Multi-scale, Multi-dimensional Reservoir Characterization Using Advanced Analytics and Machine Learning

Roozbeh Koochak, B. Eng.

A thesis submitted for the degree of Doctor of Philosophy (PHD)

Discipline of Mining and Petroleum Engineering

Faculty of Science, Engineering and Technology

The University of Adelaide

April 23

Table of Contents

Abstract

This thesis aims to investigate novel approaches in the field of Machine learning and advanced data analytics that can handle large data volumes and open new doors in the field of reservoir characterization.

To begin, a new approach for rock typing is introduced using fractal theory where conventional resistivity logs are the only required data. Fractal analysis of resistivity logs showed that the fractal dimension of these logs which is a measure of the variability of the signal, is related to the complexity of the rock fabric. the fractal dimension of multiple deep resistivity logs in the Cooper Basin, Australia was measured and compared with the fabric structure of cores from same intervals. The results showed that the fractal dimension of resistivity logs increases from 1.14 to 1.29 Ohm-meter for clean to shaly sands respectively, indicating that the fractal dimension increases with complexity of rock texture.

The thesis continues with a machine learning application to augment/automate facies classification using resistivity image logs. Given the complexity of the application, a supervised learning strategy in combination with transfer learning was used to train a deep convolutional neural network on available data. The results show that in the absence of other information/logs, the trained network can detect image facies with a testing accuracy of 82% form electric image logs and a proposed post-processing method increases the final categorization accuracy even further.

An important step in reservoir characterization is understanding and quantification of uncertainty in reservoir models. In the next section a novel Generative Adversarial Network (GAN) architecture is introduced which can generate realistic geological models while

maintaining the variability of the generated dataset. The concept of mode collapse and its adverse effect on variability is addressed in detail. The new architecture is applied to a binary channelized permeability distribution and the results compared with those generated by Deep Convolutional GAN (DCGAN) and Wasserstein GAN with gradient penalty (WGAN-GP). The results show that the proposed architecture significantly enhances variability and reduces the spatial bias induced by mode collapse, outperforming both DCGAN and WGAN-GP in the application of generating subsurface property distributions.

Finally, an advanced analytics technique for efficient history matching is proposed in the appendix. In this part of the thesis, an ensemble of surrogates (proxies) with generation-based model-management embedded in CMA-ES is proposed to reduce the number of simulation calls efficiently, while maintaining the history marching accuracy. History matching for a real field problem with 59 variables and PUNQ-S3 with eight variables was conducted via a standard CMA-ES and the proposed surrogate-assisted CMA-ES. The results showed that up to 65% and 50% less simulation calls for case#1 and case#2 were required.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Roozbeh Koochak 02/04/2023

Acknowledgement

I am grateful to all those whose belief in me, encouragement, and support have been the driving force behind my achievements.

Sahba, my loving wife, your endless patience, understanding, and belief in me have been the bedrock of my success. Your constant support and sacrifices made it possible for me to focus on my research and overcome challenges. Your presence by my side has been a source of strength, motivation, and joy, and I am forever grateful for your unwavering love.

To my dear mother, Farideh, and siblings, Atousa, Reza, and Parisa, your consistent support, encouragement, and understanding have been invaluable. Your belief in my abilities and unwavering support, both emotionally and practically, have been instrumental in my accomplishments. Your presence in my life has filled me with gratitude and motivation to excel.

I would like to extend my deepest appreciation to my esteemed supervisors, Dr. Manoucheher Haghighi, Dr. Mohammad Sayyafzadeh, and Dr. Mark Buch. Your expert guidance, knowledge, and dedication have shaped my research, broadened my horizons, and enriched my understanding of the subject matter. Your tireless efforts to provide me with valuable feedback, constructive criticism, and mentorship have played a crucial role in my academic growth and the successful completion of my Ph.D. I am truly grateful for the opportunities you have given me to explore and contribute to the field.

Lastly, I would like to thank my dear friends, Dr. Ali Nadian, Colin Jordan, and Dr. Martin Roberts for their unwavering support, interest in my work, and willingness to lend a helping hand whenever needed. Your intellectual discussions, feedback, and collaborative spirit have been instrumental in shaping my ideas and sharpening my research focus.

Thesis by publication

This is a thesis by publication and is composed of multiple pieces of work. This includes 4 publications in total. 2 published in peer-reviewed journals, 1 submitted for publication to a peer-reviewed journal and 1 conference paper. These are the details below:

Published Peer-reviewed Journal Papers:

Koochak, R., Haghighi, M., Sayyafzadeh, M. and Bunch, M., 2018. Rock typing and facies identification using fractal theory and conventional petrophysical logs. *The APPEA Journal*, *58*(1), pp.102-111.

Koochak, R., Sayyafzadeh, M., Nadian, A., Bunch, M. and Haghighi, M., 2022. A variability aware GAN for improving spatial representativeness of discrete geobodies. *Computers & Geosciences*, *166*, p.105188.

Submitted to Peer-reviewed journal for publication:

Roozbeh Koochak, Ali Nadian Ghomsheh, Manouchehr Haghighi, Mark Bunch, Mohammad Sayyafzadeh, A transfer learning approach for facies prediction using resistivity image well logs. Submitted to Geoscience Frontiers for publication.

Published Conference paper:

Sayyafzadeh, M., Koochak, R. and Barley, M., 2018, September. Accelerating cma-es in history matching problems using an ensemble of surrogates with generation-based management. In *ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery* (Vol. 2018, No. 1, pp. 1-15). EAGE Publications BV.

1. Introduction

1.1 Problem statement

Reservoir characterization includes all techniques and methods that enhance our understanding of the geologic, Geo-chemical and petrophysical controls of fluid flow. it is a constant process that Continuously evolves from field discovery to development and production down to abandonment phase. Throughout these stages, knowledge of the subsurface continually enhances as new data is received. This makes characterization of the reservoir a dynamic process. The challenge is that the dynamic properties of the reservoir such as Pressure, Saturation or Porosity continually change as the reservoir is produced. This together with the fact that most of the data measurements are indirect along with error of measurement, causes uncertainty in the reservoir characterization. Adequate characterization of a reservoir is increasingly important for optimising field development, reservoir evaluation and production. Also, in recent years with novel subsurface applications such as carbon capture and sequestration or hydrogen storage, detailed characterization is imperative to successful operations throughout the life a field. An integrated approach for characterization and/or modelling of the reservoir can tear down traditional disciplinary divides and lead to better understanding and handling of uncertainties in the reservoir. With this statement in mind, the methodologies presented in this thesis have multidisciplinary applications where the disciplines of Petrophysics, Reservoir Engineering and Geology have been brought together. Reservoir characterization generally involves estimating reservoir parameters at different locations by correlating collected data from a wide variety of sources. Sources of data for subsurface reservoirs include cores, well logs, Seismic surveys, production data and outcrop analogues. These sources of data all have different scales and vary in dimensionality. Logs have

high resolution but on a small scale, generally, 1D or 2D in dimension. Well logs provide a vertically high-resolution model at the well locations. However, the distribution of well locations is sparse and biased towards proven section of the field. Seismic data on the other hand is large scale and covers extensive areas but with low resolution. Dimensionality of this data type is 2D, 3D and sometime 4D. Combining well logs with geophysical and geological data will provide the necessary constrains required to extrapolate the high resolution well data beyond where they are measured. Core data, as 3-dimension datatype, is classed as hard data and presents the most accurate information with highest resolution however their availability is limited as the acquisition of cores, is expensive and requires a lot of effort. The most effective utilisation of these data sources is, therefore, combining these different data sources. This approach results in the best and most complete description of reservoir which is generally referred to as integration modelling. Integration of different data sources enhances understanding of the reservoir, reduces uncertainties and mitigates risks. The goal of characterization is to develop different models that can be used in analytical or numerical evaluation methods.

The oil and gas industry are experiencing a surge in the amount of data they are receiving from their fields. Field data, in recent years, has expanded in volume, velocity and complexity. This includes significant increase in the number of sensors in the field and in the pipelines, connected through Internet of Things (IoT). Resolution and sampling rate of wireline data has increased and complex data such as 4D seismic are becoming more common and more frequently recorded. Advent of technologies such as fibre optics has made access to these data virtually instant. The industry has always been overwhelmed with large quantities of data but was never able to make efficient and productive use of this data. Traditional methods in subsurface energy generally need to compromise between data size and complexity on one

hand and fidelity of the model (they are used to construct) on the other. In recent years Data analytics and Artificial Intelligence (AI) - Machine Learning (ML) more specific in this thesis - have introduced new methodologies that not only handles large and complex data types but is able to process the data much faster than traditional methods. This in part is also driven by advances in commodity hardware. Furthermore, ability of these methods to identify and learn features and patterns in the data makes them a great tool to unlock new insights, extract more information, develop new usage and leverage and optimize untapped data.

In this work, the aim is to develop and implement advanced analytics and machine learning techniques with a focus on reservoir characterization using multi-scale, multi-dimensional data. Firstly, a novel application of fractal dimension of resistivity logs is presented. In this one-dimensional, small-scale application porosity and permeability derived from cores was rock typed into categories and the fractal dimension of the corresponding deep resistivity log was calculated. The correlation between the rock fabric and fractal dimension of the resistivity logs was investigated and a new method to make direct use of these logs in rock typing is proposed. This research investigates the effect of pore structure on the variability of resistivity logs. It takes a step further than just interpreting resistivity logs based on change in average value and signature and reveals information at pore scale. In this study only the effect of pore structure on the variability and fractal dimension of logs was investigated. Further, research is required to determine the effect of other factors like fluid in the rock or different shale types. Bearing in mind that fractal dimension is rather a quality parameter and a flag for change in pore structure. Then in a 2-dimensional small-scale application, a machine learning algorithm complete with practical data pipeline is introduced to augment/automate the tedious process of resistivity image log interpretation. In this technique, a convolutional neural network (CNN) is trained to learn interpreted facies categories in one well, then the trained network is used

to detect the learned facies categories in a newly drilled well. The ability of CNNs to learn and identify features in geological facies images is thoroughly investigated by studying the confusion matrix. Multiple CNN architectures are compared and challenges in the application are identified, and solutions presented. Next, we apply machine learning to a 2-dimensional large-scale application. In this study, Generative Adversarial Networks (GANs) are used to quantify uncertainty in a field or basin wide scale. The concept of Training Image (TI) is reviewed. Methodology to generate network training data from a single TI is presented and compared with other traditional geo-statistics methods. The variability of generated realizations (which is detrimental in geological uncertainty quantification) using GANs is thoroughly investigated including the concept of mode collapse and the effect of input training data. A novel architecture specialized for maintaining the variability of geological realizations at the same level as the input training data is presented.

Working with large data sets and complex algorithm, computation efficiency and cost becomes an important factor in popularity of a characterization method. We enhance the computational efficiency of a Covariance Matric Adaptation Evolutionary Strategy (CMA-ES) by proposing an online learning scheme to update an ensemble of proxies. The effective ness of the technique was evaluated on two different history matching cases and other techniques such as generation-based model management and evolution control were examined.

1.2 Thesis Structure

This is a thesis by publication. Chapter 1 begins with an introduction to reservoir characterization and its challenges. It describes how novel advanced analytics techniques and machine learning algorithms can enhance the process of reservoir characterization and describes the aims of this thesis. The contribution of each publication to this thesis, is also

presented in this chapter. Chapter 2 follows with a comprehensive literature review. Subsequent chapters including 3, 4, 5 are composed of peer-reviewed journal papers. Chapter 3 demonstrates that the fractal dimension of resistivity logs is indicative of the complexity of rock pore fabric, and therefore can be used to define rock types. In chapter 4 a machine learning application for automated/ augmented facies classification using Convolutional Neural Networks is presented. Chapter 5 investigates a novel GAN architecture for generating realistic and statistically faithful geological realizations. Chapter 6 summarises the thesis with summary, conclusions, and future work. Finally, a conference paper in the appendix closes the thesis with an advanced analytics technique for efficient history matching. In the appendix an ensemble of surrogates with generation-based model-management embedded in CMA-ES is proposed to reduce the number of simulation calls while maintaining the history matching accuracy. Table 1 gives an overview of chapter numbers, their titles and the papers contributing to them along with publication status.

Table 1: Thesis structure

Chapter Number	Chapter Title	Paper number	Status
3	Rock typing and facies identification using fractal theory and conventional petrophysical logs	1	Published
4	A transfer learning approach for facies prediction using resistivity image well logs	2	Submitted
5	A variability aware GAN for improving spatial representativeness of discrete geobodies	3	Published
Appendix	Accelerating CMA-ES In History Matching Problems Using an Ensemble of Surrogates with Generation-Based Management	4	Published

1.3 Contribution of each publication to this thesis

The overall theme of this thesis is applications of machine learning and advanced analytics in reservoir characterization with a focus on multi-scale multi-dimensional data. Advanced

analytics and state-of-the-art techniques such as machine learning are utilized to better understand uncertainty, develop new methods, and potentially extract more information and propose new usage of the traditional data, while considering, different data sources with different scales and dimensionality. Furthermore, computational efficiency and automation is addressed where these techniques are adaptable.

This work begins by introducing a new application for conventional resistivity logs which are one dimensional and on a centimetre scale. In the paper titled "Rock typing and facies identification using fractal theory and conventional petrophysical logs" a technique to use fractal dimension of resistivity logs for rock typing and flow unit classification is proposed. Rock typing is an integral part of reservoir characterization. In this process the reservoir is subdivided into layers based on similar properties and flow points. In other words, rock fabric of each layer, that is, pore throat dimensions, geometry, size, distribution, and capillary pressures must be similar. This enhances flow behaviour modelling and, significantly reduces uncertainty and risk of predicting production and/or injection in the field. In this study, porosity and permeability measured from cores were correlated with fractal dimension of corresponding deep resistivity logs. A methodology to determine fractal dimension from 1D data is proposed and the propagation of ions and electric current in the rock fabric along with its relationship and effect to fractal dimension of the resistivity logs is investigated. Traditionally, only the change in average of resistivity log over an interval is utilized when interpreting resistivity logs. The results of this investigation show that further information can be derived from these logs. For example, presence of layered beds with thicknesses less than the resolution of the tool can be flagged using is effect of these beds on fractal dimension of the resistivity log. While this is not possible using conventional interpretation methods. This study is presented in detail in chapter 3.

Next, a Machine learning technique in facies identification is investigated in the paper titled "A transfer learning approach for facies prediction using resistivity image well logs". Resistivity image logs provide a high-resolution, 2-dimensional image of the borehole inner wall. These logs contribute significantly to the characterization and understanding of structural interpretation of the log interval. Interpretation of these logs is quite time-consuming complex and subject to the interpreter's experience specially in cases where the geological setting is complex. There is a great need for an automated workflow to augment and optimise this manual process. In this study a methodology to train a convolutional neural network on already interpreted logs in one well is presented and the trained model is used to rapidly detect the interpreted facies in a newly drilled well. A comparison of different network architectures is carried out and their confusion matrix is interpreted. Prominent challenges when applying convolutional neural networks to resistivity image logs were discussed and solutions to these challenges are suggested. This investigation showed that convolutional neural networks are able to detect details in the resistivity image logs that can be used to distinguish and classify the facies. The accuracy at which this is achieved using image logs only was quantified. Although the investigation concluded that other supplementary logs are needed for facies classification, it also showed that resistivity image logs contribute significantly to the task. Two novel pre and post processing methods were presented as well. The data preprocessing methodology maximizes the training images extracted from logs. While the post processing method enhances the accuracy and bed boundary detection resolution even further. This study is delivered in chapter 4.

Lack of data is an intrinsic issue that hinders accurate characterization of reservoirs. Often sparse data from biased locations in the field are extrapolated to extensive areas. To enhance

this process extra information can be derived from a conceptual image contributed by the geologist known as the Training Image (TI). Furthermore, Correct characterization and use of this data requires adopting methodologies that quantify and propagate uncertainty in the process. In the paper titled "A variability aware GAN for improving spatial representativeness of discrete Geobodies", a Novel Generative Adversarial Network (GAN) architecture is proposed to generate model realisations that are geologically sound (realistic) and statistically faithful. GANs are a family of deep generative models (DGM) that have shown great potential in generating realizations of geological structures. This class of Neural Networks (NN), Can approximate high dimensional probability distributions from sample data. In this study, we evaluate the Ability of GANs in generating geological realisations that are (1) visually acceptable (2) preserve the statistics of the training image and (3) maintain the variability of the structure. Analysis of Distance (ANODI) is applied to quantify how well multiple point statistic of the realizations are preserved and probability maps are utilized to quantify the variability of the realizations. The effect of input data on variability is investigated by evaluating the proposed network on two different data sets. The proposed architecture outperformed other popular networks and significantly enhanced variability and reduced spatial bias wile better preserving multiple point statistics of the realisations. A sensitivity analysis on the key parameters of the proposed architecture was also conducted. As the application of generative models increases in generating geological realizations this study presents a universal methodology to maintain the variability of the output realizations. The investigation and findings of this study constitute chapter 5.

Characterization of a reservoir model by tuning to one dimensional production data is referred to as History matching. This type of characterization is generally a nonlinear problem that does not have a closed-form solution. Characterization of numerical models – which integrates

numerous data sources with different dimensions and scales - by history matching is often an optimization problem with many parameters. As such, it is computationally expensive and time consuming. Advanced analytics and machine learning techniques can be adapted to enhance computation efficiency of history matching problems. Since this study is not directly a characterization technique, it has been attached as an appendix to this thesis. In the paper titled "Accelerating CMA-ES In History Matching Problems Using an Ensemble of Surrogates with Generation-Based Management" a surrogate-assisted Covariance matrix adaptation evolutionary strategy (CMA-ES) is proposed that reliably accelerates history matching and significantly reduces computation. The algorithm was tested on two simulation cases. The effectiveness of an ensemble of surrogates to breakdown the complexity of the fitness function was investigated. Furthermore, an online learning scheme to continuously improve the fidelity of the proxies over the history matching process is presented. Other techniques such as generation-based model management and evolution control were applied. The details and results are presented in the Appendix.

2. Literature Review

Reservoir characterization is a combination of methodologies associated with geostatistics, geophysics, petrophysics, geology and reservoir engineering (Jia et al., 2012). The main goals of reservoir characterization research are to aid field development and reservoir management teams in describing the reservoir in sufficient detail and developing 3D/4D data for reservoir development planning. Equipped with this information, higher recoveries with fewer wells in better positions at minimum cost can be obtained through optimization, increasing reserves, improving stimulation and completion practices and reducing to a minimum uncertainty, in production forecasts (Haldorsen and Damsleth, 1993, Johnston, 2004, Phillips, 1996). Large

quantities of data in different scales are collected during the life of a well, which are subsequently utilised for reservoir characterization.

Advanced analytics and machine learning play a key role in deciphering these data. In this work fractal analysis, generative adversarial networks and convolutional neural networks have been utilized to enhance workflows in reservoir characterization.

2.1 Fractal Analysis

Since the introduction of Fractal analysis by Benoit Mandelbrot, many fields of science have been revolutionized by this concept, and subsurface characterization is no exception. While statistical methods continue to be useful, they ignore the fundamental concept that geological structures are not developed through random processes, but rather deterministic causes (Hardy et al., 1996).

2.1.1 Fractal character of wireline logs

Hewett in 1986 proved that porosity logs were fractal (Hewett, 1986). Using this the fractal character Hewett generated porosity distributions that enhanced fluid flow simulation. A similar approach was utilised by other authors (Aasum and Kelkar, 1991, Berta et al., 1994, Crane and Tubman, 1990, Emanuel et al., 1988, Hardy, 1992, Hewett and Behrens, 1990, Lozada-Zumaeta et al., 2012, Perez and Chopra, 1997). Avnir et al. (1985) showed that the pore structure of rocks has fractal character. Others confirmed this finding in sandstones as well (Katz and Thompson, 1985, Krohn and Thompson, 1986). Methods and techniques to calculate the fractal dimension of these structures was also presented in these studies. Pang and North (1996) in their work suggested that fractal character of well logs is related to the stratigraphic heterogeneity. Their results showed that more heterogenous rocks produced logs with higher fractal dimension. Shen et al. (1998) suggested that the fractal dimension of pore structure can

be used to classify rock type. They investigated 22 cores and concluded that the fractal dimension can be related to oil recovery at water breakthrough and irreducible water. Wang and Mou (2014) measured the fractal dimension of various wireline log data, including compensated neutron, density, gamma ray and acoustic logs for 108 wells, and found corresponding relationships between fractal dimension of logs and texture of volcanic rocks.

2.2 Convolutional neural networks

Convolutional Neural Networks (ConvNets or CNNs) are rather well known for their applications in classification and computer vision tasks. The main components of these networks are the convolutional layer, pooling layer, nonlinear activation layer and the fully connected layer. Among these component the most significant is the convolutional layer (Alzubaidi et al., 2021). These layers not only capture the spatial dependencies of the image, but also use sparce connections and parameter sharing to avoid parameter size being too large (Chen et al., 2021). The role of the pooling layer is to subsample the feature maps to smaller ones (Gu et al., 2018) and make the feature maps more robust to single neuron errors (Liu et al., 2017). The nonlinear activation layer introduces nonlinearity to the network which significantly enhances the network performance (Gu et al., 2018). The fully connected layer is usually located at the end of the network and is the classifier section and categorizes the extracted local information (Sainath et al., 2013).

2.2.1 Development of Convolutional Neural Networks

The history of deep CNN's began with the appearance of LeNet Introduced by (LeCun et al., 1995). AlexNet was developed and built based on LeNet-5 (Krizhevsky et al., 2017). With the appearance of AlexNet for the first time learned features surpassed traditional manual feature extraction. In 2014, an innovative CNN design with a modularized network was introduced This

design was called the Visual Geometry Group(VGG) (Simonyan and Zisserman, 2014). The main drawback of this network was utilizing a large number of parameters which resulted in high computational cost. Prior to the advent of the inception network, the simplest way to enhance CNN performance was to increase their size. This caused two major issues, it made the models prone to overfitting and dramatically increased the computational cost. The Inception network (Szegedy et al., 2015) was designed to address the above issues. The inception module computes 1X1, 3X3 and 5X5 convolutions within the same module, therefore it acts as a multi-scale feature extractor. The output of these filters are then stacked along the channel dimension and fed to the next layer in the network (Zaccone et al., 2017). ResNet was introduced to address network degradation. With network depth increasing, accuracy is saturated and then degrades rapidly (He et al., 2016) this is referred to as "Network Degradation". The residual connection in ResNet is a method to break "degradation" and enable deep neural networks to achieve high accuracy (Orhan and Pitkow, 2017, Chen et al., 2021).

2.2.2 Transfer learning

Although machine learning has shown great success in different fields of science, it still has limitations. The ideal scenario for machine learning is when labelled data is abundant. However, this is rarely the case in real-word scenarios (Zhuang et al., 2020). In reality, and more so in subsurface engineering labelled data is scarce. Furthermore, large networks required for learning features of complex datasets are computationally expensive to train. Transfer learning has proven to be a very effective technique to handle the above issues in the field of image recognition and image classification (Tang, 2018). In deep convolutional neural networks, with many layers, the lower layers detect the features while the final layers detect the class of the

image (Tang, 2018). In the case of image classification, transfer learning allows us to train the classifier layers of the CNN with our specific set of images requiring far less training images and computation power (Shao et al., 2014).

2.2.3 Application of CNN to image logs

High-resolution borehole image logs, provide detailed information on lithology, sedimentary textures, paleo-flow directions, structural dip analysis, in situ stress analysis and fracture evaluation (Lai et al., 2018, Nie et al., 2013, Kosari et al., 2015, Brekke et al., 2017, Ameen, 2014). These logs contribute significantly to the geological understanding of the logged interval (Lai et al., 2018, Folkestad et al., 2012) and allow for structural features to be identified at resolutions of only a few millimetres (Ja'fari et al., 2012). Machine learning and more specifically computer vision techniques can be used to automate or augment facies classification at well locations. Gupta et al. (2019) used a UNET architecture to pick induced and natural fractures along with sedimentary surfaces from image logs. Lima et al. (2019) proposed an unsupervised neural network model for pattern recognition and facies categorization using borehole image logs. They used a non-linear autoencoder for representational learning to reconstruct the original training images and applied cluster analysis for categorization of the facies. Using the concept of transfer learning Lefranc et al. (2021), trained a ResNet architecture on synthetic training data produced by models. However, their method is still quite laborious since it requires sedimentary dips to be picked manually at high density.

2.3 Generative adversarial networks

Creating a geological model requires detailed characterization of the reservoir. However geological models are unable to fully describe the reservoir due to some parameters being

uncertain or unknown. This uncertainty in geological models must be adequately quantified and considered in the applications of the model. understanding uncertainty in reservoir models involves generating multiple realisations where governing equations are then solved for each realisation to yield an understanding of the probability distribution of the model response (Chan and Elsheikh, 2020). There has always been a desire to generate geological models that not only honour the hard data but are visually acceptable and statistically faithful. To this aim, different methodologies have evolved over time. Starting with stochastic methods followed by Multiple point statistic methods that produced superior more realistic realizations. Recently Deep Generative Models have been introduced and show great potential in this field.

2.3.1 Stochastic methods for generating realizations

Traditionally, stochastic simulation methods are the most popular methods for generating realisations. Object-based methods and pixel-based methods are the two main subcategories of stochastic simulation methods (Bai and Tahmasebi, 2020). The first category uses Boolean object-based algorithms to characterise the simulated area by placing objects that resemble geological features. Models generated using these algorithms are significantly more realistic from a geological perspective. However, Conditioning to hard data is quite challenging in these algorithms (Bai and Tahmasebi, 2020, Strebelle, 2002). Pixel-based Algorithms, Conduct the simulation pixel by pixel. These algorithms do not pose this challenge and are relatively easier to condition to hard data. examples of pixel-based algorithms are sequential gaussian simulation and truncated gaussian simulation. The challenge with these algorithms is that they are based on two point statistics which is inadequate to reproduce complex geological features (Tahmasebi, 2018) such as curvilinear geometries (e.g. Sinuous channels) which are inherent in many geological structures (Marini et al., 2018). As a result, when such geological systems

are being simulated using these algorithms, they appear geologically unrealistic. Other pixel-based methods emerged in the literature such as Truncated pluri-gaussian (Le Loc'h et al.) to overcome some of these limitations and better preserved prior geological understanding (Astrakova and Oliver, 2015). The main difficulty with applying this method is the inference of the variogram models for the underlying multi gaussian functions (Mariethoz et al., 2009). Multiple point Statistics (MPS) methods were introduced to address the problem of generating realistic models of complex geological structures. These methods require large number of samples which is generally not available in earth sciences (Marini et al., 2018). The extra information can be provided by the geologist. The geologist's insight and understanding of the targeted region is conceptualised in an image known as the Training Image (TI) (Meerschman et al., 2013). The TI is designed under expert knowledge to replicate expected patterns and statistical features of the subsurface. The MPS methods then generate realisations that resemble this training image and honour any other available data. In this scenario the TI plays and important role and guides the outcome. An example of such technique is the Single Normal Equation Simulation (SNESIM). This algorithm scans the TI and stores the probability of all pattern occurrences in a search tree. The probabilities are then retrieved based on existing data to generate realisations. The challenge with these methods is that inversion using these methods is computationally expensive. There are re-parameterisation techniques available however the models provided using these techniques do not agree well with the TI.

2.3.2 Deep Generative methods

A more recent technique for generating realisations of geological models in the literature is Deep Generative Models (DGM). DGMs are a class of Neural Networks (NN) that can approximate high dimensional probability distributions when trained on sufficient samples

from the desired distribution. The trained model can then be used to generate realisations from the underlying distribution (Ruthotto and Haber, 2021). A further advantage of DGM's is that they re-parameterise the realisations by mapping the data to a latent space. Variational Auto-Encoders (VAE) are an example of generative models that use variational Bayesian inference to approximate the probability density function of a dataset of samples. They generally consist of an encoder and a decoder. The encoder takes in sample training data and reduces it to a latent space by passing it through layers with decreasing dimensionality. The decoder then samples from the latent space and reproduces the initial sample (Sami and Mobin, 2019). Laloy et al. (2017) Showed that inversion using VAEs produces superior results compared to MPS-based inversion methods. Canchumuni et al. (2021) Combined a VAE with an Ensemble Smoother with Multiple Data Assimilation to history match production data and reported that trained VAE resulted in noisy facies reconstruction. In general VAEs have a lower generative accuracy compared to Generative Adversarial Networks (Lopez-Alvis et al., 2021) and are prone to fail at learning intractable or highly complex probability distributions (Sami and Mobin, 2019). This is because, VAEs have a simpler architecture with only one loss function (Kullback-Leibler divergence). As described above VAEs produce their output by compressing the input to a latent space. Generative Adversarial Networks (GAN) on the other hand, search for a balance point between the Discriminator and Generator in their two-player game, where one tried to trick the other. As such, GANs have a more complex architecture and use two loss functions.

2.3.3 Generative adversarial networks in subsurface applications

Generative adversarial networks (GAN) are a family of deep-learning-based generative models where the paradigm of unsupervised learning is used in their training process. The GAN

architecture has two sub-models: Generator and Discriminator. Generator is used to generate new plausible examples from the problem domain. The Discriminator is the sub-model that is used for training the Generator. The architecture has the advantage that after being trained, the sub-models can be used as standalone models for data generation or classification. The training process of a GAN is based on a game theoretic scenario in which the generator competes against an adversary. The generator network directly produces samples from a fixed-length random vector, referred to as latent space, while the discriminator network, the adversary, attempts to distinguish between samples drawn from the training data (real data), and the generated samples.

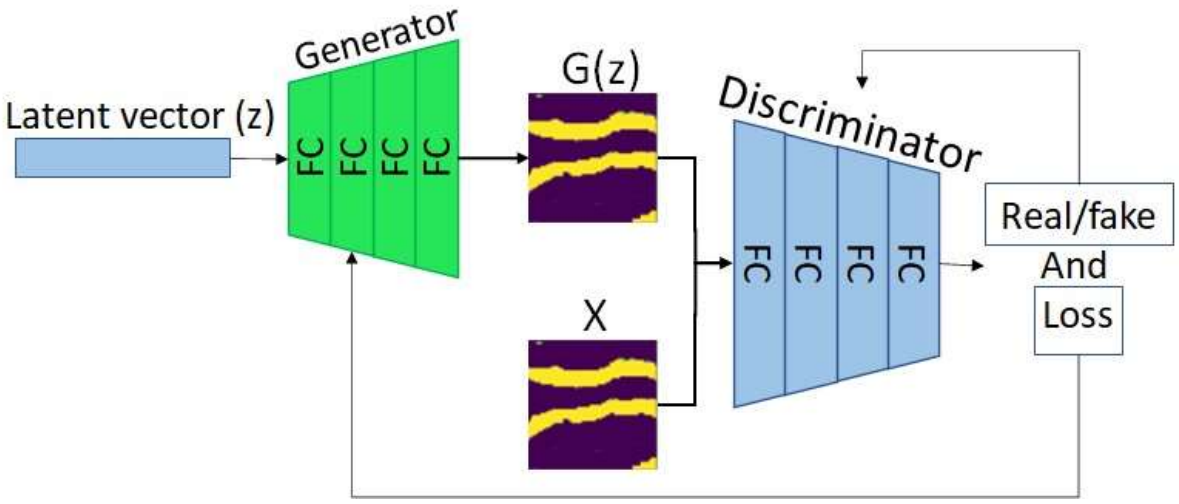

Figure 1: Vanilla GAN workflow

Figure 1 illustrates the basic workflow of a vanilla GAN. Based on the figure, latent vector z, is fed to the Generator G to generate sample $G(z)$. FC stands for Fully Connected layers. It is conventional to feed a batch of latent vectors to generate a batch of samples every time. The Discriminator is trained to classify $G(z)$ and real samples, X. This is accomplished by maximizing the probability assigned to real and generated data at the output of the discriminator.

Among DGMs, Generative Adversarial Networks (GAN) introduced by Goodfellow et al. (2014) have been applied to a variety of subsurface applications. GANs have been used to generate realistic stochastic samples of porous media (Mosser et al., 2017, Mosser et al., 2018b). Mosser et al. (2018b) Trained a modified version of the DCGAN (Radford et al., 2015) on randomly extracted images from a micro-CT. They reported that computed two-point statistics an effective property, showed excellent agreement between the GAN results and segments of the micro-CT image. However, the generated samples showed less variation compared to the input training dataset. Mode collapse was mentioned as one of possible reasons. Mode collapse is defined as the case whereby the Generator can only generate one type of sample or a small set of distinct samples. This is caused when the Generator learns to generate samples from few modes of the data distribution but ignores other modes although they are present in the input training data. This effect when generating geological realizations, will result in the Generator producing multiple samples/realizations with one or more features replicated in common locations. The realizations overall are therefore spatially biased towards those features. There are a variety of reasons for this phenomenon e.g., vanishing gradients, Discriminator overfitting, poor network architecture, etc. some solutions to reduce this effect include using Wasserstein loss or progressively increasing resolution of generated images during training. In chapter 5 of this thesis a solution suited to geological realizations is presented.

Other subsurface applications of GANs includes application to seismic data. A Cycle-GAN was used to perform stratigraphic seismic inversion based on a velocity model (Mosser et al., 2018a). In Mosser et al. (2020), a DCGAN in combination with a numerical solution of the acoustic inverse problem was trained to parametrize geological heterogeneities. Mode collapse and its adverse effects on the study were discussed and use of alternative networks

that handle mode collapse more efficiently were suggested. GANs were also used to reconstruct models retrieved by iterative geostatistical seismic inversion (Azevedo et al., 2020). In this study, two GAN networks DCGAN and WGAN (Arjovsky et al., 2017) were trained on two datasets, a set of binary facies and a continuous (P-wave propagation velocity) dataset where an infill painting methodology (Yeh et al., 2017) was used for reconstruction of images. Laloy et al. (2018) used a GAN architecture to perform MPS based geostatistical inversion for parameter estimation. They proposed a 3D extension to the original 2D spatial GAN (SGAN) (Jetchev et al., 2016) and incorporated it in a Markov chain Monte Carlo (McMc). Other examples of applications of GAN in the literature are re-parametrization (Chan and Elsheikh, 2019a, Chan and Elsheikh, 2020). GANs have shown great potential in generating realistic realizations of geological models. Multiple studies in the literature have used GANs to generate realizations conditioned to hard data (Chan and Elsheikh, 2019b, Dupont et al., 2018, Mosser et al., 2018a, Zhang et al., 2021). Different methodologies have been used to condition geological models to hard data. Some used semantic inpainting methodology (Yeh et al., 2017) to condition the model (Dupont et al., 2018). In this methodology, a GAN is trained to generate geological realizations. The loss of the GAN during training is defined such that the Generator outputs the closest realization in the latent space manifold whereby this realization honours the hard data. Others like (Chan and Elsheikh, 2019b) proposed an inference network to be added to the Generator to further refine the generator function space to produce only conditioned realizations. Zhang et al. (2021) used a U-net architecture to generate conditioned realizations. In this study the authors adopted a methodology where a loss term was formulated to maximize the distance between generated images with respect to corresponding latent vectors (Yang et al., 2019). Razak and Jafarpour (2020a, 2020b) and

Mosser et al. (2019) have gone a step further in their work by calibrating reservoir models using non-linear production data.

Although GANs promise vast potential in generating realistic geological realizations, Mode collapse is an important issue to thoroughly investigate when using GANs. Mode collapse can induce significant error in uncertainty quantification of a model. Furthermore, conditioning to hard data is not a trivial task when using GANs. The available methods have the potential to reduce variability of the generated realizations. Currently research in this field is focused on rectifying the above-mentioned limitations of GANs. It is expected that more applications and usage of the reparameterization capability of GANS will be introduced.

Finally, to close this section, pros and cons of the mentioned geostatistical methods for generating geological realizations are presented in Table 2.

Table 2: pros and Cons of different methods for generating geological realizations.

Methodology	Variation	pros	cons
Stochastic methods	Pixel-based	• Easy to implement with little computation cost. • Easy to condition to hard data	• Cannot produce visually acceptable and realistic realizations. • Requires inference of the data variogram
	Object-based	• Realizations are more realistic compared to pixel-based methods	• difficult to condition to hard data
Multiple point statistics methods	N/A	• visually realistic realizations • algorithm uses TI which can incorporate geologists' expertise	• algorithms are computationally and memory intensive. • do not re-parametrize the realization
Deep Generative methods	VAEs	• easy to train. • results are visually acceptable	• fail at learning highly intractable distributions.
	GANs	• results are visually acceptable. • Can learn complex distributions	• results affected by Mode collapse. • more research required around hard data conditioning

3. Rock typing and facies identification using fractal theory and conventional petrophysical logs

Roozbeh Koochak, Manouchehr Haghighi, Mohammad Sayyafzadeh, Mark Bunch

APPEA journal, 2018

Statement of Authorship

Title of Paper	**Rock typing and facies identification using fractal theory and conventional petrophysical logs**
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	**Koochak, R., Haghighi, M., Sayyafzadeh, M. and Bunch, M., 2018. Rock typing and facies identification using fractal theory and conventional petrophysical logs. The APPEA Journal, 58(1), pp.102-111.**

Principal Author

Name of Principal Author (Candidate)	**Roozbeh Koochak**		
Contribution to the Paper	**Conceptualization, development of algorithm, experimentation, analysis, coding, original manuscript, corresponding author duties**		
Overall percentage (%)	**70%**		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	**11/05/2023**

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	**Manouchehr Haghighi**		
Contribution to the Paper	**Conceptualization, original idea,Analysis of results, discussions, revisions, supervision**		
Signature		Date	18/05/2023

Name of Co-Author	**Mohammad Sayyafzadeh**		
Contribution to the Paper	**Discussions, revisions, supervision**		
Signature		Date	**18/05/2023**

Name of Co-Author	**Mark Bunch**		
Contribution to the Paper	**Discussions, revisions, supervision**		
Signature		Date	19/05/2023

Please cut and paste additional co-author panels here as required.

The APPEA Journal 2018, **58**, 102–111
https://doi.org/10.1071/AJ17087

Peer reviewed paper

Rock typing and facies identification using fractal theory and conventional petrophysical logs

Roozbeh Koochak[A,B], *Manouchehr Haghighi*[A], *Mohammad Sayyafzadeh*[A] and *Mark Bunch*[A]

[A]Australian School of Petroleum, Santos Petroleum Engineering Building, University of Adelaide, SA 5005, Australia.
[B]Corresponding author. Email: Roozbeh.koochak@adelaide.edu.au

Abstract. Rock typing or subdivision of a reservoir either vertically or laterally is an important task in reservoir characterisation and production prediction. Different depositional environments and diagenetic effects create rocks with different grain size distribution and grain sorting. Rock typing and zonation is usually made by analysing log data and core data (mercury injection capillary pressure and permeability measurement). In this paper, we introduce a new technique (approach) for rock typing using fractal theory in which resistivity logs are the only required data.

Since resistivity logs are sensitive to rock texture, in this study, deep conventional resistivity logs are used from eight different wells. Fractal theory is applied to our log data to seek any meaningful relationship between the variability of resistivity logs and complexity of rock fabric. Fractal theory has been previously used in many stochastic processes which have common features on multiple scales. The fractal property of a system is usually characterised by a fractal dimension. Therefore, the fractal dimension of all the resistivity logs is obtained.

The results of our case studies in the Cooper Basin of Australia show that the fractal dimension of resistivity logs increases from 1.14 to 1.29 for clean to shaly sand respectively, indicating that the fractal dimension increases with complexity of rock texture. The fractal dimension of resistivity logs is indicative of the complexity of pore fabric, and therefore can be used to define rock types.

Keywords: Cooper Basin, fractal geometry, Higuchi's fractal dimension method, resistivity well logs, rock typing.

Received 14 December 2017, accepted 15 February 2018, published online 28 May 2018

Introduction

Flow prediction of reservoirs requires identification and modelling of rock types. This results in finding the optimum location of well placement for production from the field. Rock typing is an integral part of reservoir characterisation which identifies different flow units (Gupta *et al*. 2017). The first step in trying to predict the behaviour of a reservoir system is characterisation of reservoir rock types to form a static reservoir model. Rock, fluid and rock–fluid properties are then attributed to up-scaled static model cells.

More accurate prediction of flow requires detailed knowledge of the heterogeneity of the reservoirs. The performance of a reservoir is controlled by intrinsic properties of fluids and the geometry of the pore system. The effective description of the reservoir requires an adequate understanding of petrophysical properties of the rock, such as porosity, permeability, capillary pressure, heterogeneity and fluid content (Porras *et al*. 1999).

Knowledge of these properties facilitates the subdivision of the reservoir into layers with similar properties according to flow point. This subdivision enhances flow behaviour modelling and reduces uncertainty in predicting production. This process is referred to as rock typing. There are many different definitions for 'rock type' in the literature; the most adopted definition is by Gunter *et al*. (1997): 'Rock typing is a method of classifying reservoir rocks into distinct units, each of which was deposited under similar geological conditions and has undergone similar diagenetic alterations'.

This definition includes depositional features along with diagenetic effects in defining a rock type. This implicitly suggests that the pore structure of a rock – i.e. pore and pore throat dimensions, geometry, size, distribution and capillary pressures – should be similar within a rock type. It has been long established in the oil and gas literature that these parameters are the main drivers of fluid flow in porous media.

In conventional reservoirs, flow unit rock typing is usually done based on porosity–permeability relations. There are numerous rock typing studies in the literature. The common evaluation techniques are:

- Physical core description of large- and small-scale features, along with core measurements of porosity and permeability; dominant pore throat diameter from mercury injection capillary pressure (MICP) data.
- Texture, composition and lithology of the rock, along with considering the deposition environment of the reservoir.
- Identification of lithofacies from log analysis complemented with core-based measurements.

Other methods include use of R35 measurements from mercury injection, proposed by Pittman (1992), in which average pore throat radius is measured from the 35% injection level of mercury. The concept of rock quality index and flow zone indicators was introduced by Amaefule *et al*. (1993). In tight sands, however, use of all the above techniques needed to be complemented with rock texture and fabric descriptions to achieve more accurate results (Rushing *et al*. 2008).

One way to infer rock types is by use of well logs. In this method, statistical features are defined for the trend of the corresponding well log, and these features are used to identify the same rock type in other wells. However, since conventional interpretations of logs do not provide any direct information on complex pore geometries, it is difficult to consider pore structure in rock typing methods using well logs.

Fractal geometry offers a new approach for interpreting well logs. Use of fractal theory in oil and gas has a significant history. In 1986, Hewett proved that porosity logs were fractal (Hewett 1986). Using the fractal character, Hewett generated porosity distributions and used them to enhance fluid flow simulation. There were other similar studies (Crane and Tubman 1990; Emanuel *et al*. 1990; Hewett and Behrens 1990; Aasum and Kelkar 1991; Hardy 1992; Hardy and Beier 1994; Perez and Chopra 1997; Lozada-Zumaeta *et al*. 2012). Some authors found that the pore structure of rocks was a fractal property (Avnir *et al*. 1985). Others investigated different sandstones and confirmed their fractal character (Katz and Thompson 1985; Krohn and Thompson 1986). These workers also suggested methods to determine the fractal dimension of these structures. Pang and North (1996) suggested that the fractal dimension of well logs is related to stratigraphic heterogeneity; more heterogeneous rock produced well logs with a larger fractal dimension. Shen *et al*. (1998) suggested that the fractal dimension of pore structure can be used to classify rock type. They investigated 22 cores and concluded that the fractal dimension can be related to oil recovery at water breakthrough and irreducible water. Wang and Mou (2014) measured the fractal dimension of various wireline log data, including compensated neutron, density, gamma ray and acoustic logs for 108 wells, and found corresponding relationships between fractal dimension of logs and texture of volcanic rocks.

In most studies, the fractal dimension of the pore structure is derived from microscopic core images and images of thin sections using image processing techniques. Furthermore, no study has considered the fractal dimension of resistivity well logs and its relationship to rock texture. Considering the effect of rock texture on the resistivity of the rock, along with the fact that pore structure of rocks is fractal, makes it logical to expect that the fractal dimension of resistivity logs is related to rock texture.

The objective of this study is to show that fractal dimension of resistivity logs can be related to rock texture, and thus can be used as a new method to identify rock types. First, a background of fractal theory and resistivity of porous media is presented, followed by details of our proposed method. Then, data and results are interpreted and discussed.

Theory background

In this section, a background in fractal theory and some facts on resistivity in porous media are presented.

Fractal theory background

A fractal is a geometric pattern that shows similarity to itself at any level of magnification (scale). The pattern that causes the self-similarity can be repeated at multiple scales to produce irregular shapes and surfaces that cannot be modelled by conventional geometry. This is the main feature that differentiates fractal geometry from Euclidian geometry. The concept of fractals was introduced by Mandelbrot and has been shown to be capable of mathematically modelling irregular natural patterns.

Before the introduction of fractal geometry, mathematicians had come across many patterns and shapes that could not be modelled by Euclidian geometry. A turning point came when Benoit Mandelbrot introduced a more comprehensive definition of dimension. He stated that the dimension of a fractal must be used as an exponent when measuring its size (Mandelbrot 1983). As a result, fractals cannot be described with integer dimensions but require fractional dimension, hence the name fractal geometry.

Mandelbrot defined fractal as ‘a shape made of parts that are similar to, or repeat the whole in some way’ (Mandelbrot 1983). This is a definition of self-similar fractals. This type of fractal is too regular to model natural phenomena. Self-affine fractals, however, are defined as objects that are statistically similar to themselves. For example, a fern leaf would look similar to itself at different scales but would not be identical. Self-affine fractals are mainly used to model time–depth sequences and spatial distributions. This property that objects can look statistically self-similar while also exhibiting some variability in detail at different length scales is the central feature of fractals in nature (Feder 1988). Fractals have been used in many areas of natural sciences, e.g. river networks, fault lines, mountain ranges, coastlines, thickness of tree trunks, heart rates and earthquakes.

The scale invariance of fractals is mathematically described by a power law. A power law distribution is the only statistical distribution that is scale invariant.

One of the best methods for mathematical modelling of self-affine fractals is fractional Brownian motion (fBm). In terms of a function of time, fBm is defined as:

$$B_H(t) = B_H(0) + \frac{1}{\Gamma\left(H+\frac{1}{2}\right)} \left\{ \int_{-\infty}^{0} \left[(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right] dB(s) + \int_{0}^{t} (t-s)^{H-\frac{1}{2}} dB(s) \right\} \quad (1)$$

The fBm is a self-similar, stationary process with long range interdependence, and has a covariance in the form of:

$$E\{B_H(t)B_H(\tau)\} = 0.5\sigma^2\left(|t|^{2H} + |\tau|^{2H} - |t-\tau|^{2H}\right) \quad (2)$$

where $b_H(t)$ is an fBm as a function of time, σ is the standard deviation and H is commonly known as the Hurst exponent and characterises the scaling behaviour of the series; H ranges between 0 and 1. When H is close to zero the time series are rough with strong variation, and when H is close to 1 they are smooth with less variation.

The H is related to the fractal dimension of a one-dimensional time series by:

$$D = 2 - H \quad (3)$$

where D is the fractal dimension.

The H also quantifies the persistence and anti-persistence trend in the time series. Persistence ($H>0.5$) means that an increasing trend is likely to be followed by an increasing trend; and anti-persistence ($H<0.5$) means that an increasing trend is likely to be followed by a decreasing trend. Fractal statistics provide a simple way of relating variations at larger scales to those at smaller scales and vice versa (Gray *et al*. 1993).

Equation 3 is one method of calculating the fractal dimension, where H can be derived using the rescaled range method, considering the time series is best modelled as fBm. Other popular techniques are used to test for fractal scaling and determination of fractal dimension of wireline logs (or any time series): variogram and spectral techniques, box-counting method, Katz, Sevcik and Higuchi methods.

The three commonly used techniques for estimating fractal dimension (i.e. rescaled range, variogram and spectral techniques) are not reliable when the underlying signal is either limited, non-Gaussian or non-stationary (Gray *et al*. 1993). These methods are severely affected by small number of samples and non-stationarity, therefore they are not suitable for analysing logging data. This is especially true for fluvial facies of the Cooper Basin where production intervals are relatively small. The box-counting method is highly sensitive to sampling frequency. In addition, in the case of wireline logs the data axes are incompatible, in other words the x (resistivity) and y (depth) axes cannot be compared with each other. Raghavendra and Dutt (2010) developed a variation of the box-counting method called the Multiresolution Box-counting Method (MBCM) that resolved this issue, but the MBCM requires a high sampling rate, making it unsuitable for analysing well logs. The same study showed that the Katz and Sevcik methods yielded poor results compared with MBCM and the Higuchi method.

Among the mentioned methods, Higuchi's algorithm is least affected by the sample number of the signal and is most suitable for analysing the fractal dimension of short-interval signals. For this reason, we used Higuchi's method to calculate the fractal dimension of resistivity logs.

Higuchi's method

Higuchi's algorithm for calculating the fractal dimension of a one-dimensional signal is described below (Higuchi 1988):

Consider a finite, discrete set of samples taken at regular interval:

$$x(1), x(2), x(3), \ldots, x(N) \quad (4)$$

where N is the total number of samples. From this given time-series new subseries denoted by x_m^k are generated.

$$x_m^k = \left\{ x(m), x(m+k), x(m+2k), \ldots, x\left(m + \left[\frac{N-m}{k}\right].k\right) \right\} \quad (m = 1, 2, \ldots, k) \quad (5)$$

where k is the scaling factor usually chosen based on the total number of samples. Both k and m are integer values, and [a] denotes the closest integer to a. For example, for $k=3$, and total sample number of $N=100$, the subseries are:

$$x_3^1; x(1), x(4), x(7), x(10), \ldots, x(97), x(100)$$

$$x_3^2; x(2), x(5), x(8), x(11), \ldots, x(95), x(98)$$

$$x_3^3; x(3), x(6), x(9), x(12), \ldots, x(96), x(99)$$

For each subseries x_m^k a corresponding length is defined as:

$$L_m(k) = \frac{1}{k}\left\{ \frac{N-1}{\left[\frac{N-m}{k}\right].k} \left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |x(m+ik) - x(m+(i-1).k)| \right) \right\} \quad (6)$$

where $\frac{N-1}{\left[\frac{N-m}{k}\right].k}$, is a normalisation factor for the x_m^k subseries. The above formula will result in k number of lengths for each x_m^k subseries. The length for k is defined as $L(k)$ and is computed as the average value of the k sets of $L_m(k)$. That is:

$$L(k) = \sum_{m=1}^{k} L_m(k) \quad (7)$$

If $L(k)$ is proportional to k^{-D}, then the initial time series is fractal with dimension D.

If $L(k)$ is plotted against k on a double logarithmic plot, the points will fall on a straight line, the slope of which is fractaldimension (D) of the time series.

Resistivity of porous media

Resistivity logs are among the very first logs to be recorded. The principal use of resistivity logs is to locate hydrocarbons; however, they can also provide information on lithology, texture, facies overpressure and source rock aspects (Rider 1986).

The resistivity of the rock matrix is usually assumed to be infinite, thus the resistivity is presumed to be a function of the pore fluid alone. This is not entirely true as the matrix plays a passive role in resistivity of the rock. This passive role is dependent on the geometry of the pores, pore connections, tortuosity and pore size distribution. Studies have shown that resistivity of fluid-filled sedimentary rocks is mostly controlled by pore structure (Archie 1942; Bigalke 2000). Wettability, saturation history and temperature also play important roles (Swanson 1985; Kumar *et al*. 2010). Pore-structure characteristics can be estimated from electrical resistivity logs and used for a better estimation of permeability (Verwer *et al*. 2011).

Any clays present play an active role in conduction of the rock. Clays conduct electricity in two ways: through pore water and through clay itself (Rider 1986). The resistivity of reservoir rocks depends on numerous factors which can collectively be described as rock fabric. Rock fabric describes the spatial and geometric configuration of the rock, and can be represented by numerical engineering values complemented with geological descriptions of the rock.

Katz and Thompson (1985) proposed that the pore spaces of sandstones are fractal and presented several measurements to support their proposal. The electrical conductivity in porous media obeys a scaling law (Toledo *et al*. 1994), which suggests that resistivity well logs have fractal character:

$$\sigma_w \propto S_w^{\frac{1}{m(3-D)}} \tag{8}$$

where, σ_w is the electrical conductivity of the porous medium, S_w is water saturation and m is Archie's cementation factor.

We propose that the fabric of a rock type affects the variability of resistivity logs, and this can be uniquely characterised by the fractal dimension of the resistivity logs.

Data

The Cooper Basin is the most prospective onshore petroleum and natural gas province in Australia. It is a sedimentary basin that formed and developed from the late Carboniferous to middle Triassic geologic periods. It unconformably overlies the Warburton Basin and unconformably underlies the Cretaceous Eromanga Basin (Gravestock and Jensen-Schmidt 1998). The basin is located across the north-east of South Australia and the south-west of Queensland. The reservoir system comprises a multi-zone high-sinuosity fluvial sandstone ranging from tight (unconventional) to good-quality conventional reservoir rocks. There is a large range of porosity and permeability in the basin due to combination of facies and burial depth. The main gas reservoir is within the Patchawarra Formation with an average porosity of 10.5% and permeability up to 2500 mD, and the Toolachee Formation with an average porosity of 12.4% and permeability up to 1995 mD (Gravestock *et al*. 1998).

Oil is produced from the low-sinuosity fluvial sand within the Tirrawarra Sandstone with average porosity of 11.1% and permeability up to 329 mD (Gravestock *et al*. 1998).

In this paper, 59 core samples from different locations (laterally) in the Cooper Basin were chosen for analysis. The data was quality controlled and cores with missing or low-quality data such as noisy or low sample-rate logs, missing intervals or failed MICP tests were excluded. Eventually eight cores, with MICP data and a full suite of logs, were analysed. The pore size distribution of the core samples were derived using MICP. Porosity and permeability of the cores were derived using routine core analysis.

Work flow

In this study, we aimed to investigate the relationship between the rock texture and the fractal dimension of resistivity logs. For this purpose, core data and corresponding resistivity logs were analysed. The core data were used to classify the cores in rock types. Rock types were defined using porosity and permeability of the cores along with pore size distribution.

The collected data were primarily quality controlled. Any cores or intervals with missing or noisy data or failed MICP tests were excluded. Based on these properties, the cores were then divided into three different rock types.

The fractal dimension of the resistivity log corresponding to each core was calculated. To calculate fractal dimension, the data need to be unimodal and stationary without imposing any specific trend. This is a criterion of homogeneity required for a meaningful fractal dimension. Homogeneous datasets exhibit a unimodal distribution, and a visual inspection of the histogram is sufficient to ensure that just a single peak exists in the data density distribution. It should be noted that resistivity data are not Gaussian but log-normal. Thus, the logarithm of resistivity values should be taken before plotting the histogram. If the histogram is not unimodal, the log interval needs to be broken down into more homogenous sections. Also, before calculation of the fractal dimension, the data need to be normalised with zero mean and unit standard deviation. Given the type of data and common sampling rates of well logs, Higuchi's method was chosen for this study. Higuchi's algorithm was coded in Matlab software and used for calculating fractal dimensions.

Finally, the fractal dimensions were compared with the defined rock types.

Results and discussion

Cores from different locations in the Cooper Basin were analysed. The cores were classified into three rock types. The fractal dimension of the corresponding resistivity log to each core was then calculated using Higuchi's method as mentioned in the work flow.

Based on the porosity and permeability measurements of the cores, along with the pore size distribution from MICP tests, the cores were classified into three main rock types. Properties of the rock types are presented in Table 1. The pore size distributions of all cores are presented in Figs 1–3. The rock types were defined as follows:

Rock type I was characterised by porosities in the range of 13–20% and permeability range of 50–300 mD. The pore throat size distribution ranged within 1–100 µm throats. About 50% of the pore radiuses were in the range of 10–100 µm. This arrangement showed that the rock type consisted of large grain sizes and the high permeability suggested good interconnectivity in the structure.

Table 1. Properties of the defined rock types

Rock type	Porosity (%)	Permeability (mD)	Fractal dimension	Description
I	13–20	50–300	1.145–1.158	Majority of pore throats in the range 10–100 µm
II	7–14	5–40	1.186–1.199	Majority of pore throats in the range 1–10 µm
III	5–14	0.096–3	1.221–1.286	Majority of pore throats in the range 0.1–1 µm

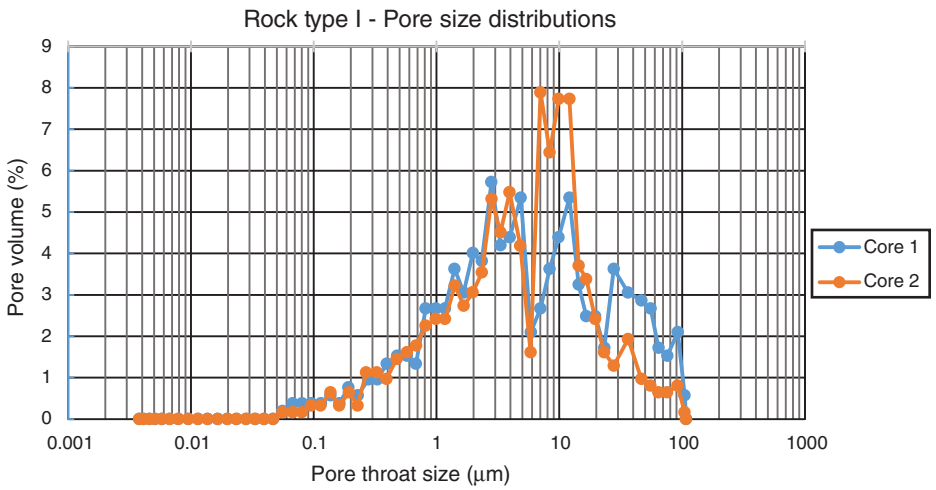

Fig. 1. Pore size distribution of the cores related to rock type I.

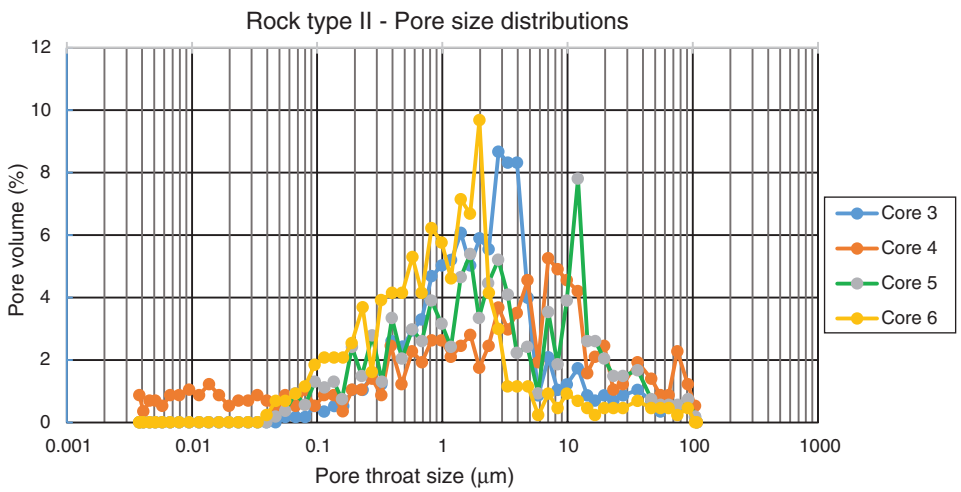

Fig. 2. Pore size distribution of the cores related to rock type II.

Rock type II featured porosities ranging within 7–14% and permeability within 5–40 mD. The majority of the pore throat sizes were within 1–10 µm, which was an order of magnitude smaller than that of type I. The pore size distribution suggested that the grain sizes in this type were smaller and reduction in permeability was due to increase in capillary pressures.

Rock type III was a tight sandstone reservoir with porosities ranging within 5–14% and permeability within 0.096–3 mD. The majority of the pore throats were 0.1–1 µm, which was an order of magnitude smaller than that of type II. Although the porosity was similar to type II, the permeability suggested that most of the porosity was ineffective. The grain sizes in this rock type were very fine and pore throats were filled with clay.

The fractal dimension of the resistivity well logs corresponding to each core was calculated using Higuchi's method. Some of the well log intervals are presented in Figs 7 and 8 as examples. In order to calculate fractal dimension of logs, the data needed to be stationary and Gaussian. To achieve this, homogenous intervals were chosen

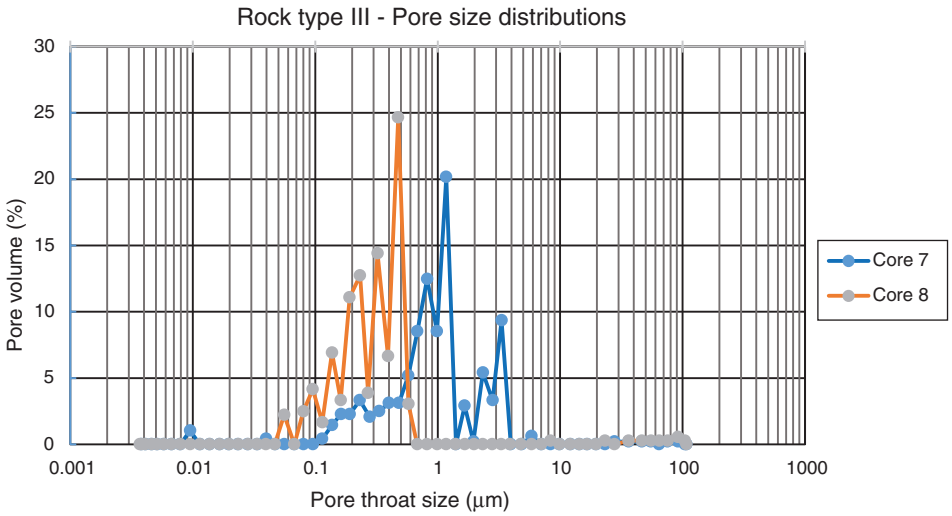

Fig. 3. Pore size distribution of the cores related to rock type III.

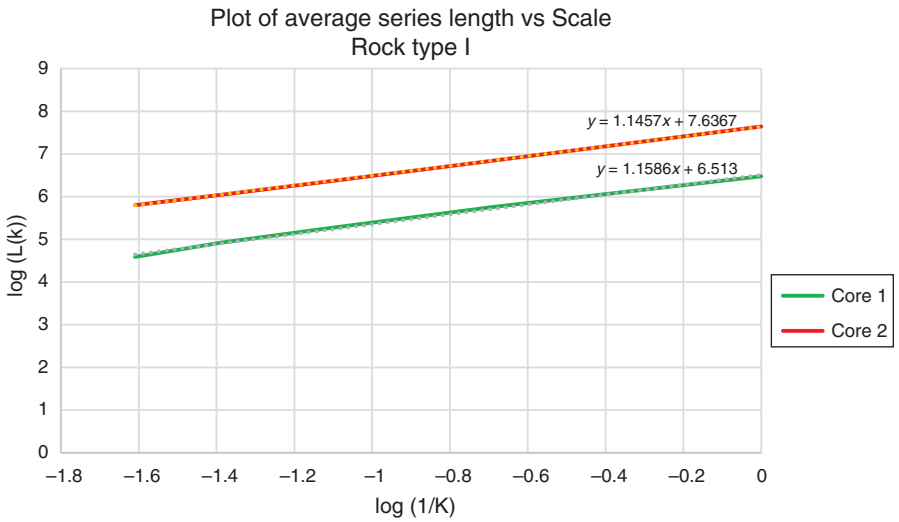

Fig. 4. Plot of scale vs length of series (well log) related to rock type I.

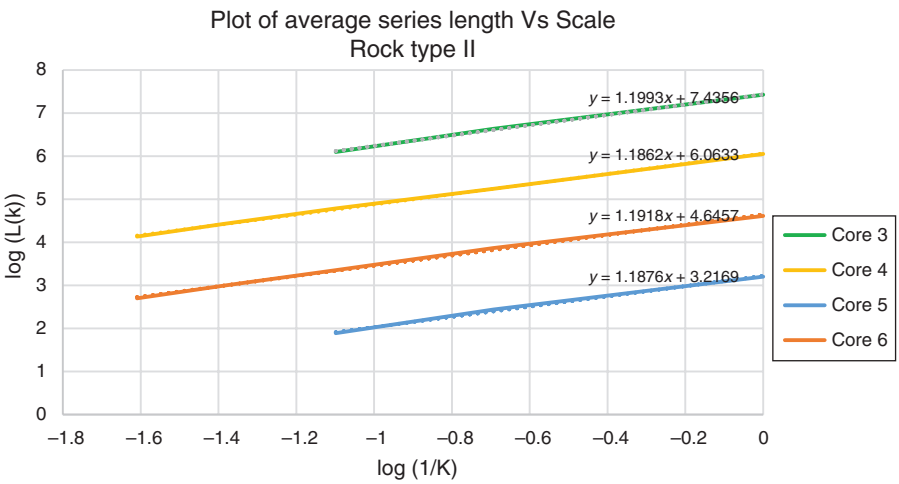

Fig. 5. Plot of scale vs length of series (well log) related to rock type II.

based on the gamma ray and sonic logs. The numerical values of resistivity curves were then extracted and normalised by subtracting the average and dividing by the standard deviation of the data.

As per Higuchi's method, series lengths were calculated for different scales k. The log of the average length of each series ($\log(L(k))$, was plotted against the log of the inverse of the scale ($1/k$). The result was a straight line, with slope equal to the fractal dimension of the resistivity log. The plots of k vs $\log(L(k)$ for every core are shown in Figs 4–6. The slope of the line was determined by a best linear fit to the points. The range of the fractal dimensions of each rock type is presented in Table 1.

As the complexity of rock structure increased (due to smaller grain sizes and/or increased clay content) and fluid flow in the media became more difficult, the fractal dimension of the corresponding resistivity log increased (Table 1). This indicates that the complexity of the rock structure was reflected in the variability of the resistivity logs.

Rock type I featured large pore sizes and high porosity overall. Large pore size indicated large, well-sorted matrix grains and the high permeability indicated that pores were well connected. The pore size distributions of the cores of type I are plotted in Fig. 1. A considerable portion of the pore throat sizes ranged within 10–100 µm. The fractal dimension of the resistivity log

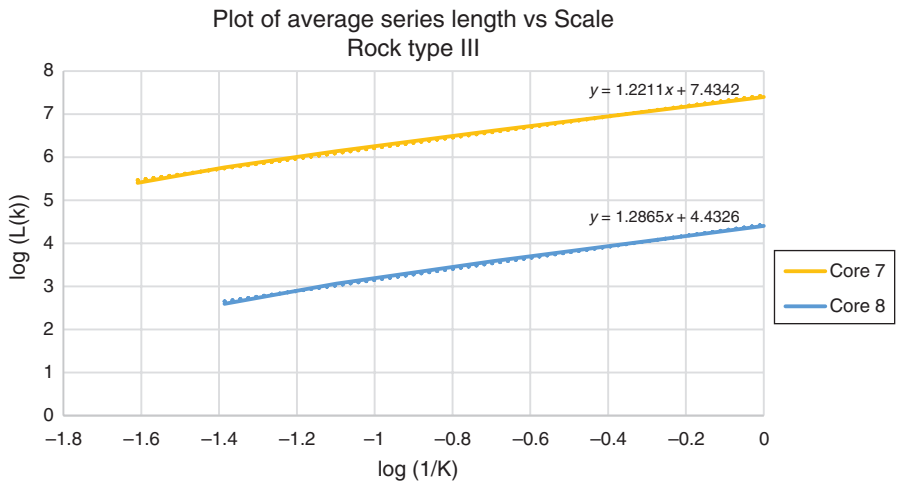

Fig. 6. Plot of scale vs length of series (well log) related to rock type III.

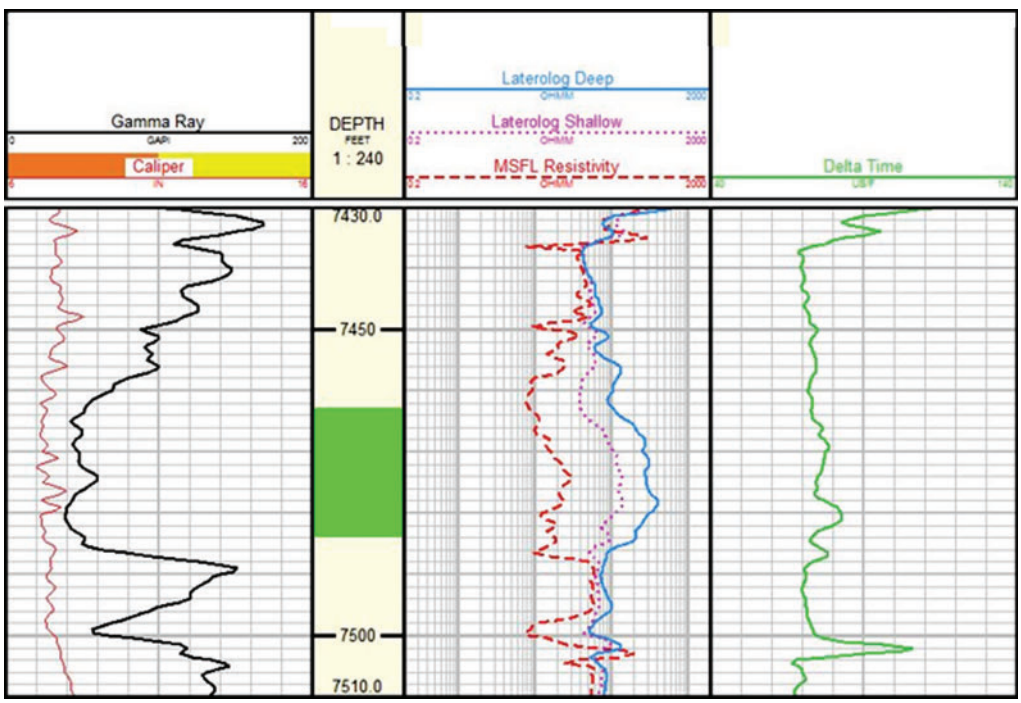

Fig. 7. Well logs corresponding to core 1.

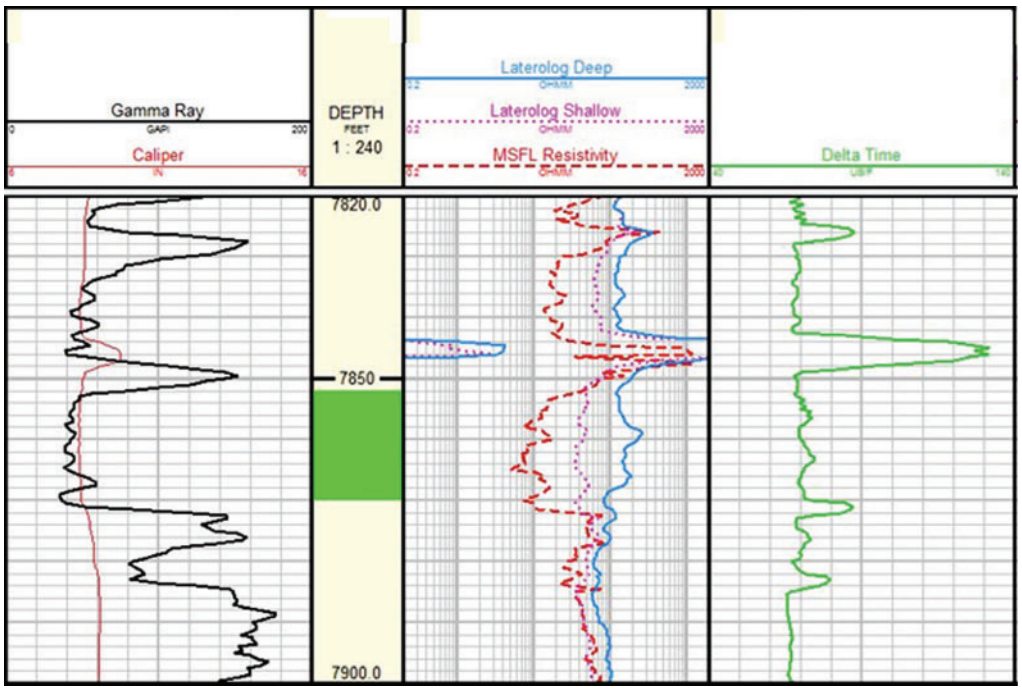

Fig. 8. Well logs corresponding to core 4.

corresponding to this rock type was in the range of 1.145–1.158, which is the lowest among the three rock types investigated.

For rock type II, pore sizes were an order of magnitude smaller than for type I. Smaller pores indicated a component of smaller grain sizes and this rock type also showed lower permeability. The pore size distribution of the cores of this rock type are shown in Fig. 2. A considerable fraction of the throat sizes were in the range of 1–10 μm. The fractal dimension of the resistivity log relating to this rock type ranged within 1.186–1.199, showing more complexity in the logs compared with type I.

Rock type III had very low permeability and the pore sizes are hypothesised to be much smaller than types I and II, as shown in the pore size distribution of the corresponding cores (Fig. 3). The gamma readings in this rock type were typically high, indicating the presence of clays, consistent with the very low permeability. Porosity in type III was similar to that of type II (ranging around 5–14%) but the majority of this porosity would be ineffective due to the clay content. The fractal dimension measured for the logs relating to type III ranged within 1.221–1.286, which was higher than both other rock types, although in this case electrical matrix surface conductivity would be an important component in the resistivity response.

The results of this study could be used to confirm chosen rock types, and also show that information could be hidden in the variability of logs which has so far been neglected.

Conclusions

In this study, core data from Cooper Basin were categorised into three rock types based on porosity, permeability, pore size distribution and core descriptions. The fractal dimension of the resistivity logs corresponding to each rock type was calculated using Higuchi's method.

The fractal dimension of resistivity logs related to rock type I, which was characterised by better sorting and bigger pore sizes, was 1.145–1.158. Rock type II was more resistant to fluid flow than type I but less than type III, and had fractal dimension range of 1.186–1.199. Finally, rock type III with the most complex pore structure showed the highest fractal dimension range of 1.221–1.286.

The results suggest that higher fractal dimensions corresponded to more complex rock fabric, i.e. as the pore structure became more resistant to fluid flow, with smaller pore sizes, more dead-end pore spaces and more tortuosity, the fractal dimension of the resistivity well logs increased. We conclude that fractal dimension of resistivity logs can be used to identify complexity of the electrical conductivity network fabric (often equivalent to the pore network fabric) and thus rock types.

This work was carried out using gas-saturated sandstone reservoir intervals. More research is needed to extend the findings to water- or oil-saturated zones and other lithologies.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors would like to acknowledge Santos and partners for providing the data used in this study.

References

Aasum, Y., and Kelkar, M. G. (1991). An application of geostatistics and fractal geometry for reservoir characterization. *SPE Formation Evaluation* **6**(01), 11–19. doi:10.2118/20257-PA

Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., and Keelan, D. K. (1993). Enhanced reservoir description: using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. In 'SPE Annual Technical Conference and Exhibition'. 3–6 October 1993, Houston, Texas. (Society of Petroleum Engineers: Houston, TX.)

Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME* **146**(01), 54–62. doi:10.2118/942054-G

Avnir, D., Farin, D., and Pfeifer, P. (1985). Surface geometric irregularity of particulate materials: The fractal approach. *Journal of Colloid and Interface Science* **103**(1), 112–123. doi:10.1016/0021-9797(85)90082-7

Bigalke, J. (2000). A study concerning the conductivity of porous rock. *Physics and Chemistry of the Earth. Part A: Solid Earth and Geodesy* **25**(2), 189–194. doi:10.1016/S1464-1895(00)00030-2

Crane, S. D., and Tubman, K. M. (1990). Reservoir variability and modeling with fractals. In '65th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers'. 23–26 September 1990, New Orleans. (Society of Petroleum Engineers: New Orleans.)

Emanuel, A. S., Alameda, G. K., Behrens, R. A., and Hewett, T. A. (1990). Reservoir performance prediction methods based on fractal geostatistics. In 'Annual Meeting of the American Association of Petroleum Geologists'. 20 March 1988, Houston, TX. (Society of Petroleum Engineers: United States.)

Feder, J. (1988). 'Fractals.' (Plenum Press: New York.)

Gravestock, D., and Jensen-Schmidt, B. (1998). Petroleum Geology of South Australia: Chapter 5: Cooper Basin—Structural Setting. Report Book, Vol. 98(9), pp. 47–66.

Gravestock, D. I., Hibburt, J., and Drexel, J. F. (1998). 'Petroleum Geology of South Australia. Volume 4: Cooper Basin.' (Petroleum Group, Primary Industries and Resources SA: Adelaide.)

Gray, T. A., Chork, C. Y., and Taggart, I. J. (1993). Pitfalls in the fractal analysis of reservoir property data. In '68th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers'. 3–6 October 1993, Houston, Texas. (Society of Petroleum Engineers: Houston, TX.)

Gunter, G. W., Finneran, J. M., Hartmann, D. J., and Miller, J. D. (1997). Early determination of reservoir flow units using an integrated petrophysical method. In 'SPE Annual Technical Conference and Exhibition'. 5–8 October 1997, San Antonio, Texas. (Society of Petroleum Engineers: United States.)

Gupta, I., Rai, C., Sondergeld, C., and Devegowda, D. (2017). Rock typing in Wolfcamp Formation. In 'SPWLA 58th Annual Logging Symposium'. 17–21 June 2017, Oklahoma City. (SPWLA 58th Annual Logging Symposium: Oklahoma.)

Hardy, H. H. (1992). The generation of reservoir property distributions in cross section for reservoir simulation based on core and outcrop photos. In 'Permian Basin Oil and Gas Recovery Conference'. 18–20 March 1992, Midland, TX. (Society of Petroleum Engineers: United States.)

Hardy, H., and Beier, R. A. (1994). 'Fractals in Reservoir Engineering.' (World Scientific.)

Hewett, T. A. (1986). 'Fractal Distributions of Reservoir Heterogeneity and Their Influence on Fluid Transport.' (Society of Petroleum Engineers.)

Hewett, T. A., and Behrens, R. A. (1990). Conditional simulation of reservoir heterogeneity with fractals. *SPE Formation Evaluation* **5**(03), 217–225. doi:10.2118/18326-PA

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D. Nonlinear Phenomena* **31**(2), 277–283. doi:10.1016/0167-2789(88)90081-4

Katz, A. J., and Thompson, A. H. (1985). Fractal sandstone pores: Implications for conductivity and pore formation. *Physical Review Letters* **54**(12), 1325–1328. doi:10.1103/PhysRevLett.54.1325

Krohn, C. E., and Thompson, A. H. (1986). Fractal sandstone pores: automated measurements using scanning-electron-microscope images. *Physical Review B: Condensed Matter and Materials Physics* **33**(9), 6366–6374. doi:10.1103/PhysRevB.33.6366

Kumar, M., Sok, R., Knackstedt, M. A., Latham, S., Senden, T. J., Sheppard, A. P., Varslot, T., and Arns, C. (2010). Mapping 3D pore scale fluid distributions: how rock resistivity is influenced by wettability and saturation history. *Petrophysics* **51**(02), 102–117.

Lozada-Zumaeta, M., Arizabalo, R. D., Ronquillo-Jarillo, G., Coconi-Morales, E., Rivera-Recillas, D., and Castrejon-Vacio, F. (2012). Distribution of petrophysical properties for sandy-clayey reservoirs by fractal interpolation. *Nonlinear Processes in Geophysics* **19**(2), 239–250. doi:10.5194/npg-19-239-2012

Mandelbrot, B. B. (1983). 'The Fractal Geometry of Nature.' Updated and augmented edn. (W.H. Freeman: New York.)

Pang, J., and North, C. P. (1996). Fractals and their applicability in geological wireline log analysis. *Journal of Petroleum Geology* **19**(3), 339–350. doi:10.1111/j.1747-5457.1996.tb00438.x

Perez, G., and Chopra, A. K. (1997). Evaluation of fractal models to describe reservoir heterogeneity and performance. *SPE Formation Evaluation* **12**(01), 65–72. doi:10.2118/22694-PA

Pittman, E. D. (1992). Relationship of porosity and permeability to various parameters derived from mercury injection-capillary pressure curves for sandstone (1). *AAPG Bulletin* **76**(2), 191–198.

Porras, J. C., Barbato, R., and Khazen, L. (1999). Reservoir flow units: a comparison between three different models in the Santa Barbara and Pirital Fields, North Monagas Area, Eastern Venezuela Basin. In 'SPE Latin American and Caribbean Petroleum Engineering Conference'. 21–23 April 1999, Caracas, Venezuela. (Society of Petroleum Engineers)

Raghavendra, B., and Dutt, N. D. (2010). Computing fractal dimension of signals using multiresolution Box-Counting method. *International Journal of Information and Mathematical Sciences* **6**(1), 50–65.

Rider, M. H. (1986). 'The Geological Interpretation of Well Logs.' (John Wiley and Sons, Inc.: New York.)

Rushing, J. A., Newsham, K. E., and Blasingame, T. A. (2008). Rock typing: keys to understanding productivity in tight gas sands. In 'SPE Unconventional Reservoirs Conference'. 10–12 February 2008, Keystone, CO. (Society of Petroleum Engineers: United States.)

Shen, P., Liu, M., and Jia, F. (1998). Application of fractal techniques in reservoir development. In 'SPE International Conference and Exhibition'. 2–6 November 1998, Beijing, China. (Society of Petroleum Engineers: Beijing)

Swanson, B. F. (1985). Microporosity in reservoir rocks: Its measurement and influence on electrical resistivity. *The Log Analyst* **26**(06).

Toledo, P. G., Novy, R. A., Davis, H. T., and Scriven, L. E. (1994). Capillary pressure, water relative permeability, electrical conductivity and capillary dispersion coefficient of fractal porous media at low wetting phase saturations. *SPE Advanced Technology Series* **2**(01), 136–141. doi:10.2118/23675-PA

Verwer, K., Eberli, G. P., and Weger, R. J. (2011). Effect of pore structure on electrical resistivity in carbonates. *AAPG Bulletin* **95**(2), 175–190. doi:10.1306/06301010047

Wang, Z.-w., and Mou, D. (2014). Fractal dimension of well logging curves associated with the texture of volcanic rocks. In 'International Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC 2014)'. Shenyang, China, 2014. (Atlantis Press: Hong Kong.)

The authors

Roozbeh Koochak is an experienced wireline logging engineer with over 10 years of experience in both cased-hole and open-hole services. He joined China National Logging Corp. in 2006 and later Baker Hughes as senior engineer in 2012. During this period he held assignments in the Middle East, Australia and China, and managed operations for two years. He is currently a Master of Philosophy Candidate in Petroleum Engineering at the University of Adelaide, working on identification of rock types using fractal theory. Roozbeh has a Bachelor of Electronics Engineering from Shahid Chamran University of Ahvaz. His research interests are reservoir modelling and simulation, petrophysics, wavelet analysis, fractal theory, uncertainty analysis, production optimisation and history matching.

Manouchehr (Manny) Haghighi is an Associate Professor of Petroleum Engineering. His research and teaching focus is on unconventional reservoirs, reservoir simulation, well testing and formation evaluation. He has supervised more than 40 MSc and 10 PhD students. Before joining the University of Adelaide in 2009, Manouchehr was Associate Professor of Petroleum Engineering at the University of Tehran. In 2000, Manouchehr established Simtech, a consulting company for integrated reservoir simulation. He has been project director of several full field simulation projects for oil and gas reservoirs. From 1995 to 2000, Manouchehr worked with the National Iranian Oil Company (NIOC) and was the director of a program for training NIOC staff at several universities in the US, UK, Canada, France, Australia and Norway. Manouchehr was a Visiting Professor at the University of Calgary during 2007–2008. Manouchehr has published more than 80 articles in peer-reviewed journals or presented in international conferences. He has served as a reviewer for different journals such as the journal of Petroleum Science and Engineering, and is a member of SPE.

Mohammad Sayyafzadeh is a lecturer in Petroleum Engineering at the University of Adelaide, where he started in 2013. His research interest is applied and computational mathematics targeting reservoir and production engineering problems. That includes computer-assisted history matching, uncertainty quantification, field development planning and reservoir flooding optimisation, modelling of unconventional resources and data analytics. Mohammad holds a BSc in Chemical Engineering, an MSc in Reservoir Engineering from Tehran Polytechnic and a PhD in Petroleum Engineering from the University of Adelaide. Mohammad has contributed in publishing 23 papers in peer-reviewed journals and conferences, and serves as a reviewer for different journals and EAGE/SPE conferences. He is the lead-investigator of a project on developing a computer-assisted history matching tool sponsored by Santos Ltd and participated as a co-investigator and research fellow in three other industrial projects.

Mark Bunch is a Senior Lecturer in Petroleum Geoscience at the Australian School of Petroleum, University of Adelaide. His present research activities are concerned with formation evaluation and seismic attribute analysis. Prior to his present role, he spent seven years with the CO2CRC as a Research Associate in reservoir characterisation, during which he worked on geological carbon storage site selection, capacity estimation and geological modelling projects in the onshore Canterbury Basin (NZ), the Gippsland and Otway basins of Victoria, the Surat Basin of Queensland and the Darling Basin of NSW. Mark spent a period of time as acting head of geomodelling for the CO2CRC Otway Basin Pilot Project and led CO2CRC storage research projects for five years. Mark holds degrees in Geology and Geophysics (BSc Hons), Hydrogeology (MSc) and a PhD in Earth Sciences (Stratigraphic Forward Modelling). Mark has also worked for the North Sea Palaeolandscapes Project and as a developer of shallow groundwater flow models to guide excavation planning.

4. A transfer learning approach for facies prediction using resistivity image well logs

Roozbeh Koochak, Ali Nadian Ghomsheh, Manouchehr Haghighi, Mark Bunch, Mohammad Sayyafzadeh,

Submitted to Geoscience Frontiers, 2023

Statement of Authorship

Title of Paper	**A transfer learning approach for facies prediction using resistivity image well logs**
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Koochak, R., Nadian, A., Bunch, M. and Haghighi, M.,Sayyafzadeh, M., 2023. A transfer learning approach for facies prediction using resistivity image well logs. submitted to Geoscience Frontiers for peer-reviewed publication**

Principal Author

Name of Principal Author (Candidate)	**Roozbeh Koochak**		
Contribution to the Paper	**Conceptualization, development of algorithm, experimentation, analysis, coding, original manuscript, corresponding author duties**		
Overall percentage (%)	**60%**		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	**17/05/2023**

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	**Ali Nadian Ghomsheh**		
Contribution to the Paper			
Signature		Date	**23/05/2023**

Name of Co-Author	Mark Bunch		
Contribution to the Paper	Discussions, revisions, sections of original manuscript, supervision		
Signature		Date	19/05/2023

Name of Co-Author	Manouchehr Haghighi		
Contribution to the Paper	Discussions, revisions, supervision		
Signature		Date	18/05/2023

Name of Co-Author	Mohammad Sayyafzadeh		
Contribution to the Paper	Analysis of results, discussions, revisions, supervision		
Signature		Date	18/05/2023

Please cut and paste additional co-author panels here as required.

A Transfer Learning Approach for Facies Prediction Using Resistivity Image Well Logs ⋆

Roozbeh Koochak[a,*], Ali Nadian Ghomsheh[b,**], Mark Bunch[c], Manouchehr Haghighi[a] and Mohammad Sayyafzadeh[a]

[a] *Discipline of Mining and Petroleum, The University of Adelaide, Australia*
[b] *Cyberspace Research Institute, Shahid Beheshti University, Tehran, Iran*
[c] *Discipline of Earth Sciences, The University of Adelaide, Australia*

ARTICLE INFO

Keywords:
Automatic facies detection
Resistivity image logs
Transfer learning
Fullbore Formation Microimager
Convolutional Neural Network

ABSTRACT

Resistivity image logs have a long history in facies detection and classification. Manual interpretation of these logs is subjective, time consuming and costly. There has always been a need to augment this process with the purpose to increase efficiency and consistency. In this study, we introduce a supervised learning technique to augment image facies detection based on electric log images alone. Being able to detect facies solely using image logs would reduce reliance on supplementary logs and reduce cost significantly. This workflow requires the geologist to label example facies in one well to be used as training data, then the trained network classifies logs at other locations. A common challenge reported in literature for application of supervised learning to facies detection is shortage of labelled data. To overcome this challenge, we propose a method of data preparation to increase the number of images from available data. The preferred machine learning technique in this study is Transfer learning. Three pretrained networks (VGG, Inception V3 and Resnet) were investigated and their output compared. In our dataset, two out of seven wells were interpreted and labelled. We used one well to train the network and the second well for validation. We introduce a post processing methodology to adapt the pre-trained classifier network to evaluating a continuous resistivity image log that not only delivers consistent results but enhances the bed-boundary detection resolution. Our proposed methodology can efficiently augment resistivity image log interpretation and reduce dependence on a particular human interpreter. The results show that in the absence of other information/logs, the trained network can detect image facies with a testing accuracy of 82% using electrical image logs alone and the proposed post-processing method increases the final categorization accuracy even further.

1. Introduction

Lithofacies or facies represent the sedimentary unit whose petrophysical properties can be distinguished from surrounding rock (Liu et al., 2021). Facies classification is a crucial step in CO2 sequestration projects, as the injection target beds are determined based on petrophysical property modelling that is guided by lithological units and features of the formation. At any particular well location, classification can be based on log data or cores. However, in between well locations the geophysical properties required for facies classification are derived from coarse-scale measurements such as seismic properties determined using surface surveys (wu2015). This study focuses on facies classification at well locations using electrical borehole image logs.

High-resolution borehole image logs, when calibrated with core data and conventional logs provide detailed insights on lithology, sedimentary textures, palaeoflow directions, structural dip analysis, in situ stress analysis and fracture evaluation (Nie et al., 2013; Ameen, 2014; Kosari et al., 2015; Brekke et al., 2017; Lai et al., 2018). These logs significantly contribute to geological understanding and structural interpretation of the logged interval (Folkestad et al., 2012; Lai et al., 2018). Sedimentological and structural features down to a few millimetres in resolution can be identified using borehole image logs (Ja'fari et al., 2012). As such, image logs are a valuable source of information for facies classification, which is an integral part of stratigraphic correlation and reservoir prediction (Liu et al., 2021).

⋆ This document is the results of the research project funded by CO2CRC.
* Corresponding author
** Principal corresponding author
roozbeh.koochak@adelaide.edu.au (R. Koochak); ali_nadian@sbu.ac.ir (A.N. Ghomsheh)
ORCID(s): 0000-0001-9447-0561 (R. Koochak); 0000-0002-2215-409X (A.N. Ghomsheh); 0000-0002-7012-1217 (M. Bunch); 0000-0001-9364-2894 (M. Haghighi); 0000-0002-4414-372X (M. Sayyafzadeh)

Electrical borehole image logs provide a two-dimensional, 360 degrees image of the borehole wall. After decades of application, geological interpretation of borehole image logs remains subjective particularly when the geological setting is complex (Lai et al., 2018). The background experience of the interpreter becomes a primary influence on the final interpretation. The process for facies identification using resistivity image logs requires an experienced geologist to visually inspect the logs and manually pick geological features using specialized application software. However, this process is tedious, time consuming and costly (Gupta et al., 2019). There is a great need for an automated workflow to augment and optimize this manual process.

Machine learning has shown promising results in augmenting facies classification at well locations. Gupta et al.(2019) used a UNET architecture to pick induced and natural fractures along with sedimentary surfaces. Labelled data from four wells was used in that study to train two separate networks. One locally trained and the other globally trained. In the locally trained network only 10% of the data in a single well were used in training and the network was used to detect features in the rest of the well. The global models are trained using labelled data from multiple wells where the model is then used to detect learned features in other wells in the region. The authors suggested that supervised-learning approaches are challenging due to sparse features and reported occurrences of false classification, without investigating the cause or offering a solution. Lima et al.(2019) proposed an unsupervised neural network model for pattern recognition and facies categorization using borehole image logs. They used a non-linear autoencoder for representational learning to reconstruct the original training images and applied cluster analysis for categorization of the facies. The methodology in their study does not guarantee returning facies categories that relate to the required geological application. Furthermore, the study does not propose how to distinguish structurally different facies that are visually similar. Using the concept of transfer learning, Lefranc et al.(2021) trained a ResNet architecture on synthetic training data produced by models. To enhance accuracy, they added noise to the synthetic data. To categorize real logs, the method requires dipping planar features (nominally sedimentary dip fabrics) to be picked manually at high density and the background resistivity must be removed so the real log is visually similar to the synthetic images. Other limitations of this method are that the training data do not cover the full range of sedimentary bedforms and the network will not be trained on features of any post depositional deformation that may have occurred. The majority of workers contributing studies to the literature have complained of a lack of reliably labelled data to use for training their models and absence of post processing methods for augmenting or categorizing additional bed boundaries. In this work we propose a methodology to increase the number of training images from available labelled data and a post processing method is developed to enhance classification using continuous logs to identify bed boundaries at a resolution of 0.3 meters. To achieve this, we trained a convolutional neural network to classify lithological facies in the Late Cretaceous succession tested by CO_2 injection and storage operations of the CO2CRC Otway Project in the Otway Basin of south eastern Australia (Bunch, 2014). A supervised learning scheme was employed to augment facies classification at well locations, using resistivity image logs. Data from two wells, CRC-1 and CRC-2 were available. Logs from both wells were quality controlled, processed and interpreted. In this study we used CRC-2 interpreted facies classes as input training data and the noisier CRC-1 image log record as the testing case. The goal of this study was to investigate whether a CNN can extract enough features from the image log data alone to be able to accurately distinguish facies classes. This would save the cost of running other supplementary logs along with borehole image logs (though in reality a standard 1D log suite will always be acquired for baseline information). As such, in this study supplementary information such as core data and conventional log data were not involved in the training of the network. Our results confirm that the distribution of the images was quite complex. This is because, (1) the training images lacked distinctive features in comparison to other datasets like those typically used for training image recognition models, such as animals or everyday objects, and (2) the relative visual similarity of image log facies categories. To overcome these challenges, we used transfer learning which is a technique that takes part of a larger trained model and uses it in a new model for a related task, without the need to access large training data and computing resources to train the original model.

Transfer learning has proven to be a very effective technique in the field of image recognition and image classification. In deep convolutional neural networks, with many layers, the lower layers detect the features while the final layers detect the class of the image (Tang, 2018). In the case of image classification, transfer learning allows us to train the classifier layers of the CNN with our specific set of images requiring far fewer training images and computational power (Tang, 2018). Transfer learning allows us to retrain the final layer of an existing model, resulting in a significant decrease in not only training time, but also the size of the dataset required. There are numerous pre-trained models used in transfer learning applications such as those named VGG, Inception and ResNet. These networks were originally trained on millions of images from thousands of classes. Retaining the classifier layers means that we can maintain the parameters (knowledge) the network has learnt during the original training and apply it to our dataset.

This results in high accuracy classification of facies with a much lower requirement for training data and demand for computing power. In this study, the three mentioned networks were applied using the input image log dataset and a post processing methodology was developed to adapt the CNN architecture to better evaluate the continuous resistivity image log. Our main goal was to explore whether CNNs are able to extract features from this modality of data alone. The remainder of this text is structured as follows: We continue with the geologic setting of the data. Next image classification using deep CNNs is reviewed. Data preparation methodology and proposed deep learning model are presented in the subsequent sections followed by results and discussions. Finally, conclusions and future work are presented.

2. Geologic setting

The CO2CRC project field demonstration site is located in south-eastern Australia, within the onshore Otway basin. The facility is to the southwest of Melbourne in the state of Victoria, between the coastal towns of Port Campbell to the east and Warrnambool to the west. The Otway Basin is a rift basin stretching northwest from southeast of South Australia to the northwest of Tasmania. The basin formed during the Antarctic-Australian separation associated with the break-up of Gondowana (Willcox and Stagg, 1990). Local tectonic activity and eustatic sea level variation has produced five major group-level sedimentary successions: the Otway, Sherbrook, Wangerrip, Nirranda and Heytesbury groups (Woollands and Wong, 2001). Deposition of the lower most Otway Group was controlled by tectonic activity prior to local relative sea level changes. The depositional environment was continental, producing first volcaniclastic sediments graduating to fluvial and shallow lacustrine sediments in the upper parts (Mishra et al., 2019; Woollands and Wong, 2001). Marine breakthrough is recorded by the Sherbrook Group, as sea level became regional base level to the depositional environment. The succession varies from marginal marine sediments at its base to those of a widespread deltaic plain environment at the top (Boyd and Gallagher, 2001; Mishra et al., 2019; Woollands and Wong, 2001). The Wangerrip group was deposited under varying conditions from low energy marginal marine to shallow marine. The depositional environment in this group is represented by a first-order transgressive-regressive cycle producing first a fining up, then a coarsening upwards succession (Morton et al., 1994; Mishra et al., 2019). The Nirranda Group developed under open marine conditions and the Heytesbury Group bears depositional environments varying from inner shelf at the base to mid-shelf at the top (Mishra et al., 2019; Woollands and Wong, 2001). The Sherbrook Group is the primary interval of interest for CO_2 storage operations of the Otway Project. The Waarre Formation – the primary petroleum reservoir within the onshore basin area – and the deep saline reservoir formations of the lower Paaratte Formation were the first and second targets for CO_2 sequestration experiments respectively (Bunch et al., 2012).

2.1. Facies categories

Six high-level facies categories were chosen for training the CNN. These capture 15 sub-classes of mainly depositional sedimentary litho-facies interpreted manually following acquisition of core and image log data. The six high-level categories are defined as Mudstone, Heterolithics, Muddy Sandstone, Clean sandstone, Cemented layers and Pebbles. The Mudstone category generally shows no visible internal laminations with common to abundant disseminated pyrite. It can sometime have stress features like borehole breakouts or drilling induced tensile fractures (which are far rarer). The Heterolithics facies category is characterized by alternating layers of centimeter to decimeter thick, moderately resistive layers with similarly thick conductive mudstones. The bed boundaries are generally sharp but in some cases are diffuse as well. Tops and bases of layers are generally well defined but when they are rather diffuse this category is difficult to distinguish from Muddy Sandstone. The Muddy Sandstone category comprises of alternating resistive (possibly sand-rich) and lightly conductive (possibly muddy) layers. The layer boundaries are generally more diffuse. The thickness of the layers ranges between centimeter and decimeters. Dispersed pyrite nodules are characteristic of this category presenting as small dark conductive patches throughout the layer. When the resistivity contrast range of the layers is high this category looks similar to Heterolithics. The presence of pyrite nodules in this category can sometime mean that it looks similar to Mudstone when images are dynamically normalized. The Clean sandstone facies generally appears in structureless (‘massive’) thicker beds occasionally with a mottled resistivity fabric. Changes in resistivity can be due to change in porosity, grain packing or cementation. This facies can look similar to Mudsone or sometimes Heterolithics when images are dynamically normalized. Cemented layers represent a distinctly post-depositional type that is indicated by highly resistive intervals with elevated density and reduced neutron porosity values. The resistivity of these beds may exceed the resistivity range of the FMI tool and saturates the

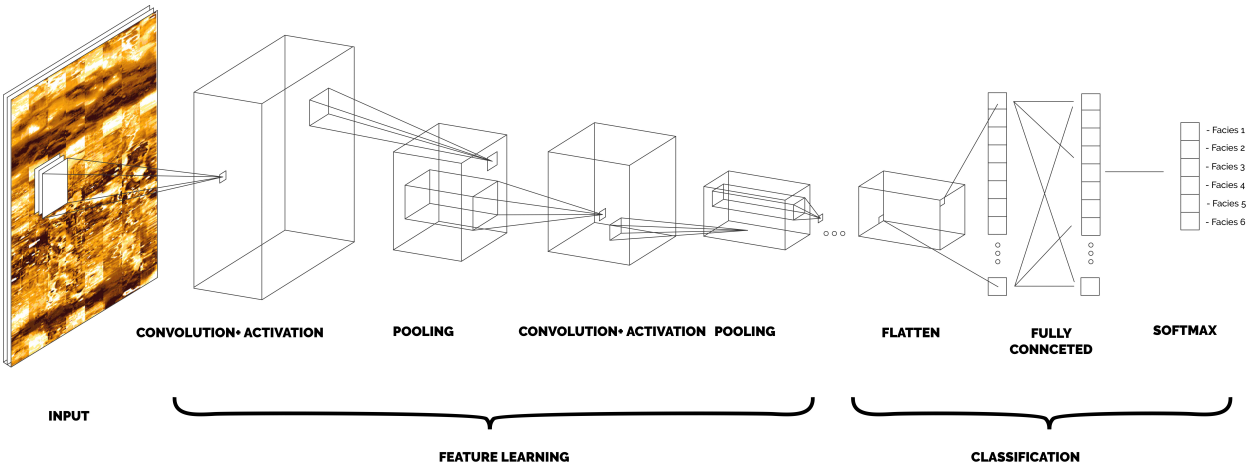

Figure 1: CNN architecture, depicting the main elements and structure

image, especially when images are statically normalized. This facies overprints primary sedimentary fabrics and is therefore quite distinct when compared to other categories.

3. Image classification using deep convolutional neural networks

CNN's have had extraordinary success in image recognition and classification applications. Image feature extraction is the basis for image classification. In general, image classification describes the whole image by extracting features using some feature learning method and then uses the classifier to identify the object category. Therefore, the method of extracting features from an image is especially important (Chen et al., 2021). Traditional image classification methods such as the Bag of Words model, extract features in a separate step (Lorente et al., 2021). The greatest success of deep CNNs is attributed to their strong feature-learning ability (Matthew Zeiler and Rob, 2014). Classification using a CNN is an end-to-end learning process. Only the original image is provided as input, the training and prediction process are carried out through the network, which produces the final output (Chen et al., 2021).

3.1. CNN Architecture

The main structure of a CNN comprises convolutional, pooling, nonlinear activation and fully connected layers. In the CNN architecture, the convolutional layer is the most significant component (Alzubaidi et al., 2021). On the one hand, it acts as the receptive field that retains the correlation between features and image pixels in the width and length direction. In other words, it can capture the spatial dependencies of the image. On the other hand, the convolutional layer repeatedly calculates the same convolution kernel at different positions using a sliding window, that is, using parameter sharing and sparse connections to avoid the parameter size being too large (Chen et al., 2021). The pooling layer subsamples the feature maps to convert large feature maps to smaller ones while maintaining the majority of dominant information at each step. This reduces network computing cost by reducing layer connections (Gu et al., 2018). The pooling layers also make the output feature map more robust to distortion or error generated by single neurons (Liu et al., 2017). The role of the non-linear activation layer in all neural networks is to map the input to the output. This introduces nonlinearity to the neural network; having a suitable nonlinear activation function can significantly enhance the performance of a network (Gu et al., 2018). The fully connected layer is generally located at the end of the CNN architecture. Inside this layer, each neuron is connected to all other neurons fulfilling the function of full connection. This layer is the classifier that categorises the local information extracted in the convolutional and pooling layers (Sainath et al., 2013). The number of neurons at the output stage is the number of required categories (Chen et al., 2021). An example of a CNN architecture is shown in Figure 1.

The history of deep CNN's began with the appearance of LeNet, which was introduced by LeCun et al. (1995). this was the first time that backpropagation was used in training CNNs. The architecture of LeNet-5 is shown in Figure 2. Krizhevsky et al. (2017) built the AlexNet model based on the LeNet-5. This network proved for the first time

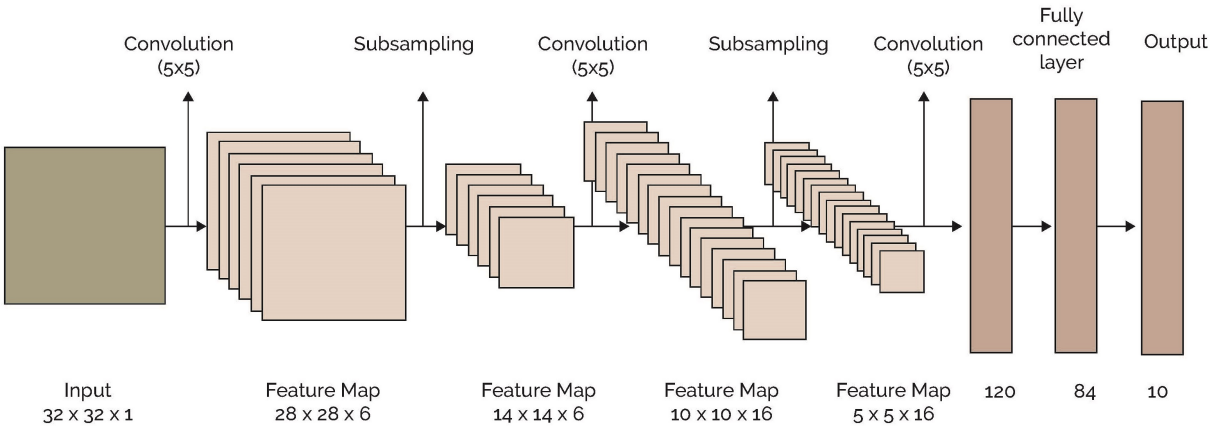

Figure 2: LeNet-5 Architecture

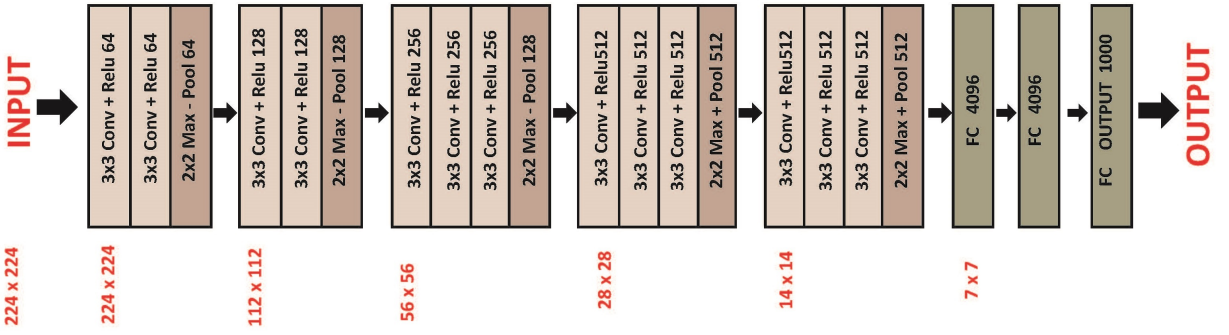

Figure 3: The Architecture of VGG-16. Convolution: Size= 3×3, Stride=1, Activation= Relu. Pooling: size= 3×3, Stride=2

that learned features can surpass manually extracted features (Chen et al., 2021). AlexNet improved CNN learning ability by increasing its depth and implementing several parameter optimization strategies (Alzubaidi et al., 2021). To enhance applicability of the CNN to a larger number of categories, the number of feature extraction stages was increased from five in LeNet to seven in AlexNet. The main drawback of AlexNet was its tendency towards overfitting due to the larger network. To alleviate this issue, overlapping pooling was introduced (Chen et al., 2021). Larger size filters were used to adapt the network to larger size input images and the RELU activation function replaced Sigmoid for faster convergence. In 2014, Simonyan and Zisserman (2014) proposed an innovative design for CNNs whereby the network was modularized. This design was called the Visual Geometry Group (VGG). The network comprised several identical convolutional layers (modules) in succession followed by a maximum pooling layer. The convolution layers maintain an unchanged image width and height while the pooling layer halves it. The VGG network has a variety of different layer structure models. Figure 3 shows the VGG-16 network layers. This network used smaller 3×3 filters and showed experimentally that parallel assignment of small filters has the same influence as larger filters. The main drawback of this is that it utilized close to 140 million parameters, which resulted in a high computational cost (Alzubaidi et al., 2021). The simplest way of enhancing CNN performance is by increasing their size. i.e., by adding more hidden layers with more channels in each layer. This approach has two major drawbacks. First, larger models are prone to overfitting especially in cases where training data is limited. Second, this approach dramatically increases the computational requirements for training. Furthermore, intuitively, visual information should be processed at multiple scales (Szegedy et al., 2015). This is because sometimes it is not possible to obtain enough useful features to perform

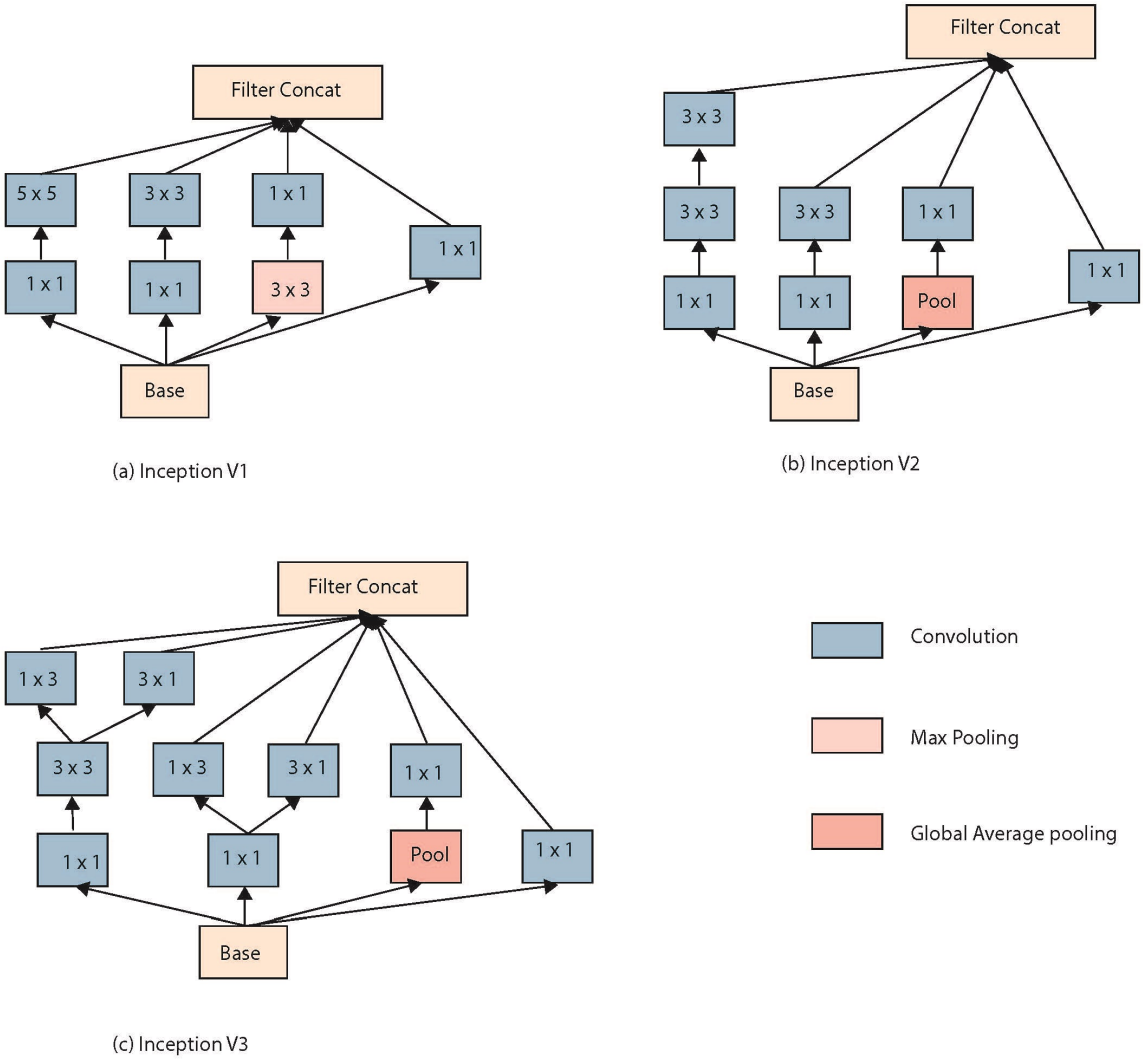

Figure 4: Inception V1 to V3 modules.

an accurate classification with a single convolution. In fact, some inputs work better with smaller convolution kernels and others with larger kernels Zaccone and Karim (2018). The Inception network was designed to address the above issues. An Inception network consists of Inception modules (shown in Figure 4) stacked on top of each other with an occasional max pooling layer (Szegedy et al., 2015). The inception module computes 1×1, 3×3 and 5×5 convolutions within the same module, therefore it acts as a multi-scale feature extractor. The outputs of these filters are then stacked along the channel dimension and fed to the next layer in the network Zaccone and Karim (2018). Inception-V3 is the third version in the series of Inception network architectures introduced by Google. It is a deep CNN used for image classification (Szegedy et al., 2015). ResNet was introduced to address network degradation. Many studies show that simply increasing the network depth does not proportionally increase its performance. With increasing network depth accuracy is saturated and degrades rapidly (He et al., 2016). This is not an issue of disappearing gradients or overfitting. In fact, deep networks with numerous layers can converge through initial and/or batch normalization. Degradation occurs when a network is not optimized well due to its increased depth. The residual connection in ResNet is a method employed to break degradation and enable deep neural networks to achieve high accuracy (Orhan and Pitkow, 2017; Chen et al., 2021). Figure 5 shows examples of the residual block used in Resnet. Figure 5(a) is an ordinary CNN

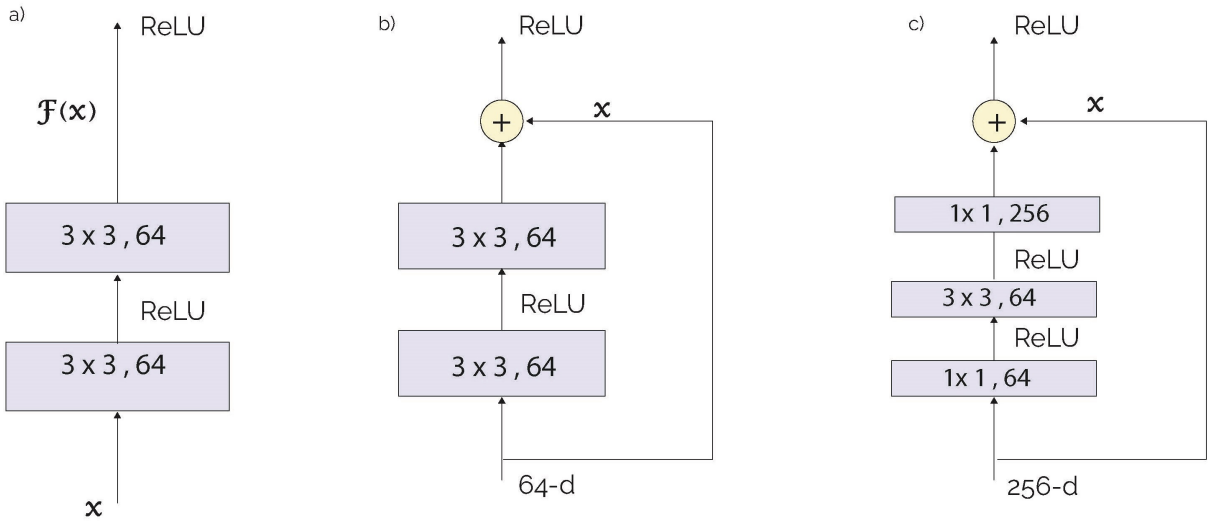

Figure 5: CNN blocks. a) ordinary CNN block b) ResNet-34 block c) deep ResNet blocks.

block. The residual block in Figure 5(b) is used in ResNet-34. When the network stack is deep, a 1×1 convolution is added at the end of the 3×3 layer to control the number of channels. The residual block in Figure 5(c) is used in ResNet-50/101/152.

In this study we trained three networks on our dataset, VGG-16, Inception-V3 and ResNet-101. All three networks architectures used Softmax as the classifier in the final layer. The input to the Softmax classifier is the embedding vector, where the length of the vector is equal to the number of classes. The SoftMax classifier scales the embeddings into probabilities. The output probabilities sum to one for all classes. Mathematically, Softmax is defined as:

$$S(y)_i = \frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_i}} \tag{1}$$

y is an input vector to Softmax function. It consists of n elements for n classes. The denominator is a normalization term. It ensures that the values of the output vector sum to one. We have used integer coding to label the training data. That is, integers [0] to [5] were used label facies one to facies six. Based on this definition of truth labels Sparce Categorical Cross Entropy loss is minimized to train the weights of the network. The mathematical notation of this loss is shown in equation 2.

$$J(w) = \sum_{i=1}^{n} t_i log\left(P_i\right) \tag{2}$$

In the above equation, w is the weights of the network. t_i is the truth value and p_i is the probability of the i^{th} class.

4. Data Preparation

CRC-1 and CRC-2 are two deep wells drilled 173 m apart (by wellhead location) for subsurface CO_2 injection experiments as part of the CO2CRC Otway Project. Data acquired at these wells as part of their drilling operations were available for network training and testing. Resistivity borehole images were acquired at both wells using Schlumberger's Fullbore Formation MicroImager (FMI). The FMI is an eight arm, pad-based micro-resistivity measurement device with pads in contact with the borehole wall. Log data were pre-processed and facies categories were interpreted throughout the logged interval. Log quality control (LQC) of raw data was undertaken to assess suitability of the raw data for processing and interpretation. Generally, the most important pre-processing of borehole image log data is speed/depth correction (electrode data alignment) and image orientation using high-resolution accelerometer

and magnetometer data (Wilson et al., 2013). For the purpose of well elevation alignment, CRC-2 log data were considered the reference case so no depth adjustment was applied. CRC-1 data were depth shifted to match CRC-2 log elevations. Regardless, lithostratigraphic variation exists between these wells despite their close proximity. During logging difficult borehole conditions such as washouts, breakouts or excessive pad pressure may cause tool sticking in the borehole and parts of the image to lose clarity or to appear stretched or compressed (Folkestad et al., 2012). An accelerometer speed correction was applied to the raw data in an attempt to remove as many speed irregularities from the image as possible. Data in both wells were normalized to produce both Statically and Dynamically normalized images. Dynamically normalized image data were generated using a sliding normalisation window of 1.0 metre. Statically normalized images were used to correlate lithological or facies changes over the entire well length, while dynamic normalized images were used for detailed comparisons of sedimentary features (Serra, 1989; Goodall et al., 1998; Wilson et al., 2013; Lai et al., 2018). The processed logs were interpreted and calibrated to information interpreted from core by an experienced expert geologist (see Lawrence et al., 2013 for details) to produce the fifteen litho-facies sub-classes that were later grouped into six categories for training. Dynamically normalized data were used to provide input training images. They were displayed as an unwrapped borehole wall image with 0° (North) at the left-hand edge, 180° (south) in the centre and 360° (North) at the right-hand edge, were converted to scalar data and exported to text format (LAS). This resulted in a continuous image without electrode pad or flap gaps in the form of a large matrix. The 192 columns of this matrix represented the reading of each pad/flap button on the FMI tool and the rows represented the data at each depth sample. The nominal vertical resolution of the FMI tool is 0.2 inches or 5 millimetres (Gaillot et al., 2007). Each value in the matrix was a resistivity value representing a single pixel in the image. To generate the training images, all intervals of the same facies category were extracted from the interpreted image based on the manual log interpretation of Lawrence et al. (2013). These sub-images were concatenated to form a single continuous image for each facies category. As a result, six images were produced, each image consisting of the same facies category from various depths of the log where generated. Each image was 192 pixel in width and the height of the image was equal to the sum of the pixels interpreted to be of that facies category as determined from the original manual log classification. To match the network input dimensions, a 192×192 window was randomly and repeatedly cropped out of each image to generate the input data. This method allows for the user to define how many images will be generated for each facies. In this work, we generated 500 images from each of the six facies category images, that is a total of 3000 input training images. The dataset was then randomly selected to produce a training dataset (80% of the images) and a testing dataset (20% of the dataset). Figure 6 shows the input dataset preparation.

5. Proposed Deep learning model

The architecture of our trained CNN requires a single 192×192×3 image as input. However, as described in the previous section, the resistivity image log is a continuous vertical record of the borehole wall. In the case of our training well data (from the CRC-2 well), the size of the image is 192×418627 and the size of the CRC-1 image is 192×680902. To adapt the CNN architecture to this image, we run a sliding 192×192 window over the log image with a stride of one pixel. At each shift of the window, we assign the output of the CNN, which is an array of six probabilities for each class, to each line of the image in that window. By the time a sliding window passes over a specific line in the image, that line is evaluated 192 times. To determine the final class for that line, we sum all the 192 arrays assigned to that line, the result is an array where each element is the sum of all probabilities assigned to the corresponding class. The sum is normalized by dividing each element by 192. The maximum probability is selected as the class for that line. Next a simple function f converts the output to a one-hot coded array. Once the algorithm is finalized, every line in the image is assigned a class. The class of any section of the log is the same as the class associated to the lines that form that section. Results show that this method enhances the continuity and accuracy of class detection and produces better bed boundary detection resolution. In this study we refer to bed boundary detection resolution as the accuracy at which the CNN detects the boundary between two consecutive facies. That is where one facies ends and the next begins. To measure the bed boundary detection resolution, the interpreted boundaries in the test log are compared with the those determined by the CNN in the same log. The proposed method can detect bed boundaries within 0.1 to 0.3 meters. Figure 7 depicts the process.

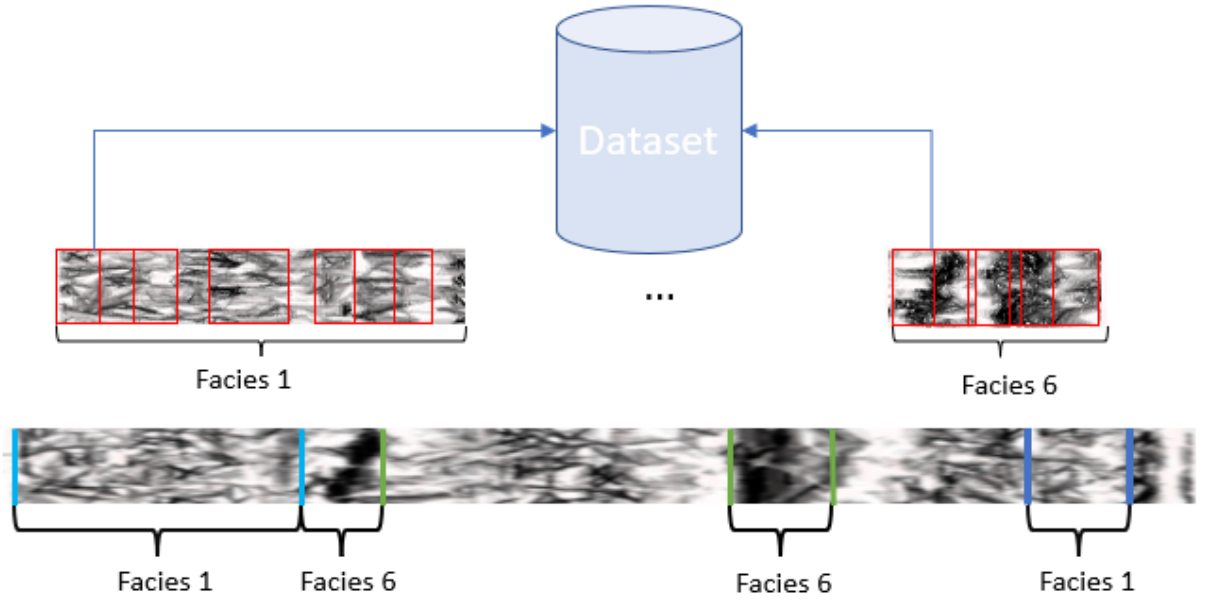

Figure 6: Data preparation methodology. First labelled facies intervals are cropped out of the original log and concatenated together. Next a random 192×192 window is used to generate images.

6. Results & Discussions

Given a set of interpreted resistivity image logs where facies categories have been determined and labelled manually by an expert (Lawrence et al., 2013), we aim to train a CNN to learn the interpreted facies and automatically detect them in a new log. To achieve this, we have trained a CNN on labelled data, and proposed a post-processing methodology to adapt the CNN architecture to evaluate a continuous resistivity image log. Our emphasis in this study is to investigate whether a CNN is able to extract enough features from each facies category to be able to accurately classify new data independent of any other supplementary logs. The logs in this study were generated using Schlumberger's FMI tool. The configuration used for logging was four pads and four flaps each hosting 24 button electrodes, all with diameters of 0.16 inches (Gaillot et al., 2007), which means the borehole wall was covered with 192 sensor buttons at every depth sample point. Therefore, the width of the log is 192 pixels. It is worth mentioning that the FMI tool does not cover the entire borehole circumference. The gaps between the pads/flaps are not sampled. This tool can be run in borehole sizes between 6.25 and 21 inches and the four pad, four flap configuration with 192 buttons covers 80% of the total borehole circumference in an 8-inch borehole (Gaillot et al., 2007). This configuration covers more area in smaller borehole sizes and less in larger ones. Thus, since the facies are generally continuous around the borehole wall, we have chosen to include all 192 pixels in the width direction, in our generated training images. On the other hand, the vertical resolution of the tool is approximately 5 millimetres. In other words, each pixel represents a vertical dimension of 5 millimetres . Most bedding boundaries show some angle of dip but the thickness of sedimentary beds is far greater than the depth sample interval. For reference, the average bed thickness of each facies category in both CRC-1 and CRC-2 wells are presented in Table 1.

Based on the thickness of our interpreted facies and the geological stratigraphy of the region we chose 0.5 meters as the depth/height window of images extracted from the concatenated facies category images that were used for training, which is equal to 192 pixels in the vertical direction. As such, 192×192 pixel dimensions were chosen for input training images. For a similar application using log data from other regions, the height of the input training images may need to be adjusted to suit the thickness of strata and resolution of the log tool in question. Any change to the height of the input image should not be detrimental to the effectiveness of the proposed methodology. When resistivity image logs are processed two normalized images are generated. A statically normalized image and a dynamically normalized image. Sedimentary structures, fractures and bioturbation are more easily observed on the dynamically normalized images,

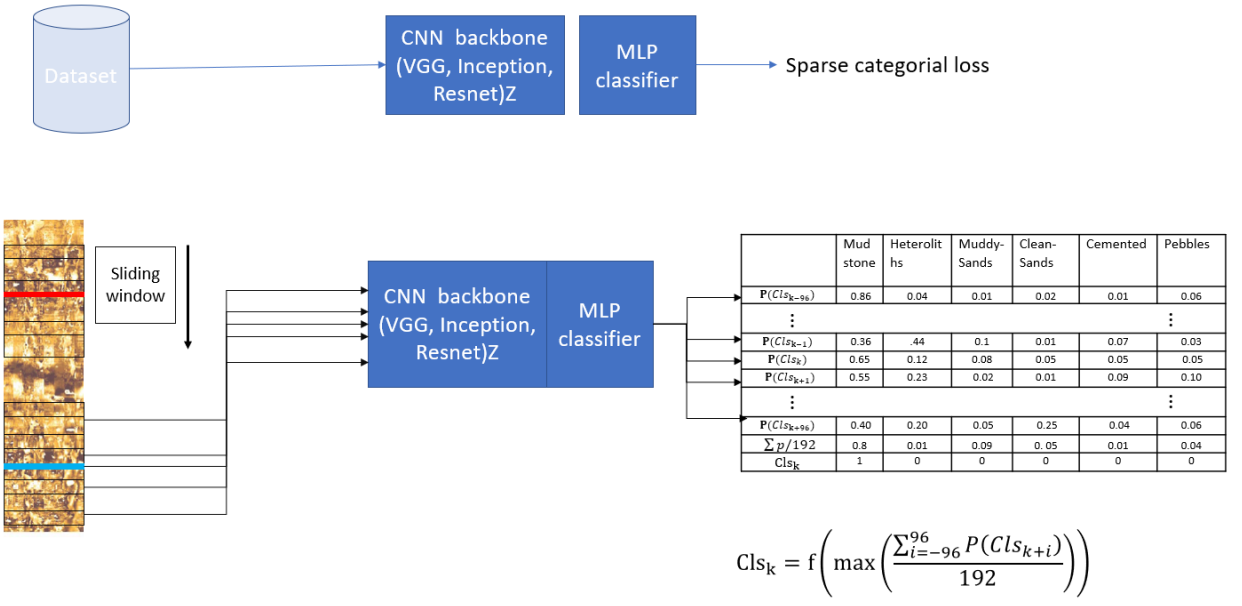

Figure 7: Post processing methodology.

Table 1
Average facies bed thickness in both CRC-1 and CRC-2 wells.

Facies	Average thickness in millimetres	
	Well CRC-1	Well CRC-2
Mudstone	1396.6	1319.6
Heterolithics	791.0	700.9
Muddy-Sands	1132.1	971.3
Clean-Sands	1166.7	767.1
Cemented	346.1	391.9
Pebbles	83.0	114.9

whereas the statically normalized images assist in identifying grain size and general lithology trends (being a better representation of absolute resistivities; (Keeton et al., 2015)). In general, dynamic normalization better accentuates local features such as small-scale sedimentary fabrics, which are key in the training process of the network. For this reason in this study, we have used the dynamic images to generate the input training dataset. The log data that were used to generate the training image data was generally of good quality with occasional noise such as drill bit effects, induced fractures, artefacts of a non-ideal borehole environment, and both stick and pull tool velocity distortion effects. When generating the images, we included the noisy intervals in the training data with the expectation that characteristic stratigraphic patterns would emerge to control the classification process. Since the noise intervals were not significant, this allows the network to learn the noise features and exclude them when classifying intervals (by attributing low and/or non-dominant probabilities for all facies categories). In a first attempt, we designed and trained a CNN from scratch. Limited labelled data and the visual simplicity of the borehole resistivity images led to the assumption that a small CNN trained from scratch would be sufficient to learn the features and the distribution of the input images. However, the testing accuracy of that network only reached a maximum of 58%. The result of this experiment showed that the training images had a complex distribution. Other contributing factors were: (1) lack of distinctive features in the training images, and; (2) the visual similarity of some facies categories in dynamically normalized images. As a result, the training data did not effectively represent the class distributions and the network could not learn and adequately discriminate rich sedimentary features. To address these issues a larger, deeper network is needed. Therefore, we

looked to the concept of transfer learning. In transfer learning an already trained model is repurposed to a new task. This technique takes advantage of the trained feature extracting convolutional layers. New fully connected classifier layers that are adapted to the new task classification categories, will be introduced to the network and trained based on the classification of the new task. Since the network has already learned to extract features, training on the new task will be significantly less computationally intensive and time consuming. A further advantage of this technique is that it can achieve a higher performance with only a small amount of data. To ensure that transfer learning is suitable for our application we tested three different pre-trained networks i.e. VGG-16, Inception-V3 and Resnet-101 on the same input training dataset. These three networks reached a testing accuracy of 67.5% , 78.1% and 82.9% respectively. Details of training and testing accuracy of the three mentioned networks are presented in Table 2.

Table 2
Training and testing Accuracy of three CNN's tested on the dataset.

CNN Name	Training Accuracy	Testing Accuracy
VGG-16	69.1%	67.5%
Inception-V3	99.7%	78.1%
ResNet-101	97.1%	82.9%

The comparison of the results showed that the accuracy trend complies with the trend on other data sets. This compatibility of accuracy trend is confirmation that transfer learning is a suitable technique for our dataset. The ResNet-101 network is the classifier of choice for this study since it produced the highest accuracy. To further evaluate the performance of the above networks, a confusion matrix for each network was developed. A confusion matrix is a visual summary of the performance of a classifier. The confusion matrices for the three networks trained on our six facies categories is shown in Figure 8. The rows and columns of the matrix are the six facies categories organised in the same order. The diagonal elements denote the correctly classified outputs. The off-diagonal elements show the incorrectly classified outcomes. The three matrixes show a similar learning pattern. Facies category '5' – Cemented, and '6' – Pebbles, are quite accurately detected by all the network s. The Cemented class returns a resistivity that may often be higher than the FMI tool capacity, thereby saturating the measurement tool. The image is therefore distinct and can be easily learned and detected by the network. An example image of the Cemented category is shown in Figure 9. Facies '3' and '4' are misclassified most often. Facies '3' is misclassified as Facies '4' 20% of the time, while Facies '4' is misclassified as Facies '3' 10% of the time in the ResNet-101 confusion matrix. These figures are higher in the VGG-16 and Inception-V3 matrices respectively, which is in line with the testing accuracy of these networks as presented in Table 2.This is due to the similarity of the sedimentary textures revealed in these two facies categories by dynamically normalized images. Conceptually, both categories represent sandstone with a mixture of mud/clay. The level of mud in the rock determines the class. Therefore, occasionally these two classes can be hard to distinguish form each other though also using the corresponding statically normalized image may help. Examples of two labelled intervals of Facies '3' and '4' are shown in Figure 10. It can be seen from the matrices that Facies '1' and '2' are occasionally misclassified as either Facies '3' or '4'. Clearly, this occurs more often for VGG-16 and Inception-V3 than for ResNet-101. These misclassifications occur because these facies are visually similar in the log even though they are geologically categorized in different categories. This limitation is due to using image logs alone which was a research question in our investigation. CNNs are able to classify visually similar categories from image logs better than the human eye, but to enhance accuracy of categorizing these facies, other data types such as supplementary 1D logs of the conventional wireline suite should be incorporated as part of the training process. This is not within the scope of the current study and therefore has not been addressed. The ultimate goal of the trained CNN is to classify the learned facies categories in a full-length resistivity image log. We have proposed a post-processing methodology to adapt the architecture of a CNN to this task. The post-processing technique is applied to the output of ResNet-101 since it produced the best transfer learning accuracy. The simplest strategy would have been to slide a 192×192 window over the resistivity image with a user-defined stride and evaluate a single image after each shift. However, given the accuracy of our best network ResNet stands at 82%, this methodology would return incorrect results 18% of the time to the post-processing step. Results were inconsistent even in thick intervals of the same facies category with incorrect designations produced too often, requiring human intervention to correct and finalize. Another imperfection of this method is the very low resolution of bed boundary detection. A proposed solution would be a moving window with a stride of one. After each shift the outcome of the CNN classifier which is an array with 6 probability elements is

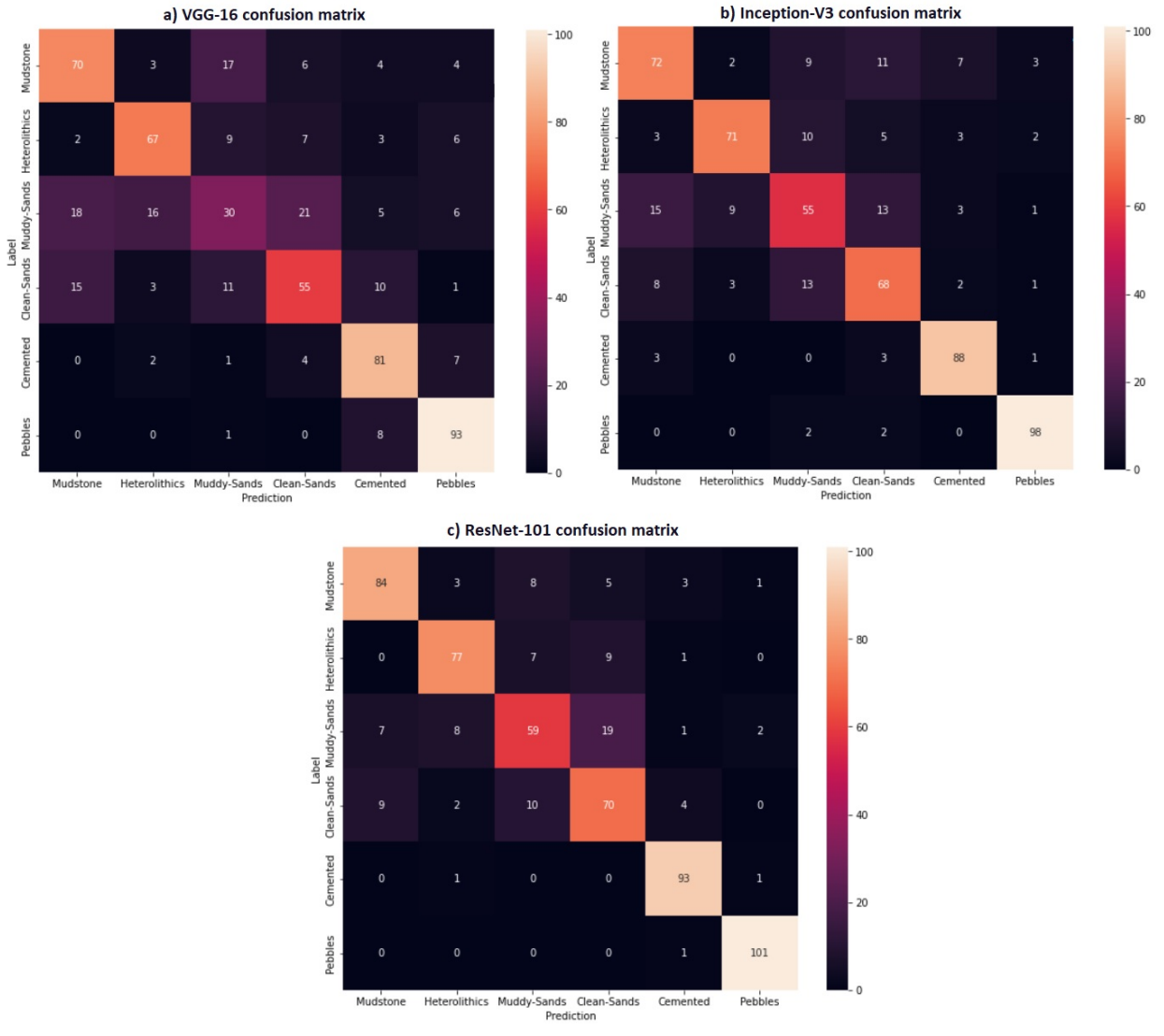

Figure 8: Confusion matrices. a) VGG-16 b)Inception-V3 c) ResNet-101.

assigned to each horizontal line in the 192×192 sliding window. As the window slides over a specific line in the image log, the line is evaluated 192 times and assigned 192 output arrays. As per the testing accuracy of the network, 82% of category designations for this line are correct, therefore the element wise sum of all the arrays will result in an array where the correct category has the highest probability and is thereby assigned to the line. Repetition of this method for every line in the image produces a significant increase in the predictive accuracy and consistency of log facies categorisation. Furthermore, bed boundary resolution is increased. We tested the network with this post-processing methodology using lower quality image log data acquired at and manually interpreted for CRC-1. This represents a 'whole-well' test that produced an accuracy of facies categorisation of 92% and a bed boundary resolution that ranges between 0.1 to 0.3 metres.

7. Conclusion

In this study, we have presented a methodology using Convolutional Neural Networks (CNN) and transfer learning to augment/automate facies classification of resistivity image logs. We have proposed a data preparation methodology that allows the user to generate more training data from the reference log. We have discussed challenges in the

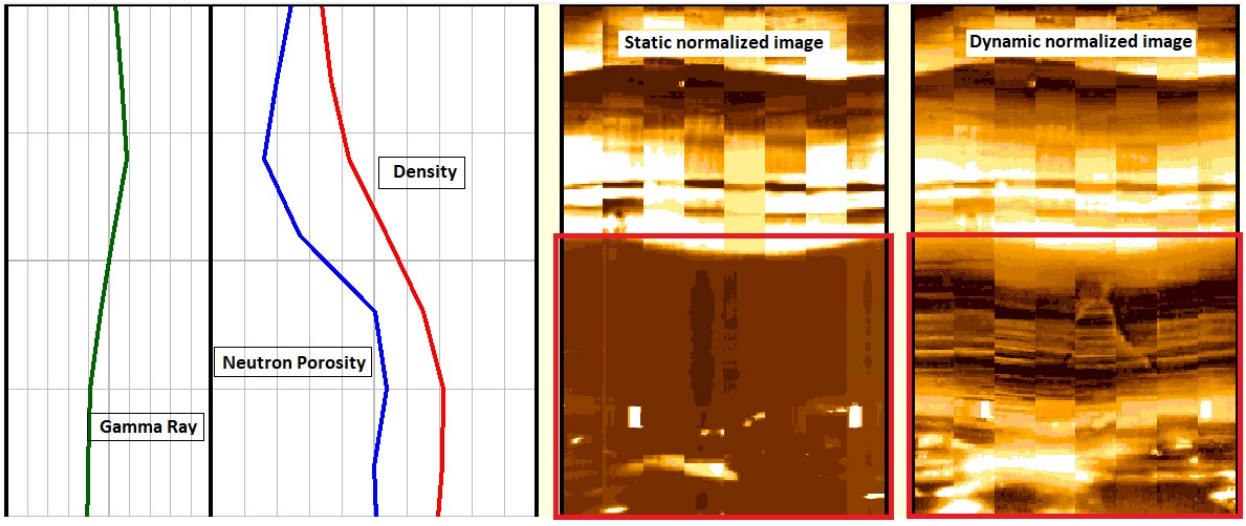

Figure 9: Example of Facies 5, Cemented.

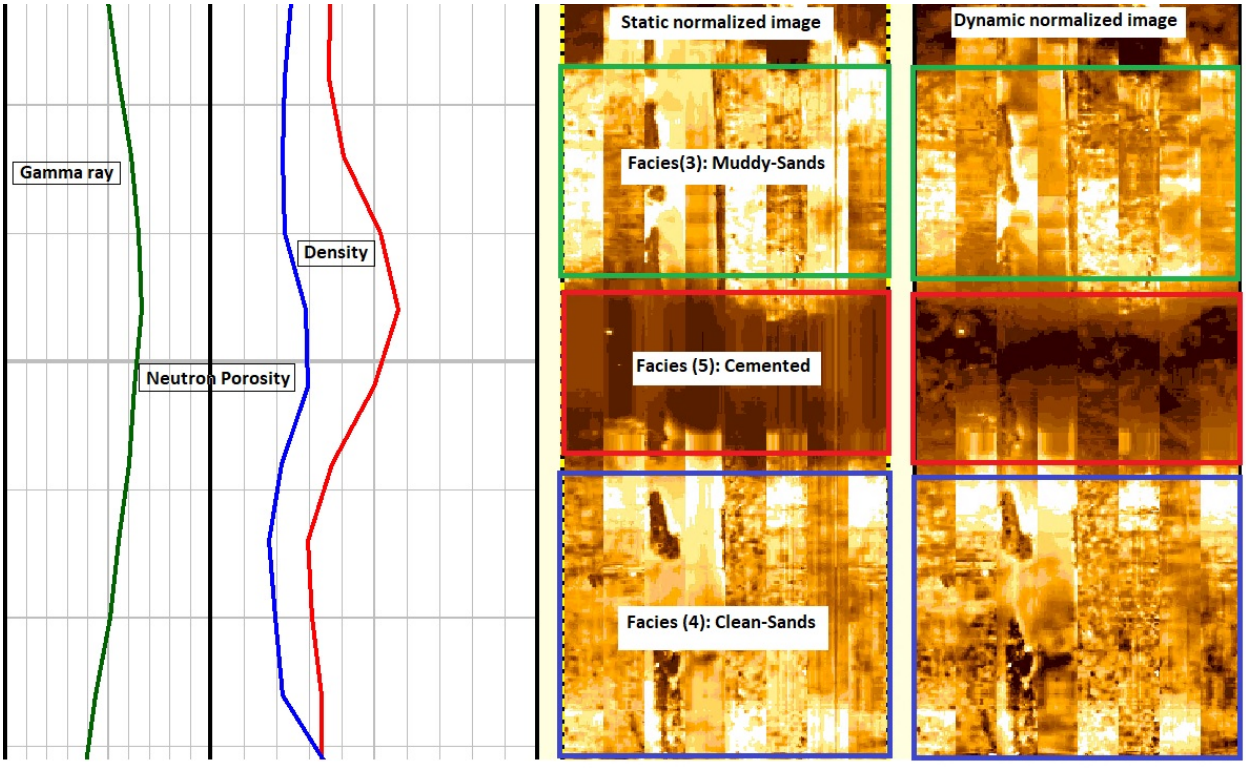

Figure 10: Example of facies 3 (Muddy-Sands) and faces 4 (Clean-Sands). some intervals of these facies are visually quite similar.

application of machine learning to facies detection and classification of image log data such as complexity of dataset distribution and inadequate labelled data, subsequently we propose transfer learning to resolve these issues. We investigated three different CNNs and showed that transfer learning is a suitable technique for our dataset. Finally, we have proposed a post-processing step to increase the accuracy of facies classification of a continuous log that also enhances bed boundary detection/resolution. Our objective in this study was to determine whether CNNs are able to detect fa-

cies classes using image log alone. The results show that in the absence of any other information and using image log only the trained network can detect facies categories with a testing accuracy of 82% and the proposed post-processing method increases the final categorization accuracy even further. We conclude that such a classification system would benefit from the use of supplementary logs. An alternative might be to combine statically and dynamically normalized images in the training dataset. The former could act as a baseline to discriminate cases of textural similarity in the latter. Otherwise, 1D wireline log data such as the 'porosity logs' (sonic velocity, bulk density, neutron porosity) might be used to provide an alternative and somewhat independent baseline discriminator for training the CNN classifier.

8. Acknowledgments

The authors would like to acknowledge and thank CO2CRC for making this research possible by providing data and funding.

Code availability section

The code used in this study, is written in Python and requires Pytorch library for deep learning. It is available for downloading at the link: https://github.com/roozbeh442/CNN-augmented-FMI-log-inerp.git

References

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L., 2021. Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. Journal of big Data 8, 1–74.

Ameen, M.S., 2014. Fracture and in-situ stress patterns and impact on performance in the khuff structural prospects, eastern offshore saudi arabia. Marine and Petroleum Geology 50, 166–184.

Boyd, G., Gallagher, S., 2001. The sedimentology and palaeoenvironments of the late cretaceous sherbrook group in the otway basin .

Brekke, H., MacEachern, J.A., Roenitz, T., Dashtgard, S.E., 2017. The use of microresistivity image logs for facies interpretations: An example in point-bar deposits of the mcmurray formation, alberta, canada. AAPG Bulletin 101, 655–682.

Bunch, M., 2014. A live test of automated facies prediction at wells for co2 storage projects .

Bunch, M., Daniel, R., Lawrence, M., Browne, G., Menacherry, S., Dance, T., Arnot, M., 2012. Multi-scale characterisation of the paaratte formation, otway basin, for co2 injection and storage. The APPEA Journal 52, 664–664.

Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., Miao, Y., 2021. Review of image classification algorithms based on convolutional neural networks. Remote Sensing 13, 4712.

Folkestad, A., Veselovsky, Z., Roberts, P., 2012. Utilising borehole image logs to interpret delta to estuarine system: A case study of the subsurface lower jurassic cook formation in the norwegian northern north sea. Marine and Petroleum Geology 29, 255–275.

Gaillot, P., Brewer, T., Pezard, P., Yeh, E.C., 2007. Borehole imaging tools—principles and applications. Scientific Drilling 5, 1–4. doi:doi:10.2204/iodp.sd.5.07S1.2007.

Goodall, T., Møller, N., Rønningsland, T., 1998. The integration of electrical image logs with core data for improved sedimentological interpretation. Geological Society, London, Special Publications 136, 237–248.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., et al., 2018. Recent advances in convolutional neural networks. Pattern recognition 77, 354–377.

Gupta, K.D., Vallega, V., Maniar, H., Marza, P., Xie, H., Ito, K., Abubakar, A., 2019. A deep-learning approach for borehole image interpretation, in: SPWLA 60th Annual Logging Symposium, OnePetro.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Ja'fari, A., Kadkhodaie-Ilkhchi, A., Sharghi, Y., Ghanavati, K., 2012. Fracture density estimation from petrophysical log data using the adaptive neuro-fuzzy inference system. Journal of Geophysics and Engineering 9, 105–114.

Keeton, G.I., Pranter, M.J., Cole, R.D., Edmund, R., 2015. Stratigraphic architecture of fluvial deposits from borehole images, spectral-gamma-ray response, and outcrop analogs, piceance basin, coloradostratigraphic architecture of fluvial deposits from borehole images. AAPG Bulletin 99, 1929–1956.

Kosari, E., Ghareh-Cheloo, S., Kadkhodaie-Ilkhchi, A., Bahroudi, A., 2015. Fracture characterization by fusion of geophysical and geomechanical data: a case study from the asmari reservoir, the central zagros fold-thrust belt. Journal of Geophysics and Engineering 12, 130–143.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. Communications of the ACM 60, 84–90.

Lai, J., Wang, G., Wang, S., Cao, J., Li, M., Pang, X., Han, C., Fan, X., Yang, L., He, Z., et al., 2018. A review on the applications of image logs in structural analysis and sedimentary characterization. Marine and Petroleum Geology 95, 139–166.

Lawrence, M., Arnot, M., Browne, G., Bunch, M., Dance, T., 2013. Geological interpretation of core and wireline data from otway project wells crc-1 and crc-2. report no. rpt12-3928. Cooperative Research Centre for Greenhouse Gas Technologies, Canberra, Australia. 76 pp. .

LeCun, Y., Jackel, L.D., Bottou, L., Cortes, C., Denker, J.S., Drucker, H., Guyon, I., Muller, U.A., Sackinger, E., Simard, P., et al., 1995. Learning algorithms for classification: A comparison on handwritten digit recognition. Neural networks: the statistical mechanics perspective 261, 2.

Lefranc, M., Bayraktar, Z., Kristensen, M., Driss, H., Le Nir, I., Marza, P., Kherroubi, J., 2021. Deep-learning-based automated sedimentary geometry characterization from borehole images. Petrophysics 62, 636–650.

Lima, L., Bize-Forest, N., Evsukoff, A., Leonhardt, R., 2019. Unsupervised deep learning for facies pattern recognition on borehole images, in: Offshore Technology Conference Brasil, OnePetro.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F.E., 2017. A survey of deep neural network architectures and their applications. Neurocomputing 234, 11–26.

Liu, X., Li, B., Li, J., Chen, X., Li, Q., Chen, Y., 2021. Semi-supervised deep autoencoder for seismic facies classification. Geophysical Prospecting 69, 1295–1315.

Lorente, Ò., Riera, I., Rana, A., 2021. Image classification with classic and deep learning techniques. arXiv preprint arXiv:2105.04895 .

Matthew Zeiler, D., Rob, F., 2014. Visualizing and understanding convolutional neural networks, ECCV.

Mishra, A., Pajank, L., Haese, R.R., 2019. High resolution characterization of lithological heterogeneity of the paaratte formation, otway basin (australia), a coastal to shallow-marine deposit. Geosciences 9, 278.

Morton, J., Hill, A., Parker, G., Tabassi, A., 1994. Towards a unified stratigraphy for the otway basin, in: Finlayson, DM (Compiler), NGMA/PESA Otway Basin Symposium, Melbourne, 1994. Extended abstracts. Australian Geological Survey Organisation Record, pp. 49–54.

Nie, X., Zou, C., Pan, L., Huang, Z., Liu, D., 2013. Fracture analysis and determination of in-situ stress direction from resistivity and acoustic image logs and core data in the wenchuan earthquake fault scientific drilling borehole-2 (50–1370 m). Tectonophysics 593, 161–171.

Orhan, A.E., Pitkow, X., 2017. Skip connections eliminate singularities. arXiv preprint arXiv:1701.09175 .

Sainath, T.N., Kingsbury, B., Mohamed, A.r., Dahl, G.E., Saon, G., Soltau, H., Beran, T., Aravkin, A.Y., Ramabhadran, B., 2013. Improvements to deep convolutional neural networks for lvcsr, in: 2013 IEEE workshop on automatic speech recognition and understanding, IEEE. pp. 315–320.

Serra, O., 1989. Formation microscanner image interpretation: Schlumberger educational services. Houston, Texas .

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9.

Tang, J., 2018. Intelligent Mobile Projects with TensorFlow: Build 10+ Artificial Intelligence Apps Using TensorFlow Mobile and Lite for IOS, Android, and Raspberry Pi. Packt Publishing Ltd.

Willcox, J., Stagg, H., 1990. Australia's southern margin: a product of oblique extension. Tectonophysics 173, 269–281.

Wilson, M.E., Lewis, D., Holland, D., Hombo, L., Goldberg, A., et al., 2013. Development of a papua new guinean onshore carbonate reservoir: a comparative borehole image (fmi) and petrographic evaluation. Marine and Petroleum Geology 44, 164–195.

Woollands, M., Wong, D., 2001. Petroleum atlas of victoria, australia. The State of Victoria, Department of Natural Resources and Environment 208.

Zaccone, G., Karim, M.R., 2018. Deep Learning with TensorFlow: Explore neural networks and build intelligent systems with Python. Packt Publishing Ltd.

5. A variability aware GAN for improving spatial representativeness of discrete geobodies

Roozbeh Koochak, Mohammad Sayyafzadeh, Ali Nadian Ghomsheh, Mark Bunch, Manouchehr Haghighi

Journal of Computers and Geosciences, 2022

Statement of Authorship

Title of Paper	**A Variability Aware GAN for Improving Spatial Representativeness of Discrete Geobodies**
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	**Koochak, R., Sayyafzadeh, M., Nadian, A., Bunch, M. and Haghighi, M., 2022. A variability aware GAN for improving spatial representativeness of discrete geobodies. Computers & Geosciences, 166, p.105188.**

Principal Author

Name of Principal Author (Candidate)	**Roozbeh Koochak**		
Contribution to the Paper	**Conceptualization, development of algorithm, experimentation, analysis, coding, original manuscript, corresponding author duties**		
Overall percentage (%)	**60%**		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	**17/05/2023**

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mohammad Sayyafzadeh		
Contribution to the Paper	Conceptualisation, theory of regularisation, analysis of results, discussions, revisions, supervision		
Signature		Date	18/05/2023

Name of Co-Author	Ali Nadian Ghomsheh		
Contribution to the Paper	Machine Learning network design and parameter optimization, discussions		
Signature		Date	23/05/2023

Name of Co-Author	Mark Bunch		
Contribution to the Paper	Discussions, revisions, supervision		
Signature		Date	19/05/2023

Name of Co-Author	Manouchehr Haghighi		
Contribution to the Paper	Discussions, revisions, supervision		
Signature		Date	18/05/2023

Please cut and paste additional co-author panels here as required.

Computers & Geosciences 166 (2022) 105188

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

A variability aware GAN for improving spatial representativeness of discrete geobodies

Roozbeh Koochak [a,*], Mohammad Sayyafzadeh [a,**], Ali Nadian [b], Mark Bunch [a], Manouchehr Haghighi [a]

[a] Australian School of Petroleum and Energy Resources, University of Adelaide, Australia
[b] Cyberspace Research Institute, Shahid Beheshti University, Tehran, Iran

ARTICLE INFO

Dataset link: https://github.com/roozbeh442/Variability-Aware-GAN.git

Keywords:
Generative Adversarial Networks
Mode collapse
Geospatial variability
Stochastic modeling
Dimensionality reduction
Non-Gaussian geostatistics

ABSTRACT

Generative Adversarial Networks (GAN) have shown great potential in not only producing acceptable realizations of geologically complex models but also successfully reparametrizing them. Training GANs is quite challenging. One such challenge is mode collapse. When generating realizations of spatial property, mode collapse causes reduction in variability, compared to the input training dataset, and thus, the realizations become spatially biased at specific locations. To address this issue, we developed a new GAN architecture where a regularization term is introduced to maintain the variability and reduce mode collapse. This is achieved by using a probability map to evaluate variability and spatial bias of generated realizations and modifying the GAN loss function to minimize this bias. We applied the new architecture to a binary channelized permeability distribution and compared the results with those generated by Deep Convolutional GAN (DCGAN) and Wasserstein GAN with gradient penalty (WGAN-GP). Our results show that the proposed architecture significantly enhances variability and reduces the spatial bias induced by mode collapse, outperforming both DCGAN and WGAN-GP in the application of generating subsurface property distributions.

1. Introduction

Geological or reservoir modeling is an intricate task requiring multiple sources of information to be combined. The modeling process usually involves the characterization of reservoir properties in extended areas using hard data from few locations. However, these models can never fully describe the reservoir since some properties of the geological system will be unknown or uncertain. The sources of uncertainty are, among others, subsurface complexity and heterogeneity, incorrect geological concept and critically sparse data sampling. Lack of hard data is an intrinsic issue that hinders accurate modeling in earth sciences. To make best use of these models, uncertainty in each included data type must be estimated and propagated through the process of generating the model. Understanding uncertainty in reservoir models usually involves generating multiple realizations of the geological model that are both geologically sound (realistic) and statistically faithful. Governing equations are then solved for each realization to get an understanding of the probability distribution of the model response (Chan and Elsheikh, 2020). Uncertainty can be reduced by conditioning the models to hard data and nonlinearly correlated production data. In this process also referred to as inversion, reparametrizing the model can significantly reduce the computational cost. Therefore, being able to reparameterize models is an important feature.

Stochastic simulation methods are the most widely used methods for generating realizations. These methods are generally divided in two categories: Object-based methods and Pixel-based (sometimes also known as Point-based) methods (Bai and Tahmasebi, 2020). Object-based simulation methods parametrize geological geometries and use Boolean object-based algorithms to place these objects on the simulated area. These methods maintain geological realism better than other methods, however, their major limitations is that conditioning the realizations to hard data is challenging (Bai and Tahmasebi, 2020; Strebelle, 2002). Pixel-based methods, such as Sequential Gaussian Simulation and Truncated Gaussian Simulation, conduct the simulation pixel by pixel. Using these techniques, it is complicated to produce curvilinear geometries such as sinuous channels, (Tahmasebi, 2018), inherent in many geological facies structures. Considering two point statistics is inadequate for reproducing complex geological structures (Marini et al., 2018), as a result realizations produced by these

* Corresponding author.
** Corresponding author.
E-mail addresses: roozbeh.koochak@adelaide.edu.au (R. Koochak), mohammad.sayyafzadeh@adelaide.edu.au (M. Sayyafzadeh).

https://doi.org/10.1016/j.cageo.2022.105188
Received 16 June 2021; Received in revised form 11 April 2022; Accepted 2 July 2022
Available online 14 July 2022

methods are unrealistic when simulating such geological systems. Truncated Pluri-Gaussian (Le Loch et al., 1994) is a pixel-based method introduced to overcome some of these limitations and better preserve prior geological understanding (Astrakova and Oliver, 2015). The basic idea in this method is to simulate several continuous gaussian fields and truncate them to produce a categorical variable. The main difficulty in applying this technique is the inference of the variogram models for the underlying multi-gaussian functions (Mariethoz et al., 2009). Multiple points statistic (MPS) methods were introduced to address the problem of generating realistic models of complex geological structures. Using multiple points requires a large number of samples which are usually not available in earth sciences (Marini et al., 2018). This extra information can be derived from a conceptual image contributed by the geologist known as the Training Image (TI). MPS methods can generate realizations that honor the training image. An example is Single Normal Equation Simulation (SNESIM). This algorithm scans the TI and stores the probability of all pattern occurrences in a search tree, the probabilities are then retrieved based on existing data to generate realizations (Strebelle and Journel, 2001). Inversion using these methods (Jäggli et al., 2017; Laloy et al., 2016) is computationally expensive. There are reparameterization techniques available, however the models provided using these techniques do not agree well with the TI (Laloy et al., 2018).

Recently, a large body of research has applied Deep Generative Models (DGM) to generate realizations of a geological model. DGMs are a class of Neural Networks (NN) that can approximate high-dimensional probability distributions when trained successfully, given sufficient samples from the desired distributions. The trained model can then be used to generate realizations from the underlying distributions (Ruthotto and Haber, 2021) and they have the added value of re-parametrizing the realizations, compared to MPS methods. Variational Auto-encoders (VAE) are an example of generative models that use variational Bayesian inference to approximate the probability density. VAEs consist of an encoder and a decoder. The encoder takes in sample training data, and passes it through layers with decreasing dimensionality, mapping the data to a latent space. The decoder then samples from the latent space and reproduces the initial data (Sami and Mobin, 2019). Inversion using VAEs has been reported to produce superior results compared to MPS-based inversion (Laloy et al., 2017). In combination with the ensemble smoother with multiple data assimilation to history match production data, Canchumuni et al. (2021) reported that the trained VAE resulted in "noisy" facies reconstructions. In general VAEs have a lower generative accuracy compared to Generative Adversarial Networks (Lopez-Alvis et al., 2021) and are prone to fail at learning intractable probability distributions (Sami and Mobin, 2019). Among DGMs, Generative Adversarial Networks (GAN) have shown great potential for generating realizations of a geological model. The basic GAN created by Goodfellow et al. (2014) consists of two neural networks, Generator and Discriminator, trained end-to-end to produce realizations of a desired distribution. Since the introduction of GAN in 2014, many GAN architectures have been developed and tailored to the specific needs of their application.

There are three main components to generating satisfactory geological realizations: 1. The realizations need to be geologically sound and visually acceptable, 2. they must preserve the statistics of the structure or reference training image (that includes univariate, two-point and multiple-point statistics) and 3. they must maintain the variability of the structure. Honoring hard data is an important feature of geological realizations, however, the focus of this study is on unconditioned models. Most of the literature documenting studies that have utilized GANs in earth science applications have reported that GANs achieve the first two components quite well, while due to mode collapse the variability tends to decrease (Chan and Elsheikh, 2020, 2019b; Azevedo et al., 2020). Training a GAN minimizes the divergence between the probability Density Function (PDF) of the training data and the PDF that the generator samples. In this process the condition where the generator captures few major modes of the input training data and ignores many small modes is referred to as mode collapse (Bang and Shim, 2018). This results in reduction in variability of the realizations that are generated. In other words, the realizations look similar in most areas of the model with minor differences. This effect will severely disturb uncertainty quantification or History matching as the generated realizations will be biased towards certain areas in the model. Some work has addressed this issue using a Wasserstein GAN, which theoretically reduces mode collapse (Arjovsky et al., 2017). In general, there are two forms of variability that need to be maintained: (1) the within-realization variability and (2) the between-realization variability (Tan et al., 2014). The within-realization variability refers to each individual realization capturing the statistics (uni-variate, two-point, multiple-point, etc.) of the geological structure (training data set), while the between-realization variability refers to the dissimilarity of the generated realizations. Mode collapse refers to the output realizations not being diverse. Therefore, variability in this paper always refers to dissimilarity between generated realizations.

In this study, to improve the variability of generated geological realizations we propose a new GAN architecture. Similar to conventional GANs, our architecture consists of a Generator and Discriminator. A regularization term has been added to the Generator loss to improve variability and reduce spatial bias. To evaluate our proposed architecture, we train a Deep Convolutional GAN (DCGAN) and a Wasserstein GAN with gradient penalty (WGAN-GP) to generate realizations based on the famous Strebelle benchmark reference training image (Strebelle and Journel, 2001) of a stationary channelized fluvial system. We then compare the results with those generated by our proposed architecture. We introduce the new architecture using this binary benchmark image since it is well known and frequently used in the literature, but the proposed methodology can potentially be extended to be applied to multiple-facies realizations. The rest of this paper is structured as follows: in Section 2, we describe some related work. A short background of GANs and their training process along with brief description of DCGAN, WGAN-GP and mode collapse in presented in Section 3. Our proposed architecture is discussed in Section 4 followed by the experimental setup in Section 5. Results and discussion, then conclusions of the study follow, together with ideas for future work.

2. Related work

In the field of geosciences several studies have been conducted to evaluate the effectiveness of GANs for generating realizations. Some GAN variants have been proposed to improve the variability of generated realizations and conditioning to hard data. Note, to prevent confusion between the geological conceptual Training Image, and GAN input training data we, from here on, refer to the conceptual geological image as Reference Training Image (RTI). To emphasize, the RTI is a conceptual image of expected spatial structures and are usually built based on prior information (Meerschman et al., 2013). The GAN input training data are a set of realizations assembled based on the RTI and used to train the GAN.

GANs have been used to generate realistic stochastic samples of porous media (Mosser et al., 2017, 2018b). In Mosser et al. (2018b) a modified version of the DCGAN (Radford et al., 2015) was trained on randomly extracted images from a micro-CT. They reported that computed two-point statistics and effective properties showed excellent agreement between the GAN results and segments of the micro-CT image. However, they observed far less variation in the generated samples compared to the input training dataset. Mode collapse was mentioned as one of possible reasons. GANs have also been applied to seismic data. A Cycle-GAN was used to perform stratigraphic seismic inversion based on a velocity model (Mosser et al., 2018c). In Mosser et al. (2020), a DCGAN was trained to parametrize geological heterogeneities and was combined with a numerical solution of the acoustic inverse problem using the adjoint method. Mode collapse was

discussed in this work and use of alternative networks that handle mode collapse more efficiently was envisioned. GANs were also used to reconstruct models retrieved by iterative geostatistical seismic inversion (Azevedo et al., 2020). Both a DCGAN and WGAN (Arjovsky et al., 2017) were trained on two types of dataset, a set of binary facies and a continuous (P-wave propagation velocity) dataset where an infill painting methodology (Yeh et al., 2017) was used for reconstruction of images. Histograms, two-point variograms and probability maps were used to evaluate the quality of the results. The authors reported a reduction in the variability of the results when generating conditioned realizations compared to unconditioned ones. Re-parametrization is an important feature of GANs. In Chan and Elsheikh (2019b) a WGAN was used to reparametrize geological realizations. The results were visually acceptable, and the flow statistics induced by the generated realizations were similar to the reference. A similar approach was used in Chan and Elsheikh (2020) to re-parametrize conditioned and unconditioned geological realizations. A GAN architecture was used in Laloy et al. (2018) to perform an MPS-based geostatistical inversion for parameter estimation. They proposed a 3D extension to the original 2D spatial GAN (SGAN) (Jetchev et al., 2016) and incorporated it in a Markov chain Monte Carlo (McMC). The authors reported that the generated realizations captured the multiple point statistics of the RTI, and the low dimensional representation allowed for efficient probabilistic inversion. Multiple studies have used GANs to generate realizations conditioned to hard data (Chan and Elsheikh, 2019a; Dupont et al., 2018; Mosser et al., 2018a; Zhang et al., 2021). While Dupont et al. (2018) used a modified version of semantic inpainting methodology to condition realizations, Chan and Elsheikh (2019a) proposed an inference network added to the Generator to further refine the generator function space so that the generated samples are conditioned to hard data. Zhang et al. (2021) used a U-net architecture to generate geological realizations conditioned to hard data. To increase the variability of the generated models they adopted a methodology where a loss term was formulated to maximize the distance between generated images with respect to corresponding latent vectors (Yang et al., 2019). The distance metric used by this paper was L1 norm. In addition to generating different geological realizations, Razak and Jafarpour (2020b,a) and Mosser et al. (2019) have gone a step further by calibrating a reservoir model using non-linear production data.

3. Background

3.1. Generative Adversarial Networks (GANs)

GANs are a family of deep-learning-based generative models where the paradigm of unsupervised learning is used in their training process. A brief explanation of supervised vs unsupervised is provided to better explain the idea behind a GAN model. In machine learning, given a set of inputs X and outputs Y as labels; a typical problem is to learn the parameters of a model that maps X → Y. The mapping is usually learned by correcting the model parameters to produce correct labels for the input set. The correction process is governed, or supervised by updating model parameters, usually in an iterative process, so as to minimize the error between model outputs and the true labels. In cases where the output set is not provided, the paradigm of unsupervised learning can be used. Given only set X, the goal is to find patterns in the data. This is a much less well-defined problem, since the desired patterns to look for are not known (Murphy, 2012). In order to make predictions about data in an unsupervised manner, models can be trained to summarize the distribution of input variables which can in turn create or generate new samples within the input distribution.

The architecture has two sub-models: Generator and Discriminator (Goodfellow et al., 2014). Generator is used to generate new plausible examples from the problem domain. The Discriminator is the sub-model that is used for training the Generator. The architecture has the advantage that after being trained, the sub-models can be used as standalone models for data generation or classification (Radford et al., 2015). The training process of a GAN is based on a game theoretic scenario in which the generator competes against an adversary. The generator network directly produces samples from a fixed-length random vector, referred to as latent space, while the discriminator network, the adversary, attempts to distinguish between samples drawn from the training data (real data), and the generated samples (Goodfellow et al., 2016).

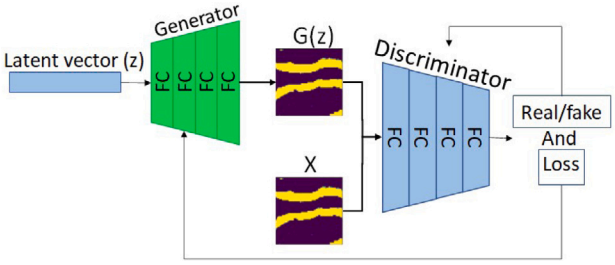

Fig. 1. Vanilla GAN architecture.

Fig. 1 illustrates the basic workflow of a vanilla GAN. Based on the figure, latent vector z, is fed to the Generator G to generate sample $G(z)$. FC stands for Fully Connected layers. It is conventional to feed a batch of latent vectors to generate a batch of samples every time. The Discriminator is trained to classify $G(z)$ and real samples, X. This is accomplished by maximizing the probability assigned to real and generated data at the output of the discriminator. From implementation point of view, this is accomplished by defining a binary classification task with labels 0 and 1 for generated and real samples. Hence, the Discriminator seeks to minimize the average binary cross entropy loss (Goodfellow et al., 2014) denoted as:

$$\mathscr{L}_D^{GAN} = -\mathbb{E}_{x\sim p_d}\left[log\left(D(x)\right)\right] - \mathbb{E}_{G(z)\sim p_g}\left[log\left(1 - D\left(G(z)\right)\right)\right] \quad (1)$$

where $\mathbb{E}$ is the expected value of sample drawn from the real data p_d or generated data p_g. $D(.)$ Denotes the output of the Discriminator which is the probability assigned to a sample being drawn from the real data. After updating the Discriminator, the Generator is trained in the GAN architecture, while keeping the discriminator frozen. In this process, a batch of samples is generated and labeled 0. The generated batch along with the real samples labels as 1s are fed to the discriminator. Here, if good samples are generated, the Discriminator produces small loss values and vice versa. The loss is then back propagated in the GAN in order to update the Generator. This process is repeated until no further improvements of the GAN is possible. Putting this into notation, the Generator loss can be represented as:

$$\mathscr{L}_G^{GAN} = \mathbb{E}_{G(z)\sim p_g}\left[log\left(1 - D\left(G(z)\right)\right)\right] \quad (2)$$

Having described the training process of GAN sub-models, the training of GAN can be summarized as a minimization–maximization (minmax) process, achieved via a minmax loss (Ko and Lin, 1995; Thekumparampil et al., 2019). The minimax GAN loss refers to the minimax simultaneous optimization of the discriminator and generator models. The Discriminator is trained to maximize the probability of assigning correct labels to samples from both the training data and generated data. At the same time, the Generator is trained to minimize $\log(1 - D\left(G(z)\right))$ (Goodfellow et al., 2014). In general, the minmax refers to an optimization strategy in two-player turn-based games for minimizing the loss or cost for the worst case of the other player. Ultimately equilibrium is achieved when $p_d = p_g$ and $D(.) = 1/2$. At this point the Discriminator is unable to distinguish between real and fake samples (Goodfellow et al., 2014).

After training, the Generator can map a random sample from the latent space into a sample from the input training data space. Therefore, new but different samples from the input training data are generated.

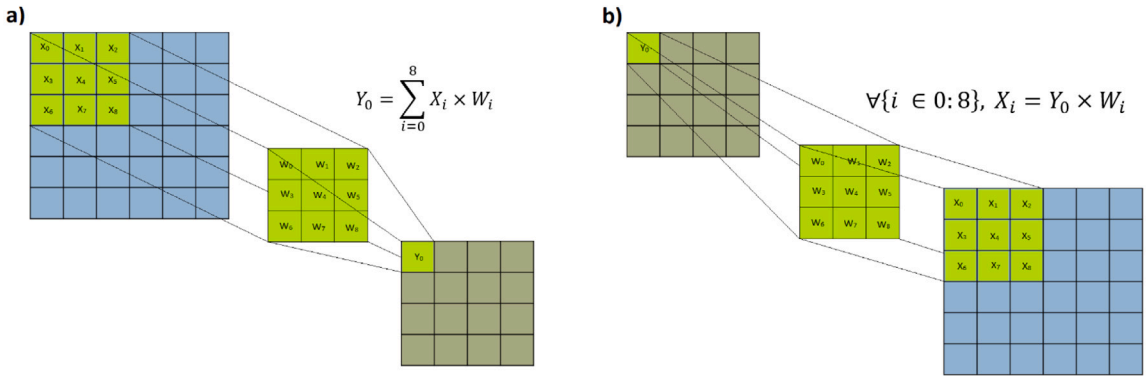

Fig. 2. Convolutional and Transposed convolutional layers. (a) In a convolutional layer the weights of the filter W map the input to a lower dimension Y. (b) in a transposed convolutional layer the filter with weights W maps its input to a higher dimension.

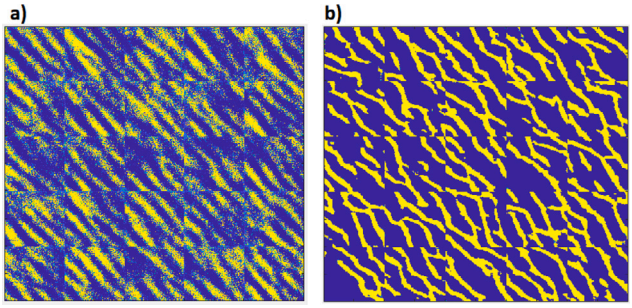

Fig. 3. Example of severe mode collapse. The image shows 25 tiled samples generated by a GAN. (a) The images are very similar and channels occur in common locations. (b) Example of how the realizations should look like with reduced mode collapse.

3.2. Deep convolutional GAN (DCGAN)

DCGAN, introduced by Radford et al. (2015) is a direct architecture extension of the GAN, designed for image related tasks, which replaces the fully connected layers of the Generator and Discriminator with convolutional layers. This is the preferred architecture when dealing with image data, as convolutional layers can extract information from images taking into account the spatial structure of the image. The layers in the generator usually try to map the input latent space into a larger space, hence transposed convolutional layers are used. On the other hand, the Discriminator tries to map both generated and real samples into a lower dimension, hence convolutional layers are considered. In both types of layer, the objective is to train convolutional filters that can extract meaningful information from the previous layer, so that the trained network can output the desired model. Fig. 2 illustrates schematically how convolutional and transposed convolutional layers operate. In the convolutional setup, the weights of the filter W map the input to a lower dimension Y, while in the transposed convolutional setup, the filter with weights W is used to map its input to a higher dimension.

3.3. Wasserstein GAN (WGAN)

The traditional GAN Loss function tries to minimize the distance between the distribution function of the generated and real data using the Kullback–Leibler (KL) divergence (Goodfellow et al., 2014). However, this loss is prone to mode collapse. This is mostly due to the well-trained Discriminator, leaving the generator with only a small value of loss available for improvement. The Wasserstein GAN (Arjovsky et al., 2017) was introduced in 2017, with a new loss function used to improve the Discriminator. Instead of the sigmoid classification loss, here the Discriminator acts like a critic and outputs real values as scores to indicate how representative the input samples are. The Discriminator loss is represented as:

$$\mathscr{L}_D^{WGAN} = -\mathbb{E}_{x\sim p_d}[D(x)] + \mathbb{E}_{G(z)\sim p_g}[D(G(z))] \tag{3}$$

This formulation is similar to the original GAN where only the log term has disappeared. Similarly, the Generator loss function is reduced to:

$$\mathscr{L}_G^{WGAN} = -\mathbb{E}_{G(z)\sim p_g}[D(G(z))] \tag{4}$$

A problem with WGAN is that the Wasserstein distance is highly interactable and the updated values of the model parameters can change rapidly causing instability in the training of GAN. To alleviate this problem, gradient clipping was proposed (Arjovsky et al., 2017). Later the clipping process was replaced with a Gradient Penalty (GP) (Gulrajani et al., 2017), resulting in a new GAN named WGAN-GP. In this scheme, the model is penalized if the gradient norms move away from the norm target value 1. The Generator loss function of the WGAN-GP is similar to WGAN, however, the GP term is added to $\mathscr{L}_D^{WGAN}$. Accordingly, $\mathscr{L}_D^{WGAN-GP}$ can be formulated as:

$$\mathscr{L}_D^{WGAN-GP} = \mathscr{L}_D^{WGAN} + \lambda\mathbb{E}_{G(z)\sim p_g}\left[\left(\|\nabla D(\alpha x + (1-\alpha G(z))\|_2 - 1\right)^2\right] \tag{5}$$

where ∇ represent the Gradient operation and α controls the contribution of real and generated samples in the penalization term.

3.4. Mode collapse

Generally, we expect GANs to generate a wide variety of samples. This is especially important in geosciences, where maintaining variability is crucial for an accurate quantification of uncertainty in predictions (e.g., forecast of hydrocarbon recovery factor). GANs have been shown to successfully generate visually acceptable samples. However due to mode collapse in the training process of a GAN, variability in the generated realizations is lower than the variability of the input data resulting in induced spatial bias. Mode collapse is defined as the case whereby the Generator can only generate one type of sample or a small set of distinct samples. This is caused when the Generator learns to generate samples from few modes of the data distribution but ignores other modes although they are present in the input training data (Srivastava et al., 2017). This effect when generating geological realizations, will result in the Generator producing multiple samples/realizations with one or more features replicated in common locations. The realizations overall are therefore spatially biased towards those features. Mode collapse can be severe where it is visually detectable in the samples or it can be partial where the effects are subtle and not visually obvious. A partial mode collapse would cause bias in the generated geological realizations where certain areas in the model will have a higher chance

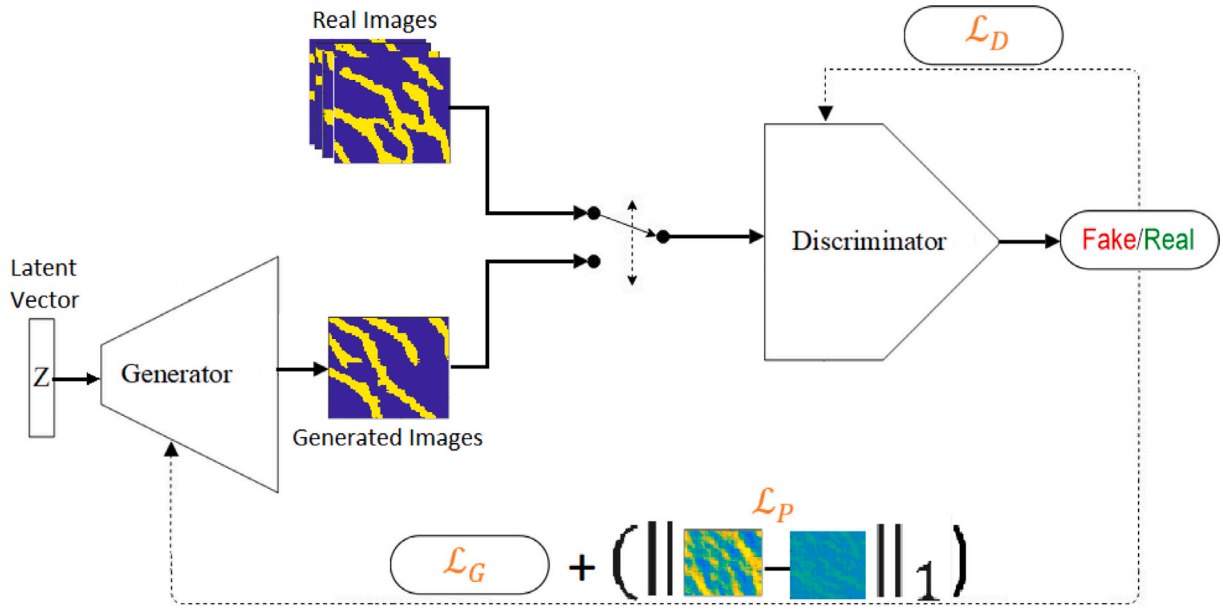

Fig. 4. Proposed framework for Variability Aware GAN(VA-GAN). In this framework, a probability map is produced for each batch of generated data in every iteration. Then probability loss is used to punish spatially biased batches.

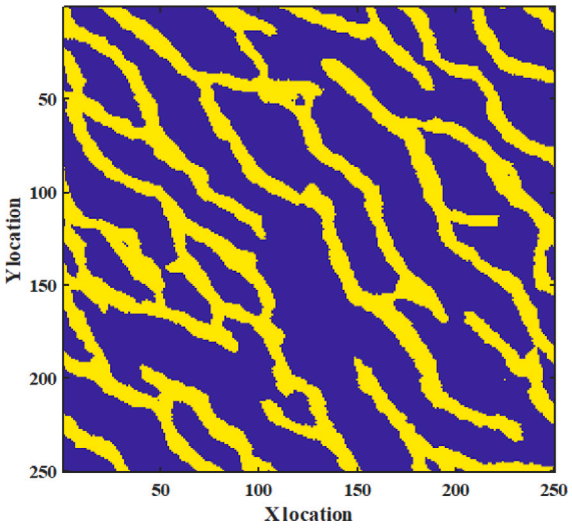

Fig. 5. Reference training image for a binary facies model of a fluvial distribution.

of showing sand while other areas are biased to shale. An example of severe mode collapse is shown in Fig. 3. In this figure, 25 tiled visually acceptable realizations of a channelized fluvial system (Fig. 3b) have been compared with 25 tiled realizations of the same geological system generated with severe mode collapse (Fig. 3a). Clearly, despite the introduced noise, under mode collapse the yellow channels are reoccurring in more or less the same locations.

4. Proposed GAN architecture

Given a distribution of geological patterns, we aim to generate realizations that are visually acceptable, honor the statistics of the distribution and maintain variability. This study focuses on maintaining the variability of the generated realizations at the same level as the input training data and reduce spatial bias as much as possible. To achieve this, we have proposed a new GAN architecture with a regularization term introduced to the Generator loss function. GANs work through the training data in batches. To investigate the variability, we generated probability maps of the input training data and GAN generated realizations. Probability maps are an excellent tool for evaluating spatial bias, since they can visually show the location of the spatial bias and quantitatively measure the severity of it. Comparison of the probability maps generated for input training data and generated realizations showed GAN results exhibit spatial bias. This comparison sparked the idea to change the GAN architecture so that in every iteration it punishes spatially biased batches.

Fig. 4 shows our proposed framework. As shown in the figure our proposed architecture consists of a Discriminator and a Generator that perform the traditional functionality. However, a regularization term is added to the generator loss to penalize batches of data that show spatial bias. This is accomplished by introducing a loss component to the Generator derived by the norm one distance between the probability map of the batch data and the probability map of the input training data. In the case of a stationary RTI, substituting with distance to a uniform map would suffice. In a non-stationary case, though, where for example the geologist prefers to assign higher probability to a feature or channel in a specific location of the map, generating the input training data probability map would accommodate for that and is a more general approach.

Assume $x_k \in \mathbb{Z}_2^{i\times j}$, where $k \in [1, 2, \dots, N_r]$ is kth binary realization sample of size $i \times j$ from a dataset $\mathbb{X} \in \mathbb{Z}_2^{i\times j\times N_r}$. N_r denotes the total number of realizations in a batch or input training dataset. Since x is a binary realization the pixel values are either 0 for background mud or 1 for channels. Summing all same index pixel values of all realizations in the dataset $\mathbb{X}$ and dividing the result by N_r, will generate the probability map $P_m(\mathbb{X}) \in \mathbb{R}^{i\times j}$ of size $i\times j$, where each pixel in $P_m(\mathbb{X})$ has a value between zero and one, which represents the probability of sand occurrence at that pixel. If each realization is represented by an $i \times j$ matrix then:

$$P_m(\mathbb{X}) = \frac{1}{N_r}\sum_{k=1}^{N_r} x_k \tag{6}$$

Probability loss is defined as:

$$\mathscr{L}_p = \|P_m(G(z)) - P_R\|_1 \tag{7}$$

where $P_m(G(z))$ is the probability map of a batch of realizations and P_R is the probability map of the input training data. We will from here on refer to this architecture as Variability Aware GAN (VA-GAN).

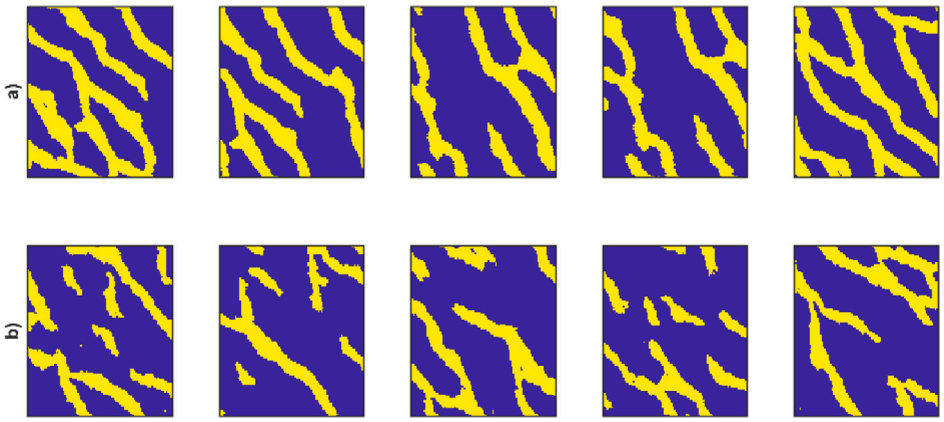

Fig. 6. Samples of input training data generated using two methods. (a) Direct Patch (b) SNESIM.

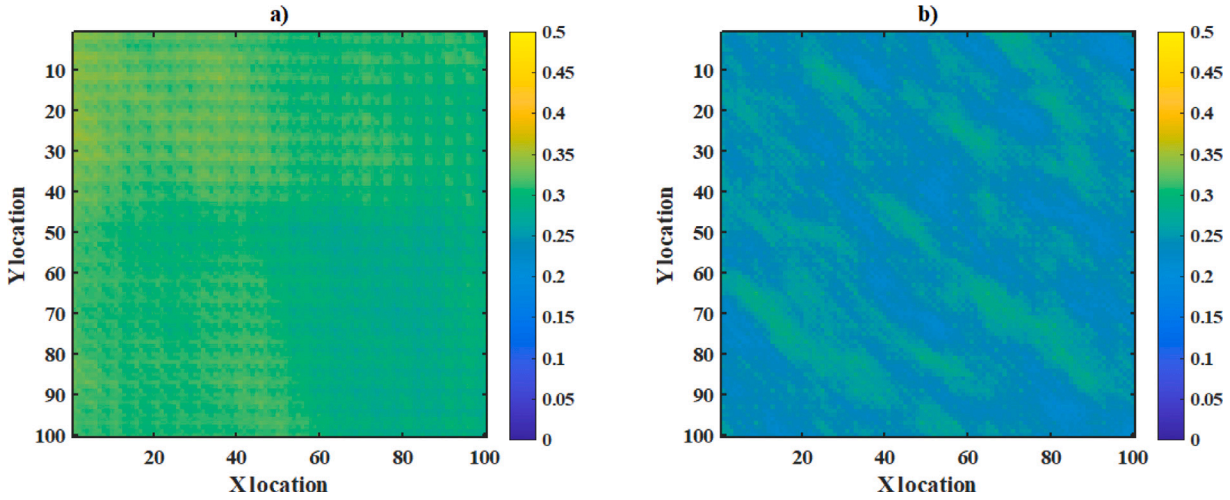

Fig. 7. Probability map of input training data. (a) Probability map of Direct Patch input training data (b) probability map of SNESIM input training images.

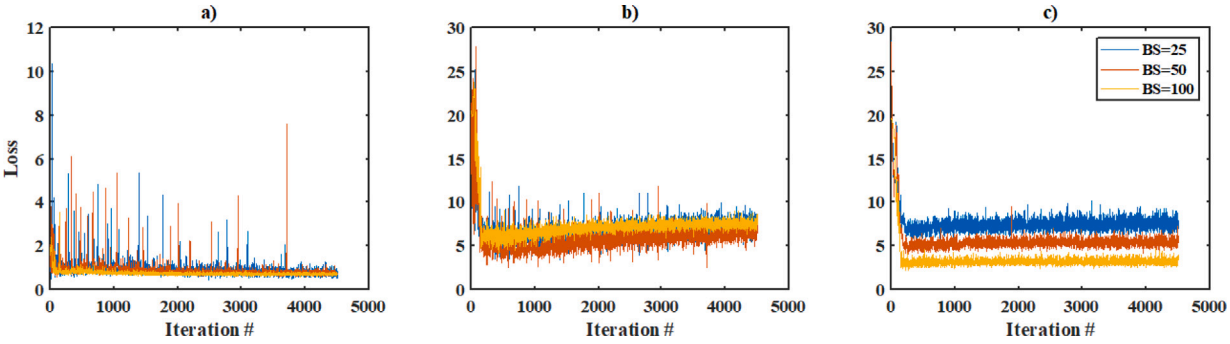

Fig. 8. Effect of three different Batch size values on VA-GAN$_{\mathscr{L}_{MinMax}}$. (a) Discriminator loss. (b) Generator loss. (c) Probability loss.

We have compared our proposed VA-GAN with two other GANs used to generate geological realizations. The DCGAN and WGAN-GP used in this study, both have convolutional layers in their structure, therefore, the main difference between the two GANs is the formulation of their loss. The DCGAN uses the traditional MinMax loss, while the WGAN-GP uses the Wasserstein loss (Wloss). We have tested VA-GAN with both minmax (VA-GAN$_{\mathscr{L}_{MinMax}}$) and Wasserstein loss (VA-GAN$_{\mathscr{L}_W}$) to evaluate its effect on variability of realizations. The Generator loss ($\mathscr{L}_G$) in VA-GAN$_{\mathscr{L}_{MinMax}}$ is:

$$\mathscr{L}_G^{\text{MinMax}} = -\mathbb{E}_{G(z)\sim p_g}\left[log\left(D\left(G(z)\right)\right)\right] + \beta\mathscr{L}_P \tag{8}$$

$\mathscr{L}_G$ in VA-GAN$_{\mathscr{L}_W}$ is:

$$\mathscr{L}_G^{\text{Wasserstein}} = -\mathbb{E}_{G(z)\sim p_g}\left[D\left(G(z)\right)\right] + \beta\mathscr{L}_P \tag{9}$$

In the above loss functions, β is a weighting factor used to establish a balance between the contribution of the first component of the loss functions and $\mathscr{L}_P$. β controls how much variability correction should be applied to the generator loss. To derive a value for β, we have conducted a sensitivity analysis. The analysis shows that the optimum outcome can be achieved when both loss components are in the same order of magnitude. For example, in DCGAN the Generator loss is in the order of $1E-1$ while the $\mathscr{L}_P$ is in the order of $1E+3$. Therefore, the best

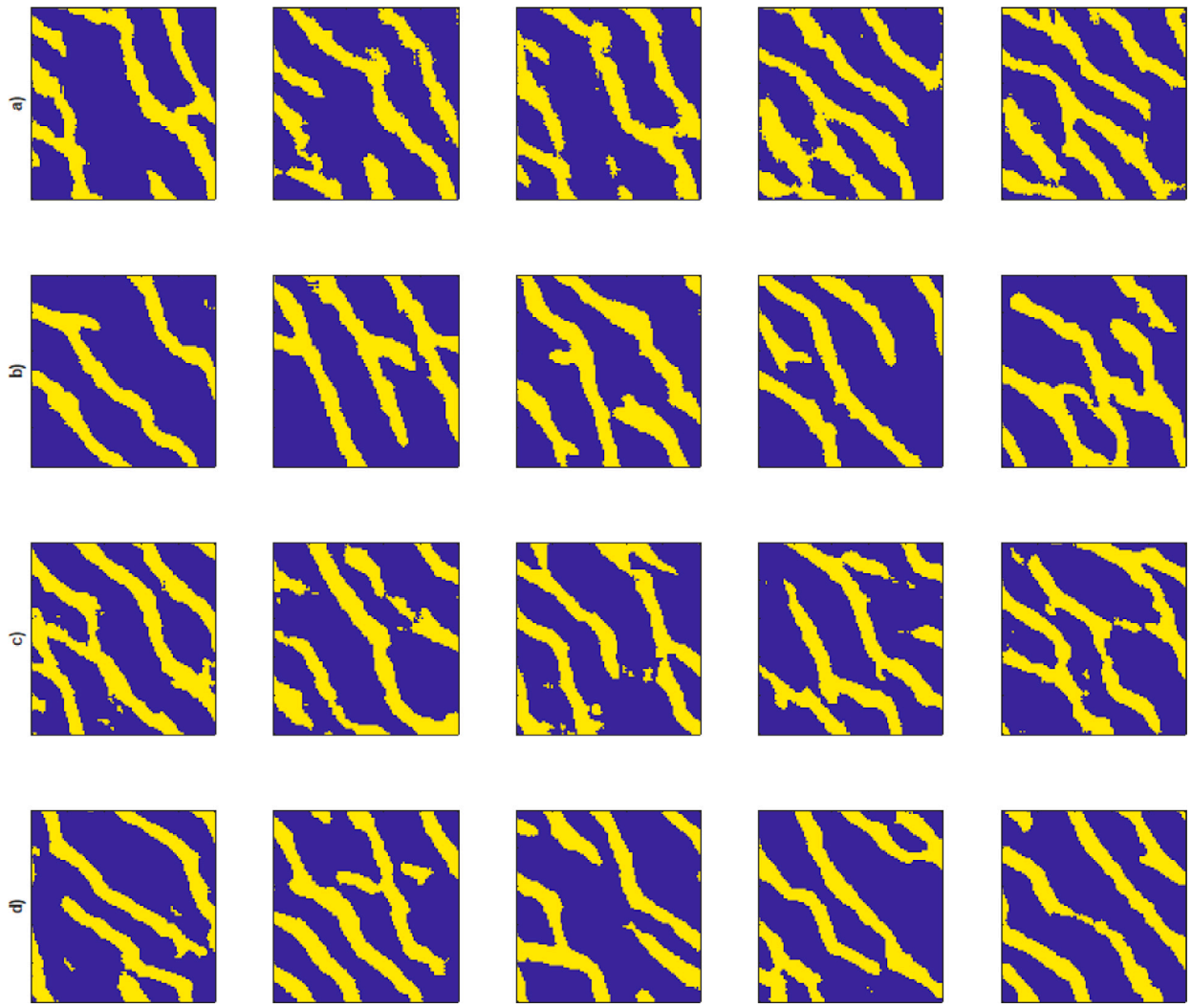

Fig. 9. Sample 100 × 100 realizations generated by different GANs trained on Direct Patch dataset. From top, (a) DCGAN (b) WGAN-GP (c) VA-GAN$_{\mathcal{L}_{MinMax}}$ (d) VA-GAN$_{\mathcal{L}_{W}}$.

β value is one that ensures both loss components have approximately similar contribution.

5. Experimental setup

5.1. Datasets and reference training image

In this study, we limit our investigations to a binary fluvial distribution of permeability zones. The GANs are trained to generate unconditional realizations of size 100 × 100. We have used a modified version of the famous Strebelle RTI (Strebelle and Journel, 2001) shown in Fig. 5 as a statistical reference to generate realizations. The RTI is a 250 × 250 pixel binary channelized permeability model exhibiting geostatistical stationarity. It consists of two facies, where yellow in Fig. 5 corresponds to channel sand and blue indicates background mud/clay-rich floodplain material. We have rotated the original channel flow direction by 45 degrees. Therefore, in our realizations the channel flow direction is along a Northwest–Southeast trend. In order to use the reference image as input training data to the GANs, we used two different methods. Our first dataset was created by sliding a 100 × 100 patch window over the reference image with a step of 4 pixels. A similar strategy was used in Azevedo et al. (2020). The step of 4 pixels was chosen to limit the number of dataset samples between 1000 to 1500 based on reports in the literature. Higher numbers would only increase the computation expense of training the network. Overall 1444 realizations of size 100 × 100 were produced. We refer to this dataset as the "Direct Patch" dataset. The second dataset was 1000 realizations generated using an MPS algorithm, SNESIM. The MPS algorithm was applied using the SGeMS toolbox (Remy et al., 2009). A similar methodology was used in Chan and Elsheikh (2020) and Liu et al. (2019). This dataset will be referred to as "SNESIM" dataset. Samples of realizations generated using the above-mentioned methods are presented in Fig. 6. There are numerous methods that can be used to generated the input training dataset (Mariethoz et al., 2010; Gravey and Mariethoz, 2020). In Cao et al. (2021) a large dataset of Strebelle RTI simulations are referenced. It has been shown that the input training data affects the generated results of GANs (Kamenshchikov and Krauledat, 2018). We trained the GANs in this study on both datasets to also investigate the effect of each method. The SNESIM was chosen for this study because it is frequently used in the literature and the Direct patch dataset and its effects on the results is a research question we aim to investigate. Replacing any of the datasets with a diffrent one would not affect the ultimate goal of the paper. Probability maps were generated for both the input training data and generated realizations. The input training data are binary images where 1 represents high permeability sand channel and 0 represents background mud. Each individual pixel in the probability map has a value between 0 and 1. This can be interpreted as the probability of sand occurring at that location. The probability map of both Direct patch and SNESIM input training data is shown in Fig. 7. These images show no spatial bias and all pixel values are close to 30% which is consistent with the univariate statistics of the RTI.

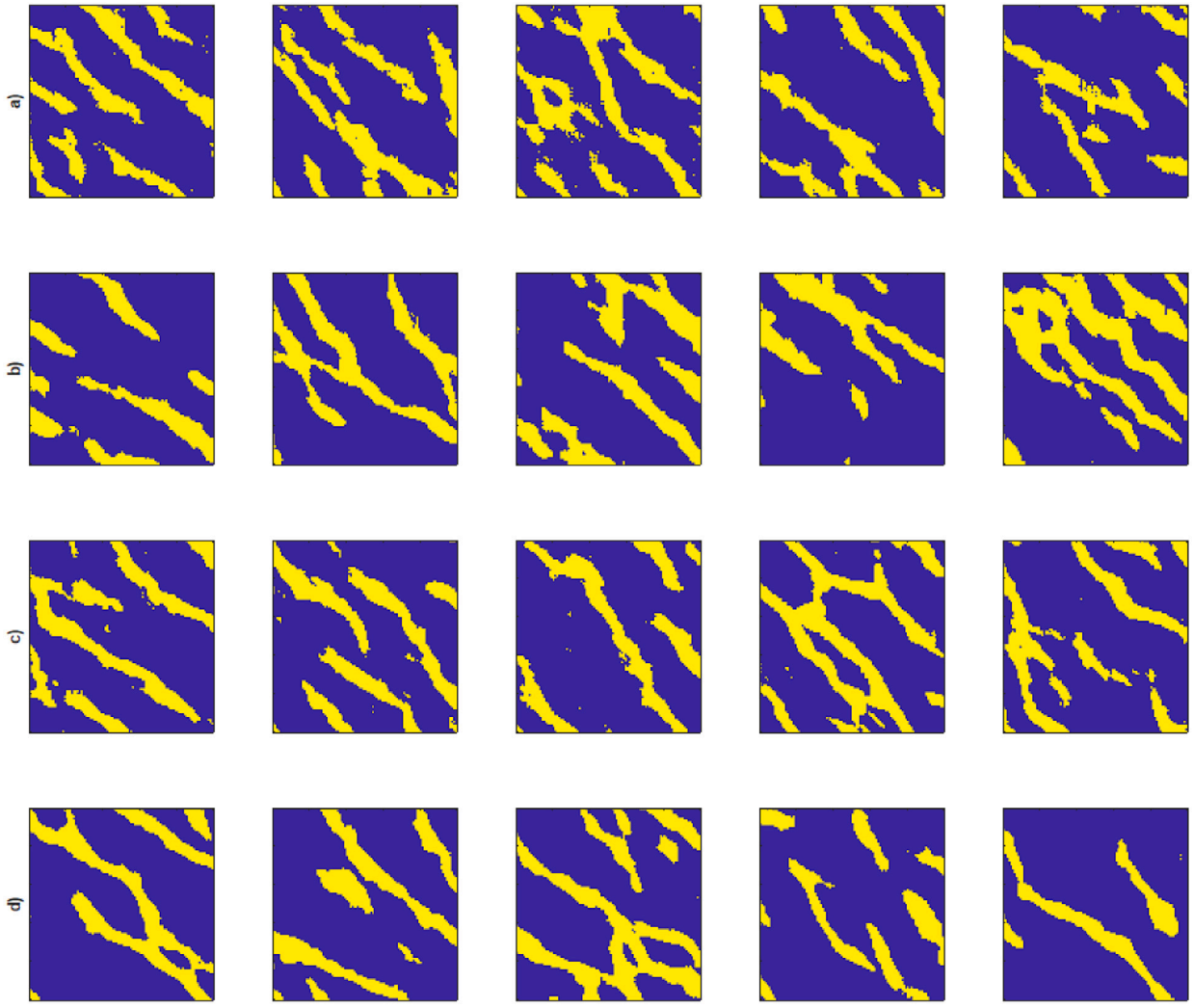

Fig. 10. Sample 100 × 100 realizations generated by different GANs trained on SNESIM dataset. From top, (a) DCGAN (b) WGAN-GP (c) VA-GAN$_{\mathscr{L}_{MinMax}}$ (d) VA-GAN$_{\mathscr{L}_{W}}$.

To generate probability maps, sets of 150 realizations were generated using a trained Generator. The output of the trained generator ranges between −1 and 1. The realizations were then rescaled and binarized.

There are many methods available for quality controlling and evaluating samples of generated realizations (Rongier et al., 2016). In this study, besides visual and qualitative evaluation of the GAN results, the method of Analysis of distances (ANODI) as described by Tan et al. (2014) was used for a more quantitative approach to evaluating the generated samples. Method of ANODI captures multiple point statistics of complex structures by calculating a multiple point histogram of patterns extracted using predefined template at different resolutions. The Jensen–Shannon divergence between the histograms is used as a measure of distance between the images. This distance is used to gauge how close the generated realizations are to the RTI. In other words, how well the generated data have preserved the multiple point statistics of the reference training image. The analysis was performed on both the input datasets and generated samples at 10 different resolutions, with a template size of 4 × 4 . To visualize the distances, Multidimensional scaling (MDS) is used. As a measure of quality of the generated data we used the average Jensen–Shannon distance of all the generated realizations to the reference training image.

5.2. GAN implementation

The DCGAN architecture used in this study, is similar to the DCGAN proposed in Radford et al. (2015). The network was customized for 100 × 100 input size images. The Discriminator uses LeakyRelu activation for all layers except Sigmoid activation as the last layer. A Dropout layer is used after the input layer with the probability of 0.5. Batch normalization was used in every layer of the discriminator. The Generator uses Relu activation for all layers except the last, where Tanh activation is used. Batch normalization was used in all layers. VA-GAN$_{\mathscr{L}_{MinMax}}$ was trained on Direct patch and SNESIM datasets to compare with DCGAN. Both networks were trained with the same hyper-parameters namely batch size of 100, learning rate of $2E-4$ and same layering in the Generator and Discriminator as described above. Adam optimizer was used for training. A sensitivity analysis showed that higher batch sizes yield better variability. As shown in Fig. 8, increasing the Batch size while training VA-GAN$_{\mathscr{L}_{MinMax}}$ reduces the probability loss while it has little effect on the Generator and Discriminator losses. Meanwhile, the probability maps showed a narrower range between the minimum and maximum pixel value and less outliers, indicating lower spatial bias and better variability for both input datasets. Batch sizes larger than 100 showed little improvement. The WGAN-GP algorithm (Gulrajani et al., 2017) used in this study has a layer structure, similar to the WGAN used by Chan and Elsheikh (2020), only input size is adjusted to 100 × 100 images. The Generator uses ReLu activation for all layers except the

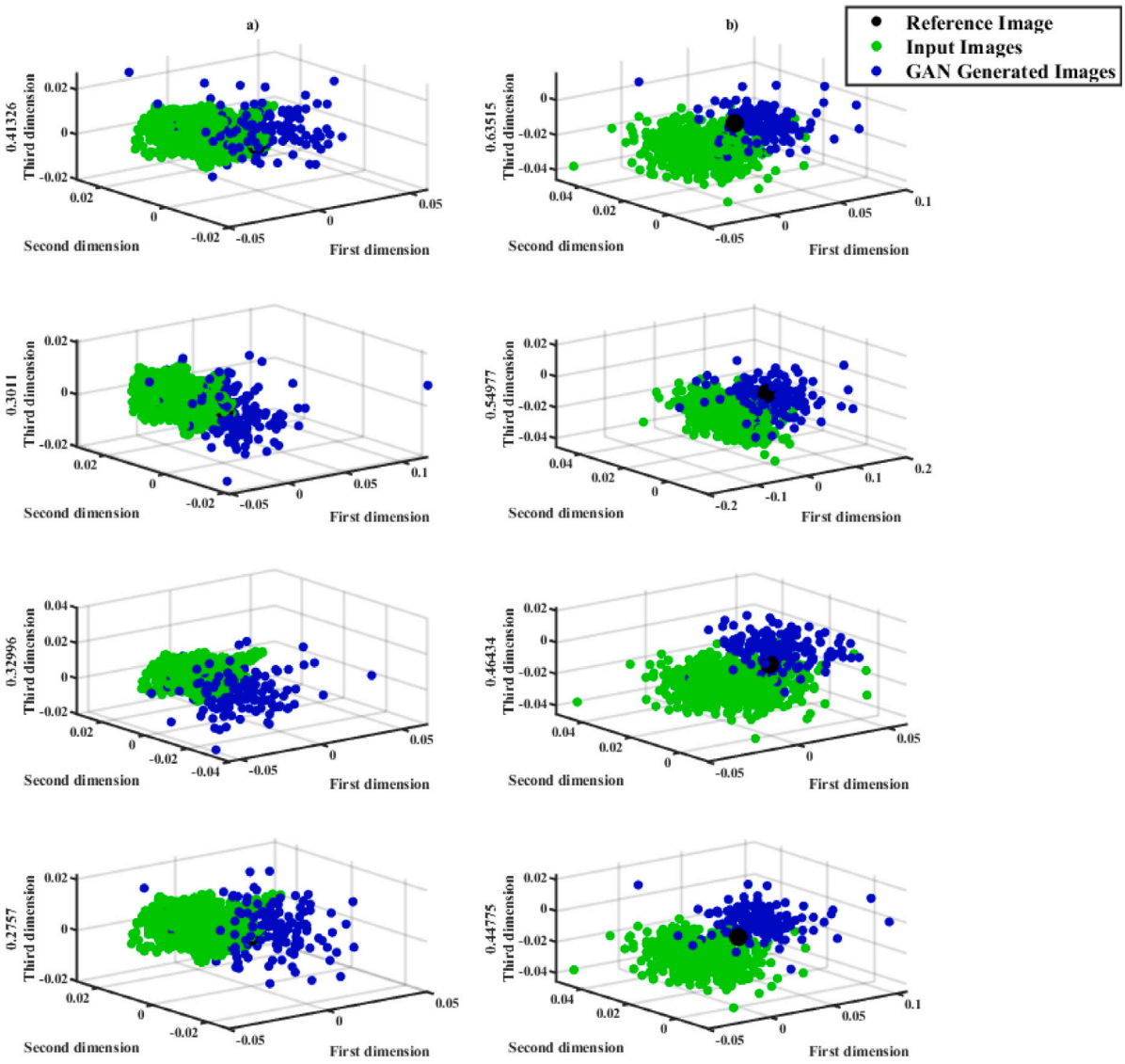

Fig. 11. MDS plots of the Jensen–Shannon distances between input datasets to RTI and generated realizations to RTI. (a) (Left column) Realizations generated from GANs trained on Direct patch dataset. (b) (right column) Realizations generated from GANs trained on SNESIM dataset. From top, Row 1: DCGAN, Row 2: WGAN-GP, Row 3: VA-GAN$_{\mathscr{L}_{MinMax}}$, Row 4: VA-GAN$_{\mathscr{L}_W}$. The number adjacent to the Z-axes of each figure is the Euclidean norm of the Jensen–Shannon distance of all generated realizations to the RTI.

final layer which has a Tanh activation. Batch normalization layers were used as well. The batch size was 32 and learning rate was $2E-4$ similar to the network trained in Chan and Elsheikh (2020). LeakyReLu activation layers along with batch normalization layers made up the Discriminator. The last layer in the Discriminator does not have any activation. In the process of training the WGAN-GP, we noticed that the batch normalization layers play an important role. The absence of these layers caused significant noise in the produced realizations. This noise in turn significantly increased the Jensen–Shannon distance of the realizations to the RTI. VA-GAN$_{\mathscr{L}_W}$ was trained with similar hyper-parameters and the variability of generated results was compared to WGAN-GP. We aimed for the GAN layers to be as close as possible to the networks used in the literature to achieve a more reliable comparison between our architecture and what has been previously used. For this reason, we have used similar layers to networks used in our Refs. Chan and Elsheikh (2020) and Radford et al. (2015). All GANs whether reference or proposed (VA-GAN) were trained from scratch and our stopping criterion was when loss stabilized rather than specific number of epochs or iterations. The latent input vector size for all the GAN architectures trained in this study is 100. The Discriminators were adjusted to input single channel (grayscale) images of size 100×100. The generator images have the same size and channel number.

6. Results

We have trained all the four GANs on two sets of input training images. A DCGAN and WGAN-GP were trained, and their results were compared with those obtained by our proposed VA-GAN$_{\mathscr{L}_{MinMax}}$ and VA-GAN$_{\mathscr{L}_W}$. Our proposed architectures, are designed to reduce spatial bias and maintain variability of the input training data. $\beta = 1E-4$ and $\beta = 1E-2$ were used for VA-GAN$_{\mathscr{L}_{MinMax}}$ and VA-GAN$_{\mathscr{L}_W}$ generator losses respectively. Visual inspection and ANODI were both used to evaluate the quality of the realizations, and probability maps were used as a measure for variability and spatial bias.

Samples of generated realizations of all four GANs trained on the Direct patch input dataset are shown in Fig. 9. Five random realizations were drawn from 150 realizations generated from each GAN. Each row in the figure corresponds to a GAN. From top, (a) DCGAN, (b) WGAN-GP (c)VA-GAN$_{\mathscr{L}_{MinMax}}$ (d)VA-GAN$_{\mathscr{L}_W}$. The realizations in the figure have preserved the binary features, channel width and continuity of the RTI. All four GANs are capable of generating visually acceptable

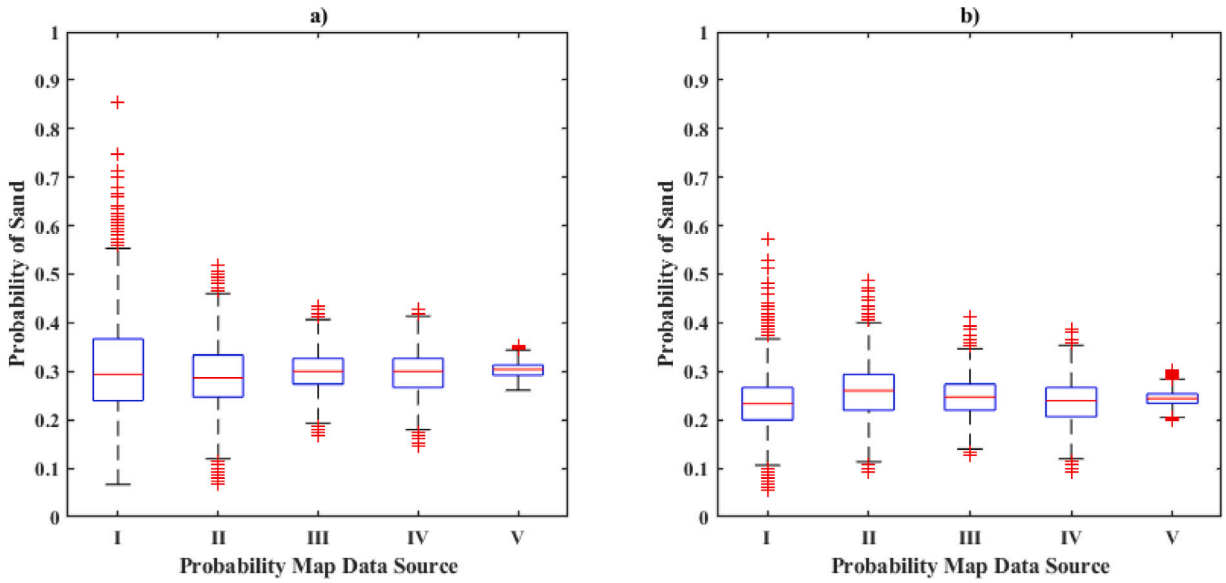

Fig. 12. Comparison of univariate statistics of Probability maps of input training data and generated realizations of GANs. The data sources for each probability map are as follows: (a) Comparison of probability maps produced from GANs trained on Direct patch dataset, (I) realizations generated from DCGAN, (II)realizations generated from WGAN-GP, (III)realizations generated from VA-GAN$_{\mathscr{L}_{MinMax}}$, (IV)realizations generated from VA-GAN$_{\mathscr{L}_W}$, (V) Direct patch input dataset. (b)Comparison of probability maps produced from GANs trained on SNESIM dataset, (I)realizations generated from DCGAN, (II)realizations generated from WGAN-GP, (III)realizations generated from VA-GAN$_{\mathscr{L}_{MinMax}}$, (IV)realizations generated from VA-GAN$_{\mathscr{L}_W}$, (V)SNESIM input dataset.

realizations that preserve qualitatively the multiple point statistics of the RTI.

Fig. 10 shows five randomly selected realizations out of a pool of 150 realizations generated by each GAN trained on the SNESIM dataset. Each row corresponds to a GAN as described above. While all samples in Fig. 10 are visually acceptable and channel sinuosity and width are consistent with the RTI, the generated samples are less continuous. GANs trained on the SNESIM dataset produce samples that occasionally show artifacts and channels are less continuous, with shorter and more dispersed channels present in most of the sample realizations. The cause of this issue is the SNESIM algorithm and the discontinuity was introduced when SNESIM was applied to generate the input training dataset. This is clearly shown in Fig. 6 and was also observed by Azevedo et al. (2020).

As second measure of quality, ANODI was applied to generated results of all the 8 settings and the input training data. The results of Jensen–Shannon distances are shown in MDS plots of Fig. 11 for better visualization. In this figure, green shows the input realizations, blue shows the generated realizations and the larger black dot represents the RTI. The left and right columns are results from GANs trained on the Direct patch and the SNESIM dataset respectively. The Euclidian norm of the Jensen–Shannon distance of all 150 generated realizations to the RTI is used to evaluate which GAN preserves the multiple point statistics of the RTI better than others. This distance for each GAN is shown next to the Z-Axes of each plot in Fig. 11 and also presented in Table 1 for convenience. In general, GANs trained on the Direct patch dataset produced statistically closer results with lower Jensen–Shannon distances to the RTI, which is in line with observing more continuous channels in the realizations. Regardless of the input training data, WGAN-GP (Fig. 11 Row2) was able to preserve better the multiple point statistics of the RTI compared to DCGAN (Fig. 11 Row1). Our proposed architecture has reduced the Jensen–Shannon distance of the realizations to RTI while enhancing variability considerably. As shown in Table 1, VA-GAN$_{\mathscr{L}_{MinMax}}$ in comparison to DCGAN and VA-GAN$_{\mathscr{L}_W}$ compared to WGAN-GP have enhance the average Jensen–Shannon distance of realizations by 19.5% and 10% respectively.

It should be noted that although DCGAN results show a higher Jensen–Shannon distance, the realization generated are still visually acceptable. However, the reduction of the Jensen–Shannon distance and generating results that better preserve multi-point statistics of the reference image is a further advantage of our proposed architecture.

Table 1
Euclidean norm of the Jensen–Shannon distances between generated realizations and the reference training image.

GAN	Direct Patch input data	SNESIM input data
DCGAN	0.41	0.64
WGAN-GP	0.30	0.55
VA-GAN$_{\mathscr{L}_{MinMax}}$	0.33	0.46
VA-GAN$_{\mathscr{L}_W}$	0.27	0.45

Probability maps were generated for both the input training datasets and the generated realizations. These maps were used to evaluate the variability and spatial bias of generated realizations. The probability maps of the input training data are shown in Fig. 7. Fig. 7a shows the probability map of the Direct Patch input data and Fig. 7b is the probability map of the SNESIM dataset. Neither of these images show any significant spatial bias. The range of the probability values at all locations are similar and there are no significant outliers. The univariate statistics of the pixel values of the maps are shown in Fig. 12. In this figure the blue box shows the 25th to 75th percentile and the red line in the middle is the median of the map pixel values. The whiskers show the range of the data and the red crosses are the outliers. The Direct patch map ranges between 26–35%, while the SNESIM map has range between 20–30%. Both maps show a narrow range with few insignificant outliers. The probability maps of the input training data were used as a reference for the GANs trained on that input data.

Probability maps of samples generated by GANs trained in this study have been presented in Fig. 13. The GANs trained on the Direct patch dataset (column a), show significantly more spatial bias compared to GANs trained on the SNESIM dataset(column b). The DCGAN structure clearly shows bias in multiple locations when trained on Direct Patch dataset (Fig. 13 a(I)). Severe outliers are observed, some areas show over 80% chance of sand occurrence and other areas with probability lower than 10% chance of sand occurrence. This can also be observed in the sample realizations (Fig. 9). A repeated trend is clearly visible in the samples generated by the DCGAN. The results of DCGAN, however,

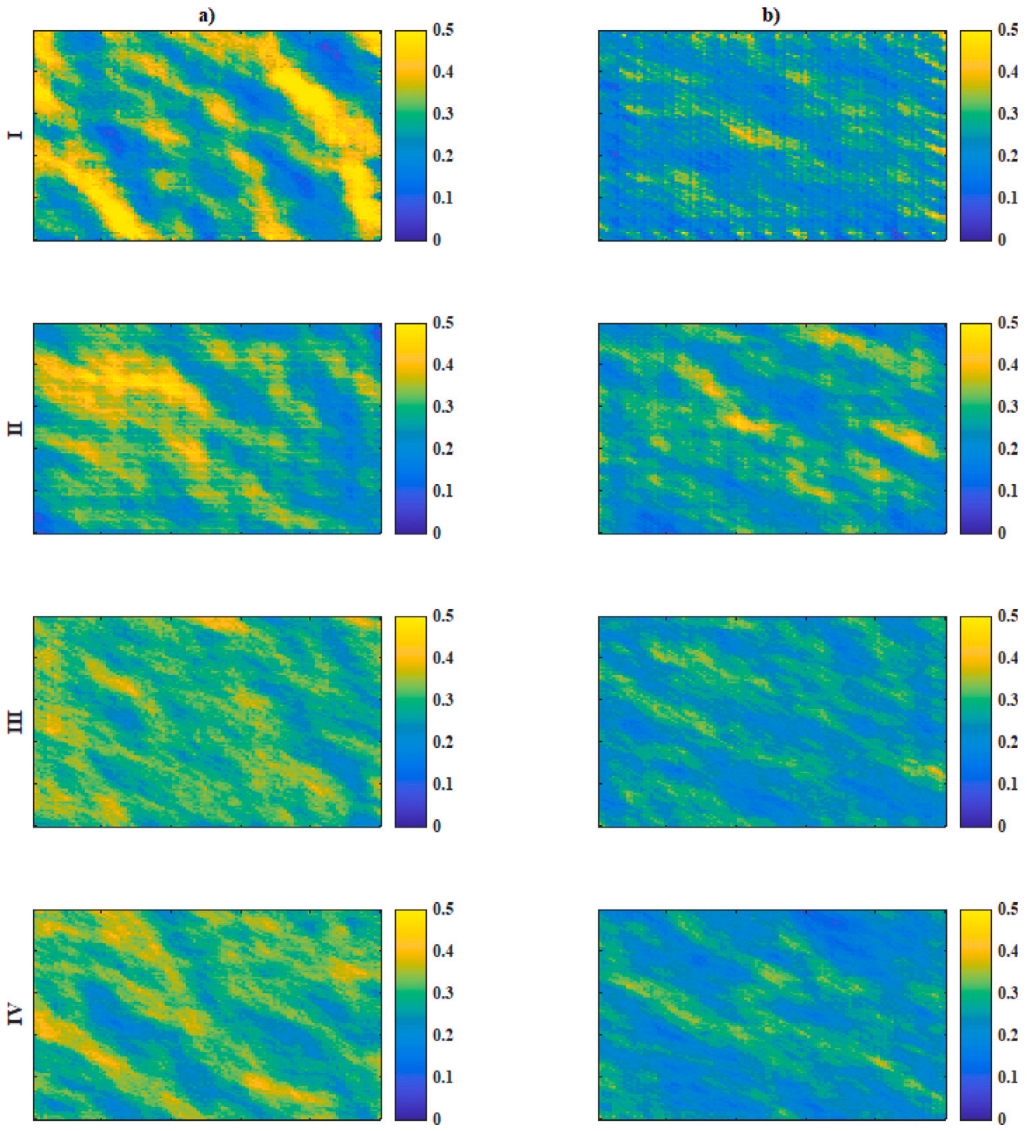

Fig. 13. Probability maps created from 150 realizations generated by different GANs. (a) (Left column) probability maps of realizations generated by GANs trained on Direct patch dataset. (b) (right column) probability maps of realizations generated by GANs trained on SNESIM dataset. (I) DCGAN, (II) WGAN-GP, (III) VA-GAN$_{\mathscr{L}_{MinMax}}$, (IV) VA-GAN$_{\mathscr{L}_W}$. Each probability map is of size 100 × 100.

are not as severe when trained on the SNESIM data set (Fig. 13 b(I)). Nevertheless, outliers over 50% and lower than 10% are present in the probability map when DCGAN is trained on SNESIM dataset. VA-GAN$_{\mathscr{L}_{MinMax}}$ significantly enhances the results, narrowing the range of probabilities and reducing the outliers (Fig. 13 Row III).

WGAN-GP algorithm, as expected and reported in the literature, reduces the effect of mode collapse. The results of this study have demonstrated this feature of the WGAN-GP (Fig. 13 Row II). Bias in the generated samples is not as severe. However, the range of probability values is wide and outliers over 50% and lower than 10% are occasionally observed. VA-GAN$_{\mathscr{L}_W}$ has enhanced the results compared to WGAN-GP as well, bringing the spatial distribution of the probability maps closer to that of the input data.

Training GANs on the Direct patch dataset produces more realistic and continuous channels. However, GANs trained on this dataset are prone to mode collapse, and spatial bias is observed in the results. Using our proposed architecture, it is possible to train a GAN with the Direct Patch input dataset, resulting in more visually continuous realizations with better preserved multiple point statistics of the RTI and at the same time maintain variability and reduce spatial bias. Naturally, the next step would be to condition the generated realizations to observed data. As mentioned in the related work section there have been a number of attempts to generate conditioned realizations. However, mode collapse and spatial bias caused by the GAN architecture and input training data often adversely affect the conditioning algorithm. The extent of this adverse effect might be offset somewhat by the tendency for a suit of conditioned realizations to exhibit less variability than unconditioned realizations. However, the specific focus in this study was on reducing the tendencies for mode collapse and spatial bias.

7. Discussions

To better demonstrate the impact of β in Eqs. (8) and (9), we have conducted a sensitivity analysis. The values of β used in this analysis are given in Table 2. Both VA-GAN$_{\mathscr{L}_{MinMax}}$ and VA-GAN$_{\mathscr{L}_W}$ were trained using the Direct patch dataset and these β values.

The proposed β is the value where both loss components are at the same order of magnitude (this is the value used to derive the results in the paper). This value was derived, using a Heuristic approach, where the first 20 iterations of training were monitored. In each iteration, both loss components were calculated and compared to determine the

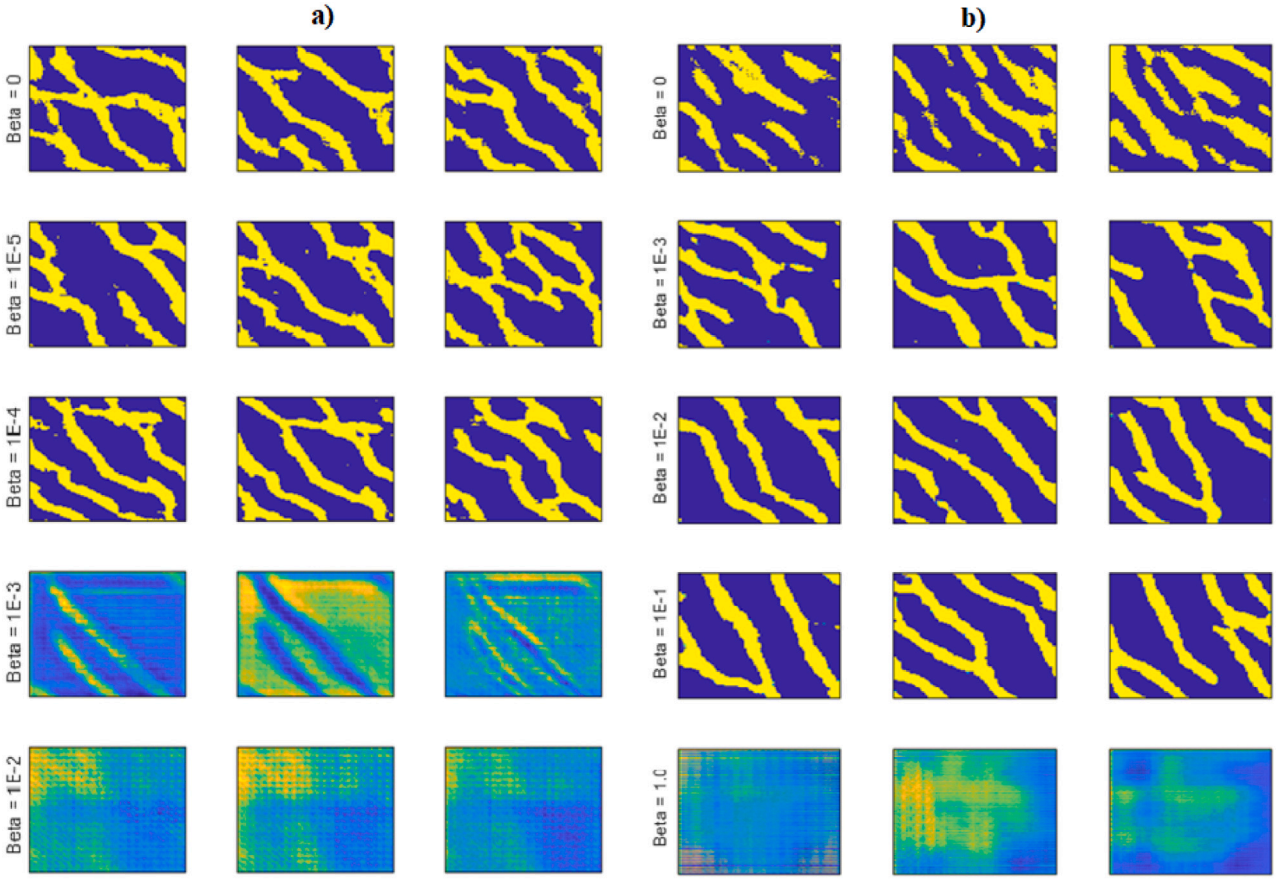

Fig. 14. Sample realizations generated from the β sensitivity analysis. β value for each set of realizations is shown on the Y axis. (a) VA-GAN$_{\mathscr{L}_{MinMax}}$, (b) VA-GAN$_{\mathscr{L}_W}$.

Table 2
Range of beta values included in the sensitivity check.

GAN	One order smaller	Proposed β	One order larger	Two orders larger
VA-GAN$_{\mathscr{L}_{MinMax}}$	$1E{-}5$	$1E{-}4$	$1E{-}3$	$1E{-}2$
VA-GAN$_{\mathscr{L}_W}$	$1E{-}3$	$1E{-}2$	$1E{-}1$	1.0

difference in the order of magnitude. β values one order of magnitude lower and one and two orders of magnitude higher compared to the proposed β were analyzed. $\beta = 0$ (regular GAN) has also been presented for better comparison. Fig. 14 shows sample realizations of the trained GANs. Three samples were randomly chosen from a pool of 150 realizations generated using each GAN. As the images show, when increasing the Beta value, the $\mathscr{L}_P$ component of loss function increases, and the focus of the loss function shifts to this component. Significantly large values will prevent the GAN from learning the features of the training data. On the other hand, as β becomes increasingly small, there is less focus on the probability loss and with extremely small values the component vanishes, therefore, results starts to resemble the regular GAN. The box plot of the probability maps generated using these realizations confirms this finding. As shown in Fig. 15, β values that do not maintain a balance between the loss components, tend to either introduce spatial bias or disturb the learning process of the GAN. In Fig. 15(a), VA-GAN$_{\mathscr{L}_W}$ still provides acceptable results with beta values one order of magnitude larger than the proposed β. This is because the GAN component of the loss in this architecture increases as the training progresses, therefore, larger values of β are still acceptable.

There are different methods that can be used to obtain the optimum beta value. Users can for example use sensitivity analysis or setup a rule to be applied every iteration based on the value of the loss components. The method of choosing beta is not detrimental to the methodology all together.

There are numerous regularization methods in the literature with the aim to stabilize GAN training and reduce Mode collapse (Kurach et al., 2019). Regularization methods generally strive to impose Lipschitz continuity on the discriminator to prevent the weights from exploding during training (Lee and Seok, 2020). We have compared the effectiveness of our proposed architecture with two regularization methods. WGAN-GP uses gradient penalty regularization. We have demonstrated, that Our proposed method can be applied to a GAN in conjunction with other regularization methods to further enhance results, as we have demonstrated with WGAN-GP (Figs. 12 and 13). We have further compared VA-GAN$_{\mathscr{L}_{MinMax}}$ with $L2$ regularized DCGAN. The results showed that while $L2$ regularization reduced outliers in the probability maps, it does not adequately reduce spatial bias. For the sake of conciseness, we have not presented the $L2$ comparison graphs here.

8. Conclusion

In this study, we developed a new GAN architecture to maximize variability of generated samples of subsurface spatial property. We applied our proposed architecture to generate unconditioned sample realizations based on the reference training image. The quality and variability of results of the conventional GANs and our proposed GANs were compared. The networks were trained on two input datasets, the Direct patch and SNESIM datasets. As quality measures, visual inspection and ANODI analysis were used. Probability maps were also applied to evaluate spatial bias and variability. Our proposed architecture significantly enhanced variability and reduced spatial bias compared to DCGAN and outperformed WGAN-GP. The results showed that the input training

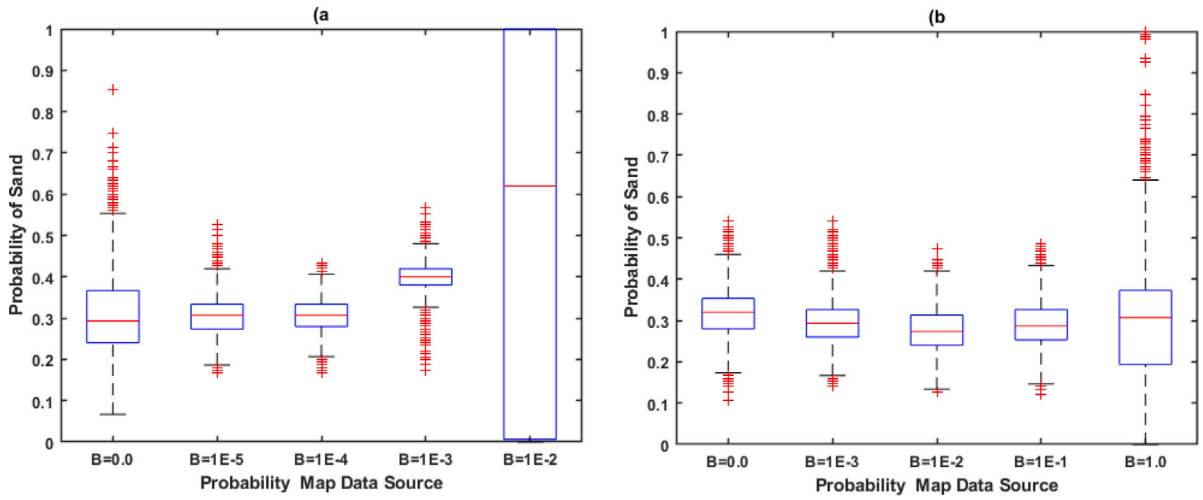

Fig. 15. Box plot of probability maps generated from the β sensitivity analysis (a)VA-GAN$_{\mathscr{L}_{MinMax}}$, (b) VA-GAN$_{\mathscr{L}_{W}}$.

dataset influenced the variability. GANs trained on the Direct patch dataset showed more spatial bias compared to GANs trained on SNESIM dataset. However, the range of values in the probability maps of our proposed architecture where narrower and closer to the probability maps of the input training data with less outliers compared to original GANs. Our proposed architecture also reduced the average Jensen–Shannon distance between the generated realizations and RTI by 19.5% and 10% compared to DCGAN and WGAN-GP respectively. This shows that the proposed architecture better preserved the multiple point statistics of the RTI. The proposed architecture can be applied to any GAN being trained on binary geological realizations.

In the future, the capability of GANs for generating Multi-facies property distributions should be investigated. It would be important to measure how the variability of a non-stationary and/or Multi-facies RTI may be preserved. A detailed comparison of GANs with current techniques would further clarify the pros and cons on these networks. In a separate avenue, there is room to research variability of conditioning algorithms and further conditioning to 'soft' geological constraints such as Isopach maps.

Code availability section

The code used in this study, is written in Matlab and requires Deep Learning tool box version 2021a or later. It is available for downloading at the link: https://github.com/roozbeh442/Variability-Aware-GAN.git

CRediT authorship contribution statement

Roozbeh Koochak: Conception and design of study, Analysis and/or interpretation of data, Writing – original draft, Writing – review & editing. **Mohammad Sayyafzadeh:** Analysis and/or interpretation of data, Writing – review & editing. **Ali Nadian:** Acquisition of data and machine learning network design, Writing – review & editing. **Mark Bunch:** Analysis and/or interpretation of data, Writing – review & editing. **Manouchehr Haghighi:** Analysis and/or interpretation of data, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Code availability section

The code used in this study, is written in Matlab and requires Deep Learning tool box version 2021a or later. It is available for downloading at the link: https://github.com/roozbeh442/Variability-Aware-GAN.git.

Acknowledgments

This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide. All authors approved the final version of the manuscript.

References

Arjovsky, M., Chintala, S., Bottou, L., 2017. Wasserstein generative adversarial networks. In: International Conference on Machine Learning. PMLR, pp. 214–223.
Astrakova, A., Oliver, D.S., 2015. Conditioning truncated pluri-Gaussian models to facies observations in ensemble-Kalman-based data assimilation. Math. Geosci. 47 (3), 345–367.
Azevedo, L., Paneiro, G., Santos, A., Soares, A., 2020. Generative adversarial network as a stochastic subsurface model reconstruction. Comput. Geosci. 24 (4), 1673–1692.
Bai, T., Tahmasebi, P., 2020. Hybrid geological modeling: Combining machine learning and multiple-point statistics. Comput. Geosci. 142, 104519.
Bang, D., Shim, H., 2018. Mggan: Solving mode collapse using manifold guided training. arXiv preprint arXiv:1804.04391.
Canchumuni, S.W., Castro, J.D., Potratz, J., Emerick, A.A., Pacheco, M.A.C., 2021. Recent developments combining ensemble smoother and deep generative networks for facies history matching. Comput. Geosci. 25 (1), 433–466.
Cao, Z., Shen, J., Gravey, M., 2021. Training image selection using recurrent neural networks: An application in hydrogeology. arXiv preprint arXiv:2101.06488.
Chan, S., Elsheikh, A.H., 2019a. Parametric generation of conditional geological realizations using generative neural networks. Comput. Geosci. 23 (5), 925–952.
Chan, S., Elsheikh, A., 2019b. Parametrization and generation of geological models with generative adversarial networks. arXiv.org.
Chan, S., Elsheikh, A.H., 2020. Parametrization of stochastic inputs using generative adversarial networks with application in geology. Front. Water 2, 5. http://dx.doi.org/10.3389/frwa.2020.00005.
Dupont, E., Zhang, T., Tilke, P., Lin, L., Bailey, W., 2018. Generating realistic geology conditioned on physical measurements with generative adversarial networks. arXiv preprint arXiv:1802.03065.
Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep Learning, Vol. 1. MIT press, Cambridge.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial networks. In: The Twenty-Eight Annual Conference on Neural Information Processing Systems(NIPS).
Gravey, M., Mariethoz, G., 2020. QuickSampling v1. 0: a robust and simplified pixel-based multiple-point simulation approach. Geosci. Model Dev. 13 (6), 2611–2630.
Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A., 2017. Improved training of wasserstein gans. arXiv preprint arXiv:1704.00028.

Jäggli, C., Straubhaar, J., Renard, P., 2017. Posterior population expansion for solving inverse problems. Water Resour. Res. 53 (4), 2902–2916.

Jetchev, N., Bergmann, U., Vollgraf, R., 2016. Texture synthesis with spatial generative adversarial networks. arXiv preprint arXiv:1611.08207.

Kamenshchikov, I., Krauledat, M., 2018. Effects of dataset properties on the training of GANs. arXiv preprint arXiv:1811.02850.

Ko, K.-I., Lin, C.-L., 1995. On the complexity of min-max optimization problems and their approximation. In: Minimax and Applications. Springer, pp. 219–239.

Kurach, K., Lučić, M., Zhai, X., Michalski, M., Gelly, S., 2019. A large-scale study on regularization and normalization in GANs. In: International Conference on Machine Learning. PMLR, pp. 3581–3590.

Laloy, E., Hérault, R., Jacques, D., Linde, N., 2018. Training-image based geostatistical inversion using a spatial generative adversarial neural network. Water Resour. Res. 54 (1), 381–406.

Laloy, E., Hérault, R., Lee, J., Jacques, D., Linde, N., 2017. Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. Adv. Water Resour. 110, 387–405.

Laloy, E., Linde, N., Jacques, D., Mariethoz, G., 2016. Merging parallel tempering with sequential geostatistical resampling for improved posterior exploration of high-dimensional subsurface categorical fields. Adv. Water Resour. 90, 57–69.

Le Loch, G., Beucher, H., Galli, A., Doligez, B., 1994. Improvement in the truncated Gaussian method: combining several Gaussian functions. In: Ecmor Iv-4th European Conference on the Mathematics of Oil Recovery. European Association of Geoscientists & Engineers, pp. cp–233.

Lee, M., Seok, J., 2020. Regularization methods for generative adversarial networks: An overview of recent studies. arXiv preprint arXiv:2005.09165.

Liu, Y., Sun, W., Durlofsky, L.J., 2019. A deep-learning-based geological parameterization for history matching complex models. Math. Geosci. 51 (6), 725–766.

Lopez-Alvis, J., Laloy, E., Nguyen, F., Hermans, T., 2021. Deep generative models in inversion: The impact of the generator's nonlinearity and development of a new approach based on a variational autoencoder. Comput. Geosci. 152, 104762.

Mariethoz, G., Renard, P., Cornaton, F., Jaquet, O., 2009. Truncated plurigaussian simulations to characterize aquifer heterogeneity. Groundwater 47 (1), 13–24.

Mariethoz, G., Renard, P., Straubhaar, J., 2010. The direct sampling method to perform multiple-point geostatistical simulations. Water Resour. Res. 46 (11).

Marini, M., Felletti, F., Beretta, G.P., Terrenghi, J., 2018. Three geostatistical methods for hydrofacies simulation ranked using a large borehole lithology dataset from the Venice Hinterland (NE Italy). Water 10 (7), 844.

Meerschman, E., Pirot, G., Mariethoz, G., Straubhaar, J., Van Meirvenne, M., Renard, P., 2013. A practical guide to performing multiple-point statistical simulations with the Direct Sampling algorithm. Comput. Geosci. 52, 307–324.

Mosser, L., Dubrule, O., Blunt, M.J., 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys. Rev. E 96 (4), 043309.

Mosser, L., Dubrule, O., Blunt, M.J., 2018a. Conditioning of three-dimensional generative adversarial networks for pore and reservoir-scale models. arXiv preprint arXiv:1802.05622.

Mosser, L., Dubrule, O., Blunt, M.J., 2018b. Stochastic reconstruction of an oolitic limestone by generative adversarial networks. Transp. Porous Media 125 (1), 81–103.

Mosser, L., Dubrule, O., Blunt, M.J., 2019. Deepflow: history matching in the space of deep generative models. arXiv preprint arXiv:1905.05749.

Mosser, L., Dubrule, O., Blunt, M.J., 2020. Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. Math. Geosci. 52 (1), 53–79.

Mosser, L., Kimman, W., Dramsch, J., Purves, S., De la Fuente Briceño, A., Ganssle, G., 2018c. Rapid seismic domain transfer: Seismic velocity inversion and modeling using deep generative neural networks. In: 80th Eage Conference and Exhibition 2018. European Association of Geoscientists & Engineers, pp. 1–5.

Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. MIT Press.

Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

Razak, S.M., Jafarpour, B., 2020a. Convolutional neural networks (CNN) for feature-based model calibration under uncertain geologic scenarios. Comput. Geosci. 24 (4), 1625–1649.

Razak, S.M., Jafarpour, B., 2020b. History matching with generative adversarial networks. In: ECMOR XVII. European Association of Geoscientists & Engineers, pp. 1–17.

Remy, N., Boucher, A., Wu, J., 2009. Applied Geostatistics with SGeMS: A User's Guide. Cambridge University Press.

Rongier, G., Collon, P., Renard, P., Straubhaar, J., Sausse, J., 2016. Comparing connected structures in ensemble of random fields. Adv. Water Resour. 96, 145–169.

Ruthotto, L., Haber, E., 2021. An introduction to deep generative modeling. arXiv preprint arXiv:2103.05180.

Sami, M., Mobin, I., 2019. A comparative study on variational autoencoders and generative adversarial networks. In: 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT). IEEE, pp. 1–5.

Srivastava, A., Valkov, L., Russell, C., Gutmann, M.U., Sutton, C., 2017. Veegan: Reducing mode collapse in gans using implicit variational learning. arXiv preprint arXiv:1705.07761.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. Math. Geol. 34 (1), 1–21.

Strebelle, S.B., Journel, A.G., 2001. Reservoir modeling using multiple-point statistics. In: SPE Annual Technical Conference and Exhibition.

Tahmasebi, P., 2018. Multiple point statistics: A review. In: Handbook of Mathematical Geosciences: Fifty Years of IAMG. pp. 613–643. http://dx.doi.org/10.1007/978-3-319-78999-6_30.

Tan, X., Tahmasebi, P., Caers, J., 2014. Comparing training-image based algorithms using an analysis of distance. Math. Geosci. 46 (2), 149–169.

Thekumparampil, K.K., Jain, P., Netrapalli, P., Oh, S., 2019. Efficient algorithms for smooth minimax optimization. arXiv preprint arXiv:1907.01543.

Yang, D., Hong, S., Jang, Y., Zhao, T., Lee, H., 2019. Diversity-sensitive conditional generative adversarial networks. arXiv preprint arXiv:1901.09024.

Yeh, R.A., Chen, C., Yian Lim, T., Schwing, A.G., Hasegawa-Johnson, M., Do, M.N., 2017. Semantic image inpainting with deep generative models. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5485–5493.

Zhang, C., Song, X., Azevedo, L., 2021. U-net generative adversarial network for subsurface facies modeling. Comput. Geosci. 25 (1), 553–573.

6. Summary, Conclusions and Future Work

The primary objective of reservoir characterization is to produce geological models that honour the available data. Adequate characterization of the reservoir results in better understanding of the subsurface which in turn increase recovery factor, leads to better reservoir management and is essential for post depletion activities such as carbon capture and sequestration or storage. In recent years the subsurface energy sector has seen a surge in data collection, however, traditional methods due to their inherent limitations are unable to fully utilize these data. This thesis investigated and introduced novel techniques in subsurface characterization using advanced analytics and machine learning that can handle large data volumes and open new doors in the field of reservoir characterization. This thesis investigated and introduced novel techniques in subsurface characterization using advanced analytics and machine learning. The thesis explored Fractal analysis, Convolutional neural networks, Generative adversarial networks and presented an enhanced optimization technique for history matching in the appendix. For each of the problems, a novel technique is proposed with the aim to enhance subsurface characterization using the available data.

The thesis began with Fractal analysis of the deep latero-log resistivity logs. A novel technique was introduced for rock typing using the Fractal dimension of these logs. The fractal dimension is a measure of the variability of the resistivity log. The research showed that as the complexity of the rock texture and fabric increases the log shows higher variability and a higher fractal dimension. The results of this investigation showed that fractal dimension of resistivity logs can be used as new parameter in reservoir rock typing. Furthermore, it introduced a new technique to extract more information from available resistivity logs that were never utilized before.

The thesis continued with introducing a methodology to automate/ augment interpretation of resistivity image logs. The research emphasised on solely using the image logs and excluded any other data types. Transfer learning was used to overcome some of the challenges faced in this process and it was proven that transfer learning is applicable to subsurface datasets. Using a trained network and novel post processing step the accuracy of the automated process exceed 90%. The aim of this study was to determine the accuracy that can be achieved using image logs alone. Generally, image logs are required to be accompanied by a suit of supplementary logs that aid the interpreter in categorizing facies. Omitting the supplementary logs would reduce the cost of the interpretation significantly. Although the result shows a high accuracy from image logs alone, this workflow would benefit from some additional data. Researching methodologies to incorporate multiple data sources in this workflow would be of great interest. Another potential application of this workflow would be denoising the logs or noise detection in the image logs.

In the next chapter, variability of geological model realizations generated using a Generative adversarial network was investigated. The input training data was produced using a conceptual image delivered by a geologist. The GAN was trained to learn the statistics of the training image and produce realizations that are realistic and statistically faithful. The variability of the images was investigated in detail and a novel method to update the loss function of the network with the aim to maintain the variability of the output samples at the same level as that of the input data. The proposed method outperformed the state-of-the-art methods in the literature used for reducing the effects of mode collapse. GANs are a recently developed technique and their application in geological model development is at the beginning of its journey. There are still

numerous challenges that need to be overcome. Mode collapse for one has a detrimental effect on uncertainty evaluation and must be evaluated and addressed when using GANs. Another important challenge is generating training data to train the network on. The method introduced in this thesis although effective, it is applicable to stationary models only. Non-stationary models or geological structures are quite challenging to train a GAN on.

Finally, in the appendix, an ensemble of surrogates (proxies) with generation-based model-management embedded in CMA-ES is proposed to reduce the number of simulation calls. The proposed method divides the likelihood term into multiple functions and represents each function with a separate surrogate. The algorithm was used to history match two cases one real with 59 uncertain variables and the other synthetic with 8 variables. The results showed that up to 65% and 50% less simulation calls for case#1 and case#2 were required.

From this investigation into development of novel advanced analytics and machine learning methods in subsurface characterization the following conclusions are drawn:

1. The fractal analysis of the deep resistivity logs revealed the correlation between the variability of resistivity logs and complexity of rock fabric, making this analysis useful for rock typing. Application of novel data analytics and machine learning techniques can introduce new features in subsurface data that have traditionally been neglected.
2. Convolutional neural networks are effective and efficient tools to analyse subsurface image data. It was shown on two occasions in this thesis that computer vision algorithms can be extended to subsurface image logs given that the limitations of the algorithm and how these limitations translate in subsurface analysis are well understood and investigated.

3. Convolutional neural networks can extract facies from resistivity images logs better than human interpretation when the resistivity logs are the only information available. However, complexity of the image distribution, lack of features in the images and similarity of categorized facies means that to reach accuracy levels similar to human full-suite interpretation, other logs and information need to be included in the network training process.
4. Given the complexity of the distribution of subsurface images, large and deep networks are required to learn the features of these images. Not only training these networks is computationally expensive, but lack of adequate data and training images makes this task even more challenging. Transfer learning has proven to be an efficient solution to both above-mentioned issues. Furthermore, I was proven that transfer learning is applicable to subsurface image datasets.
5. Generative adversarial networks (GAN) can adequately learn and generate geological models that are realistic and statistically faithful. These networks however, depending on their design can suffer from mode collapse which will hinder and reduce the variability of the generated models.
6. Providing input training data that is sufficiently representative of the distribution the network will learn, in variability, dataset size and image statistics is key to a successful training. The input training data is detrimental in the training of the network.

This research provided novel techniques for subsurface characterization and answered important questions in that field. Non-the-less this research can be extended in several ways. Below are suggestions for future work:

1. Introduction of a novel technique to incorporate other data types such as one-dimensional logs and geologists experience along with resistivity image logs when training a convolutional neural network with the aim to maximize the accuracy of automatic interpretation.
2. In depth investigation of a methodology to allow for neural networks to detect and flag noisy intervals in resistivity image logs. This would significantly reduce interpretation time. Furthermore, this feature can be used to develop online quality monitoring and noise detection capabilities.
3. Research the ability of GANs in generating multi-facies property distribution. Evaluate the variability of the generated samples as the distribution increases in complexity. This research would give GANs much wider application in geological modelling and would be a great step towards more practical applications.
4. Research novel methodology to generate training input data from a non-stationary and/or multi-facies training image while maintaining variability and statistical features of the image. This would significantly broaden the application of GANs in geological modelling.

7. Appendix: Accelerating CMA-ES In History Matching Problems Using An Ensemble Of Surrogates With Generation-Based Management

Mohammad Sayyafzadeh, **Roozbeh Koochak**, Mathew Barley

ECMOR XVI conference, 2018

Statement of Authorship

Title of Paper	**Accelerating CMA-ES In History Matching Problems Using An Ensemble Of Surrogates With Generation-Based Management**
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	**Sayyafzadeh, M., Koochak, R. and Barley, M., 2018, September. Accelerating cma-es in history matching problems using an ensemble of surrogates with generation-based management. In ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery (Vol. 2018, No. 1, pp. 1-15). EAGE Publications BV.**

Principal Author

Name of Principal Author (Candidate)	**Roozbeh Koochak**		
Contribution to the Paper	**Coding, experimentation, analysis of results, development of algorithm, conceptualization, revision**		
Overall percentage (%)	**45%**		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	**11/05/2023**

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	**Mohammad Sayyafzadeh**		
Contribution to the Paper	**Formulation, conceptualization, development of methodology and algorithm, analysis of results, discussions, revisions, original manuscript, supervision**		
Signature		Date	**18/05/2023**

Name of Co-Author	Mathew Barley		
Contribution to the Paper	Analysis of results, discussions, revisions		
Signature		Date	17-5-2023

Please cut and paste additional co-author panels here as required.

P007

Accelerating CMA-ES In History Matching Problems Using An Ensemble Of Surrogates With Generation-Based Management

M. Sayyafzadeh* (University of Adelaide), R. Koochak (The University of Adelaide), M. Barley (Santos Limited)

Summary

Because of the quasi-gradient update embedded in CMA-ES algorithm, it can outperform most of the population-based algorithms, from a convergence speed standpoint. However, due to the computationally expensive fitness function associated with history matching, the reduction of function (simulation) calls can be favourable.

In this study, an ensemble of surrogates (proxies) with generation-based model-management is proposed to reduce the number of simulation calls efficaciously. Since the fitness function is highly nonlinear, an ensemble of surrogates (Gaussian process) is utilised. The likelihood term is divided into multiple functions, and each is represented via a separate surrogate. This improved the response surface fitting.

In generation-based management, a stochastically selected measure (surrogate or reservoir-simulation) should be used to evaluate all the individuals of each generation. CMA-ES requires ranking of the individuals to select the parents. Therefore, the generation-based model-management fits well in CMA-ES, as surrogates are normally better in ranking the individuals than approximating the fitness.

History matching for a real problem with 59 variables and PUNQ-S3 with eight variables was conducted via a standard CMA-ES and the proposed surrogate-assisted CMA-ES. The results showed that up to 65% and 50% less simulation calls for case#1 and case#2 were required.

Introduction

History matching problems are typically nonlinear and do not have closed-form solutions. An iterative approach should be applied to find a solution (or multiple plausible solutions) for such problems. The uncertain parameters are first identified, and then, a quantitative measure is formulated and applied to distinguish between the models (solutions), based on observed data reproduction by a numerical simulation. After these two steps, the tuning process starts in which new solutions are generated and evaluated iteratively until one of the predefined stopping criteria is met. The tuning process is typically conducted by an optimisation rule in which a model (a set of parameters) is sought that has the best fitness according to the formulated objective function. Data assimilation is an alternative approach in which observed data integrated into the model by an assimilation technique, for instance, Ensemble Kalman Filter (EnKF) or Ensemble Smoother. In the current work, the optimisation approach is studied.

The choice of optimisation algorithm has been a matter of research. Given that history matching problems are dissimilar depending on the parameterisation, observed data type, fluid flow characteristics and other factors, one algorithm does not perform the same for every problem. Thus far, many algorithms have been analysed. Where the correlation between uncertain parameters and observed data are highly nonlinear, gradient-based and variable metric algorithms may perform incompetently (Oliver and Chen, 2011). On the other hand, population-based algorithms, such as Genetic Algorithm (Romero and Carter, 2001, Sayyafzadeh et al., 2012b), Differential Evolution (Hajizadeh et al., 2010), Particle Swarm Optimisation (Mohamed et al., 2010) and Artificial Bee Colony (Sayyafzadeh et al., 2012a), can work well in highly nonlinear and discrete problems. However, all these population-based algorithms, similar to classical stochastic algorithms, have a common drawback. They require many fitness function calls to search the solution space for the optimum point, which makes these algorithms computationally demanding (Oliver and Chen, 2011).

Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) is relatively faster than the other population-based algorithms, due to its quasi-gradient search embedded into the algorithm. This algorithm has been used in several studies for well placement and production optimisation (Forouzanfar et al., 2016, Awotunde and Naranjo, 2014, Bouzarkouna et al., 2012). However, less attention has been paid to CMA-ES in history matching problems. Schule-Riegert et al. (2009) used CMA-ES in conjunction with EnKF in which CMA-ES adjusts the initial ensemble for data assimilation. One of the reasons that CMA-ES did not gain much popularity in history matching is its computational costs. This limitation makes CMA-ES less attractive, particularly compared to assimilation techniques which can handle many variables and require much less computation.

In this study, to reduce the computational costs of CMA-ES, a surrogate-assisted CMA-ES is proposed in which an ensemble of surrogates is trained (and re-trained) and replaces the exact function (reservoir simulation) partially. A generation-based evolution-control (model management) technique is used to control the use of reservoir simulation, during the optimisation (tuning). The role of evolution-control is to avoid potential misdirection, as a result of relying solely on the approximation measure, and at the same time, to reduce the computations, as a result of using the exact measure (reservoir simulation).

Surrogate modelling, also known as proxy modelling and response surface modelling, has a long history in Petroleum engineering. The emphasis in the literature was generally on the improvement of surrogate-modelling by applying different offline-learning sampling strategies (design of experiments) and using different surrogates, e.g., artificial neural network (ANN), spline and polynomial regression. Zubarev (2009) wrote a review article about these techniques, applied in Oil&Gas industry. Online-learning techniques have been less studied. Maschio and Schiozer (2014) used ANN retrained throughout the process; their selection criterion was based on the distance of the samples. Sayyafzadeh (2017) used a self-adaptive proxy with an online-learning scheme, for well-placement optimisation problems. In another study, a similar technique was implemented in a parallel tempered Metropolis-Hastings algorithm for approximating posterior density function (Sayyafzadeh, 2016). In these online-

learning studies, promising results were obtained which are in line with other disciplines' outcomes (Razavi et al., 2012, Jin, 2011).

Due to the high nonlinearity between uncertain parameters and reservoir response, in this work, the objective function is broken down into multiple functions, and therefore an ensemble of surrogates (Gaussian process) is used to approximate the exact function. A similar proxy-ensemble approach, opposed to a single-proxy approach, was investigated by He et al. (2015) in which promising results were obtained.

In this study, the main focus is on designing a role (workflow) to optimally use the reservoir simulation during the optimisation by CMA-ES, for real history matching problems. The following concepts and techniques are combined in the workflow to accelerate CMA-ES, surrogate-ensemble, online-learning scheme and generation-based evolution-control. It is discussed why such a combination does not deteriorate the search performance of CMA-ES and can result in fitting history matched models, with reasonable computation costs. In the following sections, history matching formulation, CMA-ES, Gaussian Process (DACE package) and the workflow of the proposed algorithm are explained. The methodology section is followed by Experimental setup section in which two cases, a semi-synthetic model and real case model, are explained. In Results section, the proposed algorithm is compared against the typical CMA-ES. In the last section, some conclusive remarks are summarised.

Problem formulation

In history matching problems, the mode(s) (m^*) of posterior probability density function, $f_M(m|D=dobs,I)$, is sought, equation 1.

$$\arg\max f_M(m|D = d_{obs}, I), subject\ to: m \in \mathfrak{M} \qquad (1)$$

m is an outcome (model) of random variable M and is a column vector with the size of N_m in $\mathfrak{M}$ (model space). m denotes a reservoir simulation input data set (initial, boundary conditions and coefficients of the governing equations). d_{obs} is an outcome of random variable D and is a column vector with the size of N_{obs} in $\mathcal{D}$. d denotes a simulated reservoir response, and d_{obs} is the actual reservoir response, gathered through its life. I is the relevant background information used to build the initial model.

With the following assumptions, the history matching problems can be formulated as equation 2; i- d_{obs} contains random measurement error which has a Gaussian distribution with a zero mean and covariance matrix of C_D, ii- the modelling error of forward operator is zero, iii- prior distribution $f_M(m|I)$ is a joint distribution of X with a uniform distribution (in this study, it is assumed that the outcomes of a priori X is equally probable) and Y with a Dirac delta distribution. X denotes the uncertain dimensions of M, and Y denotes the remaining (known) dimensions of M. If a re-parametrisation technique is applied, x will be the coordinates in the new system.

$$\arg\min \psi(x), subject\ to: x \in \mathcal{X} \qquad (2)$$

$$\psi(x) = \frac{1}{2}(h(x,y) - d_{obs})^T {C_D}^{-1}(h(x,y) - d_{obs}) \qquad (3)$$

h is the forward problem that maps $\mathfrak{M}$ onto $\mathcal{D}$ (h: $\mathfrak{M} \rightarrow \mathcal{D}$). Equation 2 is the same as maximum likelihood estimation (MLE). A numerical optimisation should be applied to find the optimum point(s) of this function. In this study, it is aimed to approximate the global optimum (the most likely model).

Covariance Matrix Adaptation Evolutionary Strategy

CMA-ES algorithm was originally introduced by Hansen and Ostermeier (2001). It is based on the evolution of organisms where mutation and selection are the primary search operators. The operators

are applied in a loop, and each loop is referred to as a generation. The algorithm continues to construct new generations until at least one of the termination criteria is met. The population of each generation is created by sampling the Multivariate Normal Distribution of the search space. In each generation, λ offsprings are ranked based on their fitness value, and the parents for the next generation are selected by a weighted average of μ of the best ranked offsprings of the current generation.

The search begins by generating λ members by sampling the multivariate normal distribution around the current mean ($\bar{x}$).

$$x_k^{g+1} \sim N\left(\bar{x}^g, {\sigma^g}^2 C^g\right) \quad for\ k = 1, \dots, \lambda \tag{4}$$

Where x_k^{g+1} is the k^{th} member of the samples generated at generation ($g+1$). $N\left(\bar{x}^g, {\sigma^g}^2 C^g\right)$ denotes the multivariate normal distribution of the search space with the mean of $\bar{x}^g$, step size of σ^g and covariance matrix of C^g at generation g. To define the complete loop, all that remains is to calculate $\bar{x}^{g+1}$,σ^{g+1} and C^{g+1} for the next generation.

To determine the new mean $\bar{x}^{g+1}$ of the search distribution, the population of generation $g+1$ ($x_1^{g+1}, \dots, x_\lambda^{g+1}$) are ranked based on their fitness.

$$\psi\left(x_{1:\lambda}^{g+1}\right) \leq \psi\left(x_{2:\lambda}^{g+1}\right) \leq \cdots \leq \psi\left(x_{\lambda:\lambda}^{g+1}\right) \tag{5}$$

Where the index i:λ denotes the ith best individual out of ($x_1^{g+1}, \dots, x_\lambda^{g+1}$). The new mean is a weighted average of μ selected best fitness ranked individuals from the sample population of generation $g+1$.

$$\bar{x}^{g+1} = \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{g+1} \tag{6}$$

w_i is the weighting for the ith individual. These weights are all greater than zero and should add up to one $\sum_{i=1}^{\mu} w_i = 1$.

The covariance matrix of the distribution is updated in such a way that increases the likelihood of the previously successful search steps. A weighted selection mechanism is used. In this mechanism the ($g+1$) generation samples ($x_1^{g+1}, \dots, x_\lambda^{g+1}$) are ranked and weighted based on their fitness, and the best μ samples are used to update the covariance matrix.

$$C_\mu^{g+1} = \sum_{i=1}^{\mu} w_i (x_{i:\lambda}^{g+1} - \bar{x}^g)(x_{i:\lambda}^{g+1} - \bar{x}^g)^T \tag{7}$$

More reliable covariance matrices would require larger population sizes, which in turn has an adverse effect on the speed of the search. To maintain reliability and speed efficiency at the same time information from previous generations is added to the calculation of covariance matrix. To assign a higher weight to recent generations a learning rate $0 < C_{cov} \leq 1$ is introduced. To increase reliability of the covariance matrix, evolution path and to control the speed of the algorithm a step size is defined.

The CMA-ES algorithm used in this study is based on one developed by Hansen (2006).

Gaussian process

In Gaussian process, a polynomial is fitted as a global function and kriging is used to estimate local deviations.

$$\acute{\psi}(x) = F(x)\beta + Z(x) \qquad (8)$$

In the above equation $\acute{\psi}(x)$ is the approximated value, $F(x)$ is the fitted polynomial with parameters β, and Z is the error derived by kriging, where the error is assumed to have a normal distribution.

In this study, a zero-degree polynomial was used as the regression model of the proxy along with a linear correlation model for kriging. A MatLab toolbox, called DACE is used to fit the response surface. For more details on the Gaussian process, the reader can refer to http://www2.imm.dtu.dk/projects/dace/dace.pdf

Workflow of the proposed algorithm

The algorithm, in the first few generations, uses only the exact function (EF), until enough samples are gathered to train the approximation function (AF) and also to allow the early exploration to be conducted by the EF. This threshold should be predefined. Once the first training takes place, the evaluation of the individuals in the subsequent generations is carried out either by the AF or the EF. If the EF is used, the individual of that generation will be added to the sample pool. The sample pool with a minimum of one new member is used to retrain the AF. This creates an online-learning scheme. The evaluations by the EF throughout the optimisation can prevent potential misdirection arising from a low-fidelity AF, and the online-learning scheme can increase the fidelity of the AF over the course of the optimisation.

In order to decide which measure (EF or AF) to be used for each individual, a generation-based evolution-control technique is used. In a generation-based control, all the individuals of one generation are either evaluated by the AF or EF. In order to maintain the online-learning aspect, a stochastic rule, based on a predefined probability, is used to select the generations that must be evaluated by the EF. To find a suitable probability, a series of sensitivity analyses were conducted, and the best value found was 0.25. This control fits well in surrogate-assisted algorithms that the selection operator performs according to the rank of solutions (like CMA-ES), instead of the actual fitness value. As mentioned above, CMA-ES needs the top μ individuals in each generation to update the mean and covariance, and the actual (exact or approximated) fitness value is not utilised directly. Generally speaking, approximation functions are better in finding the rank of solutions (sorting), instead of approximating the fitness. This is shown in the results section, and the rank correlation between the evaluation by the AF and EF is used as an indication for the fidelity of AF.

The trend of fitness improvement in CMA-ES is not monotonic, as the best ever-found solution is not being transferred to the subsequent generations, i.e., there are noises in the trend. The noisiness caused by implementing a generation-based control does not have a severe impact on the search and the introduced disruptions (if any) can be treated as mutations.

$\psi(x)$ is the exact function. C_D is usually a diagonal matrix, and its entries are variances, σ_i^2 (measurements errors corresponding to d_{obs_i}). With this assumption, $\psi(x)$ can be split as below.

$$\psi(x) = \frac{1}{2}(h(x,y) - d_{obs})^T {C_D}^{-1}(h(x,y) - d_{obs}) = \frac{1}{2}\sum_{i=1}^{N_{obs}} \left(\frac{h(x,y)_i - d_{obs_i}}{\sigma_i}\right)^2 \qquad (9)$$

$$\psi(x) = \frac{1}{2}\left(\sum_{i=1}^{N_{obs}/k} \left(\frac{h(x,y)_i - d_{obs_i}}{\sigma_i}\right)^2 + \cdots + \sum_{i=\frac{(k-1)N_{obs}}{k}+1}^{N_{obs}} \left(\frac{h(x,y)_i - d_{obs_i}}{\sigma_i}\right)^2\right) \qquad (10)$$

$\psi_j(x)$ is defined as equation 11.

$$\psi_j(x) = \sum_{i=\frac{(j-1)N_{obs}}{k}+1}^{\frac{jN_{obs}}{k}} \left(\frac{h(x,y)_i - d_{obs_i}}{\sigma_i}\right)^2 \qquad (11)$$

Therefore,

$$\psi(x) = \sum_{j=1}^{k} \psi_j(x) \qquad (12)$$

Each of $\psi_j(x)$ is approximated with a separate proxy, which are used to approximate $\psi(x)$.

$$\acute{\psi}(x) = \sum_{j=1}^{k} \acute{\psi}_j(x) \qquad (13)$$

If the number of observations is small, k can be equal to N_{obs}, i.e., one proxy per misfit. However, where N_{obs} is large, the time required to train the proxies may exceed the time required for the evaluation via the EF. Hence, in such cases, the number of proxies (k) should be less than N_{obs}. In order to define k, either a sensitivity analysis can be conducted to find an optimal number of proxies or the observations can be categorised based on time, property or well. For instance, the misfit corresponding to bottomhole pressures (BHPs) of one well can be one proxy and water-cuts another proxy.

In this study, DACE package is used to obtain the $\acute{\psi}_j(x)$. The natural logarithm of $\psi_j(x)$ is used and approximated by DACE, as it may have a smoother landscape, compared to $\psi_j(x)$ (Zubarev, 2009).

CMA-ES starts with a random vector between zero and one, and given the scale of each dimension of x might be different, a transformation of variables as below is used.

$$z = \frac{L_B + U_B}{2} + \left(\left(\frac{U_B - L_B}{2} e^t\right) \odot I_{N_x}\right) x \qquad (14)$$

U_B and L_B are column vectors representing upper and lower bounds for each dimension of x. I_{N_x} is the identity matrix with a size of N_x. e^t is (1,1,…,1) with the size of N_x.

CMA-ES creates solutions that might be unfeasible (out of bounds). In this study, to handle this issue, a penalty term is defined, and if a solution is not within the predefined range, EF or AF will not be used to evaluate that solution, and instead, a fitness value based on Euclidian distance from L_B, if lower than L_B, or from the upper bound U_B, if exceeding for the U_B, will be assigned (dimension-based). This method directs the search towards feasible solution space.

To save time in the training part, all the k proxies are trained in parallel, by assigning the computation of each to one core. In addition, given that the generation-based proxy control does not have any impact on parallelisability of CMA-ES, for generations that should be evaluated by the EF, up to λ simulations could be carried out simultaneously. The proposed workflow was implemented in our in-house interface, Flexible assisted-History Matching interface (FaHMi).

Experiential setup

In order to evaluate the proposed algorithm, two history matching problems, PUNQ-S3 model and a real model are used.

PUNQ-S3

PUNQ-S3 is a semi-synthetic model used to benchmark history-matching algorithms. It is a part of a real field performed by Elf Exploration Production. The reservoir has a fairly strong aquifer, from two sides (southern and eastern), and from the other two sides, it has no-flow boundary. The reservoir has a gas cap in the dome. There are 6 producers in the reservoir which are producing oil for 16.5 years. The location of these wells can be seen in the Fig. 1. The history data contains three percent Gaussian noise with zero mean. The total number of observations are 2296 through 82 time reports, and all are used to tune the model. The observation data are cumulative field oil production, cumulative field gas production, cumulative field water production, wells water-cut, wells oil production rate, wells bottomhole pressure and wells gas-oil ratio. The simulation model is black-oil with three phases and consists of 19×28×5 grid blocks with 1761 active.

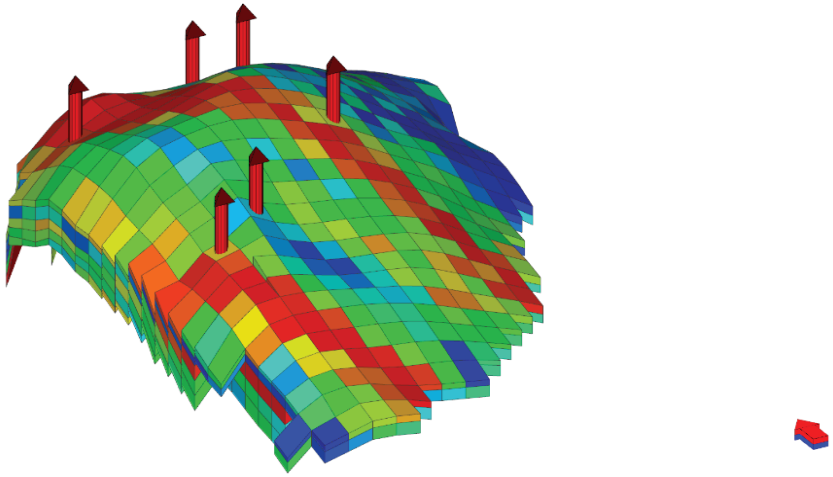

Figure 1 *PUNQ-S3 reservoir model.*

It is assumed that the parameters listed in table 1 are unknown. The true values were used to generate the noise-free history.

Table 1 *PUNQ-S3 uncertain parameters.*

Parameter	*True*	*Initial*	*LB*	*UB*	*Distribution*
Multiplier Perm X	*1*	*0.1*	*0.1*	*2*	*Uniform*
Multiplier Perm Y	*1*	*0.1*	*0.1*	*2*	*Uniform*
Multiplier Perm Z	*1*	*0.1*	*0.1*	*2*	*Uniform*
Multiplier Poro	*1*	*0.1*	*0.1*	*2*	*Uniform*
Aquifer#1 Perm (mD)	*137.5*	*500*	*1*	*800*	*Uniform*
Aquifer#2 Perm (mD)	*137.5*	*500*	*1*	*800*	*Uniform*
Water Oil Contact Depth (m)	*2395*	*2410*	*2385*	*2420*	*Uniform*
Rock Compressibility (1/bar)	*0.00045*	*0.0006*	*0.0001*	*0.0006*	*Uniform*

X Model

X is a light oil field currently under production. There are 10 producers in the reservoir and 6 water injectors. The history period is 20 months, gathered daily, and water injection started after 9 months. Multiple crests, some with unknown gas cap presence, and significant variations in solution gas content through the field, guarantee that history matching of the field will be problematic.

Wells in the simulation model are controlled by historical fluid volume rates. For this exercise, 97 data points from all the observations were used as a measure for history matching goodness. These points were selected by an expert familiar with the field and data. These 97 points include gas-oil ratio, bottomhole pressure, tubing head pressure and water-cuts of the wells and gridblocks pressures

(measured using downhole tools and simulated using pseudo wells) over different timesteps. For each data point, a tolerance was defined. Half of the tolerance was assumed as σ_i.

The simulation model is black-oil with three phases and consists of 79×145×86 grid blocks of which 480k are active. The model has two SCAL regions, and there are multiple gas-oil contacts. The reservoir consists of three main geological layers, 1, 5 and 6. There are some faults that create discontinuities within all the three geological layers. The reservoir is fairly heterogeneous, and permeability range is from a few m-Darcy to more than 500 m-Darcy. The boundaries of the reservoirs are no-flow.

Table 2 *X model uncertain parameters.*

Parameter	*Initial*	*LB*	*UB*	*Distribution*
Well Transmissibility factor (x25)	*1*	*0.33*	*3*	*Uniform*
Multiplier Poro Layer-1 North	*0.5*	*0.5*	*2*	*Uniform*
Multiplier Poro Layer-1 South	*0.7*	*0.7*	*2*	*Uniform*
Multiplier Poro Layer-5&6 (x4)	*1*	*0.5*	*2*	*Uniform*
k_v/k_h (x3)	*1*	*0.2*	*5*	*Uniform*
Gas-Oil-Contact Layer-1 (ft)	*10700*	*10200*	*10802*	*Uniform*
Gas-Oil-Contact Layer-5 (ft)	*11425*	*11410*	*11435*	*Uniform*
Gas-Oil-Contact Layer-6 North (ft)	*11590*	*11525*	*11615*	*Uniform*
Gas-Oil-Contact Layer-6 South (ft)	*11670*	*11420*	*11720*	*Uniform*
Gas-Oil-Contact Layer-6 Updip (ft)	*11570*	*11525*	*11615*	*Uniform*
3-phase Corey's (region#1) (x6)	-	*[0.05 2 1.5 -10 1 1]*	*[0.3 3.5 3 -5.5 3 2.5]*	*Uniform*
3-phase Corey's (region#2) (x6)	-	*[0.05 2 1.5 -10 1 1]*	*[0.3 3.5 3 -5.5 3 2.5]*	*Uniform*
Scaled Critical Gas Saturations (x3)	*0.025*	*0*	*0.05*	*Uniform*
Dissolved gas-oil ratio vs Depth (Layer-6) Depth (ft)	*11650*	*11640*	*11650*	*Uniform*
Dissolved gas-oil ratio vs Depth (Layer-6) Rs (MSCF/STB)	*2.65*	*2*	*5*	*Uniform*
X-dir Transmissibility Multiplier (Layer-6) across a line	*1*	*1E-6*	*1*	*Uniform*
Vaporized Oil-Gas Ratio versus depth (STB/MSCF)	*0.048*	*0.04*	*0.2*	*Uniform*
k_x, k_y, k_z multiplier for Layer 1.	*1*	*0.5*	*5*	*Uniform*

The parameters listed in table 2 are the uncertain parameters. The total number of uncertain parameters is 59. There are 25 well transmissibility factors, one for each injector and two for each producer (one for 5th geological layer and one for 6th geological layer), except for one well which was completed only in the top layer (layer-1). Six porosity multipliers were used, two for each geological layer (one for southern part and one for northern part). There are three vertical to horizontal permeability ratios, one for each layer. Five gas-oil-contacts were considered as matching parameters at different layers and different locations. Gas-oil and water-oil relative permeability parameters were also considered as uncertain parameters for each region. Three scaled critical gas saturations. Dissolved gas-oil ratio versus Depth in layer 6 was assumed unknown which requires two parameters to define it. A multiplier for transmissibility in x-dir in layer 6 between producer 7 and 8 is another uncertain parameter. Vaporised Oil-Gas Ratio is another uncertain parameter, however, its value is assumed independent of depth. A multiplier for permeability in all three directions for the first geological layer is the last uncertain parameter.

The simulation of the model takes in average 400 seconds using a 20 core workstation.

Results

The results of each history matching problem are studies separately.

PUNQ-S3 results

History matching was carried out 100 times with the conventional CMA-ES (without proxy). For each run, a random initial mean $\bar{x}^0$ was used. The maximum number of simulations for each run was set to

500. By repeating the optimisation, the impact of randomness on the assessment reduces. The average and standard deviation of $\psi(x)$ versus the number of the EF (reservoir simulations) calls are used to assess the effect of adding proxies. The solid lines in Fig. 2 & 3 correspond to the results of the unassisted-algorithm.

History matching was also carried out 100 times with the proposed surrogates' employment. CMA-ES configurations were exactly the same as the previous one. For each run, similarly, a random initial mean was used. The maximum number of simulations for each run was also the same. The surrogates were not applied until 100 samples are gathered, and after that the probability for using the EF to evaluate the individuals of a generation was set to 0.25. Following each generation where the EF was applied, the surrogates were retrained with the updated sample set. The average and standard deviation of $\psi(x)$ versus the number of the EF (reservoir simulations) calls are compared against those obtained by the unassisted-algorithm. The dashed lines in Fig. 2 & 3 correspond to the results of the assisted-algorithm.

The comparison shows that for any predefined computation (between 100 and 500 simulations), the proposed algorithm resulted into a model with a much lower misfit value, compared to the one obtained by the unassisted-CMA-ES. For instance, in order to reach a misfit of 5×10^4, the unassisted-algorithm, in average, needed almost 500 simulations, while the same quality model was obtained by the proposed algorithm with less than 250 simulations. It is expected that both algorithms eventually get to the same misfit.

Although the initial models were not the same, both algorithms performed (in average) very similarly in the early stages when the EF was the only measure (before 100 simulations).

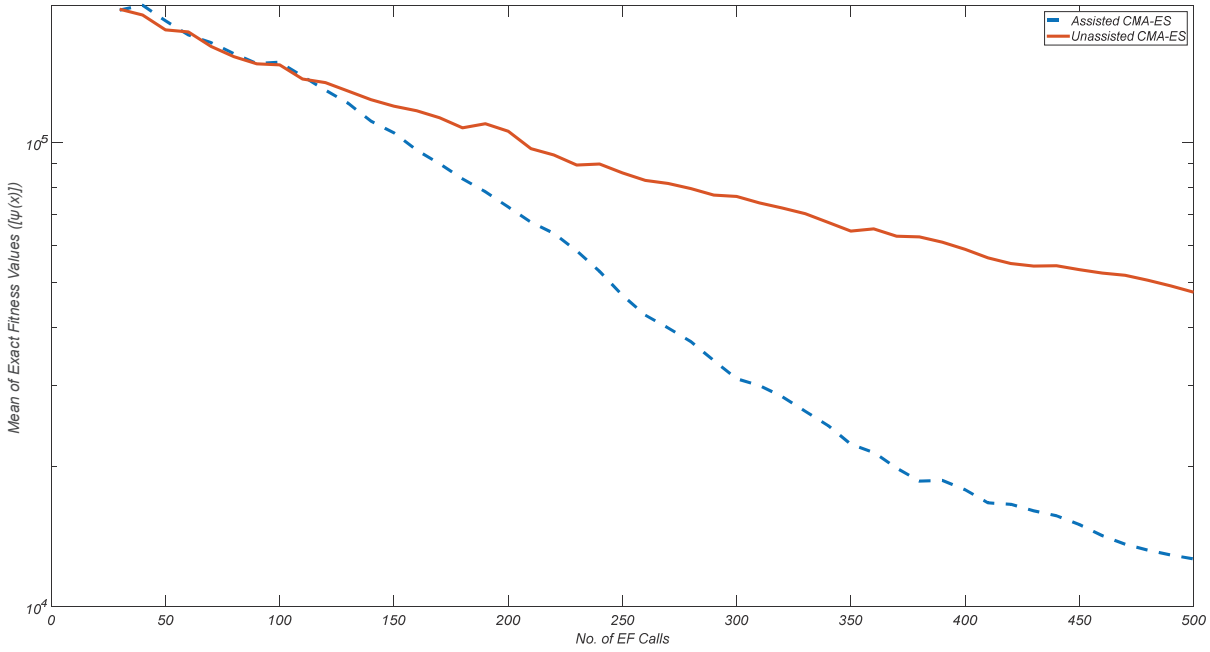

Figure 2 *Mean of $\psi(x)$ versus number of simulations- PUNQ-S3.*

The standard deviations were also compared to ensure the reliability of the proposed workflow. As it can be seen, it has a lower standard deviation, compared to the unassisted-algorithm for almost any number of simulations.

In this exercise, all the observations were used to calculate misfit, and since the number of observations was large, it was not practical to have 2296 proxies. Hence, $\psi(x)$ was split into 28 functions (one for each property) and one proxy was used per function. In order to analyse the improvement by introducing

the surrogate ensemble, compared to a single surrogate, 80 percent of the initial sample set (with the size of 100) were selected at random to train i- a single surrogate (approximating $\psi(x)$ directly) and ii- an ensemble of surrogates (approximating $\psi_1(x) \dots \psi_{28}(x)$), and they were used to predict $\psi(x)$ for the remaining (20 percent) samples. Spearman's rank correlation was calculated to compare the two approaches. This correlation was calculated for all the 100 runs and shown in Fig. 4.

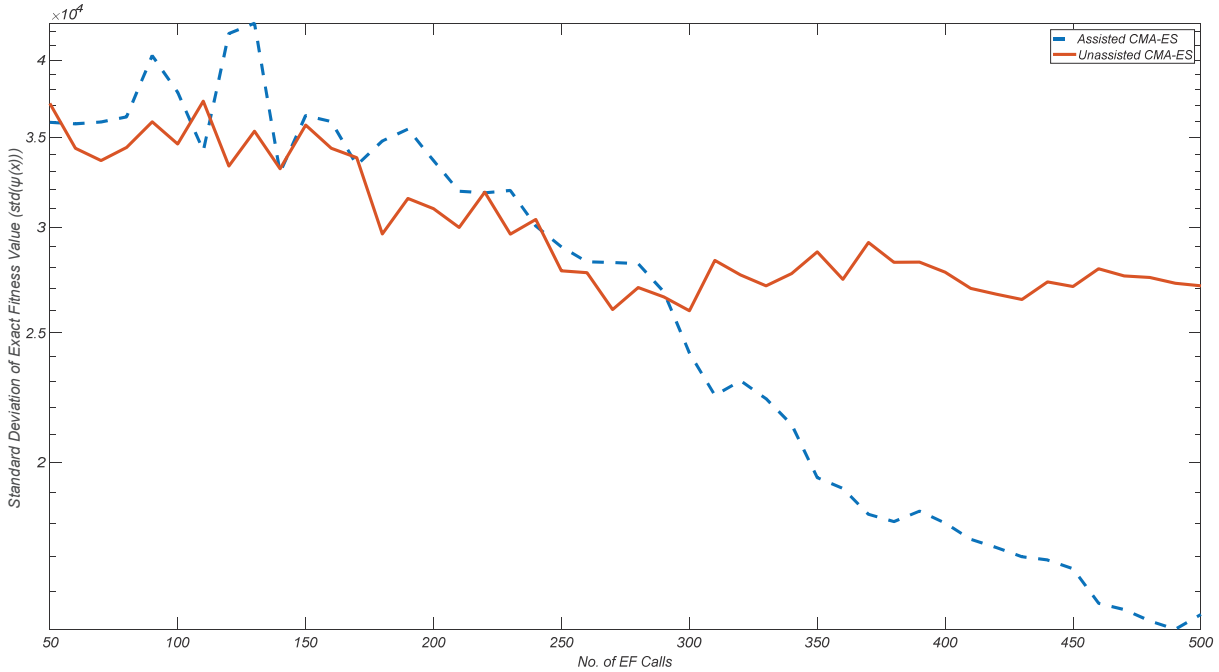

Figure 3 *Standard Deviation of $\psi(x)$ versus number of simulations- PUNQ-S3.*

As it can be seen, the one with surrogate ensemble has a much higher rank correlation, compared to the classical single proxy. The mean of rank correlation for the surrogate ensemble is 0.87, while this value is 0.45 for the single surrogate. The high rank correlation indicates that the ranking of individuals obtained by the AF for those generations where the AF is used, are close to the ranking obtained by the EF. The similar approach was used to find the optimal DACE configurations.

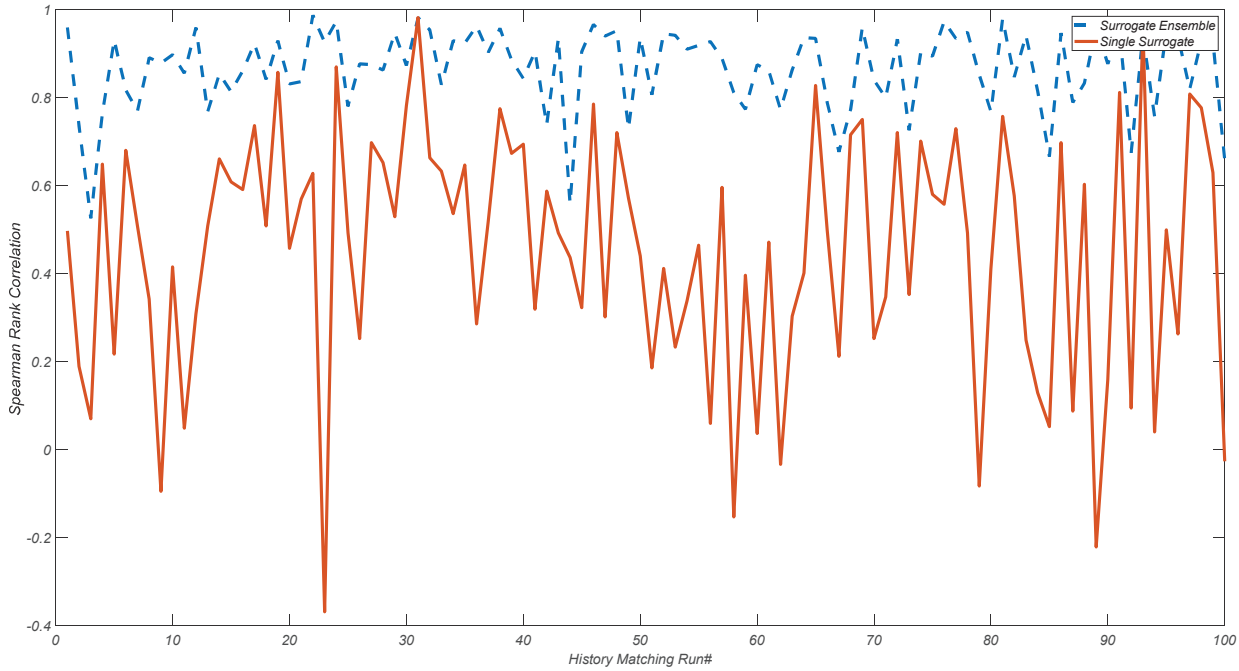

Figure 4 *Rank correlation for all 100 history matching times – PUNQ-S3.*

The best and average solution of 100 runs are shown in table 3. The solutions (the best and the average) obtained by the assisted algorithm are closer to the true solution, compared to those obtained by the unassisted-algorithm.

Table 3 *PUNQ-S3 results.*

Parameter	*True*	*Best of Assisted*	*Best of Unassisted*	*Mean of Assisted Solutions*	*Mean of Unassisted Solutions*
Multiplier Perm X	*1*	*1.028511*	*0.828452*	*0.897789*	*0.821246*
Multiplier Perm Y	*1*	*1.024132*	*0.821276*	*0.983703*	*0.930722*
Multiplier Perm Z	*1*	*0.941926*	*1.105564*	*1.198156*	*1.228287*
Multiplier Poro	*1*	*0.99097*	*1.024686*	*1.016029*	*1.015636*
Aquifer#1 Perm (mD)	*137.5*	*116.0391*	*176.9859*	*126.1072*	*282.2366*
Aquifer#2 Perm (mD)	*137.5*	*193.473*	*134.1614*	*397.443*	*419.4391*
Water Oil Contact Depth (m)	*2395*	*2395.229*	*2395.008*	*2395.308*	*2396.343*
Rock Compressibility (1/bar)	*0.00045*	*0.000431*	*0.000474*	*0.000463*	*0.000469*

X Model results

Due to high computational time, history matching was conducted only three times, once with each of the following algorithms, unassisted-CMA-ES, assisted-CMA-ES starting after 100 simulations and assisted-CMA-ES starting after 200 simulations. In all of them, the same initialisation was used, i.e., in the assisted ones, the optimisation did not start from the beginning, and instead it resumed the unassisted one from 100 and 200 simulations. By this mean, the impact of randomness on the assessment can be neglected.

The configurations of DACE and CMA-ES are the same as the ones used for PUNQ-S3 model. In this exercise, 97 observations were used to calculate misfit. Since the number of observations are limited, one surrogate per misfit was applied (97 in total). The maximum number of simulations of 500 was used for all the three algorithms. The computations associated with the training of the surrogates was negligible, compared to the simulations. Hence, the criterion for assessment was the number of simulations.

Fig. 5 shows the result of history matching with the three algorithms. The exact misfit obtained by the assisted-algorithm starting from 100 simulations was the best. Both assisted algorithms resulted into a model with a lower misfit, compared to the unassisted CMA-ES, and almost at any number of simulations, the solutions delivered by the assisted ones have a lower misfit. This indicates that adding surrogates with the proposed implementation can outperform the unassisted algorithm, even for relatively complex history matching problems. Between the two assisted algorithms, the one starting from 100 simulations outperformed the other one. This can indicate that applying surrogates with the proposed method does not misdirect the search, even if implemented from the early stage of search. This is because of the model management applied. Due to the noise in the trend, it is hard to quantify the computation reduction. However, to provide an approximate value, for instance, to get the misfit (natural logarithm) of 7.25, the unassisted CMA-ES called the simulation almost 400 times, while the assisted ones (after 200 and 100 simulations) needed roughly 250 and 150, respectively.

As it can be seen, a significant improvement in most of data points were obtained. However, some of the points are still not in the acceptable tolerance. This exercise was continuing up to 2000 simulation, and no better solutions were found. This can indicate that to find a better solution, the simulation model (particularly well models), history data or the parameterisation might need a revisit, which does not fall in the scope of this work.

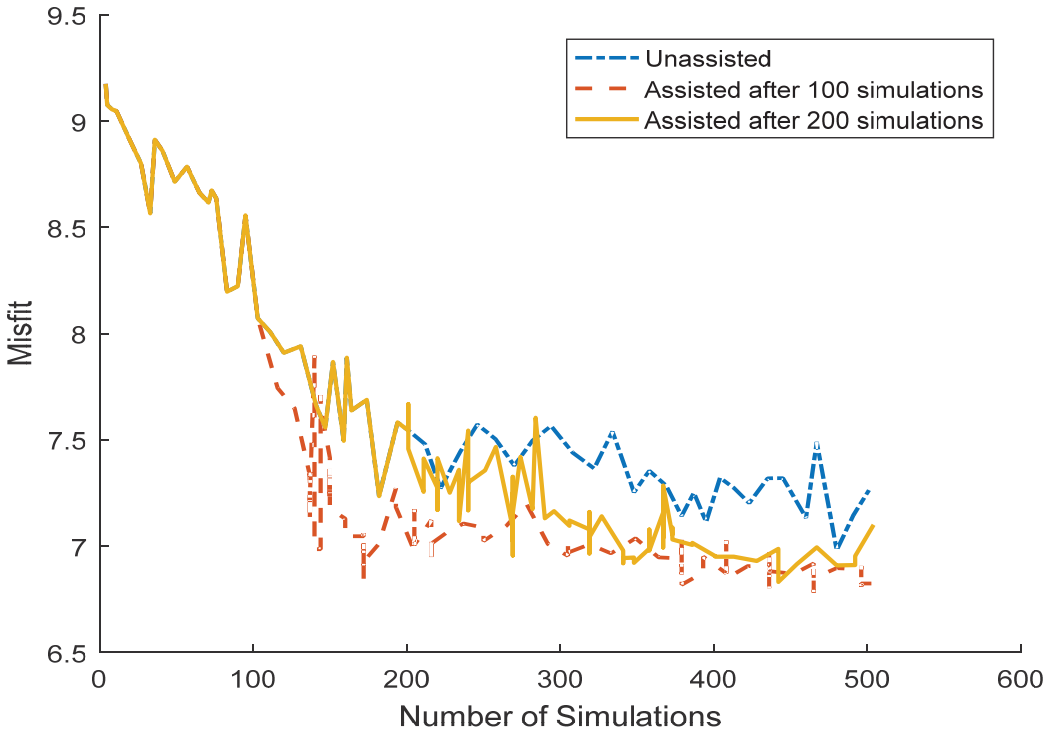

Figure 5 *$ln(\psi(x))$ versus number of simulations – X Model.*

The following figures show the simulation using the history matched model (obtained by the assisted CMA-ES, starting from 100) and the initial model, against the historical data. The green dots are the assigned tolerances for those 97 points from the history. The history data are the blue lines, the simulation results using the initial models are shown by red, and the yellow ones are the best history matched solutions found by the proposed algorithm.

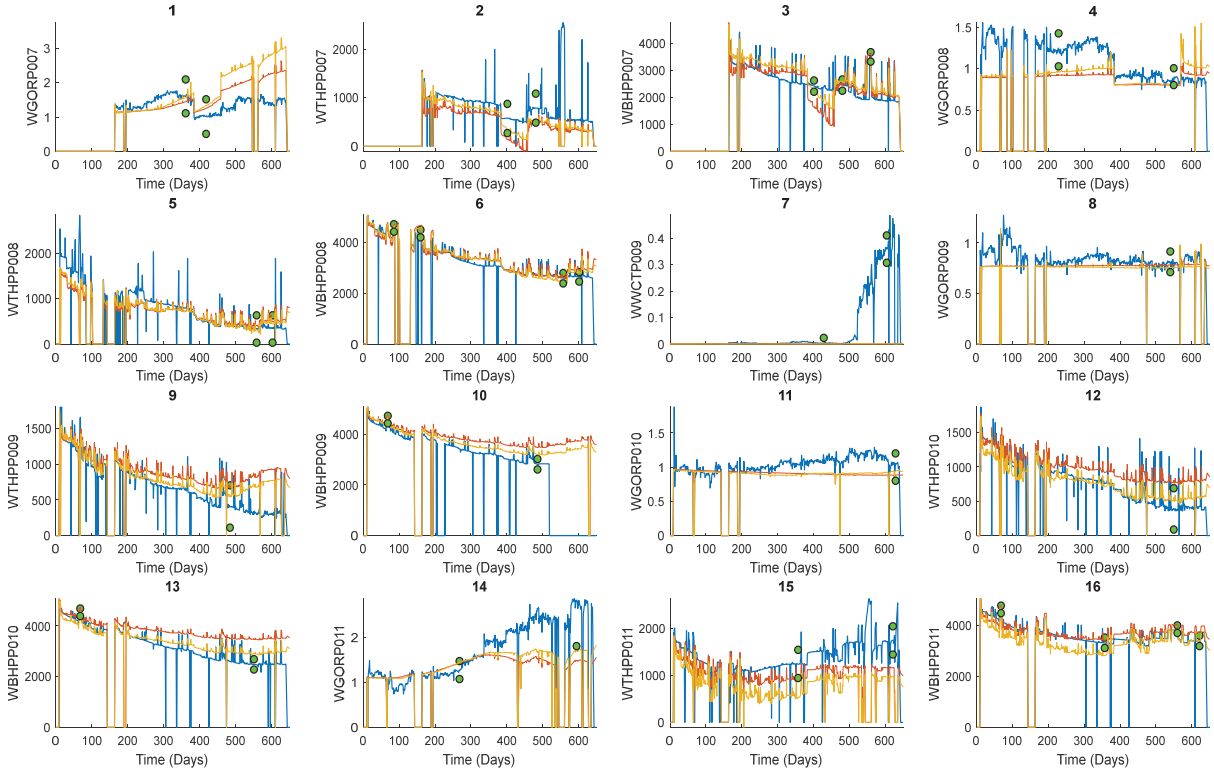

Figure 6-a *History matching results – X Model.*

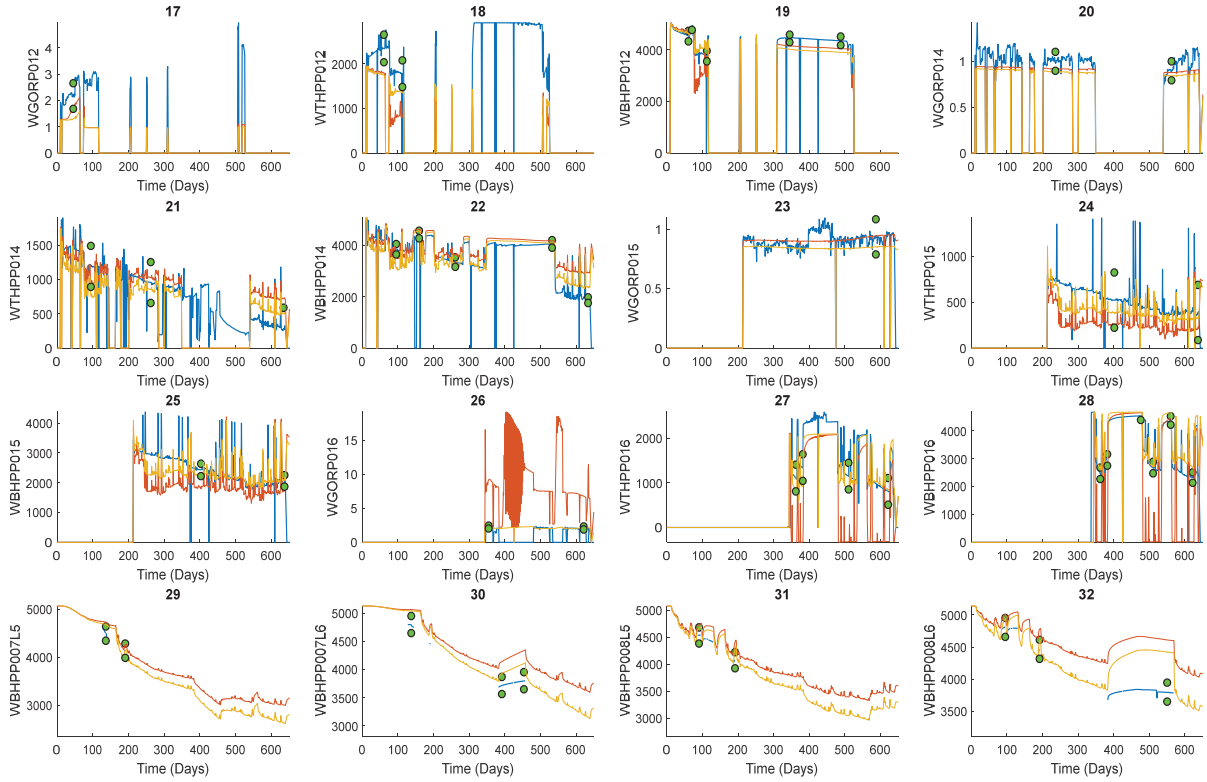

Figure 6-b *History matching results – X Model.*

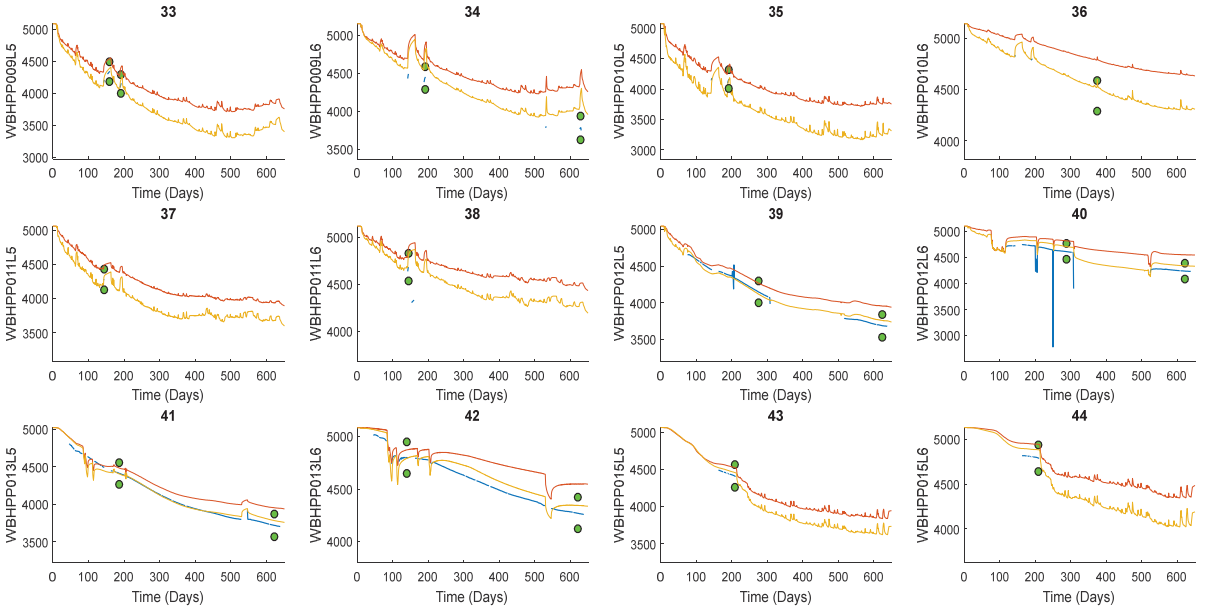

Figure 6-c *History matching results – X Model.*

Conclusions

In this study, an online-learning ensemble-proxy-based algorithm was introduced. The technique was implemented in CMA-ES and tested using two history matching cases. In order to break down the complexity of the fitness function, in terms of response surface modelling, the exact fitness function was split into multiple functions which were individually modelled via a separate proxy (Gaussian process). A significant improvement (measured by rank correlation) was obtained by the proxy-

ensemble approach, compared to a single proxy modelling. In the proposed workflow, a combination of the following techniques was used, generation-based model management with a predefined probability of 0.25, ensemble of proxies, online-learning which continuously improves the fidelity of the proxies over the course of history matching. Using the proposed implementation of proxies into CMA-ES, a substantial and reliable acceleration was observed, and the computations reduced up to 50% and 65% for the PUNQ-S3 model and X Model, respectively. The proposed workflow fits well in any population-based algorithm that does not require the actual fitness value to direct the search.

Acknowledgements

Authors are grateful to Santos Ltd. for providing the real field dataset and giving the permission to publish the study. The funding for this work was provided by Santos Ltd and Australian School of Petroleum (ASP), through the project "Full-Parameterised History Matching by Stochastic Wavelet Bases". The authors also wish to acknowledge Schlumberger, Rock Flow Dynamics and Mathworks for supporting this study by providing licenses for their software

References

AWOTUNDE, A. A. & NARANJO, C. 2014. Well Placement Optimization Constrained to Minimum Well Spacing. *SPE Latin America and Caribbean Petroleum Engineering Conference.* Maracaibo, Venezuela: Society of Petroleum Engineers.

BOUZARKOUNA, Z., DING, D. Y. & AUGER, A. 2012. Well placement optimization with the covariance matrix adaptation evolution strategy and meta-models. *Computational Geosciences,* 16**,** 75-92.

FOROUZANFAR, F., POQUIOMA, W. E. & REYNOLDS, A. C. 2016. Simultaneous and Sequential Estimation of Optimal Placement and Controls of Wells With a Covariance Matrix Adaptation Algorithm.

HAJIZADEH, Y., CHRISTIE, M. A. & DEMYANOV, V. 2010. History matching with differential evolution approach; a look at new search strategies. *SPE EUROPEC/EAGE Annual Conference and Exhibition.* Barcelona, Spain: Society of Petroleum Engineers.

HANSEN, N. 2006. The CMA Evolution Strategy: A Comparing Review. *In:* LOZANO, J. A., LARRAÑAGA, P., INZA, I. & BENGOETXEA, E. (eds.) *Towards a New Evolutionary Computation: Advances in the Estimation of Distribution Algorithms.* Berlin, Heidelberg: Springer Berlin Heidelberg.

HANSEN, N. & OSTERMEIER, A. 2001. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evolutionary Computation,* 9**,** 159-195.

HE, J., XIE, J., WEN, X.-H. & CHEN, W. 2015. Improved Proxy For History Matching Using Proxy-for-data Approach And Reduced Order Modeling. *SPE Western Regional Meeting.* Garden Grove, California, USA: Society of Petroleum Engineers.

JIN, Y. 2011. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation,* 1**,** 61-70.

MASCHIO, C. & SCHIOZER, D. J. 2014. Bayesian history matching using artificial neural network and Markov Chain Monte Carlo. *Journal of Petroleum Science and Engineering,* 123**,** 62-71.

MOHAMED, L., CHRISTIE, M. A. & DEMYANOV, V. 2010. Reservoir Model History Matching with Particle Swarms: Variants Study. *SPE Oil and Gas India Conference and Exhibition.* Mumbai, India: Society of Petroleum Engineers.

OLIVER, D. S. & CHEN, Y. 2011. Recent progress on reservoir history matching: a review. *Computational Geosciences,* 15**,** 185-221.

RAZAVI, S., TOLSON, B. A. & BURN, D. H. 2012. Review of surrogate modeling in water resources. *Water Resources Research,* 48.

ROMERO, C. E. & CARTER, J. N. 2001. Using genetic algorithms for reservoir characterisation. *Journal of Petroleum Science and Engineering,* 31**,** 113-123.

SAYYAFZADEH, M. Uncertainty quantification using a self-supervised surrogate-assisted parallel Metropolis-Hastings algorithm. ECMOR XIV-15th European Conference on the Mathematics of Oil Recovery, 2016.

SAYYAFZADEH, M. 2017. Reducing the computation time of well placement optimisation problems using self-adaptive metamodelling. *Journal of Petroleum Science and Engineering,* 151**,** 143-158.

SAYYAFZADEH, M., HAGHIGHI, M., BOLOURI, K. & ARJOMAND, E. 2012a. Reservoir characterisation using artificial bee colony optimisation. *APPEA Journal***,** 115-128.

SAYYAFZADEH, M., HAGHIGHI, M. & CARTER, J. N. 2012b. Regularization in History Matching Using Multi-Objective Genetic Algorithm and Bayesian Framework. *SPE Europec/EAGE Annual Conference.* Copenhagen, Denmark: Society of Petroleum Engineers.

SCHULZE-RIEGERT, R. W., KROSCHE, M., PAJONK, O. & MUSTAFA, H. 2009. Data Assimilation Coupled to Evolutionary Algorithms—A Case Example in History Matching. *SPE/EAGE Reservoir Characterization and Simulation Conference.* Abu Dhabi, UAE: Society of Petroleum Engineers.

ZUBAREV, D. I. 2009. Pros and Cons of Applying Proxy-models as a Substitute for Full Reservoir Simulations. *SPE Annual Technical Conference and Exhibition.* New Orleans, Louisiana: Society of Petroleum Engineers.

References

AASUM, Y. & KELKAR, M. G. 1991. An application of geostatistics and fractal geometry for reservoir characterization. *SPE Formation Evaluation,* 6**,** 11-19.

ALZUBAIDI, L., ZHANG, J., HUMAIDI, A. J., AL-DUJAILI, A., DUAN, Y., AL-SHAMMA, O., SANTAMARÍA, J., FADHEL, M. A., AL-AMIDIE, M. & FARHAN, L. 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data,* 8**,** 1-74.

AMEEN, M. S. 2014. Fracture and in-situ stress patterns and impact on performance in the Khuff structural prospects, eastern offshore Saudi Arabia. *Marine and Petroleum Geology,* 50**,** 166-184.

ARJOVSKY, M., CHINTALA, S. & BOTTOU, L. Wasserstein generative adversarial networks. 2017 2017. PMLR, 214-223.

ASTRAKOVA, A. & OLIVER, D. S. 2015. Conditioning truncated pluri-Gaussian models to facies observations in ensemble-Kalman-based data assimilation. *Mathematical Geosciences,* 47**,** 345-367.

AVNIR, D., FARIN, D. & PFEIFER, P. 1985. Surface geometric irregularity of particulate materials: The fractal approach. *Journal of Colloid And Interface Science,* 103**,** 112-123.

AZEVEDO, L., PANEIRO, G., SANTOS, A. & SOARES, A. 2020. Generative adversarial network as a stochastic subsurface model reconstruction. *Computational Geosciences,* 24**,** 1673-1692.

BAI, T. & TAHMASEBI, P. 2020. Hybrid geological modeling: Combining machine learning and multiple-point statistics. *Computers & Geosciences,* 142**,** 104519.

BERTA, D., HARDY, H. & BEIER, R. Fractal distributions of reservoir properties and their use in reservoir simulation. International petroleum conference and exhibition of Mexico, 1994. OnePetro.

BREKKE, H., MACEACHERN, J. A., ROENITZ, T. & DASHTGARD, S. E. 2017. The use of microresistivity image logs for facies interpretations: An example in point-bar deposits of the McMurray Formation, Alberta, Canada. *AAPG Bulletin,* 101**,** 655-682.

CANCHUMUNI, S. W., CASTRO, J. D., POTRATZ, J., EMERICK, A. A. & PACHECO, M. A. C. 2021. Recent developments combining ensemble smoother and deep generative networks for facies history matching. *Computational Geosciences,* 25**,** 433-466.

CHAN, S. & ELSHEIKH, A. 2019a. Parametrization and generation of geological models with generative adversarial networks. *arXiv.org*.

CHAN, S. & ELSHEIKH, A. H. 2019b. Parametric generation of conditional geological realizations using generative neural networks. *Computational Geosciences,* 23**,** 925-952.

CHAN, S. & ELSHEIKH, A. H. 2020. Parametrization of Stochastic Inputs Using Generative Adversarial Networks With Application in Geology. *Frontiers in Water,* 2**,** 5.

CHEN, L., LI, S., BAI, Q., YANG, J., JIANG, S. & MIAO, Y. 2021. Review of image classification algorithms based on convolutional neural networks. *Remote Sensing,* 13**,** 4712.

CRANE, S. D. & TUBMAN, K. M. 1990. Reservoir Variability and Modeling With Fractals. 65th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, September 23-26 1990 New Orleans. Society of Petroleum Engineers: New Orleans: Society of Petroleum Engineers.

DUPONT, E., ZHANG, T., TILKE, P., LIN, L. & BAILEY, W. 2018. Generating Realistic Geology Conditioned on Physical Measurements with Generative Adversarial Networks. *arXiv.org*.

EMANUEL, A. S., ALAMEDA, G. K., BEHRENS, R. A. & HEWETT, T. A. 1988. Reservoir Performance Prediction Methods Based on Fractal Geostatistics. Annual meeting of the American Association of Petroleum Geologists, March 20 1988 Houston, TX. Society of Petroleum Engineers: United States: Society of Petroleum Engineers.

FOLKESTAD, A., VESELOVSKY, Z. & ROBERTS, P. 2012. Utilising borehole image logs to interpret delta to estuarine system: A case study of the subsurface Lower Jurassic Cook Formation in the Norwegian northern North Sea. *Marine and Petroleum Geology,* 29**,** 255-275.

GOODFELLOW, I. J., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDE-FARLEY, D., OZAIR, S., COURVILLE, A. & BENGIO, Y. Generative adversarial networks. The twenty-eight annual conference on neural information processing systems(NIPS), 2014 Montreal, Canada.

GU, J., WANG, Z., KUEN, J., MA, L., SHAHROUDY, A., SHUAI, B., LIU, T., WANG, X., WANG, G. & CAI, J. 2018. Recent advances in convolutional neural networks. *Pattern recognition,* 77**,** 354-377.

GUPTA, K. D., VALLEGA, V., MANIAR, H., MARZA, P., XIE, H., ITO, K. & ABUBAKAR, A. A deep-learning approach for borehole image interpretation. SPWLA 60th Annual Logging Symposium, 2019. OnePetro.

HALDORSEN, H. H. & DAMSLETH, E. 1993. Challenges in reservoir characterization: GEOHORIZONS. *AAPG bulletin,* 77**,** 541-551.

HARDY, H., BEIER, R. & KORVIN, G. 1996. Fractals in reservoir engineering. *Journal of Hydrology-Amsterdam,* 176**,** 290-292.

HARDY, H. H. 1992. The Generation of Reservoir Property Distributions in Cross Section for Reservoir Simulation Based on Core and Outcrop Photos. Permian Basin Oil and Gas Recovery Conference, March 18-20 1992 Midland, TX. Society of Petroleum Engineers: United States: Society of Petroleum Engineers.

HE, K., ZHANG, X., REN, S. & SUN, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 770-778.

HEWETT, T. A. 1986. Fractal Distributions of Reservoir Heterogeneity and Their Influence on Fluid Transport. Society of Petroleum Engineers.

HEWETT, T. A. & BEHRENS, R. A. 1990. Conditional Simulation of Reservoir Heterogeneity With Fractals. *SPE Formation Evaluation,* 5**,** 217-225.

JA'FARI, A., KADKHODAIE-ILKHCHI, A., SHARGHI, Y. & GHANAVATI, K. 2012. Fracture density estimation from petrophysical log data using the adaptive neuro-fuzzy inference system. *Journal of Geophysics and Engineering,* 9**,** 105-114.

JETCHEV, N., BERGMANN, U. & VOLLGRAF, R. 2016. Texture synthesis with spatial generative adversarial networks. *arXiv preprint arXiv:1611.08207*.

JIA, A., HE, D. & JIA, C. 2012. Advances and challenges of reservoir characterization: A review of the current state-of-the-art. *Earth Sciences,* 1.

JOHNSTON, D. 2004. Reservoir characterization improves stimulation, completion practices. *Oil & gas journal,* 102**,** 60-63.

KATZ, A. J. & THOMPSON, A. H. 1985. Fractal Sandstone Pores: Implications for Conductivity and Pore Formation. *Physical Review Letters,* 54**,** 1325-1328.

KOSARI, E., GHAREH-CHELOO, S., KADKHODAIE-ILKHCHI, A. & BAHROUDI, A. 2015. Fracture characterization by fusion of geophysical and geomechanical data: a case study from

the Asmari reservoir, the Central Zagros fold-thrust belt. *Journal of Geophysics and Engineering,* 12**,** 130-143.

KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G. E. 2017. Imagenet classification with deep convolutional neural networks. *Communications of the ACM,* 60**,** 84-90.

KROHN, C. E. & THOMPSON, A. H. 1986. Fractal sandstone pores: Automated measurements using scanning-electron-microscope images. *Physical Review B,* 33**,** 6366-6374.

LAI, J., WANG, G., WANG, S., CAO, J., LI, M., PANG, X., HAN, C., FAN, X., YANG, L. & HE, Z. 2018. A review on the applications of image logs in structural analysis and sedimentary characterization. *Marine and Petroleum Geology,* 95**,** 139-166.

LALOY, E., HÉRAULT, R., JACQUES, D. & LINDE, N. 2018. Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research,* 54**,** 381-406.

LALOY, E., HÉRAULT, R., LEE, J., JACQUES, D. & LINDE, N. 2017. Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. *Advances in water resources,* 110**,** 387-405.

LE LOC'H, G., BEUCHER, H., GALLI, A. & DOLIGEZ, B. Improvement in the truncated Gaussian method: combining several Gaussian functions. 1994. European Association of Geoscientists & Engineers, cp-233.

LECUN, Y., JACKEL, L. D., BOTTOU, L., CORTES, C., DENKER, J. S., DRUCKER, H., GUYON, I., MULLER, U. A., SACKINGER, E. & SIMARD, P. 1995. Learning algorithms for classification: A comparison on handwritten digit recognition. *Neural networks: the statistical mechanics perspective,* 261**,** 2.

LEFRANC, M., BAYRAKTAR, Z., KRISTENSEN, M., DRISS, H., LE NIR, I., MARZA, P. & KHERROUBI, J. 2021. Deep-Learning-Based Automated Sedimentary Geometry Characterization From Borehole Images. *Petrophysics-The SPWLA Journal of Formation Evaluation and Reservoir Description,* 62**,** 636-650.

LIMA, L., BIZE-FOREST, N., EVSUKOFF, A. & LEONHARDT, R. Unsupervised Deep Learning for Facies Pattern Recognition on Borehole Images. Offshore Technology Conference Brasil, 2019. OnePetro.

LIU, W., WANG, Z., LIU, X., ZENG, N., LIU, Y. & ALSAADI, F. E. 2017. A survey of deep neural network architectures and their applications. *Neurocomputing,* 234**,** 11-26.

LOPEZ-ALVIS, J., LALOY, E., NGUYEN, F. & HERMANS, T. 2021. Deep generative models in inversion: The impact of the generator's nonlinearity and development of a new approach based on a variational autoencoder. *Computers & Geosciences,* 152**,** 104762.

LOZADA-ZUMAETA, M., ARIZABALO, R. D., RONQUILLO-JARILLO, G., COCONI-MORALES, E., RIVERA-RECILLAS, D. & CASTREJON-VACIO, F. 2012. Distribution of petrophysical properties for sandy-clayey reservoirs by fractal interpolation. *Nonlinear Process Geophys.,* 19**,** 239-250.

MARIETHOZ, G., RENARD, P., CORNATON, F. & JAQUET, O. 2009. Truncated plurigaussian simulations to characterize aquifer heterogeneity. *Groundwater,* 47**,** 13-24.

MARINI, M., FELLETTI, F., BERETTA, G. P. & TERRENGHI, J. 2018. Three Geostatistical methods for hydrofacies simulation ranked using a large borehole lithology dataset from the Venice Hinterland (NE Italy). *Water,* 10**,** 844.

MEERSCHMAN, E., PIROT, G., MARIETHOZ, G., STRAUBHAAR, J., VAN MEIRVENNE, M. & RENARD, P. 2013. A practical guide to performing multiple-point statistical simulations with the Direct Sampling algorithm. *Computers & Geosciences,* 52**,** 307-324.

MOSSER, L., DUBRULE, O. & BLUNT, M. J. 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Physical Review E,* 96**,** 043309.

MOSSER, L., DUBRULE, O. & BLUNT, M. J. 2018a. Conditioning of three-dimensional generative adversarial networks for pore and reservoir-scale models. *arXiv preprint arXiv:1802.05622*.

MOSSER, L., DUBRULE, O. & BLUNT, M. J. 2018b. Stochastic reconstruction of an oolitic limestone by generative adversarial networks. *Transport in Porous Media,* 125**,** 81-103.

MOSSER, L., DUBRULE, O. & BLUNT, M. J. 2019. Deepflow: history matching in the space of deep generative models. *arXiv preprint arXiv:1905.05749*.

MOSSER, L., DUBRULE, O. & BLUNT, M. J. 2020. Stochastic seismic waveform inversion using generative adversarial networks as a geological prior. *Mathematical Geosciences,* 52**,** 53-79.

NIE, X., ZOU, C., PAN, L., HUANG, Z. & LIU, D. 2013. Fracture analysis and determination of in-situ stress direction from resistivity and acoustic image logs and core data in the Wenchuan Earthquake Fault Scientific Drilling Borehole-2 (50–1370 m). *Tectonophysics,* 593**,** 161-171.

ORHAN, A. E. & PITKOW, X. 2017. Skip connections eliminate singularities. *arXiv preprint arXiv:1701.09175*.

PANG, J. & NORTH, C. P. 1996. FRACTALS AND THEIR APPLICABILITY IN GEOLOGICAL WIRELINE LOG ANALYSIS. *Journal of Petroleum Geology,* 19**,** 339-350.

PEREZ, G. & CHOPRA, A. K. 1997. Evaluation of Fractal Models To Describe Reservoir Heterogeneity and Performance. *SPE Formation Evaluation,* 12**,** 65-72.

PHILLIPS, C. 1996. Enhanced thermal recovery and reservoir characterization.

RADFORD, A., METZ, L. & CHINTALA, S. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

RAZAK, S. M. & JAFARPOUR, B. 2020a. Convolutional neural networks (CNN) for feature-based model calibration under uncertain geologic scenarios. *Computational Geosciences,* 24**,** 1625-1649.

RAZAK, S. M. & JAFARPOUR, B. 2020. History Matching with Generative Adversarial Networks. 2020 2020b. European Association of Geoscientists & Engineers, 1-17.

RUTHOTTO, L. & HABER, E. 2021. An introduction to deep generative modeling. *GAMM-Mitteilungen***,** e202100008.

SAINATH, T. N., KINGSBURY, B., MOHAMED, A.-R., DAHL, G. E., SAON, G., SOLTAU, H., BERAN, T., ARAVKIN, A. Y. & RAMABHADRAN, B. Improvements to deep convolutional neural networks for LVCSR. 2013 IEEE workshop on automatic speech recognition and understanding, 2013. IEEE, 315-320.

SAMI, M. & MOBIN, I. A comparative study on variational autoencoders and generative adversarial networks. 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT), 2019. IEEE, 1-5.

SHAO, L., ZHU, F. & LI, X. 2014. Transfer learning for visual categorization: A survey. *IEEE transactions on neural networks and learning systems,* 26**,** 1019-1034.

SHEN, P., LIU, M. & JIA, F. 1998. Application of Fractal Techniques in Reservoir Development. SPE International Conference and Exhibition, November 2-6 1998 Beijing, China. Society of Petroleum Engineers : Beijing: Society of Petroleum Engineers.

SIMONYAN, K. & ZISSERMAN, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

STREBELLE, S. 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical geology,* 34**,** 1-21.

SZEGEDY, C., LIU, W., JIA, Y., SERMANET, P., REED, S., ANGUELOV, D., ERHAN, D., VANHOUCKE, V. & RABINOVICH, A. Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015. 1-9.

TAHMASEBI, P. 2018. Multiple Point Statistics: A Review. *In:* DAYA SAGAR, B. S., CHENG, Q. & AGTERBERG, F. (eds.) *Handbook of Mathematical Geosciences: Fifty Years of IAMG.* Cham: Springer International Publishing.

TANG, J. 2018. *Intelligent Mobile Projects with TensorFlow: Build 10+ Artificial Intelligence Apps Using TensorFlow Mobile and Lite for IOS, Android, and Raspberry Pi*, Packt Publishing Ltd.

WANG, Z.-W. & MOU, D. 2015. Fractal Dimension of well logging curves associated with the texture of volcanic rocks. Internationa Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC 2014), 2014 Shenyang, China. Atlantis Press: Hong Kong: Atlantis Press.

YANG, D., HONG, S., JANG, Y., ZHAO, T. & LEE, H. 2019. Diversity-sensitive conditional generative adversarial networks. *arXiv preprint arXiv:1901.09024*.

YEH, R. A., CHEN, C., YIAN LIM, T., SCHWING, A. G., HASEGAWA-JOHNSON, M. & DO, M. N. Semantic image inpainting with deep generative models. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017 2017. 5485-5493.

ZACCONE, G., KARIM, M. R. & MENSHAWY, A. 2017. *Deep learning with TensorFlow*, Packt Publishing Ltd.

ZHANG, C., SONG, X. & AZEVEDO, L. 2021. U-net generative adversarial network for subsurface facies modeling. *Computational Geosciences,* 25**,** 553-573.

ZHUANG, F., QI, Z., DUAN, K., XI, D., ZHU, Y., ZHU, H., XIONG, H. & HE, Q. 2020. A comprehensive survey on transfer learning. *Proceedings of the IEEE,* 109**,** 43-76.