The University Of Sheffield.

Machine Learning Based Data-Driven Methods for Modelling and Simulation of Pressure Dynamics and Fluid Flow in Natural Gas Reservoirs

Aliyuda Ali

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Automatic Control and Systems Engineering
Faculty of Engineering
The University of Sheffield

Supervisors: Dr Lingzhong Guo
Prof Daniel Coca

July 2021

Abstract

Among the fossil fuels, natural gas is the cleanest compared to oil and coal. Identifying pressure trends and patterns associated with natural gas production and reservoir fluid characterisation in subsurface workflows is a major challenge in the oil and gas industry. Traditionally used methods could be employed to identify pressure dynamics and, among these methods, numerical reservoir simulation approach exhibits promising potential as a tool for studying complex multiscale reservoir problems. Managing reservoir subsurface workflows based on numerical simulation, however, poses a great computational challenge and involves numerous simulations that require the discretisation of complex nonlinear partial differential equations represented by parameter spaces of extremely high dimension. Data-driven modelling is changing the landscape of reservoir subsurface workflows and making a key difference as a promising technique for identifying patterns and predicting future trends in the oil and gas industry. Thus, this thesis aims to develop effective machine learning data-driven based methods for modelling and simulation of pressure and fluid flow dynamics in natural gas reservoirs. Firstly, an adaptive neuro-fuzzy model that captures the pressure dynamics and gas production trend under different conditions of varied BHPs for single-phase gas condensate reservoir is developed. This is followed by developing a hybrid model that combines the capabilities of adaptive neuro-fuzzy and subtractive clustering techniques to improve the dewpoint pressure predictions for gas condensate reservoirs. Next, a data-driven model based on dynamic mode decomposition algorithm is developed to decompose and reconstruct a pressure field of a depleted reservoir model that mimics the behaviour of an underground natural gas storage. Lastly, a new dynamic mode learning method is developed to examine the effect of gas injection and water flooding on pressure and fluid flow dynamics of multiphase reservoirs. Experimental results show that the outputs of the developed models are in good agreement with reference data.

Acknowledgements

Firstly, I would like to thank God Almighty for His mercy that keeps me and sufficient grace that sustains me even in the midst of challenges. May His Holy name be praised now and forever, amen.

I am most thankful to my supervisors, Dr Lingzhong Guo and Prof Daniel Coca for their exceptional support and the opportunity given to me to carry out this study under their supervision. Their expertise, enthusiasm, and exacting attention to detail have been a motivation to me and kept my work on track during the period of my PhD programme.

I am also thankful to the Petroleum Technology Development Fund (PTDF) of the Federal Republic of Nigeria for providing me with funding to carry out this study.

I am equally grateful to Dr Ivan Benemerito, Dr Uchenna Diala, and Dr Yuzhu Guo who offered me assistance in different ways especially in my early days of this study.

Special thanks to my caring wife (Yosi Faith), beloved son (Aziel), parents, and members of Sheffield Evangelical Presbyterian Church for continuous encouragement and support during the period of this study.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BHP:	Bottom hole pressure
EOS:	Equations of state
DML:	Dynamic mode learning
UNGS:	Underground natural gas storage
PVT:	Pressure-Volume-Temperature
BOAST-NFR:	Black oil applied simulation tool for naturally fractured reservoirs
DOE:	Department of energy
IMPES:	Implicit pressure - explicit saturation
MRST:	MATLAB reservoir simulation toolbox
CCE:	Constant composition expansion
CVD:	Constant volume depletion
SRK:	Soave-Redlich-Kwong
PR:	Peng-Robinson
VLE:	Vapour-liquid equilibrium
ANN:	Artificial neural network
RBF:	Radial basis function
GA:	Genetic algorithm
GCR:	Gas condensate reservoir
SVM:	Support vector machine
LSSVM:	Least square support vector machine
CA:	Simulated annealing
AI:	Artificial Intelligence
ANFIS:	Adaptive neuro-fuzzy inference system
FIS:	Fuzzy inference system
MF:	Membership function
RMSE:	Root mean squared error
DMD:	Dynamic mode decomposition
SPE:	Society of Petroleum Engineers
SVD:	Singular value decomposition

POD:	Proper orthogonal decomposition
PCA:	Principal component analysis
MSE:	Mean squared error
Tcm:	Trillion cubic meters
Bcm:	Billion cubic meters
PDEs:	Partial differential equations
ODEs:	Ordinary differential equations
ML:	Machine learning

Declaration

I, Aliyuda Ali, hereby confirm that this Thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means. I confirm that this work has not been previously presented for an award at this, or any other, university. I acknowledge that the following articles have been written from this study:

A. Ali and L. Guo. Adaptive neuro-fuzzy approach for prediction of dewpoint pressure for gas condensate reservoirs. Published in Petroleum Science and Technology vol. 38, no. 9, pp. 673–681, 2020, doi: 10.1080/10916466.2020.1769655.

A. Ali and L. Guo. Neuro-adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir. Presented at 13th IFAC Workshop on Adaptive and Learning Control Systems - ALCOS 2019, 04 – 06 December 2019, Winchester, UK.

A. Ali and L. Guo. Data-driven based investigation of pressure dynamics in underground hydrocarbon reservoirs. Presented at the 5th Annual CDT Conference in Energy Storage and Its Applications, 12 – 13 Jan. 2021, Sheffield, UK.

Chapter One: Introduction

1.1 Background and Motivation

As the world population keeps on rising and industrial and residential activities increase, so also the demand for energy [1], [2], [3]. One of the fundamental issues every country is trying to address today, is how to strike a balance between the energy and population size [4], [5]. Thus, demand for energy by the world increases as population size increases and economy grows. Study has shown as in Figure 1.1 that over the past decades, world demand for energy has been steadily increasing [6].

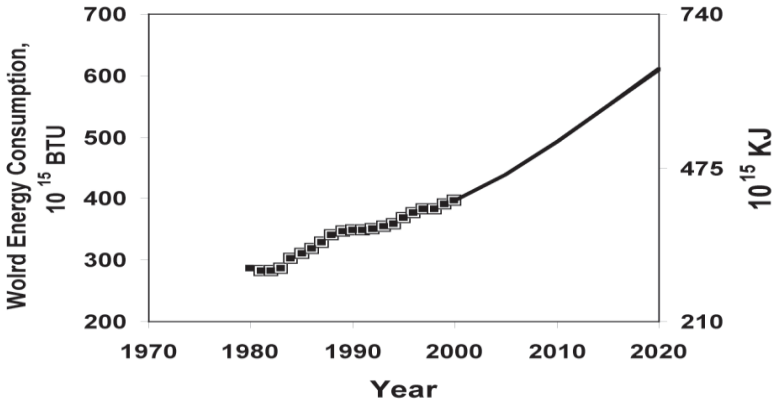

Figure 1.1: World-wide energy demand [6]

To meet the rapidly increasing global energy demand, various forms of energy sources need to apply available and develop new techniques to meet this increasing demand [7]. So far, renewable energy (solar, wind, hydro, tidal, geothermal and biomass) has accounted for up to 35% of the overall energy supply [8], [9], [10]. Nuclear energy is another source, however, the possible risk of contamination associated with nuclear energy is high [11]. Fossil fuels, which include oil, natural gas and coal, supply almost 65% of the total world energy [12], [13]. Until alternative sources of energy that are economically viable are found to support the world's energy-hungry way of life, hydrocarbon fuels in conventional and unconventional hydrocarbon reservoirs will continue to account for the large proportion of the world energy supply to meet the increasing energy demand in the next several decades [14] [15],.

Concurrently, now that climate change is recognized as one of the major challenges facing mankind, attempts headed for achieving a low-carbon energy mix have to be considered in

every parts of the globe [16], [17]. Thus, this study is motivated by the fact that, compared to other available fossil fuels (oil and coal), natural gas is found to be the cleanest of the hydrocarbon fuels and the fuel of choice for energy efficiency having not only low carbon emission, but also providing a solution to world's increasing energy demand for economic growth and environmental challenges in a safe and sustainable manner [6], [18], [19].

1.1.1 Natural Gas Consumption

It has been observed that among the fossil fuels, world demand for natural gas has been steadily increasing over the last forty years [6], as shown in Figure 1.2.

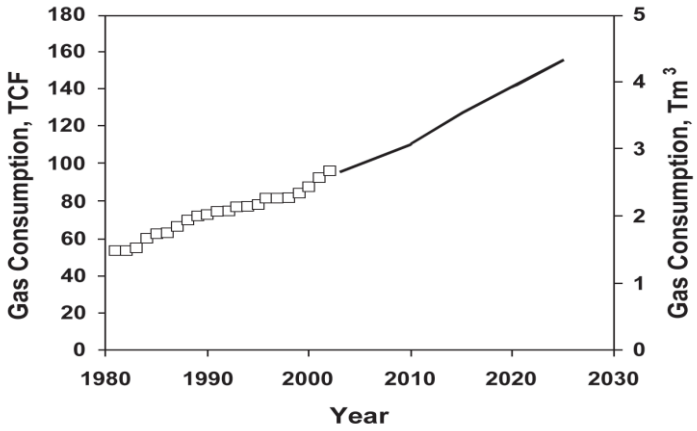

Figure 1.2: Historic demand for natural gas [6]

Considering the future energy demand by the world, power generation is expected to be the most demand. Expectation for power demand over the next 20 years is estimated to grow at the rate of more than 2% per year [6]. According to [20], power generation using natural gas is considered a better option than using any other fossil fuels for the following reasons: lower operating costs, lower CO_2 and NOx emissions, short construction time, higher operational flexibility, lower capital investment, no SOx or particulate emissions and higher thermal efficiency. The 2020 statistical report on natural gas information released by International Energy Agency [21] further disclosed that, production of natural gas globally has for the first time, hit a new high in 2019 by breaking the 4 Trillion cubic meters (Tcm) threshold with 4088 Billion cubic meters (Bcm), produced, +3.3% in comparison to 2018 [21]. The report also added that, since the financial crisis, production of natural gas has been constantly increasing at an annually compounded increase rate of 2.7%.

Industrial and residential activities remain the major factors for a large proportion of natural gas consumption in developed countries, and thus are targets of energy efficiency plans [22], [23], [24]. Industrial consumption of natural gas, both as a feedstock and fuel is set to rise at an average yearly rate of 3% which represents 46% of the increase in global consumption to 2024 [21]. As concerns and campaigns about climate change and air quality keep growing in today's world, renewable energy is growing but at a slow rate and energy resources with low-carbon are difficult to explore in some regions [25], [26]. Burning natural gas has been considered as a "bridging technology" along the long path to establishing energy systems based largely on renewables, this is why natural gas is still regarded to have major increase potential in the predictable feature [27]. Thus, it is worthwhile mentioning that if policies to cap carbon emission have to be made by world governments, then extensive fuel replacement from coal to natural gas is likely [28], [29], and any effort aimed to optimize the development of gas fields, maximize gas production and minimize its severe loss should be considered a positive contribution to a more sustainable world.

1.1.2 Natural Gas Reservoirs

Natural gas reservoirs are naturally occurring formations of rocks that trap and hold natural gas beneath the subsurface [6]. Conventional natural gas reservoirs are characterised by three main parts: the source rock, the reservoir rock, and the cap rock. The source rock is referred to as the less-permeable and compact rock that contains the kerogen that the gas forms from. Examples of source rocks are tight limestones, shales, and fine-grained. Reservoir rock refers to the porous, permeable rock layer or layers that hold the natural gas. Examples of reservoir rocks are sandstones, sand, and dolomites. The natural gas migrates from the source rock into the reservoir rock because of the pressure difference between the reservoir rocks, which are at lower pressures, and the source rocks, which are compressed by the weight of overlying rocks. Recovery from the reservoir is affected mainly by the natural expansion of the gas. The cap rock refers to an impervious rock that seals the top and sides of the reservoir rock so that the natural gas is trapped in the reservoir to prevent it from escaping upward or laterally. For a natural gas reservoir to exist, gas from the source rock must migrate into the reservoir rock, which takes millions of years.

In the oil and natural gas industry, natural gas reservoir modelling refers to the creation of a geological model that can be computationally simulated in order to improve the estimation of reserves and make decisions as regards field development, future production prediction, and evaluating scenarios with respect to reservoir management plan. Thus, a natural gas reservoir model represents the physical space of the reservoir in the form of a collection of discrete cells, depicted by a grid which may be regular or irregular. The collection of these discrete cells is usually three-dimensional, although two-dimensional and one-dimensional are sometimes used. Values of attributes such as rock properties (porosity, permeability, compressibility) and fluid properties (density, viscosity, compressibility) are associated with each cell within the grid. Other components such as reservoir pressure, temperature and wells (production and injection) with their associated properties are also added to the reservoir model. The main aim of creating a natural gas reservoir model is to provide a static description of the real reservoir prior to production.

Natural gas reservoir simulation refers to the process of utilising geological model with its associated physics and mathematics, as well as computer programming to estimate the fluids dynamics and field performance under various operating strategies. Thus, the goal of natural gas reservoir simulation is to gain insight into the recovery processes of the reservoir. The simulation process starts with building a reservoir model that describes rock and fluid properties and choosing certain numerical features of the grid such as number and size of cells. This is followed by setting up the appropriate field wells controls (bottom hole pressure constraints, injection rates, etc). These controls are what drive the model during the simulation process. Next, the output needed to be printed to file is chosen. The output can be cumulative gas production over time, average reservoir pressure as a function of time, gas saturation over time, etc. The simulation process is then carried out under various operating strategies. In the process of simulating a natural gas reservoir, the inputs to the simulator can be adjusted in such a way that a better fit to the actual reservoir performance is achieved.

1.1.3 Problem Statement

Decrease in reservoir pressure can result in underperforming production which can have a tremendous impact on the estimated well deliverability and the entire field performance [30], [31]. As such, early prediction of pressure dynamics in relation to production alterations and taking the necessary measures to control or alleviate such issues are essential for maximizing natural gas reservoir resources [32]. Techniques for learning pressure and fluid flow dynamics in natural gas reservoirs are of prime significance in developing reliable predictive models for managing reserves.

Conventionally, numerical simulation is the standard in the oil and gas industry for predicting dynamic parameters (such as production rate, injection rate, pressure drop) and quantifying uncertainties in reservoirs. Characterizing reservoirs and managing workflows using this approach involve numerous simulations for optimizing production, enhancing oil/gas recovery and history matching [33]. Flow of fluid in the reservoir is governed by complex nonlinear partial differential equations (PDEs), which in practice, are spatially discretized into a high-dimensional set of nonlinear ordinary differential equations (ODEs) [34]. For consistent representation of flow dynamics and subsurface geology, grid blocks in very large numbers are required and cumbersome algorithms are employed for their spatiotemporal solutions. It requires thousands of simulations even with advanced algorithms to achieve optimal solutions when solving with nonlinear constraints [35]. The complexity in physics associated with reservoir pressure dynamics and the multiscale nature of the rock and fluid properties present challenges in achieving better predictive models. In comparison to numerical simulation, in which model set up is laborious and implementation is time consuming, a machine learning based data-driven model that not only scales down computational complexity but also, offers accuracy in a short time without compromising results would be of great benefit.

1.2 Aim and Objectives

The challenges faced by physical models to adequately describe fluid/rock interactions and accurately capture flow dynamics on complex geometries prompted a remarkable interest in seeking for alternative solutions via machine learning models [36] [37]. Thus, the aim of this

study is to develop novel machine learning based data-driven methods for accurate modelling and simulation of pressure dynamics and fluid flow in natural gas reservoirs for achieving optimal production that will guide the implementation of effective energy policies in the natural gas industry. To achieve the aforementioned aim, the study foresees the following objectives:

i. To investigate the effect of bottom hole pressure (BHP) on production performance of a single-phase gas condensate reservoir model.
ii. To improve the prediction of dewpoint pressure for gas condensate reservoirs based on machine learning method.
iii. To examine the effect of gas injection on pressure dynamics of underground natural gas storage in depleted reservoirs.
iv. To develop a novel data-driven method that analyses the effect of injecting gas and water on the pressure dynamics of multiphase reservoir models, and compare its performance to existing mainstream methods.

1.3 Contributions

The contributions of this work to the body of knowledge can be summarised as follows:

Firstly, a machine learning model that regenerates the numerical simulation results of a single-phase gas condensate reservoir model for both production rates and pressure drop at different BHPs is developed. The novelty of this method is its ability to incorporate physics-based petro-physical data in the model development process which distinguishes it from existing methods.

Next, a hybrid machine learning model for accurate prediction of dewpoint pressure for gas condensate reservoirs is developed. The novelty of this method is its capability to adapt to changing situations and generate the desired output even if it encounters a situation that is different from the history enclosed in the historical datasets. This is followed by developing a data-driven method that decomposes and reconstructs a pressure field of a depleted reservoir model that mimics the behaviour of an underground natural gas storage. The novelty of this method is its ability to address problems of complexity and demand for high performance computing observed in numerical simulation.

Finally, a novel method called Dynamic Mode Learning (DML) that aims to provide an efficient alternative approach for learning and decomposing flow dynamics in multiphase reservoir models that produce under secondary recovery is proposed. The novelty of this method is its ability to address the limitations of expensive computational demand and variations in final results that are observed in existing methods.

1.4 Scope of the Work and Organization of Chapters

This work focuses on development of machine learning data-driven based methods and their applications to single-phase and multiphase gas condensate reservoir models, underground natural gas storage in depleted reservoirs, and multiphase reservoir models producing under secondary recovery. The remaining part of this thesis is organized as follow:

Chapter two presents a survey of literature on hydrocarbon reservoirs, techniques for simulation of hydrocarbon reservoirs and prediction of gas condensate dewpoint pressure, and an overview of UNGS facilities.

Chapter three titled "Neuro-adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir", presents a neuro-adaptive learning method for predicting the production performance and pressure dynamics of a single-phase gas condensation reservoir model. The machine learning model developed in this chapter addresses the challenges of inflexibility, long development time and high cost of development that are observed in conventional numerical reservoir simulators.

Chapter four titled "Adaptive neuro-fuzzy approach for prediction of dewpoint pressure for gas condensate reservoirs", presents a hybrid adaptive neuro-fuzzy method for prediction of dewpoint pressure for gas condensate reservoirs. The developed adaptive machine learning model addresses the challenges of high cost and long-time of dewpoint pressure prediction that are observed in experimental approach. Also, the challenges of improper characterization of components and cumbersomeness that are observed in conventional equations of state (EOS) approach are addressed.

Chapter five titled "Dynamic mode decomposition for modelling of pressure dynamics in underground natural gas storage reservoirs", presents a data-driven method for investigation

of pressure dynamics in underground natural gas storage reservoirs. The proposed data-driven model is able to learn the pressure field dynamics and approximate average pressure of the depleted reservoir model over time in relation to natural gas injection/storage with a very high accuracy.

Chapter six titled “Dynamic mode learning method for analysis of pressure and fluid phase dynamics in multiphase reservoir models”, presents a novel dynamic mode learning (DML) method for analysis of pressure and fluid phase dynamics in multiphase reservoir models. The performance of the proposed DML method is illustrated on pressure field and fluid phase distribution data sets generated from direct numerical simulations of two multiphase reservoir models.

Based on the presentations and discussion of results, conclusions and limitations of this study are drawn, with direction for future work in chapter seven.

1.5 Publications

The following papers were written in the course of this work.

1. A. Ali and L. Guo. Adaptive neuro-fuzzy approach for prediction of dewpoint pressure for gas condensate reservoirs. *Pet. Sci. Technol.*, vol. 38, no. 9, pp. 673–681, 2020, doi: 10.1080/10916466.2020.1769655.
2. A. Ali and L. Guo. Neuro-adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir. Presented at 13th IFAC Workshop on Adaptive and Learning Control Systems - ALCOS 2019, 04 – 06 December 2019, Winchester, UK. *IFAC-PapersOnLine*, vol. 52, no. 29, pp. 122–127, 2019, doi: 10.1016/j.ifacol.2019.12.632.
3. A. Ali and L. Guo. Data-driven based investigation of pressure dynamics in underground hydrocarbon reservoirs. Presented at 5th Energy Storage and its Applications Conference, 12 – 13 January 2021, Sheffield, UK. *Energy Reports* 7(5):104-110, May 2021. DOI: 10.1016/j.egyr.2021.02.036

References

[1] R. Cisneros-Dévora et al.. Molecular modeling and experimental testing on the comparative performance of macrocyclic and linear zwitterionic surfactants for the emulsified heavy oil recovery. Fuel, vol. 285, no. 3, 2021, doi: 10.1016/j.fuel.2020.119123.

[2] P. V. Elumalai, M. Nambiraj, M. Parthasarathy, D. Balasubramanian, V. Hariharan, and J. Jayakar. Experimental investigation to reduce environmental pollutants using biofuel nano-water emulsion in thermal barrier coated engine. *Fuel*, vol. 285, no. 9, p. 119200, 2021, doi: 10.1016/j.fuel.2020.119200.

[3] J. M. Colom-Díaz, Á. Millera, R. Bilbao, and M. U. Alzueta. New results of H_2S oxidation at high pressures. Experiments and kinetic modeling. *Fuel*, vol. 285, no. 10, p. 119261, 2021, doi: 10.1016/j.fuel.2020.119261.

[4] M. Škare and S. Blažević. Population and economic growth: A review essay. *Amfiteatru Econ.*, vol. 17, no. 40, pp. 1036–1053, 2015.

[5] H. A. Sadeghi. Investigating the role of balance of population and economic growth in achieving sustainable economic growth (case study: upper-middle income countries during 1985 to 2016). *Int. J. Appl. Res. Manag. Econ.*, no. 9, pp. 1–8, 2018, doi: 10.33422/ijarme.2018.09.10.

[6] M. Kelkar. *Natural gas production engineering*. 1st Ed., pp 1-92, 2008. Tulsa: PennWell.

[7] A. R. López *et al.*. Solar PV generation in Colombia - A qualitative and quantitative approach to analyze the potential of solar energy market. *Renew. Energy*, vol. 148, pp. 1266–1279, 2020, doi: 10.1016/j.renene.2019.10.066.

[8] Z. Zhong, A. Y. Sun, Y. Wang, and B. Ren. Predicting field production rates for waterflooding using a machine learning-based proxy model. *J. Pet. Sci. Eng.*, vol. 194, p. 107574, 2020, doi: 10.1016/j.petrol.2020.107574.

[9] M. Lumbreras and R. Garay. Energy & economic assessment of façade-integrated solar thermal systems combined with ultra-low temperature district-heating. *Renew. Energy*, vol. 159, pp. 1000–1014, 2020, doi: 10.1016/j.renene.2020.06.019.

[10] M. Child, C. Kemfert, D. Bogdanov, and C. Breyer. Flexible electricity generation,

grid exchange and storage for the transition to a 100% renewable energy system in Europe. *Renew. Energy*, vol. 139, pp. 80–101, 2019, doi: 10.1016/j.renene.2019.02.077.

[11] S. Suman. Hybrid nuclear-renewable energy systems: A review. *J. Clean. Prod.*, vol. 181, pp. 166–177, 2018, doi: 10.1016/j.jclepro.2018.01.262.

[12] S. S. Neshat, R. Okuno, and G. A. Pope. Simulation of solvent treatments for fluid blockage removal in tight formations using coupled three-phase flash and capillary pressure models. *J. Pet. Sci. Eng.*, vol. 195, no. 5, 2019, p. 107442, 2020, doi: 10.1016/j.petrol.2020.107442.

[13] N. Abas, A. Kalair, and N. Khan. Review of fossil fuels and future energy technologies. *Futures*, vol. 69, no. 3, pp. 31–49, 2015, doi: 10.1016/j.futures.2015.03.003.

[14] K. Ziemelis. Hydrocarbon reservoirs. *Nature*, vol. 426, no. 6964, p. 317, 2003, doi: 10.1038/426317a.

[15] A. M. Hassan, M. Ayoub, M. Eissa, H. Bruining, and P. Zitha. Study of surface complexation modeling on a novel hybrid enhanced oil recovery (EOR) method; smart-water assisted foam-flooding. *J. Pet. Sci. Eng.*, vol. 195, no. 3, p. 107563, 2020, doi: 10.1016/j.petrol.2020.107563.

[16] K. J. Feeley, C. Bravo-Avila, B. Fadrique, T. M. Perez, and D. Zuleta. Climate-driven changes in the composition of New World plant communities. *Nat. Clim. Chang.*, vol. 10, no. 10, 2020, doi: 10.1038/s41558-020-0873-2.

[17] S. V. Hanssen, V. Daioglou, Z. J. N. Steinmann, J. C. Doelman, D. P. Van Vuuren, and M. A. J. Huijbregts. The climate change mitigation potential of bioenergy with carbon capture and storage. *Nat. Clim. Chang.*, 2020, doi: 10.1038/s41558-020-0885-y.

[18] F.-Y. Liang, M. Ryvak, S. Sayeed, and N. Zhao. The role of natural gas as a primary fuel in the near future, including comparisons of acquisition, transmission and waste handling costs of as with competitive alternatives. *Chem. Cent. J.*, vol. 6, p. S4, 2012, doi: 10.1186/1752-153x-6-s1-s4.

[19] W. Guo, F. Feng, G. Song, J. Xiao, and L. Shen. Simulation and energy performance

assessment of CO2 removal from crude synthetic natural gas via physical absorption process. *J. Nat. Gas Chem.*, vol. 21, no. 6, pp. 633–638, 2012, doi: 10.1016/S1003-9953(11)60412-X.

[20] P. Cedigaz (2017). Natural gas in the world. *Summ. Int. Asso. of Nat Gas*, pp, 2 - 41.

[21] International Energy Agency (IEA). Natural gas information - 2019 Final Edition. *Database Documentation*, pp. 1 - 50, 2020.

[22] J. I. Chowdhury, Y. Hu, I. Haltas, N. Balta-Ozkan, G. Matthew, and L. Varga. Reducing industrial energy demand in the UK: A review of energy efficiency technologies and energy saving potential in selected sectors. *Renew. Sustain. Energy Rev.*, vol. 94, no. 2, pp. 1153–1178, 2018, doi: 10.1016/j.rser.2018.06.040.

[23] W. Zappa and M. van den Broek. Analysing the potential of integrating wind and solar power in Europe using spatial optimisation under various scenarios. *Renew. Sustain. Energy Rev.*, vol. 94, no. 8, pp. 1192–1216, 2018, doi: 10.1016/j.rser.2018.05.071.

[24] B. J. van Ruijven, E. De Cian, and I. Sue Wing. Amplification of future energy demand growth due to climate change. *Nat. Commun.*, vol. 10, no. 1, pp. 1–12, 2019, doi: 10.1038/s41467-019-10399-3.

[25] S. Chu and A. Majumdar. Opportunities and challenges for a sustainable energy future. *Nature*, vol. 488, no. 7411, pp. 294–303, 2012, doi: 10.1038/nature11475.

[26] S. Shah, V. Venkatramanan, and R. Prasad. Sustainable green technologies for environmental management. *Sustain. Green Technol. Environ. Manag.*, no. 2, pp. 1–303, 2019, doi: 10.1007/978-981-13-2772-8.

[27] F. Crotogino. Traditional Bulk Energy Storage—Coal and underground natural gas and oil storage. In T. M. Letcher. *Storing Energy: with Special Reference to Renewable Energy Sources*. 1st Ed., pp 391 - 409, 2016. Netherlands: Elsevier.

[28] K. Hayhoe, H. S. Kheshgi, A. K. Jain, and D. J. Wuebbles. Substitution of natural gas for coal: climatic effects of utility sector emissions. *Climatic Change*, vol. 54, no. 1-2, pp. 107 - 139, 2002, doi: 10.1023/A:1015737505552.

[29] A. Safari, N. Das, O. Langhelle, J. Roy, and M. Assadi. Natural gas: A transition fuel for sustainable energy system transformation. *Energy Sci. Eng.*, vol. 7, no. 4, pp.

1075–1094, 2019, doi: 10.1002/ese3.380.

[30] M. Arabloo, A. Shokrollahi, F. Gharagheizi, and A. H. Mohammadi. Toward a predictive model for estimating dew point pressure in gas condensate systems. *Fuel Process. Technol.*, vol. 116, pp. 317–324, 2013, doi: 10.1016/j.fuproc.2013.07.005.

[31] E. M. E. M. Shokir. Dewpoint pressure model for gas condensate reservoirs based on genetic programming. *Soc. Pet. Eng. - SPE Gas Technol. Symp.*, vol. 1, no. 4, pp. 178–190, 2008.

[32] A. Skauge, G. S. Haskjold, P. A. Ormehaug, and M. G. Aara. Studies of production under bubblepoint. no. 8, 2014, doi: 10.3997/2214-4609.201406340.

[33] Y. Zhang, J. He, C. Yang, J. Xie, R. Fitzmorris, and X. H. Wen. A physics-based data-driven model for history matching, prediction, and characterization of unconventional reservoirs. *SPE J.*, vol. 23, no. 4, pp. 1105–1125, 2018, doi: 10.2118/191126-pa.

[34] K. Aziz, and A. Setari. Applied Sci. *Petroleum reservoir simulation*. 1st Ed., pp. 357 - 372, 1979. London: Applied Sci. Publishers.

[35] Z. Liu, F. Forouzanfar, and Y. Zhao. Comparison of SQP and AL algorithms for deterministic constrained production optimization of hydrocarbon reservoirs. *J. Pet. Sci. Eng.*, vol. 171, no. 10, pp. 542–557, 2018, doi: 10.1016/j.petrol.2018.06.063.

[36] A. Ali and L. Guo. Neuro-adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir. *IFAC-PapersOnLine*, vol. 52, no. 29, pp. 122–127, 2019, doi: 10.1016/j.ifacol.2019.12.632.

[37] F. Alenezi and S. Mohaghegh. A data-driven smart proxy model for a comprehensive reservoir simulation. *4th Saudi Int. Conf. Inf. Technol. (Big Data Anal. KACSTIT*, 2016, doi: 10.1109/KACSTIT.2016.7756063.

Chapter Two: Literature Review

2.1 Introduction

This chapter takes a critical survey of literature on hydrocarbon reservoirs, techniques for simulation of hydrocarbon reservoirs, and prediction of gas condensate dewpoint pressure. The chapter begins with looking at the classification of hydrocarbon reservoirs and simulation studies of hydrocarbon reservoirs in general, then the survey is narrowed to performance and flow behaviour of gas condensate reservoirs as well as techniques used for predicting the dewpoint pressure of gas condensate systems. The chapter ends with a discussion of the drawbacks and gaps identified in the literature and how the present study attempts to address such drawbacks and fill in the gap.

2.2 Hydrocarbon Reservoirs

Hydrocarbon reservoirs are rocks that have sufficient void space to store oil and/or gas and sometimes are referred to as multicomponent systems. The composition of fluids in hydrocarbon reservoirs is strongly related to the source, history, and present reservoir conditions [1]. Hydrocarbon reservoirs most at times are classified based on their fluid compositions, petro-physical properties, and the phase behaviour exhibited by the mixture [1]. Table 2.1 gives a typical classification of hydrocarbon reservoirs based on their fluid compositions as dry gas, wet gas, gas condensate, near-critical oil, volatile oil, and black oil. It is therefore evident from Table 2.1 that fluid compositions and other physical properties differentiate one reservoir from the other. Fluid composition, temperature, and pressure are essential parameters whose conditions determine the phase behaviour of these systems.

2.2.1 Hydrocarbon Reservoir Modelling and Simulation

Reservoir modelling refers to the creation of a geological model that can be computationally simulated in order to improve the estimation of reserves and make decisions as regards field development, future production prediction, and evaluating scenarios with respect to reservoir management plan. When designing a reservoir model, it is essential that the conceptual reservoir architecture be captured by the model and that the key structural and depositional elements have been correctly selected. The authors in [2] defined simulation of hydrocarbon reservoir performance as “the construction and operation of a model whose behaviour

assumes the appearance of actual reservoir behaviour." Such a model itself can be mathematical or physical (such as laboratory sand pack). A mathematical reservoir model is a set of equations (subject to certain assumptions) that describe relationship among various parameters/variables which in turn, describes the physical processes/operations active in the reservoir. Even though the model itself may at some point lack the reality of the reservoir, its valid simulation behaviour assumes the working principle of the actual reservoir. The purpose of reservoir simulation is to estimate field performance (for example, gas recovery) under one or various producing plans.

Table 2.1: Composition of different reservoir fluids [3]

	Composition (mol %)					
Component	Dry Gas	Wet Gas	Gas Condensate	Near-Critical Oil	Volatile Oil	Black Oil
CO_2	0.10	1.41	2.37	1.30	0.93	0.02
N_2	2.07	0.25	0.31	0.56	0.21	0.34
C_1	86.12	92.46	73.19	69.44	58.77	34.62
C_2	5.91	3.18	7.80	7.88	7.57	4.11
C_3	3.58	1.01	3.55	4.26	4.09	1.01
i-C_4	1.72	0.28	0.71	0.89	0.91	0.76
n-C_4		0.24	1.45	2.14	2.09	0.49
i-C_5	0.50	0.13	0.64	0.90	0.77	0.43
n-C_5		0.08	0.68	1.13	1.15	0.21
$C_{6(S)}$		0.14	1.09	1.46	1.75	1.61
C_{7+}		0.82	8.21	10.04	21.76	56.40

While a field can be produced only once, at reasonable expense, a valid reservoir model can be produced or run many times at considerable expense over a short period of time. Results observed/collected from the model that represent different producing conditions can help in selection of an optimal set of producing conditions for the reservoir.

2.2.2 Black Oil Reservoir Modelling and Simulation

Black-oil simulators are models that are used to determine the production strategies and management of reservoirs. They essentially account for three phases: oil, water, and gas where the respective phases are assumed to have no interaction among themselves. Hydrocarbon reservoirs can be modelled using black oil simulators. When a black oil model is used to model a reservoir, the reservoir fluid Pressure-Volume-Temperature (PVT) properties are generated as functions of saturation and pressure [4]. The model consists of fluid components; gas, oil, and water at standard conditions, which are distributed among three distinct fluid phases such as gas, oil, and water, respectively. In black oil model, gas may exist as free gas or solution gas while water and oil are considered to be immiscible. The basic assumptions behind the black oil simulation are: (i) that all of the resulting fluid PVT behaviour is only a function of pressure and temperature, and (ii) reservoir fluids are in thermodynamic equilibrium throughout the reservoir and maintain constant reservoir temperature [5] [6]. Governing differential flow equations for modelling three-phase black oil are developed by combining EOS, Darcy's law, and conservation of mass. Given by [7], these equations can be expressed as follows:

$$A_x \frac{\delta}{\delta x}\left(\frac{k_{rg}}{\mu_g B_g}\frac{\delta \phi_g}{\delta x} + \frac{R_{so}k_{ro}}{\mu_o B_o}\right) + A_y \frac{\delta}{\delta y}\left(\frac{k_{rg}\delta \phi_g}{\delta y} + \frac{R_{so}k_{ro}}{\mu_o B_o}\frac{\delta \phi_o}{\delta y}\right) + A_z \frac{\delta}{\delta z}\left(\frac{k_{rg}}{\mu_g B_g}\frac{\delta \phi_g}{\delta z} + \frac{R_{so}k_{ro}}{\mu_o B_o}\frac{\delta \phi_o}{\delta z}\right) = V_b \frac{\delta}{\delta t}\left\{\phi\left(\frac{S_g}{B_g} + \frac{R_{so}S_o}{B_o}\right)\right\} + q_g \quad (2.1)$$

$$A_x \frac{\delta}{\delta x}\left(\frac{k_{ro}}{\mu_o B_o}\frac{\delta \phi_o}{\delta x}\right) + A_y \frac{\delta}{\delta y}\left(\frac{k_{ro}}{\mu_o B_o}\frac{\delta \phi_o}{\delta y}\right) + A_z \left(\frac{k_{ro}}{\mu_o B_o}\frac{\delta \phi_o}{\delta z}\right) = V_b \frac{\delta}{\delta t}\left(\varphi \frac{S_o}{B_o}\right) + q_o \quad (2.2)$$

$$A_x \frac{\delta}{\delta x}\left(\frac{k_{rw}}{\mu_w B_w}\frac{\delta \phi_w}{\delta x}\right) + A_y \frac{\delta}{\delta y}\left(\frac{k_{rw}}{\mu_w B_w}\frac{\delta \phi_w}{\delta y}\right) + A_z \frac{\delta}{\delta z}\left(\frac{k_{rw}}{\mu_w B_w}\frac{\delta \phi_w}{\delta z}\right) = V_b \frac{\delta}{\delta t} \geq \left(\phi \frac{S_w}{B_w}\right) + q_w \quad (2..3)$$

where *x*, *y* and *z* stand for the coordinates of the model; A_x, A_y, and A_z represent the cross sectional areas normal to *x*, *y*, and *z* directions respectively; subscripts *g*, *o*, and *w* represent gas phase, oil phase, and water phase, respectively; k_{rg}, k_{ro}, and k_{rw} denote relative permeabilities of the phases, B_g, B_o, and B_w denote formation volume factors; μ_g, μ_o, and μ_w

represent respective phase viscosities; S_g, S_o, and S_w denote phase saturations; q_g, q_o, and q_w represent production rates of gas, oil, and water, respectively; φ stands for the porosity; V_b stands for the bulk volume, and R_{so} denotes solution gas oil ratio. The fluid potentials are $\phi_g, \phi_o,$ and ϕ_w which are expressed as follows:

$$\phi_g = P_g + \rho_g g \tag{2.4}$$

$$\phi_o = P_o + \rho_o g \tag{2.5}$$

$$\phi_w = P_w + \rho_w g \tag{2.6}$$

where P_g, P_o, and P_w denote phase pressures; $\rho_g, \rho_o,$ and ρ_w denote phase densities of gas, oil, and water, respectively, g denotes acceleration due to gravity, and h denotes hydraulic pressure head.

2.2.3 Compositional Reservoir Modelling and Simulation

While black oil simulators represent reservoir fluids at stock-tank volume conditions and do not put fluid composition into consideration, compositional simulators on the other hand represent their equations at reservoir conditions [8]. Compositional models are developed in terms of moles of the individual components that made up the mixture [9]. In reality, there are a large number of components in a reservoir fluid sample, and developers of compositional simulators most at times represent the first few (say, three to twelve) components precisely. However, the remaining components, such as C_{7+} are grouped together into what is referred to as pseudo component. In this manner, the average properties of the remaining components are represented by the properties of such a pseudo component. Cubic equations of state (EOS) are mostly used to represent the fluid properties of the pseudo component. With vaporization test data available, compositional model developers normally modify the EOS properties of one or more of the components to closely represent the test data [10].

2.3 Gas Condensate Reservoirs

Gas-condensate reservoirs are considered an important source of natural gas which for long have been discovered as reservoir type that exhibits the most complex thermodynamic and intricate flow behaviours [16] - [19]. These reservoirs are characterized by producing both

natural gas and condensate liquid at surface. Fluid properties and phase behaviour play a critical role in gas condensate reservoirs in comparison to other reservoir type. A gas condensate reservoir is a system composed of mostly C_1, intermediate components (C_2, C_3, C_4, C_5, C_6), and a slight portion of heavier components (C_{7+}). It also contains nonhydrocarbon components such as CO_2, N_2 and sometimes H_2S. These components are naturally situated underground at elevated conditions of pressure and temperature as shown in Figure 2.1.

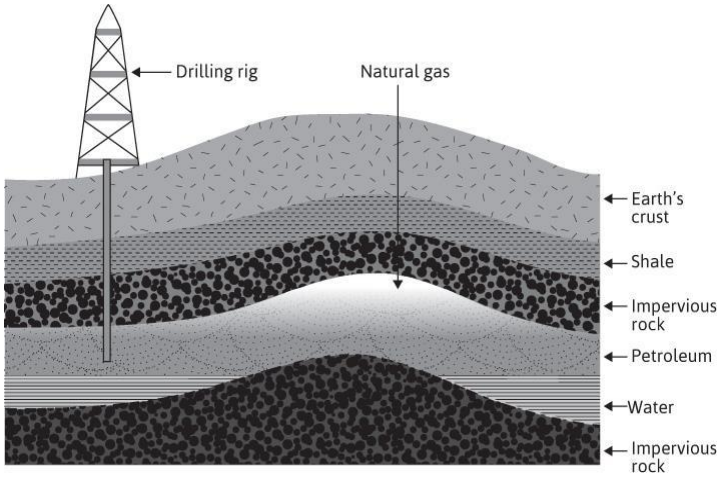

Figure 2.1: Occurrence of hydrocarbon fluid under the earth's surface [20]

To produce from such a reservoir, components must embark on a complicated journey from beneath the earth, passing through a great deal of intermediate stages, to be finally dumped into the atmosphere. One of the major challenges faced while interpreting reservoir performance for gas condensate reservoirs is the difficulty to understand the flow and phase behaviours due to the following reasons: (i) continuous composition variation with pressure as the reservoir produces, (ii) diverse and nonlinear nature of reservoir parameters associated with the subsurface system, and (iii) complexity of the reservoir fluid composition [21] - [23]. At initial reservoir conditions, a gas condensate reservoir is a single-phase (gaseous state) fluid. As the reservoir produces, formation temperature usually does not change, but pressure decreases. The reservoir experiences decline in pressure isothermally from the reservoir boundary to the well as production continues [24]. When the pressure in a gas condensate reservoir decreases to a certain point, called the dewpoint, a condensate (liquid) phase rich in heavy ends drops out of the solution [25]. At the first instance, the condensate that dropped

out in the reservoir does not flow until the assembled condensate saturation rises above the critical condensate saturation and in this case, both gas and condensate flow. Because the condensate accommodates most of the heavier components, this results to a loss of valuable hydrocarbons [26].

2.3.1 Single-Phase Gas Condensate Reservoir Performance

The phase behaviour of a Gas Condensate reservoir is a plot of pressure versus temperature that determines whether the reservoir fluid at a given pressure and temperature consists of a single-phase (gas), or two-phase (gas and liquid). The flow and phase behaviours of a gas condensate reservoir fluid can best be described with the help of a phase diagram which is unique for each reservoir fluid having different compositions. Figure 2.2 shows a Pressure-Temperature phase diagram of a gas condensate system. For this type of reservoirs, the reservoir temperature is between the critical temperature and the cricondentherm, and within this region is a line of isothermal reduction of reservoir pressure that describes pressure drops as the reservoir produces. The cricondentherm is the maximum temperature at which two phases can coexist in equilibrium. Therefore, when the reservoir pressure is above the dewpoint pressure, the reservoir is said to be in a single phase (gaseous stage).

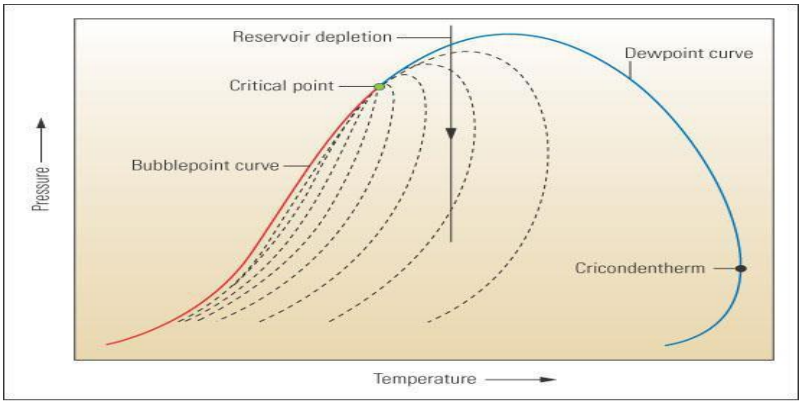

Figure 2.2: A typical phase diagram of a gas condensate reservoir [25]

2.3.2 Two-Phase Gas Condensate Reservoir Performance

As gas condensate reservoir produces, formation temperature usually does not change, but pressure decreases [27]. Significant change in reservoir pressure as a result of production may result in the volatilization of lighter components as well as condensation of heavier components in the reservoir [28] At this stage, the reservoir experiences pressure decline isothermally from the reservoir boundary to the producing well. Dropping of the Bottom-

Hole Flowing Pressure (BHFP) below the dew-point pressure, causes condensate to drop out of the gas and form a bank of liquid around the well. When the pressure further decreases isothermally, instead of having gas coming out of the reservoir, liquid condenses from the gas. This is what makes the gas condensate special and sometimes referred to as “retrograde gas”. At the first instance, the condensate that drops does not flow until the accumulated condensate saturation exceeds the critical condensate saturation and, in this case, both gas and condensate flow.

2.3.3 Formation Damage, Retrograde Condensation and Condensate Blockage

In petroleum engineering, the terms skin effect damage and formation damage have been used to describe various impairments associated to well productivity [29] - [31]. Formation damage can simply be understood as a kind of phenomenon that obstructs the normal flow of fluids towards the surface and such a production impairment can take place anywhere in the production system, from the near-wellbore region of the rock matrix to perforations and into the formation [32], [33]. One of these production impairments that has been notably reported in gas condensate systems is the retrograde condensation. Retrograde condensation is a special case of relative permeability effect. Retrograde condensation can simply be understood as a phenomenon where a condensate (liquid phase) forms from a rich gas as a result of pressure drop resulting in creation of another phase and causing condensate blockage that results in reducing permeability to gas substantially as shown in Fig. 2.3 [34].

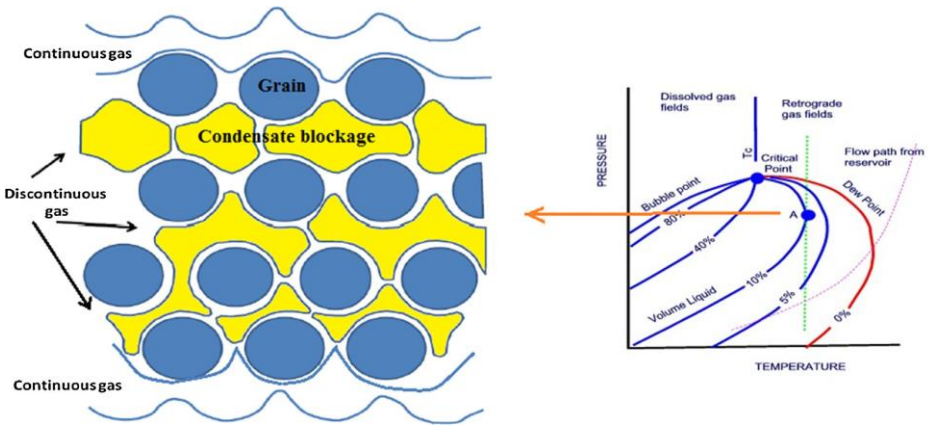

Figure 2.3: Condensate blockage formation in pore spaces of reservoir [34]

The reduction of gas permeability as a result of condensate bank is referred to as condensate blockage (or condensate banking). Condensate blockage has a negative effect in production history of a gas condensate system. When retrograde condensation occurs, serious loss in well productivity can be recorded. In certain cases, the loss in productivity can be as high as 80% [35]. This phenomenon usually occurs near the wellbore when a dewpoint pressure is reached.

Producing a well below the dewpoint causes a reduction to the relative permeability as liquids condense around the wellbore and this can be observed in significant decline in well productivity [32], [36], and [37]. Retrograde condensation is generally classified by the mechanism of its formation as either natural or induced. Natural retrograde condensations are those that take place mainly when the reservoir is being produced based on primary production method. Induced retrograde condensations on the other hand, are those that take place as a result of an external operation that was accomplished on the well such as injection operation or stimulation treatment, well completion, drilling, repair, and so on.

2.3.4 Dewpoint Pressure Prediction for Gas Condensate Reservoirs

Prediction of thermodynamic properties such as the dewpoint pressure of gas condensate reservoirs plays a vital role in interpreting reservoir performance and understanding the fluid phase behaviour. The dewpoint pressure is considered as one of the most important quantities capable of characterizing and predicting the future performance of gas condensate reservoirs. The dewpoint pressure can be defined as the pressure at which a substantial amount of gas phase exists in equilibrium with an infinitesimal amount of liquid phase [38]. As such, accurate determination and prediction of gas condensate dewpoint pressure is essential for fluid characterization, gas reservoir performance calculations, and the design of production systems [39]. When discussing the dewpoint pressure of hydrocarbon mixtures, it is important to know that there are two types of dewpoint pressures which must be distinguished [40]. The first type which is often known as normal dewpoint pressure, occurs when dry gas is compressed to the point where liquid first forms. Normal dewpoint usually occurs at low pressures (below atmospheric pressure) as such, would be of no interest in hydrocarbon reservoir performance. The second type which sometimes is referred to as retrograde

dewpoint pressure, occurs when the pressure in a gas mixture containing heavy hydrocarbons in solution is decreased until a liquid is formed. Retrograde dewpoint pressure is of great significance in interpreting hydrocarbon reservoir performance and behaviour, as such, is the property referred to throughout this study. Following is a brief description of some well-known techniques used to determine the dewpoint pressure of gas condensate reservoirs.

2.3.4.1 Constant Composition Expansion

Dewpoint pressure of a gas condensate fluid can traditionally be determined experimentally in a laboratory through a process called Constant Composition Expansion (CCE), using a visual window-type PVT cell. Figure 2.4 shows a CCE experimental procedure for determination of dewpoint pressure for gas condensate reservoirs.

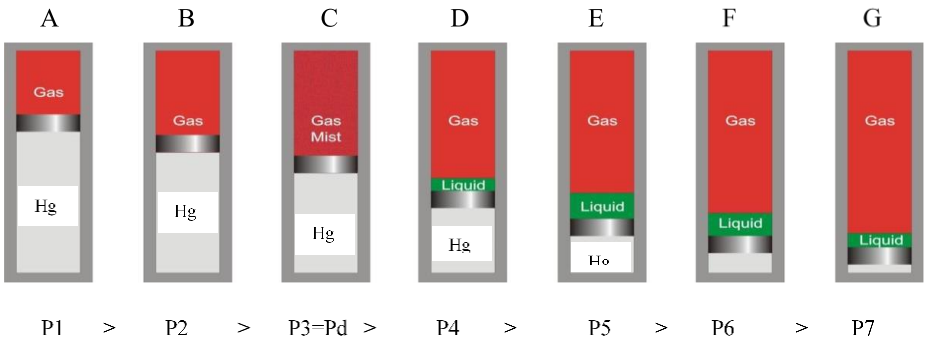

Figure 2.4: Constant composition expansion procedure

The CCE experimental procedure as shown schematically in Figure 2.4 involves placing a hydrocarbon fluid sample at reservoir temperature and pressure in a visual PVT cell. At the first instance, the sample is placed at a pressure in excess of the initial reservoir pressure (Figure 2.4, A). Keeping the temperature constant, the pressure is then reduced in steps by removing mercury from the cell, and for each pressure change, the change in the total hydrocarbon volume, V_t is measured. The corresponding volume and the saturation pressure (dew-point pressure) are observed and recorded and used as a reference volume V_{sat}.

The volume of the hydrocarbon system which is reported as the ratio of the reference volume to saturation volume is termed the relative volume. The relative volume is a function of the cell pressure and is expressed mathematically as:

$$V_{rel} = \frac{V_t}{V_{sat}} \tag{2.7}$$

where, V_{rel} stands for relative volume, V_t denotes total hydrocarbon volume, and V_{sat} denotes volume at saturation pressure. This test is sometimes called pressure-volume relations. Throughout the experiment, the composition of the entire hydrocarbon sample in the cell remains fixed and no hydrocarbon material is removed from the cell, hence, the name constant composition expansion.

2.3.4.2 Constant Volume Depletion Test

The constant volume depletion (CVD) test is another experimental method for determining the dewpoint pressure for gas condensate fluids [41]. The method begins at the dewpoint pressure of the gas condensate and measures the saturation volume, V_{sat}, at the dew point. Decrease in pressure and increase in volume leads to formation of two-phase (gas-condensate). To keep the total volume of the phases equal at a constant pressure, gas is depleted from a valve at the top of the cylinder as shown in Figure 2.5.

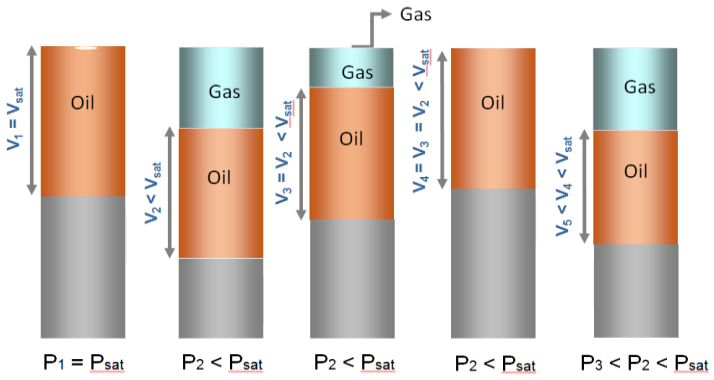

Figure 2.5: Schematic diagram of CVD test

The percentage of the liquid dropout as to the saturation volume as well as the percentage of the depleted gas as to the original gas are measured. The volatile condensate as they may change with each stage in the productions or the PVT properties of the gas condensate are then obtained.

2.3.4.3 Equations of State (EOS) Approach

Equations relating pressure, volume, and temperature (PVT) of a compound, fluid or mixture are referred to as Equations of State (EOS) [3]. These equations describe the state, volumetric and phase behaviour of pure compounds and mixtures under certain thermodynamic conditions. Using the EOS to predict thermodynamic quantities requires critical properties and acentric factor of each component present in the mixture. In this study, the most widely

used Soave-Redlich-Kwong (SRK) and Peng-Robinson (PR) EOS for dewpoint pressure prediction are reviewed.

Soave-Redlich-Kwong (SRK) EOS

The RK EOS is expressed as

$$P = \frac{RT}{V-b} - \frac{a}{V(V+b)} \tag{2.8}$$

where *P* is the pressure, *R* the universal gas constant, *V* is the molar volume and *T* the absolute temperature. In an attempt to improve the vapour-liquid equilibrium (VLE) predictions of the RK EOS, Soave [42] modified the RK EOS by introducing a dimensionless term, α, which resulted to the SRK EOS as follows:

$$P = \frac{RT}{V-b} - \frac{a\alpha}{V(V+b)}. \tag{2.9}$$

At dewpoint pressure, the mole fractions of the vapour phase, y_i, equals to unity as shown in (2.10)

$$\sum_{i=1}^{N} y_i = 1. \tag{2.10}$$

The vapour phase mole fractions can be expressed in terms of K-values as

$$\sum_{i=1}^{N} y_i = \sum_{i=1}^{N} \frac{y_i}{K_i} = 1 \tag{2.11}$$

where *N* represents the number of components in the entire mixture, y_i is the mole fraction of component *i* in the mixture when the first drop of liquid phase appears, and K_i denotes K-value of component *i* which is the ratio of the mole fraction of component *i* in the gas phase to the mole fraction of component *i* in liquid phase. Thus, the dewpoint pressure can be calculated using the following expression,

$$P_d = \sum_{i=1}^{N} = \frac{f_i^V}{\varphi_i^L} \tag{2.12}$$

where f stands for the vapour fugacity of the *i*th component, which is calculated from the EOS, and φ stands for the liquid fugacity coefficient.

Peng-Robinson (PR) EOS

An equation that contains two constants was proposed by Peng and Robinson. PR EOS created great expectations for improved VLE predictions in petroleum fluids. The PR EOS is expressed as

$$P = \frac{RT}{V-b} - \frac{a}{V(V+b)+(V-b)}. \tag{2.13}$$

The Peng-Robinson and Soave-Redlich-Kwong equations are the two most widely used cubic EOS for VLE and phase predictions of hydrocarbon fluids. Calculation of the parameters is what makes the difference between SRK-EOS and PR-EOS.

2.3.4.4 Equilibrium Ratio (K-values) Approach

Another classic method for calculation of dewpoint pressure of multicomponent systems is the K-value method. The K-value of a component in a mixture is defined as the ratio of the component mole fraction in the gas phase to that in the liquid phase [43]. K-values are thermodynamic quantities that depend on temperature, pressure and composition of the mixture and is expressed as

$$K_i = \frac{y_i}{x_i} \tag{2.14}$$

where y stands for the mole fraction of component *i* in the gas phase, and x stands for the mole fraction of component *i* in the liquid phase. For a gas condensate system, this approach requires that the fugacity of each component must be the same in both phases. That is,

$$f_i^V = f_i^L (i = 1,2,\ldots,N). \tag{2.15}$$

The superscripts *V* and *L* express vapour, and liquid phases, respectively, the fugacity of component *i* denoted by f_i can be calculated by EOS, *N* represents the total number of components in the system. To calculate the dewpoint pressure, the equilibrium condition in equation 2.15 must be met, as well as the following expression,

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} \frac{y_i}{K_i} = 1 \tag{2.16}$$

where y_i denotes the mole fraction of component *i* in the mixture when the first drop of liquid phase appears. Equation 2.16 is nonlinear and can be solved iteratively using initial guesses of equilibrium ratios, K_i as defined in equation 2.14, and initial pressure. The value of K_i can be corrected by

$$K_i = \left(\frac{y_i}{x_i}\right)\left(\frac{f_i^L}{f_i^V}\right) = \frac{\phi_i^L}{\phi_i^V} \tag{2.17}$$

where ϕ stands for the fugacity coefficient of component *i* and can be calculated using EOS.

2.3.4.5 Dewpoint Pressure Prediction and Reservoir Flow Characterization Using Data-Driven Based Machine Learning Methods

In recent years, data-driven techniques have been widely applied in the field of reservoir modeling and simulation, and they are highly regarded for their strong ability to account for nonlinearity of fluid flow in porous media. The authors in [44] and [39] applied Artificial Neural Networks (ANN) to predict gas condensate dewpoint pressure, and evaluate the dynamics of reservoir parameters under different conditions of CO_2 injections, respectively. In [34], multigene genetic programming was applied to predict the dewpoint pressure of gas condensate fluids. In their approach, the authors used 158 experimental data points of Constant Volume Depletion (CVD) tests and consider the dewpoint pressure as a function of reservoir temperature, molecular weight of heavier components, mole fraction of heavier, intermediate and volatile components. In [45] Radial Basis Function (RBF) network coupled with Genetic Algorithm (GA) was used to predict the dewpoint pressures of gas condensate fluids using 562 experimental data. The authors in [46] presented an approach based on Least Square Support Vector Machine (LSSVM) coupled with Simulated Annealing (CA) to predict the dewpoint pressures of gas condensate fluids using a total of 562 experimental data.

Even though the data-driven based machine learning approaches applied in previous studies present reasonable estimations of the dewpoint pressures for certain gas condensate reservoirs, however, they solely rely on huge amount of training data for accurate results and

do not put into consideration any particular prior knowledge of the characteristics or physics of the relationships involved in gas condensate reservoirs. As such, they are lacking in term of good generalization ability, and are unable to explain the general working principle of gas condensate fluids. Thus, they can be described as black box models.

Another noticeable drawback with the data-driven based machine learning methods applied previously is that, the models reported to address the dewpoint pressure estimation were developed based on the common assumptions that when a computational intelligence model is developed and trained using a huge amount of historical datasets, the resulting model will, to some extent, handle similar situations during their operations. However, in many real-world operations, these assumptions may tend to fail as in most cases as the amount of previously available data used in developing and training the model may not be sufficient to represent the underlying system. Furthermore, the environment as well as the system and its constituents may change over time. In such a situation, it becomes difficult for a model that is developed and trained using only iterative algorithms to cope with such changes and process the data efficiently.

Therefore, this study aims to address the aforementioned challenges and limitations of the data-driven based machine learning methods previously applied for modelling of pressure dynamics and fluid flow in nature gas reservoirs. In particular, this work utilises machine learning approach that combines the power of Fuzzy Inference System that transforms stream data into a knowledge-based inference, and the learning ability of ANN that learns the patterns and behaviours of the data, thereby resulting to adaptive models that are capable of adapting to changing situations and generating the desired result even if they encounter a situation that is different from the history enclosed in the historical datasets.

2.4 Reservoir Simulation Software/Platform

A large number of programs/applications are available for reservoir simulation. These applications can be classified broadly into two categories namely: open-source and commercial. Following is a brief description (in alphabetical order) of the most well-known applications for hydrocarbon reservoir simulation:

i. BOAST-NFR: Black Oil Applied Simulation Tool for Naturally Fractured Reservoirs (BOAST-NFR) is an open-source package for reservoir simulation published by the United States Department of Energy (DOE) and available for educational purposes [11]. It is a three-dimensional, three-phase, black oil simulator that employs the implicit pressure - explicit saturation (IMPES) formulation for solving partial differential flow equations by the way of finding the pressure distribution for a given time step first, then computes the saturation distribution for the same time step isothermally.

ii. ECLIPSE: This is a well-known commercial reservoir simulator that was originally developed by ECL and currently owned, maintained and marketed by a division of Schlumberger. In addition to add-on options, the software contains tools for streamline, compositional and black oil simulation.

iii. INTERSECT: This is a commercial high-resolution reservoir simulator developed by Schlumberger that can be used for various reservoir simulation tasks such as modelling of highly heterogeneous formations and complex geological structures.

MATLAB Reservoir Simulation Toolbox (MRST). The MRST is an open-source software for modelling and simulation of reservoir fluid flow and transport in porous media. It was developed by the Computational Geosciences research group at SINTEF Digital, Norway [12]. The software has a wide international user and also, external modules developed by other research groups from University of Bergen, Heriot-Watt University, TU Delf, and Norwegian University of Science and Technology. Similar to MATLAB, MRST is not only a simulator, but also, developed as a tool that enables researchers to rapidly prototype and demonstrate new modelling concepts and simulation methods. Thus, the software offers a wide range of computational methods and data structures that one can combine to develop a custom-made tool for reservoir modelling and simulation. Furthermore, MRST provides quite in-depth compositional and black oil reservoir simulators that are capable of simulating industry-standard models as well as graphical user interfaces designed for post-processing simulation results. The MRST software is structured into two major sections as follows: (a) a small core module providing functionality and fundamental data structures and, (b) a large collection of add-on modules providing solvers, discretization, and a wide variety of

workflow tools and simulators [13] - [15]. In this work, MRST is used to carry out numerical simulations for all the reservoirs models presented in this thesis.

2.5 Challenges and Motivations

The importance of reservoir modelling and simulation to estimate field performance under one or various producing plans in the oil and gas industry cannot be underestimated. Experimental approaches for predicting gas condensate dewpoint pressure are time-consuming, intrusive, costly, and sometimes dangerous. Even though black oil simulators can be used to model fluid in three phases (gas, oil and water) and are suitable for studies of numerous problems, it is worthwhile mentioning that black oil simulation has some drawbacks in areas of application. Black oil simulation approach gives no consideration to hydrocarbon fluid composition and in areas where mass transfer between phases is important, black oil simulators are not suitable for studying such cases. For instance, black oil simulators are not suitable for studying problems related to volatile oil and gas condensate reservoirs because in these reservoirs, the physical properties and fluid composition of the phases change with pressure [11].

On the other hand, compositional simulations are considered suitable for modelling and simulating lean gas cycling in the presence of oil by the way of causing the oil to vaporize in the lean gas. However, it should be noted that the pseudo component approach in compositional simulators normally results in lumping all of the heavier fluid components together. As such, this approach leads to over predicting the vaporization of the pseudo component. To reduce the errors caused by pseudo component vaporization, a good compositional model should be able to accommodate all the components that constitute the reservoir fluid under study.

It has been observed and reported that calculating dewpoint pressure using SRK-EOS and PR-EOS yields different results [47]. As such, to calculate the dewpoint pressure using the EOS, consideration must be given to the choice of the EOS. In addition, use of EOS requires that the heptane plus fraction be split into several pseudo-fractions and each of the sub fractions be characterized [43] [48]. Furthermore, when using EOS to predict other physical quantities, it must be first, tuned properly to match some experimental data for a specific

reservoir fluid [49]. In doing so, quite often the influence of composition on dewpoint pressure is neglected. Another drawback of using the EOS to predict thermodynamic property and phase behaviour of reservoir fluids is the difficulty of characterizing the plus fraction properly. Since the hydrocarbon plus fraction comprises a significant portion of the reservoir fluid, improper characterization of this component will lead to errors in predicting some quantities that depend on them.

Equilibrium ratios (K-values) approach defines pressure as a function of K-values and involves trial and error. Trial and error approach associated with K-values while calculating the dewpoint pressure makes the method unreliable and most of K-values approaches at high pressures are inaccurate [48]. It has also been observed that K-values are dependent on pressure, temperature and composition and their use adds nothing to the accuracy of predicting dewpoint pressure [40]. In addition, it has been reported that K-values approach is only suitable for calculating dewpoint pressure at low or moderate pressures. At high pressures, the method tends to be slow and imprecise [43].

This chapter presents a critical survey of literature in relation to hydrocarbon reservoir classification and simulation techniques, gas condensate reservoir performance, and dewpoint pressure prediction for gas condensate reservoirs. Drawbacks of existing methods in learning pressure dynamics and predicting gas condensate dewpoint pressure are identified, and how the present work attempts to address such drawbacks are presented. The next chapter presents a physics-informed intelligent method for predicting the pressure dynamics and production trend of a single-phase gas condensate reservoir model with varied BHFPs.

References

[1] W. Lyons. *Working Guide to Petroleum and Natural Gas Production Engineering*. 1st Ed., pp 391 - 409, 2010. Burlington: Gulf Publishing

[2] R. P. Batycky, M. R. Thiele, K.H. Coats, A. Grindheim, D. Ponting, J. E. Killough, T. Settari, L. K. Thomas, J. Wallis, J.W. Watts and C. H. Whitson (1994). Reservoir Simulation. In: L. W. Lake (1965). *Petroleum Enginnering Handbook*, vol. V. SPE, Zulia, pp 1399-1465.

[3] C. H. Whitson and M. R. Brule. *Phase Behavior*. *Soc. Pet. Eng.* Vol. 20, pp. 47-65, 2000.

[4] K. H. Coats, L. K. Thomas, and R. G. Pierson. Compositional and black oil reservoir simulation. *Soc. Pet. Eng.*, pp. 149–162, 1995, doi: 10.2118/29111-MS.

[5] B. Li, Z. Chen, and G. Huan, "The sequential method for the black-oil reservoir simulation on unstructured grids," *J. Comput. Phys.*, vol. 192, no. 1, pp. 36–72, 2003, doi: 10.1016/S0021-9991(03)00346-2.

[6] Z. E. Heinemann. *Introduction to Reservoir Simulation*, vol 4, PHDG, pp 2-104, 2001. New York.

[7] Z. Chen, G. Huan and Y. Ma (2015). *Computational Methods for Mutiphase Flows in Porous Media*, vol. 1, pp 9-46, 2015. Philadelphia: SIAM.

[8] M. Ahmadi, M. Sharifi, and A. Hashemi. Comparison of simulation methods in gas condensate reservoirs. *Pet. Sci. Technol.*, vol. 32, no. 7, pp. 761–771, 2014, doi: 10.1080/10916466.2011.604063.

[9] Y. Wang and J. E. Killough. A New Approach to Load Balance for Parallel / Compositional Simulation Based on Reservoir-Model Overdecomposition. *SPE J.*, vol. 19, no. 02, pp. 304–315, 2014.

[10] B. M. Tayyebi, M. Chenani, and S. N. Kashkooli. Oil and Gas Reservoirs Management Principles: Challenges and Strategies for Optimal Preservation of Hydrocarbon Reserves in the Common Fields. *Int. J. Acad. Res. Econ. Manag. Sci.*, vol. 3, no. 4, pp. 52–64, 2014, doi: 10.6007/IJAREMS/v3-i4/1067.

[11] J. G. Almengor *et al.*. *User's Guide and Documentation Manual For BOAST-NFR For Excel.* Mewbourne School of Petroleum and Geological Engineering, The University of Oklahoma, 2002.

[12] S. Krogstad, K. A. Lie, O. Møyner, H. M. Nilsen, X. Raynaud, and B. Skaflestad. MRST-AD – an open-source framework for rapid prototyping. *SPE Reserv. Simul. Symp. 23-25 February, Houston, Texas, USA*, 2015.

[13] K.-A. Lie. *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST)*. 1st ed, pp. 597 – 630, 2019. Cambridge: Cambridge University Press.

[14] K. A. Lie, S. Krogstad, I. S. Ligaarden, J. R. Natvig, H. M. Nilsen, and B. Skaflestad, "Open-source MATLAB implementation of consistent discretisations on complex grids," *Comput. Geosci.*, vol. 16, no. 2, pp. 297–322, 2012, doi: 10.1007/s10596-011-9244-4.

[15] Lie, K.-A. 2015a. JOLT 1: *Introduction to MRST*. SINTEF ICT / ICME, Stanford University. www.sintef.no/mrst-jolts.

[16] S. Al-Rbeawi. Analysis of pressure behaviors and flow regimes of naturally and hydraulically fractured unconventional gas reservoirs using multi-linear flow regimes approach. *J. Nat. Gas Sci. Eng.*, vol. 45, pp. 637–658, 2017, doi: 10.1016/j.jngse.2017.06.026.

[17] M. A. Ahmadi and M. Ebadi. Evolving smart approach for determination dew point pressure through condensate gas reservoirs. *Fuel*, vol. 117, no. PARTB, pp. 1074–1084, 2014, doi: 10.1016/j.fuel.2013.10.010.

[18] M. Sheydaeemehr, B. Sedaeesola, and A. Vatani. Gas-condensate production improvement using wettability alteration: A giant gas condensate field case study. *J. Nat. Gas Sci. Eng.*, vol. 21, pp. 201–208, 2014, doi: 10.1016/j.jngse.2014.07.011.

[19] M. Mahdaviara, N. A. Menad, M. H. Ghazanfari, and A. Hemmati-Sarapardeh. Modeling relative permeability of gas condensate reservoirs: Advanced computational frameworks. *J. Pet. Sci. Eng.*, vol. 189, no. September 2019, p. 106929, 2020, doi: 10.1016/j.petrol.2020.106929.

[20] R. E. Terry and J. B. Rogers (2014). Introduction to petroleum reservoirs and reservoir engineering. In: A. B. Zolotukhin and J. R. Ursin. *Applied Petroleum Reservoir Engineering*, 3rd ed., pp 1-14, 2000. New Jersey: Wiley.

[21] J. Shi, L. Huang, X. Li, and K. Sepehrnoori. Production forecasting of gas condensate

well considering fluid phase behavior in the reservoir and wellbore. *J. Nat. Gas Sci. Eng.*, vol. 24, pp. 279–290, 2015, doi: 10.1016/j.jngse.2015.03.033.

[22] M. Alarouj, O. Alomair, and A. Elsharkawy. Gas condensate reservoirs: Characterization and calculation of dew-point pressure. *Pet. Explor. Dev.*, vol. 47, no. 5, pp. 1091–1102, 2020, doi: 10.1016/S1876-3804(20)60120-3.

[23] R. Mokhtari, F. Varzandeh, and M. R. Rahimpour. Well productivity in an Iranian gas-condensate reservoir: A case study. *J. Nat. Gas Sci. Eng.*, vol. 14, pp. 66–76, 2013, doi: 10.1016/j.jngse.2013.05.006.

[24] O. Ajagbe and M. Fahes. Establishing screening criteria for field application of wettability alteration in gas-condensate reservoirs. *J. Pet. Sci. Eng.*, vol. 193, no. January, p. 107342, 2020, doi: 10.1016/j.petrol.2020.107342.

[25] L. Fan, B. Harris, and A. Jamaluddin. Understanding Gas-Condensate Reservoirs. *Oilf. Rev.* pp 14–27, 2005.

[26] H. Kaydani, A. Hagizadeh, and A. Mohebbi. A dew point pressure model for gas condensate reservoirs based on an artificial neural network. *Pet. Sci. Technol.*, vol. 31, no. 12, pp. 1228–1237, 2013, doi: 10.1080/10916466.2010.540616.

[27] A. Rahimzadeh, M. Bazargan, R. Darvishi, and A. H. Mohammadi. Condensate blockage study in gas condensate reservoir. *J. Nat. Gas Sci. Eng.*, vol. 33, pp. 634–643, 2016, doi: 10.1016/j.jngse.2016.05.048.

[28] R. A. Dawe and C. A. Grattoni. Fluid flow behaviour of gas-condensate and near-miscible fluids at the pore scale. *J. Pet. Sci. Eng.*, vol. 55, no. 3–4, pp. 228–236, 2007, doi: 10.1016/j.petrol.2006.08.009.

[29] R. F. Krueger. SPE 17459 An Overview of formation damage and well productivity in oilfield operations: An update. *Soc. Pet. Eng.*, no. February, pp. 535–553, 1988, doi: 10.2118/10029-PA.

[30] M. M. Sharma and Y. C. Yortsos. Permeability impairment due to fines migration in sandstones. *Proc. SPE Form. Damage Control Symp.*, 1986, doi: 10.2118/14819-MS.

[31] D. H. GKAY and R. W. REX. Formation damage in sandstones caused by clay dispersion and migration. *Clays Clay Miner.*, no. 1960, pp. 355–366, 1966, doi: 10.1016/B978-0-08-011908-3.50033-5.

[32] M. J. Economides and K. G. Nolte. *Reservoir Stimulation*. 2nd ed, pp. 208 – 226, 1989. New Jersey: Wiley.

[33] S. Karami, A. H. Saeedi Dehaghani, and S. A. Hossein Seyed Mousavi. Condensate blockage removal using microwave and ultrasonic waves: Discussion on rock mechanical and electrical properties. *J. Pet. Sci. Eng.*, vol. 193, no. March, p. 107309, 2020, doi: 10.1016/j.petrol.2020.107309.

[34] H. Kaydani, A. Mohebbi, and A. Hajizadeh. Dew point pressure model for gas condensate reservoirs based on multi-gene genetic programming approach. *Appl. Soft Comput. J.*, vol. 47, pp. 168–178, 2016, doi: 10.1016/j.asoc.2016.05.049.

[35] L. Fan, B. Harris, and A. Jamaluddin. Understanding gas-condensate reservoirs. *Oilf. Rev.*, 10(4):16-25, 2005.

[36] E. Meehan. An analysis of rate-sensitive skin in gas wells. *Soc. Pet. Eng., pp 1 – 13*, 1983.

[37] S. B. Hinchman. Productivity loss in gas condensate reservoirs. *Soc. Pet. Eng.*, pp 1 -8, 1965.

[38] M. Arabloo, A. Shokrollahi, F. Gharagheizi, and A. H. Mohammadi. Toward a predictive model for estimating dew point pressure in gas condensate systems. *Fuel Process. Technol.*, vol. 116, pp. 317–324, 2013, doi: 10.1016/j.fuproc.2013.07.005.

[39] H. Kaydani, A. Hagizadeh, and A. Mohebbi. A dew point pressure model for gas condensate reservoirs based on an artificial neural network. *Pet. Sci. Technol.*, vol. 31, no. 12, pp. 1228–1237, 2013, doi: 10.1080/10916466.2010.540616.

[40] L. K. Nemeth and H. T. Kennedy. A Correlation of dewpoint pressure with fluid composition and temperature. *Soc. Pet. Eng. J.*, vol. 7, no. 02, pp. 99–104, 1967, doi: 10.2118/1477-PA.

[41] M. Arabloo and S. Rafiee-Taghanaki. SVM modeling of the constant volume depletion (CVD) behavior of gas condensate reservoirs. *J. Nat. Gas Sci. Eng.*, vol. 21, pp. 1148–1155, 2014, doi: 10.1016/j.jngse.2014.11.002.

[42] G. Soave. Equilibrium constants from a modified Redlich-Kwong equation of state. *Chem. Eng. Sci.*, vol. 27, no. 6, pp. 1197–1203, 1972, doi: 10.1016/0009-2509(72)80096-4.

[43] D. H. Xu, A. Danesh, and A. C. Todd. An accelerated successive substitution method for calculation of saturation pressure of multicomponent fluids. *Fluid Phase Equilib.*, vol. 72, no. C, pp. 15–24, 1992, doi: 10.1016/0378-3812(92)85016-2.

[44] S. Amini and S. Mohaghegh. Application of machine learning and artificial intelligence in proxy modeling for fluid flow in porous media. *Fluids*, vol. 4, no. 3, pp. 1–17, 2019, doi: 10.3390/fluids4030126.

[45] H. Rostami-Hosseinkhani, F. Esmaeilzadeh, and D. Mowla. Application of expert systems for accurate determination of dew-point pressure of gas condensate reservoirs. *J. Nat. Gas Sci. Eng.*, vol. 18, pp. 296–303, 2014, doi: 10.1016/j.jngse.2014.02.009.

[46] S. Rafiee-Taghanaki, M. Arabloo, A. Chamkalani, M. Amani, M. H. Zargari, and M. R. Adelzadeh. Implementation of SVM framework to estimate PVT properties of reservoir oil. *Fluid Phase Equilib.*, vol. 346, pp. 25–32, 2013, doi: 10.1016/j.fluid.2013.02.012.

[47] E. M. E. M. Shokir. Dewpoint pressure model for gas condensate reservoirs based on genetic programming. *Energy and Fuels*, vol. 22, no. 5, pp. 3194–3200, 2008, doi: 10.1021/ef800225b.

[48] M. A. Ahmadi and A. Elsharkawy. Robust correlation to predict dew point pressure of gas condensate reservoirs. *Petroleum*, vol. 3, no. 3, pp. 340–347, 2017, doi: 10.1016/j.petlm.2016.05.001.

[49] A. M. Elsharkawy. Predicting the dew point pressure for gas condensate reservoirs:

Empirical models and equations of state. *Fluid Phase Equilib.*, vol. 193, no. 1–2, pp. 147–165, 2002, doi: 10.1016/S0378-3812(01)00724-5.

Chapter Three: Neuro-Adaptive Learning Approach for Predicting Production Performance and Pressure Dynamics of Gas Condensation Reservoir

3.1 Introduction

In Reservoir Engineering, state-of-the-art data analysis enables engineers to characterize reservoirs and plan for developing fields. This allows production companies to save huge amounts that would otherwise be allocated to reservoir modelling and simulation, as well as well testing. Numerical reservoir modelling and simulation is the standard use in industries today for comprehensive study of oil and gas fields. However, the inflexible behaviour, development time, and cost of numerical simulators are major challenges to production engineers, managers and modellers. On the other hand, Machine Learning (ML) based reservoir models are characterised with low cost of development, short development time and fast track analysis, and flexibility to estimate the uncertainties normally found in numerical simulators. In this chapter, an approach for controlling gas production rate and reservoir pressure drop in gas condensate reservoir is described. Numerical simulations of production rate and pressure drop were carried out first. An adaptive Neuro-Fuzzy system was then developed and trained using some parts of the numerical simulation results. This ML-based system is validated and tested with the other part of the numerical simulation results that had not been used during the training. The developed system regenerates the numerical simulation results for both production rates and pressure drop at different Bottom Hole Pressures (BHPs) with very high accuracy (>98%). Results of this chapter showed that ML-based reservoir simulation can be considered as a vital tool for production engineers, managers, and modellers for a quick and more informed decision as regards field development plans that can meet operational targets.

3.2 Problem Statement and Model of GCR

Numerical simulators are the standard today in oil and gas industries to simulate reservoirs. Most of these simulators are developed based on Implicit Pressure, Explicit Saturation (IMPES) method where pressure and transport equations are solved in separate steps using numerical techniques that are designed to utilize special features of such equations. Conventionally, developing a gas condensate field and investigating its production behaviour

has several challenges: 1. Understanding the complexity of the system and its physics at microscopic level, 2. Up-scaling such complexities and physics to reservoir simulation scale, 3. Modelling and incorporating the petro-physical properties into the simulation model. Apart from the aforementioned challenges, another concern when trying to model a gas condensate field is the fact that the economic production of these special reservoirs depends on the number of phases that exist in the reservoir. In line with this, modelling a gas condensate field becomes even more complicated with the existence of two phases (gas and condensate).

With recent interest and enthusiasm in oil and gas industry towards intelligent fields, smart wells, and analysis of data for process optimization and decision-making, AI/ML techniques are making firm steps toward becoming more preferred in the widespread of the oil and gas industry [6]. However, existing machine learning methods applied to investigate problems related to gas condensate reservoirs are limited in the literature. Furthermore, the available ones are focused on prediction of dewpoint pressure where most of them rely on measured data for characterizing reservoirs and give no consideration to production trend and pressure dynamics which are essential for accurate prediction of gas condensate dewpoint pressure and other reservoir quantities. These limited existing machine learning techniques for dewpoint prediction include multi-gene genetic programming [7], artificial neural network [8], least square support vector machine [3], genetic programming [9], and expert systems [10]. Apart from their reliance on not physics-based data, they also consider dewpoint pressure as a function of only fluid composition, reservoir temperature, specific gravity and molecular weight of heptane plus (C_{7+}). Thus, they are also limited given that they neglect some relevant geometric and flow parameters such as bottom-hole and reservoir pressures, reservoir geometries, fluid density, viscosity and compressibility, porosity, and permeability. These parameters have impact on analysing reservoir performance and thus any model that would be used to accurately characterize a reservoir for managerial decision and operational plan, should incorporate these parameters.

Therefore, the objective of this chapter is to address the aforementioned challenges and limitations of existing techniques by demonstrating the potentiality of ML-based reservoir simulation systems that can capture production and pressure trends of a gas condensate reservoir based on physics-informed data.

A gas condensate reservoir (GCR) is a system composed of predominantly methane and other light hydrocarbons with a slight portion of heavier components situated underground at elevated conditions of pressure and temperature [1]. To produce from such a reservoir, components must embark on a complicated journey from beneath the earth, passing through a great deal of intermediate stages, to be finally dumped into our atmosphere. A model of GCR is presented Figure 3.1.

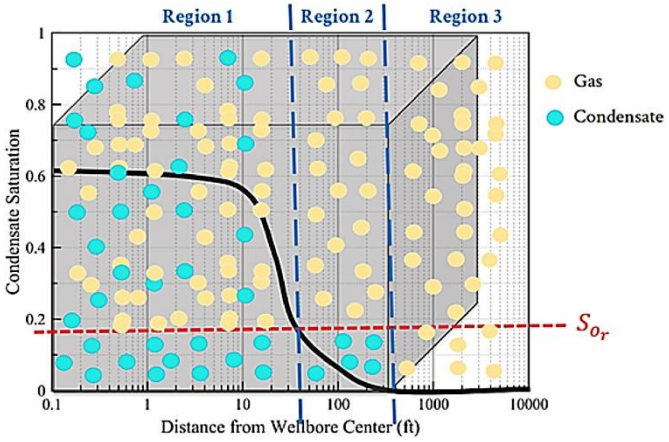

Figure 3.1: A model of gas condensate reservoir [2]

A model of GCR can be explained using three regions as shown in Figure 3.1. At stage of discovery, a gas condensate reservoir is a single-phase (gaseous state) fluid and the reservoir pressure is above the dewpoint pressure. This is depicted in region 3 of Figure 3.1. As production commences, formation temperature mostly remains unchanged, but pressure drops as the gas components begin to move towards the wellbore. Thus, the reservoir experiences decline in pressure from the boundaries to the well as production continues. When the pressure within the reservoir drops to a certain point, called the dewpoint, condensates drop out of the mixture as predicted in region 2 of Figure 3.1. At this stage, the reservoir splits from single phase (gas) to two-phase (gas and condensate). The dewpoint pressure is considered as one of the most valuable parameters capable of describing and predicting the future behaviour of gas condensate systems. It is defined as the pressure at which a massive quantity of gas exists in equilibrium with a minute quantity of liquid [3]. Continuous production below the dewpoint causes an increase in the level of the condensates which in consequence, obstructs the flow of gas towards the wellbore as depicted in region 1 of Figure 3.1. When pressure decreases isothermally as a result of continuous production, instead of having gas coming out of the system, liquid condenses from the gas. This is what

makes the gas condensate reservoirs special and are sometimes referred to as “retrograde systems”. To model and simulate a GCR, fluid composition is used to compute fluid properties such as fluid density, fluid viscosity, fluid compressibility. These variables together with reservoir initial pressure, temperature, and rock properties (porosity, permeability, rock compressibility) are used as input variables to the GCR model. Depending on the aim of the study, the output variable can the average reservoir pressure as a function of time, cumulative gas production over time, gas saturation over time, well pressure (well head or bottom hole pressure) over time.

In addition to their unique retrograde behaviours, these unique reservoirs also differ in their thermodynamic and flow behaviour from other common oil and gas reservoirs due to the fact that they exist as a single gas phase at the time of discovery [4]. This key fact nearly always governs the development and operation plans for recovery of hydrocarbons from gas condensate reservoirs [5]. Control over pressure in gas condensate reservoirs is necessary due to the fact that production declines significantly when reservoir pressure falls below the dewpoint pressure. Reliable prediction of petro-physical quantities such as the dewpoint pressure over the life of these reservoirs requires an accurate knowledge and clear understanding of the flow behaviour of gas condensate fluid through reservoir rock. Hence, for optimum gas condensate engineering computations such as predicting productivity trend and pressure dynamics, a suitable model that represent the realistic operations of gas condensate reservoir is essential.

The rest of this chapter is organised as follows. Section 3.3 develops a numerical simulation model using MATLAB Reservoir Simulation Toolbox [11], and tests the model by using a fluid data of Pars gas condensate field [12]. Section 3.4 describes the development of the Neuro-Adaptive system. Section 3.5 discusses the results of both the numerical simulation and the Neuro-Adaptive system. Section 3.6 ends the chapter with conclusions and future work.

3.3 Numerical Modelling and Simulation Implementation

This section presents the major components/objects of the numerical simulation model for prediction of production performance and pressure dynamics in gas condensate reservoirs. These components are sufficient to represent the physical trend of gas condensate reservoirs

which might not only be of benefit to production engineers, but also build confidence among operational team and lessens the risk of complex model errors. The major steps involved in developing the numerical simulation model are summarized as follows:

- Set up reservoir model with rock properties
- Define fluid and production well
- Impose vertical equilibrium and solve initial pressure distribution
- Compute mapping between faces and cells
- Define discrete operators, flow, and well equations
- Define simulation parameters
- Simulate reservoir at different conditions and visualize results
- Save simulation results.

Figure 3.2 shows a schematic diagram that illustrates the steps involved in simulating the reservoir model numerically.

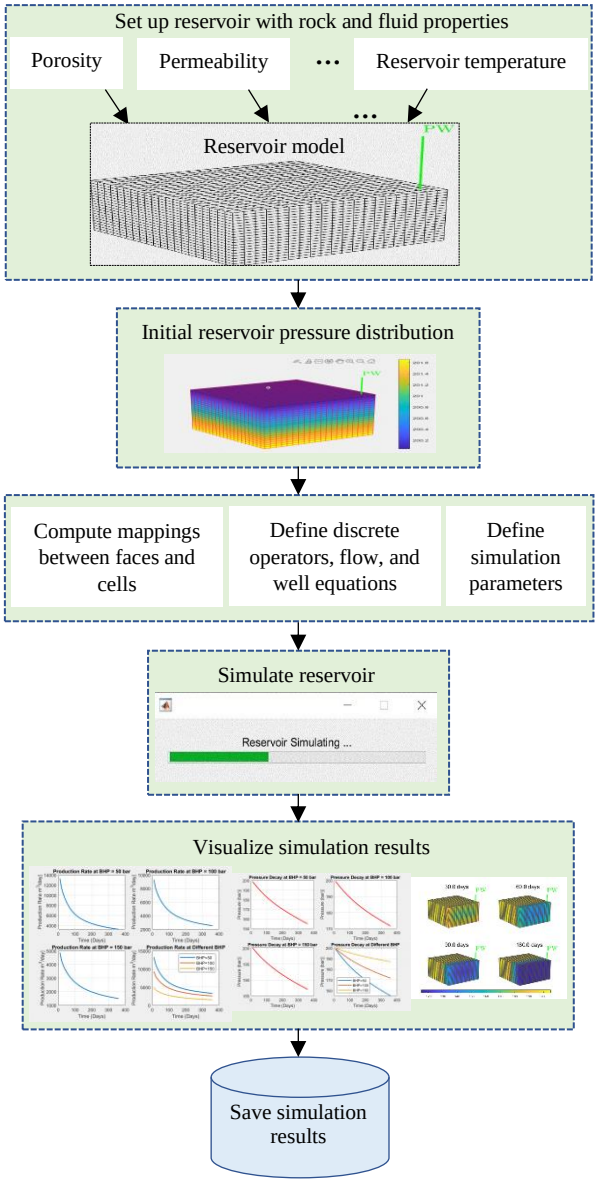

Figure 3.2: Schematic illustration of the numerical simulation development process

3.3.1 Reservoir Rock and Fluid Properties

The reservoir model in this chapter is assumed to be a homogeneous 500 × 500 × 25 m^3 box with rock properties represented on a regular 20 × 20 × 5 Cartesian grid which consists of set of grid cells in three-dimensional volumetric form that allows the distributions of petro-physical parameters. The petro-physical parameters are values stored inside each grid cell which in turn, defined the rock properties. Parameters that characterize the reservoir rock and fluid properties used in developing the numerical simulation model are presented in Table 3.1. The fluid composition used for the numerical simulation is given in Table 3.2, representing the fluid composition of Pars field.

Table 3.1: Reservoir rock and fluid parameters.

Porosity (%)	15
Permeability (mD)	0.5
Rock compressibility (bar^{-1})	0.0156
Fluid density (kg/m^3)	173
Fluid viscosity (cP)	0.0221
Fluid compressibility (bar^{-1})	0.00286
Reservoir pressure (bar)	200
Reservoir temperature (^0F)	212

Table 3.2: Fluid composition used for numerical simulation [12].

Component	**C_1**	**C_2**	**C_3**	**iC_4**	**nC_4**	**iC_5**
Composition (%)	89.24	2.28	0.51	0.12	0.13	0.06
Component	**nC_5**	**C_6**	**C_{7+}**	**N_2**	**CO_2**	**H_2S**
Composition (%)	0.04	0.06	0.24	1.7	4.96	0.66

3.3.2 Equations Governing Fluid Flow

In this section, fluid flow in gas condensate reservoir is descried by the way of combining the equation of conservation of mass (3.1) and Darcy's law of fluid flow in porous media (3.2).

$$\frac{\partial}{\partial t}(\varepsilon\rho) + \nabla \cdot (\rho v) = pw \tag{3.1}$$

$$v = -\frac{K}{\mu}(\nabla P - g\rho\nabla z) \tag{3.2}$$

where $\varepsilon, \rho, v, K, \mu,\ g, z, P,$ and pw represent porosity, fluid density, flow velocity, permeability, fluid viscosity, gravity, reservoir vertical direction, reservoir pressure, and production well respectively. Before production commences, we assume the reservoir is at hydrostatic equilibrium as such, for the reservoir pressure to be distributed within the reservoir, the following equation must be satisfied.

$$\frac{dP}{dz} = \rho g(P) \tag{3.3}$$

The hydrostatic distribution within the reservoir is computed numerically with respect to a fixed point $p(z_0) = P$. This distributes the initial pressure in all the cells of the reservoir grid and sets the reservoir at its initial condition. Figure 3.3 shows a plot of the reservoir with a single production well and initial pressure distributed within the grid.

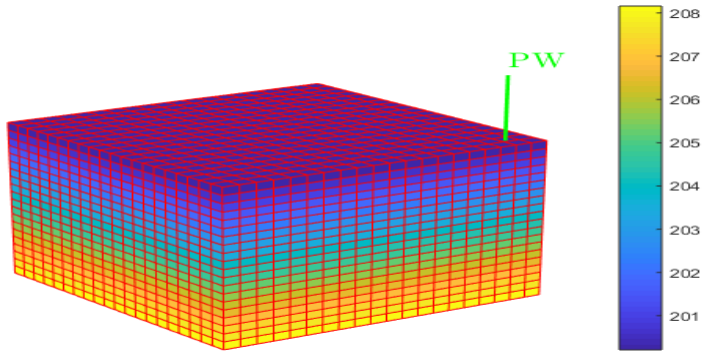

Figure 3.3: Reservoir model with production well and initial pressure distribution

3.3.3 The Production Well

The production well appears as a source term in Eq. (3.1). Hence, we need to establish a connection between the production well and the reservoir. This is accomplished by the way of defining an equation that accounts for the rate of flow within the grid cells in which the reservoir is connected to the production well. By approximating fluid density and enabling computations at the wellbore, the pressure $P[c_{on}]$ in connection c_{on} of production well $P_w[c_{on}]$ is given as

$$P_c[c_{on}] = P_{bh}[N_w(c_{on})] + g\Delta z[c_{on}]\rho(P_{bh}[N_w(c_{on})]) \tag{3.4}$$

where $\Delta z[c_{on}]$ stands for distance from the well connection to the wellbore. Using Production Indices PI, we relate the average pressure within the grid cells to the pressure at the well connection and write the equation for the mass flow-rate as follows

$$Q_c[c_{on}] = \frac{\rho(P[N_c(c_{on})])}{\mu} PI[c_{on}](P_c[c_{on}] - P[N_c(c_{on})]) \tag{3.5}$$

where $P[N_c(c_{on})]$ represents the pressure in the cell $N_c(c_{on})$ containing connection c_{on}.

3.4 Neuro-Adaptive Learning Approach for Gas Condensate Reservoirs Proxy Model Development

3.4.1 ML-Based Proxy Modelling

A proxy model is a model that mimics the outputs of a more complex model in a manner that is simpler and resource-saving. Therefore, a proxy model can simply be seen as a model that models another model. This technique is becoming popular and of great help in oil and gas industry due to their reduced size and short run-time which are essential for master development planning and operating a field [13], [14]. This section considers Neuro-Adaptive Learning approach suitable for proxy modelling applicable to gas condensate field by using Adaptive Neuro-Fuzzy Inference System (ANFIS) to train a Fuzzy Inference System (FIS) which in turn regenerates the outputs of a numerical simulation model. The proxy model developed in this chapter uses a hybrid (a combination of back propagation algorithm and least square) optimization method for adjusting or turning the parameters of the model's Membership Functions (MFs).

3.4.2 ML-Based Proxy Model Development Procedure

Generation of input/output data set is the starting point and the most important step when developing a Neuro-Adaptive Learning System. In this chapter, Tables 3.1 shows the parameters which were used in the numerical simulation model as input data. The data were used to simulate the reservoir for 360 days (approximately 1 years). Upon running the simulations at different bottom hole pressures, spatiotemporal database was generated for each simulation. These databases were developed by extracting the static and dynamic variables from the numerical simulations. In this context, static variable refers to a variable that does not change with time (porosity, permeability, and temperature), whereas dynamic variable refers to a variable that changes with time (timestep, bottom hole pressure, gas production rate, and average reservoir pressure). Two ML models based on Sugeno fuzzy inference systems (FISs) were then developed and trained using ANFIS to predict the gas production rate and reservoir pressure change, respectively. Diagrams of input and output

mappings for the two ML-based models for predictions of gas production rate and reservoir pressure change are shown in Figures 3.4 and 3.5, respectively.

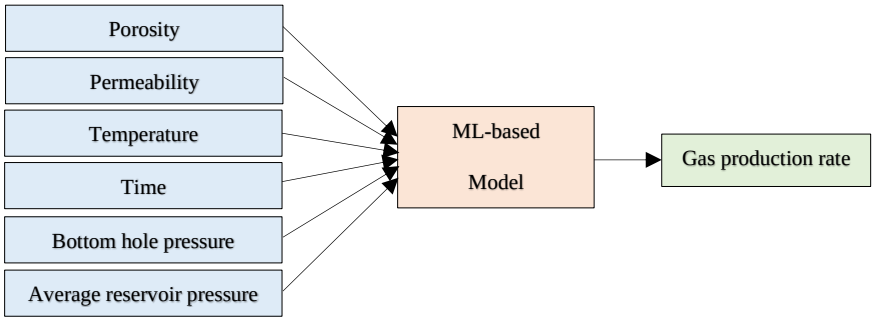

Figure 3.4: ML-based model input and output mapping for prediction of gas production rate with time.

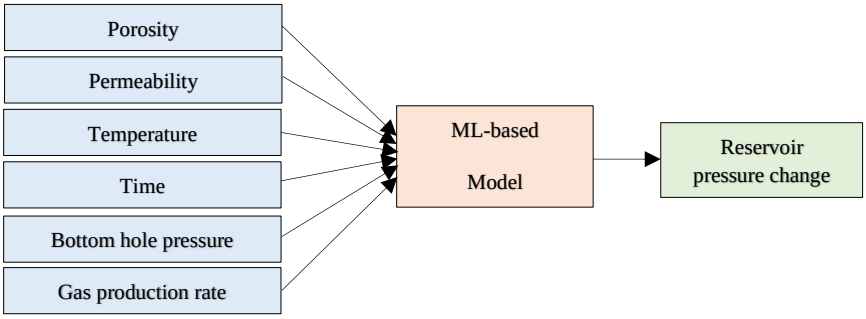

Figure 3.5: ML-based model input and output mapping for prediction of reservoir pressure change with time.

A flow diagram that shows the basic steps involved in developing the ML-based models is given in Figure 3.6.

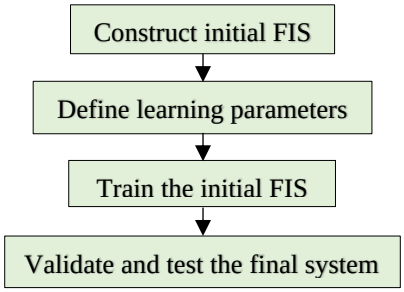

Figure 3.6: Basic workflow of the ML-based model development.

The first step involves generating an initial FIS model. Here, a Sugeno-type FIS is defined using a grid partitioning method. Five Gaussian membership functions (Gaussmf) were specified, and the initial rule-based FIS structure with the aforementioned properties was

constructed using the input/output data. Training the initial FIS involves tuning its membership function parameters such that the model fits the training data. To achieve this, ANFIS training options were configured and used to train the initial FIS. For simplicity, the ANFIS training process is explained as follows. Consider a Sugeno-type FIS with two inputs (x_1, and x_2) and one output (f). The fuzzy rules of such a FIS can be expressed as follows:

1. If x_1 is *A1* and x_2 is *B1*, then $f_1 = p_1 x_1 + q_1 x_2 + r_1$

2. If x_1 is *A2* and x_2 is *B2*, then $f_2 = p_2 x_1 + q_2 x_2 + r_2$

where *A1*, *A2*, *B1*, and *B2* represent the membership functions (MFs), and (p_i, q_i, r_i) denote set of consequent parameters. An ANFIS structure is a network consisting of five layers. The membership functions (MFs) *A1*, *A2*, *B1*, and *B2* are contained in the first layer. The degree of each input is determined by the MFs in accordance with the shape of the MFs. In the second layer, product operation is performed on the membership degrees for the purpose of computing the firing strength of each rule. In layer three, the normalised firing strength of each rule is computed. Layer four is a linear polynomial equation that can be expressed as

$$f_i = p_i x_1 + q_i x_2 + r_i \tag{3.6}$$

Layer five is the last layer that sums the outputs of rules computed in the previous layers. In this chapter, a hybrid optimization method consisting of backpropagation and least squares estimation was used in the training process. ANFIS computes the outputs of the nodes until the fourth layer where least squares estimation is used to update the consequent parameters and errors are propagated backward until the first layer where gradient descent is used to tune MFs parameters. This process keeps iterating until the training goal is achieved or the designated number of iterations is reached. Thus, the backpropagation is for the parameters associated with the input MFs, whereas the least square estimation is for the parameters associated with the output MFs.

To this end, the first intelligent model was constructed and trained using production rate data at BHP = 50, validated, and tested using production data at BHP = 100, and BHP = 150, respectively. The second intelligent model was constructed and trained using pressure drop

data at BHP = 150, validated, and tested using pressure drop data at BHP = 100, and BHP = 50, respectively. Figure 3.7 shows all the stages of the neuro-adaptive proxy model development workflow.

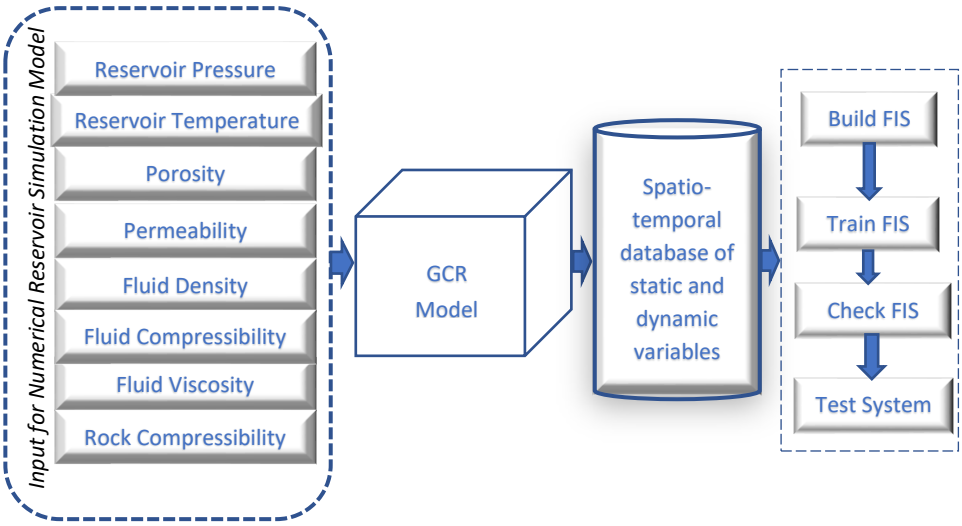

Figure 3.7: Neuro-adaptive proxy model development workflow

3.5 Results and Discussion

Two Neuro-Adaptive systems were developed to reproduce the results of numerical simulations for production rate and pressure drop with time at different BHPs. The first system was developed to regenerate the numerical simulation results of gas production for 360 days and the second system was designed to regenerate the pressure drop within the reservoir for the same time of production.

3.5.1 Results of Numerical Simulation Model

Two cases were considered in simulating the reservoir to investigate the dynamics of gas production rate and pressure drop over time.

Case 1: This case investigates the effect of BHP on gas production rate. Three simulations were carried out by setting the BHP to 50 bar, 100 bar, and 150 bar. Figure 3.8 shows the results of gas production rate with time for the three BHPs used. The results in Figure 3.8 show that at low BHP of 50 bar, the reservoir produced high volume of gas at earlier production with approximately >13,000 m^3/day. However, as the reservoir produces over time, the rate of production decreases and at the end of the simulation period (day 360),

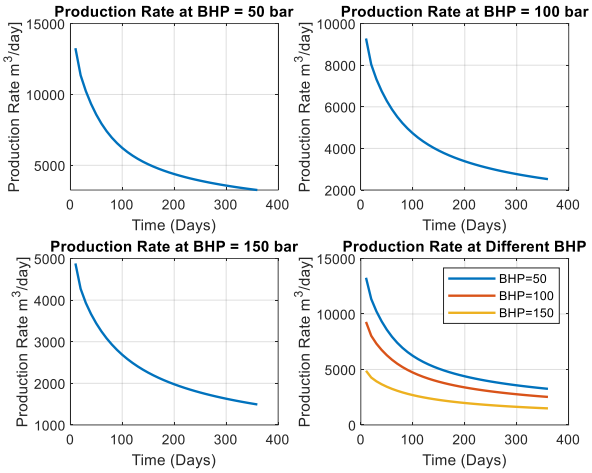

Figure 3.8: Production rate with time at different BHPs.

the reservoir produced approximately 3,500 m^3/day. At moderate BHP of 100 bar, it can be observed that the rate at which the reservoir produces has decreased. With this setting, the reservoir produced approximately > 9,000 m^3/day at the start of production and at the end of the simulation period, the reservoir produced approximately 2,500 m^3/day. At high BHP of 150 bar, it can be observed that the rate at which the reservoir produces has decreased significantly. At this BHP, the reservoir produced approximately 5,000 m^3/day at the start. However, as production continuous with time, it can be observed that the reservoir produced approximately 1,500 m^3/day at the end of the simulation period. The results in this case reveal that the ability of a well to deliver natural gas depends on the difference between the formation pressure and the BHP. It is worth mentioning that the initial formation pressure (reservoir pressure) in this simulation was set to 200 bar. Thus, when the difference between the formation pressure and the BHP is high, natural gas can be produced at a very high rate. In other words, for a natural gas reservoir to deliver at high rate, the formation pressure within the reservoir must be sufficient enough to drive the flow of the natural gas through the formation into the well bore. As the BHP pressure keeps on increasing, formation pressure decreases and this causes the flow of natural gas into the well bore to decrease, leading to low production rate [1].

Case 2: This case investigates the effect of BHP on pressure dynamic as the reservoir produces over time. In the manner, three simulations were carried out by setting the BHP to

50 bar, 100 bar, and 150 bar. Figure 3.9 shows the results of reservoir pressure change with time for the three BHPs used.

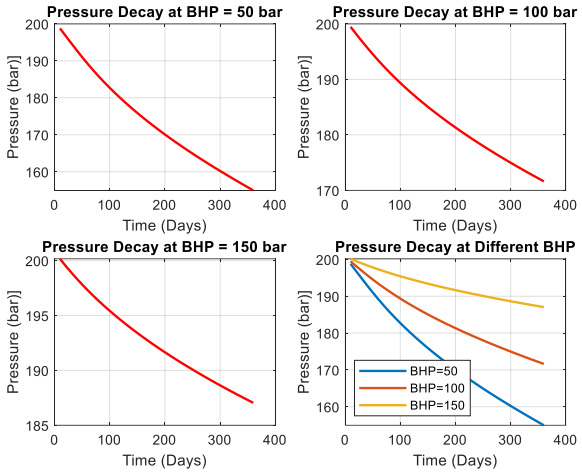

Figure 3.9: Pressure drop with time at different BHPs.

The results in Figure 3.9 show that at low BHP of 50 bar, the reservoir experiences a significant decline in pressure. It can be seen that initially, the pressure within the formation is 200 bar, however, at the end of the simulation period, the pressure within the formation has reduced to approximately 155 bar. At moderate BHP of 100 bar, it can be observed that the rate at which the reservoir pressure drops has decreased to approximately 171 bar at the end of the simulation period. At high BHP of 150 bar, it can be observed that the reservoir experiences insignificant decline in pressure. With this setting, it can be observed that the reservoir has dropped to approximately 187 bar at the end of the simulation period. The results of this case reveal that the significant decline in reservoir pressure at low BHP of 50 can be attributed back to the high volume of gas produced at low BHP as discussed in Figure 3.5. In the same vein, the insignificant decline in reservoir pressure at high BHP can be attributed to the low volume of gas produced at high BHP as earlier discussed in Figure 3.8. In summary, the results of the two simulation cases can be stated as, when the initial reservoir pressure is far greater than the BHP, the reservoir produces at high rate and consequently, experiences significant pressure drop, whereas, when the difference between the initial reservoir pressure and the BHP is not much, the reservoir produces at low rate and consequently, experiences insignificant pressure drop. For the sake of visualisation,

snapshots of pressure decay within the reservoir cells for some selected days as the reservoir produces at BHP = 50 bar are shown in Figure 3.10.

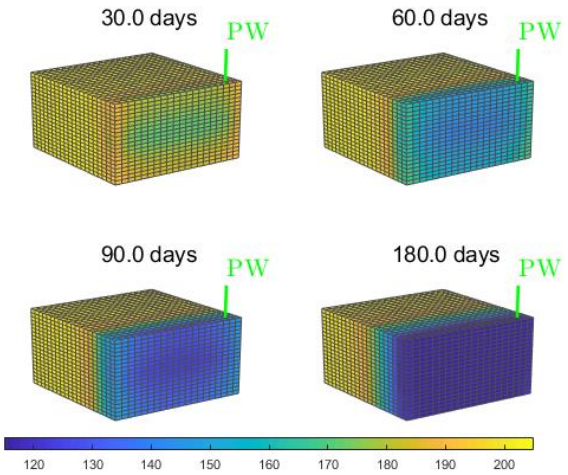

Figure 3.10: Snapshots of pressure decay with time at BHP = 50 bar.

3.5.2 Results of Neuro-Adaptive Learning System

As stated in section 3.4.2, two ML-based models were developed, one for the gas production rate data and the other for the pressure drop data. Figure 3.11 shows the results of the Neuro-Adaptive systems that were designed and trained with production rate data at BHP at BHP = 50, and with pressure drop data at BHP = 150, respectively.

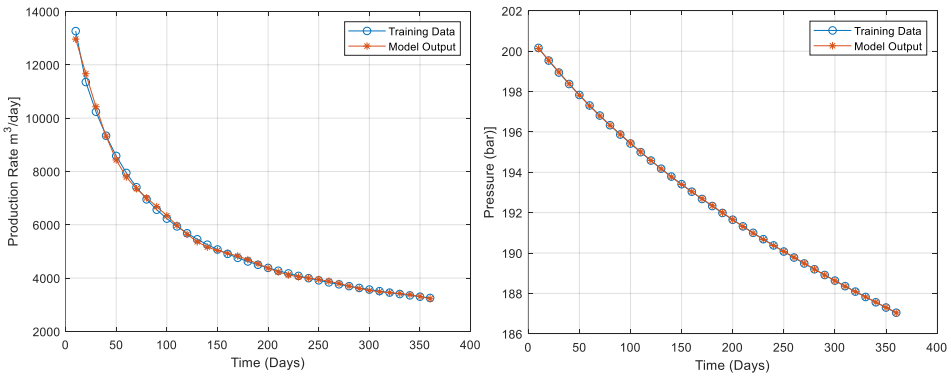

Figure 3.11: Plot of neuro-adaptive system trained with (left) production rate data at BHP = 50 bar, and (right) pressure drop data at BHP = 150 bar.

After training the two systems, they were then checked with a data set of production rate at BHP = 100, and pressure drop at BHP = 100, respectively. Figure 3.12 compares the numerical simulation result with the outputs of the two systems for production rate and pressure drop.

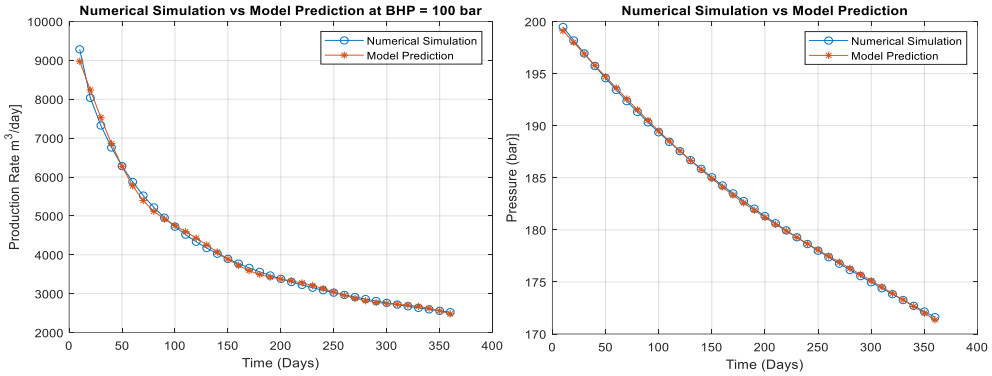

Figure 3. 12: Plot of neuro-adaptive system validated with (left) production rate data at BHP = 100 bar, and (right) pressure drop data at BHP = 100 bar.

The two Neuro-Adaptive systems were then tested by running blind simulations with data sets of production rate at BHP = 150, and pressure drop at BHP = 50, respectively. Figure 3.13 compares the numerical simulation result with the systems' predictions.

It can be seen from Figures 3.11 through 3.13 that the outputs of the Neuro-Adaptive systems for both production rate and pressure drop predictions demonstrate a very high accuracy in terms of training, checking, and testing when compared to the outputs of the numerical simulations.

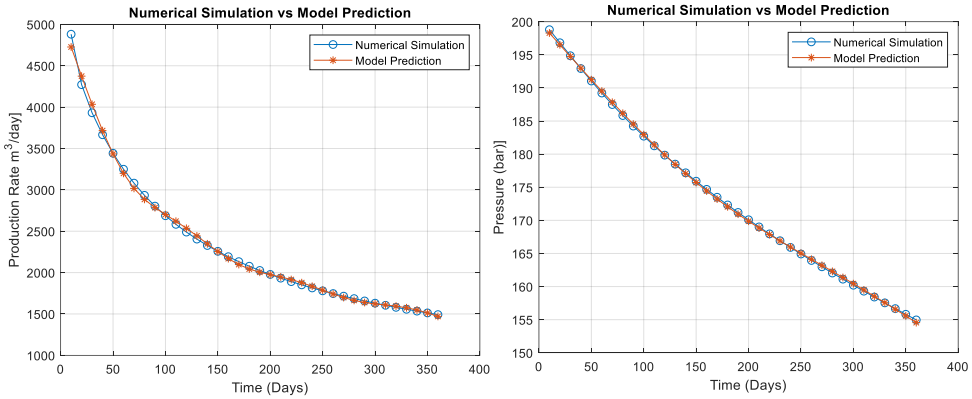

Figure 3. 13: Plot of neuro-adaptive system tested with (left) production rate data at BHP = 150 bar, and (right) pressure drop data at BHP = 50 bar.

3.5.3 Model Performance Evaluation Metrics

For the purpose of evaluating the performance of the machine learning methods, two classic statistical quantities are selected namely, coefficient of determination (R^2) and root mean squared error (RMSE). The R^2 is used to assess the goodness-of-fit and is expressed as follows [15]:

$$R^2 = \left[1 - \frac{\frac{1}{N}\sum_{t=1}^{N}(\hat{y}_t - y_t)^2}{var(y)}\right] \tag{3.7}$$

where $\hat{y}_t \; and \; y_t$ represent the predicted value and the actual value at time t, respectively. N represents the total number of data points/samples in the dataset. R^2 is a statistical tool that indicates how well a model replicates the actual outcomes. It has a range from 0 to 1 where 1 signifies perfect prediction accuracy, and 0 signifies complete failure of prediction. The closer the R^2 value to 1, the better the prediction accuracy and, the closer the R^2 value to 0, the worse the prediction accuracy. The RMSE is the square root of the average sum of deviation of predicted values and actual values. The RMSE is very sensitive to a very small or large error value and hence, has a good reflection for precision [16]. The RMSE can be computed as follows:

$$RMSE = \sqrt{\frac{1}{N} \times \sum_{t=1}^{N}(\hat{y}_t - y_t)^2} \tag{3.8}$$

The RMSE has values from 0 and above. A RMSE value of 0 signifies a perfect and error-free prediction, while the farther the value from 0, the greater the prediction error and the worse the prediction accuracy. Using the two metrics described above, performance of the proposed ML-based model on both production rate and pressure drop datasets is evaluated by comparing its results with two machine learning mainstream methods namely, ANN and SVM. For the purpose of implementing the ANN and SVM algorithms, the spatiotemporal data generated from running the numerical simulation using three different bottom hole pressure values (50 bar, 100 bar, and 150 bar) were compiled for both gas production rate and reservoir pressure change. This gives a total of 108 data samples for each model (gas production rate and reservoir pressure change). The entire data were randomly partitioned into 70 %, 15%, and 15%, for training, validation, and testing, respectively for each model.

The extracted static and dynamic variables as shown in Figures 3.3 and 3.4 are used to map the input and output for the gas production rate and reservoir pressure change, respectively. Grid search technique was used to obtain the optimal parameter values for each of the two comparison algorithms. Table 3.3 presents a detailed description of parameters used for implementing the ANN and SVM algorithms.

Table 3.3: Parameters used for implementing ANN and SVM algorithms.

Algorithm	**Hyperparameters**
ANN	Number of hidden layers: 1, number of neurons: 13, training function: Levenberg-Marquardt, activation function: Tangent sigmoid, performance function: mean squared error, initial training gain: 0.001, training gain decrease factor: 0.1, training gain increase factor: 10, maximum training gain: 1e10, minimum performance gradient: 1e-7, maximum number epochs: 500.
SVM	Kernel function: Gaussian, kernel scale: 3.7, box constrain: 0.741, Epsilon: 0.074, solver: Sequential Minimal Optimization (SMO).

The results of the comparison for gas production rate and reservoir pressure drop are presented in Tables 3.4 and 3.5, respectively.

Table 3.4: Production rate model performance evaluation.

Method	Training		Validation		Testing	
	R^2	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
ANN	0.9578	493.81	0.9575	508.78	0.9521	510.15
SVM	0.9369	648.07	0.9329	656.88	0.9290	714.34
ANFIS	0.9898	240.98	0.9869	301.79	0.9811	327.49

Table 3.5: Pressure drop model performance evaluation.

Method	Training		Validation		Testing	
	R^2	**RMSE**	**R^2**	**RMSE**	**R^2**	**RMSE**
ANN	0.9888	1.6250	0.9825	1.5501	0.9762	1.7815
SVM	0.9376	2.9001	0.9231	3.1319	0.9209	3.5350
ANFIS	0.9960	0.7577	0.9917	1.0654	0.9901	1.3321

It can be observed from Tables 3.4 and 3.5 that the proposed ML-based model exhibits better performance over ANN and SVM in terms of training, validation, and testing for both gas production rate and reservoir pressure drop, respectively. The performance of the proposed ML-based model can be attributed to its ability to integrate both neural network and fuzzy logic principles into its model development. Thus, outperforming the stand-alone ANN and SVM methods.

3.6 Conclusions

In this chapter, Neuro-Adaptive systems that mimic and regenerate results of numerical simulations for production rate and pressure drop trends in gas condensate reservoir are developed. Results from numerical simulations showed that varying the BHP can be a good idea of controlling reservoir flow, production rate, as well as pressure drop within the system. The novelty of this chapter is the incorporation of physics-based petro-physical data in the model development which distinguishes the proposed models from existing ones. Two performance metrics namely, coefficient of determination (R^2) and root mean squared error (RMSE) were used to evaluate the performance of the proposed approach. The performance of the proposed neuro-adaptive model is compared with two machine learning methods namely, ANN and SVM. The proposed neuro-adaptive approach exhibits better performance in terms of training, validation, and testing on both production rate and pressure drop datasets over ANN and SVM methods. The reason as to why the proposed neuro-adaptive method performed better can be attributed to its adaptive capability to transform stream data into a knowledge-based inference, and learn the patterns and behaviours of the data, thereby resulting to adaptive model that is capable of adapting to changing situations and generating

the desired result. The methodology and approach described for the Neuro-Adaptive model development can be used to evaluate a green or brown field. Since the method depends on data sets generated from numerical simulations, it may require the existence of a numerical simulator for development of green fields while for brown fields, this approach can be applied independently based on historical data. Given that this approach depends on the number of rules generated by a FIS, this approach can be computationally intensive on systems with low specifications. Steps have been taken to address this drawback and apply the extended approach to predict the dewpoint pressure and to examine more feasible problems of two-phase simulations and control for gas condensate reservoirs.

References

[1] L. Fan, B. Harris, and A. Jamaluddin. Understanding gas-condensate reservoirs. *Oilf. Rev.* 10(4):16-25, 2005.

[2] A. Rahimzadeh, M. Bazargan, R. Darvishi, and A. H. Mohammadi. Condensate blockage study in gas condensate reservoir. *J. Nat. Gas Sci. Eng.*, vol. 33, pp. 634–643, 2016, doi: 10.1016/j.jngse.2016.05.048.

[3] M. Arabloo, A. Shokrollahi, F. Gharagheizi, and A. H. Mohammadi. Toward a predictive model for estimating dew point pressure in gas condensate systems. *Fuel Process. Technol.*, vol. 116, pp. 317–324, 2013, doi: 10.1016/j.fuproc.2013.07.005.

[4] S. M. Dessouky and S. Gomaa. Prediction of dew point pressure in gas condensate reservoirs based on a prediction of dew point pressure in gas condensate reservoirs based on a combination of gene expression programming (GEP) and multiple regression analysis. *Pet Petro Chem Eng J,* vol. 2, no 5, pp 1 - 16, 2018.

[5] O. O. Agbaka, and J. A. Dala. Onshore condensate well production testing and analysis : A tool for optimal reservoir management. *Society of Petroleum Engineers,* vol. 18, no. 4, pp 1–15, 2016.

[6] S. D. Mohaghegh. Recent developments in application of Artificial Intelligence in petroleum engineering. *Distinguished author series*, part 2, pp 86 - 91, 2005.

[7] H. Kaydani, A. Mohebbi, and A. Hajizadeh. Dew point pressure model for gas

condensate reservoirs based on multi-gene genetic programming approach. *Appl. Soft Comput. J.*, vol. 47, pp. 168–178, 2016, doi: 10.1016/j.asoc.2016.05.049.

[8] H. Kaydani, A. Hagizadeh, and A. Mohebbi. A dew point pressure model for gas condensate reservoirs based on an artificial neural network. *Pet. Sci. Technol.*, vol. 31, no. 12, pp. 1228–1237, 2013, doi: 10.1080/10916466.2010.540616.

[9] E. M. E. M. Shokir. Dewpoint pressure model for gas condensate reservoirs based on genetic programming. *Energy and Fuels*, vol. 22, no. 5, pp. 3194–3200, 2008, doi: 10.1021/ef800225b.

[10] H. Rostami-Hosseinkhani, F. Esmaeilzadeh, and D. Mowla. Application of expert systems for accurate determination of dew-point pressure of gas condensate reservoirs. *J. Nat. Gas Sci. Eng.*, vol. 18, pp. 296–303, 2014, doi: 10.1016/j.jngse.2014.02.009.

[11] K.-A. Lie. *An introduction to reservoir simulation using MATLAB: User Guide for the Matlab Reservoir Simulation Toolbox (MRST).*1st ed. Oslo, Norway: SINTEF ICT, pp 442 - 472, 2016.

[12] F. Nematizadeh, and A. Betpolice. Pars offshore gas field development. *Eupropean offshore petroleum conference and exhibition*. pp 363 - 371, 1978.

[13] S. Esmaili, S. D. Mohaghegh, and D. Kalantari. Fast track analysis of shale numerical models. *Society of Petro. Engr*, 16:2, pp 1 - 11, 2012.

[14] S. Mohaghegh and W. V. U. Virtual-Intelligence Applications in Petroleum Engineering. *Distinguished author series*: Part 3 — Fuzzy Logic, pp. 82–87, 2000.

[15] A. C. Cameron and F. A. G. Windmeijer. An R-squared measure of goodness of fit for some common nonlinear regression models. *J. Econometrics*, vol. 77, pp. 329–342, 1996.

[16] C. Voyant, G. Notton, S. Kalogirou, and M. L. Nivet. Machine learning methods for solar radiation forecasting : *A review. Renew. Energy*, vol. 105, pp. 569–582, 2017, doi: 10.1016/j.renene.2016.12.095.

Chapter Four: Adaptive Neuro-Fuzzy Approach for Prediction of Dewpoint Pressure for Gas Condensate Reservoirs

4.1 Introduction

Dewpoint pressure is an important parameter for reservoir management and characterization. Gas condensate reservoirs experience significant reduction in productivity when initial reservoir pressure decreases below the dewpoint pressure. As such, an effective and efficient methods for prediction of this thermodynamic quantity are crucial for operational plans. In this chapter, a hybrid Artificial Intelligence model, based on adaptive neuro-fuzzy approach, for the prediction of gas condensate dewpoint pressure is presented. The proposed model combines the learning ability of Artificial Neural Network (ANN) and the capability of rule-based Fuzzy Inference System (FIS). Firstly, fuzzy subtractive clustering technique is applied to a set of measured input/output data to identify an initial system based on extracted extract groupings from the dataset. The generated system is then trained using Adaptive Neuro-Fuzzy Inference System (ANFIS) after which model validation and testing were performed. The performance of the proposed model is compared with previously reported methods including empirical correlations, Soave-Redlich-Kwong, and Peng-Robinson equations of states. The results show that our proposed model outperforms the previous and existing methods with up to 99% accuracy and with the least Root Mean Square Error (RMSE) of 2.188 for some selected fluid samples. The better performance of the proposed neuro-adaptive method over the previously reported methods can be attributed to its adaptive capability of transforming stream data into a knowledge-based inference, and learning the patterns and behaviours of the data, thereby resulting to adaptive model that is capable of adapting to changing situations and generating the desired result even if it encounters a situation that is different from the history enclosed in the historical datasets. The approach presented in chapter three has similar adaptive capability with the one presented in this chapter.

4.2 Existing Methods of Dewpoint Pressure Prediction

Design of intelligent techniques for development of gas condensate fields and accurate prediction of thermodynamic properties for gas condensate reservoirs has been and is still

one of the active areas of research in petroleum industry [1] – [7]. Conventionally, dewpoint pressure of a gas condensate fluid can be estimated experimentally using Constant Composition Expansion (CCE) or Constant Volume Depletion (CVD) methods, however, experimental determination of dewpoint pressure at reservoir temperature for gas condensate reservoir is costly and time intensive [8], [9]. Other methods include Equilibrium Ratio (K-Value) approach and the use of Equation of States (EOS) such as the Soave-Redlich-Kwong (SRK) and Peng-Robinson (PR) EOSs [10]. Equilibrium Ratio (K-Value) approach is a trial-and-error method. Even though, this method can estimate dewpoint pressure to some reasonable extent, nevertheless, they are only suitable for estimating dewpoint pressure at low and moderate pressures. SRK and PR EOSs have been widely used in the petroleum industry for predicting thermodynamic properties. However, these traditional methods have some drawbacks which include improper characterization of the heptane hydrocarbon plus fraction.

With recent interest and enthusiasm in petroleum industry toward intelligent fields, smart wells and analysis of data for process optimization and decision-making, AI techniques are making firm steps toward becoming more preferred in the widespread of the oil and gas industry [11] – [13]. However, some of the studies conducted to address the dewpoint pressure estimation have been developed based on the common assumptions that when a computational intelligence model is developed and trained using a huge amount of historical datasets, the resulting model will, to some extent, handle similar situations during their operations [14] – [16].

In many real-world operations, these assumptions may tend to fail as in most cases, the amount of previously available data used in developing and training the model may not be sufficient to represent the underlying system. Furthermore, the environment as well as the system and its constituents may change over time. In such a situation, it becomes difficult for a model that is developed and trained using only iterative algorithms to cope with such changes and process the data efficiently.

Therefore, the aim of this chapter is to address the aforementioned challenges and limitations of both conventional and existing techniques by demonstrating the potentiality of a hybrid

AI-based approach that combines the power of FIS that transforms stream data into a knowledge-based inference, and the learning ability of ANN that learns the patterns and behaviours of the data, thereby resulting to a model that is capable of adapting to changing situations and generating the desired output even if it encounters a situation that is different from the history enclosed in the historical datasets.

4.3 Adaptive Neuro-Fuzzy Model Development for Dewpoint Pressure Prediction

Generation of input/output data set is the starting point and the most important step when developing an adaptive Neuro-Fuzzy model. In this chapter, 168 published fluid samples drawn from measured, experimental, and filed data for gas condensate systems in [10], [15] - [23] were used in developing the intelligent model. Each set of data has 1 output (dewpoint pressure) and 13 input parameters. Table 4.1 shows the range of the input/output variables used in developing the model.

Table 4.1: Range of variables used for model development.

Component	Minimum	Maximum	Average
Methane - C_1	0.3344	0.9668	0.7280
Ethane - C_2	0.0037	1.092	0.0879
Propane - C_3	0.0018	0.601	0.0510
Butane - C_4	0.0018	0.203	0.0315
Pentane - C_5	0.0006	0.085	0.0172
Hexane - C_6	0.0005	0.0846	0.0132
Heptane plus - C_{7+}	0.0019	0.138	0.0389
Nitrogen - N_2	0.0000	0.4322	0.0137
Carbon Dioxide - CO_2	0.0000	0.124	0.0205
Hydrogen Sulphide - H_2S	0.000	0.1255	0.0054
SG C_{7+}	0.0747	0.959	0.7740
MW C_{7+}	100.0	253	142.9
Temperature (^{0}F)	40.00	337	217.4
Dewpoint pressure (psia)	1110	11830	3792.26

The proposed model defines gas condensate dewpoint pressure as a function of fluid composition, reservoir temperature, specific gravity and molecular weight of heptane plus fraction as follows:

$$P_d = f(Z_i, \gamma C_{7+}, MW_{C7+}, T) \tag{4.1}$$

where P_d stands for dewpoint pressure in psia, Z_i represents fluid composition in mole fraction, γC_{7+} denotes specific gravity of heptane plus, MW_{C7+} and T represent molecular weight of heptane plus and reservoir temperature, respectively.

4.3.1 Data Clustering

Clustering of data set forms the basis of various classification and system modelling algorithms. The purpose of data clustering is to extract natural formations or groupings of data from a data set, thereby resulting to a concise representation of the behaviour of the system under study. Even though clustering is, in most cases associated to classification problems, here, we use the subtractive clustering estimation technique proposed by [24] to address the problem of generating many rules which result to a computationally intensive complex model.

Having a collection of n data points $(x_1, x_2, x_3, \dots, x_n)$ in M-dimensional space, we consider each data point as a candidate for cluster centre and define a density measure at data point x_i as

$$P_i = \sum_{j=1}^{n} exp\left[-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right] \tag{4.2}$$

where r_a is a positive constant that defines the neighbourhood of data points. Thus, a density measure of a data point is a function of its separation to all other data points. After computing the density measure of each data point, the first cluster centre is then selected. This cluster centre is the data point that has the highest density measure. If we consider the position of the first cluster centre as x_1^* and its density measure as P_1^*, then the density measure of each data point x_i is then revised by the following equation

$$P_i = P_i - P_1^* exp\left[\frac{\|x_i - x_1^*\|^2}{(r_b/2)^2}\right] \quad (4.3)$$

where r_b is a positive constant. Here, an amount of density measure from each data point is subtracted as a function of its separation from the first cluster centre. Data points closer to the first cluster x_1^* will certainly have reduced density measure and hence, are expected to be picked as the next cluster centre. After revising the density measure of all data points in accordance with (4.3), the data point that has the highest remaining density measure is then picked as the next cluster centre, x_2^*. This procedure is repeated until a number of cluster centres required to develop the model are obtained.

4.3.2 Adaptive Neuro-Fuzzy System Modelling

Application of ANFIS to predict properties of great importance for engineering calculations in chemical and petroleum engineering has been reported in the literature [25] - [29]. In this present study, cluster centres are generated first, after which a Sugeno-type FIS is then designed and trained using ANFIS. Figure 4.1 is an illustration of an ANFIS system which we will use for simplicity to describe the procedure used in developing our proposed model.

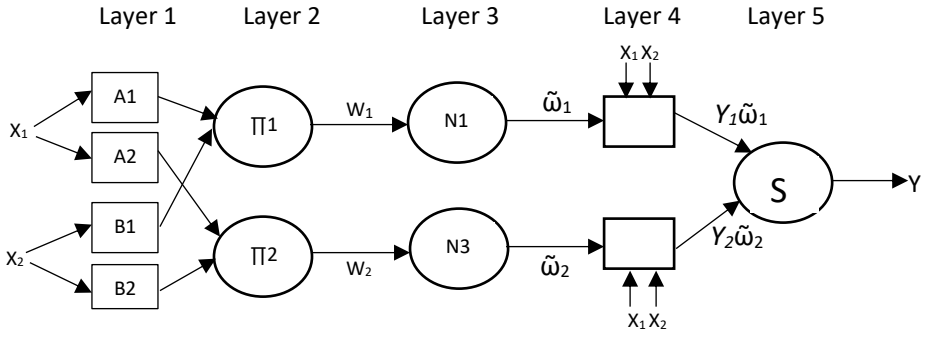

Figure 4.1: ANFIS structure of two input variables and one output variable

Figure 4.1 is a description of ANFIS model with two inputs (x1, and x2) and one output (Y). Each input variable has two Membership Functions (A1, A2, and B1, B2) and the system has two fuzzy rules (∏1 and ∏2). Similar to the rules illustrated in chapter three, the fuzzy rules of such a system can be expressed as follows:

Rule 1:

If X_1 is A_1 and X_2 is B_1, then $Y_1 = p_1X_1 + q_1X_2 + r_1$

Rule 2:

If X_1 is A_2 and X_2 is B_2, then $Y_2 = p_2X_1 + q_2X_2 + r_2$.

Each layer of Figure 4.1 can briefly be explained in accordance to [25] as follows:

Layer 1

This layer is known as the Fuzzification layer. In this layer, each node i is an adaptive node having a node function as

$$O_{1,i} = \mu_{Ai}(x), \quad \text{for } i = 1,2 \tag{4.4}$$

where x stands for the input to node i and μ_{Ai} denotes the MF that describes the linguistic label such as "low" or "high". Therefore, each node in this layer denotes a membership based on the following Gaussian function.

$$\mu_{Ai}(x) = exp\left[\left\{\left(\frac{x-c_i}{a_i}\right)^2\right\}^{b_i}\right] \tag{4.5}$$

where parameters $a, b,$ and c represent the premise parameters that are responsible for changing the shape of the MF.

Layer 2

This layer is known as the rule layer. In this layer, every node is a fixed node with a label $\prod$, and its output is the product of all the incoming signals given as

$$O_{2,i} = w_i = \mu_{Ai}(x_1) * \mu_{Bi}(x_2), \quad \text{for } i = 1,2 \tag{4.6}$$

here, the input variable x_1 with a linguistic value A_i is multiplied with the input variable x_2 with a linguistic value B_i and the resulting output w_i is the firing weight of the rule $\prod 1$. Therefore, each node in this layer represents the firing weight of a rule.

Layer 3

This layer is referred to as the normalization layer and every node is a fixed node with a label N. Every node in this layer is responsible for normalizing the weight of all rules based on the following equation:

$$O_{3,i} = \tilde{\omega}_i = \frac{w_i}{w_1 + w_2}, \quad \text{for } i = 1,2 \tag{4.7}$$

where $\tilde{\omega}_i$ stands for the node's output that computes the ration of the *i*th rule's firing weight to the sum of all rule's firing weights. Therefore, the outputs of this layer are known as normalized firing weights.

Layer 4

This layer is referred to as the defuzzification layer. In this layer, every node i is an adaptive node having a node function as:

$$O_{4,i} = \tilde{\omega}_{iY_i} = \tilde{\omega}_i(p_i x_1 + q_i x_2 + r_i) \tag{4.8}$$

where $\tilde{\omega}_i$ is a normalized firing weight from layer 3 and (p_i, q_i, r_i) is a set of parameters of this node. In this layer, parameters are referred to as consequent parameters.

Layer 5

This layer is referred to as the output layer. It is characterized by a single node with a label S and is responsible for computing the overall output of the system which is the summation of all incoming signals as follows:

$$O_{5,i} = \sum_i \tilde{\omega}_i \, Y_i = \frac{\sum_i w_i Y_i}{\sum_i w_i} \tag{4.9}$$

In the developed model, the ANFIS was trained using a hybrid (feedforward and feed backward) learning algorithms where the MF parameters are identified and tuned. In the forward feed, least squares method was used to identify the consequent parameters on layer 4 while in the backward feed, errors were propagated backwardly and the premise parameters were updated using gradient descent method. Table 4.2 shows the properties of the developed hybrid intelligent model.

Table 4.2: Specifications of the developed intelligent model.

Parameter	**Description/value**
Number of input variables	13
Number of output variable	1
Fuzzy Inference System	Sugeno-type
Input Membership Function	Gaussian
Output Membership Function	Linear
Range of influence of cluster centers	0.275
Number of cluster centres	62
Number of fuzzy rules	62
Optimization method	Hybrid
Number of epochs for training	20
Initial step size	0.1
Step size decrease rate	0.081

4.4 Results and Discussion

The entire data were randomly partitioned into 70 %, 15%, and 15%, for training, validation, and testing, respectively. Hold-out validation technique was used for validating and testing the proposed model. The model's training performance for prediction of dewpoint pressures based on Gaussian-based MF is shown in Figure 4.2.

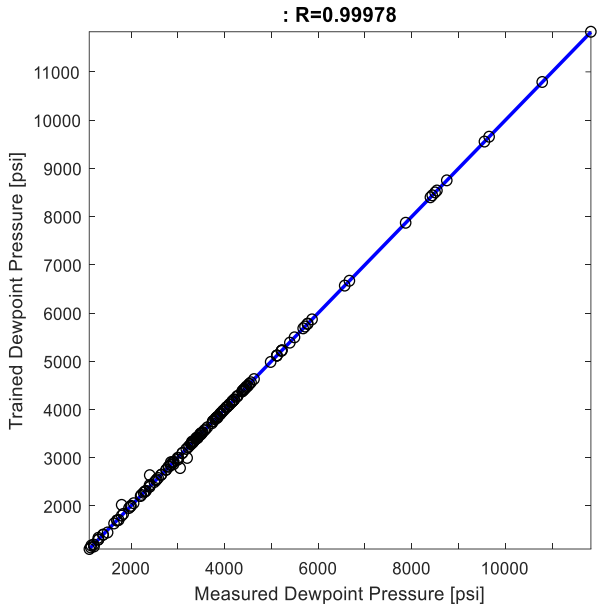

Figure 4.2: Comparison of trained and measured dewpoint pressure.

As presented in Figure 4.2, the ANFIS model training was performed and the result of the training showed a correlation coefficient of 0.99978.

The trained model is then validated by the validating data set. The primary reason for performing this validation is to test the generalization capability of the proposed model at each epoch and to ensure that the model does not overfit the training data. Figure 4.3 shows the result of validating the model and the result depicts the reliability and effectiveness of the model by tracking the given input/output validating data set very well with a correlation coefficient of 0.99967.

Next, the model was tested using the testing data set to ascertain its predictability performance. Figure 4.4 shows the result of this test where the model performs very well in predicting the dewpoint pressures with a correlation coefficient of 0.9997.

To evaluate the performance of the proposed intelligent model for prediction of gas condensate dewpoint pressure, comparison of the model's result was made with the results of other related studies using a statistical quantity known as the Root Mean Squared Error (RMSE).

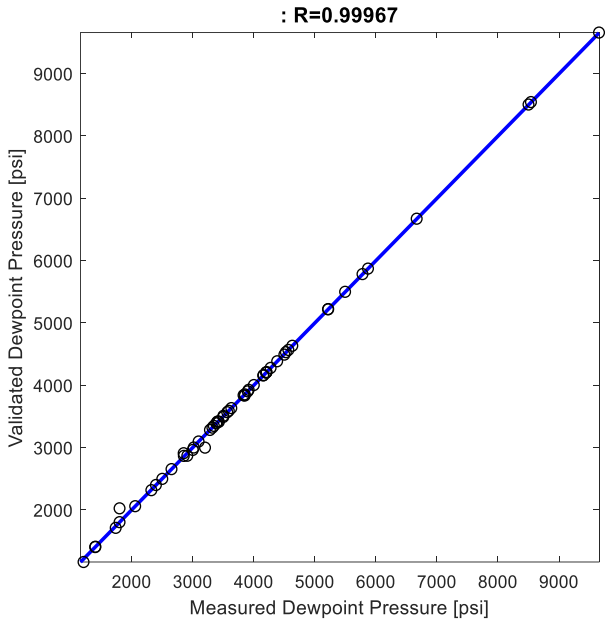

Figure 4.3: Comparison of validated and measured dewpoint pressures

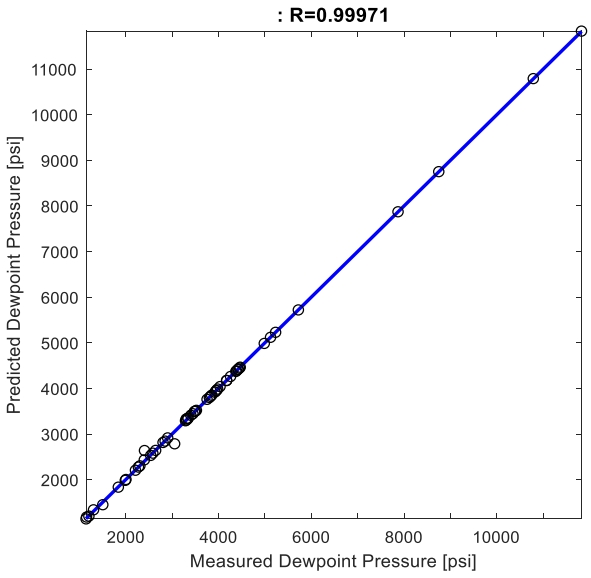

Figure 4.4: Comparison of predicted and measure dewpoint pressure.

The statistical quantity RMSE is defined by the following equations [29]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Pd_i - \widehat{Pd}_i\right)^2} \tag{4.10}$$

where $Pd_i =$ measured dewpoint pressure of fluid sample i

$\widehat{Pd}_i$ = predicted dewpoint pressure of fluid sample i

n = number of fluid samples.

From the testing data set, measured dewpoint pressures of 14 fluid samples [10] were used to compare the performance of the proposed model with seven methods reported in the literature [10]. From Table 4.3, the first three approaches reported by Elsharkawy, Organic & Golding, Nemeth & Kennedy are empirical correlation methods, whereas the remaining four approaches are Soave-Redlich-Kwong, and Peng-Robinson equations of states. Table 4.3 compares the RMSE of this study with the seven methods reported in the literature.

Table 4.3: Comparative analysis of results of dewpoint prediction

Measured Dewpoint	**Elsharkawy**	**Organic & Golding**	**Nemeth & Kennedy**	**SRK EOS 1**	**SRK EOS 2**	**PR EOS 1**	**PR EOS 2**	**This Study**
3095	3220	2650	2823	3104	3077	2833	2815	3087
3337	6541	2750	4144	6808	4739	5947	4909	3337
2651	2655	2620	2792	2911	2902	2682	2675	2651
11830	12081	7800	5545	12914	12689	11738	11426	11831
3345	3345	3850	2507	3342	2790	3206	2603	3345
8750	8643	7055	9136	9850	9599	9172	9138	8750
5780	5808	3468	5492	5602	6500	5096	5427	5780
5229	5349	4270	5062	5147	5141	5330	5307	5229
4203	4270	3674	4054	4152	4157	4256	4258	4202
4173	4170	3740	3983	4153	4152	4243	4242	4174
5219	5179	4054	4932	4823	4830	4709	4712	5219
4172	4173	3808	4061	4089	4098	3855	3863	4172
4160	4139	3520	4063	4055	4105	4159	4235	4160
7871	6734	4800	7178	6208	6762	6183	6374	7871
RMSE	**913.0**	**1646**	**1728**	**1117**	**638.8**	**877.6**	**662.9**	**2.188**

It can be observed from the comparative analysis of the dewpoint pressure prediction performances (Table 4.3) that the developed intelligent model of this chapter proves more accurate compared to other methods.

4.5 Conclusions

In this chapter, a hybrid intelligent model based on adaptive neuro-fuzzy inference system has been developed for improved prediction of dewpoint pressure for gas condensate reservoirs. The performance of the developed model has been examined in comparison with different studies conducted for dewpoint pressure prediction. Results of the proposed model developed in this chapter show a very good performance in terms of accuracy with the lowest root mean squared error. The proposed intelligent model can be considered a vital tool of help to production engineers, managers, and modellers for a quick and more informed decision as regards gas condensate field development plans that can meet operational targets. The main advantage of the approach utilized in this chapter is its ability to describe the overall complex behaviour of a nonlinear dynamic parameter (dewpoint pressure) of a system by combining the system's sub-linear parameters (input variables).

References

[1] M. M. Ghiasi, A. Shahdi, P. Barati, and M. Arabloo. Robust modeling approach for estimation of compressibility factor in retrograde gas condensate systems. *Ind. Eng. Chem. Res.*, vol. 53, no. 32, pp. 12872–12887, 2014, doi: 10.1021/ie404269b.

[2] H. A. Zamani, S. Rafiee-Taghanaki, M. Karimi, M. Arabloo, and A. Dadashi. Implementing ANFIS for prediction of reservoir oil solution gas-oil ratio. *J. Nat. Gas Sci. Eng.*, vol. 25, pp. 325–334, 2015, doi: 10.1016/j.jngse.2015.04.008.

[3] S. Rafiee-Taghanaki, M. Arabloo, A. Chamkalani, M. Amani, M. H. Zargari, and M. R. Adelzadeh. Implementation of SVM framework to estimate PVT properties of reservoir oil. *Fluid Phase Equilib.*, vol. 346, pp. 25–32, 2013, doi: 10.1016/j.fluid.2013.02.012.

[4] J. Ye. Artificial neural network modeling of methanol production from syngas. *Pet. Sci. Technol.*, vol. 37, no. 6, pp. 629–632, 2019, doi: 10.1080/10916466.2018.1560321.

[5] S. M. Bassir and M. Madani. A new model for predicting asphaltene precipitation of diluted crude oil by implementing LSSVM-CSA algorithm. *Pet. Sci. Technol.*, vol.

37, no. 22, pp. 2252–2259, 2019, doi: 10.1080/10916466.2019.1632896.

[6] M. Sun. Prediction of viscosity of branched alkanes using gene expression programing. *Pet. Sci. Technol.*, vol. 36, no. 23, pp. 2049–2056, 2018, doi: 10.1080/10916466.2018.1531029.

[7] H. Saghafi and M. Arabloo. Development of genetic programming (GP) models for gas condensate compressibility factor determination below dew point pressure. *J. Pet. Sci. Eng.*, vol. 171, no. 6, pp. 890–904, 2018, doi: 10.1016/j.petrol.2018.08.020.

[8] M. Arabloo and S. Rafiee-Taghanaki. SVM modeling of the constant volume depletion (CVD) behavior of gas condensate reservoirs. *J. Nat. Gas Sci. Eng.*, vol. 21, pp. 1148–1155, 2014, doi: 10.1016/j.jngse.2014.11.002.

[9] A. Najafi-Marghmaleki, A. Tatar, A. Barati-Harooni, M. Arabloo, S. Rafiee-Taghanaki, and A. H. Mohammadi. Reliable modeling of constant volume depletion (CVD) behaviors in gas condensate reservoirs. *Fuel*, vol. 231, no. 5, pp. 146–156, 2018, doi: 10.1016/j.fuel.2018.04.130.

[10] A. M. Elsharkawy. Predicting the dew point pressure for gas condensate reservoirs: Empirical models and equations of state. *Fluid Phase Equilib.*, vol. 193, no. 1–2, pp. 147–165, 2002, doi: 10.1016/S0378-3812(01)00724-5.

[11] P. Bakhtiari Manesh, S. shahryari, and A. Bemani. Utilization of Grid partitioning based Fuzzy inference system approach as a novel method to estimate solubility of hydrocarbons in carbon dioxide. *Pet. Sci. Technol.*, vol. 36, no. 5, pp. 350–355, 2018, doi: 10.1080/10916466.2018.1425718.

[12] L. Cheng, Y. Wang, Z. Zhang, and W. Gao. Prediction of gas composition obtained from steam-gasification of residual oil using an Artificial Neural Network (ANN) model. *Pet. Sci. Technol.*, vol. 37, no. 6, pp. 641–644, 2019, doi: 10.1080/10916466.2018.1560324.

[13] F. Ajorkaran and A. Cheshmeh Sefidi. Application of RBF-ANN in prediction of natural gas density in different operational conditions. *Pet. Sci. Technol.*, vol. 37, no. 22, pp. 2246–2251, 2019, doi: 10.1080/10916466.2018.1476888.

[14] E. M. E. M. Shokir. Dewpoint pressure model for gas condensate reservoirs based on genetic programming. *Soc. Pet. Eng. - SPE Gas Technol. Symp.*, vol. 1, no. 4, pp. 178–190, 2008.

[15] H. Kaydani, A. Mohebbi, and A. Hajizadeh. Dew point pressure model for gas condensate reservoirs based on multi-gene genetic programming approach. *Appl. Soft Comput. J.*, vol. 47, pp. 168–178, 2016, doi: 10.1016/j.asoc.2016.05.049.

[16] E. An, S. Gomaa, and D. Sm. Prediction of Dew Point Pressure in Gas Condensate

Reservoirs Based On a Combination of Gene Expression Programming (GEP) and Multiple Regression Analysis. *Petro. & Petrochem. Eng.*, vol. 2 no. 5, pp 1–16, 2018.

[17] D. Galatro and F. Marín-Cordero. Considerations for the dew point calculation in rich natural gas. *J. Nat. Gas Sci. Eng.*, vol. 18, pp. 112–119, 2014, doi: 10.1016/j.jngse.2014.02.002.

[18] H. Kaydani, A. Hagizadeh, and A. Mohebbi. A dew point pressure model for gas condensate reservoirs based on an artificial neural network. *Pet. Sci. Technol.*, vol. 31, no. 12, pp. 1228–1237, 2013, doi: 10.1080/10916466.2010.540616.

[19] K. Nasrifar, O. Bolland, and M. Moshfeghian. Predicting natural gas dew points from 15 equations of state. *Energy and Fuels*, vol. 19, no. 2, pp. 561–572, 2005, doi: 10.1021/ef0498465.

[20] S. M. J. Majidi, A. Shokrollahi, M. Arabloo, R. Mahdikhani-Soleymanloo, and M. Masihi. Evolving an accurate model based on machine learning approach for prediction of dew-point pressure in gas condensate reservoirs. *Chem. Eng. Res. Des.*, vol. 92, no. 5, pp. 891–902, 2014, doi: 10.1016/j.cherd.2013.08.014.

[21] F. Nematizadeh and A. Betpolice. Pars offshore gas field development. *Soc. Pet. Eng. - SPE Eur. Pet. Conf.* pp. 363–370, *1978*, doi: 10.2523/8127-ms.

[22] D. E. Kenyon and G. A. Behie. Third Spe Comparative Solution Project: Gas Cycling of Retrograde Condensate Reservoirs. *JPT, J. Pet. Technol.*, vol. 39, no. 8, pp. 981–997, 1987, doi: 10.2118/12278-PA.

[23] Kelkar, M. *Natural gas production engineering*. 1st ed. Tulsa: PennWell, pp 1–92, 2008.

[24] Chiu, S. L. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, vol. 2, no. 3, pp 267–78, 1994, doi:10.3233/IFS-1994-2306.

[25] M. M. Ghiasi, M. Arabloo, A. H. Mohammadi, and T. Barghi. Application of ANFIS soft computing technique in modeling the CO2 capture with MEA, DEA, and TEA aqueous solutions. *Int. J. Greenh. Gas Control*, vol. 49, pp. 47–54, 2016, doi: 10.1016/j.ijggc.2016.02.015.

[26] H. Saghafi and M. Arabloo. Estimation of carbon dioxide equilibrium adsorption isotherms using adaptive neuro-fuzzy inference systems (ANFIS) and regression models. *Environ. Prog. Sustain. Energy*, vol. 36, no. 5, pp. 1374–1382, 2017, doi: 10.1002/ep.12581.

[27] S. Sharifian, M. Madadkhani, M. Rahimi, M. Mir, and A. Baghban. QSPR based ANFIS model for predicting standard molar chemical exergy of organic materials.

Pet. Sci. Technol., vol. 37, no. 21, pp. 2174–2181, 2019, doi: 10.1080/10916466.2018.1496100.

[28] M. Chen, J. Sasanipour, S. A. Kiaian Mousavy, E. Khajeh, and M. Kamyab. Estimating phase behavior of the asphaltene precipitation by GA-ANFIS approach. *Pet. Sci. Technol.*, vol. 36, no. 19, pp. 1582–1588, 2018, doi: 10.1080/10916466.2018.1493503.

[29] A. Fayazi, M. Arabloo, and A. H. Mohammadi. Efficient estimation of natural gas compressibility factor using a rigorous method. *J. Nat. Gas Sci. and Eng.*, 16:8–17. doi:10.1016/j.jngse.2013.10.004.

Chapter Five: Dynamic Mode Decomposition for Modelling of Pressure Dynamics in Underground Natural Gas Storage Reservoirs

5.1 Introduction

The process of storing natural gas in underground geological formations involves applying pressure to force the gas into and out of the porous and permeable reservoirs. In response to gas extraction/withdrawal and storage/injection, the reservoir compresses and expands as a major consequences of fluid pore pressure variations. The major challenge associated with this type of energy systems is learning the pore pressure variations within the grid as fluid is being injected and/or withdrawn. As such, it is essential to identify a realistic model that accounts for the pore pressure variations at any point in time. In this chapter, a data-driven technique called Dynamic Mode Decomposition (DMD) to investigate the pressure dynamics of an underground hydrocarbon reservoir model in relation to natural gas injection/storage is presented. For demonstration purpose, a hydrocarbon reservoir model using a benchmark data of the first Society of Petroleum Engineers (SPE1) Comparative Solution Project was implement first. The reservoir model was simulated using varied specifications of porosity and permeability distributions within the reservoir layers. Experiment was then performed using different number of modes on the pressure field datasets generated from the numerical simulations. It is shown that DMD is capable of reconstructing the reservoir pressure field and approximating the average reservoir pressure change over time with the best result obtained using 15 number of modes. Given that underground natural gas storage reservoirs are already developed depleted reservoirs, it is recommended that DMD could serve as a reliable tool for fast evaluation of pressure dynamics of underground natural gas storage given its low complexity and insignificant loss of prediction accuracy.

5.2 Existing Methods and Problem Statement

Given that energy storage plays a vital contribution to energy security in the present energy systems, the need for storing energy in bulk to strike a balance between supply and demand is essential [1], [2]. Underground natural gas storage (UNGS) plays a crucial role in ensuring that any excess gas delivered during the hot season (summer months) is available to meet the

raised demand of the cold season (winter months). It requires sufficient time to explore, produce, and transport natural gas and the produced quantity that arrives its destination is sometimes not needed instantly, thus, to sustain that produced quantity, it is stored via injection into underground storage facilities [3]. Industrial and residential operations are responsible for a large proportion of natural gas consumption in developed countries, and thus are targets of energy efficiency plans [4]. According to [5], two major reasons that necessitate natural gas storage are: first, meeting seasonal consumption requirements and second, as a security against unanticipated supply interruptions. Most facilities for storing natural gas are constructed in deep underground formations, mostly including depleted gas and oil reservoirs, aquifer reservoirs, and salt caverns [6]. Currently, UNGS in depleted reservoirs has the highest percentage of natural gas storage globally, aquifers represent the second-largest UNGS, and salt caverns, particularly caverns left after mining activities in salt formations, represent the third-largest UNGS [7]. Among the major types of UNGS facilities, depleted hydrocarbon reservoirs are considered economical due to their known geological characteristics compared to aquifers and salt caverns [8]. In comparison to large tanks which are limited to relatively low pressures when used to store natural gas on the surface, UNGSs have the advantage of accommodating natural gas at high pressures, thus, providing a larger and safer storage capacity [9]. Storage of natural gas in depleted hydrocarbon reservoirs has been considered a strategic practice to meet the increasing seasonal demand of natural gas in different parts of the world. In 1915, the first successful underground storage of natural gas in a depleted hydrocarbon reservoir occurred in Ontario, Canada. Since then, several of such facilities have been built in North America, Canada, Europe, Asia-Oceanic, Middle East, Argentina, and other parts of the world [10]. Underground natural gas storage may be regarded as a long-term confinement of natural gas within geological formations. Hence, two of the most fundamental static parameters of an underground storage facility are porosity (its ability to keep natural gas for future use) and permeability (its ability to transmit gas into and out of the formation). According to [3], the process of storing natural gas in geological formations involves applying pressure to force the gas into and out of the porous and permeable reservoir. As natural gas is being injected into the geological formation, pressure is being built up within the formation, thereby making the geological formation becoming a

type of pressurized natural gas container. As described in [10], the proportion of natural gas that can be stored and withdrawn during a normal cycle of a depleted reservoir is referred to as working gas while the proportion of the natural gas that must remain in the reservoir to maintain pressure within the reservoir is called base gas (also known as cushion gas).

As described in [11], during the normal operation of depleted storage reservoir, base gas remains permanent within the reservoir to maintain pressure required to drive the natural gas into the well. Working gas which is referred to as the volume of natural gas available for withdrawal during the normal functioning of the storage reservoir, is injected during storage and this process causes the pressure in the reservoir to increase. This makes the pressure within the storage reservoir becomes high. Converting a depleted hydrocarbon (oil or natural gas) reservoir from production facility to storage facility takes advantage of using an already developed reservoir with existing wells, pipeline connections, as well as extraction and distribution equipments that were leftover when the reservoir was productive [12]. Having these equipments in place cuts down the cost of converting depleted oil or natural gas reservoirs into storage facilities and thereby making depleted reservoirs, on average, the easiest and cheapest to develop, operate and maintain compared to other types of UNGS facilities such as caverns and aquifers [13].

Much of the practices in UNGS technology are borrowed from oil and gas reservoir engineering [10]. Conventionally, numerical simulation is the standard in the oil and gas industry for predicting dynamic parameters (such as production rate, injection rate, pressure dynamics), and quantifying uncertainties in reservoirs. Characterizing reservoirs and managing workflows using this approach involves numerous simulations for optimizing production, enhancing oil/gas recovery, and history matching [14]. Flow of fluid in the reservoir is governed by complex nonlinear partial differential equations (PDEs), which in practice, are spatially discretized into a high-dimensional set of nonlinear ordinary differential equations (ODEs) [15]. For consistent representation of flow dynamics and subsurface geology, grid blocks in very large numbers are required and cumbersome algorithms are employed for their spatiotemporal solutions. It requires thousands of simulations even with advanced algorithms to achieve optimal solutions when solving with nonlinear constraints [16]. The complexity in physics associated with reservoir fluid flow

and the multiscale nature of the rock and fluid properties present challenges in achieving better predictive models. In comparison to numerical simulation, in which model set up is laborious and implementation is time consuming, a data-driven model that not only scales down computational complexity but also, offers accuracy in a short time without compromising results would be of great benefit [17]. The challenges faced by physical models to adequately describe fluid/rock interactions and accurately capture flow dynamics on complex geometries prompted a remarkable interest in seeking for alternative solutions via data-driven models [18], [19].

Recently, application of machine learning and data-driven techniques to track and predict dynamic parameters have been receiving attentions in the energy sector [18] - [20]. In recent years, a large body of research called Dynamic Mode Decomposition (DMD) has emerged around modal decomposition and machine learning methods. DMD originated as a new promising tool in the fluid dynamics community to discover spatiotemporal meaningful structures from high-dimensional fluids data [21]. The evolving success of DMD arises from the fact that it is a data-driven and equation-free technique that is capable of discovering spatiotemporal meaningful patterns that may be used for diagnosis, control, state estimation and future-state prediction of complex dynamical systems [22].

Given that the UNGS industry borrowed much of its knowledge from oil and gas reservoir engineering [10], the present chapter focuses on applying DMD to investigate the pressure dynamics in an underground hydrocarbon reservoir model that operates under natural gas injection that mimics the process of storing/injecting working gas in depleted storage reservoirs.

The remaining sections of this chapter are organized as follows: section three presents materials and methods with focus on first, a brief description of the reservoir model used as a reference to the proposed DMD method. The section then describes the DMD algorithm utilized in this study. Section four presents computational results of both the reference model and the data-driven model. Section five ends the chapter with conclusions on strengths and weaknesses of the proposed DMD method, as well as next steps toward future work.

5.3 The Reference Model and DMD Algorithm

5.3.1 Numerical Reservoir Simulation: The Reference Model

The first step towards achieving the aim of this study is the development of a numerical reservoir simulation model to simulate the process of storing natural gas (working gas) in a depleted reservoir. The reservoir model which serves as a reference model is developed using the available dynamic and static data of the first SPE Comparative Solution Project [23]. The first SPE Comparative Solution Project was organized by [23] and is a description of a depletion problem with natural gas injection into a 10×10×3 reservoir model with an injector and a producer in diagonally opposite corners. The natural gas injection well was completed in layer 1 and is located at grid point (1, 1) while the producing well was completed in layer 3 and is located at grid point (10, 10). In this chapter, experiments were performed using three different specifications of rock properties (porosity and permeability) distributed within the reservoir layers to observe the effect of rock properties on the reservoir pressure as natural gas is being injected. Initially, the reservoir is undersaturated with an initial formation pressure of 330 bar and zero free gas (Sg = 0.0). As mentioned earlier, the geological structure of the reservoir model is a three-dimensional formation that consists of 10 × 10 grid blocks in the x-y dimension and 3 layers in the z dimension. Thus, the numerical reservoir simulation model consists of 10 ×10 ×3 = 300 grid blocks (cells). For each of the three rock property distributions, the reservoir was simulated for 1200 days in 120 time-steps. This yielded a 300×120 matrix that contains 36,000 grid cells in the spatiotemporal database for each of the rock property distributions. Each grid cell in the spatiotemporal database contains some dynamics of the reservoir's parameters in a given run and given time step. For the purpose of this chapter, the pressure field datasets of the gas injection reservoir model for the three cases were retrieved and utilized by the proposed data-driven technique. For the sake of simplicity, the simulation process neglects the effect of interface pressure, and as for inlet boundary condition, a bottom-hole pressure is assigned to the injection well.

5.3.2 Dynamic Mode Decomposition (DMD)

The proposed method utilizes the capability of DMD as a data-driven modal reduction

technique to provide accurate reconstruction of reservoir pressure field and approximation of average pressure change from the numerical simulation data or production/experimental data. From the DMD perspective, the pressure field data collected from the numerical simulation of the reservoir model in this study were arranged in two snapshot matrices as follows:

$$Y = \begin{bmatrix} | & | & | \\ y_1 & \cdots & y_{m-1} \\ | & | & | \end{bmatrix}, \tag{5.1a}$$

$$Y' = \begin{bmatrix} | & | & | \\ y_2 & \cdots & y_m \\ | & | & | \end{bmatrix}. \tag{5.1b}$$

These two matrices have large number of rows than columns, that is, $n \gg m$ and consist of the states of the system and their columns were captured in equal-spaced time, with a time step Δt. Each $Y_i = Y(i\Delta t)$ is a vector with gridlock components *c*, as such, $Y, Y' \in \mathbb{R}^{c\times(m-1)}$. Using data from the numerical simulations, DMD algorithm attempts to compute the eigendecomposition of a linear operator matrix A defined as

$$AY \approx Y'. \tag{5.2}$$

This implies that

$$A = Y'Y^{\dagger} \tag{5.3}$$

here, $Y^{\dagger}$ stands for the Moore-Penrose pseudo-inverse of *Y*. To get an estimate of matrix *A*, Singular Value Decomposition (SVD) has to be computed on the full-snapshot matrix *Y* as follows

$$Y = USV^{*} \tag{5.4}$$

where $U \in \mathbb{R}^{n\times r}, S \in \mathbb{R}^{r\times r}, V \in \mathbb{R}^{m\times r}$ and $r \leq m$ stands for the rank of the data matrix *Y*. The columns of *U* are referred to as POD modes, and they satisfy $U^{*} \cdot U = I$. In the same manner, columns of *V* are orthonormal, and satisfy $V^{*} \cdot V = I$. The diagonal of *S* contains the

singular values of matrix Y. The full matrix A can be acquired by solving the pseudo-inverse of Y as follows

$$A = Y'VS^{-1}U^{*}. \tag{5.5}$$

Here, the eigenvectors of A represent the dynamic modes (ϕ_i) with distinctive eigenvalues (λ_i). However, in many applications, the matrix A will be a high-dimensional matrix and computing its eigendecomposition becomes computationally expensive. For this reason, a reduced matrix $\tilde{A}$ which have the same nonzero eigenvalues as A is computed as

$$\tilde{A} = U^{*}AU = U^{*}Y'\mathrm{V}S^{-1}. \tag{5.6}$$

Thus, the spectral decomposition of $\tilde{A}$ can be computed as

$$\tilde{A}W = W\Lambda. \tag{5.7}$$

Here, the DMD eigenvalues are the elements of the diagonal matrix Λ and the eigenvectors of $\tilde{A}$ are represented by the columns of W. The high-dimensional DMD modes ϕ_D can then be reconstructed by using W of the reduced system and the snapshot matrix Y' as follows

$$\phi_D = Y'\mathrm{V}S^{-1}W. \tag{5.8}$$

Here, the eigenvectors of the high-dimensional matrix A are these DMD modes which correspond to the eigenvalues in Λ. Finally, average reservoir pressure at each time step can be approximated from the reconstructed pressure field by taking the average of all pressures in all grid cells at a particular time step.

5.4 Computational Results and Discussion

5.4.1 Numerical Simulation of Natural Gas Injection/Storage

As mentioned in section 5.3.1, three cases were considered using three different distributions of rock properties (porosity and permeability). Case one considers a reservoir model with uniform porosity and uniform permeability distributions within the grid cells. Case two considers a reservoir model with heterogeneous porosity and uniform permeability distributions within the grid cells. Case three considers a reservoir model with both porosity

and permeability heterogeneously distributed within the grid cells. Tables 5.1 to 5.3 show the different rock property distributions used for the three cases considered in this study.

Table 5.1: Uniform porosity and uniform permeability distribution.

Reservoir Layer	**Porosity (%)**	**Permeability (mD)**
Layer 1	0.2	250
Layer 2	0.2	250
Layer 3	0.2	250

Table 5.2: Heterogeneous porosity and uniform permeability distribution.

Reservoir Layer	**Porosity (%)**	**Permeability (mD)**
Layer 1	0.3	250
Layer 2	0.1	250
Layer 3	0.2	250

Table 5.3: Heterogeneous porosity and heterogeneous permeability distribution.

Reservoir Layer	**Porosity (%)**	**Permeability (mD)**
Layer 1	0.3	500
Layer 2	0.1	50
Layer 3	0.2	200

For each of the three cases shown in Tables 5.1 to 5.3, the reservoir model was simulated for a period of 1200 days of injecting/storing natural gas into the formation while the production well remained closed. Results of the reservoir's pressure field that form the spatiotemporal database were compiled and saved. Plots of the reservoir model depicting the injection well (I), production well (P), porosity, and permeability distributions within the reservoir layers for the three cases are shown in Figures 5.1 to 5.3.

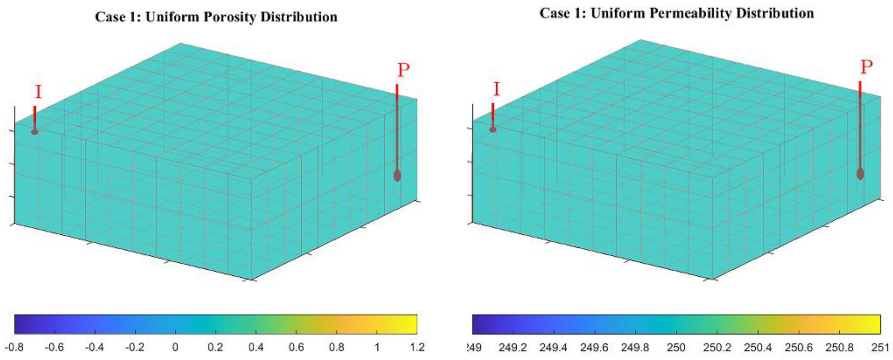

Figure 5.1: Porosity and permeability uniformly distributed within reservoir layers.

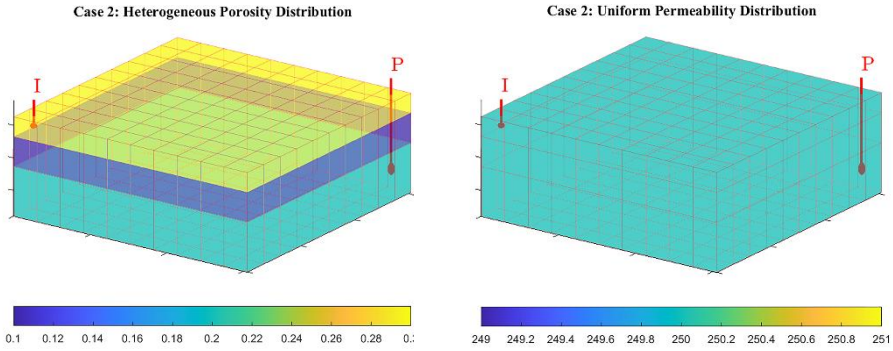

Figure 5.2: Heterogeneous porosity and uniform permeability distribution within reservoir layers.

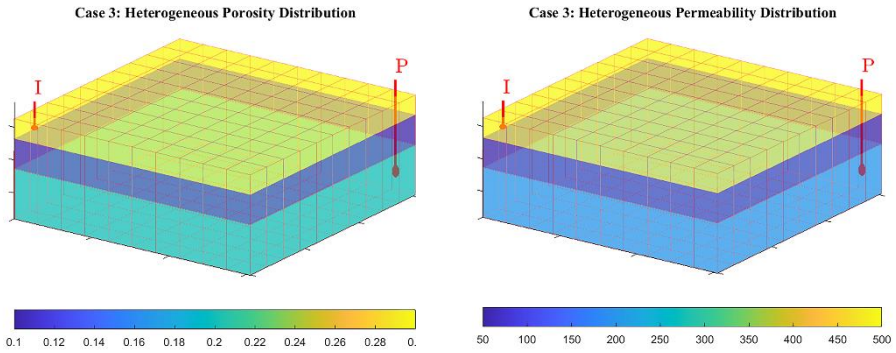

Figure 5.3: Both porosity and permeability heterogeneously distributed within reservoir layers.

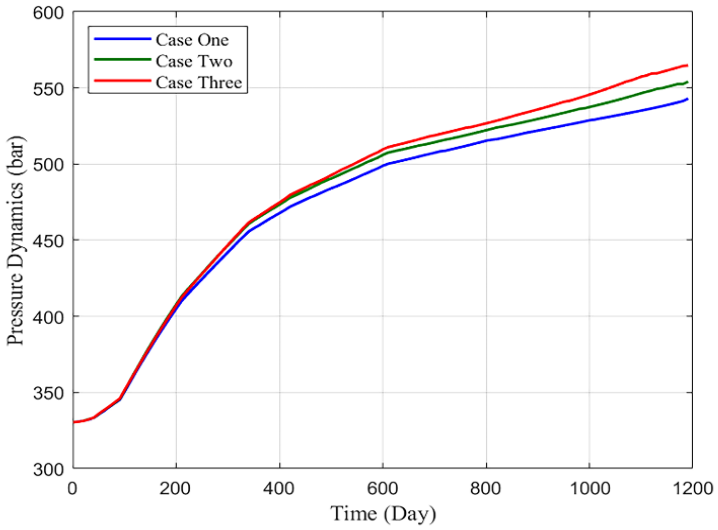

Figure 5.4: Average reservoir pressure change with time for the three cases of rock property distributions.

Results of average reservoir pressure change over time for the three case are presented in Figure 5.4. As shown in Figure 5.4, it can be observed that for all the three cases of rock property distributions, as natural gas is being injected/stored into the reservoir model, the formation pressure keeps rising within the reservoir, thereby making the reservoir a sort of pressurized gas storage as it expands to accommodate the injected natural gas. As regards the effect of rock properties on the pressure dynamics, it can be observed in particular that in case one (uniform porosity and permeability), the formation pressure rose up from 330 bar to approximately 543 bar at the end of the simulation period. In case two (heterogeneous porosity and uniform permeability), it can be observed that the formation pressure rose up from 330 bar to approximately 554 bar at the end of the simulation period.

In case three (heterogeneous porosity and permeability), it can be observed that the formation pressure rose up from 330 bar to approximately 565 bar at the end of the simulation period. The formation pressure within the reservoir can be associated with the amount of natural gas the reservoir accommodated at the end of the injection period [11]. As such, it can be seen that when both porosity and permeability are heterogeneously distributed within the reservoir layers, much gas is accommodated within the reservoir than when either or both of the rock properties are uniformly distributed within the reservoir layers. This is not strange as regards the roles of porosity and permeability in natural gas reservoirs. Reservoir porosity which is

the measure of the void spaces determines the ability of the reservoir to hold/store fluid whereas, reservoir permeability which is the measure of the connection between pores determines the ability of the reservoir to transmit fluid. Thus, the size of the void spaces and the connections between them play a vital role in allowing natural gas reservoirs to store and transmit fluid. Results in Figure 5.4 implies that when porosity and permeability are heterogeneously distributed within the reservoir layers, the reservoir store and transmits much natural gas than when either or both porosity and permeability are uniformly distributed within the reservoir layers. To gain an insight of how the pressure evolves within the reservoir grid cells as natural gas is being injected, plot of the reservoir's pore pressure evolution for some selected days from the time that storage starts for case two simulation is shown in Figure 5.5.

5.4.2 Model Performance Evaluation Metrics

Before presenting the experimental results of the proposed DMD model, a brief presentation of the metrics used to evaluate the performance of the model is given. For the purpose of evaluating the proposed DMD model performance, two classic statistical quantities are selected namely, mean squared error (MSE) and root mean squared error (RMSE). The MSE is a tool that measures the degree of change by revealing the average of the quadratic sum of the variation of predicted value and observed value [24] and is expressed as:

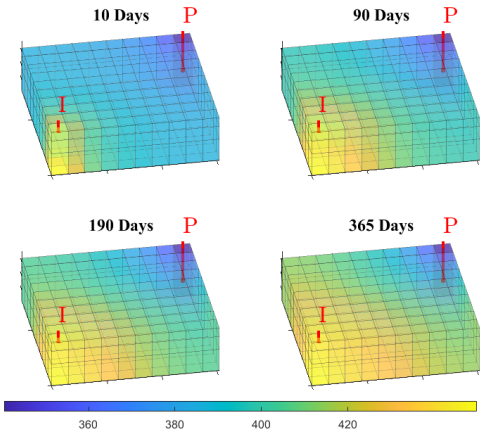

Figure 5.5: Reservoir pore pressure variations for some selected days of natural gas storage.

$$MSE = \frac{1}{N} \times \sum_{t=1}^{N} (\hat{y}_t - y_t)^2. \tag{5.9}$$

where $\hat{y}_t \; and \; y_t$ represent the predicted value and the observed value at time t, respectively. N represents the total number of data points/samples in the dataset. The RMSE has been explained earlier in section 3.5.3.

Both MSE and RMSE have values from 0 and above. A MSE or RMSE value of 0 signifies a perfect and error-free prediction, while the farther the value from 0, the greater the prediction error and the worse the prediction accuracy.

5.4.3 DMD Reconstruction of Pressure Field and Prediction of Average Reservoir Pressure Dynamics

Next step involves applying DMD on the spatiotemporal pressure field datasets to reconstruct the reservoir pressure field and approximate the average reservoir pressure change over time. Recall that the reservoir was simulate using three different cases of rock properties as described in section 5.4.1. Results of the numerical simulation gave three sets of data for the three cases. To validate the effectiveness and ability of DMD algorithm in reconstructing the reservoir pressure field and approximating average reservoir pressure, experiments were performed on the three different data sets using varied number of modes. Results of the experiment are presented in Table 5.4.

Table 5.4: DMD performance evaluation on reservoir pressure field data using varied number of modes.

	Reservoir's Pressure Field Reconstruction		**Average Reservoir Pressure Approximation**	
	MSE	**RMSE**	**MSE**	**RMSE**
		5 Modes		
Case 1	57.0504	7.5532	50.5379	7.1090
Case 2	5.1220	2.2632	4.6852	2.1645
Case 3	48.9100	6.9936	41.5799	6.4482
		10 Modes		
Case 1	36.5079	6.0422	33.2166	5.7634

Case 2	2.4903	1.5781	2.4190	1.5553
Case 3	10.1393	3.1842	7.8096	2.7946
		15 Modes		
Case 1	4.7534	2.1802	3.4812	1.8658
Case 2	1.2766	1.1299	1.2361	1.1118
Case 3	4.0760	2.0189	2.6932	1.6411
		20 Modes		
Case 1	6.8079	2.6092	5.1770	2.2753
Case 2	1.3176	1.1479	1.2794	1.1311
Case 3	8.7691	2.9613	5.7869	2.4056
		25 Modes		
Case 1	14.2286	3.7721	13.9880	3.7400
Case 2	1.5841	1.2586	1.4503	1.2043
Case 3	9.1536	3.0255	6.3095	2.5119
		30 Modes		
Case 1	36.2391	6.0199	35.9150	5.9929
Case 2	2.0836	1.4435	2.2500	1.3180
Case 3	55.7081	7.4638	41.5124	6.4430

It can be seen from Table 5.4 that on general note, pressure field reconstruction and average reservoir pressure approximation errors for all the three cases decrease as the number of modes increases from 5 to 15. However, from modes 20 to 30, the errors increase for all the three cases. In other words, accuracy of all the three cases improves as the number of modes increases from 5 to 15. Experimental results shown in Table 5.4 show that DMD algorithm performed well with least MSE and RMSE with 15 modes for both reservoir pressure field reconstruction and average reservoir pressure approximation on all the three data sets. In particular, it can be observed that with 15 number of modes, DMD recorded the least MSE and RMSE on case two (heterogeneous porosity and uniform permeability) dataset. For the sake of visualization, plot that compares average reservoir pressure field as reference data

with DMD best outputs with 15 number of modes for all the three cases is presented in Figure 5.6.

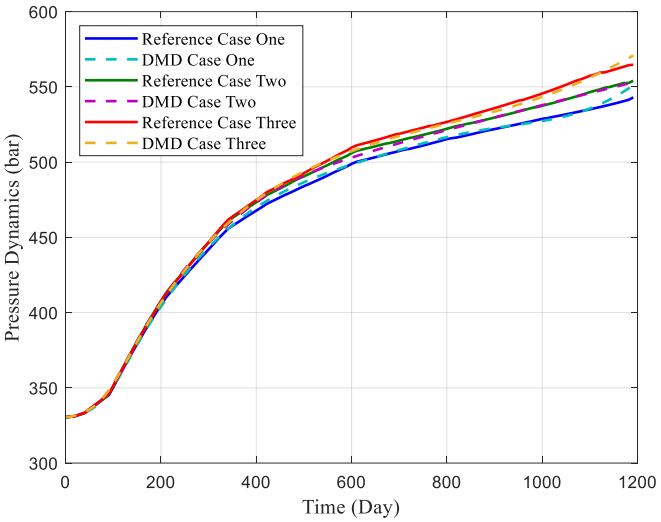

Figure 5.6: Comparison of reference data with DMD outputs with 15 modes for all the three cases.

From Figure 5.6, it can be observed that DMD output for case two has a better match with its reference data from day one to the end of the simulation period compared to DMD outputs for case one and three. Figure 5.7 Compares the reservoir's pore pressure variations versus DMD outputs for some selected days are shown in Figure 5.5

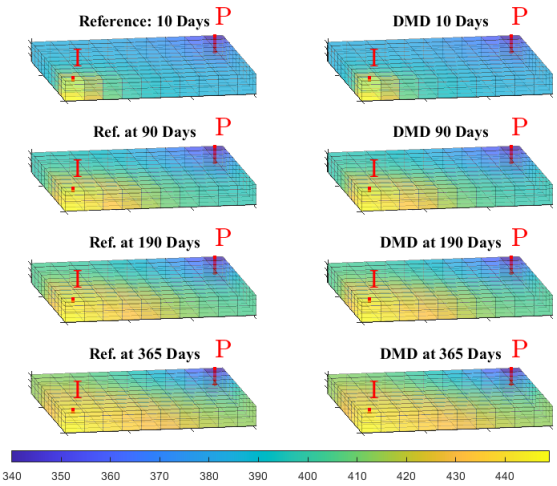

Figure 5.7: Comparison of reservoir pore pressure variations versus DMD outputs for the selected days shown in Figure 5.5.

It can be observed from Figures 5.7 that the pressure evolutions within the reservoir grid cells generated from numerical simulation for some selected days are in good agreement with the ones generated by DMD.

Plots that compare MSE and RMSE for pressure field reconstruction and average reservoir pressure approximation on all the three datasets are presented in Figures 5.8 and 5.9, respectively.

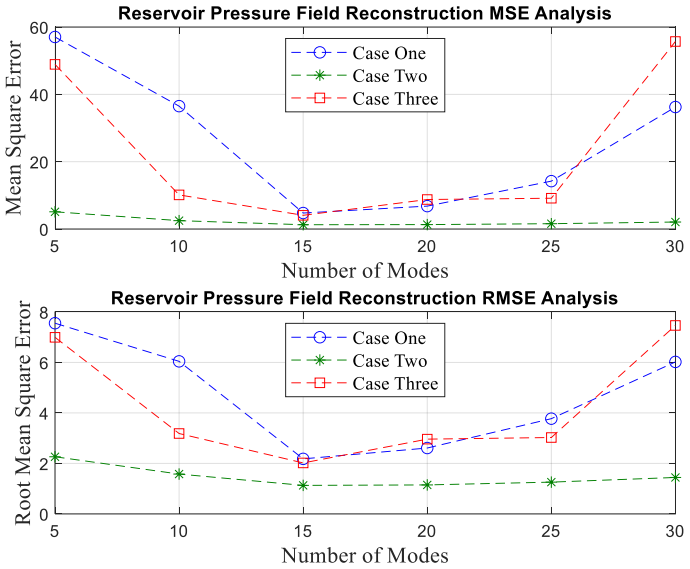

Figure 5.8: Comparison of DMD reservoir pressure field reconstruction errors.

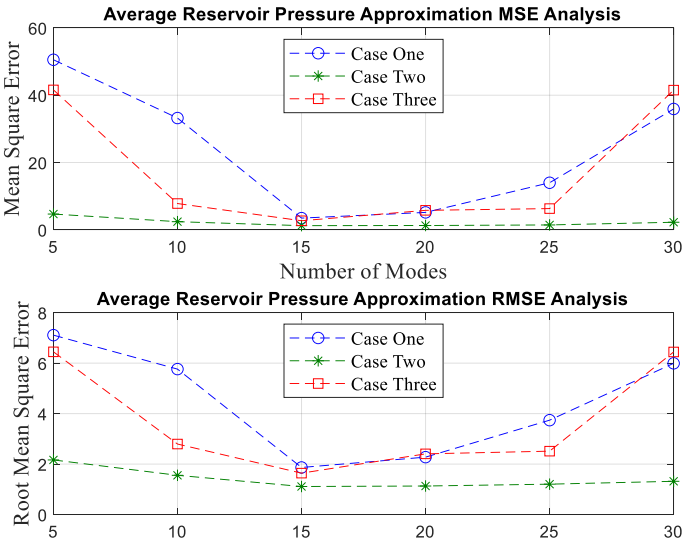

Figure 5.9: Comparison of DMD average reservoir pressure approximation errors.

From Figure 5.8, it can be observed that with 15 modes DMD is able to reconstruct the reservoir pressure field with the least MSE and RMSE of 1.2766 and 1.1299, respectively on case two dataset. In the same manner, it can be observed from Figure 5.9 that with 15 modes, DMD is able to approximate the average reservoir pressure with the least MSE and RMSE of 1.2361 and 1.1118, respectively on case two dataset.

5.5 Conclusion

In this chapter, a data-driven method that decomposes and reconstructs a pressure field of a depleted reservoir model that mimics the behaviour of an underground natural gas storage is proposed. The reservoir model was first simulated numerically using three different cases/distributions of rock properties as natural gas is being injected into the formation. Spatiotemporal datasets that represent the reservoir's pressure field for the three cases simulated were then retrieved. Results of applying DMD on the pressure field data show that the proposed data-driven technique is capable of reconstructing the reservoir pressure field and approximating the average reservoir pressure over time with 15 number of modes for all the three cases considered in this study. In particular, DMD recorded the least mean square error (MSE) and root mean square error (RMSE) for both pressure field reconstruction and average reservoir pressure approximation with 15 modes on case two (heterogeneous porosity and uniform permeability) dataset. Considering that storing natural gas in depleted reservoirs is more economical, the proposed technique can be used as a reliable tool for monitoring the pressure dynamics of underground natural gas storage in depleted reservoirs. As strengths and weaknesses are better understood from this method, it is worth mentioning that in standard DMD, one of the noticeable drawbacks is computation of singular values performed on the high-dimensional full-state matrix which is computationally burdensome. Also, in standard DMD, better results are obtained when the number of columns of the full-state matrix is less than the number of rows. On the contrary, when the number of columns of the full-state matrix is greater than the number of rows, part of the columns that carries spatial information might be lost in the process of implementing the method. Even though, this might not be a problem when dealing with overdetermined systems, nevertheless, this is a drawback when dealing with underdetermined systems as information regarding low level

dynamics in spatiotemporal data might not be analyzed. The next chapter will focus on developing a method that will address these notable drawbacks.

References

[1] Y. Jiang, L. Kang, and Y. Liu. Optimal configuration of battery energy storage system with multiple types of batteries based on supply-demand characteristics. *Energy*, vol. 206, p. 118093, 2020, doi: 10.1016/j.energy.2020.118093.

[2] Y. Wang, R. Das, G. Putrus, and R. Kotter. Economic evaluation of photovoltaic and energy storage technologies for future domestic energy systems – A case study of the UK. *Energy*, vol. 203, p. 117826, 2020, doi: 10.1016/j.energy.2020.117826.

[3] H. Plaat. Underground gas storage: Why and how. *Geol. Soc. Spec. Publ.*, vol. 313, pp. 25–37, 2009, doi: 10.1144/SP313.4.

[4] International Energy Agency. "Gas analysis and forecast to 2014", Exec. Summ., 1-8, 2019.

[5] P. Teatini, N. Castelletto, M. Ferronato, G. Gambolati, C. Janna, E. Cairo, D. Marzorati, D. Colombo, A. Ferretti, A. Bagliani, & F. Bottazzi. Geomechanical response to seasonal gas storage in depleted reservoirs : A case study in the Po River basin, Italy. *J. Geophysical research*, vol. 116, pp. 1–21, 2011. https://doi.org/10.1029/2010JF001793.

[6] W. Liu, X. Zhang, J. Fan, Y. Li, and L. Wang. Evaluation of potential for salt cavern gas storage and integration of brine extraction: Cavern utilization, Yangtze River Delta region. *Nat. Resour. Res.*, vol. 29, no. 5, pp. 3275–3290, 2020, doi: 10.1007/s11053-020-09640-4.

[7] Natural Gas Supply Association (NGSA), Storage of natural gas. Retrieved July 27, 2020 from http://naturalgas.org/naturalgas/storage/.

[8] M. Healey. Underground gas storage in Cheshire – the costain experience. White paper, pp. 1-9, 2008.

[9] H. Peng *et al.*. Computed tomography analysis on cyclic fatigue and damage

properties of rock salt under gas pressure. *Int. J. Fatigue*, vol. 134, no. October 2019, p. 105523, 2020, doi: 10.1016/j.ijfatigue.2020.105523.

[10] F. Verga. What's conventional and what's special in a reservoir study for underground gas storage. *Energies*, vol. 11, no. 5, 2018, doi: 10.3390/en11051245.

[11] D. J. Evans. An appraisal of underground gas storage technologies and incidents , for the development of risk assessment methodology. *Br. Geol. Surv. Open Rep. OR/07/023*, p. 287, 2007, doi: 10.1016/j.jfca.2010.03.028.

[12] U. M. Mgbaja and N. Enwere. Reservoir characterization, simulation & estimation of storage capacity of depleted reservoirs in Niger Delta for underground natural gas storage. *Soc. Pet. Eng. - Niger. Annu. Int. Conf. Exhib. 2017*, pp. 1373–1387, 2017.

[13] F. Verga. What's conventional and what's special in a reservoir study for underground gas storage. *Energies*, vol. 11, no. 5, 2018, doi: 10.3390/en11051245.

[14] Y. Zhang, J. He, C. Yang, J. Xie, R. Fitzmorris, and X. H. Wen. A physics-based data-driven model for history matching, prediction, and characterization of unconventional reservoirs. *SPE J.*, vol. 23, no. 4, pp. 1105–1125, 2018, doi: 10.2118/191126-pa.

[15] K. Aziz and A. Settari. *Petroleum reservoir simulation.* 1st ed. Essex: Applied Sci. Pub. Lmt., vol. 1, pp. 42 - 47, 1979.

[16] Z. Liu, F. Forouzanfar, and Y. Zhao. Comparison of SQP and AL algorithms for deterministic constrained production optimization of hydrocarbon reservoirs. *J. Pet. Sci. Eng.*, vol. 171, no. October 2017, pp. 542–557, 2018, doi: 10.1016/j.petrol.2018.06.063.

[17] M. Hasan and P. Goharshenasan. A new hybrid method based on Fuzzy Logic for maximum power point tracking of photovoltaic systems. *Energy Reports*, vol. 6, pp. 1619–1632, 2020, doi: 10.1016/j.egyr.2020.06.010.

[18] A. Ali and L. Guo, "Neuro-adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir," *IFAC-*

PapersOnLine, vol. 52, no. 29, pp. 122–127, 2019, doi: 10.1016/j.ifacol.2019.12.632.

[19] A. Ali and L. Guo, "Adaptive neuro-fuzzy approach for prediction of dewpoint pressure for gas condensate reservoirs," *Pet. Sci. Technol.*, vol. 38, no. 9, pp. 673–681, 2020, doi: 10.1080/10916466.2020.1769655.

[20] K. Aliyuda, J. Howell, A. Hartley, and A. Ali, Stratigraphic controls on hydrocarbon recovery in clastic reservoirs of the Norwegian continental shelf, *Pet. Geosci.*, pp. petgeo2019-133, 2020, doi: 10.1144/petgeo2019-133.

[21] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.*, vol. 656, pp. 5–28, 2010, doi: 10.1017/S0022112010001217.

[22] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz. On dynamic mode decomposition: Theory and applications. *J. Comput. Dyn.*, vol. 1, no. 2, pp. 391–421, 2014, doi: 10.3934/jcd.2014.1.391.

[23] A. S. Odeh. Comparison of solutions to a three-dimensional black-oil reservoir simulation problem. *J Pet Technol*, vol. 33, no. 01. pp. 13 - 25, 1981, doi: 10.2118/9723-PA.

[24] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecasting*, vol. 22, pp. 679–688, 2006, doi: 10.1016/j.ijforecast.2006.03.001.

Chapter Six: Dynamic Mode Learning Method for Analysis of Pressure and Fluid Phase Dynamics in Multiphase Reservoirs

6.1 Introduction

Secondary recovery involves injecting water or gas into reservoirs to maintain or boost the pressure and sustain production levels at viable rates. Accurate tracking of pressure dynamics and fluid phase distribution as reservoirs produce under secondary production is one of the challenging tasks in reservoir simulation. In this chapter, a novel method called Dynamic Mode Learning (DML) that aims to provide an efficient alternative approach for learning and decomposing flow dynamics in multiphase reservoir models that produce under secondary recovery is proposed. Existing algorithms suffer from expensive computational demand and unpredictability of reproducing results. The new DML model is developed in the form of a learning system by first, constructing a simple, fast and efficient learning system that extracts important features from original full-state data and places them in a low-dimensional representation as extracted features. The extracted features are then used to reduce the original high-dimensional data after which dynamic modes are computed on the reduced data. In this new technique, once the parameters of the learning system are configured and the original high-dimensional data is reduced, all other computations are performed on the reduced data and the final results are reproduced efficiently at each run. The performance of the proposed DML method is illustrated on pressure field and fluid phase distribution data sets generated from direct numerical simulations of two multiphase reservoir models. Experimental results performed on four different data sets reveal that the proposed DML method exhibits better and effective performance over standard and compressed dynamic mode decomposition (DMD) mainstream algorithms.

6.2 Existing Methods and Problem Statement

As the world population keeps on rising and industrial and residential activities increase, so also the demand for energy. To meet the rapidly increasing global energy demand, various forms of energy sources need to employ available and new techniques to meet this increasing demand. So far, renewable energy (solar, wind, hydro, tidal, geothermal and biomass) has

accounted for up to 35% of the overall energy supply [1]. Nuclear energy is another source, however, the possible risk of contamination associated with nuclear energy is high [2]. Fossil fuels, which include oil, natural gas and coal, supply almost 65% of the total world energy. Fossil fuels in conventional and unconventional hydrocarbon reservoirs will continue to account for the large proportion of the world energy supply in the next several decades [3], [4], [5]. Managing hydrocarbon reservoir workflows normally involve numerous simulations for optimizing production, enhancing oil recovery and history matching [6], [7], [8], [9]. Flow of fluid in porous media is governed by complex nonlinear partial differential equations (PDEs), which in practice, are spatially discretized into a high-dimensional set of nonlinear ordinary differential equations (ODEs) [10], [11]. For consistent representation of flow dynamics and subsurface geology, grid blocks in very large numbers are required and cumbersome algorithms are employed for their spatiotemporal solutions. It requires thousands of simulations even with advanced algorithms to achieve optimal solutions when solving with nonlinear constraints [12], [13]. The complexity in physics associated with reservoir multiphase fluid flow and the multiscale nature of the rock and fluid properties present challenges in achieving better predictive models.

At early stage of production, most hydrocarbon reservoirs produce under primary recovery where a reservoir (formation) pressure forces the fluid into the well. However, since production is usually accompanied by a decrease in reservoir pressure, primary recovery via natural lift soon comes to an end. When a large portion of the oil or gas in a reservoir cannot be recovered by primary production, a method known as secondary or enhanced recovery is employed to pressurize the reservoir. Secondary recovery is achieved by injecting water (waterflooding) or gas (gas flooding) into the reservoir to displace produced fluids and hence maintain or boost the reservoir pressure [14], [15]. The performance of a reservoir model operating under secondary recovery can be significantly influenced by different types of parameters. These parameters include static reservoir parameters (such as porosity, permeability, grid location), dynamic reservoir parameters (such as reservoir pressure, fluid composition, fluid saturation, relative permeability, and well data (such as injection rate, production rate, well radius and well patterns [16], [1]. Conventionally, numerical simulation is used to quantify uncertainties and find the best set of parameters that give the best

performing model. However, the higher the number of parameters, the more cumbersome the reservoir model becomes and consequently, the more computationally expensive the entire simulation process becomes. In comparison to numerical simulation, in which model set up is laborious and implementation is time consuming, a data-driven model that not only scales down computational complexity but also, offers accuracy without compromising results would be of great benefit [17], [18], [19]. Data-driven based machine learning techniques have been successfully applied in reservoir characterization and engineering to provide solutions to challenges that include estimation of flow rates of oil and gas in multiphase production systems [20], well testing and sensitivity analysis [21], [22], prediction of well performance [23], lithology identification [24], to mention but a few.

In recent years, a large body of research called Dynamic Mode Decomposition (DMD) has emerged around modal decomposition and machine learning methods [25], [26], [27], [28], [29]. Developed by Schmid [30], DMD originated as a new promising tool in the fluid dynamics community to discover spatiotemporal meaningful structures from high-dimensional fluids data. The evolving success of DMD arises from the fact that it is a data-driven and equation-free technique that is capable of discovering spatiotemporal meaningful patterns that may be used for diagnosis, control, state estimation and future-state prediction of complex dynamical systems [31]. Showing that DMD is connected to the underlying nonlinear dynamics via Koopman operator theory, the method quickly gained acceptance as a technique for analyzing data from nonlinear dynamical systems [32], [33]. Even though DMD was initially introduced in the fluid dynamics/mechanics community, the method has proved to be a powerful tool that encompasses a wide range of techniques, from statistical regression and machine learning to computer vision and compressive sensing [34]. Several studies have applied DMD and in some cases, modifications have been performed on the standard DMD to address specific problems. The authors in [35] applied DMD via dictionary learning to separate video streams into background and foreground information which resulted in extracting dynamics and detecting moving objects in videos. In [36], DMD was coupled with convolutional neural network to decompose and visualize a flow around circular cylinder at various Reynold numbers. The authors in [25] demonstrated the capability of integrating DMD with multi-resolution analysis to decompose video streams into multi-

time scale objects and features. In [37], comparison was made between DMD and deep learning techniques for analysis of two-phase flow data generated via numerical simulations and experimental setups. In this application, DMD outperformed the deep learning techniques. In [38], DMD was applied to analyze voltage of a power system and in [39], time frequency spectral analysis of nonlinear flows was performed using DMD. In [40], the authors extended the application of DMD by introducing local DMD and sparse DMD. They demonstrated the application of sparse DMD on a single-phase reservoir model to capture reservoir behavior and applied the local DMD to create a proxy model for application in a hydraulic fracturing process.

Being a data-driven technique and despite its successful application in diverse areas, it is observed that standard DMD works best when the number of columns of the snapshot matrix, which is also the number of time steps, is smaller than the number of rows, which is also the number of observations, measurements, or samples for the purpose of retaining the spatial information wherever possible. In standard DMD, Singular Value Decomposition (SVD) is performed on the full-state snapshot data. For instance, performing SVD on a $m \times n$ snapshot matrix M results in three matrices $U,\ S,$ and V, where U is an $m \times m$ unitary matrix, S is an $m \times n$ diagonal matrix, and V is an $n \times n$ unitary matrix. The entries of S are referred to as the singular values of M, the columns of U and V are referred to as the left-singular and right-singular vectors of M, respectively. After performing SVD on the full-state snapshot matrix M, one needs to truncate the components of the three matrices U, S and V according to the rank of the snapshot matrix. In a situation like this, if the number of columns of the full-state snapshot matrix M is greater than the number of rows, part of the columns of U (left-singular vector) that carries spatial information will be lost. Even though, this might not be a problem when dealing with overdetermined systems, nevertheless, this is a drawback when dealing with underdetermined systems as information regarding low level dynamics in spatiotemporal data might not be analyzed. A modified version of standard DMD called compressed DMD produces almost similar result as standard DMD but at low cost of computation [41]. Compressed DMD integrates standard DMD with compressive sensing to achieve results by reconstructing a full-state snapshot from a random under sampling of the full-state data [42]. The basic idea behind compressed DMD method is to construct a

measurement matrix $C \in \mathbb{R}^{p \times n}$ of random samples first. The measurement matric C is then used to compressed the full-state snapshot matrices after which mode decomposition is computed on the compressed representation of the original data. It is worth noting that the measurement matric $C \in \mathbb{R}^{p \times n}$ which is the cornerstone of compressed DMD is formed by drawing $p \times n$ independent random samples, where p stands for the number of samples randomly drawn. As such, the measurement matrix C is a kind of random sensing matrix whose number of samples (rows) randomly generated is equal to the value of p. To compress the full-state snapshot matrices, the algorithm uses the measurement matrix C to generate compressed matrices whose number of rows are randomly picked from the full-state snapshot matrices without replacement and are equal to the value of p. Even though, with compressed DMD, computational cost is reduced compared to standard DMD, nevertheless, a noticeable drawback with compressed DMD is that, the compression of the full-state snapshot data is performed by a measurement matrix whose components are randomly generated, as such, due to the randomness nature of the measurement matrix, anytime the algorithm is ran, its components change and in turn, the overall result changes.

This chapter presents a new technique known as Dynamic Mode Learning (DML) that is developed based on extracting the characteristics of the full-state data and using the learned and extracted features to develop a model that accurately capture dynamics and provide behavior analysis of the full-state system. To achieve this, a simple and fast learning system that extracts important features of the full-state data is developed first. The extracted features that contain the characteristics and underlying dynamics in the full-state data are then used to reduce the full-state data, after which SVD and dynamic modes are computed on the resulting reduced data. Finally, the full-state system is reconstructed from the eigenvalues and dynamic modes of the reduced system. As such, this approach is aimed at developing a model that accurately capture the underlying dynamics in the full-state system.

6.2.1 Novelty and Contribution

The main contribution of this chapter is the successful development of a novel DML model that addresses the drawbacks of standard and compressed DMD mainstream algorithms. By learning and extracting the features of the full-state data, the proposed DML model eliminates

the need for computing dynamic modes on the full-state data as in standard DMD and rather, results in computing dynamic modes on a reduced data that contains the underlying dynamics of the full-state data. Furthermore, by using a learning system to extract features that are fixed and used to reduce the full-state data, the proposed DML model also addresses the unpredictability of results that arise from randomness nature of measurement matrix as identified in compressed DMD. The eigenvalues on the reduced data matrix of the proposed technique are shown to be similar as the full-state data eigenvalues. In this new technique, once the parameters of the learning system are configured and the system is trained on the full-state data, its output (extracted features) remain fixed and no matter how many times the final model developed from the extracted features is ran, overall result remains unchanged. The performance of the proposed DML model developed in this chapter is validated using four different data sets that represent pressure dynamics and fluid phase distribution of multiphase reservoir models. Firstly, the technique is applied to a pressure field data of a 3-D multiphase reservoir model that describes gas injection/oil production scenario. In this first application, the data was generated by simulating the reservoir model using the benchmark data of the first Society of Petroleum Engineers (SPE1) Comparative Solution Project. The aim of the first application is to assess the learning capability of the proposed technique on the pressure dynamics of the reservoir model at each grid block over time. Secondly, the new technique is applied on a pressure field data generated from simulating another different 3-D multiphase reservoir model that describes a water injection scenario to capture the pressure dynamics of the reservoir model. Third and fourth applications are performed on the water and gas saturation data sets generated from simulating the water injection reservoir model. The aim of the third and fourth applications is to validate the ability of the proposed model in capturing and reconstructing the fluid phase distribution data generated from simulating the 3-D multiphase reservoir model that describes a water injection scenario. In simulating the water injection reservoir model, relative permeabilities are sampled from the SPE3 benchmark data. Experiments were performed by running the proposed DML model on the four data sets generated from direct numerical simulations using varied number of modes. Comparing the performance of the proposed DML model in terms of relative errors of

prediction and reconstruction with those of standard and compressed DMD baseline algorithms, the new DML model proved to be more accurate with insignificant errors.

The remaining parts of this chapter are organized as follows: Section three presents materials and methods used to develop the proposed DML model. Section four presents the formulation and implementation of the two reservoir models whose data were used to validate the proposed DML model. Section five presents experimental results and discussion. Section six ends the chapter with conclusion and focus for future work.

6.3 Dynamic Mode Learning Model Development

This section presents the details of the proposed DML model development. This includes a description of the learning system that extracts features of the full-state data, a detailed description of the proposed DML method working principle, and a step-by-step algorithm that describes the development of the proposed DML method.

6.3.1 Feature Extraction

The proposed DML method is developed based on the idea of future extraction in high-dimensional data [43]. The main idea behind the proposed DML model is learning and extracting important features of the full-state data first, and then using the extracted features to reduce the full-state data. In contrast to the traditional compressed DMD that constructs a measurement matrix from random samples and uses it to compress the full-state data, we first developed a simple learning system that extracts and maps the underlying features of the full-state data to a low-dimensional space and saves them as a set of extracted features. This set of extracted features are then use to reduce the full-state data after which dynamic modes are computed on the reduced data. Suppose we present the full-state data which contains n samples as input data X and project it using $g(XW_i + b_i)$, to transform it into ith extracted features, F_i, where $g: R^n \rightarrow R^n$ represents the transfer function, W_i stands for the weight matrix, and b_i stands for the bias vector. The concatenation of all the first i sets of extracted features is denoted as $F^i \equiv [F_1, \dots, F_i]$. If $L \in R^{k \times n}$ is defined as the output matrix where k is the number of nodes in the extraction layer, then for any n samples of the full-state data the learning system generates k samples of extracted features which can be expressed as follows

$$F_i = g(XW_i + b_i), \quad i = 1, \dots, k \tag{6.1}$$

where W_i and b_i are randomly generated from the normal uniform distributions within the interval of [-1, 1]. For the transfer function, log-sigmoid is chosen to establish the samples of extracted features and to improve the generalization ability of the system, L_2-norm weight regulirizer is added to the transfer function. The architecture of the feature extraction system that describes its working process is shown in Figure 6.1.

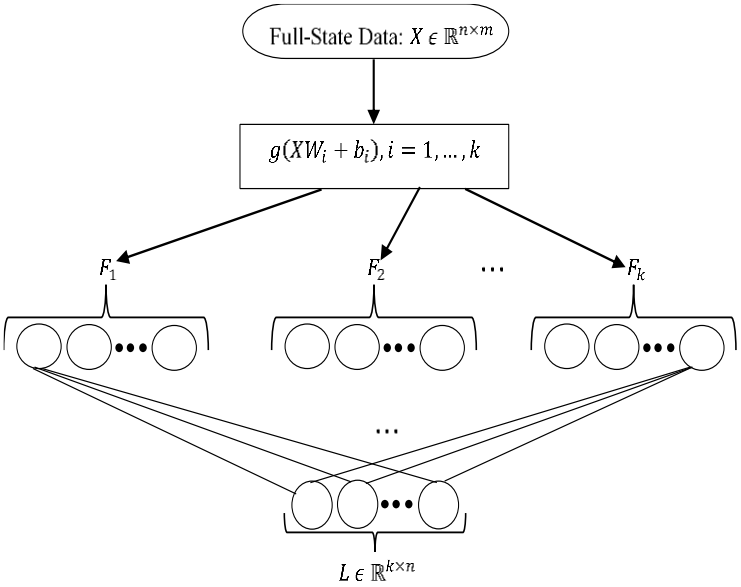

Figure 6.1: Illustration of the workflow process of the feature extraction system.

Once the system extracts the features and placed them in $L \in \mathbb{R}^{k \times n}$, the extracted features are then used to reduce the full-state data after which the dynamic modes are computed on the reduced data. As such, computationally expensive SVD on the high-dimensional original data is bypassed, and is rather, performed on the reduced snapshot data. Also, it is worth noting that the number of samples/rows in the reduced snapshot data will be equal to the number of samples/rows in the output layer of the feature extraction system (that is, the value of *k*).

6.3.2 Dynamic Mode Learning (DML) Method

In this section, a detailed description of the proposed DML method is presented. Suppose that a high-dimensional data $X \in R^{n \times m}$ is generated and collected from numerical simulations or experiments, the features of the original data X are first extracted using (6.1) and placed

in a low-dimensional matrix, $L \epsilon R^{k \times n}$. Next, the original data are arranged in two snapshot matrices as follows

$$X = \begin{bmatrix} | & | & | \\ x_1 & \dots & x_{m-1} \\ | & | & | \end{bmatrix}, \tag{6.2a}$$

$$X' = \begin{bmatrix} | & | & | \\ x_2 & \dots & x_m \\ | & | & | \end{bmatrix}. \tag{6.2b}$$

These two matrices have large number of rows than columns, that is, $n \gg m$ and consist of the states of the system and their columns were captured in equal-spaced time, with a time step Δt. Each $X_i = X(i\Delta t)$ is a vector with components c, as such, $X, X' \in \mathbb{R}^{c \times (m-1)}$. Dynamic mode learning method attempts to construct a linear dynamical system

$$X_{t+1} \approx AX_t \tag{6.3}$$

and thus

$$X' \approx AX. \tag{6.4}$$

It is interesting to realize that the least-squares solution of (6.4) leads to

$$A = X'X^{\dagger} \tag{6.5}$$

here, $X^{\dagger}$ stands for the Moore-Penrose pseudo-inverse of X. To get an estimate of matrix A, DML method uses the matrix that contains the extracted features to reduce the snapshot matrices as follows

$$X_R = L * X \tag{6.6a}$$

$$X'_R = L * X', \tag{6.6b}$$

then Singular Value Decomposition (SVD) is computed on the reduced snapshot matrix X_R as follows

$$X_R = USV^* \tag{6.7}$$

where $U \in \mathbb{R}^{n\times r}, S \in \mathbb{R}^{r\times r}, V \in \mathbb{R}^{m\times r}$ and $r \leq m$ stands for the rank of truncating the snapshot matrix X_R. The columns of *U* are referred to as POD modes, and they satisfy $U^* \cdot U = I$. In the same manner, columns of *V* are orthonormal, and satisfy $V^* \cdot V = I$. The diagonal of *S* contains the singular values of matrix X_R. The full matrix *A* then can be acquired by solving the pseudo-inverse of X_R as follows

$$A = X_R' V S^{-1} U^*. \tag{6.8}$$

In DML method, the interest is in the leading eigenvalues *r* and eigenvectors of *A*, for this reason, *A* is therefore projected onto the POD modes in *U* as follows

$$\tilde{A} = U X_R' V S^{-1} \tag{6.9}$$

It is worth mentioning that the solution of (6.9) is the least squares fit of A. The point here is, instead of working on the full matrix *A*, we directly computed the reduced-order approximation $\tilde{A}$ in such a way that the full matrix *A* and the reduced-order matrix $\tilde{A}$ have the same nonzero eigenvalues. Thus, the spectral decomposition of $\tilde{A}$ can be computed as

$$\tilde{A} H = H\Lambda. \tag{6.10}$$

Here, the DML eigenvalues are the elements of the diagonal matrix Λ and the eigenvectors of $\tilde{A}$ are represented by the columns of *H*. The dynamic mode ϕ can then be obtained by using *W* of the reduced system and the snapshot matrix X' as follows

$$\phi_L = X' V S^{-1} H. \tag{6.11}$$

Algorithm 6.1 summarizes the steps involved in the DML method.

Algorithm 6.1: Proposed DML method development

Input: Original data matrix

Output: Dynamic modes $\boldsymbol{\phi_L}$ and eigenvalues $\boldsymbol{\Lambda}$

1: Extract features of the original data and save them in $\boldsymbol{L} \in \mathbb{R}^{\boldsymbol{k}\times\boldsymbol{n}}$

2: Arrange the original data as snapshot matrices $\mathbf{X}$ **and** $\boldsymbol{X'}$

3: Reduce the snapshot matrices using the extracted features as $\boldsymbol{X_R = L * Y},\ \boldsymbol{X'_R = L * Y'}$

4: Compute SVD on the reduced snapshot matrix $\boldsymbol{X_R}$ to rank *r* as $\boldsymbol{USV} = \mathbf{svd}(\boldsymbol{X_R}, \boldsymbol{r})$

5: Obtain the full matrix $\boldsymbol{A}$ **as** $\boldsymbol{A = X'_R VS^{-1} U^*}$

5: Compute least squares fit as $\boldsymbol{\widetilde{A} = UX'_R VS^{-1}}$

6: Compute the eigen-decomposition of $\boldsymbol{\widetilde{A}}$ as $\boldsymbol{H, \Lambda} = \mathbf{eig}(\boldsymbol{\widetilde{A}})$

7: Compute dynamic modes as $\boldsymbol{\phi_L = X' VS^{-1} H}$

6.4 Numerical Reservoir Simulation Setup

This section presents the formulation and implementation of the two reservoir models that were used as reference and whose data were used to validate the proposed DML method.

6.4.1. Formulation of Multiphase Reservoir Model

Fluid flow in porous media can be described by a system of partial differential equations. The system must contain constitutive equations that relate various physical quantities, boundary conditions, as well as source terms that illustrate the forces driving the fluid flow. In the following, we describe in brief the multiphase flow model used in this chapter. For the sake of brevity, details are skipped but a more extensive discussion can be found in [44]. Consideration is given to a generic system for multiphase flow and the mass conservation equation for each phase α is written as

$$\frac{\partial}{\partial t}(\varepsilon \rho_\alpha S_\alpha) + \nabla \cdot (\rho_\alpha v_\alpha) = \rho_\alpha q_\alpha, \tag{6.12}$$

where ε represents rock porosity, ρ_α denotes fluid phase density, S_α represents fluid phase saturation, ∇ stands for divergence operator, v_α denotes flow velocity of the fluid phase, q_α denotes volumetric source/sink term, and t stands for time. By using the concept of relative permeabilities, the extension of Darcy's law to multiphase flow is applied as follows:

$$v_a = -\frac{Kk_{r\alpha}}{\mu_\alpha}(\nabla p_\alpha - g\rho_\alpha \nabla h), \tag{6.13}$$

where K denotes permeability, k_{ra} stands for phase relative permeability, μ_α denotes fluid phase viscosity, ∇ stands for gradient operator, p_α denotes phase pressure, g stands for gravity, and h stands for reservoir depth.

To compute the approximate solutions, discrete derivative operators were introduced for all derivatives and the resulting system of fully implicit, discrete flow equations for phase α is written as

$$\frac{(\varepsilon S_\alpha \rho_\alpha)^{n+1} - (\varepsilon S_\alpha \rho_\alpha)^n}{\Delta t^n} + div(\rho v)_\alpha^{n+1} = (\rho q)_\alpha^{n+1}, \tag{6.14a}$$

$$v_\alpha^{n+1} = \frac{Kk_{ra}}{\mu_\alpha^{n+1}}\{grad(p_\alpha^{n+1}) - g\rho_\alpha^{n+1} grad(h)\}, \tag{6.14b}$$

where S_α, ε and p_α denote vectors with one saturation, one porosity, and one pressure value per cell, respectively, v_α denotes vector of fluxes for phase α per face, and so forth. The superscript denotes discrete times at which the unknown states of the reservoir are computed, Δt stands for the interval between two points in time, and div and $grad$ denote discrete operators for divergence and gradient operators, respectively. The general multiphase model is completed by imposing the saturation constraint. Equation (6.14) is nonlinear which is solved numerically using a fully-implicit two-point finite-volume procedure to generate spatiotemporal evolution of the reservoirs' pressure and fluid phase saturation at any point in time.

6.4.2 Gas Injection Reservoir Model

The gas injection reservoir model used as the first reference model is implemented using the benchmark data of the first SPE Comparative Solution Project [45]. The SPE 1 benchmark

is a description of a depletion problem with gas injection in a $10 \times 10 \times 3$ reservoir model with an injector and a producer wells completed in diagonally opposite corners. The gas injection well was completed in layer 1 and is located at grid point (1, 1) while the producing well was completed in layer 3 and is located at grid point (10, 10). The reservoir has a porosity of 0.3 uniformly distributed within the grid blocks, whereas the permeability is heterogeneous with values 500, 50, and 200 mD in layers 1, 2, and 3, respectively with respective thicknesses of 20, 30, and 50 ft. Initially, the reservoir is undersaturated with a constant pressure field in each layer, a homogeneous mixture of water ($S_w = 0.12$), and oil ($S_o = 0.88$), and zero free gas $(S_g = 0.0)$ throughout the reservoir model. Detailed data that describes the petrophysical and PVT properties as well as the relative permeability of the reservoir model can be found in [45]. As mentioned above, the geological structure of the gas injection reservoir model is a three-dimensional formation that consists of 10×10 grid blocks in the $x - y$ dimension and 3 layers in the z dimension. Thus, the reservoir model consists of $10 \times 10 \times 3 = 300$ grid blocks (cells). To generate the reservoir's data at any grid cell in time, the reservoir is numerically simulated for 1200 days in 120 time-steps. The result of the numerical simulation yielded a 300×120 matrix that contains 36,000 records in the spatiotemporal database. Each record in the spatiotemporal database contains information about the reservoir's static and dynamic parameters in a single grid block in a given run and given time step. For the purpose of this chapter, the pressure field data of the gas injection reservoir model is retrieved and utilized as original data by the proposed DML model and comparison algorithms.

6.4.3 Water Injection Reservoir Model

For the second reference model, a three-phase flow model with water injection in a $9 \times 9 \times 4$ reservoir model that describes the effect of injector and producer wells scenario is set up. The reservoir model has a rock compressibility of 6.89e-8/bar and the initial reservoir pressure of 200 bar. Initially, the bottom layer is filled with water, the two middle layers are filled with oil, and the upper layer is filled with gas. Water is injected from the lower southwest corner, and from a cell at the northeast corner of the middle layers, fluids are produced by setting the pressure in this cell lower than the initial average reservoir pressure. The reservoir model has

a porosity of 0.13 and permeability of 30 mD uniformly distributed within the grid blocks, with relative permeabilities sampled from SPE3 benchmark data [46]. The fluids properties are defined as follows: water has viscosity, compressibility, and density of 1 cP, 1.37e-7/bar, and 1014 kg/m^3, respectively. Oil has viscosity, compressibility, and density of 0.5 cP, 6.89e-7/bar, and 850 kg/m^3, respectively. Gas has viscosity, compressibility, and density of 0.015 cP, 6.89e-5/bar, and 1.2 kg/m^3, respectively. When the reservoir starts production, the reservoir pressure decreases and the compressed gas expands so that a large proportion of it is driven out of the formation through the oil layer. Considering that gas is more mobile than oil, this causes the gas to flow and pass the oil instead of displacing it. Thus, only a minor proportion of the oil will be produced. To improve oil production, water is then injected into the water zone (bottom layer) which boost the reservoir pressure and in turn, enhance oil production. As mentioned above, the geological structure of the water injection reservoir model is a three-dimensional formation that consists of 9×9 grid blocks in the $x - y$ dimension and 4 layers in the z dimension. Thus, the reservoir model consists of $9 \times 9 \times 4 = 324$ grid blocks (cells). To generate the reservoir's data at any grid cell in time, the reservoir is numerically simulated for 1095 days (3 years) in 220 time-steps. The result of the numerical simulation yielded a 324×220 snapshot matrix that contains 71,280 records in the spatiotemporal database. For the purpose of this study, the reservoir's pressure field data, water saturation field data and gas saturation field data were retrieved and utilized by the proposed DML model and comparison algorithms.

6.5 Experimental Results and Discussion

In this section, results of experiments for the purpose of verifying the proposed DML model are presented. To validate the effectiveness of the proposed model, experiments are performed on the four different data sets generated from simulating the two multiphase flow reservoir models. Two of the data sets are for reservoir pressure fields and the other two are for reservoir fluid phase (water and gas saturations) distribution. To prove the performance of the proposed DML model, its ability to reconstruct reservoir pressure field, approximate average reservoir pressure change, reconstruct water saturation field, and reconstruct gas saturation field are compared to existing mainstream standard and compressed DMD methods. Each of the three algorithms mentioned above was evaluated using 10, 15, 20, and

25 number of modes at a time. Two classic statistical quantities are selected namely, mean square error (MSE), and root mean squared error (RMSE) to evaluate the performance of the three algorithms. The aim of performing the experiments is to measure the errors generated by each algorithm on each of the tasks performed on the original data sets.

Recall that the proposed DML algorithm uses the extracted features representation matrix $L \in \mathbb{R}^{k \times n}$ to reduce the full-state snapshot matrices after which mode decomposition is computed on the reduced data. In the proposed DML algorithm, the number of rows/samples in the reduced snapshot data is equal to the number of rows/samples in $L \in \mathbb{R}^{k \times n}$ which in turn, is equal to the number of nodes in the output layer (value of k). Recall also, that in compressed DMD, a measurement matrix $C \in \mathbb{R}^{p \times n}$ is used to compress the snapshot matrices after which mode decomposition is applied on the compressed data. In compressed DMD, the number of rows/samples in the compressed snapshot data is equal to the number of rows/samples in the measurement matrix which is the number of random samples generated by the measurement matrix (value of p). For fair comparison, we set the value of p for the compressed DMD method, and k for the proposed DML model to 25 each. Thus, $C \in \mathbb{R}^{p \times n}$ and $L \in \mathbb{R}^{k \times n}$ have the same number of rows/samples so that the resulting compressed snapshot matrices that the compressed DMD algorithm utilized have the same rows/samples with the reduced snapshot matrices utilized by the proposed DML algorithm. For the standard DMD algorithm, no compression or reduction of the full-sate data is needed, thus, SVD is applied directly on the full-state snapshots.

6.5.1 Experiment on The Gas Injection Reservoir Pressure Field Data

First experiment is performed on the pressure field data generated from simulating the gas injection reservoir model. The original data contains 300 samples generated in 120 time steps. Thus, the original data is represented in a 300×120 matrix. To gain an insight of how the pressure evolves within the reservoir grid cells as the reservoir produces while natural gas is being injected, plot of the reservoir's average pressure change over time is shown in Figure 6.2 and the reservoir's pore pressure evolution for some selected days of the simulation period is shown in Figure 6.3.

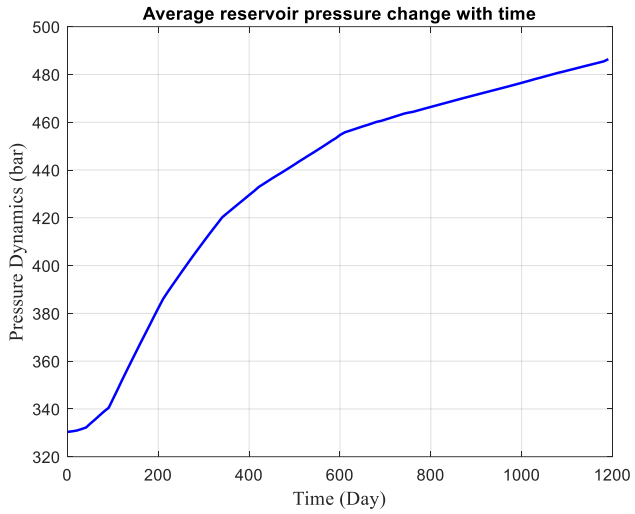

Figure 6.2: Plot of average reservoir pressure change with time for gas injection reservoir simulation.

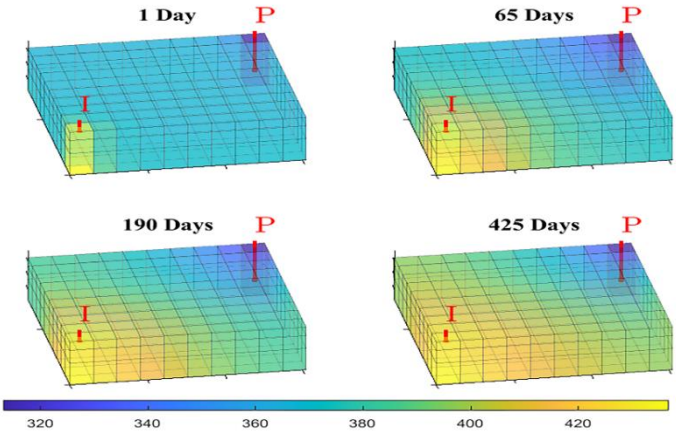

Figure 6.3: Reservoir pore pressure variations for some selected days of gas injection reservoir simulation.

Projecting the original data on the feature extraction system described in section 6.3.1 and setting the value of *k* to 25, not only reduces the dimension of the data, but also captures the characteristics of the data and stored them in 25×300 matrix. This matrix is then used to reduce the full-state 300×120 matrix after which dynamics modes are computed on the reduced data. Table 6.1 presents the experimental results performed on the pressure field data of the gas injection reservoir model by the three methods using 10, 15, 20, and 25 number of modes. From Table 6.1, it can be seen that on general note, pressure field reconstruction and

average reservoir pressure approximation errors for all the three methods decrease as the number of modes increases. In other words, accuracy for all the three method improves as the number of modes increases. However, it is worth noting that the proposed DML model exhibits better performance with all the number of modes over standard and compressed DMD methods. In particular, it can be observed that with 15 modes, the proposed DML model is able to record < 1 MSE and < 1 RMSE for both pressure field reconstruction and average reservoir pressure approximation while standard and compressed DMD methods have not recorded such results even with 25 modes. Overall performance ranking for this first experiment is DML model, followed by standard DMD, then compressed DMD.

Table 6.1: Comparison of algorithm performance evaluation on gas injection reservoir pressure field data using different number of modes.

	Reservoir's Pressure Field Reconstruction		**Average Reservoir Pressure Approximation**	
	MSE	**RMSE**	**MSE**	**RMSE**
		10 Modes		
Standard DMD	26.7476	5.1718	25.8519	5.0845
Compressed DMD	27.2875	5.2237	26.8314	5.1799
DML model	10.5455	3.2474	10.2619	3.2034
		15 Modes		
Standard DMD	9.1087	3.0181	8.9316	2.9886
Compressed DMD	12.3192	3.5099	11.8441	3.4415
DML model	0.3289	0.5735	0.2873	0.5360
		20 Modes		
Standard DMD	6.8884	2.6246	6.7435	2.5968
Compressed DMD	7.0209	2.6497	6.7113	2.5906
DML model	0.4286	0.6547	0.3841	0.6198

	25 Modes			
Standard DMD	2.2879	1.5126	2.1944	1.4813
Compressed DMD	3.5791	1.8919	3.4204	1.8494
DML model	0.3093	0.5561	0.2706	0.5202

6.5.2 Experiment on The Water Injection Reservoir Pressure Field Data

Second experiment is performed on the pressure field data generated from simulating the water injection reservoir model. The original data contains 324 samples generated in 220 time steps. Thus, the original data is represented in a 324×220 matrix. To gain an insight of how the pressure evolves within the reservoir grid cells as the reservoir produces while water is being injected, plot of the reservoir's average pressure change over time is shown in Figure 6.4.

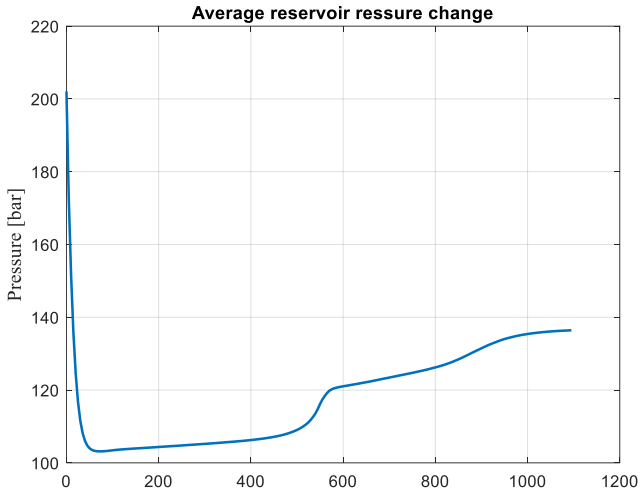

Figure 6.4: Plot of average reservoir pressure change with time for water injection reservoir simulation.

Projecting the original data on the feature extraction system described in section 6.3.1 and setting the value of *k* to 25 reduces the dimension of the data and captures the characteristics of the data which are then stored in a 25×324 matrix. This matrix is then used to reduce the full-state 324×220 matrix after which dynamics modes are computed on the reduced data. Table 6.2 presents the competitive experimental results performed on the pressure field data of the water injection reservoir model by the three methods using 10, 15, 20, and 25

number of modes. From Table 6.2, it can be seen that on general note, pressure field reconstruction and average reservoir pressure approximation errors for all the three methods decrease as the number of modes increases. In other words, accuracy for all the three methods improves as the number of modes increases. However, it is can be noticed that the proposed DML model exhibits better performance with all number of modes over standard and compressed DMD methods. In particular, it can be observed here that with 10 modes, the proposed DML model is able to record < 1 MSE and < 1 RMSE for both pressure field reconstruction and average reservoir pressure approximation. Standard DMD recorded < 1 MSE and < 1 RMSE for both pressure field reconstruction and average reservoir pressure approximation with 15 modes, and finally, with 20 modes, compressed DMD recorded < 1 MSE and < 1 RMSE for both pressure field reconstruction and average reservoir pressure approximation. Overall performance ranking for this second experiment is DML model, followed by standard DMD, then compressed DMD.

Table 6.2: Comparison of algorithm performance evaluation on water injection reservoir pressure field data using different number of modes.

	Reservoir's Pressure Field Reconstruction		**Average Reservoir Pressure Approximation**	
	MSE	**RMSE**	**MSE**	**RMSE**
		10 Modes		
Standard DMD	2.1018	1.4498	2.0737	1.4400
Compressed DMD	5.0937	2.2569	5.0217	2.2409
DML model	0.3392	0.5824	0.3359	0.5796
		15 Modes		
Standard DMD	0.8332	0.9128	0.8232	0.9073
Compressed DMD	2.0252	1.4231	1.9834	1.4083
DML model	0.1099	0.3315	0.1087	0.3297
		20 Modes		

Standard DMD	0.3328	0.5769	0.3294	0.5739
Compressed DMD	0.4742	0.6886	0.4687	0.6846
DML model	0.0780	0.2793	0.0772	0.2778
		25 Modes		
Standard DMD	0.1056	0.3250	0.1045	0.3232
Compressed DMD	0.3682	0.6068	0.3636	0.6030
DML model	0.0452	0.2126	0.0447	0.2114

6.5.3 Experiment on Fluid Phase Distribution Data

The third and fourth experiments are performed on the fluid phase distribution data generated from simulating the water injection reservoir model. These data consist of two sub-data sets namely, water saturation field and gas saturation field data. Each of these fields contain 324 samples generated in 220 time steps, thus, the water saturation and gas saturation fields data are represented in 324×220 matrix each. These data sets are projected on the feature extraction system one at a time by setting the value of *k* to 25. Both data sets were reduced in dimensions and for each, their characteristics were captured and stored in a 25×324 matrix. This matrix is then used to reduce the full-state 324×220 matrix for each of the data sets, after which dynamics modes are computed on the reduced data. Table 6.3 presents the competitive experimental results performed on the water and gas saturation fields data. It can be noticed that all the three methods recorded insignificant errors of < 1 with all modes, this is not surprising due the fact that the saturation values for both water and gas are numbers that lie between 0 and 1. From Table 6.3, it can be seen on general note that, for all the three methods, reconstruction errors decrease as the number of modes increases for both water and gas saturation fields. In other words, accuracy of all the three method improves as the number of modes increases. However, it is can be noticed in particular that the proposed DML model exhibits better performance with all number of modes over standard and compressed DMD methods. Overall performance ranking for this experiment is DML model, followed by standard DMD, then compressed DMD.

Table 6.3: Comparison of algorithm performance evaluation on water and gas saturation fields data using different number of modes.

	Water Saturation Field Reconstruction		**Gas Saturation Field Reconstruction**	
	MSE	**RMSE**	**MSE**	**RMSE**
		10 Modes		
Standard DMD	1.95e-05	0.0398	5.87e-05	0.0608
Compressed DMD	4.25e-05	0.0536	1.80e-04	0.1072
DML model	5.33e-07	0.0068	2.41e-05	0.0466
		15 Modes		
Standard DMD	1.96e-06	0.0121	2.14e-05	0.0406
Compressed DMD	1.02e-05	0.0271	7.30e-05	0.0729
DML model	1.35e-07	0.0028	8.74e-06	0.0299
		20 Modes		
Standard DMD	3.00e-07	0.0047	9.88e-06	0.0307
Compressed DMD	1.67e-06	0.0115	2.80e-05	0.0471
DML model	4.21e-08	0.0017	3.89e-06	0.0216
		25 Modes		
Standard DMD	6.84e-08	0.0019	4.05e-06	0.0220
Compressed DMD	2.97e-07	0.0043	1.18e-05	0.0291
DML model	2.50e-08	0.0014	2.53e-06	0.0185

For the sake of visualization, plots that compare the prediction of average reservoir pressure over time by all the three algorithms using 15 number of modes on the pressure field data of

the gas and water injection reservoir models are presented in Figures. 6.5 and 6.6, respectively.

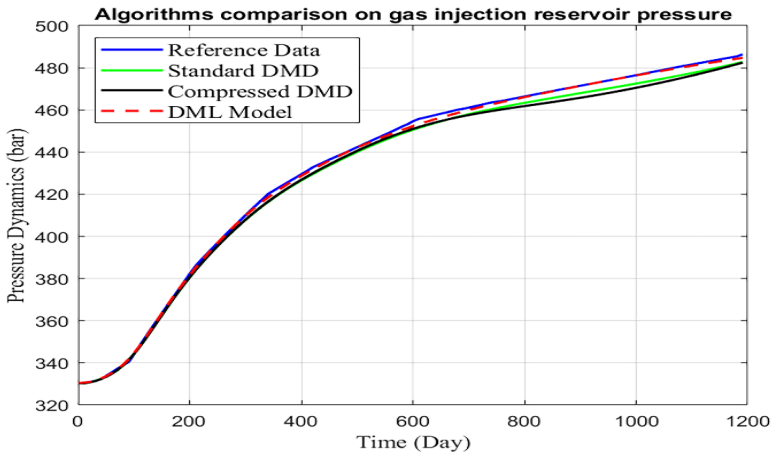

Figure 6.5: Comparison of algorithms performance for average reservoir pressure prediction on gas injection reservoir data.

Figure 6.5 reports the effect of gas injection on the reservoir's pressure dynamics over the simulation period. It can be observed from Figure 6.5 that as gas is being injected through the injection well into the reservoir model, the formation pressure keeps raising above the

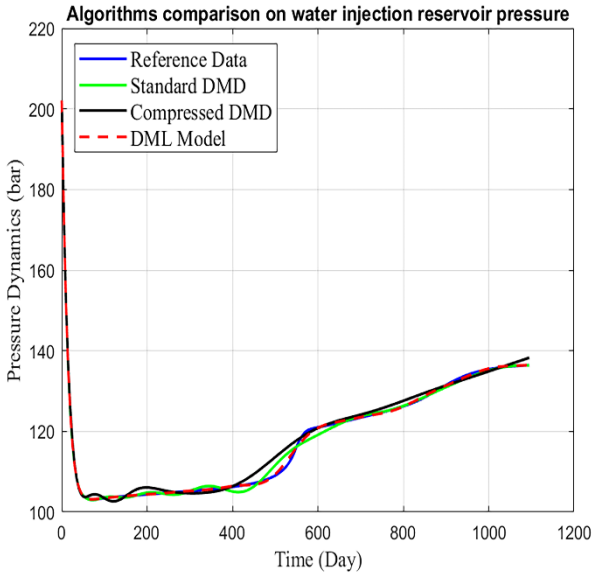

Figure 6.6: Comparison of algorithms performance for average reservoir pressure prediction on water injection reservoir data.

initial pressure and thereby enhancing the recovery of oil through the production well. As for the ability of the three algorithms on the reservoir's pressure dynamics for the simulation period, it can be noticed that from day 1 to approximately 300 days, all algorithms perform well to capture the reservoir's pressure dynamics. However, after 300 days, standard and

compressed DMD algorithms begin to lose track of the reservoir's pressure dynamics and their deviations continue to increase to the end of the simulation period. On the other hand, it can be observed that the proposed DML algorithm is able to capture the reservoir's pressure dynamics from day 1 to the end of the simulation period with a slight deviation at around day 570 to day 690.

Figure 6.6 reports the effect of water injection on the reservoirs' pressure dynamics over the simulation period. It can be noticed from Figure 6.3 that from day 1 to approximately 45 days, the reservoir's pressure quickly depleted from 200 bar to approximately 105 bar. However, after that period, the reservoir's pressure was able to be maintained by injecting water into the formation and towards the end of the simulation period (day 1095), the reservoir's pressure increases significantly to approximately 137 bar as the water front approaches the producer. As for the ability of the three algorithms on the reservoir's pressure dynamics for the simulation period, it can be noticed that from day 1 to approximately day 45, all the three algorithms were able to capture the reservoir's pressure dynamics. However, from that point to the end of the simulation period, compressed DMD lost track of the reservoir's pressure dynamics while standard DMD lost track of the pressure dynamics from approximately day 310 to day 655. On the other hand, the proposed DML model is able to keep track of the reservoir's pressure dynamics from day 1 to the end of the simulation period.

In Figure 6.7, comparison is made between the original data of the gas injection reservoir pore pressure variations for some selected days and the ones reconstructed by the proposed DML model using 15 number of modes. In Figure 6.8, comparison is also made between the original data of the water injection reservoir fluid phase distribution for some selected days and the ones reconstructed by the proposed DML model using 15 number of modes. In Figure 6.7, the effect of gas injection into the reservoir as discussed above can be seen as the formation pressure keeps rising and evolving within the grid cells. It can be noticed that the reservoir's pore pressure variations reconstructed by the proposed DML model is in good agreement with the reference data. Figure 6.8 reports the fluid phase distribution within the reservoir grid cells as water is being injected into the reservoir.

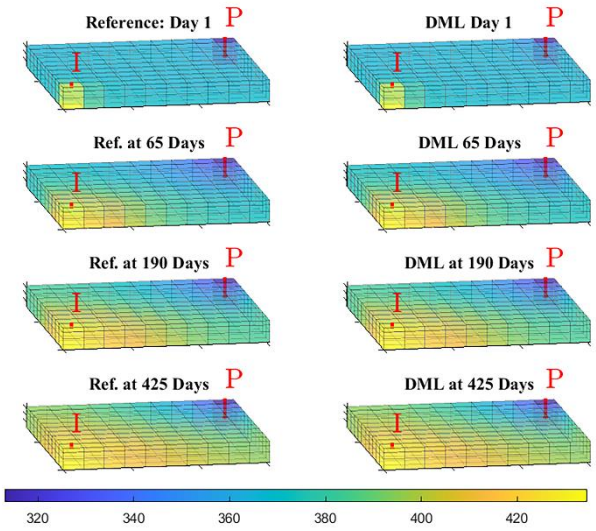

Figure 6.7: Comparison of gas injection reservoir pore pressure variations for some selected days between reference data and DML model.

It can be observed that the fluid phase distribution reconstructed by the proposed DML model is in good agreement with the reference data. Figure 6.8: Comparison of water injection reservoir fluid phase distribution for some selected days between reference data and DML model

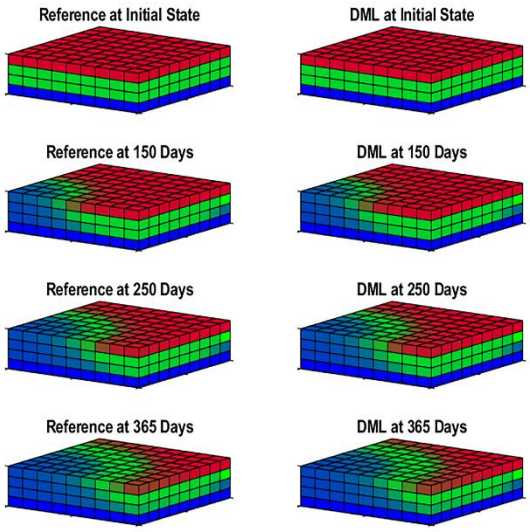

Figure 6.8: Comparison of water injection reservoir fluid phase distribution for some selected days between reference data and DML model.

Furthermore, comparisons are made between the true eigenvalues and the ones generated by the proposed DML model using 15 modes on the four data sets utilized in this study and the results are presented in Figures 6.9 through 6.12. It can be observed in Figures 6.9 to 6.12 that the eigenvalues generated by the proposed DML model match the true eigenvalues. This confirms that the proposed DML model is able to capture the dynamics in the original data sets faithfully. confirms that the proposed DML model is able to capture the dynamics in the original data sets faithfully.

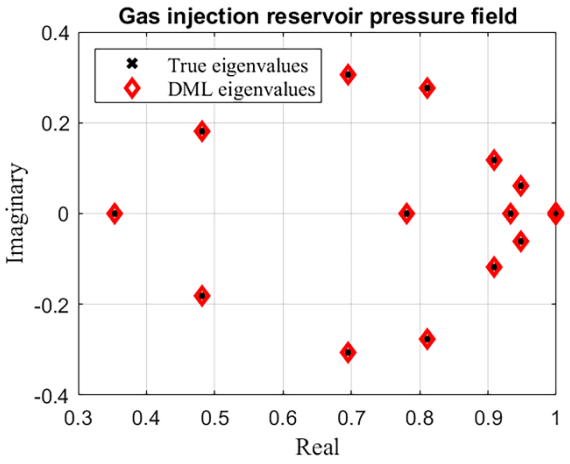

Figure 6.9: Comparison of true eigenvalues and the ones generated by the proposed DML model for gas injection reservoir pressure data.

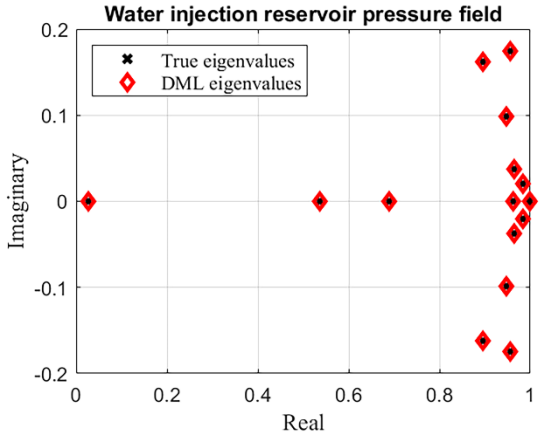

Figure 6.10: Comparison of true eigenvalues and the ones generated by the proposed DML model for water injection reservoir pressure data.

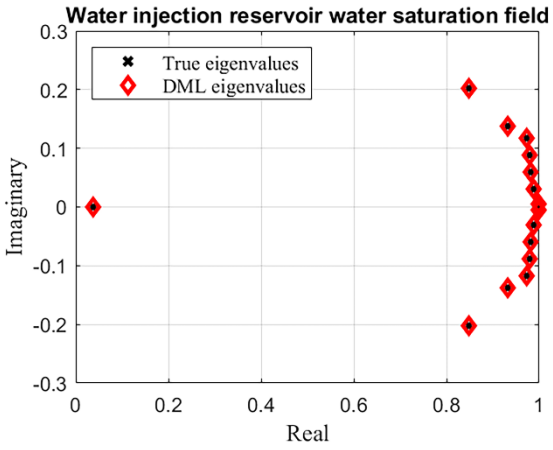

Figure 6.11: Comparison of true eigenvalues and the ones generated by the proposed DML model for water injection reservoir water saturation data.

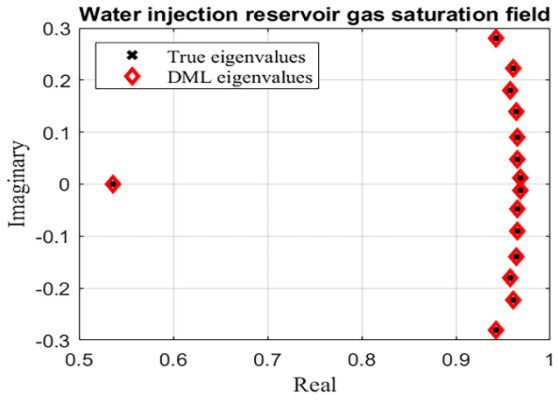

Figure 6.12: Comparison of true eigenvalues and the ones generated by the proposed DML model for water injection gas saturation data.

6.6 Conclusions

In this chapter, DML method that efficiently learns and decomposes dynamic modes in high-dimensional data is proposed. Developed based on the idea of dimensionality reduction and feature extraction, the proposed model eliminates the computationally expensive SVD of standard DMD and addresses the issue of random sampling in compressed DMD. The proposed model is developed by first, constructing a simple system that extracts significant features whose output that contains the underlying dynamics of the original data is used to reduce the full-state data, after which dynamic modes are computed on the reduced data. As such, the success of the proposed model can be attributed to feature extraction performed

first on the original data before further computations. Performance of the proposed model is validated on four different data sets generated from direct numerical simulations of two different multiphase reservoir models. Ability of the proposed model to keep track of reservoir pressure dynamics and reconstruct reservoir's pore pressure variations and fluid phase distribution are compared to mainstream algorithms namely, standard and compressed DMD methods. Experiments performed on the four data sets reveal that the proposed DML model exhibits better performance with the least prediction and reconstruction errors over standard and compressed DMD methods. Furthermore, eigenvalues generated by the proposed DML model are shown to match the true eigenvalues of the four different data sets utilized in this chapter. This shows that it is possible to apply the proposed model to high-dimensional systems in porous media, fluid dynamics, and to other spatiotemporal measurements. As focus for future research, consideration will be given to combining the proposed technique with other DMD innovations such as DMD with control.

References

[1] Z. Zhong, A. Y. Sun, Y. Wang, and B. Ren. Predicting field production rates for waterflooding using a machine learning-based proxy model. *J. Pet. Sci. Eng.*, vol. 194, pp. 107574, 2020, doi: 10.1016/j.petrol.2020.107574.

[2] S. Suman. Hybrid nuclear-renewable energy systems: A review. *J. Clean. Prod.*, vol. 181, pp. 166–177, 2018, doi: 10.1016/j.jclepro.2018.01.262.

[3] S. S. Neshat, R. Okuno, and G. A. Pope. Simulation of solvent treatments for fluid blockage removal in tight formations using coupled three-phase flash and capillary pressure models. *J. Pet. Sci. Eng.*, vol. 195, p. 107442, 2020, doi: 10.1016/j.petrol.2020.107442.

[4] J. M. Colom-Díaz, Á. Millera, R. Bilbao, and M. U. Alzueta. New results of H2S oxidation at high pressures. Experiments and kinetic modeling. *Fuel*, vol. 285, 119261, 2021, doi: 10.1016/j.fuel.2020.119261.

[5] A. M. Hassan, M. Ayoub, M. Eissa, H. Bruining, and P. Zitha. Study of surface complexation modeling on a novel hybrid enhanced oil recovery (EOR) method;

smart-water assisted foam-flooding. *J. Pet. Sci. Eng.*, vol. 195, p. 107563, 2020, doi: 10.1016/j.petrol.2020.107563.

[6] A. Salehi, G. Hetz, F. Olalotiti, N. Sorek, H. Darabi, and D. Castineira. A comprehensive adaptive forecasting framework for optimum field development planning. *Soc. Pet. Eng. - SPE Reserv. Simul. Conf. 2019, RSC 2019*, 2019, doi: 10.2118/193914-ms.

[7] P. Lerlertpakdee, B. Jafarpour, and E. Gildin. Efficient production optimization with flow-network models. *SPE J.*, vol. 19, no. 6, pp. 1083–1095, 2014, doi: 10.2118/170241-PA.

[8] Z. Guo, A. C. Reynolds, and H. Zhao. Waterflooding optimization with the INSIM-FT data-driven model. *Comput. Geosci.*, vol. 22, no. 3, pp. 745–761, 2018, doi: 10.1007/s10596-018-9723-y.

[9] Y. Zhang, J. He, C. Yang, J. Xie, R. Fitzmorris, and X. H. Wen. A physics-based data-driven model for history matching, prediction, and characterization of unconventional reservoirs. *SPE J.*, vol. 23, no. 4, pp. 1105–1125, 2018, doi: 10.2118/191126-pa.

[10] K. Aziz, and A. Setari. Applied Sci. *Petroleum Reservoir Simulation*. First Edition, pp. 357 - 372, 1979. London: Applied Sci. Publishers.

[11] J. E. Aarnes, S. Krogstad, K-A, Lie. A hierarchical multiscale method for two-phase flow based upon mixed finite elements and nonuniform coarse grids. Multiscale Model Simul, vol 5, no, 2, pp. 337 - 363, 2006, doi: 10.1137/050634566.

[12] Z. Liu, F. Forouzanfar, and Y. Zhao. Comparison of SQP and AL algorithms for deterministic constrained production optimization of hydrocarbon reservoirs. *J. Pet. Sci. Eng.*, vol. 171, pp. 542–557, 2018, doi: 10.1016/j.petrol.2018.06.063.

[13] Z. Liu and A. C. Reynolds. An SQP-filter algorithm with an improved stochastic gradient for robust life-cycle optimization problems with nonlinear constraints. *Soc. Pet. Eng. - SPE Reserv. Simul. Conf. 2019, RSC 2019*, no. 1988, 2019, doi:

10.2118/193925-ms.

[14] E. Artun. Performance assessment and forecasting of cyclic gas injection into a hydraulically fractured well using data analytics and machine learning. *J. Pet. Sci. Eng.*, vol. 195, p. 107768, 2020, doi: 10.1016/j.petrol.2020.107768.

[15] J. Wei, X. Zhou, J. Zhou, J. Li, and A. Wang. Recovery efficiency of tight oil reservoirs with different injection fluids: An experimental investigation of oil-water distribution feature. *J. Pet. Sci. Eng.*, vol. 195, p. 107678, 2020, doi: 10.1016/j.petrol.2020.107678.

[16] A. Ali and L. Guo. Neuro-adaptive learning approach for predicting production performance and pressure dynamics of gas condensation reservoir. *IFAC-PapersOnLine*, vol. 52, no. 29, pp. 122–127, 2019, doi: 10.1016/j.ifacol.2019.12.632.

[17] X. Luo and T. Bhakta. Automatic and adaptive localization for ensemble-based history matching. *J. Pet. Sci. Eng.*, vol. 184, p. 106559, 2020, doi: 10.1016/j.petrol.2019.106559.

[18] A. Ali and L. Guo. Adaptive neuro-fuzzy approach for prediction of dewpoint pressure for gas condensate reservoirs. *Pet. Sci. Technol.*, vol. 38, no. 9, pp. 673–681, 2020, doi: 10.1080/10916466.2020.1769655.

[19] S. Mishra, K. R. Bukkarapu, and A. Krishnasamy. A composition based approach to predict density, viscosity and surface tension of biodiesel fuels. *Fuel*, vol. 285, p. 119056, 2021, doi: 10.1016/j.fuel.2020.119056.

[20] T. Bikmukhametov and J. Jäschke. First principles and machine learning virtual flow metering: A literature review. *J. Pet. Sci. Eng.*, vol. 184, p. 106487, 2020, doi: 10.1016/j.petrol.2019.106487.

[21] A. T. Tunkiel, D. Sui, and T. Wiktorski. Data-driven sensitivity analysis of complex machine learning models: A case study of directional drilling. *J. Pet. Sci. Eng.*, vol. 195, 2020, doi: 10.1016/j.petrol.2020.107630.

[22] X. Liu *et al.*. Automatic well test interpretation based on convolutional neural network for infinite reservoir. *J. Pet. Sci. Eng.*, vol. 195, p. 107618, 2020, doi: 10.1016/j.petrol.2020.107618.

[23] E. D. Attanasi, P. A. Freeman, and T. C. Coburn. Well predictive performance of play-wide and Subarea Random Forest models for Bakken productivity. *J. Pet. Sci. Eng.*, vol. 191, p. 107150, 2020, doi: 10.1016/j.petrol.2020.107150.

[24] Z. Li *et al.*. Semi-supervised learning for lithology identification using Laplacian support vector machine. *J. Pet. Sci. Eng.*, vol. 195, p. 107510, 2020, doi: 10.1016/j.petrol.2020.107510.

[25] J. N. Kutz, X. Fu, S. L. Brunton, and N. B. Erichson. Multi-resolution dynamic mode decomposition for foreground / background separation and object tracking. 2015 EEE International Conference on Computer Vision Workshop, pp. 1-9, doi: 10.1109/ICCVW.2015.122.

[26] M. H. Naderi, H. Eivazi, and V. Esfahanian. New method for dynamic mode decomposition of flows over moving structures based on machine learning (hybrid dynamic mode decomposition). *Phys. Fluids*, vol. 31, no. 12, 2019, doi: 10.1063/1.5128341.

[27] M. R. Jovanović, P. J. Schmid, and J. W. Nichols. Sparsity-promoting dynamic mode decomposition. *Phys. Fluids*, vol. 26, no. 2, pp. 1–22, 2014, doi: 10.1063/1.4863670.

[28] K. Fujii and Y. Kawahara. Supervised dynamic mode decomposition via multitask learning. *Pattern Recognit. Lett.*, vol. 122, pp. 7–13, 2019, doi: 10.1016/j.patrec.2019.02.010.

[29] J. N. Kutz, X. Fu, and S. L. Brunton. Multiresolution dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.*, vol. 15, no. 2, pp. 713–735, 2016, doi: 10.1137/15M1023543.

[30] P. J. Schmid. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.*, vol. 656, pp. 5–28, 2010, doi: 10.1017/S0022112010001217.

[31] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: Data-driven modeling of complex systems*. First Edition, pp. 1-22, 2016. Philadelphia, SIAM.

[32] C. W. Rowley, I. Mezic, S. Bagheri, P. Schlatter, and D. S. Henningson. Spectral analysis of nonlinear flows. *J. Fluid Mech*., pp. 1-13, 2009, doi: 10.1017/S0022112009992059.

[33] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley. A data-driven approximation of the koopman operator: extending dynamic mode decomposition. *Journal of Nonlinear Science*, vol 25, pp.1307–1346, 2015.

[34] J. N. Kutz. *Data-driven modeling & scientific computation*. First Edition, pp. 506-515, 2013. Oxford, Oxford University Press.

[35] I. U. Haq, K. Fujii, and Y. Kawahara. Dynamic mode decomposition via dictionary learning for foreground modeling in videos. *Comput. Vis. Image Underst*., vol. 199, no. 6, pp 1-10, 2020, doi: 10.1016/j.cviu.2020.103022.

[36] T. Murata, K. Fukami, and K. Fukagata. Nonlinear mode decomposition with convolutional neural networks for fluid dynamics. *J. Fluid Mech*., pp. 1–15, 2019, doi: 10.1017/jfm.2019.822.

[37] E. M. Ramos, G. M. Darze, F. R. T. do Nascimento, J. L. H. Faccini, and G. A. Giraldi. Comparison of dynamic mode decomposition and deep learning techniques for two-phase flows analysis. *Flow, Turbulence and Combustion*, pp. 1-35, 2020, doi:10.1007/s10494-020-00151-z.

[38] Y. Susuki and K. Sako. Data-based voltage analysis of power systems via delay embedding and extended dynamic mode decomposition. *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 221–226, 2018, doi: 10.1016/j.ifacol.2018.11.705.

[39] S. M. Hirsh, B. W. Brunton, and J. N. Kutz. Data-driven spatiotemporal modal decomposition for time frequency analysis. *Appl. Comput. Harmon. Anal*., vol. 49, no. 3, pp. 771–790, 2020, doi: 10.1016/j.acha.2020.06.005.

[40] A. Bao, E. Gildin, A. Narasingam, and J. S. Kwon. Data-driven model reduction for coupled flow and geomechanics based on DMD methods. *Fluids*, vol. 4, no. 3, pp. 1-22, 2019, doi: 10.3390/fluids4030138.

[41] N. B. Erichson, S. L. Brunton, and J. N. Kutz. Compressed dynamic mode decomposition for background modeling. *J. Real-Time Image Process*., vol. 16, no. 5, pp. 1479–1492, 2019, doi: 10.1007/s11554-016-0655-2.

[42] S. L. Brunton, J. L. Proctor, J. H. Tu, and J. N. Kutz. Compressed sensing and dynamic mode decomposition. *J. Comput. Dyn.*, vol. 2, no. 2, pp. 165–191, 2015, doi: 10.3934/jcd.2015002.

[43] G. Isabelle, G. Steve, N. Masoud and Z. Lotfi (Eds.). *Feature Extraction Foundations and Applications*. Springer-Verlag, Berlin, 2006. doi:10.1007/978-3-540-35488-8.

[44] Lie, K-A. *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST)*. First Edition, pp. 231-246, 2019. Cambridge, Cambridge University Press. doi: doi: 10.1017/9781108591416.

[45] A. S. Odeh. Comparison of solutions to a three-dimensional black-oil reservoir simulation problem. *J Pet Technol*, vol. 33, no. 01. pp. 13 - 25, 1981, doi: 10.2118/9723-PA.

[46] D. E. Kenyon and G. A. Behie. Third SPE Comparative Solution Project: Gas Cycling of Retrograde Condensate Reservoirs. *JPT, J. Pet. Technol*., vol. 39, no. 8, pp. 981–997, 1987, doi: 10.2118/12278-PA.

Chapter Seven: Conclusions and Future Work

7.1 Introduction

This study illustrates the effectiveness of machine learning data-driven based methods in modelling and simulation of pressure and fluid flow dynamics of hydrocarbon reservoirs. Data driven methods are gaining popularity among researchers in the oil and natural gas industry due to their ability to capture nonlinear behaviour of high-dimensional parameters that are difficult to discover directly from complex dynamical subsurface systems. This study aims to develop data-driven machine learning methods for improving the prediction of pressure and fluid flow dynamics in natural gas reservoirs. First part of the study investigates the effect of BHP on production performance of a single-phase gas condensate reservoir model (chapter 3). Second part improves the dewpoint pressure predictions for gas condensate reservoirs over traditional techniques (chapter 4). Next, the study examines the pressure behaviour of underground natural gas storage in depleted reservoir model as natural gas is being injected/stored in the reservoir (chapter 5). Finally, the effects of gas injection and water flooding on pressure and fluid flow dynamics of multiphase reservoir models were investigated (chapter 6). As strengths and weaknesses are better understood from these methods, this chapter summarises the major findings, pinpoints the notable drawbacks of these methods, and highlights directions for future work.

7.2 Summary

The objectives of this study were achieved by conducting different studies whose findings were presented in chapters three to six. In chapter three, an AI-based model that captures the pressure dynamics and gas production trend under different conditions of varied BHPs for single-phase gas condensate reservoir model has been developed. Findings of this chapter reveals that at low BHP, the reservoir model under study produces high volume of gas at earlier production stage, however, as production continuous over time, the rate of production decreases. Also, findings show that at low BHP the reservoir experiences a significant decline in pressure. The physics-informed AI-based model developed was able to capture the pressure dynamics and gas production trend with better accuracy in comparison to other state of the art methods. A notable drawback of the AI-based model developed in chapter three is

its reliance on large number of rules for better accuracy which in turn, leads to a computationally intensive model.

In chapter four, an AI-based model that improves the prediction of dewpoint pressure for gas condensate fluids has been developed. Comparing the performance of the proposed AI-based model with traditional methods reveals that the developed model improves the dewpoint pressure prediction by achieving up to 99% accuracy when compared to other traditional methods. To address the drawback noticed in chapter three, the AI-based model developed in chapter four used subtractive clustering technique to identify an initial system based on extracted formations or groupings from the dataset after which training, validation, testing of the final model were performed using different partitioned datasets. Thus, the AI-based model presented in chapter four can be described as an extended version of the one presented in chapter three.

In chapter five, a data-driven method that decomposes and reconstructs a pressure field of a depleted reservoir model that mimics the behaviour of an underground natural gas storage is proposed. The data-driven model developed in chapter five is based on DMD approach. To evaluate the effectiveness of the proposed DMD model, three experiments were performed using different number of DMD modes on different datasets of varied rock properties generated from numerical simulations. Findings show that the proposed DMD model performed better with the least error in reconstructing the high-dimensional pressure field and approximating average reservoir pressure using 15 number of modes. A notable drawback of DMD is the computation of singular values performed on the high-dimensional full-state matrix which is computationally burdensome. Also, it is observed that in standard DMD, better results are obtained when the number of columns of the full-state matrix is less than the number of rows. On the contrary, when the number of columns of the full-state matrix is greater than the number of rows, part of the columns that carries spatial information might be lost in the process of implementing the method.

To address the drawbacks of DMD noticed in chapter five and even that of unpredictability of reproducing results noticed in compressed DMD (an extended version of DMD), this study developed a novel method called Dynamic Mode Learning (DML) and presented it in chapter

six. The new DML model is developed in the form of a learning system by first, constructing a simple, fast and efficient learning system that extracts important features from the original high-dimensional data and places them in a low-dimensional representation as extracted features. The extracted features are then used to reduce the original data after which dynamic modes are computed on the reduced data. By learning and extracting the features of the full-state data, the proposed DML method eliminates the need for computing dynamic modes on the full-state data as in standard DMD and rather, results in computing dynamic modes on a reduced data that contains the underlying dynamics of the full-state data. Furthermore, by using a learning system to extract features that are fixed and used them to reduce the full-state data, the proposed DML method also addresses the unpredictability of results that arise from randomness nature of measurement matrix as identified in compressed DMD. To evaluate the effectiveness of the proposed DML method, data were obtained from numerical simulations of SPE benchmark reservoir models. Experiments performed on pressure field and fluid phase distribution datasets generated from direct numerical simulations of two multiphase reservoir models show that the proposed DML method outperforms both standard and compressed DMD in terms of reservoir pressure field reconstruction, average reservoir pressure approximation, water saturation field reconstruction, and gas saturation field reconstruction.

7.3 Limitations

Even though the DML method developed in chapter six performed better over the state-of-the-art methods on different datasets, nevertheless, the method lacks some features as listed below.

Given that the success of the DML method depends on the performance of the initial learning system that extracts the features of the high-dimensional data, the proposed method might not be effective when it encounters noisy data or data with missing values. In other words, the initial learning system performed well because in this study, the data were generated from direct numerical simulations, as such, the data might not be noisy and contain no missing values. However, in a situation where the data are obtained experimentally or historically and contain some noise or missing values, the proposed DML method might not be effective due

to the fact that the initial learning system development process neglects data cleaning which is an important aspect in machine learning model development process.

Unlike other modal decomposition methods such as PCA and POD where components are in order of variance significance and energy dominance, respectively, In DML method, there is no unique correct manner to rank eigenvalue importance. Thus, it might be difficult to determine which modes are the most physically relevant.

The new DML method was developed and tested on domains that involve spatiotemporal data to identify dynamics. In such domains, observations/measurements are made at neighbouring locations/spaces and time stamps are mutually correlated with each other. Thus, in a domain where independence exists among observations/measurements and time stamps are not correlated, the new DML method might not produce the desired results.

7.4 Directions for Future Work

- Incorporate the aspect of data cleaning in the process of developing the initial learning system for DML method. This will delimit the application of the new method and improves its generalisability even if it encounters noisy data or data with missing values obtained experimentally or historically.
- Examine the capability of DML method to identify and extrapolate the dynamics of spatiotemporal systems from transient behavior to permanent, and constructing efficient data-driven reduced order models for spatiotemporal systems.
- Examine the capability of the developed DML method to develop predictive models that incorporate both temporal and spatial information to model relationships among spatiotemporal variables applicable to domains of practical applications such as traffic dynamics, climate science, neuroscience, environmental science, health care, social media, among others

Appendix A: Nomenclature

CO_2 : Carbon dioxide

NOx : Nitrogen oxides

SOx : Sulphur oxides

A_x : Cross sectional area normal to x-direction (m^2)

A_y : Cross sectional area normal to y-direction (m^2)

A_z : Cross sectional area normal to z-direction (m^2)

g : gas phase (fraction)

o : oil phase (fraction)

w : water phase (fraction)

k_{rg} : relative permeabilities of gas (fraction)

k_{ro} : relative permeabilities of oil (fraction)

k_{rw} : relative permeabilities of water (fraction)

μ_g : gas viscosity (cP)

μ_o : oil viscosity (cP)

μ_w : water viscosity (cP)

B_o : oil formation volume factor (bbls/STB)

B_g : gas formation volume factor (ft^3/scf)

B_w : water formation volume factor (bbls/STB)

S_g : gas saturation (fraction)

S_o : oil saturation (fraction)

S_w : water saturation (fraction)

q_g : gas production rate (MMcf/d)

q_o : oil production rate (MMbbl/d)

q_w : water production rate (MMbbl/d)

P_o : oil pressure (bar)

P_g : gas pressure (bar)

P_w : water pressure (bar)

ρ_g : gas density (kg/m^3)

ρ_o : oil density (kg/m^3)

ρ_w : water density (kg/m^3)

C_1 : methane

C_2 : ethane

C_3 : propane

iC_4 : isobutane

nC_4 : butane

iC_5 : isopentane

nC_5 : pentane

C_6 : hexane

C_{7+} : heptane plus

N_2 : Nitrogen

H_2S : Hydrogen sulfide

SG : specific gravity

MW : molecular weight (g/mole)

V_{rel} : relative volume (m^3)

V_t : total volume (m^3)

V_{sat} : volume at saturation pressure (m^3)

P : pressure (bar)

R : universal gas constant (m^3 bar K−1 mol−1)

T : reservoir temperature (^{0}F)

V : molar volume (m^3)

y_i : mole fractions of vapour phase (fraction)

x_i : mole fractions of liquid phase (fraction)

P_d : measured dewpoint pressure (psi)

$\hat{P}_d$: predicted dewpoint pressure (psi)

f_i^V : vapour fugacity (bar)

f_i^L : liquid fugacity (bar)

φ_i^L : liquid fugacity coefficient

φ_i^V : vapour fugacity coefficient

K_i : K-values

mD : MilliDarcy

Cp : centiPoise

^{0}F : degree Fahrenheit

ε : porosity (percentage)

v : flow velocity (m^3/s)

K : permeability (mD)

PI : production index

R^2 : coefficient of determination

N : total number of data points

Z_i : mole fraction of component i

γC_{7+} : specific gravity of heptane plus

MW_{C7+} : molecular weight of heptane plus

r_a : positive constant

$O_{1,i}$: fuzzification layer

$O_{2,i}$: rule layer

$O_{3,i}$: normalization layer

$\tilde{\omega}_i$: node's output of component i

$O_{4,i}$: defuzzification layer

$O_{5,i}$: output layer

ϕ_D DMD mode

F_i : extracted feature i

W_i : weight matrix

b_i : bias vector

ϕ_L : DML mode