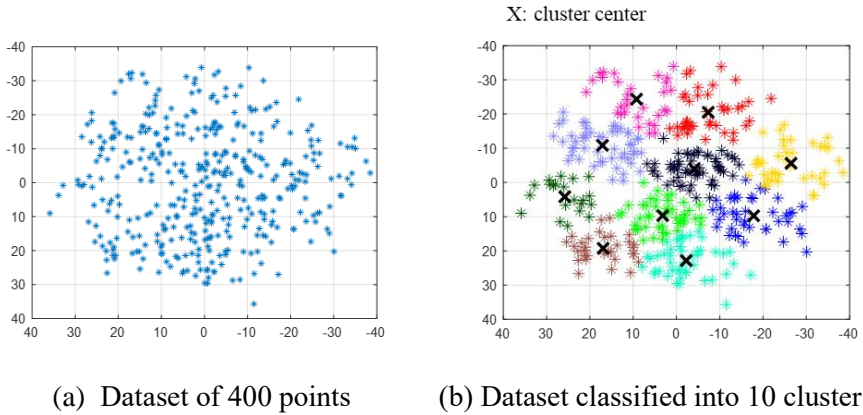
Figure 2.6 An example of k-means clustering with 10 clusters

2.2.2 Ensemble ranking method

Ensemble ranking method is a technique for organizing reservoir models in a ranking manner depending on their response data. This method is based on clustering results. After reservoir models are classified into clusters by k-means clustering, a reservoir model closest to a center of a cluster in a reduced dimension space is chosen for every cluster. These chosen models are representative models for the clusters. Production responses of the representative models are computed by a forward simulation. With the production responses, root mean square error (RMSE) are calculated with true observed data by using Equation (2.5).

$$RMSE = \sqrt{\frac{1}{N_{ob}} \sum_{i=1}^{N_{ob}} (p_i - \widehat{p_i})} \qquad (2.5)$$

where, p is production data of a well from a representative model, and $\hat{p}$ is the true observed data of a well. N_{ob} is the number of observation data.

Using calculated production error values, a score is calculated for each cluster by using Equation (2.6).

$$score_k = \frac{\dfrac{1}{RMSE_{min}}}{\dfrac{1}{RMSE_k}} \qquad (2.6)$$

where, k is a cluster index. Figure 2.7 shows the scoring process described above for one production well. For the visualization purpose, reservoir models

are projected in a 3D space using 3 DCT coefficients.

This scoring procedure is performed for every well. Therefore, every reservoir model has as many scores as the number of production wells. Models are ranked by the sum of their scores in descending order. A certain amount of reservoir models from the highest scored are selected.
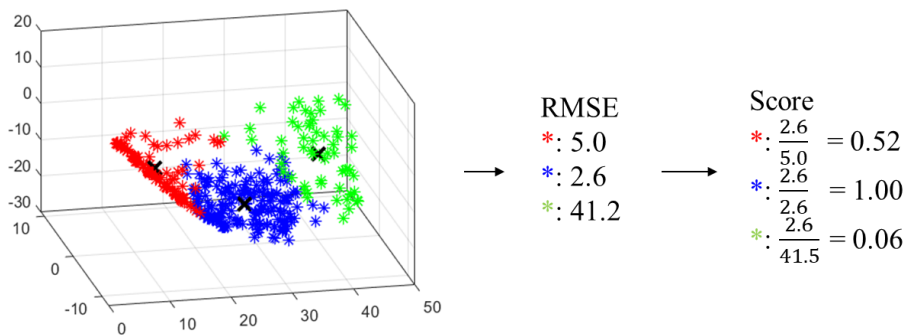
Figure 2.7 An example of calculating scores for reservoir models

2.3 Model generation

2.3.1 Generative adversarial network

GAN is a popular deep learning algorithm for data generation. It was proposed by Goodfellow et al. (2014). The main goal of GAN is to learn a distribution of data and generate new data that follows the distribution. Unlike other neural networks, GAN is composed of two networks: a generator and a discriminator. A typical structure of GAN is shown in Figure 2.8.

In GAN, a generator and a discriminator are trained in an adversarial manner for its own objective. The objective of a generator is to fool a discriminator by producing fake data that are indistinguishable from real data, while a discriminator tries to identify if data are fake or real. With a generator of well-trained GAN, it is possible to generate new data very similar to data used for training. The loss function of GAN is defined as Equation (2.7).

$$\min_{G} \max_{D} V(D, G) = E_x[\log D(x)] + E_z \left[\log \left(1 - D\big(G(z)\big)\right)\right] \quad (2.7)$$

where, x is real data used for training, and z is latent variables.

When a discriminator receives real data, it aims for D(x) to be 1. When the discriminator receives generated data by a generator, it aims for D(G(z)) to be 0, whereas a generator tries to make D(G(z)) to be 1. Therefore, the loss function of a discriminator is $\max_{D}[\log D(x) + \log \left(1 - D\big(G(z)\big)\right)]$, and the loss function of a generator is $\min_{G}[\log \left(1 - D\big(G(z)\big)\right)]$.

In this study, among variants of GAN, deep convolutional generative

adversarial network (DCGAN) is adopted to generate new reservoir models by using selected models as input. DCGAN was proposed by Radford et al. (2015), and it is known to be a successful GAN design for handling image data (Guo et al., 2019; Liu et al., 2019). The main difference of GAN and DCGAN is the architecture. DCGAN replaces fully connected layers with convolutional layers, as shown in Figure 2.9. The structure of DCGAN used for this study is based on the guideline by Radford et al. (2015).
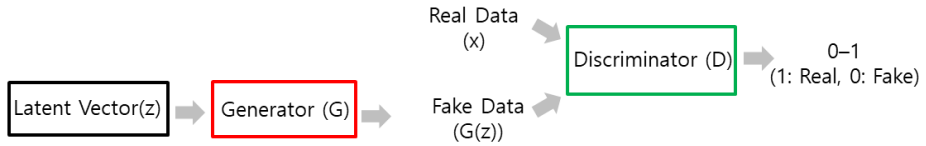
Figure 2.8 Structure of GAN
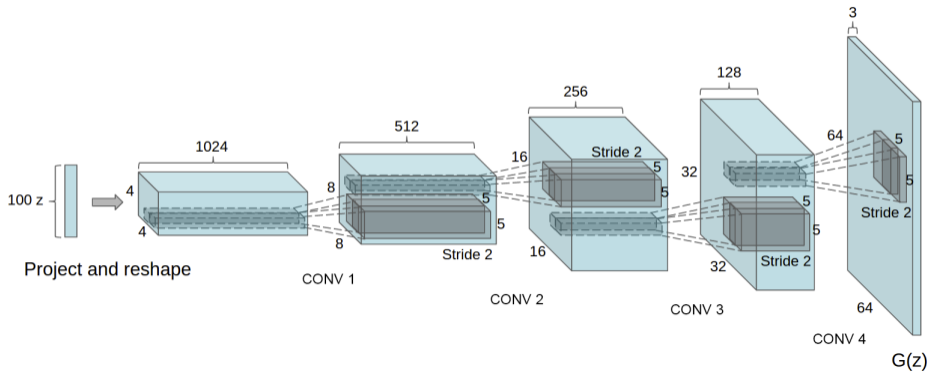
Figure 2.9 Structure of a generator in DCGAN (Radford et al., 2015)

2.3.2   Transfer learning

Transfer learning is a useful technique for deep learning to solve tasks efficiently. It is to train a new network with new data by utilizing knowledge from a network which has been pre-trained for a related task. Transfer learning is performed by transferring parameters from a pre-trained network. It uses the parameters as initial parameters for training with new data rather than starting from scratch. Despite of numerous ways of re-training in transfer learning, fine-tuning is the most intuitive and effective strategy (Wang et al., 2018; Mo et al., 2020). It allows all transferred parameters to be updated during training instead of freezing some of the parameters (Oquab et al., 2014).

In this study, fine-tuning is applied to enhance performance of training DCGAN with selected reservoir models. When training DCGAN with only selected models from scratch, it results in poor training due to its vast number of parameters and limited data. This issue is overcome by pre-training DCGAN with every reservoir model and updating it with selected models. With this fine-tuned DCGAN, new models that are similar to the selected ones can be stably generated.

3. Model regeneration scheme

The proposed method is to perform reliable uncertainty quantification of channel reservoirs without a model inversion method by a model regeneration scheme. The model regeneration scheme consists of 3 parts: feature extraction, model selection, and model regeneration. These parts are denoted as red, blue, and green respectively in Figure 3.1 which shows the procedure of the proposed method.

The first part of the model regeneration scheme is to extract features of near-wellbore areas of production wells in channel reservoir models. Drainage area localization is, first, performed to an ensemble of 400 initial reservoir models. The reservoir models are divided into smaller sections as many as the number of production wells. The sections represent near-wellbore areas. Permeability values of the models within each section are assigned to a corresponding production well.

DCT is then applied to compress assigned near-wellbore data without losing its features. Among coefficients obtained from DCT, 6 coefficients of the lowest frequency cosine functions are used to project the models in a reduced dimension space. This allows to differentiate the models by features of a near-wellbore area.

The second part of the model regeneration scheme is to select reservoir models similar to a true model by organizing them in ranking. In the reduced dimension space, the reservoir models are classified by k-means clustering. A model closest to a center of a cluster is decided to be a representative model for the cluster.

Representative models are then computed by a forward simulator, Eclipse from Schlumberger, and production responses of the models are gained. RMSE values are calculated with the production responses and true observed data. Based on the RMSE value of the representative models, a score for each cluster is calculated. This scoring process is implement for every production well, and the reservoir models are ranked by the sum of their scores in descending order. Top 50 reservoir models are selected, which are the most similar models to the true model.

The last part of the model regeneration scheme is to train DCGAN in order to generate new channel reservoir models. DCGAN is pre-trained with the 400 reservoir models. The parameters from the pre-trained DCGAN are then set to be initial parameters and fine-tuned with the selected 50 models. After the train, 50 new reservoir models are generated.

Details of the structure of DCGAN used in this study can be found in Table 3.1. Adam optimizer is adopted for DCGAN training, and learning parameters are set to 0.0002, 0.5 and 0.999 for learning rate, beta1, and beta2, respectively (Kingma and Ba, 2014). Also, training epoch is 300, and one epoch iterates 8 times of training. Due to faster convergence of a discriminator when training DCGAN with channel reservoir models (Kang and Choe, 2020), training ratio of the generator and the discriminator is set to be 8:1.

With this structure of DCGAN, pre-training with 400 models is performed. The same structure is utilized for transfer learning. However, when fine-tuning transferred parameters, it trains for only 30 epochs. Tensorflow 1.14 is implemented for DCGAN, and specification of the machine used for this study is as follows: Intel i5-4670 CPU 3.40 GHz, 16 RAM, and GTX 1660 6 GB.

In the proposed method, after new reservoir models are generated, the first and second part of the model regeneration scheme are repeated with the selected 50 initial models and generated new 50 models. Final 10 models are selected to predict future production responses with reduced uncertainty.
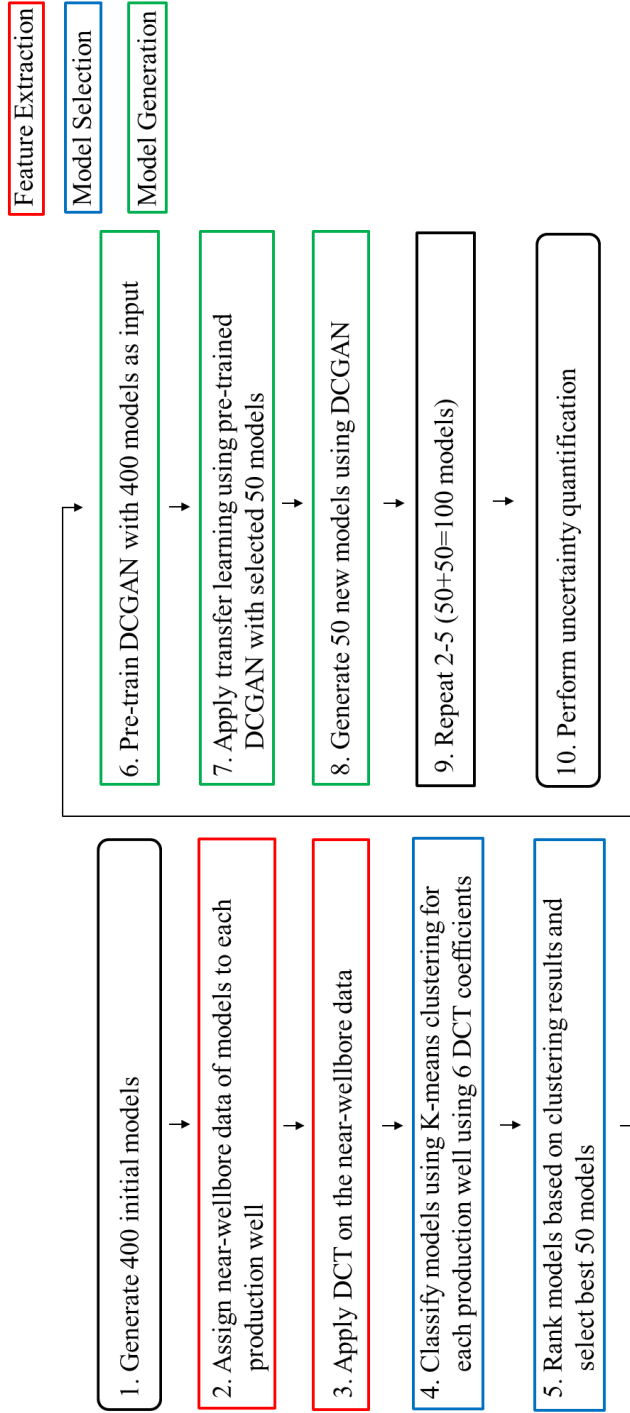
Figure 3.1 The flow chart of the proposed method

Table 3.1 Structure of DCGAN for this study

Generator		Discriminator	
Layer	**Remarks**	**Layer**	**Remarks**
Input layer	(1, 1, 100)	Input layer	(64, 64, 1)
1st hidden layer	(4, 4, 1024) kernel size: 4 strides: 1	1st hidden layer	(32, 32, 128) kernel size: 4 strides: 2
2nd hidden layer	(8, 8, 512) kernel size: 4 strides: 2	2nd hidden layer	(16, 16, 256) kernel size: 4 strides: 2
3rd hidden layer	(16, 16, 256) kernel size: 4 strides: 2	3rd hidden layer	(8, 8, 512) kernel size: 4 strides: 2
4th hidden layer	(32, 32, 128) kernel size: 4 strides: 2	4th hidden layer	(4, 4, 1024) kernel size: 4 strides: 2
Output layer	(64, 64, 1) kernel size: 4 strides: 2	Output layer	(1, 1, 1) kernel size: 4 strides: 1

4. Uncertainty quantification of channel reservoirs

The proposed method is applied to 3 cases of a 2D synthetic channel field. To validate the proposed method, the 3 cases are different in size, channel trend, and design of waterflooding which is often conducted at an early stage of field development in channel fields to acknowledge connectivity of channels.

4.1 Case 1

401 channel reservoir models are generated by using SNESIM (single normal equation simulation) module in SGeMS. The size of the models is 21 by 21 by 1, and each grid is a cubic of 50ft. Figure 4.1 shows the training image (TI) and 9 core sample data used for generating the models. An average permeability of the models is 1000 md for sand and 1 md for shale. Other petrophysical parameters of the models can be found in Table 4.1.

One of the 401 reservoir models is assumed to be the reference model, in Figure 4.2. The other 400 models are assumed to be initial models of an ensemble. The average permeability distribution of the 400 models in Figure 4.2 implies that the models have different trends of channels.

There are 8 production wells and 1 injection well in a waterflooding pattern of inverted nine-spot. Figure 4.3 shows oil production rates and watercuts of the 400 initial models in the 8 producers for 1500 days. The production responses are very high in uncertainty. The average oil rates and watercuts in the majority of the wells are far different from the reference model.
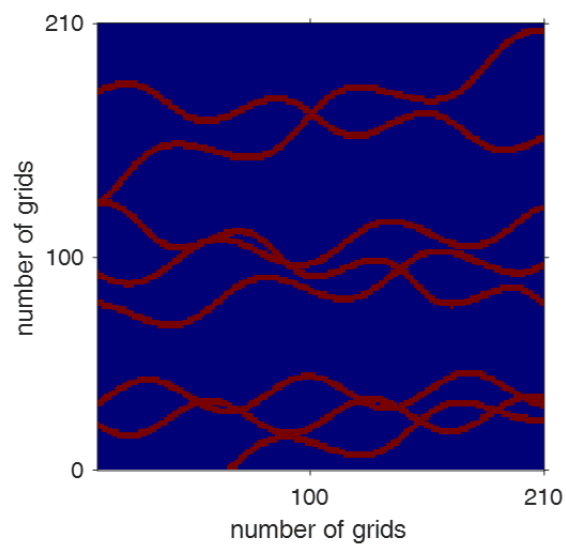
(a) Training image
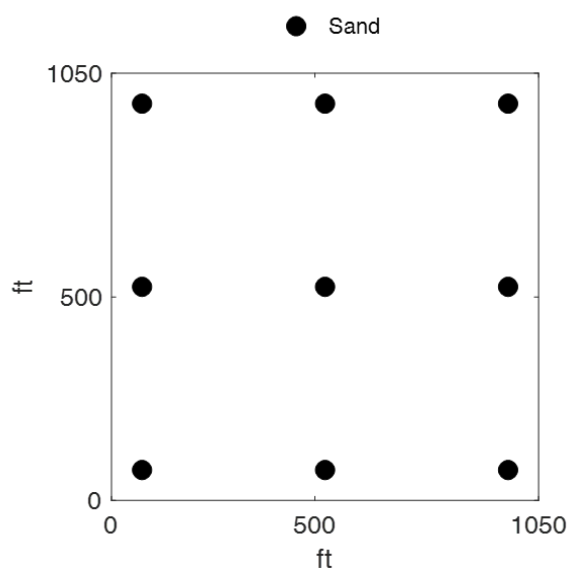
(b) Core sample data

Figure 4.1 Geological information for model generation for Case 1

Table 4.1 Petrophysical parameters for the simulation

Parameters		Values
Initial reservoir pressure, psia		2,000
Initial water saturation, fraction		0.25
Initial porosity, fraction		0.2
Rock compressibility, 1/psi		3.00E-05 at 2,000 psia
Formation volume factor, rb/STB	Oil	1.012 at 0 psia 1.011 at 1,000 psia 1.010 at 2,000 psia
	Water	1 at 2,000 psia
Fluid compressibility, 1/psi	Oil	1.00E-06
	Water	5.00E-07
Fluid viscosity, cp	Oil	3
	Water	1
Fluid density, lb/ft^3	Oil	48.62
	Water	62.31
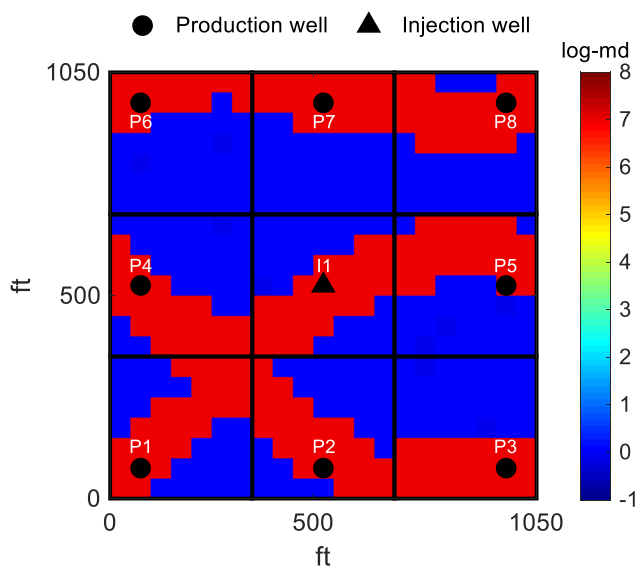
(a) Reference model
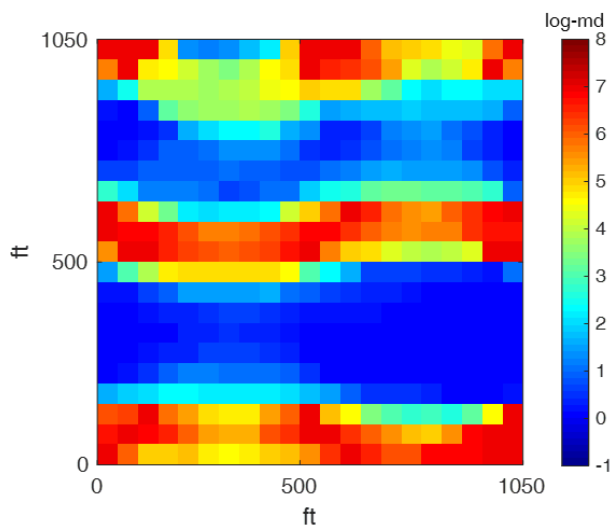
(b) 400 initial models

Figure 4.2 Reference model and the mean of permeability values for Case 1
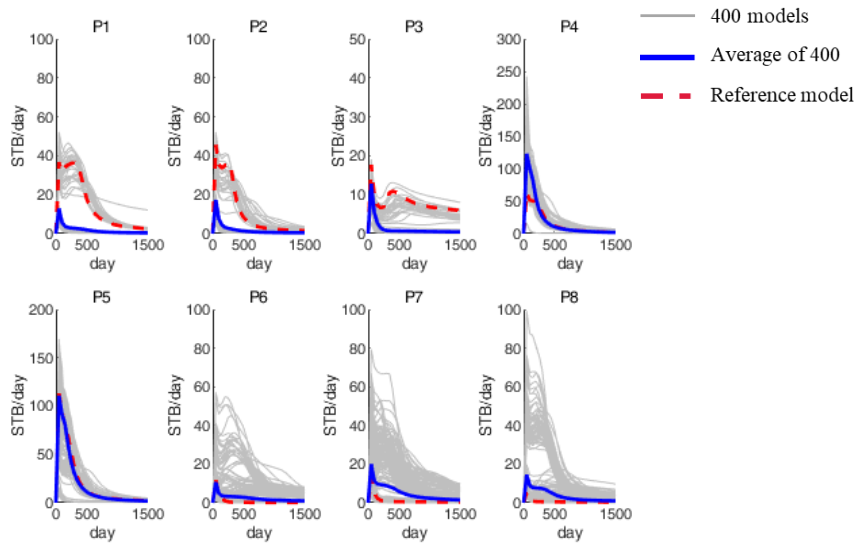
(a) Oil rates
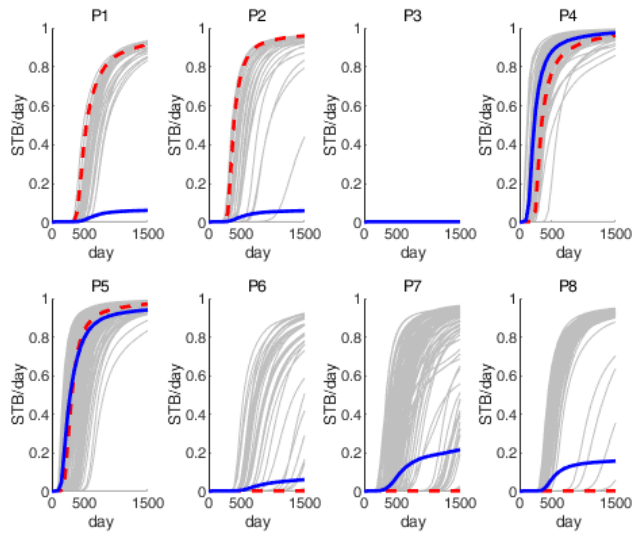
(b) Watercuts

Figure 4.3 Production responses of the 400 initial models for Case 1

The feature extraction and model selection of the model regeneration scheme are performed by utilizing production of 500 days. The bottomhole pressure limit for the producers and the injection rate for the injector are set to be 400 psia and 200 STB/day for production simulation. Figure 4.4 shows the result of ranking the 400 models.

By comparing the average permeability distribution of rank 1-100, 101-200, 201-300, and 301-400, it is certain that the top ranked reservoir models have very similar trends of channels to the reference model. The top 50 ranked models are selected and used for future prediction of production for 1000 days. In Figure 4.5, the 50 selected models predict future production similar to the production of the reference with reduced uncertainty. These 50 models are utilized as input data for DCGAN.

Figure 4.6 and 4.7 demonstrate necessity of transfer learning. When training DCGAN with the 50 selected models from scratch, the generator and discriminator do not converge. This leads to poor quality of output data, as shown in Figure 4.6. On the other hand, the both networks stably converge as training epoch increases with the 400 models. The generator from this training can output desirable data.

High score

Low score

(a) Reference

(b) Rank 1-100

(c) Rank 101-200

(d) Rank 201-300

(e) Rank 301-400
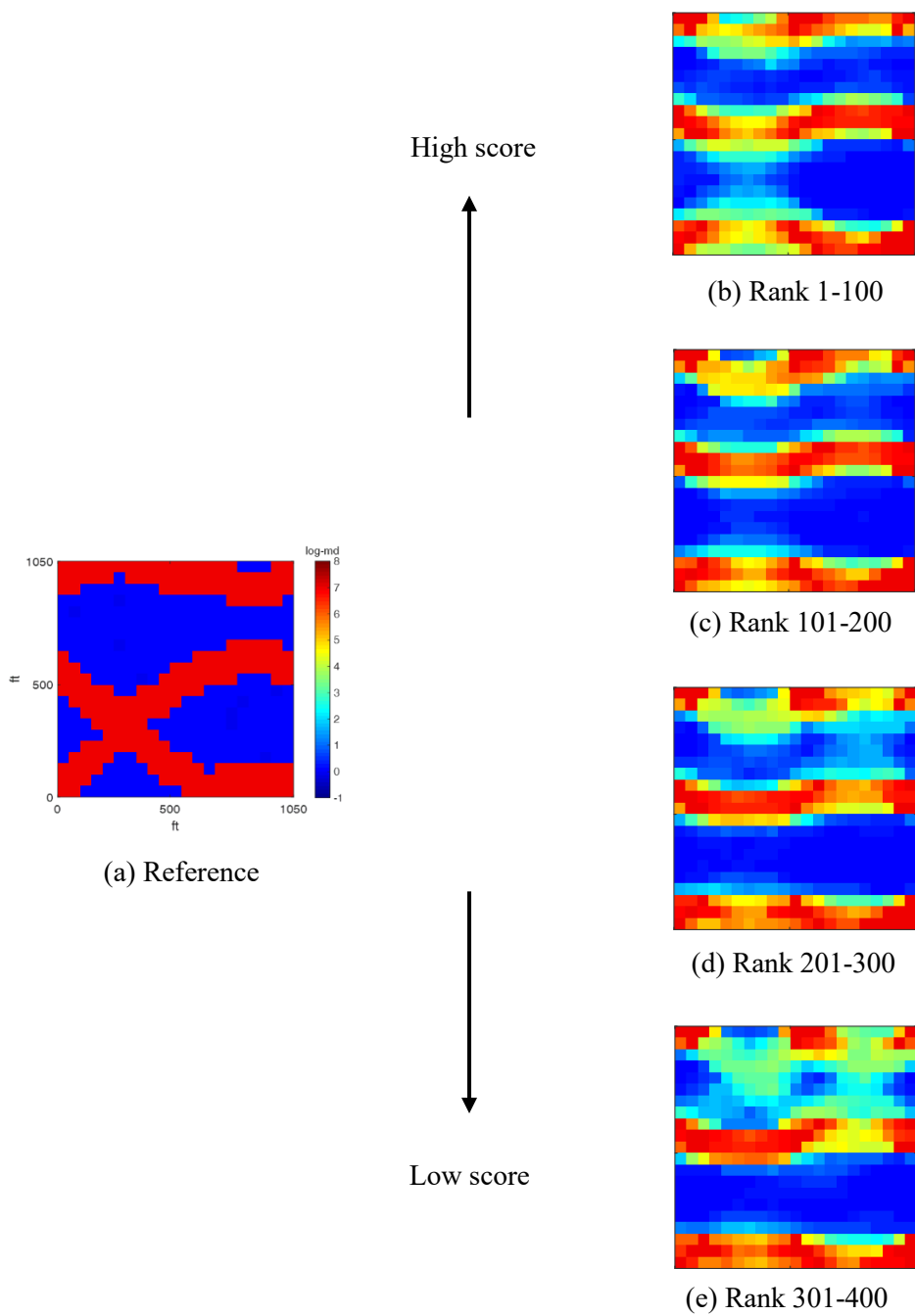
Figure 4.4 The result of ranking models for Case 1
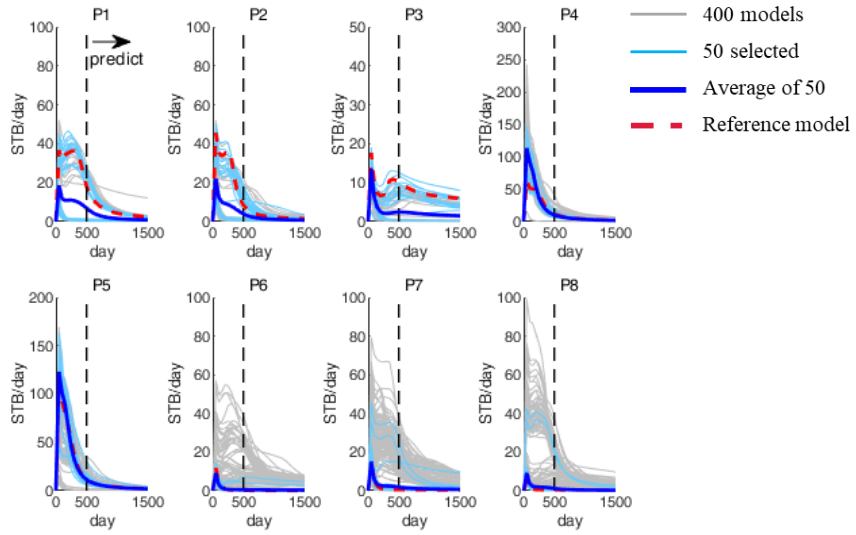
(a) Oil rates
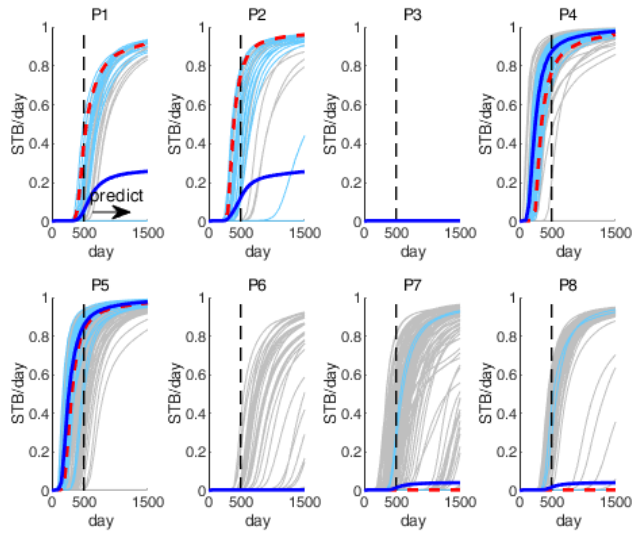
(b) Watercuts

Figure 4.5 Production responses of the top 50 models for Case 1
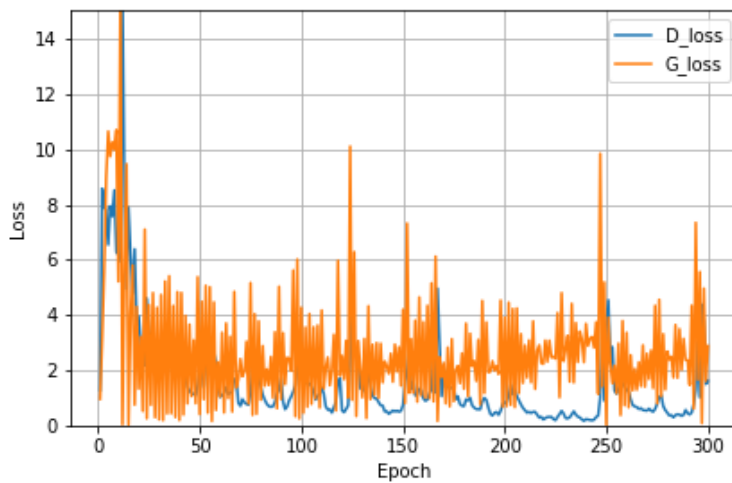
(a) Training result
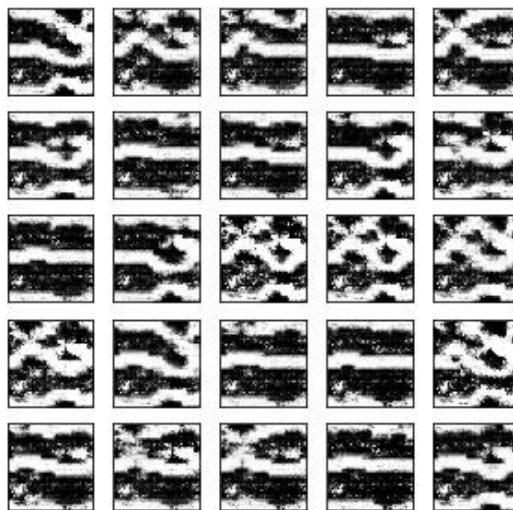
(b) 25 generated models

Figure 4.6 The result of DCGAN training with the 50 models

without transfer learning
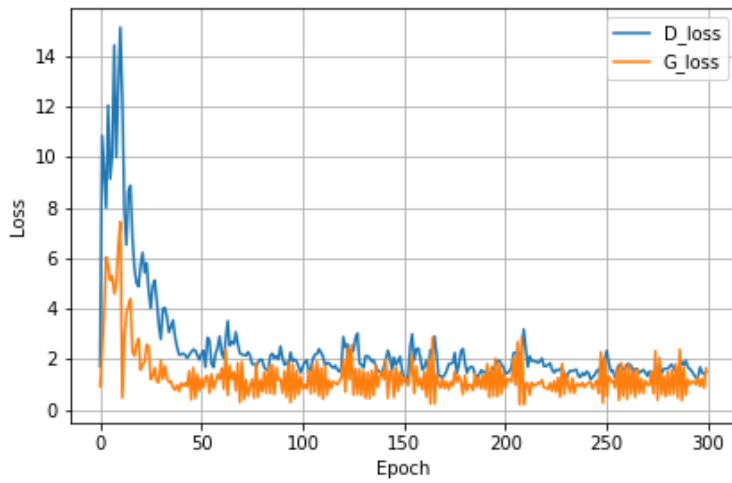
(a) Training result
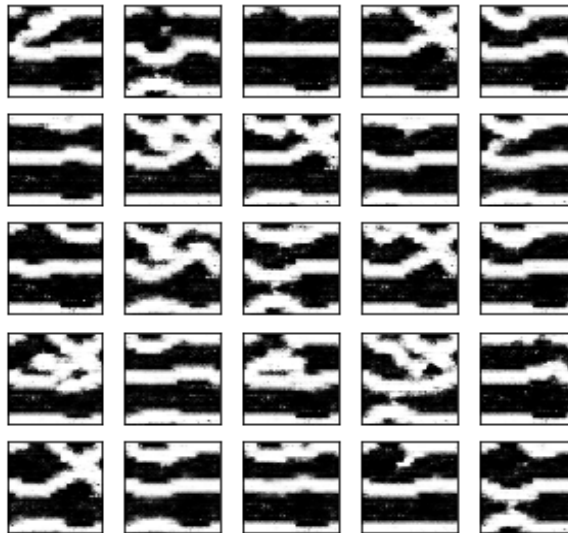
(b) 25 generated models

Figure 4.7 The result of DCGAN training with the 400 models

After training DCGAN with the 400 initial models, the pre-trained DCGAN is fine-tuned with the 50 selected models. The average of the 50 selected models and new 50 generated models can be found in Figure 4.8. These 100 models are ranked, and top 10 ranked models are selected by repeating the feature extraction and model selection method.

The result of the proposed method, Case 1a, is compared to 2 other methods: Case 1b and Case 1c. Case 1b selects top 10 ranked models from the 400 initial models as final models. Case 1c selects 10 models by using the information of the whole permeability fields.

In Case 1c, the permeability values of the models are reduced by principal component analysis. 2 principal components are used to project the models in a reduced dimension space. K-means clustering is then performed to classify them, and a representative model of clusters is computed by ECLIPSE. After calculating RMSE of the representative models, the cluster of the represent model that has the lowest RMSE is decided to be the cluster that the models similar to the reference model. The closest models from the center of the cluster are selected as final 10 models.

Figure 4.9 shows the average permeability distribution of the final 10 models for Case 1a, 1b, and 1c. Although the final models for the all 3 cases have similar channel patterns to the reference model, the result of Case 1a has the strongest connection of the main channels. As shown in Table 4.2, every selected model in Case 1a has the correct channel connection. However, in Case 1b, 5 of the models have the connection of the main channels, and only 1 model has it in Case 1c.
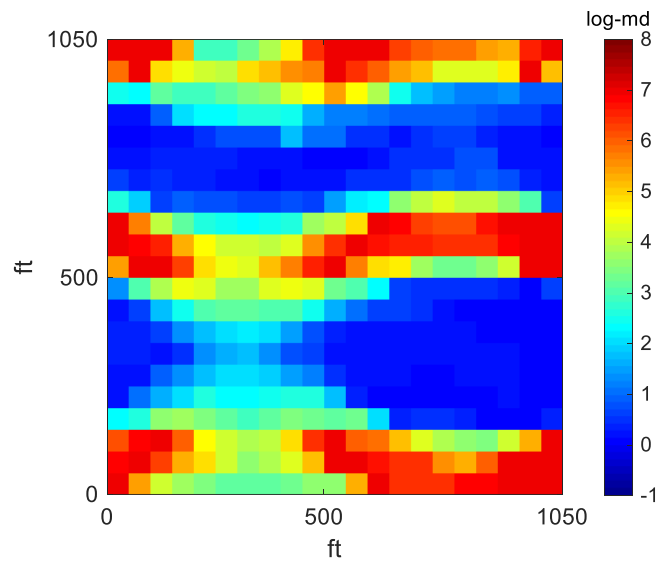
(a) 50 selected models
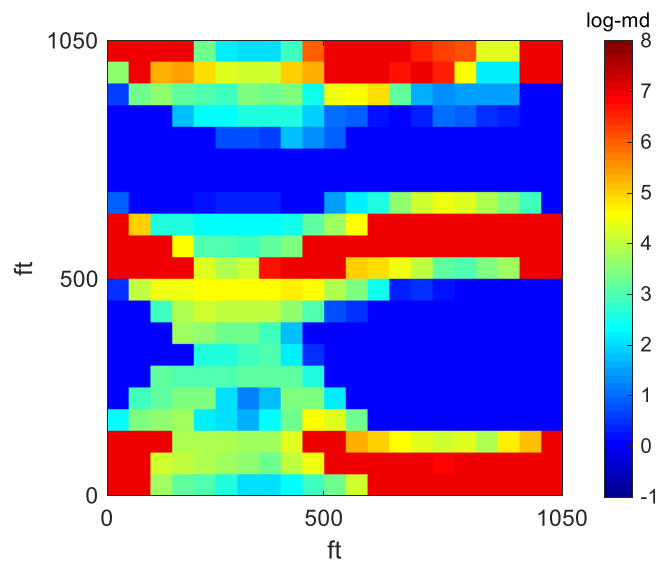
(b) 50 generated models

Figure 4.8 The mean of permeability values of the selected and generated models for Case 1

(a) reference model   (b) 400 initial models

(c) Case 1a   (d) Case 1b   (e) Case 1c
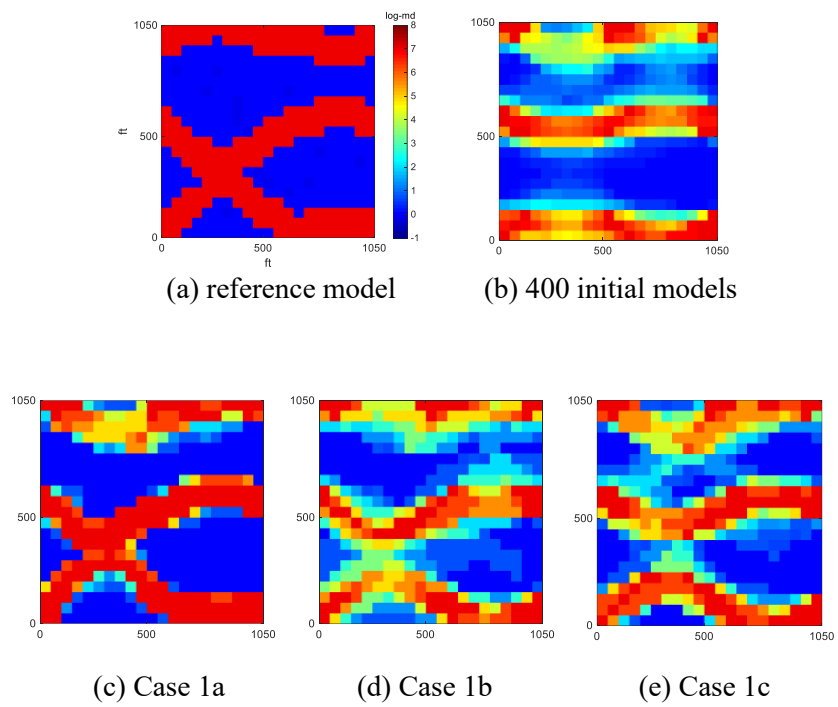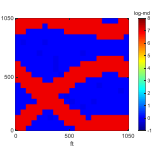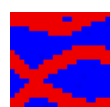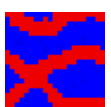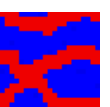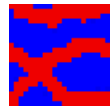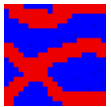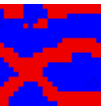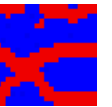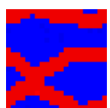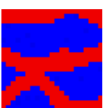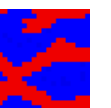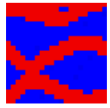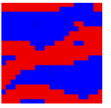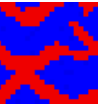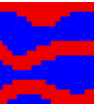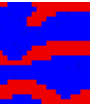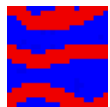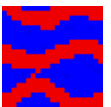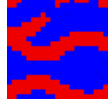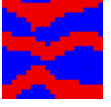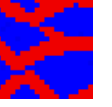
Figure 4.9 The mean of permeability values of the final 10 models for Case 1

Table 4.2 Permeability field of the final 10 models for Case 1

Reference	10 selected reservoir models				
	Case 1a				
	Case 1b				
	Case 1c				

Future prediction of production for 1000 days with the final models are analyzed for Case 1a, 1b, and 1c in Figure 4.10, 4.11, and 4.12. Figure 4.10 shows that Case 1a results in reliable forecasts with a reduced uncertainty range. In Figure 4.11, an uncertain range of the final models in Case 1b covers the production of the reference model. However, the difference between the reference production curves and the average production curves of the models shows large uncertainty, especially in P1 and P2. Figure 4.12 shows further increase in the uncertainty with the final modes in Case 1c.
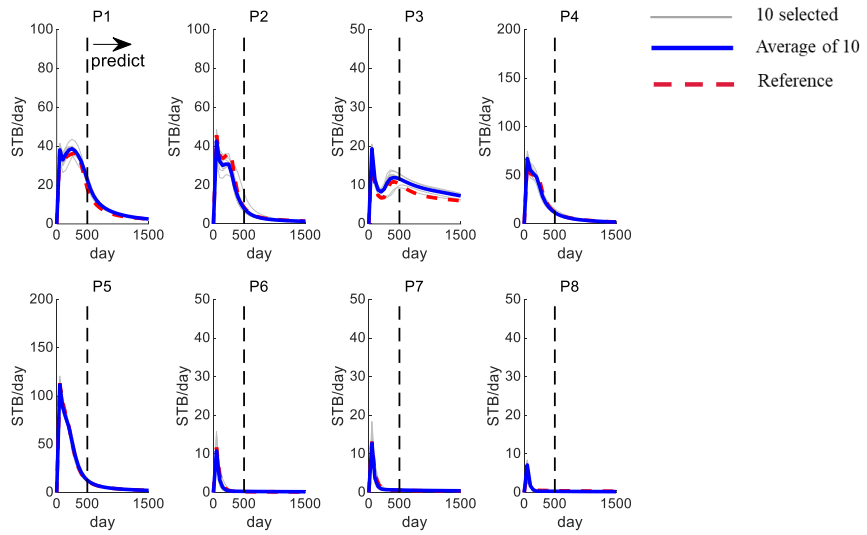
(a) Oil rates
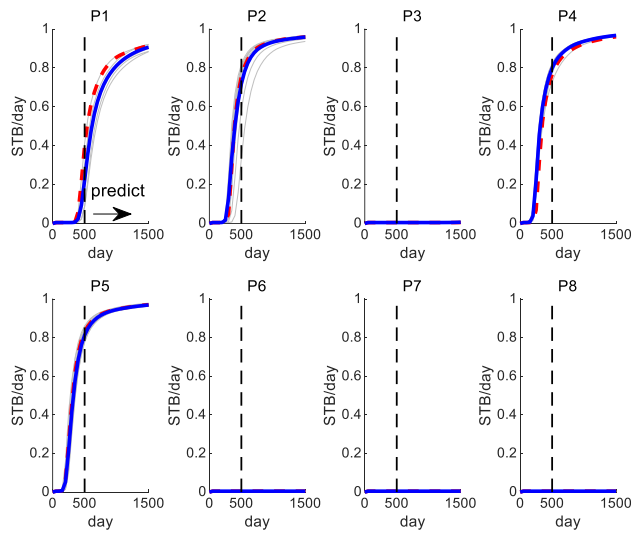
(b) Watercuts

Figure 4.10 Production responses of the final 10 models in Case 1a
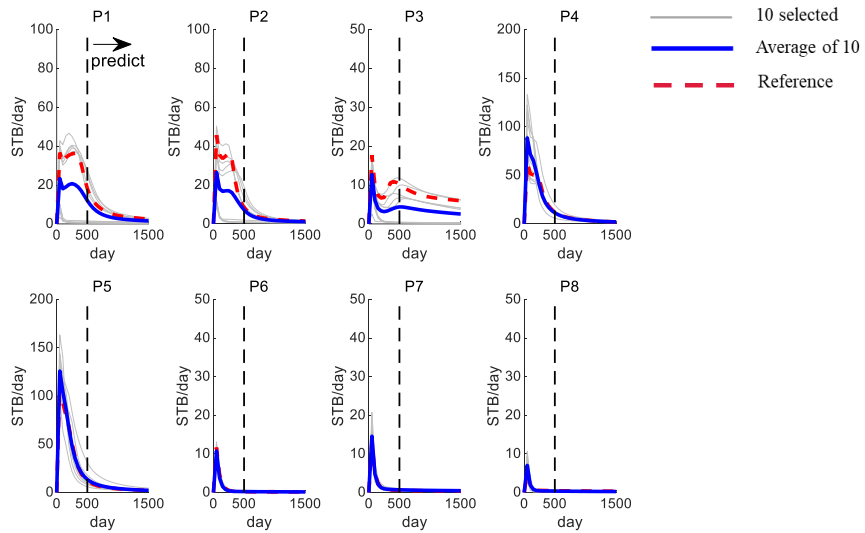
(a) Oil rates
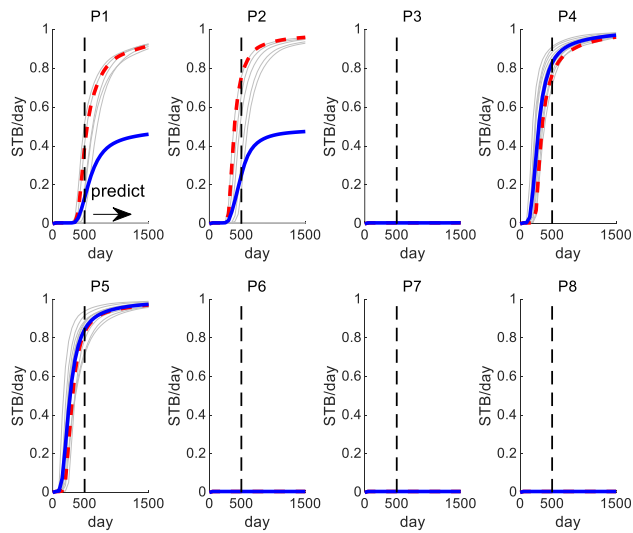
(b) Watercuts

Figure 4.11 Production responses of the final 10 models in Case 1b
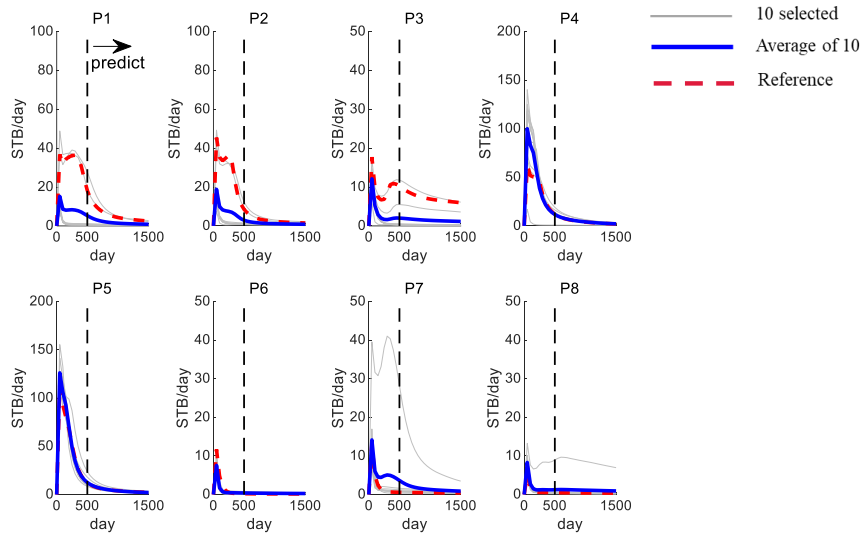
(a) Oil rates
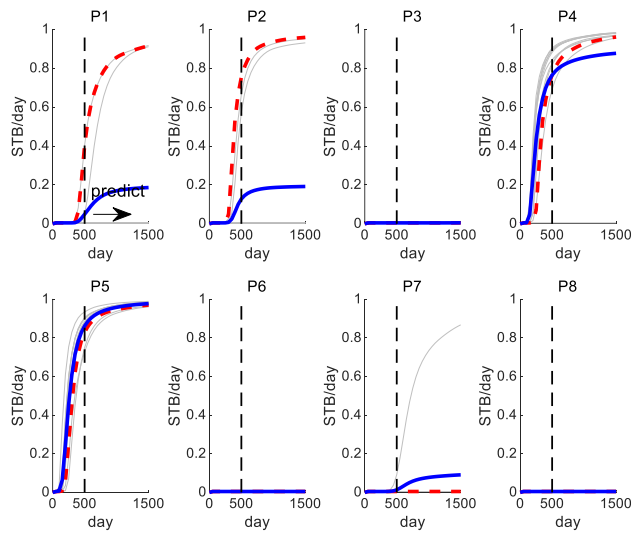
(b) Watercuts

Figure 4.12 Production responses of the final 10 models in Case 1c

4.2   Case 2

The proposed method is tested with another ensemble of channel reservoir models that have the size of 36 by 36 by 1. 401 models are generated by using SNESIM and geological information shown in Figure 4.13. The other petrophysical parameters of the models are the same as Case 1. Figure 4.14 shows the reference model and the average permeability distribution of the 400 models. There are 8 production wells and 8 injection wells in a waterflooding pattern of staggered line drive.

The feature extraction and model selection procedure are implemented by utilizing production of 500 days. The bottomhole pressure limit for production and the water injection rate for are set to be 400 psia and 200 STB/day for reservoir simulation. Figure 4.15 shows the result of ranking the 400 models. By comparing the average permeability distribution of the ranked models, the connection of the channels gets weaker as the score of the models goes down.

After fine-tuning DCGAN with the 50 selected models by utilizing the parameters of the pre-trained DCGAN, 50 new reservoir models are generated. The average permeability values of the 50 selected models and 50 generated models are shown in Figure 4.16. The feature extraction and model selection are repeated with the 100 models to organize them in a ranking manner. Top 10 ranked models are then selected for production forecasts.
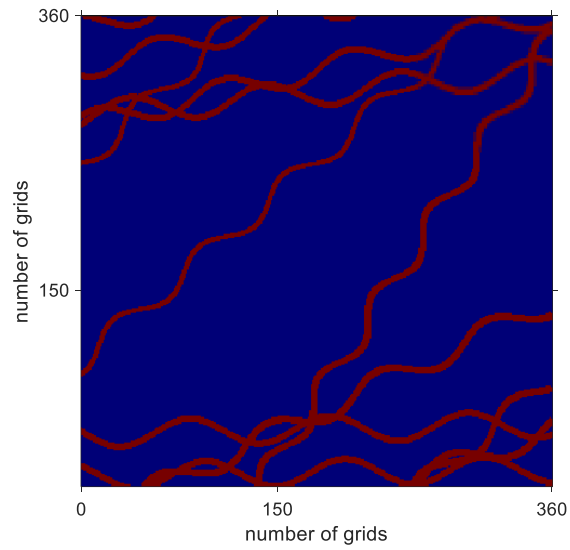
(a) Training image
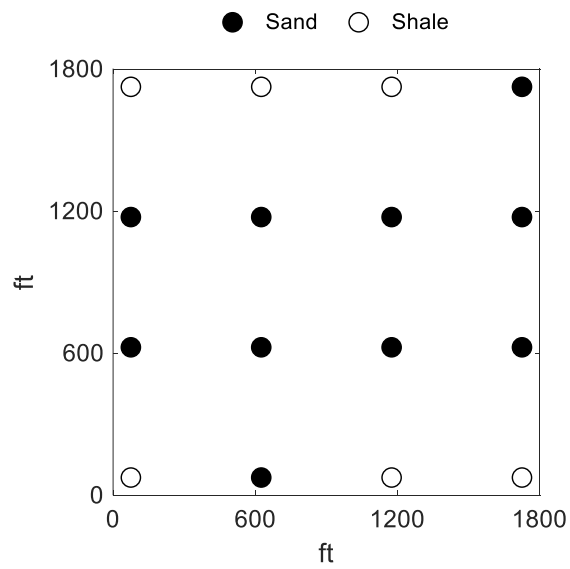
(b) Core sample data

Figure 4.13 Geological information for model generation for Case 2
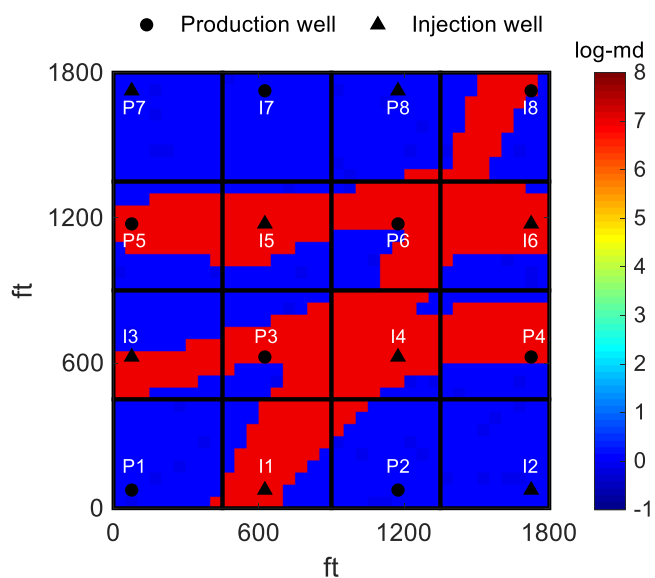
(a) Reference model
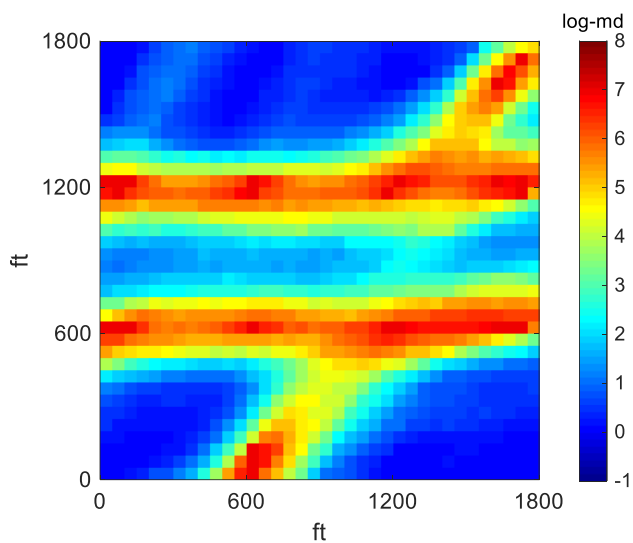
(b) 400 initial models

Figure 4.14 Reference model and the mean of permeability values for Case 2

High score

(b) Rank 1-100

(c) Rank 101-200

(a) Reference

(d) Rank 201-300

Low score

(e) Rank 301-400
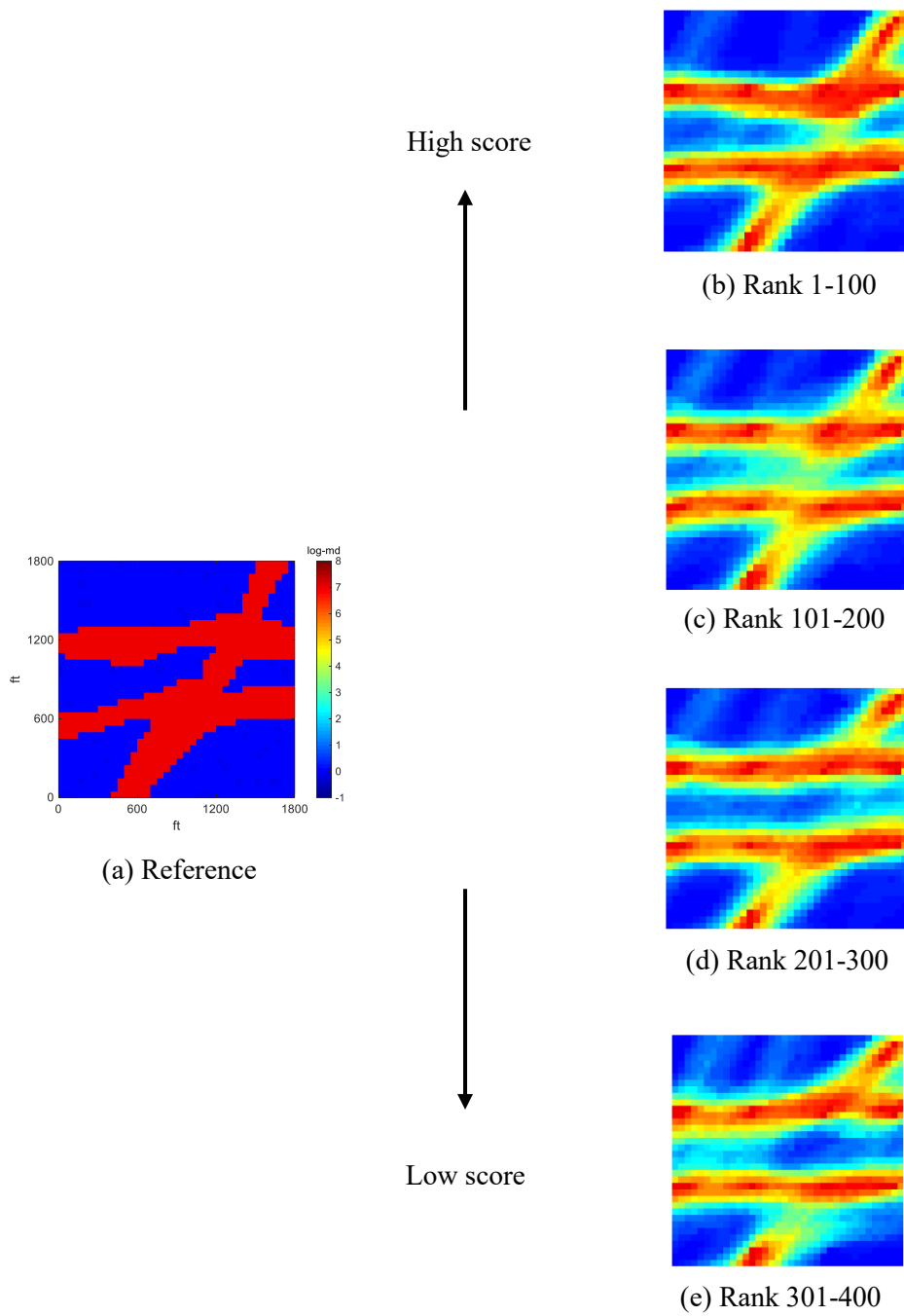
Figure 4.15 The result of ranking models for Case 2

(a) 50 selected models

(b) 50 generated models
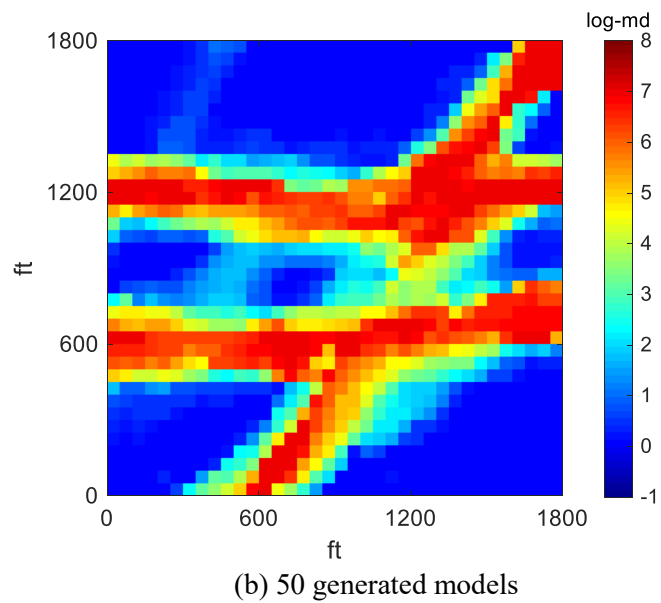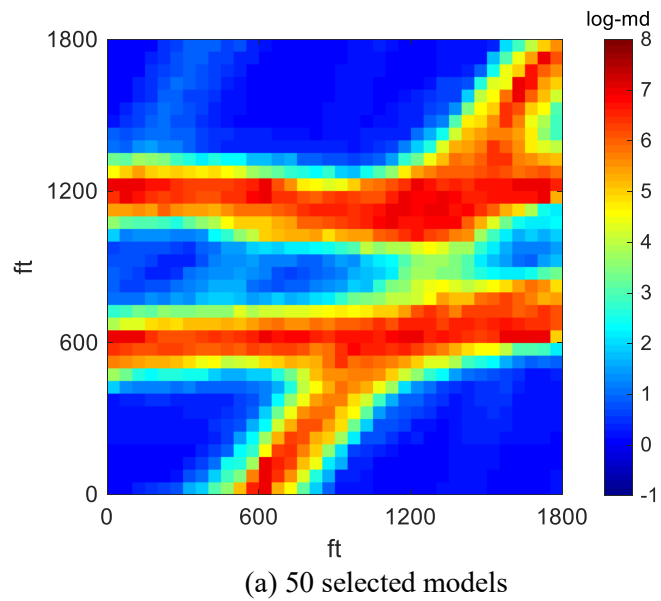
Figure 4.16 The mean of permeability values of the selected and generated
models for Case 2

The final selected models by the proposed method, Case 2a, is compared to final models selected by 2 other methods: Case 2b and Case 2c. The model selection methods for Case 2b and Case 2c are identical to Case 1b and Case 1c. Figure 4.17 represents the average permeability distribution of the final 10 models for Case 2a, 2b, and 2c. While overall patterns of the channels are characterized for the 3 case, the connection between the two parallel channels seems to be very week in Case 2c. Furthermore, Figure 4.17 shows that this connection appears more vividly in Cases 2a compared to Case 2b.

This observation can be supported by the permeability fields of the final models for the 3 case, in Table 4.3. In Case 2b and 2c, less than half of the models contain correct information of the channel patterns, whereas in Case 2a, the majority of the final models very similar channel patterns to the reference model. Using these fields, production forecasts are also compared for the 3 cases.

In Figure 4.18, future production responses of 1000 days are predicted for Case 2a. It is certain that the final 10 models in Case 2a allows to forecast the production reliably with reduced uncertainty range. In P2, water breakthrough does not occur before 500 days of production, which is the data used for the model selection. Nonetheless, Figure 4.18 shows that using the final models, the watercut behavior in P2 can be effectively predicted.

This cannot be found in Case 2b and Case 2c. In Case 2b, Figure 4.19 indicates that not only the uncertainty range in P2 is greater, but also there is a model that predicts water breakthrough in P1, which is incorrect. In Figure 4.20, significant increase in uncertainty of production forecasts with the final model in Case 2c can be found.

(a) reference model   (b) 400 initial models

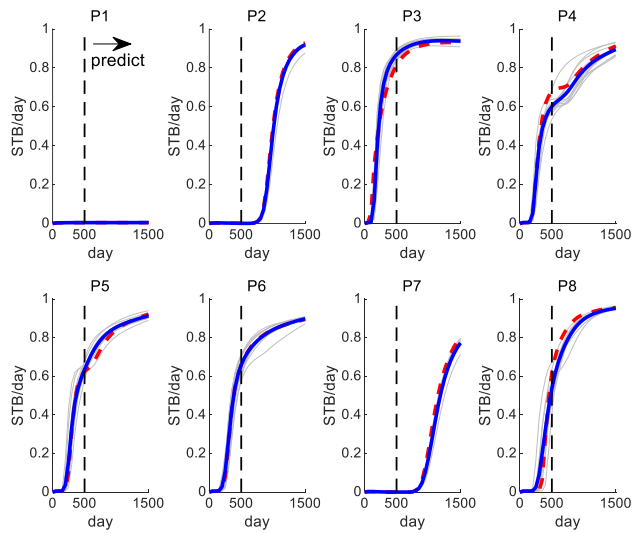(c) Case 2a   (d) Case 2b   (e) Case 2c
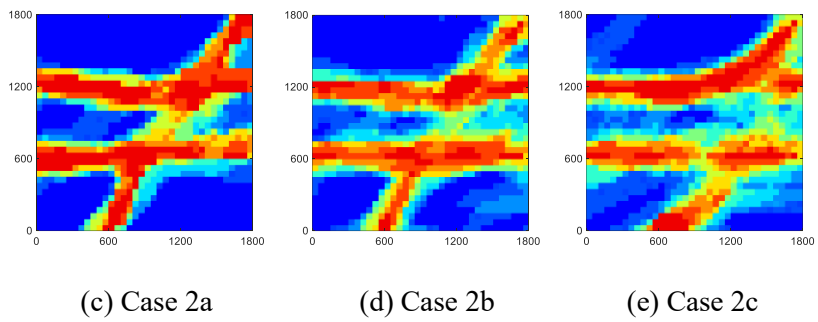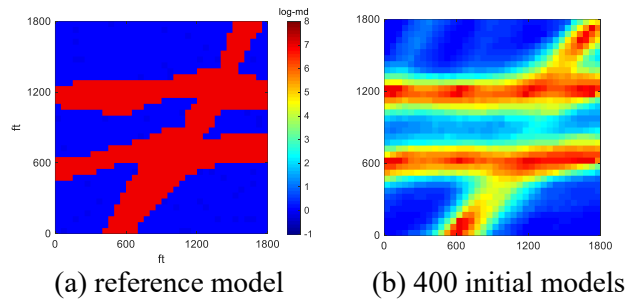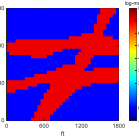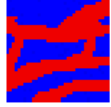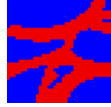
Figure 4.17 The mean of permeability values of the final 10 models for Case 2

Table 4.3 Permeability field of the final 10 models for Case 2

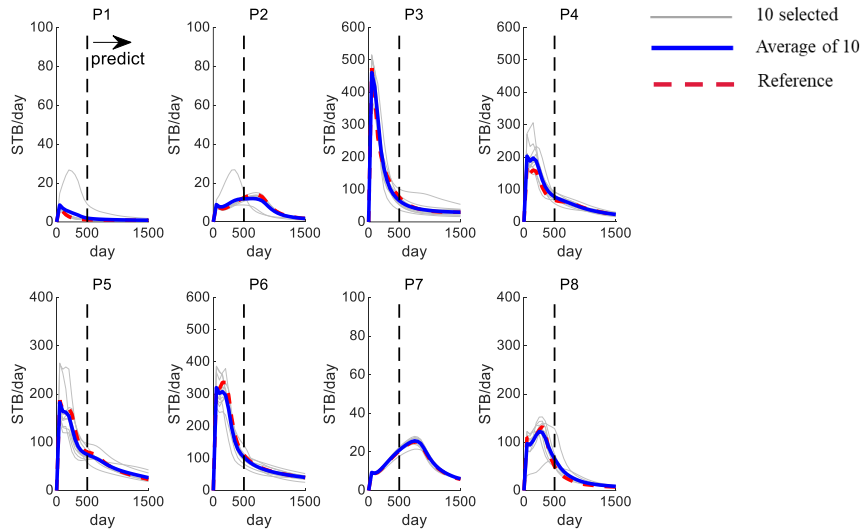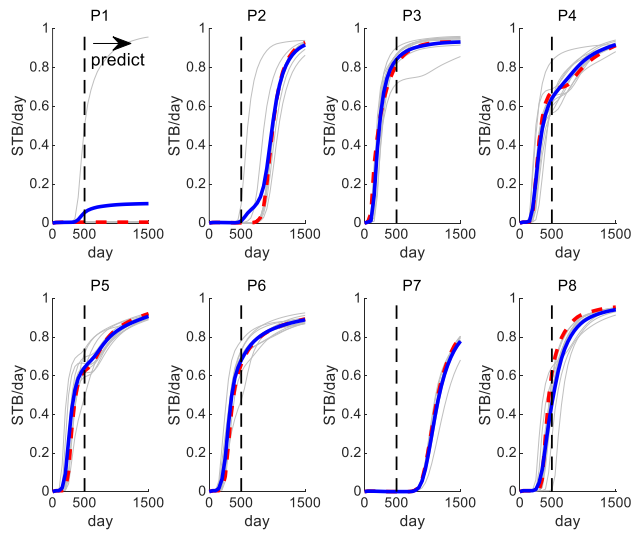| Reference	10 selected reservoir models					
	Case 2a					
	Case 2b					
	Case 2c					

(a) Oil rates

(b) Watercuts

Figure 4.18 Production responses of the final 10 models in Case 2a

(a) Oil rates

(b) Watercuts

Figure 4.19 Production responses of the final 10 models in Case 2b
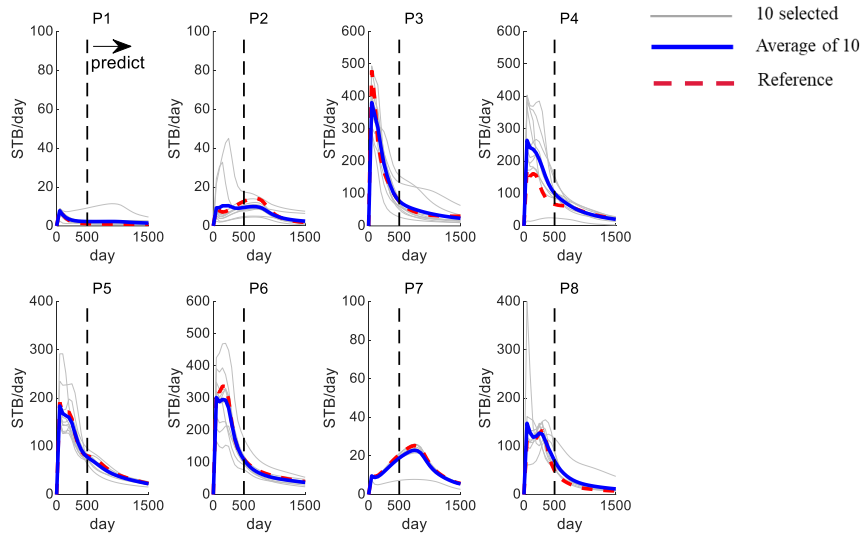
(a) Oil rates
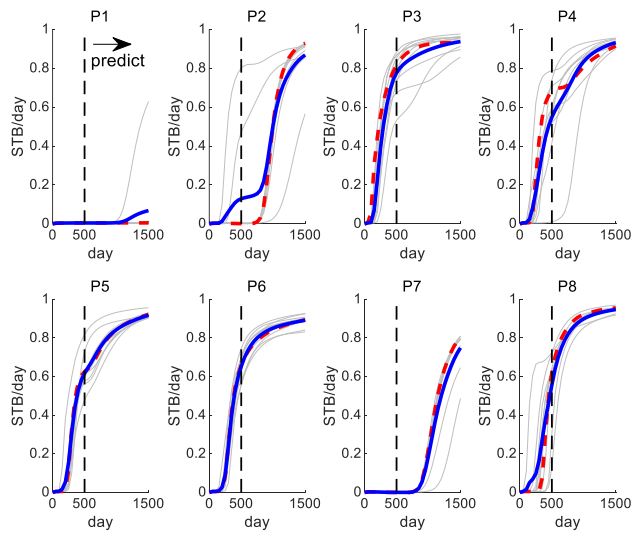
(b) Watercuts

Figure 4.20 Production responses of the final 10 models in Case 2c

4.3 Case 3

The proposed method is performed with more complicated channel reservoir models with line drive wateringflooding. SNESIM and geological information shown in Figure 4.21 are utilized for generating 401 models, and the size of the models is 55 by 55 by 1. The petrophysical parameters of the models are the same as Case 1 and 2. The reference model and the average permeability distribution of the 400 models can be found in Figure 4.22. There are 15 production wells and 10 injection wells in a waterflooding pattern of direct line drive.

The feature extraction and model selection of the model regeneration schem are implemented with production of 500 days obtained by ECLIPSE. The production wells are operated under a constraint of 400 psia for the bottomhole pressure limit. The injection wells are operated under a constraint of 300 STB/day for the water injection rate. The result of ranking the 400 models is analyzed in Figure 4.23. As shown in this figure, there is a tendency of the models with lower scores to have different connections of the channels compared to the reference model.

50 new reservoir models are generated by the fine-tuned DCGAN with the 50 selected models. Figure 4.24 represents the average permeability values of the 50 selected models and 50 generated models. After the feature extraction and model selection are repeated to rank these 100 models, top 10 ranked models are selected and utilized for predicting future production.
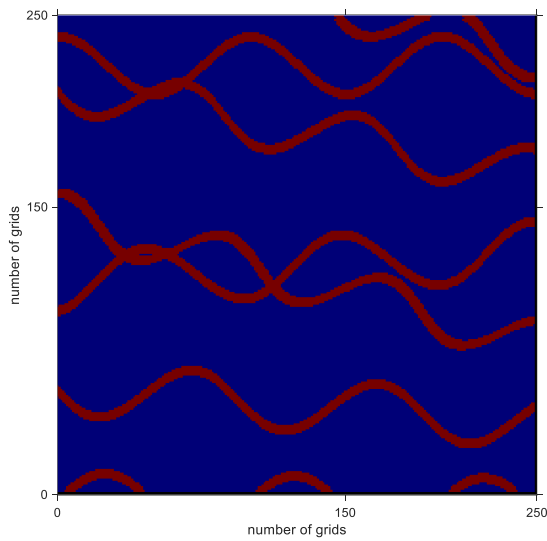
(a) Training image
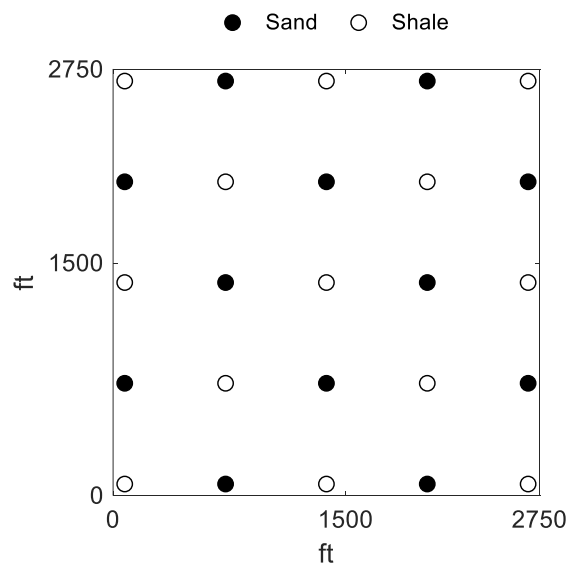
(b) Core sample data

Figure 4.21 Geological information for model generation for Case 3
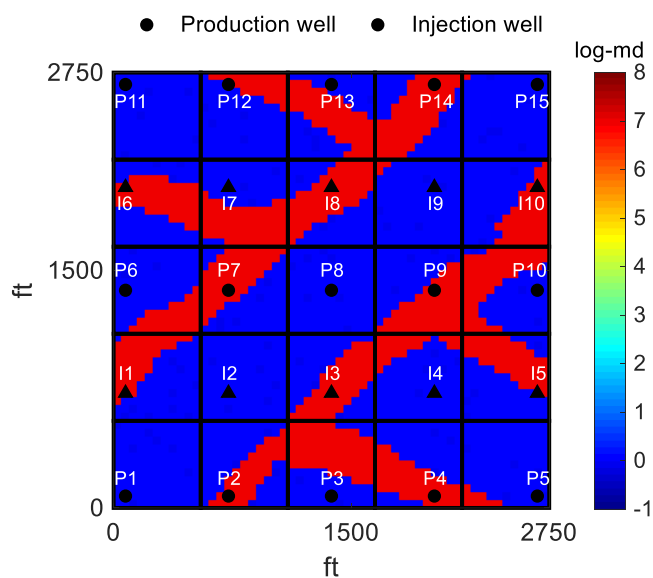
(a) Reference model
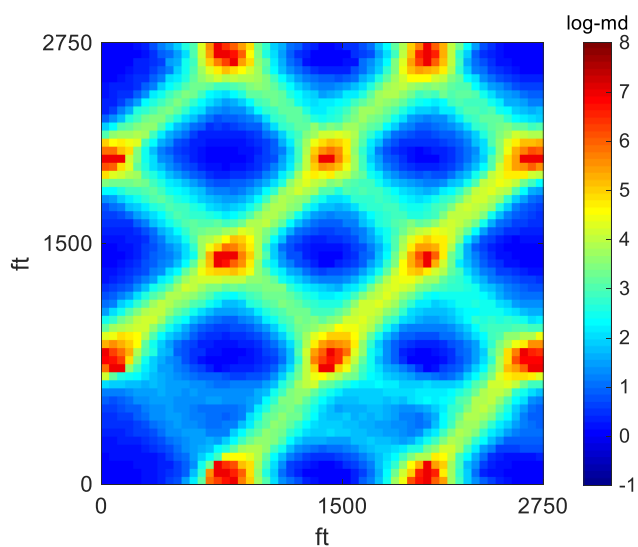
(b) 400 initial models

Figure 4.22 Reference model and the mean of permeability values for Case 3

High score

(b) Rank 1-100

(c) Rank 101-200

(a) Reference

(d) Rank 201-300

Low score

(e) Rank 301-400
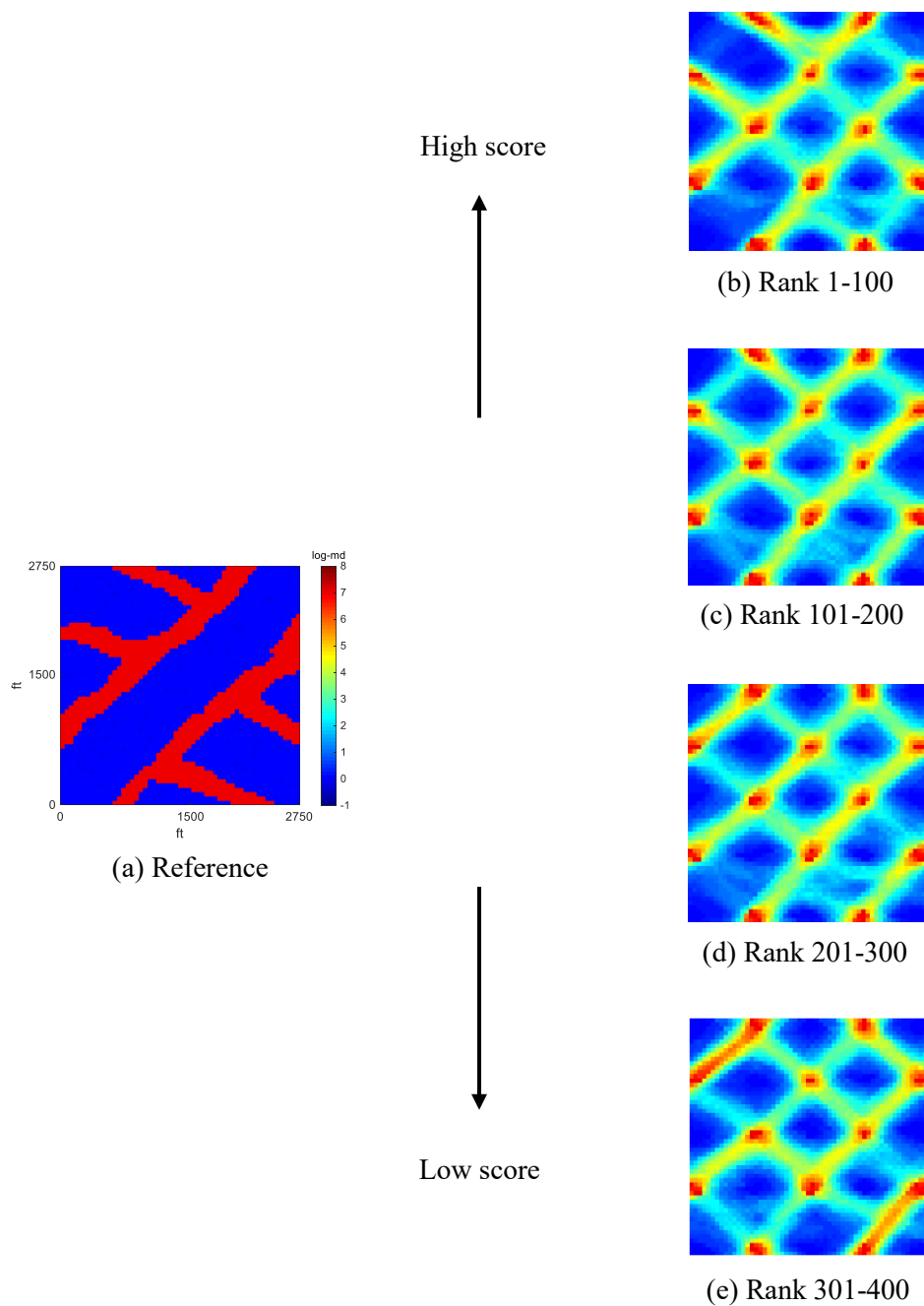
Figure 4.23 The result of ranking models for Case 3
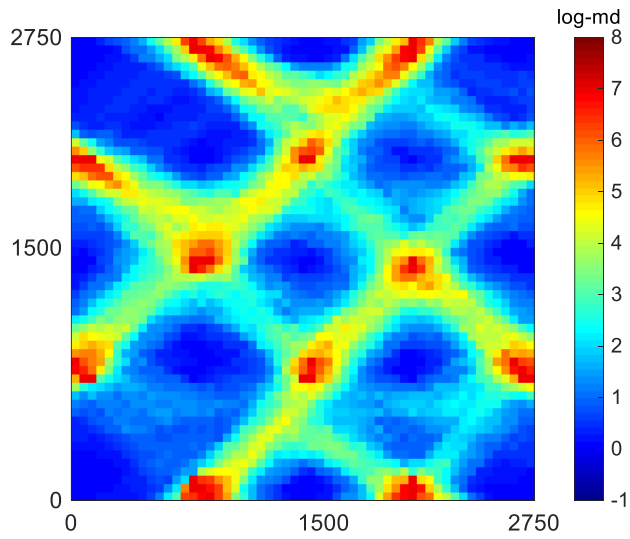
(a) 50 selected models
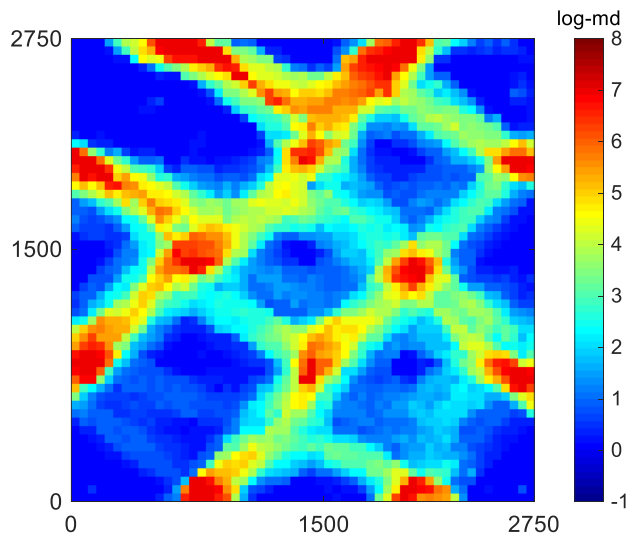
(b) 50 generated models

Figure 4.24 The mean of permeability values of the selected and generated models for Case 3

Identical to Case 1 and 2, permeability fields and production forecasts of the final models selected by the proposed method, Case 3a, are compared to the results of 2 other methods: Case 3b and Case 3c. The method for these are the same as the previous cases. According Figure 4.25, the final models in Case 3c have the very different trend of the channels, compare to the reference model. The average permeability distribution of the final models in Case 3a has a correct overall patterns of the channels located in the north west region of the reservoir. The connections of the other channels seem to be weak.

In Table 4.4, the permeability fields of the final models for the 3 cases can be found. Among the final selected models in Case 3c, none of the model has similar channel trends to the reference, whereas in Case 3b, there is 1 model that matches with the reference. Although Figure 4.25 shows that the final models with relatively correct channel characteristic are selected in Case 3a, only 2 of the models has a strong similarity in channel patterns with the reference, as shown in Table 4.4.

Future production responses of 1000 days with the final models for the 3 cases are analyzed in Figure 4.26, 4.27, and 4.28. Production data of only 6 wells are presented in these 3 figures. The 6 production wells are the ones with production rates above 100 STB/day. By comparing Figure 4.27 and 4.28, the uncertainty range of the future production seems to be smaller when the method for Case 3b is applied, especially in P2 and P4. Figure 4.26 shows that there is reduction in the overall uncertainty in Case 3a, compared to Case 3b. However, the decrease in the uncertainty is insignificant. Therefore, further analysis is performed.

(a) reference model　　(b) 400 initial models

(c) Case 3a　　(d) Case 3b　　(e) Case 3c
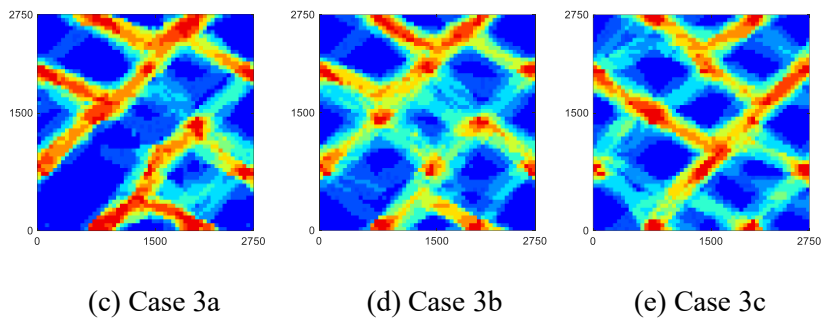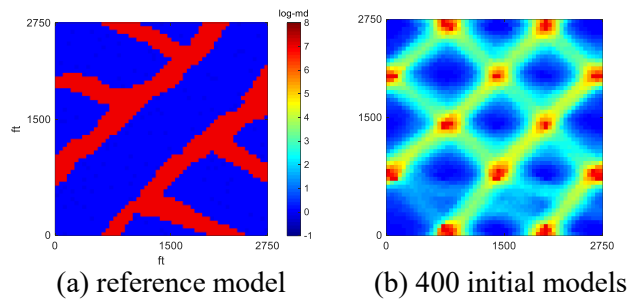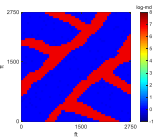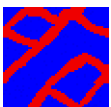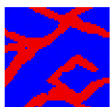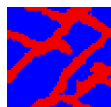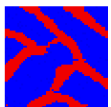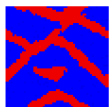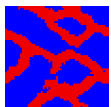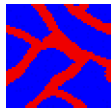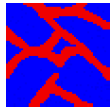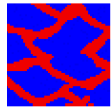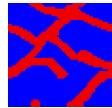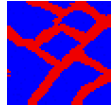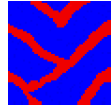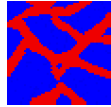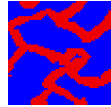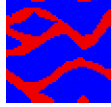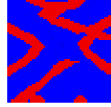
Figure 4.25 The mean of permeability values of the final 10 models
for Case 3a, 3b, 3c

Table 4.4 Permeability field of the final 10 models for Case 3a, 3b, 3c

(a) Oil rates

(b) Watercuts

Figure 4.26 Production responses of the final 10 models in Case 3a
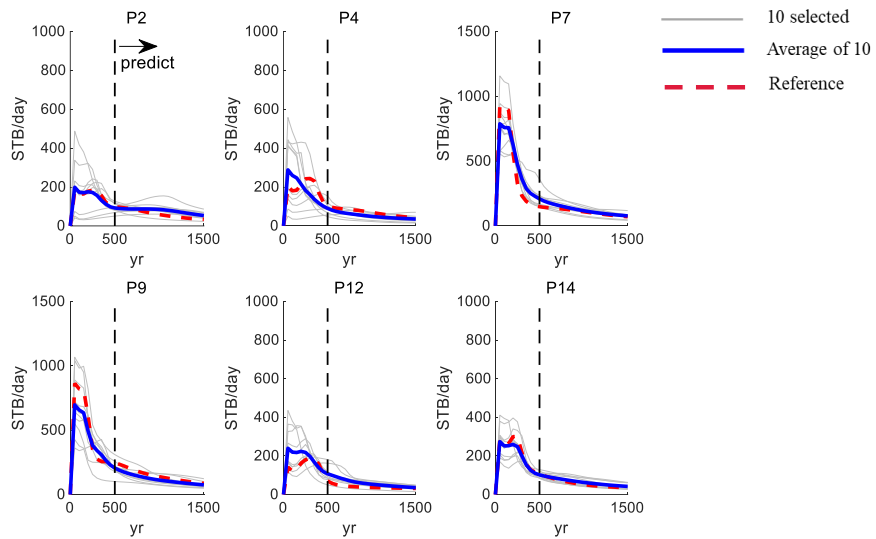
(a) Oil rates
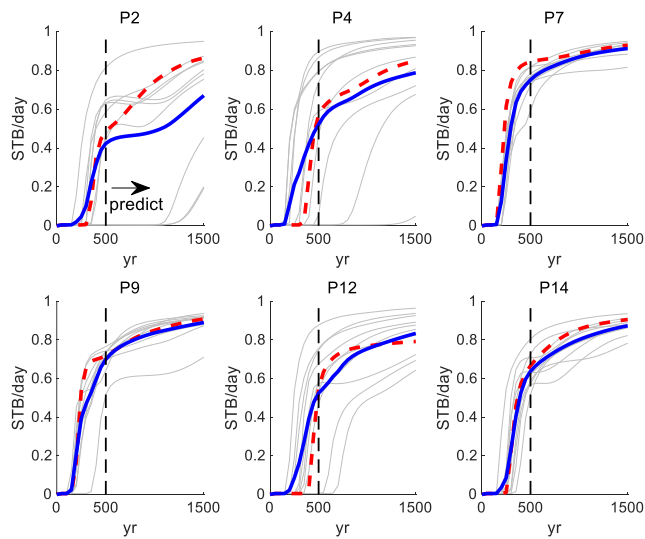
(b) Watercuts

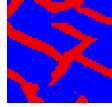Figure 4.27 Production responses of the final 10 models in Case 3b

(a) Oil rates

(b) Watercuts

Figure 4.28 Production responses of the final 10 models in Case 3c

Having only 1 reservoir model that matches with the reference model for Case 3b and 2 models for Case 3a infers that there are very few models that have similar channel characteristics in an ensemble of the initial models. To understand the effectiveness of generating new reservoir models by DCGAN, 2 additional cases are created: Case 3d and Case 3e. They are identical to Case 3a, expect the number of new generated models. In Case 3d, 10 final models are selected from 50 selected and 100 generated models while in case 3e, 10 final models are selected from 50 selects and 150 generated models.

Figure 4.29 shows that as the number of the new generated models are increased, the average permeability distribution of the final models becomes stronger in the correct channel connections. This observation is also found in Table 4.5. The more reservoir models are generated by DCGAN, the more models that are very similar to the reference model are selected. While there are 2 models that match with the reference model in Case 3a, there are 3 models and 5 models like them in Case 3d and 3e, respectively. When generating more models by DCGAN, not only the number of the models that match with the reference increases, but also the models with greater similarity of the channel characteristics to the reference model are created.

Figure 4.30 and 4.31 show the future forecasts of 1000 days for Case 3d and 3e. Comparing Figure 4.26 and 4.30 allows to identify the effectiveness of increasing the number of new generated models for reducing in uncertainty. The uncertainty range in the majority wells are significantly decreased in Case 3d, compared to 3a. Figure 4.31 demonstrates further reduction in the uncertainty range in Case 3e, which leads to reliable future forecasts.

(a) reference model

(b) 400 initial models
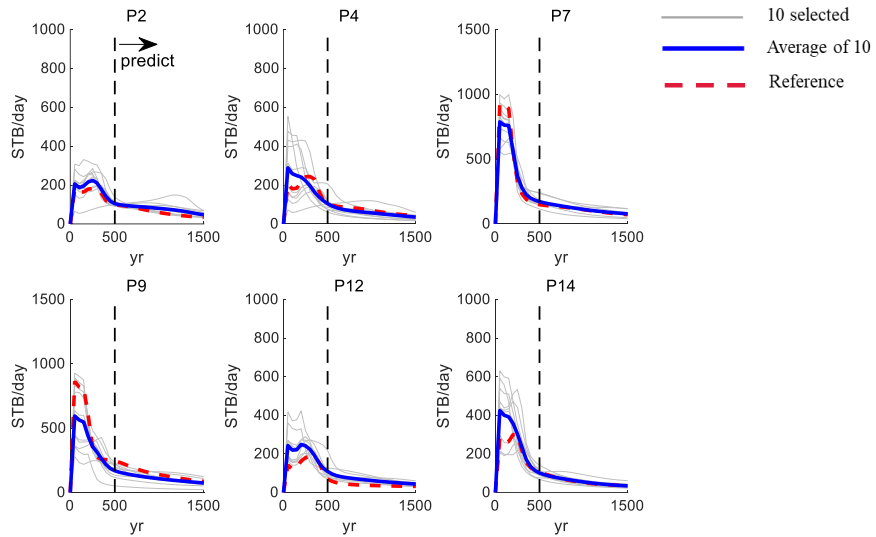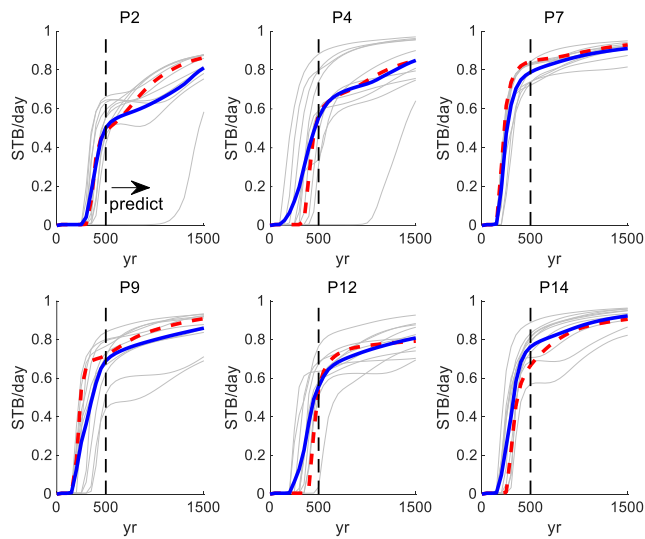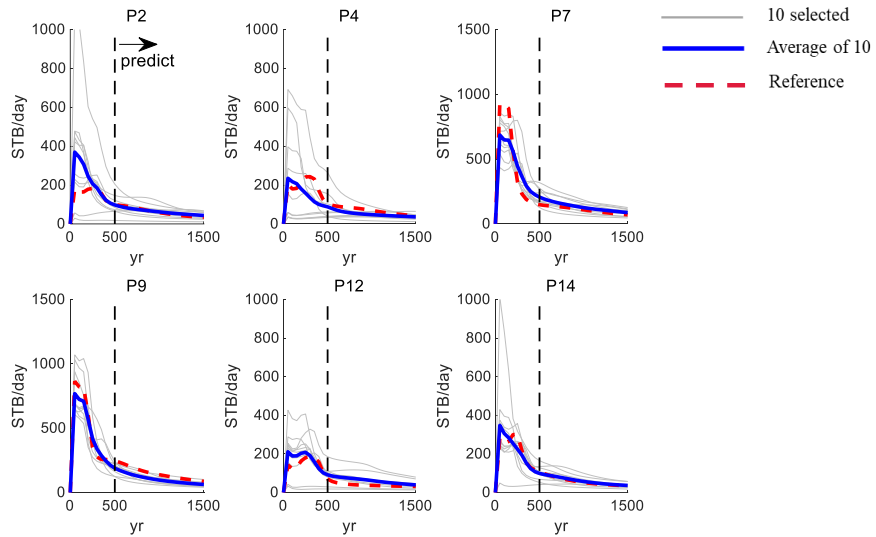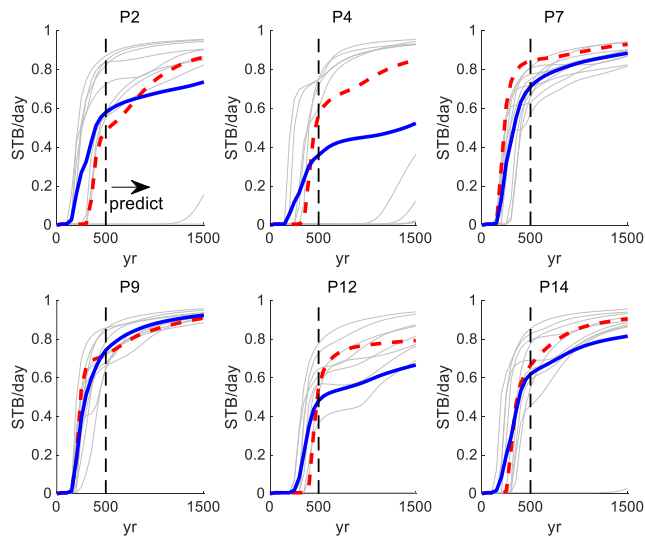
(c) Case 3a
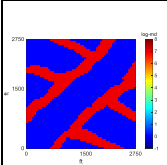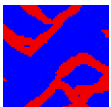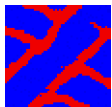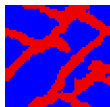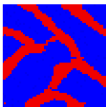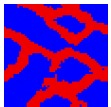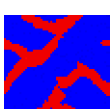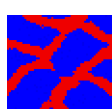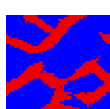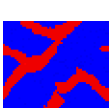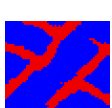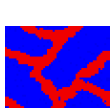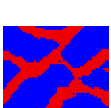
(d) Case 3d

(e) Case 3e
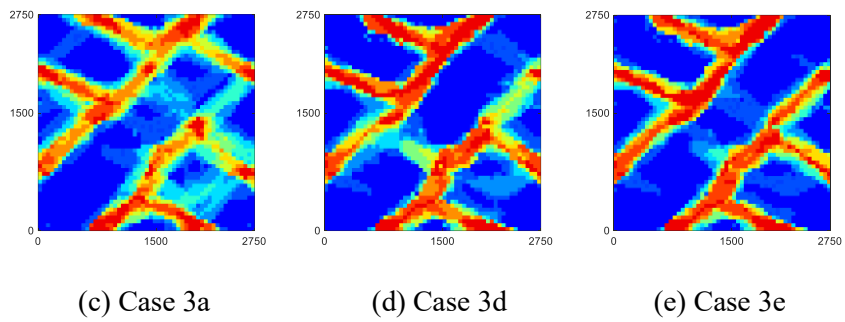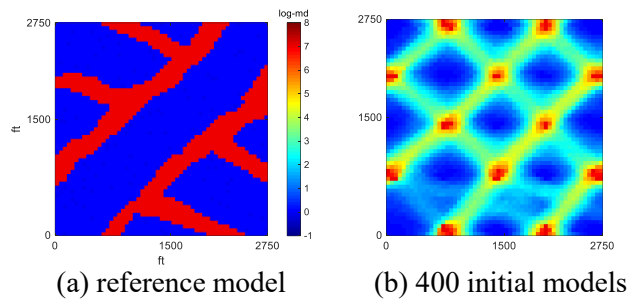
Figure 4.29 The mean of permeability values of the final 10 models
for Case 3a, 3d, 3e

Table 4.5 Permeability field of the final 10 models for Case 3a, 3d, 3e

(a) Oil rates

(b) Watercuts

Figure 4.30 Production responses of the final 10 models in Case 3d
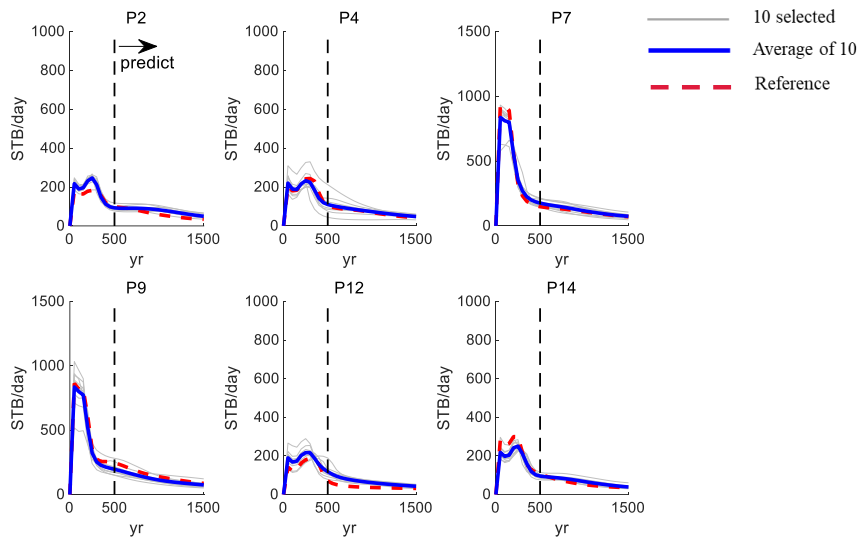
(a) Oil rates
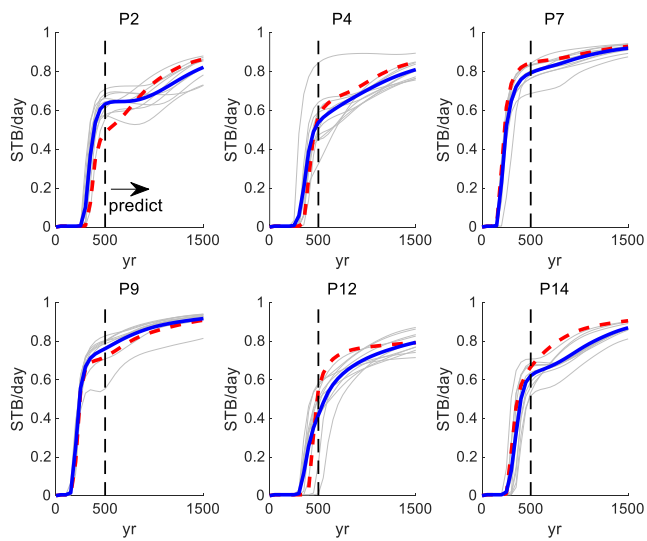
(b) Watercuts

Figure 4.31 Production responses of the final 10 models in Case 3e

5. Conclusions

This research introduces a model regeneration scheme for reliable uncertainty quantification of channel reservoirs. The scheme is to effectively extract features of channels in near-wellbore area, to select reservoir models from an ensemble, and to regenerate new models. In the feature extraction process, drainage area localization is implemented on the reservoir models to gain near-wellbore data. A geological 2D reservoir is divided into smaller sections as many as the number of production wells in the reservoir. Features of the obtained near-wellbore data are extracted by using DCT. With coefficients of DCT, the models are differentiated for every near-wellbore area of the production wells.

In the model selection process, K-means clustering is performed to classify them into clusters with the extracted information of the features. Based on production responses of a representative model of each cluster, a score is calculated and assigned to reservoir models within the cluster. By implementing this process for every production well, the reservoir models are organized in ranking by the sum of their scores in descending order. The top ranked reservoir models are selected,

In the model generation process, DCGAN is trained with the selected models to generated reservoir models that do not previously exist. To increase the performance of DCGAN training, transfer learning is applied by fine-tuning a pre-trained DCGAN. Pre-training DCGAN is performed with the all initial models in the ensemble.

After the model generation process, the feature extraction and model

selection process are repeated with the selected and generated models. By applying the proposed method with the model regeneration scheme to 3 different 2D channel reservoirs, the following conclusions are drawn.

1. Utilizing information of main trends of channels in near-wellbores improves characterization of local channel connectivity and patterns. By comparing several cases in this study, it is certain that selecting models with near-wellbore data allows better local channel characterization and improves the performance of the production forecasts.

2. Evaluating models for production wells individually allows to organize the models by degree of similarity to a true model. The model ranking result shows a vivid trend of the channel characteristic change depending on their ranking. While the highest ranked models have the most proper characteristics, its similarity decreases as the ranking goes down.

3. Under the circumstance of high uncertainty in channel reservoirs, generating new models with selected models increases probability of the existence of models similar to a true model. As the number of the generated models increases, the characterization results are improved with the decreased uncertainty in the production forecasts. More proper models are selected as final models (Case 3).

In this study, the proposed method is analyzed by using the 2D channel reservoir models with waterflooding patterns. For channel reservoirs in which wells are not evenly distributed, the method can be improved with streamline simulation. Tracing streamlines can be able to help reasonable decision of drainage area for the wells. Furthermore, the proposed method can be developed for reliable uncertainty quantification of 3D channel reservoirs. By increasing the dimension of DCT, it can extract features of 3D channels. Also, replacing the convolution layers used in DCGAN to 3D convolution layers can allow to generate 3D reservoir models.

Reference

Chang, Y., Stordal, A.S., and Valestrand, R. 2016. Integrated work flow of preserving facies realism in history matching: application to the brugge field *SPE Journal* 21 (4), 1413-1424.

Choe, J. 2013. Geostatistics, revised ed. Seoul, Korea: Sigma Press.

Choe, J. 2017. Offshore Drilling Engineering, 2nd ed. Seoul, Korea: CIR Press.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. 2014. Generative adversarial networks. *Advances in Neural Inforamtion Processing Systems*.

Guo, Z., Wan, Y., and Ye, H. 2019. A data imputation method for multivariate time series based on generative adversarial network. *Neurocomputing* 360, 185–197.

Jafarpour, B. and McLaughlin, D.B. 2007. Efficient Permeability Parameterization with the Discrete Cosine Transform, SPE Reservoir Simulation Symposium, Houston, TX, Feb. 26–28, *SPE* Paper No. SPE-106453-MS.

Jo, H., Jung, H., Ahn, J., Lee, K., and Choe, J. 2017. History Matching of Channel Reservoirs Using Ensemble Kalman Filter with Continuous Update of

Channel Information. *Energy Exploration & Exploitation* 35 (1), 3-23.

Jung, H., Jo, H., Kim, S., Lee, K., and Choe, J. 2018. Geological model sampling using PCA-assisted support vector machine for reliable channel reservoir characterization. *Journal of Petroleum Science and Engineering* 167, 396-405.

Jung, H., Jo, H., Lee, K., and Choe, J. 2017a. Characterization of Various Channel Fields Using an Initial Ensemble Selection Scheme and Covariance Localization. ASME. *J. Energy Resour. Technol*. 139(6): 062906.

Jung, H., Jo, H., Kim, S., Lee, K., and Choe, J. 2017b. Recursive Update of Channel Information for Reliable History Matching of Channel Reservoirs Using EnKF with DCT. *Journal of Petroleum Science and Engineering* 154, 19-37.

Jung, S. and Choe, J. 2010. Stochastic Estimation of Oil Production by History Matching with Ensemble Kalman Filter. *Energy Sources, Part A* 32 (10), 952-961.

Jung, S. and Choe, J. 2012. Reservoir characterization using a streamline-assisted ensemble Kalman filter with covariance localization. *Energy Exploration & Exploitation* 30 (4), 645-660.

Kang, B. and Choe, J. 2017. Regeneration of Initial Ensembles with Facies Analysis for Efficient History Matching. *Journal of Energy Resources*

Technology 139 (4), 042903.

Kang, B. and Choe, J. 2020. Uncertainty quantification of channel reservoirs assisted by cluster analysis and deep convolutional generative adversarial networks. *Journal of Petroleum Science and Engineering* 187, 106742.

Kang, B., Jung, H., Jeong, H., and Choe, J. 2019. Characterization of three-dimensional channel reservoirs using ensemble Kalman filter assisted by principal component analysis. *Petroleum Science* 17, 182-195.

Kim, S., Lee, C., Lee, K., and Choe, J. 2016. Aquifer characterization of gas reservoirs using ensemble kalman filter and covariance localization. *Journal of Petroleum Science and Engineering* 146, 446-456.

Kim, S., Lee, C., Lee, K., and Choe, J. 2016. Characterization of Channelized Gas Reservoirs Using Ensemble Kalman Filter With Application of Discrete Cosine Transformation. *Energy Exploration & Exploitation* 34 (2), 319-336.

Lee, K., Lim, J., Choe, J., and Lee, H.S. 2017. Regeneration of Channelized Reservoirs Using History-matched Facies-probability Map Without Inverse Scheme. *Journal of Petroleum Science and Engineering* 149, 340-350.

Lee, K., Lim, K., Ahn, S., and Kim, J. 2018. Feature extraction using a deep learning algorithm for uncertainty quantification of channelized reservoirs. *Journal of Petroleum Science and Engineering* 171, 1007-1022.

Lee, J. and Choe, J. 2016. Reliable Reservoir Characterization and History Matching Using a Pattern Recognition Based Distance. Presented at the ASME 2016 35TH International Conference on Ocean, Offshore and Arctic Engineering, Busan, Korea, 19-24 June.

Noguchi, A. and Harada, T. 2019. Image Generation from Small Datasets via Batch Statistics Adaptation. arXiv:1904.01774.

Radford, A., Metz, L., and Chintala, S. 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. arXiv:1511.06434.

Shin, Y., Jeong, H., and Choe, J. 2010. Reservoir Characterization Using an EnKF and a Non-parametric Approach for Highly Non-Gaussian Permeability Fields. *Energy Sources, Part A* 32 (16), 1569-1578.

Wang, Y., Wu, C., Gonzalez-Garcia, A., Berga, D., Herranz, L., Khan, F., and Weijer, J. 2018. Transferring GANs: generating images from limited data. arXiv:1805.01677.

Wang, Y., Wu, C., Herranz, L., Weijer, J., Gonzalez-Garcia, A., and Raducanu, B. 2018. Transferring GANs: generating images from limited data. arXiv:1805.01677.

Yeo, M., Jung, S., and Choe, J. 2014. Covariance Matrix Localization Using

Drainage Area in an Ensemble Kalman Filter, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 36:19, 2154-2165

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. 2014. How transferable are features in deep neural networks?. arXiv:1411.1792.

Zhou, H., Li, L., and Franssen, H. 2012. Pattern recognition in a bimodal aquifer using the normal-score ensemble Kalman Filter. *Mathematical Geosciences* 44(2), 169–185.

국문초록

딥러닝 기반 모델 재생성 기법을 이용한 채널저류층의 불확실성 평가

불균질성이 높은 채널저류층에서의 생산거동은 채널의 특징에 큰 영향을 받는다. 기존에는 생산자료를 매칭하는 역산기법을 통해 저류층 모델을 교정하였다. 하지만 이 기법은 채널저류층에 적용할 경우 지질학적 특징 손실과 초기모델에 큰 영향을 받는 한계가 있다.

따라서 본 연구에서는 생산자료를 매칭하는 전통적인 역산과정 없이 효과적으로 채널저류층의 불확실성을 평가할 수 있는 모델 재생성 기법을 제안한다. 모델 재생성 기법은 특징추출, 모델선정, 모델생성으로 구성되어 있다. 특징추출에서는 배수구역 지역화와 이산 코사인변환법을 적용하여 유정 인근지역의 채널의 특징을 추출한다. 모델선정에서는 K-평균 군집법과 앙상블순위 기법으로 참조모델과 유사한 모델들을 선정한다. 모델생성에서는 생성적 적대 신경망과 전이학습을 통해 새로운 모델을 생성한다.

제안방법은 특징추출과 모델선정으로 참조모델과 유사한 모델들을 선정하고 이들을 사용하여 새로운 모델을 생성한다. 특징추출과 모델선정을 반복하여 선정한 모델과 생성한 모델로부터 최종 모델을 선정한고 미래 생산량을 예측하여 그 불확실성을 평가한다. 제안방법을 세 개의 2차원 케이스에 적용하였다. 적용 결과, 제안방법은 효과적으로 채널과 유정 인근지역의 특징을 파악할 수 있는 것을 확인하였다. 이를 통해 참조모델과 유사한 모델들을 생성함으로써 미래 생산량 예측성능을 향상시킬 수 있는 것을 확인하였다.