KHAZAR UNIVERSITY

School: Graduate School of Science, Art and Technology

Department: Petroleum Engineering

Specialty: Development of Oil and Gas Fields

MASTER'S THESIS

MACHINE LEARNING-BASED RESERVOIR PERFORMANCE MODELLING AND OPTIMIZATION FOR IMPROVED PRODUCTION FORECASTING AND DECISION-MAKING

Student: _____ Khalid Orujov Nasraddin

Supervisor: _____ Assoc. Prof. Dr. Grigorii Penkov

BAKU – 2025

XƏZƏR UNİVERSİTETİ

Fakültə: Təbiət elmləri, Sənət və Texnologiya yüksək təhsil

Departament: Neft-qaz Mühəndisliyi

İxtisas: Neft və qaz yataqlarının işlənməsi

MAGİSTR DİSSERTASİYA İŞİ

TƏKMİLLƏŞDİRİLMİŞ İSTEHSALIN PROQNOZLAŞDIRILMASI VƏ QƏRARLARIN QƏBULU ÜÇÜN MAŞIN TƏLİMİNƏ ƏSASLANAN YATAQ PERFORMANSININ MODELLƏŞDİRİLMƏSİ VƏ OPTİMALLAŞDIRILMASI

İddiaçı: _____ Xalid Orucov Nəsrəddin oğlu

Elmi Rəhbər: _____ Dosent Dr. Grigorii Penkov

BAKI – 2025

TABLE OF CONTENTS

INTRODUCTION

Actuality is that the oil and gas industry has been impacted by digitalization and artificial intelligence. Examples of the applications of these technologies include well drilling, predictive maintenance execution, and the establishment of digital fields. For several decades, numerical reservoir models and "digital twins" have been employed to estimate hydrocarbon production volumes. Conversely, recent advancements in computational power and artificial intelligence have enabled oil and gas companies to create "digital twins" of reservoirs, which are model ensembles that encompass the uncertainty range in static data (including petrophysics and geological structure), dynamic data (such as oil or gas properties), and economic factors (such as capital and operating expenditures). The application of machine learning and artificial intelligence enhances hydrocarbon production predictions under uncertainty. This is achieved by calibrating the model ensembles to the observed data. Estimating and predicting petroleum production is a significant difficulty for the upstream petroleum sector. The forecasting of production resources can assist engineers in multiple ways, including the analysis of production impacts, the preparation of schedules, the allocation of resources, and the formulation of project decisions. Nonetheless, the intricacies of data, coupled with restricted analytical insights, render the endeavor exceedingly formidable (Sagheer & Kotb, 2018). Modelling different scenarios may require an excessive amount of time, potentially leading to missed opportunities (Sagheer & Kotb, 2018). Hydrocarbon production forecasting includes estimating ultimate recoverable reserves and calculating oil production profiles, which are essential for business planning, asset appraisal, and decision-making in the oil and gas sector. Methods such as Decline Curve Analysis, Computer Simulations, Material Balance, Volumetric Calculations, and Pressure Transient Analysis have been utilized in various scenarios to accomplish production forecasting objectives. The advent of big data and advancements in computing technology have facilitated the creation of data-driven methodologies to address challenges in the oil and gas sector.

The implementation of enhanced data management protocols has led to increased interest and utilization of machine learning models across multiple sectors, including oil and gas. Production forecasting is characterized by a time series examination of the well's output rate. This is regarded as a time-dependent issue. Time series forecasting entails anticipating future behaviours of systems. It relies on historical data and the present condition of the system (Sagheer, 2018). Statistical

approaches have been implemented to predict oil production rates from the wells. Techniques like ARIMA have been implemented to establish a balanced, precise, and dependable approach for forecasting petroleum output (Ediger, et al., 2006). The nonlinear characteristics of the production data and features necessitate the utilization of nonlinear time series models. Various neural network models have been examined to simulate the production behaviour of a reservoir due to its nonlinear dynamics. Machine learning techniques, including Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU), have been utilized for forecasting (Sagheer & Kotb, 2019; Alimohammadi & Rahmanifard, 2020). These studies have demonstrated the viability of deep neural networks. Traditional time-series forecasting techniques such as Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), and Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) have been effectively utilized for predicting petroleum production. They excel in predicting unconventional gas and oil reservoirs characterized by regular shifting patterns (Gupta et al., 2014; Morgan, 2018). Gupta et al. (2014) endeavoured to deploy neural networks (NN) to predict production rates in unconventional resources. The reduction in production is recorded during the neural networks training process and was utilized in the production data during the forecasting phase. These models were chosen because to their low data intensity and their capacity for automation, allowing application across multiple wells. In 2021, Al-Shabandar et al. utilized a deep gated recurrent neural network for forecasting petroleum production.

The proposed model featured a low-complexity design and the ability to monitor long-interval time series datasets. Both numerical and categorical characteristics were utilized to predict oil output. Empirical testing demonstrated that Deep Gated Neural Networks might yield satisfactory results for a time series forecasting challenge. The results indicate that Deep Gated Neural Networks get superior computational efficiency regarding training duration. The proposed model demonstrated superior performance compared to existing common models.

The purpose of the research is to apply the Machine Learning model to determine the production trend of the well based on different influencing parameters during this thesis. This study presents the development of prediction models utilizing statistical approaches, tree-based ensembles (XGBoost) and a deep learning algorithm, specifically LSTM, along with an evaluation and comparison of their applications and results. The efficacy of statistical approaches was assessed in comparison to tree-based ensemble ML model, XGBoost and the deep learning model derived from LSTM.

The objectives of the research are to determine the production trend of the reservoir and the changes through the life of the field will be a part of the evaluation of this thesis. During the study the following objectives are tried to be achieved:

- The review of the different papers related to production forecasting techniques and Machine learning applications for forecasting well performance.
- Preparing the methodology of the research for data collection and processing, applying ML for production forecasting.
- Increasing the accuracy of the LSTM model for forecasting and comparing it with the real production data and trivial methods.
- Applying XGBoost model for comparing the results obtained through LSTM model and selecting optimal model for reservoir management and decision-making.

CHAPTER I. LITERATURE REVIEW

Machine Learning Techniques for Oil Well Production Forecasting

Forecasting the production of oil well has always been a challenge because of the complexity of geological heterogeneity, nonlinearity in reservoir dynamics, and variety in conditions of the operations. Trivial methods like DCA (decline curve analysis) and numerical models for reservoirs simulation face with these issues and these complexities demand ML techniques which are able to model nonlinear correlations and handle big datasets.

Modern studies have distributed that models of machine learning may vividly improvise the accuracy in forecasting through the integration of various types of data, such as production history, features of the geology, inputs of completion design, and variations of operations.

It is a good illustration that one of ensemble-based ML models has been developed through the utilization of data of 80000 wells for unconventional reservoirs in Northern America, by integrating data related to geology, completion and production history for the forecast. This technique performed better in comparison to trivial techniques, specifically for the wells whose have production history no more than 1 year via the capture of the trends of the decline and modifications in the operations. The models of ensemble demonstrated robustness in prediction of not only short-term, but also long-term profiles of production, highlighting the positives sides in the integration of multi-source data into the frameworks of ML.

In a similar manner, the framework for the double-stage forecasting has been suggested that wells have been classified into low and high-yield selections on the basis of cumulative production, followed by the application of dedicated models of regression to each selection for enhancing the accuracy of the prediction. This technique addresses the essential variations amongst the wells formed due to the differences in geology and development strategy, which are the factors that are often overlooked in common models.

Deep Learning Models

The techniques of Deep Learning, specifically RNNs (recurrent neural networks) and their types like LSTM (long short-term memory) networks, have been popular in the forecast of oil production because of the ability of theirs for modelling temporal correlations of sequential data of production.

In one of the studies, which has applied Aramco's dataset in the building deep learning and ML models, such as RNN, ANN, XGBoost, by succeeding the R^2 scores of 0.98, 0.97 and 0.96, accordingly. These outputs show the models of deep learning may replicate the results of reservoir simulation closely, by lowering the time for the computation very drastically from hours, even days to several minutes. This makes these models heftily suitable for the rapid forecast of the production and making decisions.

Another integrated model, which is the combination of deep learning and other techniques of ML, was built for improvising the efficiency and accuracy of the forecasting. Feature extraction and prediction of time-series were integrated into this model, by distributing superb results in the capture of the complexity of production dynamics.

1.1. IMPACT OF DIGITALIZATION IN FORECASTING PRODUCTION AND ANALYSIS OF PROJECT DECISIONS

Digitalization has shown its significant impact on production forecasting and making decisions (Clemens & Viechtbauer-Gruber, 2020) in the projects.

According to Accenture (2017), forty percent of upstream oil and gas businesses are concerned that they may run out of competition in the digital race. The other side of the coin is that businesses that can successfully build advanced data analytics methodology and possess deep analytics expertise that is supported by a specialized talent strategy are outperforming their competitors (Bughin et al. 2017, Bisson et al. 2018). Companies in the oil and gas industry have been impacted by digitalization in a variety of ways. The transition from corrective to predictive maintenance was accomplished with the help of big data analytics, and the utilization of data analytics has resulted in improvements to offshore operations. Additionally, digitalization is influencing drilling operations, seawater treatment, seismic interpretation, and subsea operations. The use of digital twins (Poddar 2018) in offshore developments and the implementation of digital oilfields (Salman et al. 2018) are two further topics. The oil and gas industry has also adopted computerized analytical methodologies in the process of forecasting the production of hydrocarbons (Shahkarami et al. 2014, Eltahan et al. 2019). According to Lucchesi (2019), the production of hydrocarbons has a significant influence on the economics of oil and gas projects. Several different approaches are utilized by oil and gas businesses to forecast the production of hydrocarbons. A Value of Information (Vol) analysis is carried out (Steineder et al. 2019) and field development options are

selected (Mantatzis et al. 2019) based on the hydrocarbon production forecast as well as other economic parameters such as the price of oil, capital expenditures (CAPEX), operating expenditures (OPEX), and the tax regime.

In the below figure, different levels for decision analysis and the forecast of production have been provided. Level 0 means no automation, while level 5 means full automation.

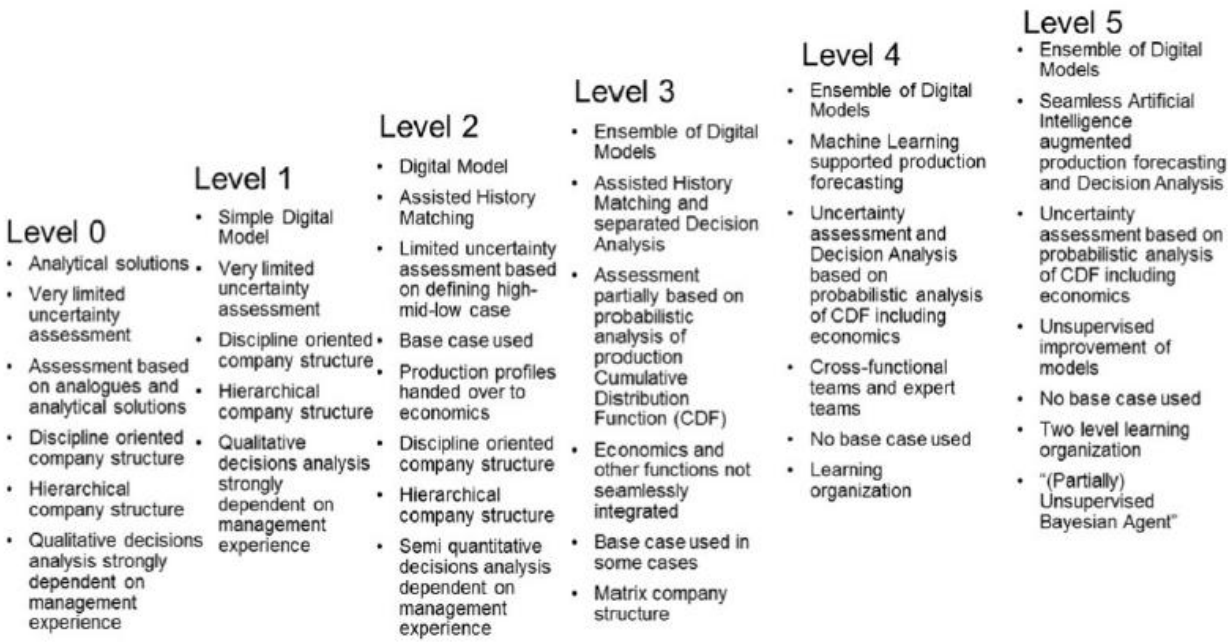

Figure 1.1.1. The levels of production forecasting & decision analysis

1.1.1. Level 0

The application of analytical solutions to the forecasting of hydrocarbon production gives this level its distinctive characteristics (Crawford 1960). The structure of oil and gas businesses was traditionally hierarchical (Tannenbaum et al. 1974) and discipline-oriented; hence, decision analysis was heavily dependent on the expertise of the teams engaged and management.

The application of general physics and engineering concepts to reservoirs is typically the basis for making judgments regarding field development (Buckley and Craze 1943). It is important to consider the costs of various appraisal methods when determining how to develop fields. Because of the constraints imposed by the analytical solutions, the uncertainty that is associated with the various development possibilities is not taken into consideration quantitatively.

1.1.2. Level 1

The investigation of many processes, such as in-situ combustion (Coats 1980), which are difficult to cover by analytical solutions, was carried out with the use of numerical models (Crookston et al. 1979). It is via the incorporation of the outcomes of the numerical models that decisions are formed (Bennett 1981). The limited processing capacity that is available results in a significant simplification of geological data (Poston and Gross 1986). There is integration of several disciplines; nevertheless, the influence of uncertainty on decisions is not quantified within this framework. In certain instances, several geological scenarios are taken into consideration and examined to determine the impact that they have on the development options available (Khandwala et al. 1984). The scope of decision analysis is restricted, and it is distinct from other fields of study. There is a hierarchical structure within the organization, and the decisions that are made are heavily influenced by the level of expertise that the associated team and management possess.

1.1.3. Level 2

As the computing power of computers continues to improve, digital models are being utilized more frequently than they were in Level 1. The performance of GPUs and other specialized processors has surged significantly, with GPU performance apparently escalating by about 7,000 times since 2003. This increase in power facilitates the implementation of intricate algorithms and extensive data processing, which are crucial for contemporary digital modeling methodologies. Consequently, activities that were once computationally prohibitive are now achievable, enabling the development and use of more advanced models.

Based on the advances in processing power, integrated studies (Bastian et al. 1998) are utilized for assisted history matching (Grussaute & Gouel 1998) and enhanced workflows. An example of this may be seen in Figure 1.1.3.1 (Chiotoroiu et al. 2013). A comprehensive data analysis that is based on analytical solutions and data visualization is the first step in the workflow that is illustrated in Illustration 2. After that, a stage that involves a global history match is carried out, and feedback loops are utilized in the numerous subsurface disciplines. Once a "history match" with the global parameters (such as reservoir pressure and field oil production) has been accomplished to a satisfactory level, diagnostic diagrams are utilized to determine which wells are the most significant to match. Through the utilization of streamline modeling, it is possible to make "minimal invasive" adjustments to the reservoir parameter to produce local matches between the wells. The quality of

the history match can be evaluated spatially, and if the result is poor, modification of the geological model can be implemented if the evaluation is unsuccessful.

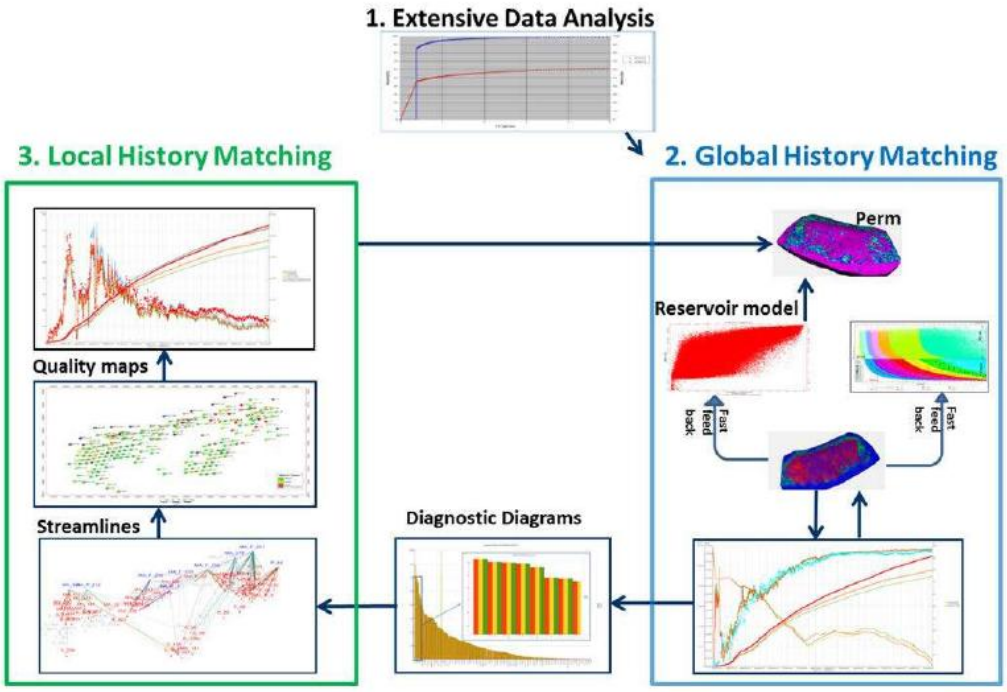

Figure 1.1.3.1. Workflow of Level 2

The ability to comprehend the physical and chemical processes involved in the feedback loops and data analysis are among the most important abilities that are required. The explanation behind this is that a significant number of parameters need to be handled and altered on a consistent basis. In addition, the ability to analyze data and use a computer is required to assist the discussion of many disciplines, as well as to consistently characterize and simulate the reservoir. The goal is to locate a model of the subsurface that is consistent and that corresponds to the data that has been measured. This is done with the intention of producing a "digital twin" of the subsurface. According to Poddar (2018), a digital twin is defined as "a virtual and simulated model of a physical asset that can be used for various purposes." When it comes to subsurface modelling, one of the instances of how a digital twin can be utilized is to investigate the impacts of varying operating conditions or drilling additional wells on the production of hydrocarbons. There are other sections of the oil and gas business that have effectively implemented digital twins, such as the simulation of surface physical and chemical processes or the simulation of plants (Nixon, and Pena 2019). According to Scheidt

et al. (2018), the utilization of a subsurface "digital twin" presents a variety of challenges. One of these challenges is that a plethora of alternative parameter combinations can lead to an acceptable "history match," but they also result in varied forecasts. Level 2 firms invest significant effort in creating a "digital twin" to showcase it in peer evaluations at stage-gate processes. The "digital twin" undergoes continuous updating as it is manipulated by the data collected over consecutive years. (Ibrahimov 2015). Limited uncertainty assessment based on high-mid-low situations is a common characteristic of Level 2, which is typically defined by this. The economics department receives hydrocarbon production profiles from the subsurface teams so that they can do decision analysis on them respectively. The restricted number of simulations and subsequent economic evaluation of each of the production profiles are the foundations upon which decision analysis is built (Evens 2000). It is necessary to have feedback loops going to both the surface and subsurface teams after the economic evaluation has been completed. The decisions are made based on the limited information that is contained in high-mid-low cases and the examination of economics separately. Since this information is not included in high-mid-low scenarios, it is not possible to conduct a quantitative study of the impact that different factors have on the development of alternatives. When it comes to field development planning, decisions are made based on the experience of the technical team and management. The outcomes of numerical modelling are utilized as a source of guidance and to calculate deterministic high-mid-low Net Present Values (NPVs) for the various development alternatives. Expected Value in Terms of Money (EMV) is only estimated in certain circumstances by employing methods such as Swanson's mean, which are approximations of the NPV distributions (Hurst et al. 2000). Field development decisions are heavily influenced by the collective experience of the individuals participating (Malhotra et al. 2004). This is because the approach has inherent flaws that make it difficult to implement. It is the goal of businesses to transition from rigid hierarchical structure to flatter hierarchies to progress from Level 1 production forecasting to Level 2 production forecasting. Flatter companies allow for more open lines of communication and collaboration, as well as enhanced involvement of teams in the decision-making process. As a result, they can make decisions more quickly since there are less layers of hierarchy.

1.1.4. Level 3

This stage involves the generation of ensembles of digital models, which are then utilized for the purpose of forecasting hydrocarbon output in the presence of uncertainty (Sieberer et al. 2019).

Figure 3 depicts an example of a workflow that fits this description. A geo-sensitive component is included in the workflow to capture static uncertainty, which encompasses a few different geological categories. Because of a dynamic reaction, the geological models are grouped together in a space that is multi-dimensional (Scheidt and Caers 2009). Different dynamic responses, such as tracers or flow pattern maps (Thiele and Baticky 2016), can be utilized to carry out the clustering process.

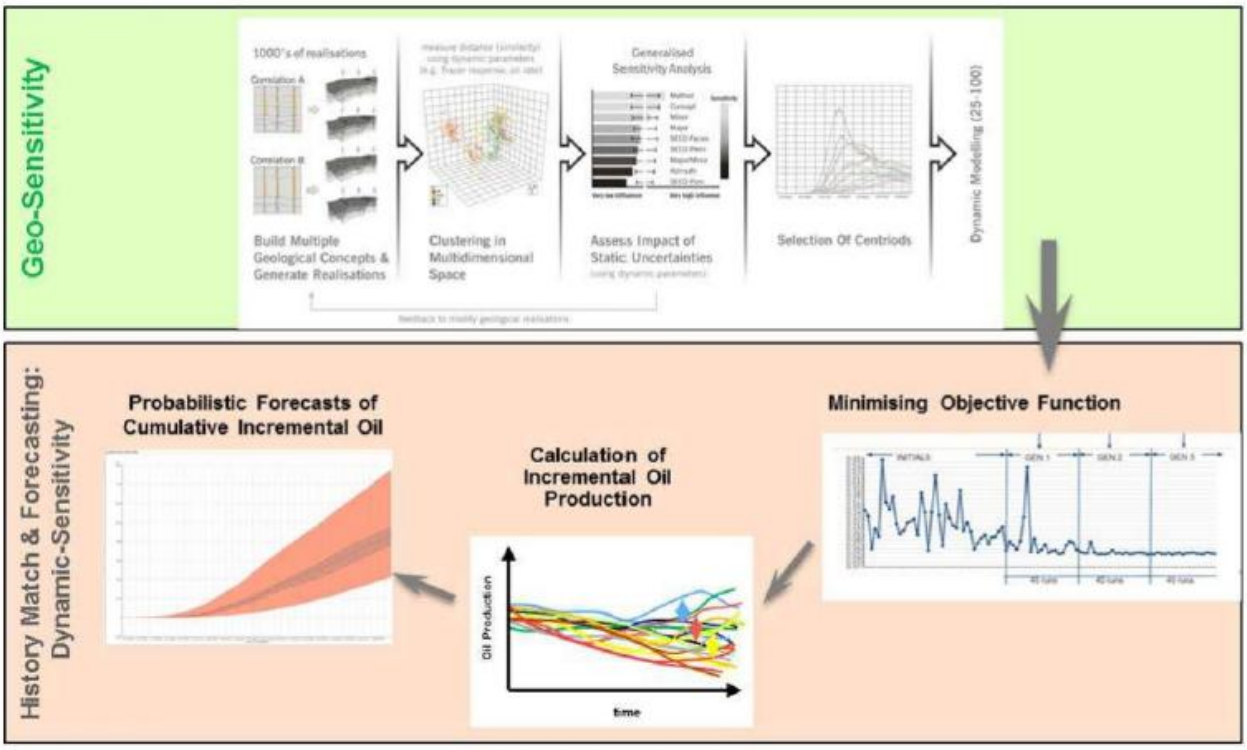

Figure 1.1.4.1. Example of workflow which consists of geo-sensitivity section

Since the clustering, a subset of the geological models that are representative of the geological uncertainty space is chosen. Additionally, the dynamic sensitivity portion of the workflow makes use of these models, which have been matched to their respective histories (Thiele and Batycky 2016). The evaluation of the physical and chemical description of the processes, the characterization of the reservoir, and the integration with several disciplines are all essential skills. In addition, it is required to evaluate the parameter ranges, and it is necessary to possess data science abilities to make use of the data that is produced by the simulations (clustering, distance-based generalized sensitivity analysis, evolutionary algorithms). The method of operation shifts from attempting to locate a calibrated "digital twin" of the subsurface, which is then utilized for forecasting, to gaining an awareness of the uncertainties and conditionalizing the models to the

data that has been observed. An alternative to the generation of a "digital twin" of the reservoir is the development of "digital siblings" that embrace the ambiguity that is present in the data. The term "digital siblings" refers to a collection of models that take specific measurements into consideration, such as pressures or manufacturing data. In contrast to "digital twins," "digital siblings" are not attempting to locate a precise digital representation of the subsurface; rather, they are attempting to evaluate the ambiguity while understanding that the challenge of matching the history is an ill-posed problem. In Level 3 companies, decision making is frequently kept distinct from the forecasting of hydrocarbon production. The production profiles are then given to the economics department so that they can analyse the number of different development choices. When it comes to improving resource allocation and increasing formal lateral communication, Level 3 organizations usually arrange their organizational structures in the form of matrix frameworks. The leadership of matrix organizations is comprised of both functional and project management, with the goal of ensuring that both technical and economic issues are addressed. According to Schnetler et al. (2015), some of the common issues that arise in such an organization include the need for a relatively large number of managers and the possibility of power disputes transpiring. The development of Net Present Value (NPV) Cumulative Distribution Functions (CDF) and the calculation of the Expected Utility (Begg et al. 2003) are two of the decision-making processes that are utilized by some Level 3 firms. These processes are utilized to take into consideration the risk attitude of the company (Sieberer et al. 2017). Because numerical modelling allows for the quantification of risk and uncertainty, these approaches result in quantitative decision making for field development planning. Calculating and optimizing the value of information is something that can be done (Steineder et al. 2019), and it is possible to quantify the impact that different factors have on decisions (Steineder and Clemens 2020). According to Mantatzis et al. 2019, decisions about field development planning are frequently made based on the Expected Utility that has been calculated. Management at the upper and top levels are responsible for making strategic decisions that cannot be measured.

1.1.5. Level 4

The utilization of probability distributions is the foundation for the fourth level of hydrocarbon production forecasts and decision analysis. Both theory-based models and data science models are utilized for the purpose of estimating hydrocarbon output, although the choice is ultimately determined by the availability of data.

The information that is contained in the data is what data-based models rely on. They require a substantial amount of data samples that are indicative of the whole. Text mining and object recognition are two examples of models that fall into this category (Karpatne et al. 2017). A plot of these kinds of models can be found in the bottom right corner of Figure 1.1.5.1. (Cao et al. (2016) and Kumar (2019)) both state that data-driven models are utilized for shale reservoirs in the process of projecting hydrocarbon production.

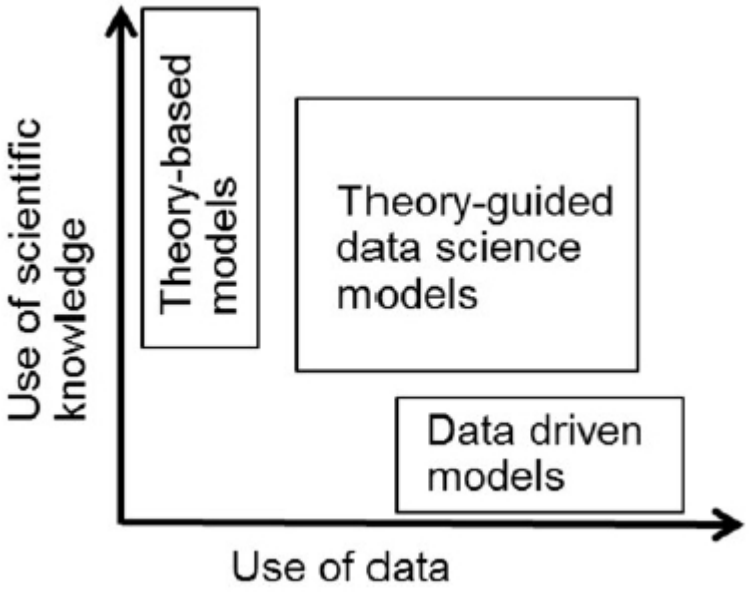

Figure 1.1.5.1. Picking between theory-based and data-driven models is contingent upon the level of scientific knowledge and the accessibility of pertinent data.

Theory-based models are derived from fundamental principles. Training examples can be generated using them to include input and output variables for the purpose of developing a data science model capable of extracting correlations between the variables (Karpatne et al. 2017). Models of this nature are shown in the upper left corner of Figure 1.1.5.1 Theory-guided data science models can be employed in situations where available scientific knowledge and data are inadequate (Karpatne et al. 2017).

The subsequent section provides a more comprehensive discussion on the integration of theory-based models with data science. Insufficient data is available for conventional hydrocarbon reservoirs to support the application of data-based models solely for hydrocarbon output forecasts.

A method within theory-based models is Bayesian Evidential Learning (BEL), which comprises the subsequent stages as outlined by Scheidt et al. (2018):

- Development of the question and specification of predictive variables
- Statement of the complexity and prior uncertainty of the model
- Monte Carlo simulation and data-driven falsification of prior uncertainty
- An investigation of sensitivity on both the data and prediction factors.
- A design for reducing the amount of uncertainty on prediction variables based on the measurements.
- Making decisions, as well as falsification and sensitivity of the posterior.

The quantitative study of a number of different possibilities is the foundation upon which Level 1.1.5.1 enterprises construct their decision-making processes for field developments (Steineder et al. 2019). For the purpose of ensuring that statistically significant numbers of models are developed and that these models may be updated in the event that new data becomes available, the quantitative analysis makes use of techniques such as machine learning.

Because decisions of this nature cannot be quantified, upper and top management are the ones who make significant strategic choices. Companies need to change from the view that people need to be directed and managed to give them explicit responsibility and authority in order to enable them to produce inventive solutions, according to Aghina et al. 2017's list of mind-set shifts that need to be done in order to enable data-centric decision making. Rapid decision-making and learning cycles need to be established, and technology needs to be viewed as a facilitator rather than a tool that supports it. Data democratization is essential in order to support the shift from Level 3 to Level 4, which means that teams need to have access to a significant volume of raw data. In order to eliminate silos and cut down on the amount of work that is duplicated, data democratization is essential (Yoder 2019, Anand and Krishna 2019). The integration of advanced analytics into the company is necessary in order to achieve major improvements in decision-making. According to Bisson et al. (2018), the characteristics that enable businesses to be ahead of their competitors (also known as "breakaway companies") in the process of creating value from digitization were investigated. They make the assertion that "breakaway companies are more likely to use sophisticated analytics techniques, such as reinforcement and deep learning, which can provide a substantial lift in value over using more traditional analytics approaches". Handscomb et al. (2016) and de Smet (2018) suggest that the organizational structure of businesses that are transitioning from Level 3 to Level 4 may evolve in the direction of organizational agility.

There has been a significant improvement in development cycles. Level 4 decision analysis

makes it possible to undertake integrated development planning, evaluation of uncertainties, and evaluation of the value of information analysis. This is in contrast to the traditional method of teams working in parallel and then developing fields. It is important to take into consideration the risk attitude of the person making the decision during this process (Steineder et al. 2019). According to Aghina et al. (2017), the democratization of data has allowed for the coverage of uncertainties in the subsurface, surface, and economic dimensions.

When it comes to decision making, a Level 4 organization has the potential to be changed into a Learning Organization. One definition of a Learning Organization is a business that provides its employees with opportunities to learn and that also undergoes constant self-improvement. Within the oil and gas business, there has been a shift toward learning organizations in the field of health, safety, and the environment. Additionally, there has been an improvement in technical abilities, which has included the use of systems thinking. Therefore, learning from previous decisions is difficult (Nandurdikar and Wallace 2011). This is because experience-based, leader-driven decision making makes it difficult to learn from past decisions. Due to the fact that such decisions are founded on quantitative data (for example, Tsuchiya 1993), level 4 decision analysis makes it possible for businesses to gain knowledge from their previous choices.

For the purpose of forecasting and making decisions regarding hydrocarbon output at the Level 4 level, profound technical capabilities are necessary. The reason for this is that the specification of the prior uncertainty ranges and the selection of the model are both quite important and depend on the abilities that the members of the team possess in terms of technology. When it comes to the use of theory-based models, the experts in the relevant fields (such as Special Core Analysis, PVT, Enhanced Oil Recovery, and Fractured Reservoirs) play a significant role. This is because these models are utilized to build the training sets for machine learning. Inaccuracies in the description of the chemical and physical processes lead to incorrect training of the machine learning models and forecasting that is not accurate.

In addition, extensive data science abilities are required in order to make use of the suitable approach and evaluate the outcomes. Overfitting, low-representation training due to insufficient data, spurious correlations, Ramsey-type correlations and the suitability of the instrumentation for the construction of interest are some of the potential pitfalls that need to be understood and avoided because they have the potential to result in incorrect decisions. According to the findings of Shepperd et al. 2019, out of the 49 papers that pertain to Machine Learning research, 22 of them have flaws that can be demonstrated. When it comes to the petroleum business, applying machine

learning takes a significant amount of expertise in order to avoid making conclusions that are incorrect.

1.1.6. Level 5

The process of decision making is increasingly being augmented by Artificial Intelligence (AI) with each passing level. A "Bayesian Agent" is an autonomous, goal-directed creature that observes and acts upon an environment (Russel and Norvig 2010). The corporations are developing this type of agent in order to better serve their customers. In the context of hydrocarbon production predictions and decision analysis, there is an illustration of a Bayesian Agent in Figure 1.1.6.1.

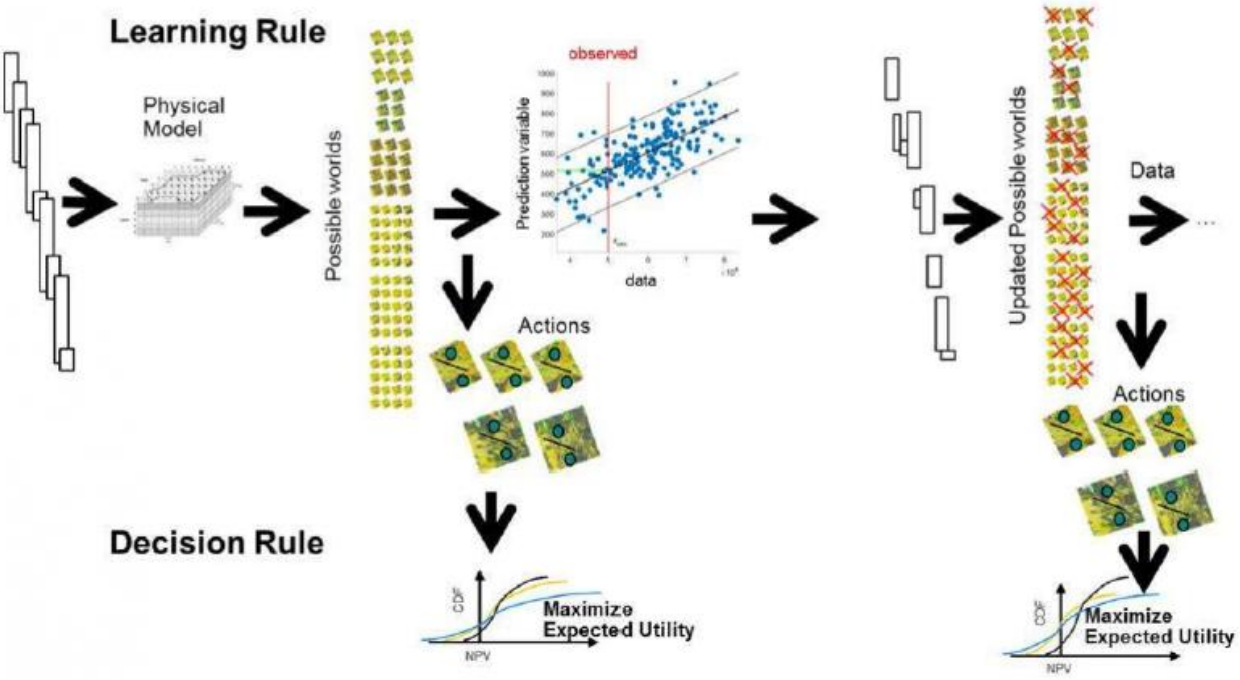

Figure 1.1.6.1. Bayesian Agent in the field of production forecasting & decision making.

Both a learning rule and a decision rule are what distinguish a Bayesian Agent from other agents. A number of different worlds that are feasible are defined by the learning rule, which is based on a variety of uncertain parameters. The conceivable worlds are realizations of the geology model, but they also include data on the economy and dynamics. In the realm of subsurface modelling, numerous worlds are formed through the utilization of numerical models (Reichstein et al. 2019). This is due to the fact that there is an insufficient amount of data from oil and gas fields to fully implement data science techniques. A number of different "worlds" are produced by the numerical models. Through the application of Al, the subsurface models, which include uncertainty, are

improved. Both generative adversarial nets (Goodfellow et al. 2014) and reinforcement learning (Silver et al. 2016) are examples of methods that could be considered in this category. The decision rule begins with the definition of actions and continues with the maximization of expected utility. The posterior distributions are constructed on the basis of further knowledge, and the possible worlds are conditioned to the data that has been observed. The anticipated utility is then maximized at that point. A comparison is made between the anticipated utility of a project conducted with and without the measurement of additional data in order to determine the value of the information. A number of tasks, including decision making, the optimization of activities, and the updating of conceivable worlds, are all carried out with the assistance of artificial intelligence. As a means of enhancing Expected Utility and decreasing economic risk, the Bayesian Agent makes recommendations regarding actions and procedures. Upper and Top Management are responsible for making strategic decisions that are not accounted for by the learning and decision rule.

There is the possibility of introducing Al enhanced portfolio optimization approaches in addition to the learning that occurs throughout the process of project creation. Agrawal and Jaiswal (2012) and Nowe et al. (2012) are two examples of how multi-agent systems can be utilized in conjunction with Game Theory (Nash 1950) to do optimization on project portfolios. As a result of the fact that these technologies require a significant proportion of the projects in the project portfolio to be covered by stochastic approaches that result in quantitative cumulative distribution functions rather than mid-high-low situations, Level 5 businesses will be able to implement them. This makes it possible to have a two-stage learning process, which includes learning within the creation of individual projects as well as making use of the learnings from the project portfolio to better individual projects. These kinds of organizations require skills that are comparable to those required by Level 4 organizations. For the purpose of enhancing the partially unsupervised Bayesian Agent, it is necessary to possess core scientific knowledge in order to be able to explain physical chemical processes in numerical models. Data labelling (and possible reinforcement learning to overcome this challenge), judging if a data set is sufficiently large and comprehensive, explainability of the Al results, generalization of the learning, distinction of causation from correlation, and bias in data and algorithms are some of the questions that need to be addressed in order to evaluate the limitations of the Bayesian Agent and the methodologies. In-depth data science skills are required in order to evaluate these limitations and address these questions (Chui et al. 2018).

Some of the difficulties associated with decision making that is aided by artificial

intelligence include "Reward Hacking" due to effects such as partially seen goals, abstract rewards, Goodhart's Law, or feedback loops, amongst other types of impacts (Amodei et al. 2016). It is possible for businesses to face risks in their financial performance, non-financial performance, legal and compliance, and reputation if they make mistakes in conceptualization, data management, model development, model implementation and model use, and decision-making processes. The increasing complexity of Al necessitates the development of substantial abilities in order to comprehend the patterns of behavior exhibited by these algorithms. In order to enhance the comprehension of artificial intelligence models, many techniques are utilized. These techniques include testing with Concept Activation Vectors, the utilization of tiny systematic perturbations (Ghorbani et al. 2019), and the utilization of Partial Dependence Plots (Guidotti et al. 2018). It is essential to use these techniques in order to comprehend the reasoning behind the decisions made by the agent system. Carvalho et al. 2019 provides a concise summary of the current state of interpretability in machine learning. In petroleum engineering, having the ability to understand Al choices is absolutely necessary in order to prevent a negative influence on the operation of the organization.

1.2. REVIEW OF INDIVIDUAL STUDIES.

The petroleum sector recently experiences the transformation stage from traditional techniques to integrated data-driven technologies, especially ML (machine learning). As the operations and explorations are getting more costly and complex, ML plays the role in the extraction of deep insights from the big datasets, automation of the time-taking workload, and enhancement of decision-making for the reservoir engineering. Amidst multitude of applications, machine learning demonstrates key roles in a couple of areas: forecast of production and decision making. The forecasting of the production is the key for management of the reservoir, by impacting on essential activities, like planning of the field development, estimation of the reserves, budgeting, and valuation of the assets. Trivially, models relying on physics and empirical methods like DCA and numerical simulators for the reservoir were significantly the main reference point. But the struggles the conventional techniques face are the heterogeneities, variations in operations, and the amount of data connected with mature fields and the unconventional types of the reservoirs.

Hence, machine learning has started to play the role as alternative approach, which is capable of capturing nonlinearity in the relationships, incorporation of various sources of data and improvisation of the accuracy of forecasting outputs.

The oil and gas operation demands very critical skills of decision-making from the optimization of well placement and the strategies of the completion to prioritization of scenarios of the development. The tools of ML are heavily applied to support in these scenarios based on the judgements of engineering. The advantage of this technique is that hidden patterns can be uncovered in historical data. The drivers of the production can be identified and improvised operational plans can be proposed, by assisting the engineers through making more detailed decisions more effectively and confidently. The review of the case studies covers 8 different applications of ML in real-life examples of machine learning applied in the forecasting of production and the process of decision-making. The papers cover different geological features from the geology of Middle East to Midland and Appalachian Basins, utilizes quite a big range of approaches of machine learning, ensemble models, neural networks and the interference of Bayesian. Detailed review per study, then the synthesis of themes, highlights the current situation of the research, determines the upsides and downsides of multiple methods and gives the direction for potential future work in this area.

1.2.1. Machine Learning for Performance Analysis in Carbonate Reservoirs

Huang et al. (2021) demonstrated an application of the ML techniques in the evaluation and optimization of the performance of the well for quite heterogeneous reservoir of carbonate in the Middle East. The research highlights the complexities in geology and operations in the typical carbonate formations, like porosity variations, complicated networks of the fracture, inconsistencies in the productivity of the well. In such reservoirs, for benchmarking and analysis of the production, by using trivial techniques through empirical rules or deterministic models, it is quite a big challenge to obtain a decisive pattern. In order to overcome this challenge, Huang et al. (2021) suggests a data-driven approach relying on PCA (principal component analysis) and k-means clustering, assisting for well classification based on their profiles of the production through the utilization of dynamic and static data. In the dataset, the combination of the geological features (permeability, porosity and net pay thickness), metrics of production (GOR, oil rate and water cut), and details of well geometry have been collected in the period of 15 years. The initial purpose of using PCA was to lower the dimensionality for 27-feature data, maintaining its main variance. The transition contributes to simplifying the task of clustering, through the projection of the data into smaller pieces of orthogonal axes which cover the key trends of the performance. The components which were reduced are followingly fed into the algorithm of K-means clustering to classify the

wells based on their performance. The actionable insights are offered by the resulting clusters. For instance, a group of wells could show the high production oil, but high value of water cut as it may have the connectivity of the reservoir to the aquifer or the poor design of the completion. Another group could distribute the stable performance in production with moderate gas rate, which is the indication of a more desirable segment of the reservoir. Through mapping the clusters spatially, the distribution of the performance across the field can be visualized and the outputs can be correlated with the features of the geology. This leads to a powerful tool which can be used to target the re-completions, redesigning the plans of drilling, and modifying the strategies of the lift. The essential part of this approach is to own the ability for the integration of disparate datasets to the framework, which is analytical and gives holistic diagnostics of the well, which are related to static and dynamics behavior. Furthermore, the need for the labeled data is removed by the unsupervised behavior of the method, which is generally not available in fields of legacy. This suggests scalability, too. The workflow can be broadened to up to thousands of wells in relatively low overhead of computation. But certain limitations exist in this study. As the major components are the combinations of input parameters. The transformation of PCA is effective for the reduction of dimensionality and may sacrifice interpretability to some extent. This makes the difficult the clusters to correlate with specific characteristics of the reservoir. Moreover, the descriptive insights are provided by clustering, hence it is not for predicting production or determining the uncertainty. This technique is consequently considered more diagnostically, and the quality and the completeness of the inputs impacts on the effectiveness of the model. In spite of those caveats, the valuable contribution of the study is highlighted through the demonstration of enhancement of the unsupervised learning on analysis of the performance development for the complex reservoir patterns. This gives the chance for wider utilization of clustering and reduction of the dimensionality in the workflows of the field optimization, specifically trivial techniques lacking. As long as the sector goes on growing with big dataset as of today, the essentiality of the frameworks will increase in order to obtain the useful pattern and assist to reach projected decision-making.

1.2.2. Bayesian Deep Decline Curve Analysis (BDDCA) for Production Forecasting

Tadjer and colleagues present (Tadjer et al. (2022)) the hybrid modelling framework for production forecasting, which is called BDDCA developed for the enhancement of the forecasting oil

production in long-term focus, while preserving the physical realism and capturing the uncertainty explicitly. The study recognizes the DCA fast and interpretable technique as traditional techniques, but when it comes to unconventional reservoirs, it provides unrealistic forecasts for long-term. Although the pure models of machine learning are powerful, they frequently lack the physical grounding, and the prediction results violate the known behavior of the reservoir. In order to highlight this dichotomy, Tadjer et al. (2022) uses the integration of DCA to neural networks' flexibility and Bayesian statistics rigor. The integrated approach operates in a couple of stages. Initially, the tool of automated machine learning is utilized to determine the correlation among the completion and geological parameters, such as proppant volume, lateral length, and cluster spacing and the features of the modified model of Arps decline. The forecast of DCA parameters is enabled by this mapping technique for the well, whose production history is not so long. The following stage is the introduction of a neural ordinary differential equation model of Bayesian (BN-ODE) that considers the oil production time-dependent system which is governed by learnt differential equations. Essentially, the above-mentioned equations are modified to match the physical behavior illustrated by decline models. The component of Bayesian gives the chance to framework in order to sample from the DCA parameters' posterior distribution; by producing the forecasts in terms of probabilistic perspective and it considers uncertainty in such a principled manner. There are around 400 horizontal wells, which have detailed production data and completion diagrams. In the training part of the model, the production rates are considered as time-series data and validated against taken well data. In order to measure the accuracy of the forecasts, MAPE (mean absolute percentage error) is utilized, for the computation of uncertainty intervals, NUTS (No-U-Turn Sampler), which is a type of Hamiltonian Monte Carlo, is applied. The outputs show that the hybrid model outperforms both types of methods, which are DCA and typical neural networks for long-term forecasts. Futhermore, the forecasts obtained via BN-ODE are more reliable, more interpretable and they can be connected to known physical mechanisms. The approach of the model to uncertainty is one of its most significant contributions. In comparison with the ML models called black box which output the estimate of the point, BDDCA model outputs the interval of the confidence, which are balanced at both theory and data. The capability of the framework in this instance is priceless to make risk-informed decisions, like reserve estimation, modelling of the financial analysis, and allocation of the capital. Moreover, the utilization of SHAP (Shapley Additive Explanations) which is for interpreting model's inputs, gives the transparent distribution and contributes to determine the main drivers of performance of the well like stage count or

proppant intensity. In spite of its sophistication, the model contains few downsides. The load of computation which is associated with training of Bayesian and sampling of NUTS is very high, which makes the model less favorable for low-resource environment and real-time utilizations. Additionally, although the physical constraints are included in the model, high-quality, granular data to demonstrate in optimal conditions are still required, and this is not something that can be available in all reservoirs. The model demands expertise in Bayesian inference, differential equations, and architecture of the neural network, so this framework contains the complexity to implement. In conclusion, this study suggests such a methodology that perfectly correlates data-driven forecasts with the knowledge of the domain. With embedding the constraints into the framework and evaluating the uncertainty via the methods of Bayesian, the standard for the forecast of the production for conventional reservoirs is elevated. This provides both accurate forecasts and a strong alternative to integrate into the workflows of the reservoir management.

1.2.3. Autoregressive and Ensemble ML Models for Forecasting Midland Basin in

Gupta et al. (2021)) investigated the utilization of ML and time-series forecasting techniques in order to automate the forecasting of the production in the Midland Basin, which is the core region within Permian Basin. This study is dedicated to a known clear operational challenge, which is to obtain the scalable, accurate, and timely forecast of the production of thousands of wells on the basis of planning of the development and budgeting quarterly. Trivial techniques like DCA are not the optimal solution for solving this task as they are manual and contain assumptions, which are human factor based. The authors of the paper delve into the problem through the integration of AR (autoregressive models) and tree-based ensemble learning, especially, ETR (extra trees regressor) to a robust forecasting pipeline. There are more than 2000 post-drills and around 350 pre-drill horizontal wells in the dataset. The input parameters include geological properties, artificial lift types, completion data, inter-well spacing, and values of past production. This kind of wide dataset gives a chance to model for understanding cross-sectional relationships and temporal patterns. AR models are applied to cover sequential correlations in the rates of production, whereas the ETR deals with non-linear relationships amidst operational and static features. The estimations are mainly generated at 5- and 30-days intervals, which aligns with the internal planning cycles of the company. One of the key learnings from the study is that models of ML outperform traditional DCA, especially in the wells, which are post-drill, after the period of transient to pseudo-steady

state transition (around 60 days). The models of DCA might match better in the early stage of production because of the aligning regime with initial flow, whereas they start to deviate in the long-term estimations. In comparison, the models of ML adapt better for changes in mid-life production and intervention operations, suggesting the accuracy for improved planning quarterly. It is noteworthy to mention that the framework of ML is completely automated, which enables fast response of re-forecasting to update the data of the field, which can be considered a major efficiency for the companies who are operators.

The study provides ML importance in the forecasting of pre-drill wells utilizing the learning, which is based on analogues, too. Through the train of the model in similar wells and context of geology, authors generate the forecasts to provide the prioritization of the drilling and planning of the capital for the locations which have not been drilled yet. The capacity is so relevant in basins, which are high activity like Permian, where the schedule of the development is so dynamic and the turnover of the data is rapid. However, like other studies, this study has certain limitations. The model of ETR is powerful but has some gaps in interpretability. This model does not give clear insight to the importance of the feature or reasoning of models, unlike SHAP-enabled trees or linear models. Moreover, the uncertainty qualification was not addressed in the study: the estimations are provided as estimates of point, which may put limitations for their utilization in risk management or probabilistic planning. The model drift or the maintenance cycle, essential assumptions for the applied models evolving in environments of the continuous operations have not been covered very well. In spite of the constraints, the study highlights the ML applicability and the relevance of the operation. Integration of ML into regular forecasting and cycles of planning can give reduced workload for humans, the enhancement of the consistency and supporting the strategical decisions. It can assist as a guide for the companies, who are looking for embedding the ML into workflows of the asset management, especially in reservoirs, which are high volume and fast-paced.

1.2.4. Machine Learning versus Type Curves in the Appalachian Basin

Cui et al. (2024) conducted comparative study using the forecasting techniques of the ML and traditional type curve in the formation of Lower Marcellus of the Basin of Appalachian. In this play, more than 2000 horizontal wells of gas were drilled. The purpose of the study is to respond to the need for the industry: development of forecasting methods and its alignment with complexity in geology, heterogeneity and the practices of the completion. Through the application of the tree-

based ML models and the techniques of the advanced explainability, Cui et al. (2024 tries to enhance the accuracy of the forecast and give useful insight for the planning strategically, especially the optimization of the acreage and the design of the completion. Typical type curves, which are the core part of the reservoir engineering, contain the generation of the decline curve, which is representative for the certain group of the wells, which locate in close operational and geological characteristics. Despite its effectiveness and intuitiveness, the method assumes that wells in type curve area (TCA) are assumed to behave similarly during the certain time. The model fails sometimes as it does not consider nuances in variations due to completion design, orientation of the well, spacing, and underlying geology. In order to address this, around 30 TCAs have been constructed across the area of the study, hyperbolic models of decline have been calibrated for each TCA and their performance against the ML approach has been compared via the usage of the gradient-boosted decision trees. The models of ML are trained on the basis of a rich dataset which contains the static data, geological features (thickness, depth, rock type), the completion design (count of the stage, loading of proppant, the volume of the fluid), and the metrics of well spacing. The targeted variable is the cumulative production of the gas in specific duration of time window (12, 24, 36, 48 months), and for each TCA, the models are separately trained. The outputs vividly distribute the superb performance of the model of ML: in ML models of 88% of TCAs, higher R^2, lower RMSE are achieved in comparison to the forecasts of type curve. Those advancements are specially pronounced in the sections where the complexity in geology or irregular spacing of wells exists, where trivial techniques try to simplify the behaviour of the production. The special innovative approach of this research is to build the ML derived RQI (rock quality index) utilizing SHAP values. This SHAP value gives the chance to the authors for the assessment of relative impact of each input parameter on the production prediction for each of well. Via the integration of these attributions spatially, the RQI is built as the continuous map of the surface, by suggesting the granular insight into the quality of the reservoir. This intelligence provides an opportunity to asset teams to get better decisions made about the acquisition of the acreage, targeting of the wells, and prioritizing the drilling sequence. Although conventional maps of reservoir quality are heavily relying on hand-picked features or the discrete cutoff, RQI evaluates the combo effects for multitude of factors learnt by the model. The strength of the study lies in the blends of accuracy of the forecast, practical utility and explainability. It describes that ML both outperform trivial methods and provides extra layers of insights, which are actionable. The application of SHAP is dedicated to one of the major barriers to adoption of ML, which is interpretability, and this assists

to develop the trust amidst engineers, which are utilized to models, which are deterministic. Furthermore, the various scenarios for completion uplift are stimulated through the variation of input parameters, hence giving the opportunity to do “what-if” analysis, which is challenging to perform with the static types of curves.

Nevertheless, the study does have some flaws: When they are far from the training data or when the training quantity is less than ideals with noise in it. For instance, in areas close to asset boundaries where there is higher geological heterogeneity as well as few historic wells. This means that the model's performance will not extend organically beyond its original range of applications, as does traditional linear fitting. Interpreters of SHAP describe how it works; yet at root, the actual model remains a black box in many respects. Additionally, uncertainty quantification is not fed into the forecast outputs--which might be critical for risk-informed planning! One weakness of this model is that we cannot predict with confidence what its effect on prediction performance will be. The other is in the decision support function. By offering a convincing argument for using machine learning to predict output in real time or at least avoid under-for forecasting large development projects that would normally be broken into small increments due to its lack of trend knowledge (i.e., "heterogeneity"), for instance, Cui et al. have revived old questions about just what we do with all our big data. Not only does their work increase both prediction voucher accuracy and decision reinforcement, but it also shows how ML can help close the chasm between field authentic righteousness in unconventional resource plays.

1.2.5. A Machine Learning and Data Analytics Approach for History Matching in a Mature Multilayered Field

One of the most challenging tasks in reservoir modeling is the history matching of a mature multilayered field that has been in production for decades. Suwito et al. (2022) have tackled just such a problem with their study set in Handil field of the Mahakam Delta area Indonesia.Thus, their research illustrates the opportunity to employ machine learning not merely for forecasting, but also speeding up simulation, quantifying uncertainty and supporting decisions.For example, while many studies using ML typically take aim at flawlessly drilling their first few wells or some peculiar kind such as 'tight', this work is aimed at a mature historic brownfield asset with extensive gauntlet-layering, complex well interactions, and thousands of historical data points.The chief

contribution of this paper lies in combining machine learning--in the form, random forest regression--with traditional reservoir simulation tools to build a proxy model. This allows for rapid evaluation of hundreds of development scenarios by replacing computationally expensive full-physics simulations with statistically-trained surrogate models.Thus the workflow is conducted through four gated phases: data conditioning and quality control, feasibility analysis, dynamic modeling and calibration, forecasting and optimization. Python scripts do the model runs and post-processing and a cloud-based dashboard offers visual back for engineering teams.The authors leverage data from wells in more than 50 reservoir zones, with logs and coring nearly 100 contact regions, and some 50 years of oil, gas, and water production. Static properties such as porosity and permeability are derived from supervised learning algorithms based on log and core data.Meanwhile, historical production trends are fit using the random forest model, which is trained to predict production outcomes as a convolution of geological inputs and development parameters.This dual capacity—of accelerating history matching and directing optimization—is key to differentiating the present study from more narrowly focused applications that harness ML.A notable facet of this paper is how it lays stress on scale and practical application. The authors take their framework and apply it to a full-field model that encompasses hundreds of wells, instead of confining themselves in a limited pilot. This shows the method's scalability nature worth and real-world worthiness.The cloud-ready dashboard means that simulation outputs can be interacted with on real time, easing collaboration and scenario exploration for a variety of engineering teams. In addition, the research is able to give quantitative measurement of the relative importance for each input parameter, aiding engineers' task in prioritizing collection and execution efforts.However, there are still some challenges. Although the random forest model performs well in interpolating known scenarios, its extrapolation capability is limited, especially in data-sparse areas or when there is an operational regime change. Reliance on historical patterns also means that novel completion designs or new well types may lie far beyond the model’s training distribution, reducing forecast reliability.Furthermore, although the model enhances computational efficiency, it does not inherently embed physical laws. Therefore its predictions must still be validated against physics-based simulations.Uncertainty quantification comes from Monte Carlo simulation. However, the treatment is essentially statistical rather than probabilistic or Bayesian, limiting its value in frameworks for risk management.Despite these shortcomings, the study makes a strong case for using ML to complement, not replace, traditional history matching-by speeding up simulation, orienting around key variables, and forging through the tangle of alternative scenarios, the ML-

aided method can impart distinct operational and strategic value. It is especially suited for mature assets where there are many data available and even incremental benefits from optimization yield big economic returns.As digital oilfield technologies continue to develop, the methodolgy proposed by Suwito et al. is an example of hybrid field planning that is based on data and physics alike.

1.2.6. Machine Learning Prediction versus Decline Curve Prediction in the Niger Delta

This study, led by Jayeola and colleagues (Jayeola et al. (2022)), provides a clear example of how machine learning may be applied in the Niger Delta Basin. this is particularly true on complex offshore oilfields where traditional statistical models such as Grand Orateur Analysis (DCA) miss some subtle trends. Focusing of 15 oil wells with eight years of daily production data, the authors use Long Short-Term Memory (LSTM) neural networks to forecast time-series production and compare their results with Arps based DCA predictions. Although conventional DCA is widely used, it is unable to produceviable forecasts under the changing, non-linear conditions that increasingly govern field operations as a saturation offshore asset matures. By contrast, LSTM may be capable of modeling long-term time-series dependencies. The system i sensitive to adjust production forecasts for results as subtle as changing gas and water output from associated reservoirs. The abundance of detailed data motivates one contrast, involving the degree of optimization obtained by theAdam algorithm versus Stochastic Gradient Descent (SGD). While SGD has traditionally been used for training neural networks, the authors argue that Adam gives significantly improved convergence rates and forecast accuracy. An LSTM model adapted with Adam showed 96% validation accuracy and lower root mean squared error (RMSE) than its SGD equivalent. This insight points up the broader signal that machine performance is not simply a function of model structure; are also importantly influenced by source data preparation and hyperparameter optimization. The ML pipeline for this study includes data normalization with MinMaxScaler, turning time-series inputs into 3D tensors suitable for LSTM input, and employing dropout layers to filter out noise. All in all, it offes a good example of best practices in ML for temporal forecasting. The results provide convincing proof of how deep learning can be used in reservoir engineering, especially when it comes to data-rich but physics-poor reservoirs. The LSTM forecasts westbound closely with actual production values and were more robust to operational noise and transient behaviour than DCA. In fields like those found in the Niger Delta oil province, where fluctuating reservoir pressure and complex well designs lead to widely differing

flow regimes, LSTM's ability to identify and learn from hidden temporal patterns has market value. The findings are particularly relevant for operators in less developed areas, where it may be impossible to carry out high-fidelity reservoir simulations due to lack of data, budgetary considerations, or simply lack of computing power. However, the study cannot answer all questions. The LSTM model, like many deep learning architectures, operates essentially as a black box—there is little insight into what variables are important or why a given forecast is made. This may create obstacles to adoption in workplaces that value transparency and traceability in engineering judgement. Nor does the study explore how these forecasts could be incorporated into wider operational workflows such as reserves calculation, field development planning, or economic modelling. There is no examination of the impact which ML-derived predictions would have on well interventions, 'field of the future' strategies, or how capital ought to be allocated. Furthermore, uncertainty quantification is entirely lacking. The model performs well in point prediction, but offers no prediction intervals or confidence bands—increasingly important requirements for decision-critical settings. Despite these shortcomings, the paper still makes a major contribution in demonstrating that post-modern ML architectures such as LSTM can bring a higher return than old style methods in real world, intricate production scenarios. Confirmation against traditional techniques makes the argument stronger for using it; especially in parts of the world like West Africa which are underappreciated by the Petroleum Engineering research community. By showing both methodological rigor and practical application, the study offers a strong case for expanding the use of ML techniques in reservoir forecasting—particularly when conventional tools do not provide much insight.

1.2.7. Data Conditioning and Machine Learning Forecasting on a North Sea Well Pad

A comprehensive methodology is described in this study, presented by Bagheri and colleagues. Data conditioning - that is, preparing data for use in production forecasting through machine learning (ML) applications or otherwise-is considered here as essential to such work from the outset.Using data from a well pad in the southern Norwegian North Sea—a multiphase well from Volve field—the authors highlight how data quality, cleaning and pre-processing can have a negative effect on later downstream machine learning models. The work stands out in this respect from many other ML studies published on the web that use petroleum data by supposing their original datum sets are clean and ready to use.The dataset spans over eight years of operation. It

contains production rates, injection data, pressure readings, and downhole temperature measurements from a number of production and injection wells, all placed in multiples throughout the field. About sixty percent (60%) of the dataset were missing or empty, or have abnormal values. The situation is common but rarely given full consideration in field data, even less so at ML conferences. The authors use z-score methods for anomaly detection and support vector regression (SVR) and multilayer perceptron (MLP) models for data impuration.By filling in gaps in time series production and injection records, they are able to improve the completeness and coherence of the datateset. The authors restate that decreasing dimensionality with principal components analysis (PCA) took place next. This addressed multicollinearity among features and retained only the most informative inputs for final forecasting models. Finally, the performance of conditioned SVR, MLP and LSTM models is compared by the authors. As expected, LSTM outperforms all others, by capturing long-term temporal dependencies that are inherent in time-series production data it scored an R^2 of 0.98 and the lowest RMSE for all comparisons.SVR, as a model for regression and also as one that is capable of filling in gaps, concluded their results under fluctuating production regimes and multi-step forecasting was not as good as might have been hoped. One of the key contributions of the paper is a well-detailed and systematic approach to data cleaning and reconstruction--an area that is often ignored but should be indispensible for successful ML deployment. By demonstrating how dirty data cripples model accuracy and how strategic imputation can improve forecasting performance the authors prove that data preprocessing is not mere slingernuther preparatory work but constitutes a major link in ML modelling pipeline.As sensor degradation, drilling platform failures and data gaps all are common situations in offshore wells and old oil fields, this issue is particularly salient.In its methodology the paper may be strong but it does not examine the extent of interpretability or apply empirical methods to costs. There is no attempt to measure how particular features bear on outcome with tools like SHAP or permutation importance used for explanation purposes. As a result the model remains mostly opaque in the eyes of its users, thus lacking in integration with either oilfield development planning scenarios or production optimization cases.Also, there are no bounds to the uncertainty of the predictions or probabilistic forecasts at all. In an operational setting (or really whenever decision makers need to comprehend risk in any form), such probabilistic forecasts might prove quite important.It is to be noted, however, that Bagheri et al. provide a significant solution; they effectively break one of the oil and gas sector's main barriers to ML use the poor quality of its data. Its approach offers guidance for others having problems similar in nature.In short, by showing that

time and effort spent in data preprocessing can directly translate into forecasting accuracy and model reliability, this paper convincingly argues for the adoption of data conditioning as an essential factor in the machine learning lifecycle. This work bridges a gap between academic model development and field applications, especially useful for entities in early stages of digital transformation.

1.2.8. Enhanced Asset Optimization Using ML, Type Wells, and RTA in the Marcellus

Haghighat and Burrough (Haghighat and Burrough (2024)) This study provides a rare masterpiece for the workshop in unconventional development: combining diziterfily techniques of machine learning with tradition restutvoirengool workshops to achieve an unprecedented height. However, relying on the cars of fired well construction through decline curve an alysis (DCA), rate transient analysis (RTA) and machine learning-based gradient boosting scope boost (XGBoost) As the authors explain distinguishes this study ceasless when compared reach its library. Running from physics-based modeling to production forecasting and economic decision-making, all these paths entail their own distinctive but largely automated steps. They start by constructing a type well from production data of 52 offset wells rooted in history which is adjusted for lateral length. This type well serves as the benchmark for per formance expectations in a given geo graphic area or basin. Next. the authors apply RTA to quantitatively understand reservoir properties such as fracture half-length, conductivity, and permeability. These parameters feed a numerical simulator that generates forward production scenarios under different development con figurations, including changes in well spacing, number of stages per well, and amount proppant used. To complement and expand these simulations, the authors train an XGBoost model with data from 300 wells. This includes geologic and engineering as well as spatial features. Crucially, they employ SHAP-based Factor Contribution Analysis (FCA) to explain the model's predictions. This not only increases the interpretability but also allows identification of diminishing returns on variables such as proppant concentration and lateral length. For example. while higher proppant loads and tighter fracture networks maximizewell EUR optimal net present value (NPV) occurs at more moderatelevels because of diminishing cost efficiency trade-offs. Such insights are critical in planning field development work so one can balance productivity with economic returns. This study offers a rare integrated blueprint for asset development planning. Instead of treating machine learning as an isolated solution in its own right, its real aim is to integrate this with well-established engineering

tools in order to give the developer a more comprehensive understanding of reservoir behavior and development outcomes. Meanwhile, the combination of RTAC and ML allows for insights coming from physics-based reasoning as well as data-driven input; whereas economic modeling links the whole flowchart up to business objectives. This is particularly useful in resource-intensive plays like the Marcellus, where momentary gains in overall design yield large changes across your stocks. However, some limitations need to be mentioned. Our model assumes uniform data accuracy and measurement uniformity among all wells, which may be different in other geographies. Even if SHAP values make things more transparent they do not totally solve one problem unique to ensemble models like XGBoost: their black-box nature. Moreover, we have not yet explicitly built any uncertainty quantification into the forecast or economic analysis. This would make it more serviceable in high-risk investment scenarios. Bearing these omissions in mind, the article puts up an extremely practical and replicable template for adding ML to the reservoir engineer or asset manager’s toolbox.

1.2.9. Implications for decision-making and research gaps

The literature reviewed consists of eight different and technically advanced works,reflecting the increasing maturity of machine learning (ML) as an effective tool in production forecasting and decision-making within petroleum engineering. These studies differ in their geological settings, model types, and objectives-- yet they all cast light on common themes and strategic applications, while underscoring the remaining challenges. Collectively, taken together, these studies sketch out a composite picture of where we stand today. Forpractitioners and researchers alike, this cross-sectional view of the field is of high value. This strategic use of diverse and integrated data is common across all the papers. From unconventional shale plays like Bakken and Marcellus in the Midland Basin, to complex offshore and mature carbonate formations such as those in Nigeria's Niger Delta or Indonesia's Gulf Mahakam. Even in an established producing area like the North Sea, it closely links the progress of ML applications with ability to synthesize dynamic production data, static geological properties, completion details and spatial context. This is exemplified by the studies of Huang et al. and Suwito et al., which blend both static and dynamic data sources to understand heterogeneity and history matching, respectively, while those of Cui et al. and Haghighat & Burrough utilize geospatial insights or feature attribution tools as decision aids. Meanwhile, in a specialized and unique contribution, Bagheri et alhighlight the importance of data

preparation itself — stressing that data quality is make-or-break for the reliability and generalization ability of models. In terms of model selection and method design, the papers cover the full spectrum of ML - from unsupervised clustering (Huang et al.), ensemble methods (Gupta et al., Suwito et al.), and time-series neural networks (Jayeola et al., Bagheri et al.), to hybrid implementations that blend physics with machine learning (Tadjer et al.) or integrate RTA and DCA with ML and economic modeling (Haghighat & Burrough). This diversity in methods reflects the need to tailor ML approaches to each individual task: unsupervised learning for discovering patterns, supervised regression forecasting, and hybrid models which incorporate domain knowledge and build trust. Explanability tools like SHAP (Cui et al., Haghighat & Burrough) are now a growing trend, which plays an especially important role in improving model transparency and acceptance among engineers. Forecasting accuracy is still an urgent goal; and all studies demonstrate that often ML models can outperform traditional techniques—particularly in capturing the non-linear, time varying patterns which DCA or type curves have difficulty with. Gupta et al. conclude that tree-based regressors are particularly good at prediction post-transient production, while Jayeola et al. and Bagheri et al. find that in managing sequence-based data which is noisy or variable LSTMs are better than other regressors. Tadjer et allgo one step further: not only do they present Bayesian neural ODEs that ensure highest accuracy while also offering forecasting probabilities, they combine physical behavior and uncertainty quantification in a single framework. However, high predictive performance alone is no longer sufficient.Increasingly, these studies look at how ML supports decision-making and operational efficiency. In practical terms, ML can improve short and midterm field operations by automating production forecasts (Gupta et al., Jayeola et al.), pointing out underperforming wells (Huang et al.), and suggesting current adjustments to artificial lift or completion strategies (Bagheri et al.). At a strategic level, ML outputs now influence well spacing decisions, completion intensity optimization, acreage valuation, and economic planning. A sophisticated example of this is provided by Haghighat & Burrough, who show how ML predictions can be combined with DCA, RTA, and NPV in a multi-layered evaluation approach–transforming ML's role from forecasting alone to full-blown optimization and capital planning. The studies reviewed also reveal that there are several gaps in research and application which need to be filled in order for ML to realize its full potential. Interpretive abilities continue to be a headache, especially with deep learning models like LSTM which are far from transparent compared with tree-based algorithms. True, Cui et al. and Haghighat & Burrough have effectively deployed SHAP for interpreting features, but others like Jayeola et

al.and Bagheri et al.provide highly accurate yet fundamentally opaque forecasts. To bridge this gap, we need farther advances in explainable AI tailored to the characteristics of reservoirs, especially in sequence-based models. Another problem that urgently needs to be addressed is uncertainty quantification. Only Tadjer et al.model forecast uncertainty explicitly, using Bayesian inference to provide confidence intervals. The remaining studies produce point forecasts, which may limit their value in risk-based decisions such as reserves booking, scenario planning, and infrastructure investment. Introducing probabilistic frameworks — whether Bayesian deep learning or Monte Carlo dropout— into time-series models and tree-based regressors is an area ripe for future exploration. Data availability and quality are further obstacles. Bagheri et al.show howcoponditioning the data can dramatically improve the performance of ML; but in many fields the data is either sparse, noisy, or incomplete. There is a need for robust data preparation workflows, uniform imputation strategies and feature engineering templates that can be transferred and modified across fields or organizations. Nor have many of these papers tackled the full lifecycle management for ML models— how one monitors models as they move into the field, retrains them with new data, or sustains them in real production environments. An investigation into “MLOps” (Machine Learning Operations) on subsurface forecasting could supply actual blueprints for establishing and maintaining robust ML tools within enterprises.

Finally, the econom in the use of databases is a maping dimension of implications. DarkHorse uses a loadstream from the feed to network data into Zabbix, while crawlers and so on getting run once each time complain to user traffic for downloading a failed document and not again after successful retrieval go through such tasks progress within your browser-at least there that's how things should be In their current version some pieced-out examples were merely collected for different interfaces, instead producing a bit of a mess. But with the directional flow chart mechanism, people find it easier to get these details nicely squared away They. also came up with a graphical tool Weka at that stage. Now many algorithms have been put into it which some intelligence-aware mathematical geeks gathered from around the world Future work should do more of this integration, e.g. by feeding ML predictions as inputs into financial models. And it needs to increase understanding That combinations also must link up development planning parameters and riskweighted asset investment strategy so that other people deeply involved in both sides of those business undertakings than the original technical work can then realize their work alongside your 'technical work-dlings' All eight of these case studies are a comprehensive and multi-faceted examination

into ML's impact on reservoir forecasting and production scheduling. They provide proof that not only razor sharp accuracy but also increased speed, capacity for scalability (scalability is another name for volume) and decision support are major benefits of ML applications. However, as this field evolves, attention must move from a single model's isolated performance to more general themes like interpretability and uncertainty. Workflows need responding properly in order for them work both efficiently and economically together. These shortcomings, therefore, are key issues in achieving the full value from ML as a critical element of modern reservoir management.

1.3. OIL PRODUCTION PREDICTION USING TIME SERIES FORECASTING AND MACHINE LEARNING TECHNIQUES

In the oil and gas industry, anticipating the oil output has remained challenging, given the importance it holds for an organization’s strategic decision-making. In the past, several empirical correlations along with different mathematical models served this purpose. Today, the shift to data-oriented extrapolation has led to the adoption of machine learning algorithms such as Random Forest (RF), Artificial Neural Network (ANN), Long Short-Term Memory neural network (LSTM), Recurrent Neural Network (RNN), and even DeepAR among others. This paper presents a comparative analysis between time series and machine learning techniques to forecast oil production. To reach this goal, the ARIMA, Prophet, Random Forest, CatBoost, and XGBoost algorithms will be used. Time series forecasting uses historical data to build predictive models, and the machine learning approach builds a model on a dataset which can reliably be utilized to make predictions. In recent years, advancement in computing technology, including data analytics, has enabled the development of precise and sophisticated models to aid in accurately projecting crude oil production. Methods based on artificial intelligence (AI) and machine learning (ML) have attracted considerable attention in this field owing to their ability to handle large volumes of data and provide accurate predictions. Various ML approaches such as Random Forest (RF), Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Recurrent Neural Network (RNN), alongside Deep-AR, have been utilized for forecasting crude oil production. These algorithms are capable of recognizing and determining the relationship between oil production and other data captured from the field through sensor devices using machine learning techniques. Temitope James Omotosho, (2024) applied a time-series forecasting techniques and a machine learning methodology to anticipate oil output. Time series forecasting is the application of a

forecasting model to estimate future values based on historical time-stamped data. Modelling is the process of constructing models by analysing previous data and applying them to generate forecasts that will inform future decision-making. A fundamental differentiation in forecasting is that the future result is completely unknown at the moment of study and can only be approximated through comprehensive examination and priors based on evidence. The objective of time series analysis is to derive valuable statistical properties (such as trend, pattern, and seasonality) from a time series, construct a model that explains these properties, employ the model for prediction, and ultimately use the knowledge acquired from the study to make informed decisions. A machine learning-based forecast is the outcome of an algorithm that has undergone training using a historical dataset. The method thereafter produces likely values for unidentified variables in every entry of the new data. The objective of prediction in machine learning is to forecast a probable dataset that aligns with the input data. This facilitates the oil and gas sector in predicting forthcoming crude oil output and market patterns. In order to accomplish optimal outcomes in machine learning prediction, businesses need to provide infrastructure to facilitate the solutions, together with high-quality data to input into the algorithm. Numerous scholars have conducted studies on the forecasting of crude oil production by employing various time series forecasting models and machine learning techniques. According to Omekara et al. (2015), the multiplicative Seasonal Autoregressive Integrated Moving Average model (SARIMA) (1, 1, 1) (0, 1, 1) was identified as the most effective model for predicting crude oil production data in Nigeria. The proposed model was recommended to the appropriate authorities for the purpose of predicting future crude oil volumes in the country. The ARIMA model was employed by Fatoki et al. (2017) to conduct an estimation of crude oil production in Nigeria. The order of the ARIMA model (1, 2, 2) appropriately corresponds to the data. The model projected a consistent upward trend in crude oil output from 2014 to 2023. However, the actual production of crude oil experienced a significant decline after 2015. In August 2016, Nigeria achieved its lowest crude oil production of 1.5 million barrels per day (mmbpd) between 2006 and 2020. In their 2015 study, Balogun and Ogunleye examined an ARIMA model with varying orders (1,1,1), (2,1,2), (2,1,1), and (2,1,0) to forecast short-term crude oil output. The findings indicated that the model with the order (1,1,1) outperformed the other models. An investigation was carried out by Sadeeq and Ahmadu (2018) to determine the most effective time series model for monthly crude oil output in Nigeria, employing the ARIMA model. ARIMA (2,1,0) (2,1,1) excelled as the optimal model for the given data. Exhibited strong predictive capabilities for monthly crude oil output. In their study, Tadjer et al. (2021) investigated the use of

machine learning-based decline curve analysis for short-term oil production forecasting. They specifically considered two prominent models, namely DeepAR and Prophet. The efficacy of the proposed models was evaluated on a specific well from the Midland fields in the United States. It was deduced that both DeepAR and Prophet analysis are valuable tools for enhancing comprehension of oil well behavior. Moreover, these strategies can effectively reduce the occurrence of over/underestimations that may arise from relying solely on a single decline curve model for forecasting. A comparative analysis was conducted by Zou et al. (2021) to assess the accuracy of shale oil production prediction using long-term and short-term memory neural network (LSTM), ARPS production decline model, and Prophet algorithm. The results demonstrate that the Prophet algorithm has superior prediction accuracy, particularly in the context of complex shale oil production.

Despite the promising results obtained by Omekara et al. (2015) and Zou et al. (2021) using ARIMA and Prophet for prediction, the scope of their methodology is restricted to univariate data. Specifically, the dataset used only the date column as its input parameter for forecasting oil production. Consequently, it does not accurately represent the other factors that influence oil production in real-life applications. The objective of this publication is to demonstrate the drawbacks of using univariate data for forecasting oil production in comparison to considering other factors. To address the difficulties of predicting oil output using a single variable, it is essential to adopt a method that considers other factors influencing oil production, such as employing a non-parametric machine learning model. Although machine learning (ML) is a relatively new approach in the petroleum business, some academics have explored its potential uses in predicting crude oil production. Liu et al (2019) discovered that conventional back propagation neural networks are unable to effectively capture the temporal correlation among data. This led to the development of a long short-term memory (LSTM) model for creation of a production prediction model that considers both production data trends and context correlations. The findings indicate that the projected output generated by the LSTM network has a strong correlation with the real output, thereby effectively representing the dynamic fluctuations in production. In their study, Luo et al. (2019) constructed non-linear models and employed Random Forest (RF) and Deep Neural Network (DNN) algorithms to predict the total oil output over a period of 6 months. The complete dataset was acquired from approximately 3600 wells located in the Eagle Ford formations. The analysis revealed key geological characteristics, including structural depth, formation thickness, and total organic carbon (TOC), as input variables that

influenced the productivity of wells in Eagle Ford. Obite et al (2021) conducted a comparative analysis of a classical model (ARIMA) and two machine learning models (ANN and RF) at the level of crude oil production modeling in Nigeria. The Artificial Neural Network (ANN) model that achieved the optimal balance between Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Nash—Sutcliffe Efficiency (NSE) parameters was employed for the prediction of crude oil output in Nigeria. The present study has conducted a comprehensive examination of machine learning and time series forecasting methodologies employed in the analysis of oil production. Furthermore, the model performance criteria employed to assess the accuracy of the forecast were elucidated, and the findings were thereafter presented and deliberated about.

1.3.1. Workflow

The input and output data that were gathered from the Volve production field in Norway served as the basis for the crude oil production data that was utilized in this academic investigation. For the period beginning in September 2007 and ending in September 2016, the data consists of daily production estimates. The data was separated into two sets, often known as the training set and the test set, in a manner that was not random but rather sequential. There is a way to prevent data leakage for future prediction, and that way is through the sequential split. Sixty percent of the data is comprised of the training set, while the remaining thirty percent is for the test set. The training set is utilized for the purpose of estimating the parameters of the model, whereas the test set is utilized for the purpose of validating the model and gaining an understanding of how well the model performs on new datasets. There were five different models that were applied to the data. These models included time series forecasting methods such as ARIMA and Prophet, as well as machine learning models such as Random Forest, CatBoost, and XGBoost algorithms. Using the RMSE, MAE, and R^2 score, the best model was chosen to represent the data.

1.3.2. Data cleaning and preprocessing

As part of the data cleaning procedure, several useless columns related to oil production forecast, including well_bore_code, npd_well_bore_code, npd_well_bore name, npd_field_code, and others were eliminated. Descriptive statistics analysis of the data revealed the presence of outliers, which were subsequently substituted with zero. The presence of numerous readings per day, which were not necessarily taken at a certain moment for some days, resulted in a significant amount of lost

data. The method employed in this study to address the missing values was to consolidate the data into a daily frequency instead of removing or replacing the missing values. To do this, the various readings for each day were grouped based on the date column and then the sum for each day was determined. Only two essential data elements were necessary for the univariate analysis: the date column and the goal variable (oil production). Hence, a single-dimensional data frame was generated for the purpose of time series prediction utilizing ARIMA and Prophet models.

Comprehending the distribution of variables is essential while performing data analysis. The Kalman Filter (KDE) is a nonparametric technique employed for the estimation of the probability density function of variables. A significant observation is that the majority of the distributions exhibit non-normality, with certain distributions being bimodal and others being left-skewed. Consequently, data transformation is essential to standardize the data, so enhancing the efficiency of the model. Presented in this context, Figures 1.3.2.1-1.3.2.7 depict the univariate distribution of each characteristic by constructing a histogram and applying a kernel density estimate (KDE) for fitting.

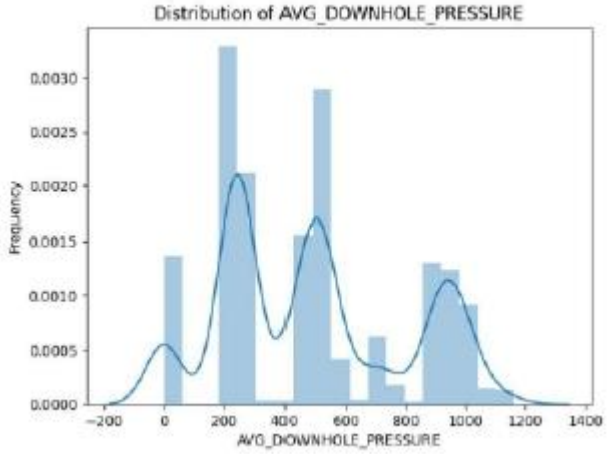

Figure 1.3.2.1. Distribution of downhole pressure

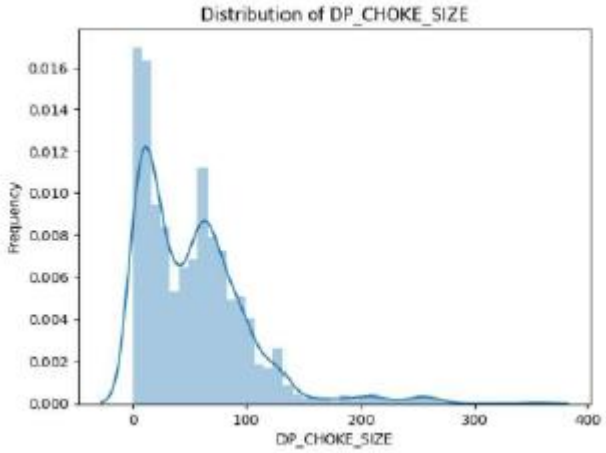

Figure 1.3.2.2. Choke size distribution

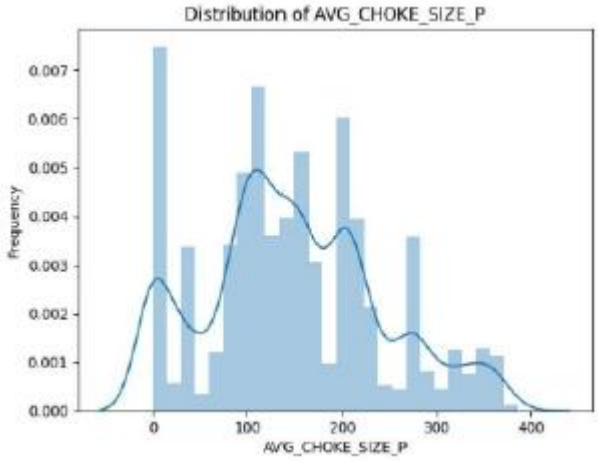

Figure 1.3.2.3. Distribution of averaged choke size

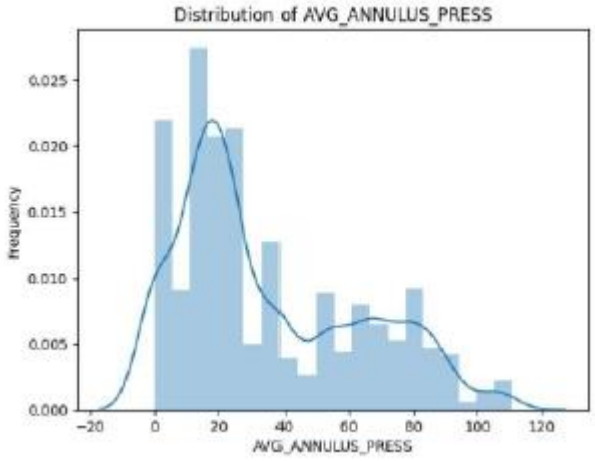

Figure 1.3.2.4. Distribution of average annulus pressure

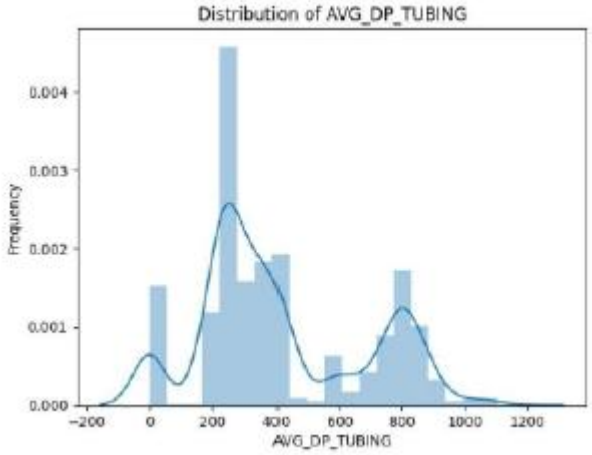

Figure 1.3.2.5. Distribution of average downhole pressure tubing pressure

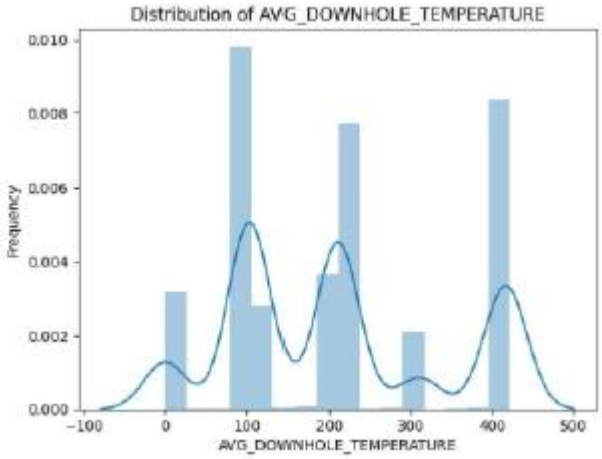

Figure 1.3.2.6. Distribution of average downhole temperature

Following the completion of the data cleaning procedure, it was imperative to analyze the association among the variables in the dataset. This measure was implemented in order to mitigate the problem of multicollinearity, which can greatly affect the precision of the findings. A heatmap was created to visually represent the correlation between the variables. The heatmap illustrated the interrelationships among the several variables in the dataset. This facilitated a rapid and effortless ascertainment of any robust correlations, which could then be taken into consideration during the study.

Through careful analysis of the heatmap (Figure 1.3.2.7), it is evident that there is a flawless correlation between the amounts of oil production and gas output. To mitigate the risks of data leaking and overfitting, the variables related to gas production and water production were

eliminated. Furthermore, as oil, gas, and water are generated simultaneously, it is not feasible to use these variables as predictors for oil production volume in real-life applications. As a result of their low feature importance and minimal effect on the accuracy of the model, the average wellhead pressure and temperature were eliminated from the dataset in order to enhance the precision of the forecast.

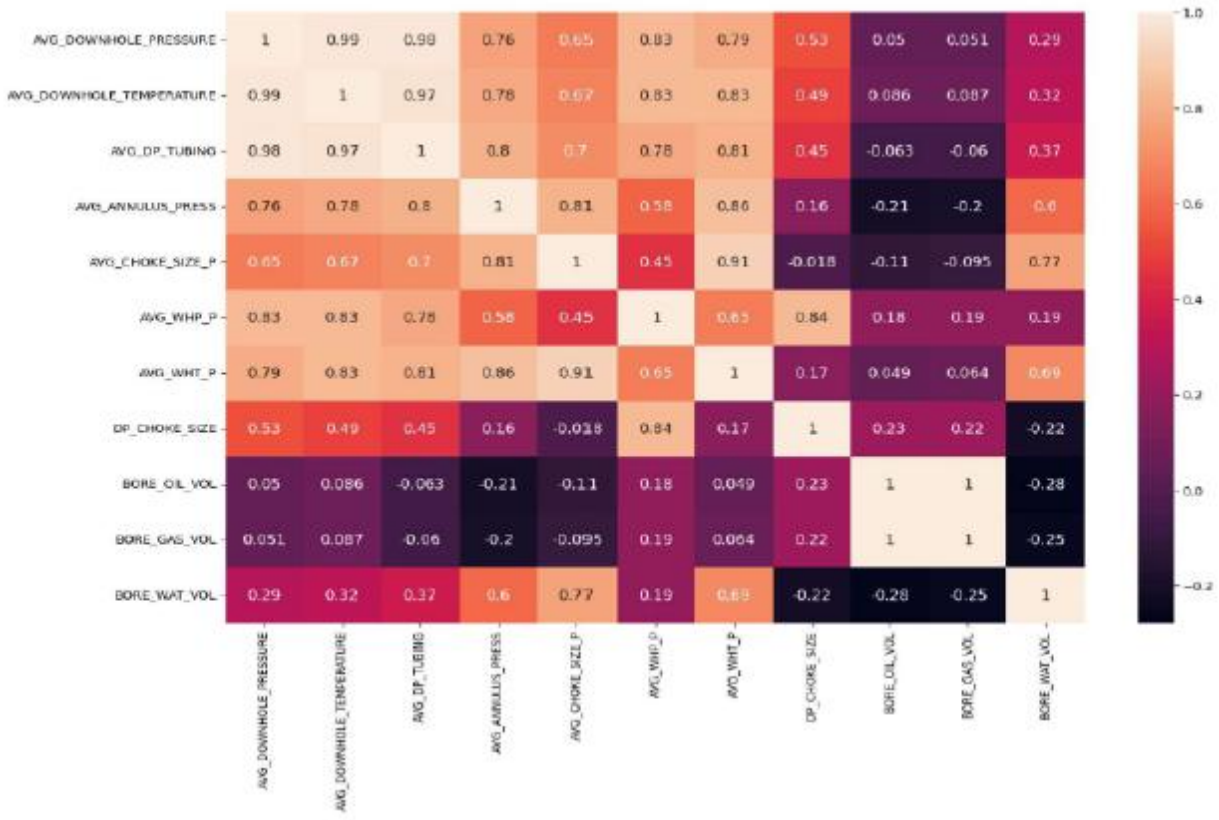

Figure 1.3.2.7. Heatmap of correlation of the dataset features

To facilitate the machine learning-based prediction, the data underwent normalization using the MinMaxScaler. Only the pertinent characteristics that resulted in improved precision in forecasting oil production were preserved in the dataset. Table 1.3.2.1 presents a concise overview of the six main characteristics that made a substantial contribution to the forecast of oil output. The precise selection of these features enabled the study to enhance the forecast accuracy and guarantee the relevance and effectiveness of the analysis.

Table 1.3.2.1. The overview of the parameters impacting on the forecast.

Feature description	Value
Average Downhole Temperature	101.06
Average Tubing Downhole Pressure	244.28
Average Choke Size	198.2
Choke Size DP Ratio	8.078
Average Downhole Pressure	242.73
Average Annulus Pressure	36.44

1.3.3. ARIMA forecasting model

The ARIMA model, which was formulated by Box and Jenkins in 1976, integrated the Moving Average (MA) and Autoregressive (AR) models for stationary population data. The formula given is ARIMA (p, d, q), where "d" represents the number of differenced data points required to reach stationarity, "p" denotes the number of lags in the Partial Autocorrelation Function (PACF) plot that surpasses the significant threshold, and "q" represents the number of lags in the Autocorrelation Function (ACF) plot that surpass the significant threshold. This topic is thoroughly elucidated in the work of Nwosu and Obite (2021). An ARIMA (p, d, q) model for a time series data X is defined by Equation 1:

$$\varphi(B)(1-B)^{d}X_t = \theta(B)Z_t \quad (1.3.3.1)$$

where:

φ (B) - characteristic polynomial of order “p”, θ(B) - characteristic polynomial of order “q”, $(1 - B)^{d}$ - differencing of order “d” of the data, X_t - observed value at time t, Z_t - random error

1.3.4. Prophet Forecasting Model

Prophet forecasting, a Bayesian nonlinear generative model for time series forecasting, was developed by the Facebook Research team (Taylor and Letham, 2007) with the aim of providing high-quality multistep-ahead forecasting. Prophet facilitates the automation of term computations in the model and mitigates forecast inaccuracies. This library predicts data using either the logistic growth model for non-linear data or the piecewise linear model for data with linear characteristics,

with the selection of the latter being the default choice. The library offers user-friendly and intuitive settings that can be readily adjusted by anyone without expertise in forecasting. The Prophet forecasting model employs the additive regression model expressed by Equation 1.3.4.1. This model consists of the following components:

Given by:

$$y(t) = g(t) + s(t) + h(t) + A_t \quad (1.3.4.1)$$

Where:

Let y(t) be the variable of interest. The figure g(t) represents a piecewise linear or logistic growth curve. Let s(t) represent periodic changes. Variable h(t) represents the impact of irregular holidays. Δ, - error term that considers any unpredictable fluctuations.

1.3.5. XGBoost Statistical Model

XGBoost or Extreme Gradient Boosting stands out amongst the rest as a go-to option and favorite in the field of machine learning, such as benchmarking appears effortless due to its uses in customized ensemble learning technique. XGBoost's algorithm works on a decision-tree model that unites the pros of both bagging and boosting methods. In Bagging, a number of decision tree models are developed using random samples of the training set, after which their predictions are statistically averaged to obtain the final result. Using this method reduces the variance within the model, making it more robust. In boosting, there is a conditioning algorithm that adds a new tree onto the model that was previously fitted on the data, with the next tree learning from the biases of the previous tree. The basic learners in boosting are weak learners who have high bias and low prediction power. But, when these are combined, they produce a learner that is low in bias and variance. Both bagging and boosting improve the strength and accuracy of decision tree models tremendously. In XGboost, the process of constructing the decision trees is iterative, with new trees correcting errors of older trees. Bagging and boosting allows the construction of multiple decision trees and therefore enabling more precise forecasts.

Additionally, XGBoost incorporates the regularization technique to reduce overfitting, so enabling it to be more robust and better at generalizing to new data. XGBoost is very effective when working with large datasets containing a lot of features because it is designed to perform best with high-dimensional data. This model is recognized for its speed and scalability, facilitating training and prediction on large datasets significantly faster.

1.3.6. The CatBoost model

Catboost is a quantile regression method with an integrated gradient boosting learning algorithm built from the ground up to deal with categorical features. Like other boosting schemes, this strategy is based on decision trees, but it has a unique approach to handling categorical values. Categorical features do not need to be one-hot encoded or preprocessed because CatBoost is able to do so automatically. Stubbs said CatBoost, along with other machine learning algorithms, does a good job handling missing values as well. Due to the implementation of a symmetric tree structure for decision tree construction, CatBoost delegates pruning and boosting. As a result, training is faster and more accurate than with other methods that use decision trees. Moreover, the innovation in calculating gradients and hessians boost training speeds. Training is done by iteratively generating/adding a set of decision trees, each with superior accuracy to the last. As the index of newly added trees becomes higher, the accuracy gets closer to that of previously added trees, meaning the addition of less accuracy. Starting criteria defines how many trees must be planted.

CatBoost also adds other optimizations like adaptive learning, early stopping, computing importance of features, and rate calculating defined on the number of features added to the model. These improve the efficiency of training, prevent overfitting, and augment model interpretability.

1.3.7. Ensemble Random Forest Model

The Random Forest approach integrates multiple decision trees and applies ensemble learning to classification and regression problems. The final output value is the average predicted value (in regression) or the class that appears most frequently (in classification) among all of them, according to Ho's definition (1998). This paper by Nwosu et al. (2021) gives an exhaustive account of the Random Forest algorithms. The Random Forest algorithm employs a subset of attributes to split nodes, which have been preselected randomly. Thus, a model which is improved over predecessors can be created. To grow each tree, the following steps are taken:

1. Start with choosing the cases: to build the tree, choose M case records from the given dataset. Assume that the dataset contains M records and you can choose them multiple times.

2. Do the same for each node: With the set of explanatory variables P, you need to choose a specific number p that is less than P. Then you make a random selection of p variables to limit yourself to.

With the optimal split algorithm, this node will be divided on the p selected variables. While growing each decision tree in the forest, p is a fixed value. If two trees are correlated, the error rate in the forest will increase. When "p" is lowered, the degree of association among trees is less.

1.3.8. Forecasting Performance measures

Precise evaluation of our models using suitable metrics is crucial. Within time series forecasting, several metrics can be employed to assess the performance of a model. The optimal model for crude oil production in the Volve Field in Norway was determined using the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2 score). The algorithm exhibiting the lowest RMSE and MAE values, together with the greatest R^2 score value, is selected as the most effective algorithm. The performance evaluation metrics can be calculated using the formulas provided in equations 1.3.8.1-1.3.8.3.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|A_i - F_i| \quad (1.3.8.1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A_i - F_i)^2} \quad (1.3.8.2)$$

$$R^2 = 1 - \frac{RSS}{TSS} \quad (1.3.8.3)$$

Where,

A_i = Actual value, F_i = Forecasted value, RSS = Residuals sum of squares, TSS = Total sum of squares

1.3.9. Machine Learning Based Prediction

The present work employed decision-tree based machine learning algorithms, which possess the ability to take into account several variables and generate precise predictions. Through the integration of reservoir conditions and production equipment data, our system attains superior accuracy and reliability in its forecasts compared to conventional time series forecasting methods. Three machine learning algorithms, namely XGBoost, CatBoost, and Random Forest, were deployed to forecast crude oil production. 70% of the data was used to fit the models, while the remaining 30% was used for evaluation. Following the training of the models using the training data, an analysis was conducted on the factors that influenced the prediction accuracy of the three

models. Analysis of the top features revealed that related features enhanced the model's accuracy, as depicted in Figure 1.3.9.1.

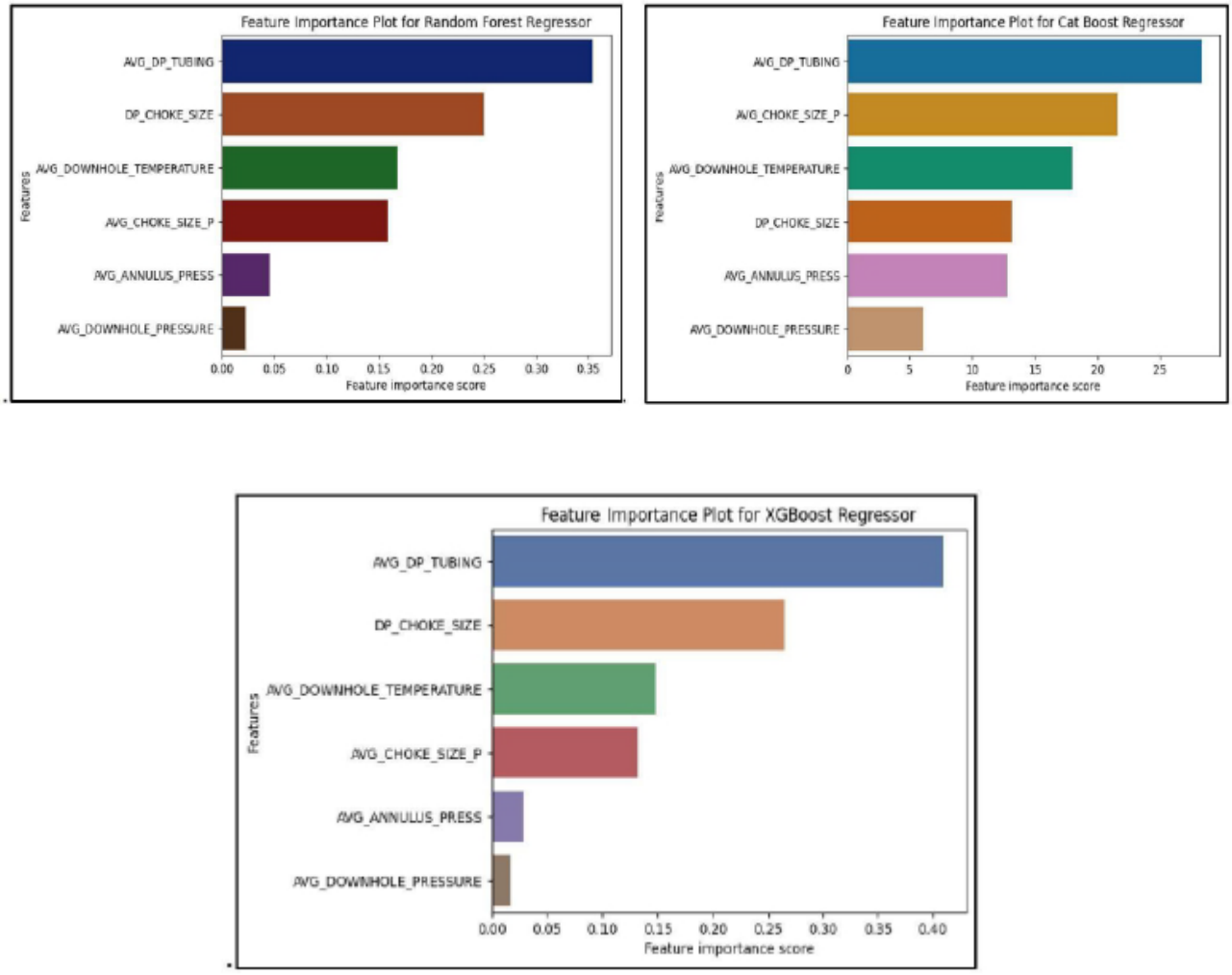

Figure 1.3.9.1. Feature importance diagram for 3 top performing models

Only six characteristics were found to be important to the correct prediction of crude oil output after meticulous iterations of training had been performed to achieve the highest possible accuracy across all three models. Figures 1.3.9.2a-1.3.9.2c illustrate the outcomes that were discovered as a consequence of the predictions made by the three models.

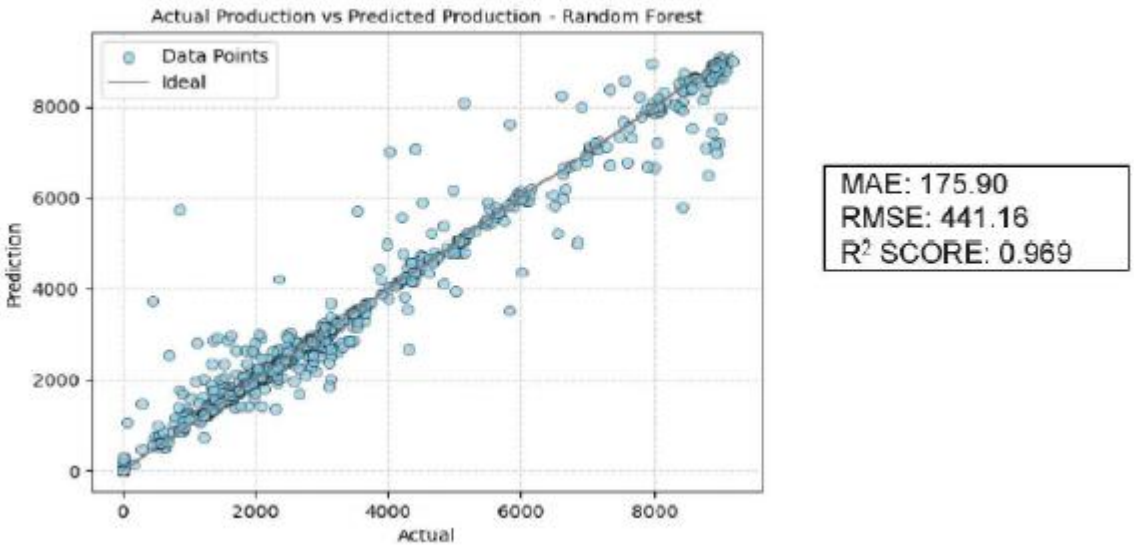

Figure 1.3.9.2a. Random Forest Algorithm performance

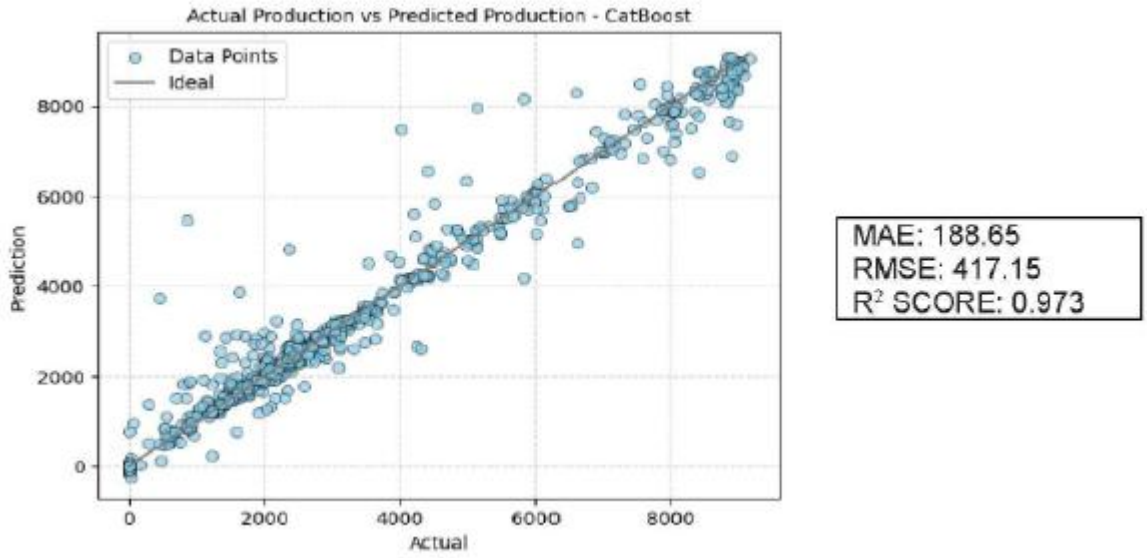

Figure 1.3.9.2b. CatBoost Algorithm Performance

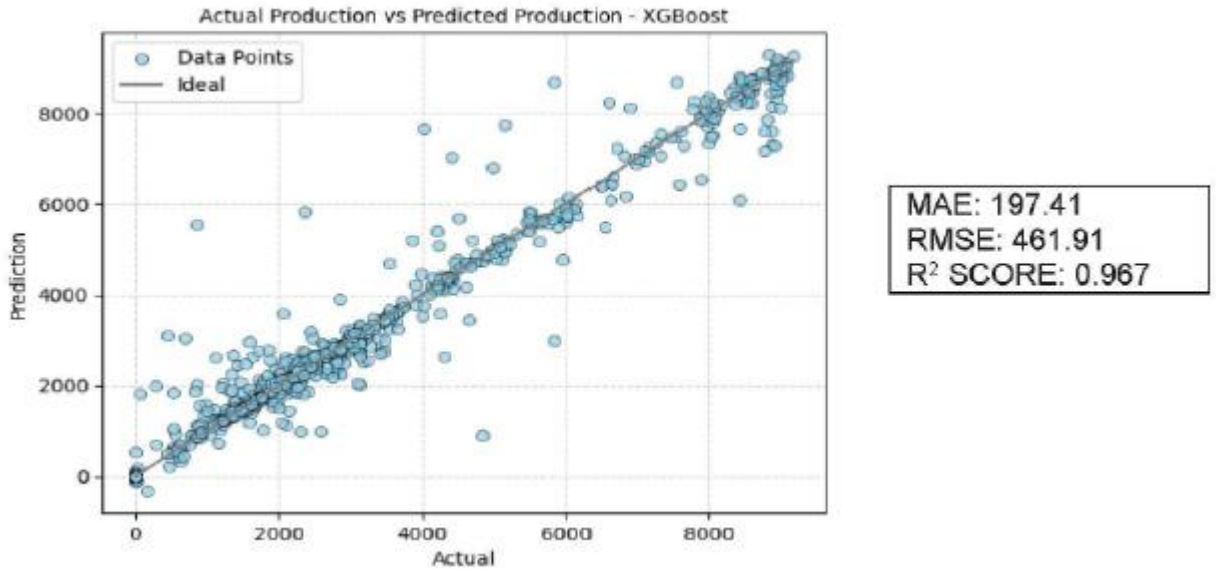

Figure 1.3.9.2c XGBoost Algorithm Performance

The analysis of Figure 1.3.9.2a:1.3.9.2c revealed that the Random Forest method exhibited the highest performance in terms of the Mean Absolute Error (MAE). This metric is crucial as it provides insight into the proximity of our forecast to the actual values. In contrast, the CatBoost algorithm exhibited superior performance in terms of the RMSE and R^2 score. Given the absence of a model exhibiting superior accuracy across the assessment measures, the prediction accuracy was enhanced by the implementation of hyperparameter tuning and stacking techniques. The data presented in Figure 1.3.9.2a:1.3.9.2c represent the outcomes of the baseline models without any optimization. In order to achieve a more optimal outcome and a corresponding improvement in accuracy, we conducted hyperparameter tweaking on each of the models using the GridSearchCV package. Table 1.3.9.1 displays the findings for the optimal hyperparameters. Random Forest continues to exhibit the highest performance in terms of Mean Absolute Error (MAE), with a limited improvement compared to the baseline model.

Table 1.3.9.1. ML model prediction results

	MAE	RMSE	R^2
XGBoost	197.41	461.91	0.967
CatBoost	184.49	413.83	0.973
Random Forest	173.7	436.74	0.971

Given the comparable prediction accuracies of the fine-tuned models and the baseline models, we opted to leverage the respective strengths of each algorithm by performing a stacking operation to see if it would result in an enhancement in accuracy.

1.3.10. Stacking

The stacking procedure was performed using the stacking regressor tool in the sklearn machine learning toolkit. The estimators used were XGBoost, CatBoost, and Random Forest, with RigdeCV serving as the final estimator. Following the stacking procedure, there was a notable improvement in the precision of the forecasts across all evaluation criteria. The outcome is depicted in Figure 1.3.10.1.

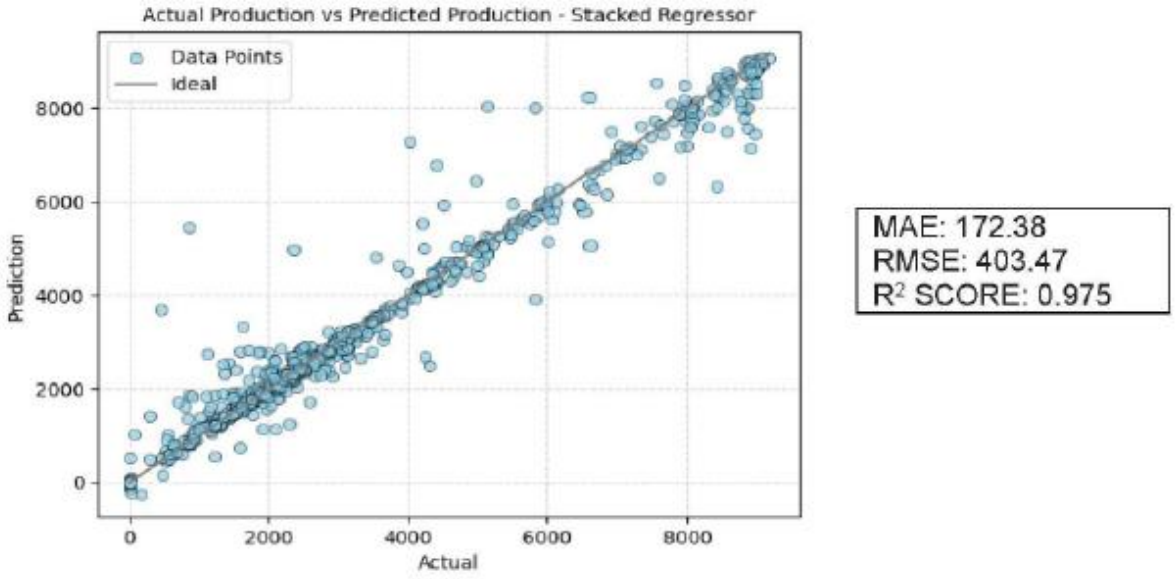

Figure 1.3.10.1. Stacking Regressor Model Performance

1.3.11. Model Comparison

Upon evaluating the several approaches employed in this work to forecast crude oil production, it has been confirmed that the stacking regressor outperforms all other techniques in terms of assessment metrics. Figure 1.3.11.1 presents a concise overview of the performance of all the models employed in this investigation, categorized by mean absolute error (MAE) and coefficient of determination (R^2 score), arranged in order of decreasing order of performance.

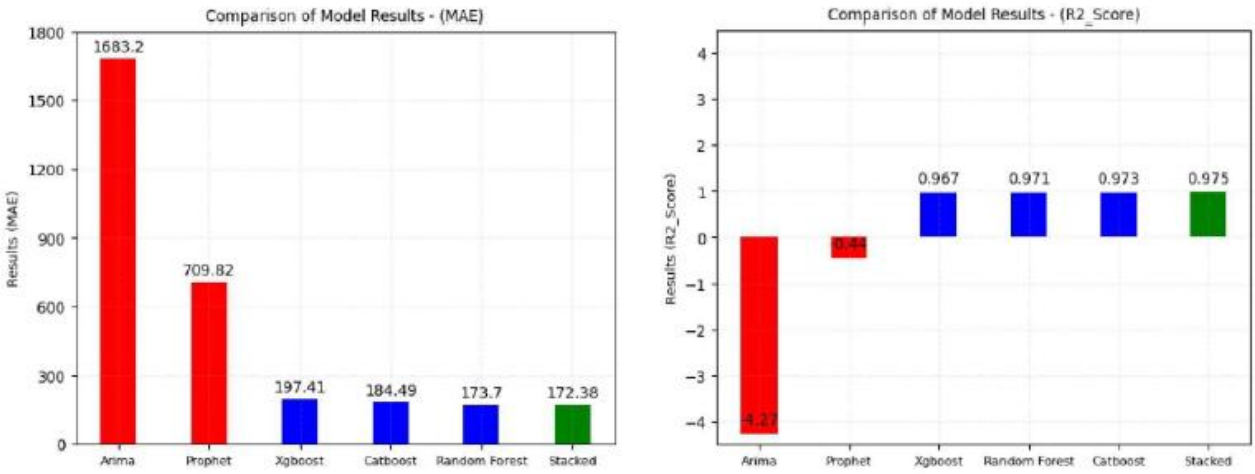

Figure 1.3.11.1. Model comparison outcomes.

Created models undergo thorough evaluation through statistical analysis and error measures, demonstrating varying levels of predictive accuracy. Considering the model performances, the LSTM model is suitable for forecasting petroleum output, effectively addressing seasonality and production anomalies throughout the reservoir's lifespan. The analysis of several statistical models suggests that, instead of relying solely on oil output as the input variable, it is essential to incorporate additional characteristics to predict oil production in a reservoir. The forecasting efficacy demonstrates that the suggested LSTM model is applicable to long-term time series predictions in the petroleum sector. The research demonstrates that selecting the appropriate optimizer is crucial for training the LSTM model. An LSTM architecture optimized by the ADAM (Ifeoluwa Jayeola, Bukola Olusola and Kale Orodu, 2022) algorithm yields enhanced training and validation accuracy across all forecasts. The suggested utilization of computational tools in forecasting issues has demonstrated itself as a robust and dependable approach for predicting the future performance of production wells. Hence, in this research, as ML model, LSTM and XGBoost model will be utilized.

CHAPTER II. METHODOLOGY

As the toolkit for making the models, Python and its libraries will be used. The core part of the model, LSTM model, through each layer's 50 memory units will be able to provide a sequence of output rather than single value. The LSTM layer can receive 3D input. A dropout layer is included after each LSTM layer to prevent overfitting and enhance generalization error. The output layer is dense and uses a linear activation function to collect input from the previous layer's neurons. The libraries required for constructing the model are:

- NumPy.
- Matplotlib.
- Pandas.
- Keras;
- Scikit-learn.

As feature engineering stuff, using data mining techniques to extract features from existing datasets improves prediction model performance by better representing the underlying problem. Scikit-Learn's MinMaxScaler was used to scale our dataset to a range of zero to one for feature scaling. The input data is converted to a 3-dimensional array with 60 timestamps and one feature for each step.

2.1. Data collection and analysis

For the goal of high accuracy results in the model, data collection and processing are the key stage at beginning of the model, while tackling data inconsistencies. As data sources, production data from Volve field wells (choke position, downhole pressure and temperature, wellhead pressure, GOR, oil rate, water rate, water injection rate, reservoir pressure and time-series data) will be utilized in the model. As parts of methodology, 3 different models will be used for the production forecasting: ARIMA, Holt Winters and LSTM models.

2.2. ARIMA Model

This study's first model was the auto-Arima model. The model used autoregressive terms, seasonal and non-seasonal differences, and moving averages for the number of lagged forecasts in the prediction equation. The general formula for describing ARIMA models is given in Equation 2.2.1 below:

$$x(t) = \sum_{i=1}^{p} a_i x(t-i) - \sum_{i=1}^{q} \beta_i \varepsilon(t-1) \quad (2.2.1)$$

The model's fit was created using Python's statsmodel module. Hyperparameters were used to optimize the auto-regressive, integrated, and moving average parameters of the generated model. The training dataset was first obtained from the train-test split, conducted during preprocessing. The auto-regressive parameter (p) and moving averages were assigned values between 1 and 3. The integrated parameter (d) was set at 0 with a maximum value of 1. The dataset's seasonality was set to 12 for True Boolean states. A maximum order of 12 was chosen for the model with 50 fits. The ARIMA model's order was determined using Akaike's Information Criterion (AIC).

An example from the ARIMA model:

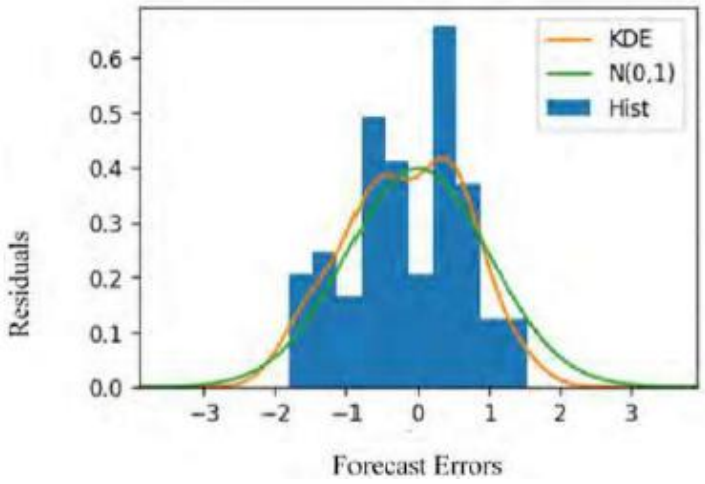

Figure 2.2.1. Normalized residual plot for this model

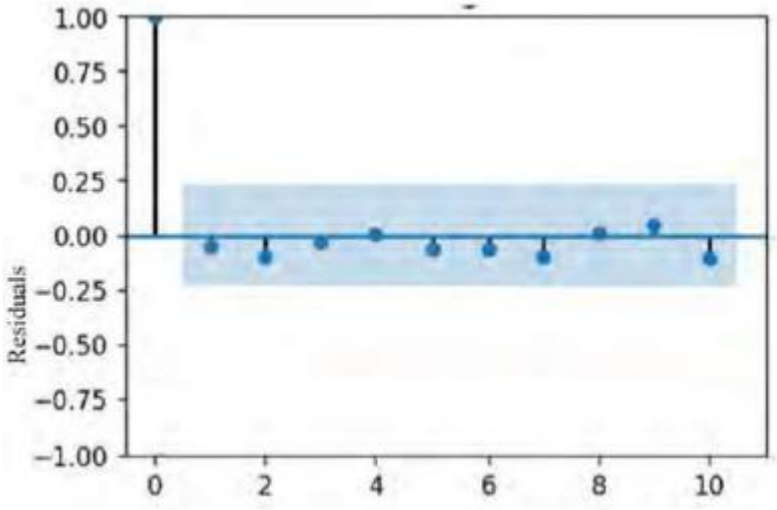

Figure 2.2.2. ARIMA model's standardized Autocorrelation residual plot.

2.3. Holt Winters

The Holt-Winters statistical model, which employs the exponential smoothing procedure, was utilized next in this investigation. Forecasting for this model was based on the previous period's forecast and actual value. Equation 2.3.1 represents the general mathematical expression for basic exponential smoothing.

$$y'_{t+1} = ay_t + (1 - a)y'_{t-1} \quad (2.3.1)$$

In this equation, y' indicates anticipated values at specific intervals, y is the actual value, and an is a smoothing factor ranging from 0 to 1. The Holt-Winters model was selected because it anticipates time series data with both trend and seasonal fluctuations. The model's hyperparameters were tweaked to get an optimal fit. The arguments supplied through the model included dependent variables, seasonal periods, trends, and seasonal states. The dependent variable was the oil production rate, which was adjusted for trend and seasonality.

Examples of the model output from the literature review:

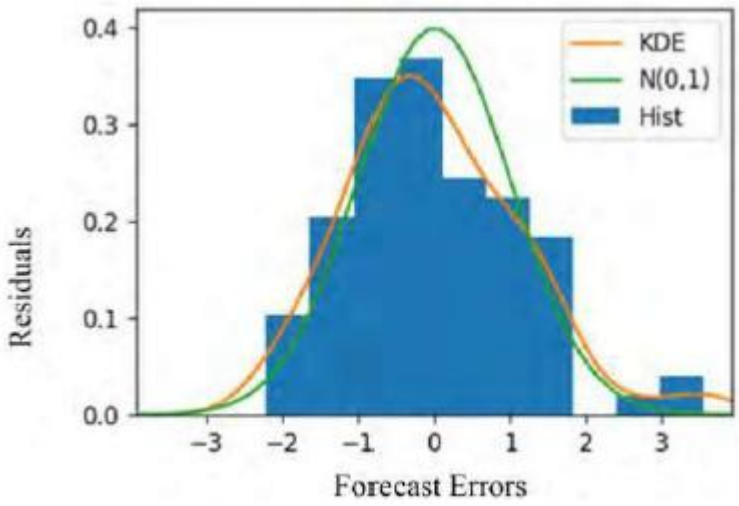

Figure 2.3.1. Distribution plot for Holt Winters model of exponential smoothing.

2.4. LSTM Model

A multivariate LSTM model was also created to anticipate oil output using many observations over a single time step/period. A stacked LSTM approach was utilized, with the amount of time steps and parallel series specified in the input layer. The number of parallel series was used to designate how many values the built model should predict in the output layer. This was counted as two. During model creation, the Mean Absolute Error served as the loss function, and Adam was used to identify the best model selection. Following model design, the training dataset was used to train the model. For this case, there were five hyperparameters: number of epochs, batch size, validation data, verbose, and a Boolean for shuffle. 400 epochs were employed, with a batch size of 36 and two hidden layers. Cross-validation was utilized to each iteration of the model creation process, and the loss for each iteration was calculated.

The training dataset was utilized to train all of the models. The approach produced results, and the models were tested for prediction using the test dataset. The models' performance was tested using statistical error metrics, including MAE, MSE, R^2, and RMSE.

The example output from the model:

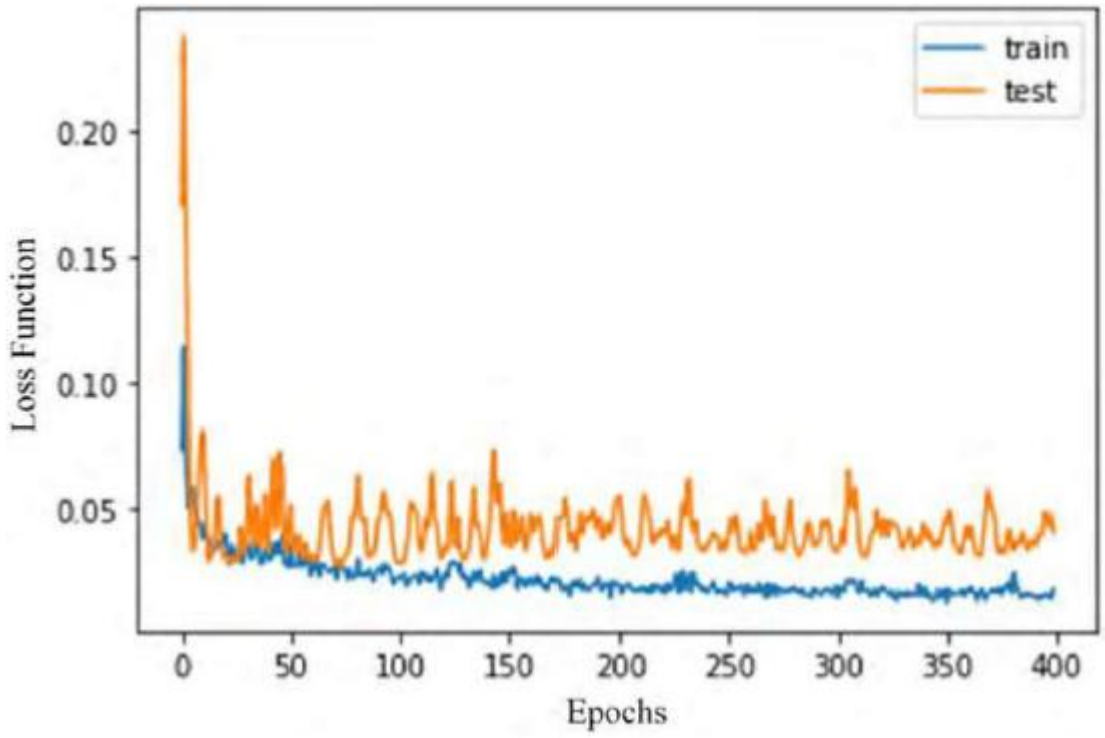

Figure 2.4.1. The losses of value in the LSTM model development.

Figure 2.4.2. Prediction performance of LSTM model.

2.5. XGBoost model

For the prediction, XGBRegressor (XGBoost) model is used, as that is good at handling complexity in production data, robust to outliers and missing parameters and supporting early stopping in order to prevent overfitting. For estimating the uncertainty, HistGradientBoostingRegressor (Scikit-

learn) is utilized. This one deals with uncertainty of the prediction, by applying confidence intervals. This model is more memory-efficient and faster than typical Gradient Boosting, handles NaN values natively and made for quantile regression in comparison to XGBoost.

2.6. Error analysis

The mean absolute percentage error compares the expected values to the actual values. A forecast model with an MAE value around zero is considered good, however zero MSE indicates no error, which is nearly impossible. The mean absolute error for this investigation was calculated using a scale similar to the time series data being simulated. The Mean Absolute Error (MAE) was determined using Equation 2.6.1:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} abs(Ei) \quad (2.6.1)$$

Where, E stands for Error.

The Mean Square Error (MSE) represents the standard deviation of residuals (prediction mistakes). It indicates the proximity of a regression line to a group of points. It calculates the squares of the absolute residuals. A lower MSE, similar to the MAE, indicates a more accurate prediction model. Equation 2.6.2 yields the Mean Square Error (MSE):

$$MAE = \left[\frac{1}{n}\sum_{i=1}^{n} E_I^2\right]^{1/2} \quad (2.6.2)$$

The Root Mean Square Error (RMSE) is the average of the mean square error values in Equation 2.6.2.

The equation 2.6.3 yielded the following relationship:

$$RMAE = \left[\frac{1}{n}\sum_{i=2}^{n} E_i^2\right]^{\frac{1}{2}} (2.6.3)$$

The coefficient of determination (R^2) assesses the correlation between the output and the goal variable.

The coefficient of determination (R^2) value is assessed on a scale of 0 to 1. The R^2 value, comparable to the coefficient of correlation, indicates the strength of a linear relationship between

expected output and target variables. Equation 2.6.1 provides the mathematical expression for the coefficient of determination:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - y')^2}{\sum_{i=1}^{n} E_i^2} \quad (2.6.1)$$

Where, y' shows the mean of the values, while y_i distributes the actual target values. The variance of the data is proportional to the numerator.

2.7. Output Generation

To develop a production forecast, the model is evaluated using the testing data set. After a few epochs, the model accurately predicts based on the shape and position of the historical time series. To evaluate our simulation results, we select an existing well that was previously unknown to the LSTM model and anticipate its production. The prediction profile from the Neural Network is compared to the actual production profile to determine the trend of the findings.

2.8. Visualization

Data visualization uses visual tools like charts, graphs, and maps to highlight trends, anomalies, and patterns in historical well data. Data visualization is the visual display of data. Matplotlib, a Python-based charting package, was utilized in this study to show projected and actual production profiles.

RESULTS

As was mentioned in the above sections, the production dataset of Volve field belonging to Equinor will be used. Before showing the results, it is good to show what has been provided from publicly available wells. Figure 3.1 shows the water and oil production from each well. The 15/9-F-5 well is on the list of both injector and producer. It was an injector, then converted into an oil producer. 15/9-F-12 and 15/9-F-14 are the producers which produced the longest time. 15/9-F-14 has been selected as one of the longest producing wells and has quite a long production history.

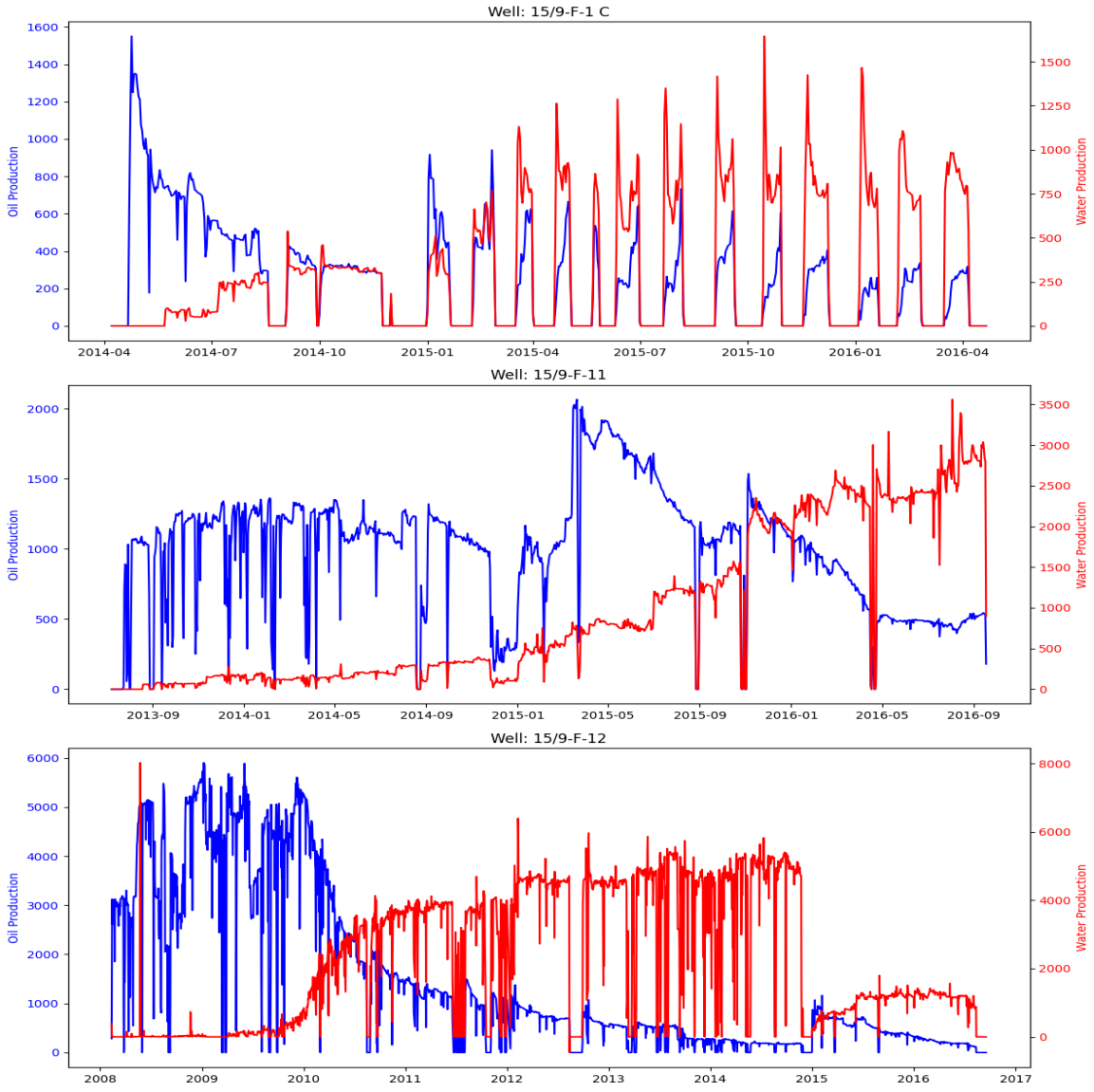

Figure 3.1. Oil and water production from the wells

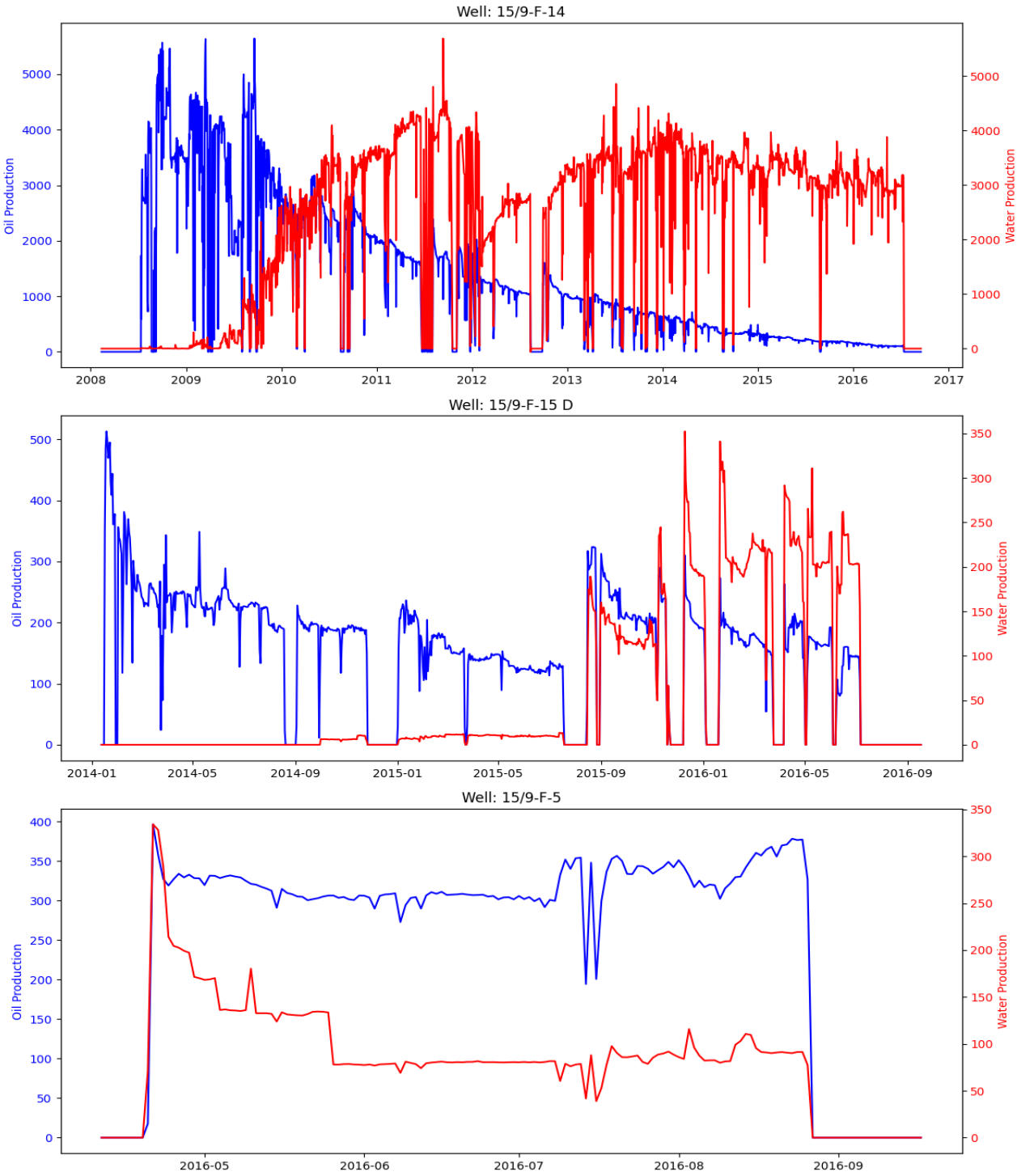

Figure 3.2. Oil and water production from the wells

If it is visualized in the form of pie chart, the contribution of produced reservoir fluids from each well can be shown below:

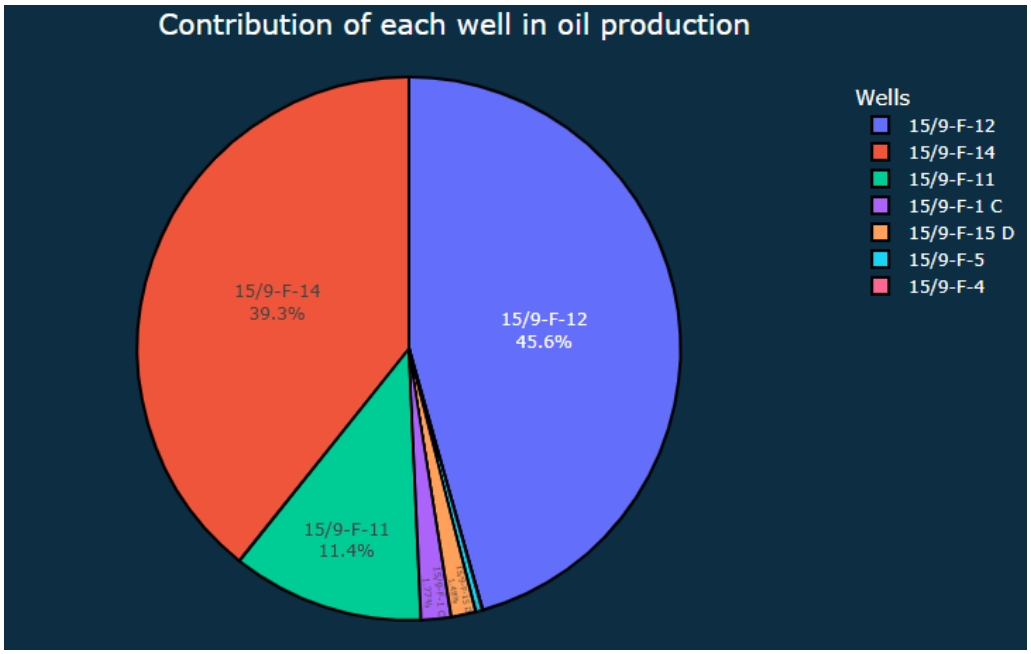

Figure 3.3. The contribution of oil production per well.

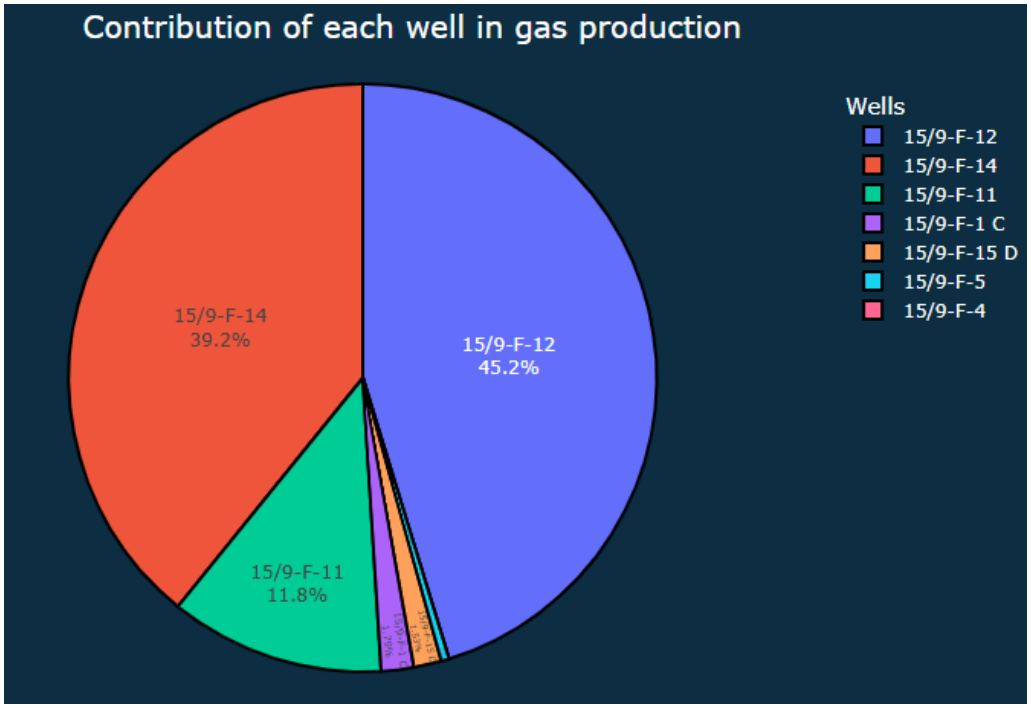

Figure 3.4. Contribution of gas production per well

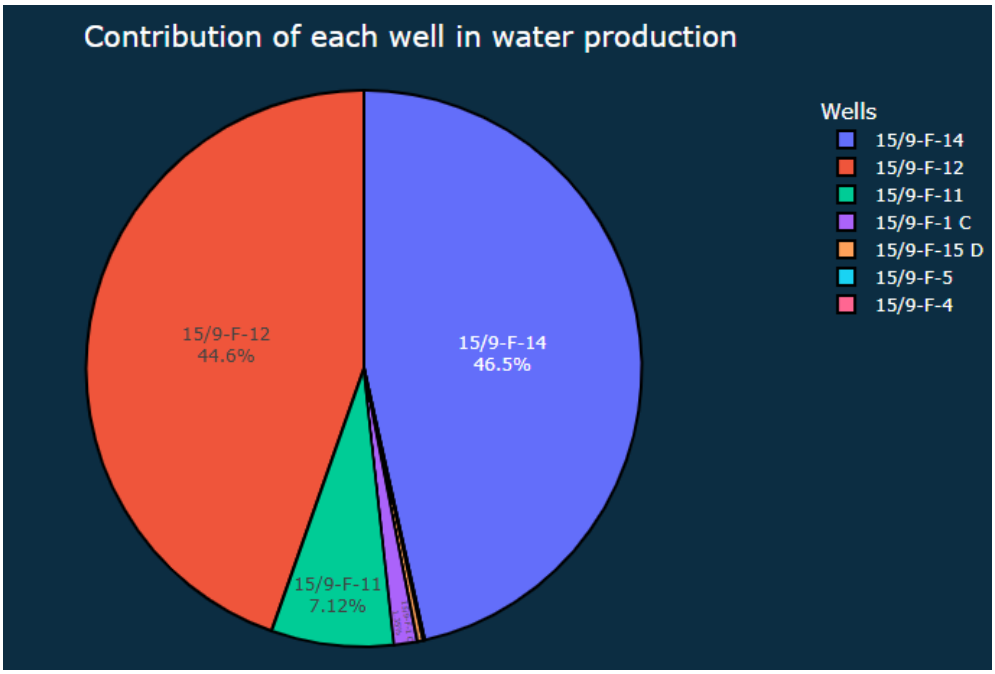

Figure 3.5. Contribution of water production per well.

For the optimization of performance of the model, the extraction of all important features from the dataset is needed. The data for the model optimization is both temporal and spatial. To build a comprehensive model of the reservoir, data extraction from surface related operational conditions, PTA results, different petrophysical logs (open and cased holes), well tests, survey of temperature, injection and production history, well design details, core data, etc are needed.

But for this research, the focus will be on Machine Learning rather than extraction of the feature. Dynamic (temporal) data will be utilized: downhole pressure, downhole temperature, tubing pressure, annulus pressure, choke size, wellhead pressure, oil, water and gas production data (Figure 3.6):

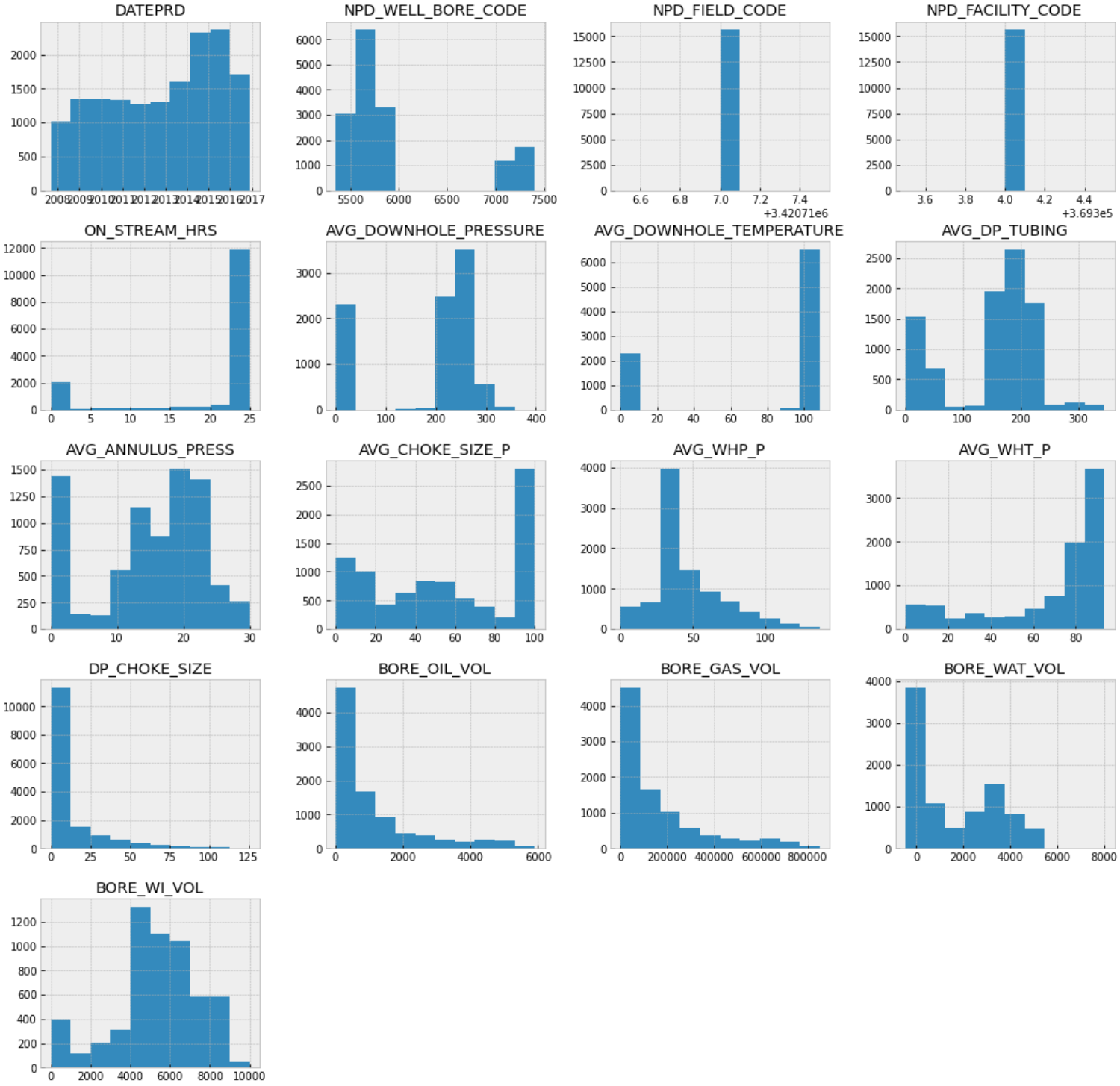

Figure 3.6. The distribution of the data needed for consideration of production prediction

For the production forecasting ML model, the production data of 15/9-F-14 will be utilized. The plots for production of oil, water and gas versus the date have been provided below:

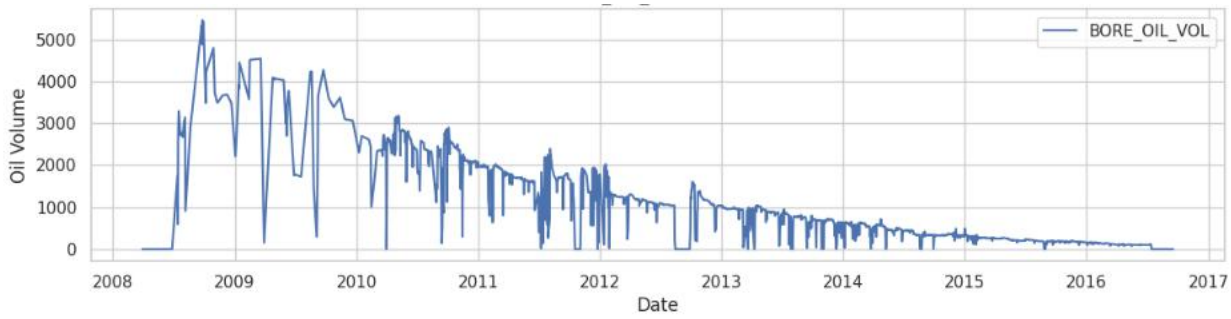

Figure 3.7. Oil production distribution of 15/9-F-14.

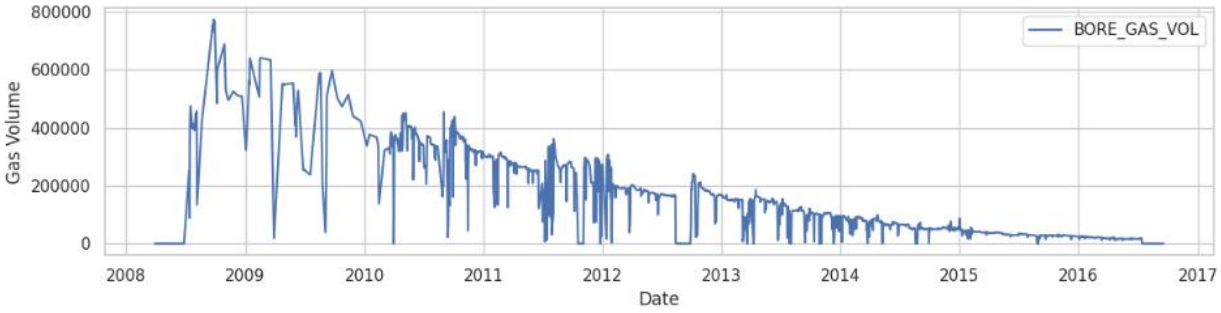

Figure 3.8. Gas production distribution of 15/9-F-14.

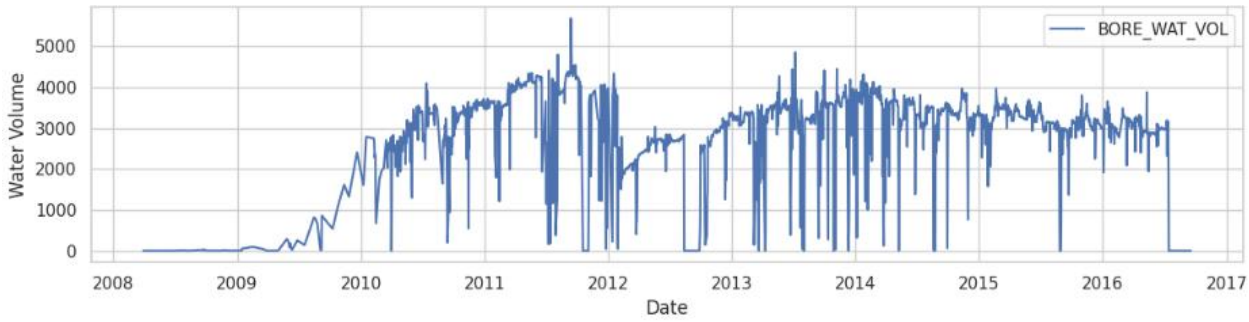

Figure 3.9. Water production distribution of 15/9-F-14.

In LSTM, the previous time step data is fed, and current time step's target is predicted. For this case, given production data will be fed into the previous step, and we will forecast the current one for water, gas and oil production. If the data will be separated into two parts: features and targets. LSTM utilizes the data type of NumPy; hence the data must be converted into DataFrame type to NumPy data type. Sklearn will be utilized to split data into 70% for training, 15% for validation, and 15% for test data.

Visualized form of the data after split is shown below plot:

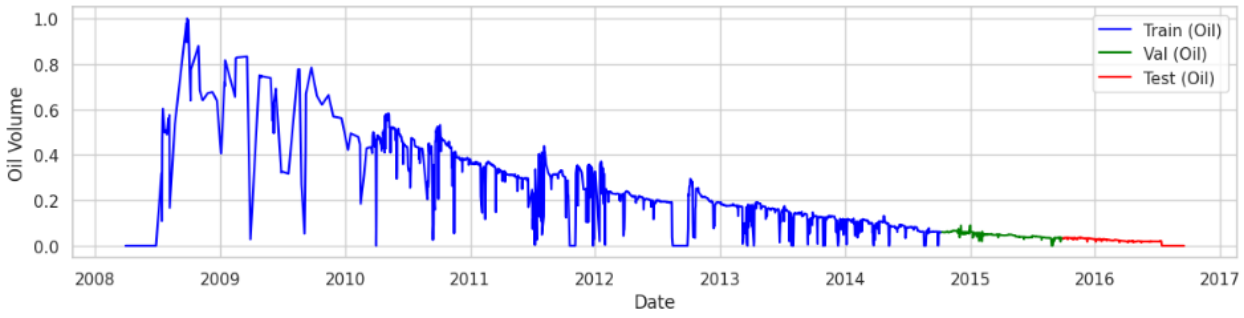

Figure 3.10. Training, validation and test data for oil rate

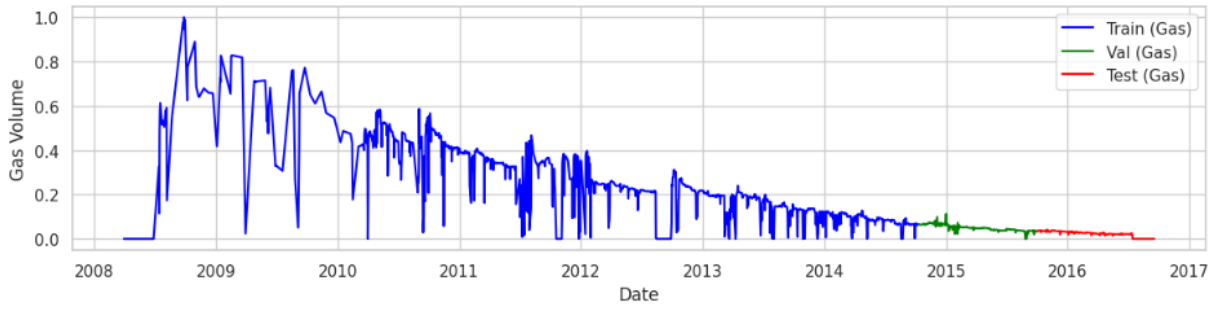

Figure 3.11. Training, validation and test data for gas rate

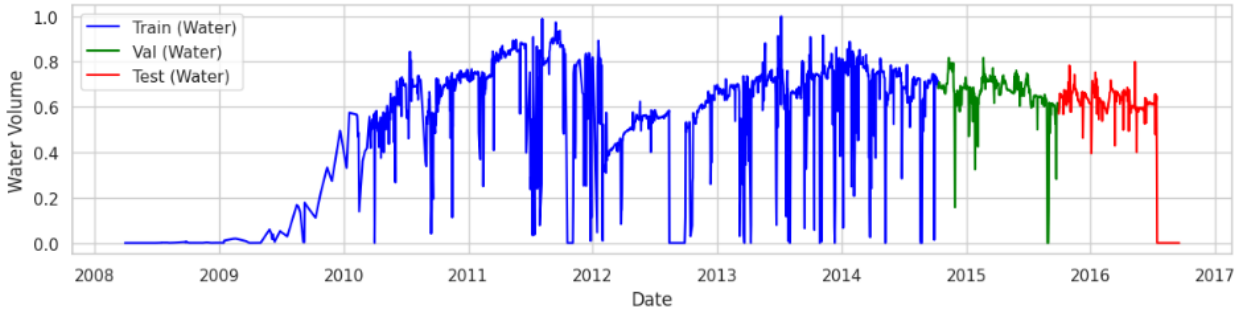

Figure 3.12. Training, validation and test data for water rate

After testing, it has been clear that utilization of 5 time steps (days) for input and only one time step (day) for output provides the desirable result. Consequently, the data will be:

- The features for the 5 days and the target for the 6^{th} day.
- The next is that all features for next 5 days (2^{nd} to 6^{th} days) and target of the 7^{th} day and in this order, these steps will continue.

To provide better understanding, for the first batch, sample data can be examined:

- As input, 5 time steps, each one containing a dozen of features, as output, single time step, which has 3 targeted values.
- It is noteworthy to mention that the last 3 data from 12 input data are related to production of oil, gas and water for the previous time step, accordingly.
- The targets for the current one then turn into input for the next step of time. The following steps of time will be continued in this pattern.

While starting training model, EarlyStopping from TensorFlow is utilized to stop the training early in case of not making significant improvements on data validation stage. After the training of the model, EarlyStopping is activated in around 220 epochs. The following plot shows the losses for training and validation:

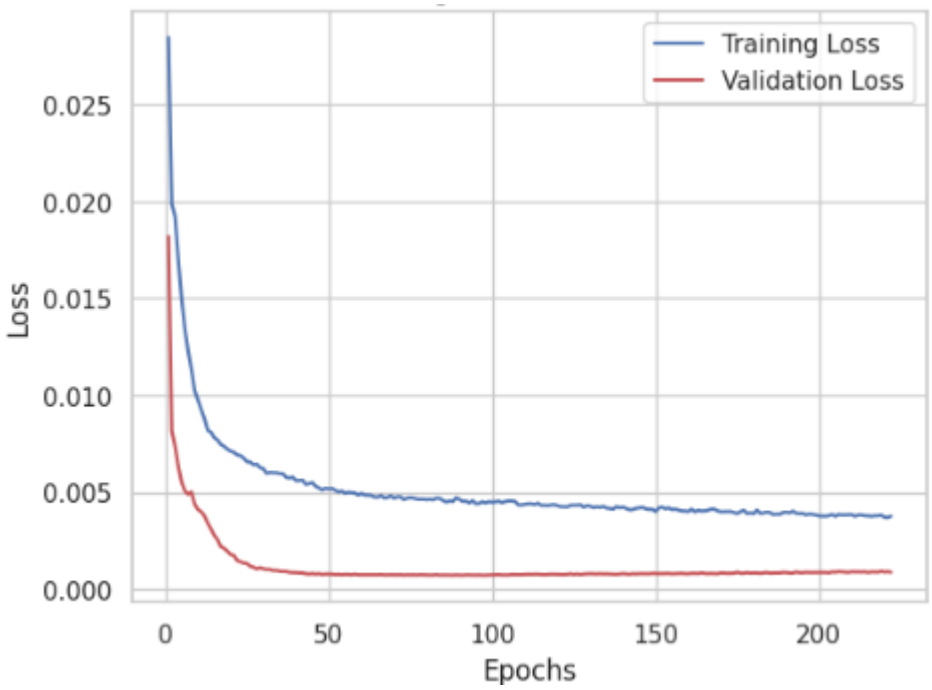

Figure 3.13. The training and validation data losses.

Now the trained model can be used for prediction for production on the test data.

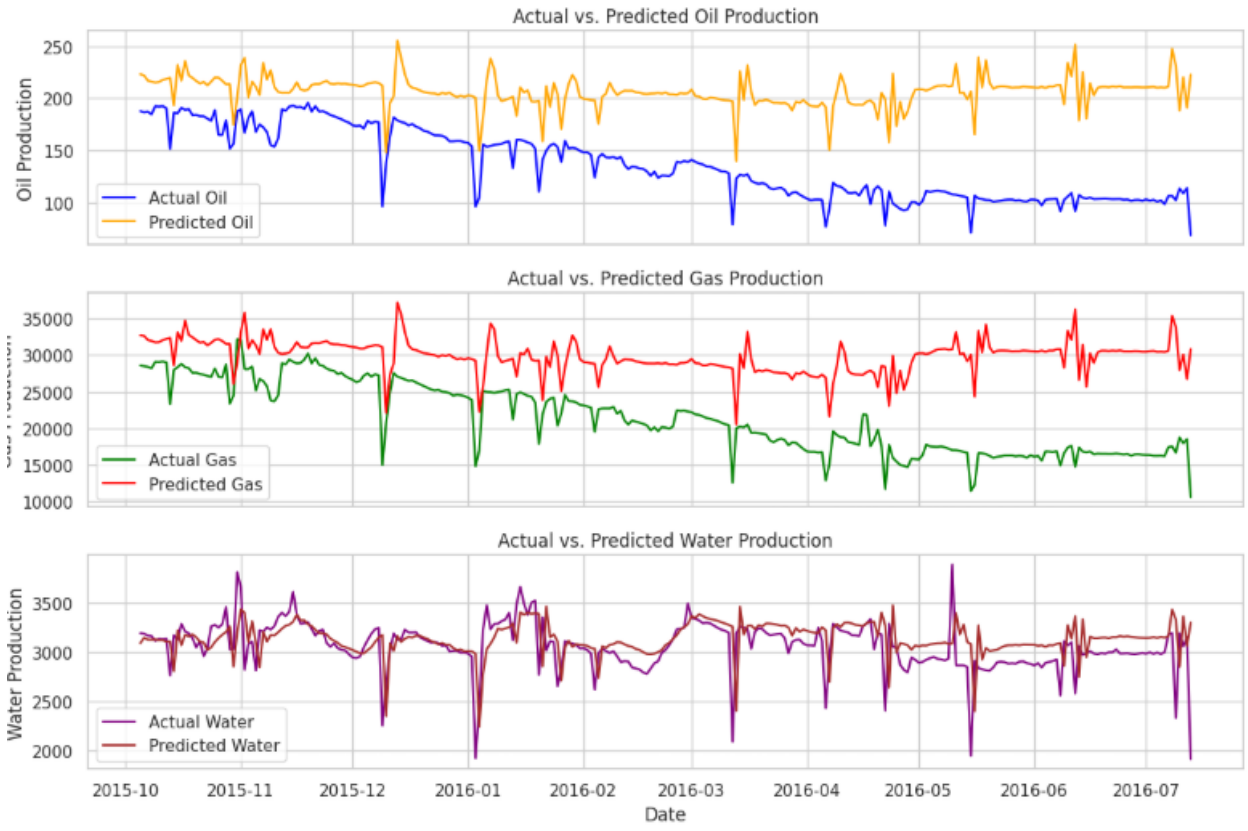

Figure 3.14. Actual and predictable data for the produced fluids.

From the plots, it is seen that the model has very good prediction for, however that for oil and gas near to the end of the production deviation starts to increase.

For the comparison, another ML model can be used, as LSTM shows some imperfections. For the prediction, XGBRegressor (XGBoost) model is used, as that is good at handling complexity in production data, robust to outliers and missing parameters and supporting early stopping in order to prevent overfitting. For estimating the uncertainty, HistGradientBoostingRegressor (Scikit-learn) is utilized. This one deals with uncertainty of the prediction, by applying confidence intervals. This model is more memory-efficient and faster than typical Gradient Boosting, handles NaN values natively and made for quantile regression in comparison to XGBoost. The combo of two models provides the following results (Figure 3.15-3.17)

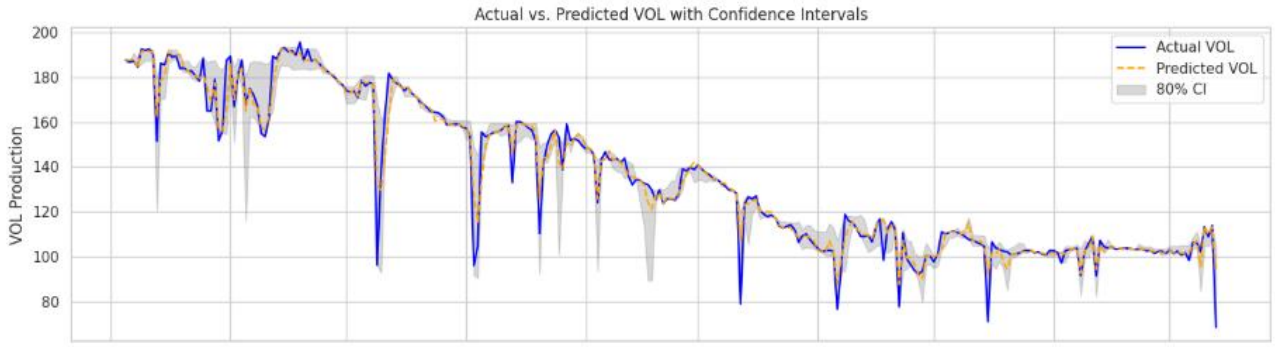

Figure 3.15. Actual and predicted production for oil with 80% confidence

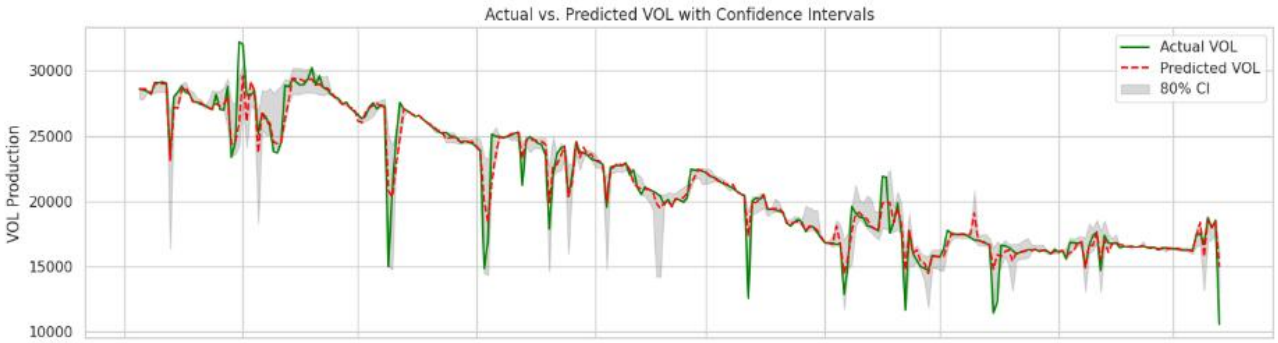

Figure 3.16. Actual and predicted production for gas with 80% confidence

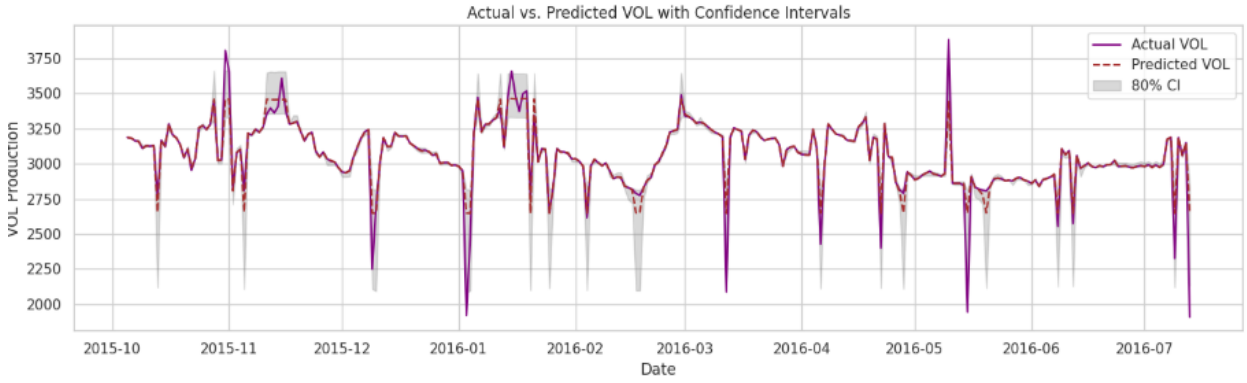

Figure 3.17. Actual and predicted production for water with 80% confidence

When it comes to the comparison of the models with numbers, the following table (Table 3.1) and the bar chart summarize all the prediction accuracies:

Table 3.1. Comparison of the models:

ML Model	Fluid	MAE	RMSE	R2
XGBoost	Oil	0.75	1.60	0.9975
	Gas	132.18	355.28	0.9941
	Water	0.71	2.88	0.9999
LSTM	Oil	7.53	12.15	0.8535
	Gas	1255.85	2007.62	0.8114
	Water	150.48	239.90	0.1069

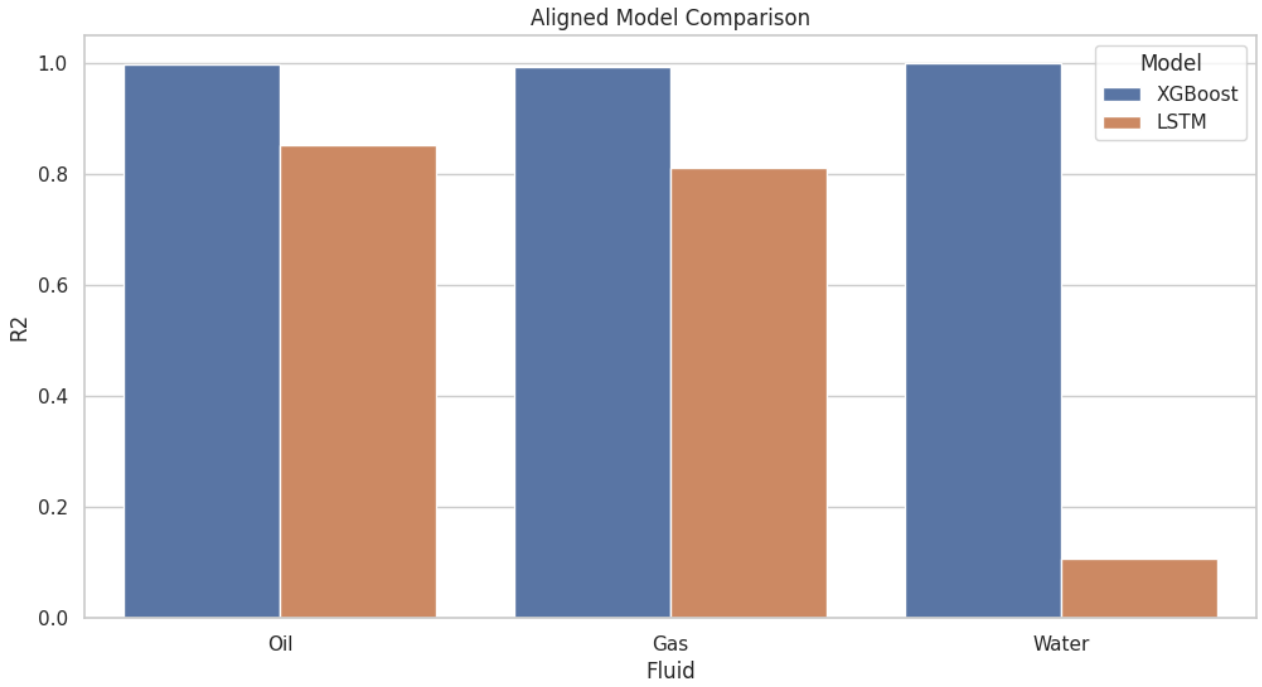

Figure 3.18. The visual distribution of model comparisons.

CONCLUSION

In this paper, we tested and compared two different methods, XGBoost and Long Short-Term Memory (LSTM), for time-series forecasting of oil, gas, and water production in petroleum reservoirs. The real-field data comparison highlights drastic differences in accuracy, computing costs, and reliability, of relevance not exclusively at field level, but also at corporate level.

The findings indicated that XGBoost had a superior performance to LSTM in terms of the target variables. It was highly accurate in oil production prediction ($R^2 = 0.997$, MAE < 1 bbl/day) with similar performances for gas ($R^2 = 0.994$, MAE = 132 Mcf/day) and water ($R^2 = 0.999$, MAE = 0.71 bbl/day). Such accuracy allows reservoir engineers to make decisions with confidence about XGBoost predictions within a short and mid-term operational plan - e.g. production allocation, well intervention scheduling, lift optimization. Furthermore, the stability of the model with multiphase flow and varying pressure demonstrates the applicability of the model in complex reservoir. On the other hand, the LSTM model tended to fail to figure out the production tuning, especially for water ($R^2 = 0.107$) and demanded more computational burden, resulting in less accurate and less efficient outputs.

From an economic perspective, the backtesting differential between the two models has measurable value. For a 10,000 BOPD producing field, a 1% improvement in forecast accuracy can save around $1 million a year (by minimizing over or underproduction and better utilizing the surface facilities). On the same note, accurate gas prediction enables limited flaring, and the potential to save up to $300,000 in annual regulatory penalties and lost revenues based on the current gas prices. Near-perfect water prediction is not only used for forward planning for injections and disposals, but it is also helping to avoid unnecessary handling charges and to reduce problems, such as scaling and corrosion, that can extend equipment life, and that often reduces capex over a decade by up to 30%.

In addition to delivering direct cost savings, the implications of this work are of interest in the context of asset planning and strategic decision making. Accurate predictions allow well workovers to be prioritized by indicating poorly producing wells or re-fracturing candidates. For instance, a single well that has the potential to increase production by 5% could yield more than $2,000,000 in additional revenue per year. Precise forecasting also underpins risk management,

such as hedging financial risk based on production certainty, and this improves the precision of reserve reporting under regulations such the SEC or NPR, lower exposure to litigation or audit risk.

Currently, this work provides insights into future deployment. Fleetwide integration of XGBoost can be rolled out in the short term across live assets by ensuring that it flows to SCADA systems and real-time dashboards to optimize daily operations. Educating field engineers to interpret SHAP values and feature importance metrics can drive ML democratization and improve the cross-collaboration between domains. In the medium term, hybrid models such as using XGBoost alongside reservoir simulators could help long-term depletion forecasting – and the underperforming LSTM may be useful for detecting anomalies as its sensitivity to 'odd' patterns becomes an advantage. More-long term strategies may involve things like incorporating these models into digital twin models or tying gas forecasting to flare power projects which are connected to sustainability and ESG initiatives.

In summary, the results strongly demonstrate that XGBoost is a better machine learning model for data-driven production forecasting than the studied reservoir setting due to its unprecedented accuracy, interpretability and model efficiency. Although deep learning remains promising for future applications—such as when a hybrid model is trained and/or a larger dataset is available-current evidence indicates a preference for ensembled methods such as XGBoost in practical deployment. The findings not only improve the technical understanding of ML model performance in reservoir forecasting but also provide a practical guideline for embedding data science in asset optimization, capital planning, and long-term field development strategy.

REFERENCES

1. Accenture. (2017). Drill Deeper into Digital. 2017 Upstream Oil and Gas Digital Trends Survey: Key Findings.

2. Bughin, J., Hazan, E., Sree Ramaswamy, P., DC, W., & Chu, M. (2017). Artificial intelligence the next digital frontier.

3. Salman, Y., Gauder, D., & Turnbull, W. B. (2018, November). Digitalization of a Giant Field–The Rumaila Story. In *Abu Dhabi International Petroleum Exhibition and Conference* (p. D031S066R002). SPE.

4. Poddar, T. (2018, March). Digital twin bridging intelligence among man, machine and environment. In *offshore technology conference Asia* (p. D022S001R003). OTC.

5. Shahkarami, A., Mohaghegh, S. D., Gholami, V., & Haghighat, S. A. (2014, April). Artificial intelligence (AI) assisted history matching. In *SPE Western Regional Meeting* (pp. SPE-169507). SPE.

6. Eltahan, E., Yu, W., Sepehrnoori, K., Kerr, E., Miao, J., & Ambrose, R. (2019, April). Modeling Naturally and Hydraulically Fractured Reservoirs with Artificial Intelligence and Assisted History Matching Methods Using Physics-Based Simulators. In *SPE Western Regional Meeting* (p. D041S013R002). SPE.

7. Lucchesi, R. D. (2019, April). Main factors impacting oil projects return: A sensitivity analysis. In *Offshore Technology Conference* (p. D041S051R001). OTC.

8. Mantatzis, K., Haake, M., Clemens, T., & Steineder, D. (2019, October). Reservoir Characterization, Incremental Production Forecasting and Decision Making of an Extended Reach Well Project Under Uncertainty. In *SPE Russian Petroleum Technology Conference* (p. D033S022R003). SPE.

9. Crawford, P. B. (1960). Factors affecting waterflood pattern performance and selection. *Journal of Petroleum Technology*, *12*(12), 11-15.

10. Tannenbaum, A. S., Kavcic, B., Rosner, M., Vianello, M., & Weiser, G. (1977). Hierarchy in organizations. *Nursing Administration Quarterly*, *1*(4), 87-88.

11 Buckley, S. E., & Craze, R. C. (1943, January). The Development and Control of Oil Reservoirs. In *Drilling and Production Practice*. OnePetro.

12. Khandwala, S. M., Rani, A. M. A., & Rahman, M. N. A. (1984, February). Field Development Planning For The Semangkok Field Offshore Pen Insular Malaysia. In *SPE Offshore South East Asia Show* (pp. SPE-12410). SPE.

13. Crookston, R. B., Culham, W. E., & Chen, W. H. (1979). A numerical simulation model for thermal recovery processes. *Society of Petroleum Engineers Journal*, *19*(01), 37-58.

14. Bennett, P. M. (1981, March). Economic Field Development Planning. In *SPE Permian Basin Oil and Gas Recovery Conference* (pp. SPE-9718). SPE.

15. Poston, S. W., & Gross, S. J. (1986). Numerical simulation of sandstone reservoir models. *SPE Reservoir Engineering*, *1*(04), 423-429.

16. Nixon, J., & Pena, E. (2019, April). The evolution of asset management: Harnessing digitalization and data analytics. In *Offshore Technology Conference* (p. D021S016R005). OTC.

17. Ibrahimov, T. (2015, November). History of history match in Azeri field. In *SPE Annual Caspian Technical Conference* (pp. SPE-177395). SPE.

18. Scheidt, C., Li, L., & Caers, J. (Eds.). (2018). *Quantifying uncertainty in subsurface systems*. John Wiley & Sons.

19. Evans, R. (2000, October). Decision analysis for integrated reservoir management. In *SPE Europec featured at EAGE Conference and Exhibition?* (pp. SPE-65148). SPE.

20] Hurst, A., Brown, G. C., & Swanson, R. I. (2000). Swanson's 30-40-30 Rule. *AAPG bulletin*, *84*(12), 1883-1891.

21. Malhotra, V., Lee, M. D., & Khurana, A. (2004, October). Decisions and uncertainty management: Expertise matters. In *SPE Asia Pacific oil and gas conference and exhibition* (pp. SPE-88511). SPE.

22. Sieberer, M., Clemens, T., Peisker, J., & Ofori, S. (2019). Polymer-flood field implementation: Pattern configuration and horizontal vs. vertical wells. *SPE Reservoir Evaluation & Engineering*, *22*(02), 577-596.

23. Scheidt, C., & Caers, J. (2009). Uncertainty quantification in reservoir performance using distances and kernel methods—application to a West Africa deepwater turbidite reservoir. *SPE Journal*, *14*(04), 680-692.

24. Thiele, M. R., & Batycky, R. P. (2016, September). Evolve: A linear workflow for quantifying reservoir uncertainty. In *SPE Annual Technical Conference and Exhibition?* (p. D021S022R007). SPE.

25. Schnetler, R., Steyn, H., & Van Staden, P. J. (2015). Characteristics of matrix structures, and their effects on project success. *South African Journal of Industrial Engineering*, *26*(1), 11-26.

26. Begg, S. H., Bratvold, R. B., & Campbell, J. M. (2003, October). Shrinks or quants: who will improve decision-making. In *SPE Annual Technical Conference and Exhibition?* (pp. SPE-84238). SPE.

27. Sieberer, M., Jamek, K., & Clemens, T. (2017). Polymer-flooding economics, from pilot to field implementation. *SPE Economics & Management*, *9*(03), 51-60.

28. Steineder, D., & Clemens, T. (2020, December). Quantifying the Importance of Decisions and the Impact of Data Acquisition on Decisions in Hydrocarbon Field Re-Developments. In *SPE Europec featured at EAGE Conference and Exhibition?* (p. D011S005R003). SPE.

29. Steineder, D., Clemens, T., Osivandi, K., & Thiele, M. R. (2019). Maximizing the Value of Information of a Horizontal Polymer Pilot Under Uncertainty Incorporating the Risk Attitude of the Decision Maker. *SPE Reservoir Evaluation & Engineering*, *22*(02), 756-774.

30. Cao, Q., Banerjee, R., Gupta, S., Li, J., Zhou, W., & Jeyachandra, B. (2016, June). Data driven production forecasting using machine learning. In *SPE Argentina Exploration and Production of unconventional resources symposium* (p. D021S006R001). SPE.

31. Kumar, A. (2019, April). A machine learning application for field planning. In *Offshore Technology Conference* (p. D021S026R005). OTC.

32. Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., ... & Kumar, V. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on knowledge and data engineering*, *29*(10), 2318-2331.

33. Aghina, W.; Ahlbiick, K. de Smet, A.; Fahrbach, C.; Handscomb, C.; Lackey, G.; Luri, M.;

Murarka, M.; Salo, O.; Seem, E. and J. Woxholth. 2017. The 5 Trademarks of Agile Organizations. McKinsey and Company. (December 2017).

34. Yoder, R. T. (2019, April). Digitalization and data democratization in offshore drilling. In *Offshore Technology Conference* (p. D032S058R008). OTC.

35. Anand, B., & Krishna, K. (2019, April). Overcoming hurdles to accelerate data-centric ways of working in the energy industry. In *Offshore Technology Conference* (p. D021S016R006). OTC.

36. Bisson, P.; Hall, B.; McCarthy, B. and K. Rifai. (2018). Breaking away: The secrets to scaling analytics. McKinsey Global Institute.

37. Handscomb, C., Sharabura, S., & Woxholth, J. (2016). The oil and gas organization of the future. *McKinsey & Company*.

38. Nandurdikar, N., & Wallace, L. (2011, October). Failure to produce: An investigation of deficiencies in production attainment. In *SPE Annual Technical Conference and Exhibition?* (pp. SPE-145437). SPE.

39. Tsuchiya, S. (1993). Artificial Intelligence and Organizational Learning — How can Al contribute to Organizational Learning? AAAI Technical Report WS-93-03.

40. Shepperd, M., Guo, Y., Li, N., Arzoky, M., Capiluppi, A., Counsell, S., ... & Yousefi, L. (2019). The prevalence of errors in machine learning experiments. In *Intelligent Data Engineering and Automated Learning–IDEAL 2019: 20th International Conference, Manchester, UK, November 14–16, 2019, Proceedings, Part I 20* (pp. 102-109). Springer International Publishing.

41. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.

42. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat, F. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195-204.

43. Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *nature*, *529*(7587), 484-489.

44. Nash Jr, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, *36*(1), 48-49.

45. Agrawal, A., & Jaiswal, D. (2012). When machine learning meets ai and game theory. *Comput. Sci*, *2012*, 1-5.

46. Nowé, A., Vrancx, P., & De Hauwere, Y. M. (2012). Game theory and multi-agent reinforcement learning. In *Reinforcement learning: State-of-the-art* (pp. 441-470). Berlin, Heidelberg: Springer Berlin Heidelberg.

47. Chui, M., Manyika, J., & Miremadi, M. (2018). What AI can and can't do (yet) for your business. *McKinsey Quarterly*, *1*(97-108), 1.

48. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.

49. Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, *8*(8), 832.

50. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, *51*(5), 1-42.

51. Clemens, T., & Viechtbauer-Gruber, M. (2020). Impact of digitalization on the way of working and skills development in hydrocarbon production forecasting and project decision analysis. *SPE Reservoir Evaluation & Engineering*, *23*(04), 1358-1372.

52. Omekara, C. O., Okereke, O. E., Ire, K. I., & Okamgba, C. O. (2015). ARIMA modeling of Nigeria crude oil production. *Journal of Energy Technologies and Policy*, *5*(10), 1-5.

53. Omotosho, T. J. (2024, August). Oil Production Prediction Using Time Series Forecasting and Machine Learning Techniques. In *SPE Nigeria Annual International Conference and Exhibition* (p. D032S028R006). SPE.

54. Fatoki, O., Ilo, H. O., & Ugwoke, P. O. (2017). Time series analysis of crude oil production in Nigeria. *Adv. Multidiscip. Sci. Res*, *3*(3), 1-10.

55. Balogun, O.S and Ogunleye, O.M. (2015). "On the Time Series Modelling of Crude Oil Exportation in Nigeria" European Journal of Academic Essays 2(3): 1-11, 2015.

56. Sadeeq, S. A., & Ahmadu, A. O. (2018). A Study of The Suitable Time Series Model for Monthly Crude Oil Production in Nigeria. *International Journal of Mathematics and Physical Sciences Research*, *6*(1), 106-112.

57. Tadjer, A., Hong, A., & Bratvold, R. B. (2021). Machine learning based decline curve analysis for short-term oil production forecast. *Energy Exploration & Exploitation*, *39*(5), 1747-1769.

58. Wan, X., Zou, Y., Wang, J., & Wang, W. (2021, August). Prediction of shale oil production based on Prophet algorithm. In *Journal of Physics: Conference Series* (Vol. 2009, No. 1, p. 012056). IOP Publishing.

59. Luo, G., Tian, Y., Sharma, A., & Ehlig-Economides, C. (2019, March). Eagle ford well insights using data-driven approaches. In *International Petroleum Technology Conference* (p. D021S026R003). IPTC.

60. Obite, C. P., Chukwu, A., Bartholomew, D. C., Nwosu, U. I., & Esiaba, G. E. (2021). Classical and machine learning modeling of crude oil production in Nigeria: Identification of an eminent model for application. *Energy Reports*, *7*, 3497-3505.

61. Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, *20*(8), 832-844.

62. Obite, C. P., Chukwu, A., Bartholomew, D. C., Nwosu, U. I., & Esiaba, G. E. (2021). Classical and machine learning modeling of crude oil production in Nigeria: Identification of an eminent model for application. *Energy Reports*, *7*, 3497-3505.

63. Taylor, S.J. and Letham, B. (2007) Forecasting at scale.

64. Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, *323*, 203-213.

65. Ediger, V. Ş., Akar, S., & Uğurlu, B. (2006). Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model. *Energy Policy*, *34*(18), 3836-3846.

66. Alimohammadi, H., Rahmanifard, H., & Chen, N. (2020, October). Multivariate time series modelling approach for production forecasting in unconventional resources. In *SPE Annual Technical Conference and Exhibition?* (p. D022S061R030). SPE.

67. Gupta, S., Fuehrer, F., & Jeyachandra, B. C. (2014, September). Production forecasting in unconventional resources using data mining and time series analysis. In *SPE Canada Unconventional Resources Conference* (p. D011S004R008). SPE.

68. Al-Shabandar, R., Jaddoa, A., Liatsis, P., & Hussain, A. J. (2020). A deep gated recurrent neural network for petroleum production forecasting. *Machine Learning with Applications, 3*, 100013.

69. Morgan, E. (2018, October). Accounting for serial autocorrelation in decline curve analysis of Marcellus shale gas wells. In *SPE Eastern Regional Meeting* (p. D043S008R001). SPE.

70. Jayeola, I., Olusola, B., & Orodu, K. (2022, August). Machine Learning Prediction Versus Decline Curve Prediction: A Niger Delta Case Study. In *SPE Nigeria Annual International Conference and Exhibition* (p. D021S009R002). SPE.

71. Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, *323*, 203-213.

72. Huang, R., Wei, C., Wang, B., Li, B., Yang, J., Wu, S., ... & Sudakov, V. (2021, November). A comprehensive machine learning approach for quantitatively analyzing development performance and optimization for a heterogeneous carbonate reservoir in Middle East. In *SPE Eastern Europe Subsurface Conference* (p. D012S001R001). SPE.

73. Tadjer, A., Hong, A., & Bratvold, R. (2022). Bayesian deep decline curve analysis: A new approach for well oil production modeling and forecasting. *SPE Reservoir Evaluation & Engineering*, *25*(03), 568-582.

74. Gupta, I., Samandarli, O., Burks, A., Jayaram, V., McMaster, D., Niederhut, D., & Cross, T. (2021, December). Autoregressive and Machine Learning Driven Production Forecasting–Midland Basin Case Study. In *Unconventional Resources Technology Conference, 26–28 July 2021* (pp. 3104-3117). Unconventional Resources Technology Conference (URTeC).

75. Cui, A., Yanke, A., Dao, T., Ye, P., Cross, T., & Davis, B. (2024, November). Machine learning vs. type curves in the Appalachian Basin: A comparative study. In *Unconventional Resources Technology Conference, 17–19 June 2024* (pp. 1882-1891). Unconventional Resources Technology Conference (URTeC).

76. Suwito, E., Sianturi, J. A. D., Irawan, A., Panjaitan, P. R., Saeed, Y., Priyanto, A. D., & Elfeel, M. A. (2022, October). Novel machine learning and data analytics approach for history matching

giant mature multilayered oil field. In *Abu Dhabi International Petroleum Exhibition and Conference* (p. D031S073R002). SPE.

77. Bagheri, M., Zhao, H., Sun, M., Huang, L., Madasu, S., Lindner, P., & Toti, G. (2020, May). Data conditioning and forecasting methodology using machine learning on production data for a well pad. In *Offshore technology conference* (p. D031S037R002). OTC.

78. Haghighat, A., & Burrough, T. (2024, November). Enhanced decision making and asset optimization for unconventional resources with type wells (DCA), RTA-based numerical modeling and machine learning. In *Unconventional Resources Technology Conference, 17–19 June 2024* (pp. 3190-3203). Unconventional Resources Technology Conference (URTeC).

79. Jamalbayov, M. A., Valiyev, N. A., Ibrahimov, K. M., Babayev, M., & Novruzova, S. H. (2024). ENERGY AND EFFICIENCY OPTIMIZATION IN SUCKER-ROD PUMPING USING DISCRETE-IMITATION MODELING CONCEPT: APPLICATION TO WELL OPERATIONS IN THE BIBI-EIBAT FIELD OF AZERBAIJAN.