

**RESERVOIR PERFORMANCE FORECASTING USING SIMULATION-BASED
UNCERTAINTY ANALYSIS IN PYTHON**



**IYANGBE JOSHUA ESOSA
ENG2002617**

**DEPARTMENT OF PETROLEUM ENGINEERING
FACULTY OF ENGINEERING
UNIVERSITY OF BENIN
BENIN CITY**

NOVEMBER 2025

**RESERVOIR PERFORMANCE FORECASTING USING SIMULATION-BASED
UNCERTAINTY ANALYSIS IN PYTHON**



**IYANGBE JOSHUA ESOSA
ENG2002617**

**A PROJECT SUBMITTED TO THE DEPARTMENT OF PETROLEUM
ENGINEERING, FACULTY OF ENGINEERING, UNIVERSITY OF BENIN.
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF
BACHELOR OF ENGINEERING (B.ENG) DEGREE.**

NOVEMBER 2025

CERTIFICATION

This is to verify that this project work was carried out by Iyangbe Joshua Esosa with the matriculation number: ENG2002617 in the department of petroleum engineering. It is adequate and satisfactory, both in scope and content, for the award of Bachelor of Engineering (B.ENG) Degree in Petroleum Engineering of the University of Benin.

Prof O. A. OLAFUYI
(Project Supervisor)

DATE

DR. OLUWASEUN TAIWO
(Project Coordinator)

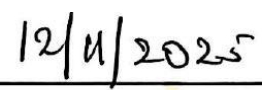
DATE

DR. IKPOMWOSA OHENHEN
(Head Of Department)

DATE



Prof. KEVIN C. IGWILO
(External Supervisor)



DATE

DEDICATION

This project is dedicated to God Almighty, the giver of knowledge and understanding for His protection, provision, guidance and strength He accorded throughout the research and academic pursuit, and for the wisdom and grace to complete this project successfully.

I also dedicate it to my family, especially my parents, for their unwavering support, prayers, and encouragement.

ACKNOWLEDGEMENT

The Almighty God surely merits my hearty thanks and gratitude, for making it possible for me to start and accomplish this work in peace. I also extend my heartfelt appreciation to my supervisor, Prof O. A. Olafuyi, for his support and mentorship. His expertise, insightful feedback, and dedication to my intellectual growth have been instrumental in shaping this project.

I am truly fortunate to have had the opportunity to work under his guidance. I would also like to acknowledge the head of the department and all staff in the Department of Petroleum Engineering, whose contributions greatly enriched the knowledge base for fostering an environment of academic excellence.

I also appreciate Mr O. Izuchukwu for his support and encouragement while working with him.

I acknowledge and appreciate Dr. Ikomwosa Ohenhen for his assistance and support

Furthermore, I would like to express my heartfelt gratitude to my family and loved ones who have been a constant source of support both emotionally and financially, as well as their constant prayers and encouragement throughout this academic journey. In particular, I want to acknowledge my mother, Mrs. Rita Iyangbe, for her unwavering belief in me and her boundless love that has been my pillar of strength.

Lastly, I would like to acknowledge all my friends and classmates within the Department of Petroleum Engineering

ABSTRACT

Traditional reservoir forecasting methods provide single deterministic predictions that fail to capture subsurface uncertainty. This research develops a Python-based simulation framework integrating material balance physics with Monte Carlo uncertainty quantification for probabilistic reservoir performance forecasting.

The framework generates P10/P50/P90 forecasts through efficient tank modeling, completing 5,000-realization uncertainty analysis in under one hour—a 50-100x speedup versus commercial simulators. Application to a representative sandstone reservoir yields 18.5 MMstb cumulative recovery (P50) with $\pm 24\%$ uncertainty (P90: 14.2 MMstb, P10: 23.1 MMstb). Sensitivity analysis identifies permeability as the dominant uncertainty driver (correlation +0.68), directly informing data acquisition priorities.

Validation confirms material balance closure within 0.5% and recovery factors consistent with published analogs. By eliminating software licensing costs and computational barriers, the framework democratizes probabilistic forecasting, transforming reservoir engineering from deterministic prediction to accessible, risk-aware decision support.

Keywords: Reservoir forecasting, uncertainty quantification, Monte Carlo simulation, Python, probabilistic reserves.

TABLE OF CONTENT

TITLE PAGE	i
CERTIFICATION	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
TABLE OF CONTENT	vi
LIST OF TABLES	xiii
LIST OF FIGURES	xiv
CHAPTER ONE	1
INTRODUCTION	1
1.1 BACKGROUND OF STUDY	1
1.2 STATEMENT OF THE PROBLEM	2
1.3 AIM AND OBJECTIVE	3
1.4 SIGNIFICANCE OF THE STUDY	4
1.5 SCOPE OF STUDY	5
1.6 LIMITATIONS	6
1.7 JUSTIFICATION OF THE STUDY	7
CHAPTER TWO	9
LITERATURE REVIEW	9
2.1 INTRODUCTION	9

2.2 DATA SOURCES AND CHARACTERISTICS FOR RESERVOIR FORECASTING	13
2.3 MACHINE LEARNING APPROACHES IN RESERVOIR ENGINEERING	16
2.3.1 Introduction to Machine Learning in Reservoir Forecasting	16
2.3.2 Traditional Machine Learning Algorithms	17
2.3.3 Artificial Neural Networks	18
2.3.4 Deep Learning Models	18
2.3.5 Hybrid Approaches	18
2.3.6 Case Studies	19
2.3.7 Challenges and Limitations	19
2.3.8 Emerging Trends	19
2.4 UNCERTAINTY QUANTIFICATION (UQ) TECHNIQUES IN RESERVOIR FORECASTING	20
2.4.1 Introduction to Uncertainty in Reservoir Forecasting	20
2.4.2 Types of Uncertainty	21
2.4.3 Traditional UQ techniques	22
2.4.4 Advanced Probabilistic Methods	23
2.4.5 Computational Efficiency in Uncertainty Quantification	25
2.4.6 Summary Table	26
2.5 VALIDATION, METRICS AND BACK-TESTING STRATEGIES	27
2.5.1 Introduction to Model Validation	27
2.5.2 Types of Validation Strategies	28

2.5.3 Back-Testing Strategies	29
2.5.4 Key Performance Metrics (Technical Overview)	29
2.5.5 Reliability, Robustness and Generalization Checks	30
2.5.6 Model Calibration vs Validation	31
2.5.7 Model Validation and Back-Testing Flow	31
2.6 CASE STUDY AND APPLICATIONS	33
2.6.1 Simulation-Based Forecasting with Uncertainty Analysis: Industry Applications	33
2.6.2 Simulation vs. Empirical Forecasting Methods	34
2.6.3 Illustrative Example: Uncertainty Analysis in a Sandstone Reservoir	34
2.7 CHALLENGES AND FUTURE DIRECTIONS IN SIMULATION-BASED UNCERTAINTY QUANTIFICATION	35
2.8 SYNTHESIS AND POSITIONING OF CURRENT STUDY	37
CHAPTER THREE	40
METHODOLOGY	40
3.1 RESEARCH DESIGN	40
3.2 DATA REQUIREMENTS AND SOURCES	41
3.3 DATA PRE-PROCESSING AND QUALITY CHECKS	43
Table 3.1: Unit Standardization for Reservoir Properties	45
3.4 MODEL SELECTION/RESERVOIR SIMULATION APPROACH	46
3.5 SIMULATION SETUP	52
3.5.1 Reservoir Representation and Grid Setup	52
3.5.2 Rock and Fluid Property Specification	53

3.5.3 Initialization and Boundary Conditions	53
3.5.4 Production Control/Control Settings	54
3.5.5 History Matching Approach	54
3.6 UNCERTAINTY IDENTIFICATION AND PARAMETER SELECTION	55
3.6.1 Geological and Petrophysical Properties	55
3.6.2 Fluid Property and PVT Uncertainty	56
3.6.3 Operational and Engineering Uncertainties	57
3.6.4 Sensitivity Ranges for Key Parameters	57
3.7 UNCERTAINTY QUANTIFICATION (UQ) METHODS	59
3.7.1 Conceptual Overview	59
3.7.2 Uncertainty Quantification Strategy	59
3.7.3 Implementation Procedure	61
3.7.4 Mathematical Representation	63
3.7.6 Python Implementation Tools	64
3.8 MODEL EVALUATION AND VALIDATION	65
3.8.1 Purpose of Model Evaluation	66
3.8.2 Validation Techniques	66
3.8.3 Model Comparison and Ranking	69
3.8.4 Visual Model Validation	70
3.8.5 Validation Deliverables	70
3.9 RISK ASSESSMENT/DECISION CRITERIA	71

3.9.1 Purpose of Risk Assessment	71
3.9.2 Probabilistic Output Analysis	72
3.9.3 Economic Risk Indicators	73
3.9.4 Decision Criteria	74
3.9.5 Visualization and Reporting	74
3.9.6 Deliverables	75
3.10 SUMMARY OF METHODOLOGICAL WORKFLOW	75
CHAPTER FOUR	78
RESULTS AND DISCUSSION	78
4.1 INTRODUCTION	78
4.2 BASE CASE RESERVOIR DESCRIPTION AND DETERMINISTIC FORECAST	79
4.2.1 Reservoir Properties and Initial Conditions	79
4.2.2 Deterministic Production Forecast	80
4.2.3 Drive Mechanism and Depletion Characteristics	82
4.3 UNCERTAINTY PROPAGATION: PROBABILISTIC FORECASTING RESULTS	83
4.3.1 Parameter Uncertainty Ranges	84
4.3.2 Probabilistic Production Forecast	85
4.3.3 Recovery Factor Distribution	87
4.4 SENSITIVITY ANALYSIS: IDENTIFYING KEY UNCERTAINTY DRIVERS	89
4.4.1 Spearman Rank Correlation Analysis	89
4.4.2 Tornado Diagram Visualisation	91

4.4.3 Implications for Data Acquisition Strategy	92
4.5 MODEL VALIDATION AND CONSISTENCY CHECKS	93
4.5.1 Material Balance Closure	93
4.5.2 Pressure Decline Realism	95
4.5.3 Production Decline Consistency with Field Analogs	96
4.5.4 Sensitivity to Numerical Parameters	97
4.6 RISK ASSESSMENT AND DECISION SUPPORT	98
4.6.1 Probabilistic Reserve Classification	98
4.6.2 Economic Scenario Analysis	99
4.6.3 Value of Information Analysis	101
4.6.4 Decision Recommendation	101
4.7 COMPARISON WITH ALTERNATIVE FORECASTING METHODS	103
4.7.1 Decline Curve Analysis (DCA) Comparison	104
4.7.2 Computational Efficiency Comparison	105
4.7.3 Interpretability and Transparency	107
4.8 LIMITATIONS AND SOURCES OF UNCERTAINTY NOT CAPTURED	108
4.8.1 Structural Model Uncertainty	108
4.8.2 Operational Uncertainties Beyond Modeled Scenarios	109
4.8.3 Geomechanical and Non-Darcy Flow Effects	109
4.8.4 Statistical Assumptions in Uncertainty Quantification	110
4.8.5 Data Quality and Representatives	111

4.9 CHAPTER SUMMARY	112
CHAPTER FIVE	114
CONCLUSION AND RECOMMENDATION	114
5.1 CONCLUSION	114
5.2 RECOMMENDATION	115
5.3 CONTRIBUTION TO KNOWLEDGE	117
REFERENCES	119

LIST OF TABLES

Table 2.2: Comparison of Uncertainty Quantification Methods	26
Table 2.5: Model Calibration vs Validation	31
Table 3.2: Rock and Fluid Properties Used in Base Model	53
Table 3.3: Geological and Petrophysical Uncertainties	55
Table 3.4: Fluid Property and PVT Uncertainties	56
Table 3.5: Operational and Engineering Uncertainties	57
Table 3.6: Sensitivity Ranges for Key Parameters	57
Table 3.7: Model Performance Comparison Metrics	69
Table 3.8: Probabilistic Forecast Results Example	72
Table 4.1: Base Case Reservoir Parameters	79
Table 4.2: Uncertain Parameter Distributions for Monte Carlo Analysis	84
Table 4.3: Sensitivity Ranking - Spearman Correlation with 20-Year Cumulative Production	89
Table 4.4: Validation Against Industry Analog Data	96
Table 4.5: Numerical Stability Tests	97
Table 4.6: Probabilistic Reserves Estimates	98
Table 4.7: Economic Risk Assessment	100
Table 4.8: Comparison of Forecasting Methods	104

LIST OF FIGURES

Figure 3.1: Simulation Workflow Diagram	51
Figure 3.2 Conceptual Simulation Setup Structure	55
Figure 3.3: Uncertainty Parameter Identification and Ranges	58
Figure 3.5: Cross-Validation Visualization	70
Figure 3.6: Integrated Workflow for Reservoir Uncertainty and Risk Analysis	77
Figure 4.1A: Oil Production Rate (Base Case Forecast)	80
Figure 4.1B: Cumulative Oil Production	81
Figure 4.1: Base Case Deterministic Production Forecast	81
Figure 4.2: Pressure and GOR Evolution - Base Case	83
Figure 4.4 (A): Cumulative Distribution Function - Ultimate Recovery Factor	87
Figure 4.4 (B): Cumulative Distribution Function - Ultimate Recovery Factor	88
Figure 4.4: Cumulative Distribution Function - Ultimate Recovery Factor	88
Figure 4.5: Tornado Diagram - Impact on 20-Year Cumulative Production	91
Figure 4.6: Material Balance Closure Error Distribution	94
Figure 4.7: Pressure Decline Comparison - Simulation vs Analytical	95

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF STUDY

Reservoir performance forecasting is one of the most important aspects of petroleum engineering because it provides a basis for estimating how much oil, gas, or water a reservoir will produce in the future. These forecasts guide decisions on field development, production planning, and investment strategies. For decades, engineers have relied on methods such as decline curve analysis, material balance techniques, and numerical reservoir simulation to predict future production trends (Arps, 1945; Aziz & Settari, 1979). While these approaches have been widely used, they are often limited by assumptions of reservoir homogeneity and simplified physics that may not hold in complex, heterogeneous formations.

The petroleum industry today is confronted with reservoirs that are increasingly more complex—tight formations, unconventional plays, and mature fields with uncertain geologic descriptions. In such settings, deterministic models that give a single forecast scenario may fail to capture the full range of possible outcomes. This becomes even more critical when considering the economic risks associated with drilling, development, and production. Uncertainty in reservoir properties such as porosity, permeability, and fluid saturations can translate into significant uncertainty in production forecasts. As a result, decision-makers often seek methods that can capture variability and provide probabilistic outcomes instead of a single “best guess” (Caers, 2005).

Over the last decade, machine learning (ML) has become a promising tool to complement or even substitute traditional approaches to reservoir forecasting. Unlike physics-based methods, ML models do not require an explicit mathematical description of the reservoir. Instead, they

learn patterns directly from historical data—production rates, well logs, seismic attributes, and reservoir characteristics—and use these patterns to predict future performance. This ability to capture complex, nonlinear relationships in large datasets makes ML particularly well-suited for petroleum applications where data is abundant but uncertainty is high (Wang & Mohaghegh, 2020). However, the application of machine learning requires extensive historical datasets and often lacks the physical interpretability needed for regulatory approval and engineering validation. For reservoirs with limited production history or where physical understanding is critical, simulation-based approaches integrated with uncertainty quantification provide a balanced framework that combines computational efficiency with reservoir physics (Aziz & Settari, 1979; Caers, 2005).

In parallel, uncertainty analysis has grown in importance as part of reservoir management. Forecasts without a measure of uncertainty can be misleading, as they provide only a single trajectory while ignoring the inherent variability of subsurface systems. Methods such as Monte Carlo simulation, bootstrapping, and probabilistic modeling help quantify uncertainty by generating a range of possible scenarios and estimating the likelihood of different outcomes. Integrating uncertainty quantification with simulation-based forecasting therefore offers a modern and powerful approach: simulation ensures forecasts are physically realistic and interpretable, while uncertainty quantification ensures they capture the range of plausible outcomes necessary for reliable decision-making (Al-Fatlawi et al., 2022).

1.2 STATEMENT OF THE PROBLEM

Despite the wide use of traditional reservoir forecasting techniques such as decline curve analysis and numerical simulation, significant challenges remain in achieving accurate and reliable predictions. Reservoirs are inherently heterogeneous, and the available data is often incomplete, uncertain, or noisy. This makes deterministic forecasts—which provide only a

single predicted outcome—unreliable for decision-making in complex field environments (Caers, 2005).

In addition, conventional simulators can be computationally expensive and require detailed inputs that are not always available, particularly in data-limited reservoirs. As a result, operators face difficulties in quantifying risks and planning long-term strategies. There is therefore a pressing need for forecasting methods that are data-driven, computationally efficient, and capable of explicitly accounting for uncertainty in order to improve decision-making and reduce financial risk (Wang & Mohaghegh, 2020).

1.3 AIM AND OBJECTIVE

The primary aim of this study is to develop a simulation-based computational framework for forecasting reservoir performance with integrated uncertainty analysis using Python. This framework is intended to provide faster, more reliable, and more informative probabilistic forecasts compared to traditional deterministic methods.

To achieve this aim, the specific objectives are to:

1. Collect and specify reservoir properties, fluid characteristics, and production parameters suitable for simulation-based forecasting.
2. Develop a Python-based reservoir simulation model that captures material balance principles and production decline behavior under varying operational constraints.
3. Incorporate uncertainty analysis methods, including Monte Carlo simulation and sensitivity analysis, to quantify the impact of uncertain reservoir parameters on production forecasts and generate probabilistic outcomes (P10/P50/P90).
4. Validate the simulation framework through material balance consistency checks and compare probabilistic forecasts with deterministic prediction methods.

5. Evaluate forecast reliability using engineering validation criteria and quantify uncertainty ranges through probabilistic performance metrics (P10/P50/P90 coverage, sensitivity rankings).

1.4 SIGNIFICANCE OF THE STUDY

This study is significant because it provides a modern, accessible approach to one of the core challenges in petroleum engineering—forecasting reservoir performance with explicit uncertainty quantification. Traditional methods such as decline curve analysis and numerical simulation, though valuable, often struggle with the complexity and uncertainty of real reservoirs. Decline curve analysis is limited by its empirical nature and assumptions about flow regimes, while commercial reservoir simulators can be computationally expensive and require proprietary software licenses (Arps, 1945; Aziz & Settari, 1979).

By developing a Python-based simulation framework, this study offers a transparent, flexible, and cost-effective alternative that maintains physical realism while enabling rapid uncertainty analysis. The framework can capture material balance relationships, production decline trends, and drive mechanisms without the overhead of commercial software, making it particularly suitable for screening studies, sensitivity analyses, and educational applications (Caers, 2005).

Equally important is the explicit integration of probabilistic uncertainty analysis. In practice, decision-makers need more than a single deterministic prediction; they require insight into the range of possible outcomes and the risks associated with each. This research addresses that gap by combining simulation-based modeling with probabilistic methods to generate forecasts that are not only accurate but also uncertainty-aware. Such tools can improve decision-making, reduce financial risks, and provide petroleum engineers with greater confidence in planning and operations (Caers, 2005).

Finally, the choice of Python as the implementation platform ensures that the framework is both accessible and practically deployable. Python's extensive scientific libraries (NumPy, SciPy, Pandas) and visualization tools (Matplotlib, Seaborn) enable efficient numerical computation and clear presentation of uncertainty results. Unlike proprietary simulation software, this open-source approach allows for full transparency in model assumptions, customization for specific reservoir types, and integration with modern data analytics workflows. The framework can therefore serve as both a practical engineering tool and an educational resource for demonstrating reservoir forecasting principles under uncertainty (Al-Fatlawi et al., 2022; Caers, 2005).

1.5 SCOPE OF STUDY

The scope of this study is limited to forecasting reservoir performance using simulation-based methods integrated with uncertainty quantification. The focus is on developing a Python-based computational framework that implements material balance principles, production decline modeling, and probabilistic forecasting. Traditional methods such as decline curve analysis and analytical solutions are referenced for validation and comparison purposes.

The simulation framework represents the reservoir as a tank model characterized by average properties (porosity, permeability, fluid PVT characteristics), which is appropriate for field-scale performance forecasting and uncertainty screening studies. While full 3D gridded reservoir simulation offers higher spatial resolution, the tank model approach provides a practical balance between physical realism and computational efficiency, particularly when the goal is to explore a wide range of uncertain parameter combinations (Dake, 1978; Ahmed, 2019). The uncertainty analysis incorporates Monte Carlo simulation and sensitivity analysis methods to explore how variations in reservoir properties (permeability, porosity, aquifer

strength) and operational parameters (well constraints, production limits) affect production forecasts.

The study explicitly integrates uncertainty analysis into the forecasting workflow. Methods such as Monte Carlo simulation, Latin Hypercube Sampling, and tornado diagram analysis will be employed to generate probabilistic outcomes and quantify forecast variability. However, the analysis will be constrained by the availability and quality of data, meaning the models will be trained and validated on the datasets provided. The study does not aim to develop new simulation algorithms or numerical methods but rather to implement and validate existing reservoir engineering principles within a Python-based probabilistic forecasting framework.

Finally, the implementation will be carried out in Python, leveraging its open-source libraries and frameworks for numerical simulation, statistical analysis, and visualization. The scope does not extend to building large-scale commercial software but instead focuses on developing a practical and adaptable forecasting framework that can be reproduced and refined in future research or field applications.

1.6 LIMITATIONS

Like any reservoir modeling study, this research is subject to certain limitations. The accuracy of the simulation-based forecasts depends on the quality and representativeness of the input parameters (rock properties, fluid PVT data, and production constraints). Since the study uses representative reservoir parameters based on industry literature and analogous field data rather than a specific proprietary dataset, the results demonstrate the methodology's capability but may not directly match any particular field's behavior. However, the framework is designed to be adaptable and can readily incorporate field-specific data when available (Caers, 2005).

Additionally, while the uncertainty analysis provides probabilistic forecasts (P10/P50/P90), it cannot account for unknown reservoir features that have not been explicitly parameterized in the model. For instance, unexpected geological discontinuities, undetected compartmentalization, or future operational changes beyond the assumed scenarios would introduce uncertainties not captured in the current framework (Oliver & Chen, 2011).

1.7 JUSTIFICATION OF THE STUDY

The justification for this study stems from the growing need for more reliable and efficient forecasting tools in petroleum engineering. Traditional reservoir forecasting techniques, while valuable, often fall short in addressing the increasing complexity and uncertainty of modern reservoirs. Simulation-based approaches integrated with Python offer a practical way to address these challenges by maintaining physical realism while enabling rapid probabilistic forecasting without the computational burden and licensing costs of commercial reservoir simulators (Aziz & Settari, 1979; Caers, 2005).

Integrating uncertainty analysis further strengthens the case for this research. In an industry where multi-million-dollar decisions hinge on production forecasts, relying solely on single, deterministic predictions can be risky. Probabilistic forecasts provide decision-makers with a clearer understanding of the range of potential outcomes, making it possible to manage risks more effectively and plan under uncertainty (Caers, 2005).

This study is therefore justified not only by its technical contribution---demonstrating the use of simulation-based methods with integrated uncertainty quantification in reservoir forecasting---but also by its practical relevance. By developing the framework in Python, a widely accessible and industry-adopted programming language, the research ensures that its

methods can be easily adapted and scaled by petroleum engineers, researchers, and industry practitioners alike (Al-Fatlawi et al., 2022).

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

Forecasting reservoir performance has always been one of the most fundamental responsibilities of petroleum engineers. Whether it is estimating how much oil and gas can ultimately be recovered, projecting future production rates, or planning facilities and investment strategies, reliable forecasts serve as the foundation for decision-making in the oil and gas industry. Over the decades, various techniques have been developed for this purpose, ranging from simple empirical models to complex numerical simulations. Each of these approaches has played an important role in the evolution of reservoir engineering, but they have also faced significant challenges—especially when applied to reservoirs characterized by heterogeneity, uncertainty, and incomplete data.

One of the earliest and most widely applied methods is decline curve analysis (DCA), introduced by J. J. Arps in 1945. This empirical technique uses production history data to identify decline trends, which are then extrapolated into the future (Arps, 1945). The simplicity of DCA made it extremely attractive in the mid-20th century, particularly at a time when computational resources were limited and engineers needed quick, practical solutions. For conventional reservoirs exhibiting boundary-dominated flow, DCA provided reasonable forecasts that were sufficient for field planning. However, as more complex reservoirs were developed—tight gas sands, shales, and heterogeneous carbonates—engineers began to realize the limitations of DCA. Its reliance on fixed exponential, hyperbolic, or harmonic decline assumptions often led to inaccurate predictions in settings where flow regimes deviated from the classical behavior (Ilk et al., 2008).

To overcome these shortcomings, engineers turned to material balance methods. Rooted in the principle of conservation of mass, material balance relates cumulative production to changes in average reservoir pressure. This method is more physically grounded than DCA, offering insights into drive mechanisms such as water drive, gas cap expansion, or solution gas drive. The strength of material balance lies in its ability to provide estimates of hydrocarbons initially in place (HCPV) without requiring detailed spatial reservoir models (Craft & Hawkins, 1959). However, this approach also has practical limitations. It depends on accurate reservoir pressure measurements—data that are not only expensive to obtain but also may not represent the average conditions in heterogeneous systems. Additionally, it assumes uniform reservoir properties, which is rarely the case in naturally fractured or stratified formations (Aziz & Settari, 1979).

As computing power advanced in the 1960s and 1970s, petroleum engineers embraced numerical reservoir simulation as the most comprehensive tool for performance forecasting. Numerical simulators, such as ECLIPSE and CMG, discretize the reservoir into thousands or even millions of grid blocks and solve governing multiphase flow equations using finite-difference methods (Aziz & Settari, 1979). This enables them to capture the complex interplay between rock and fluid properties, well controls, and geological features. For example, in giant offshore fields like Ghawar in Saudi Arabia and Prudhoe Bay in Alaska, simulation models have been critical in testing development scenarios, designing waterfloods, and evaluating enhanced oil recovery (EOR) schemes (Lake, 1989). Despite their sophistication, however, simulators are far from perfect. They are extremely data-hungry, requiring detailed input on porosity, permeability, relative permeability curves, and fluid properties. In many reservoirs, these inputs are uncertain or unavailable. Furthermore, full-field simulations can be computationally expensive, often taking hours or days to run a single forecast scenario. This

makes it impractical to perform extensive uncertainty analyses or rapid forecasting under tight deadlines (Oliver & Chen, 2011).

Another pressing challenge across all these forecasting approaches is the problem of uncertainty. Subsurface reservoirs are inherently heterogeneous and only partially observed. Core samples and well logs provide valuable data, but they represent a tiny fraction of the reservoir volume. Seismic data gives broader coverage but with lower resolution. As Caers (2005) noted, these uncertainties propagate into reservoir models, making forecasts less reliable. Deterministic models that produce a single "best estimate" often give decision-makers a false sense of confidence, ignoring the wide range of possible outcomes that could result from unknown geological or operational variables. This has led to increased emphasis on probabilistic approaches, such as Monte Carlo simulation, stochastic reservoir modeling, and Bayesian inference, which aim to quantify the uncertainty in predictions and present decision-makers with probability distributions instead of single numbers (Deutsch, 2002).

In parallel with these developments, advances in computational methods and data analytics have begun to complement traditional reservoir forecasting approaches. Machine learning techniques, for instance, have shown potential in specific applications such as pattern recognition in well performance and rapid proxy modeling. However, these methods typically require extensive historical datasets and often lack the physical interpretability needed for regulatory workflows. Instead of explicitly simulating the physics of fluid flow, ML learns patterns directly from data—whether from production histories, well logs, or seismic attributes. For example, artificial neural networks (ANNs) have been applied to predict production decline in unconventional reservoirs, while support vector machines (SVMs) and random forests have been used to identify patterns in well performance (Wang & Mohaghegh, 2020). Unlike DCA or material balance, ML does not require strict assumptions about flow regimes, and unlike

simulation, it can deliver forecasts rapidly once trained. The drawback, however, is that ML models require large, high-quality datasets and can sometimes act as "black boxes," making it difficult to interpret their predictions. An alternative path that has gained traction involves enhancing traditional simulation approaches with robust uncertainty quantification, combining the physical realism of material balance and flow equations with probabilistic methods that explicitly capture parameter uncertainty. This hybrid approach maintains engineering interpretability while providing the risk-informed forecasts that modern decision-making requires (Caers, 2005; Oliver & Chen, 2011).

The integration of uncertainty quantification with machine learning is an especially promising direction. By combining the predictive power of ML with probabilistic frameworks such as Bayesian neural networks or ensemble methods, researchers can provide forecasts that not only predict expected outcomes but also quantify the uncertainty surrounding those predictions (Al-Fatlawi et al., 2022). This is particularly important in modern petroleum engineering, where investment decisions involve billions of dollars, and the ability to weigh risks against potential rewards is critical.

In summary, the evolution of reservoir forecasting reflects a progressive refinement of methods, from simple empirical curves to physics-based numerical simulation, and now toward probabilistic frameworks that explicitly quantify uncertainty. Each advancement has responded to growing reservoir complexity and the need for more reliable risk assessment. The current emphasis on simulation-based uncertainty quantification recognizes that deterministic forecasts, while valuable, cannot adequately support high-stakes decisions without accompanying measures of confidence and variability. Modern forecasting frameworks must therefore balance physical realism, computational practicality, and transparent uncertainty representation—principles that form the foundation of this study.

2.2 DATA SOURCES AND CHARACTERISTICS FOR RESERVOIR FORECASTING

Data is at the core of any reservoir forecasting study, whether using physics-based simulation or data-driven approaches. For simulation-based models, data quality directly affects parameter initialization, validation, and uncertainty characterization. While machine learning models learn patterns directly from data, simulation-based approaches use data to constrain physical parameters and validate model behavior provided to them. This means that the quantity, quality, and diversity of available data determine not only the accuracy of predictions but also the reliability of the models when applied in real-world decision-making (Wang & Mohaghegh, 2020). Reservoir forecasting, therefore, depends on a broad spectrum of data sources—ranging from routine field measurements to highly specialized laboratory analyses.

- **Production Data:** Production data is one of the most important and widely used inputs in reservoir forecasting. This includes historical measurements of oil, gas, and water production rates, as well as bottom-hole pressures and flowing wellhead pressures. Because production data directly reflects reservoir behavior under changing conditions, it provides critical signals for model training. Decline curve analysis, for instance, relies entirely on production history, while material balance methods use cumulative production and pressure trends to infer hydrocarbons in place. In forecasting applications, production data often serves as both input and output: past production rates can be used as features, while future rates are the prediction targets. Studies have shown that incorporating time-series production data into recurrent neural networks (RNNs) or long short-term memory networks (LSTMs) significantly improves forecast accuracy in unconventional reservoirs (Fang et al., 2019). However, production data is not without challenges. It can be noisy, incomplete, and affected by operational constraints such as choke adjustments, artificial lift installations, or workovers.

Filtering out these effects to capture the underlying reservoir performance is critical in ensuring the data's usefulness for ML models (Narsilio et al., 2017).

- **Well Logs:** Well log data is another fundamental source of information, providing insights into subsurface lithology and rock properties. Logs such as gamma ray, resistivity, density, neutron porosity, and sonic travel time are routinely recorded in most wells. They are often combined to estimate porosity, permeability, water saturation, and lithofacies distributions—parameters that strongly influence reservoir performance. Machine learning methods have been widely applied to interpret well logs, such as using supervised classification algorithms to identify lithofacies or regression models to predict permeability from log responses (Bhattacharya & Carr, 2016). For forecasting purposes, log-derived properties are valuable features that can help ML models link geological characteristics with observed production trends. For example, reservoirs with high-porosity sands are expected to sustain higher production, while intervals with shales may contribute little. The main challenge with log data lies in heterogeneity: logs are available only in drilled wells, and their resolution can vary across different logging runs. Integrating log data across fields with diverse acquisition standards requires careful preprocessing and normalization.
- **Seismic Data:** Seismic surveys provide large-scale information about reservoir structure and stratigraphy. Unlike well logs, which are localized, seismic data covers wide areas and helps map faults, horizons, and depositional features. Attributes derived from seismic, such as amplitude, impedance, or coherence, can be linked with porosity and fluid content, offering indirect indicators of reservoir quality. In forecasting applications, seismic data is typically used to enhance spatial understanding of reservoir heterogeneity. Techniques such as convolutional neural networks (CNNs) have been applied to seismic interpretation, automatically identifying

faults, channels, or lithological facies (Di et al., 2018). In forecasting, seismic-derived attributes can be integrated with well-level data to create richer feature sets for ML models. However, seismic data introduces its own uncertainties due to noise, resolution limitations, and interpretation biases. Additionally, seismic-to-well ties often require calibration, and misalignment between datasets can reduce ML accuracy if not handled carefully (Zhang & Curtis, 2020).

- **Core and PVT Data:** Core samples and laboratory fluid analyses provide ground truth information about rock and fluid properties. Core analysis offers direct measurements of porosity, permeability, relative permeability, and capillary pressure, while PVT (pressure–volume–temperature) analysis characterizes the behavior of reservoir fluids under varying pressure and temperature conditions. These datasets, although sparse, are critical for understanding reservoir flow dynamics. In forecasting applications, core and PVT data are often used to calibrate or validate models rather than as direct training inputs. Because cores are expensive and collected only in select wells, they represent a limited portion of the reservoir. Nonetheless, when integrated with log and production data, they can significantly improve the robustness of ML predictions. For example, permeability values derived from cores can be used to label training datasets in supervised learning models designed to predict permeability from well logs (Amaefule et al., 1993).
- **Operation and Field Data:** Reservoir performance is influenced not only by subsurface properties but also by operational decisions. Data related to well design (e.g., horizontal vs. vertical wells, completion strategies, hydraulic fracturing stages), surface constraints (pipeline capacity, separator conditions), and operational histories (shut-ins, artificial lift installations, or workovers) all affect production trends. Machine learning models that exclude these operational variables may incorrectly attribute

production changes to reservoir properties alone, leading to biased forecasts. For instance, a sudden drop in production might be due to scaling or equipment malfunction rather than reservoir depletion. Including operational metadata in training datasets helps ML distinguish between subsurface-driven and surface-driven behaviors (Jahanbakhshi & Mohaghegh, 2019).

- **Data Quality, Preprocessing and Integration Challenges:** While the diversity of data sources provides rich opportunities for forecasting, it also creates significant challenges in terms of data integration and preprocessing. Reservoir datasets are often inconsistent, with missing values, different scales, and varying levels of uncertainty. Machine learning requires datasets to be cleaned, normalized, and structured into formats suitable for training. Common preprocessing steps include outlier removal, interpolation of missing data, feature scaling, and dimensionality reduction through techniques like principal component analysis (PCA). Another challenge lies in temporal and spatial alignment. Production data is continuous and time-dependent, while log and seismic data are static snapshots. Merging these datasets into coherent ML frameworks requires careful feature engineering. Moreover, the volume of data can be overwhelming: modern seismic surveys generate terabytes of information, while production data can span decades. Efficient handling, storage, and selection of relevant features are therefore crucial for building reliable models (Goodfellow et al., 2016).

2.3 MACHINE LEARNING APPROACHES IN RESERVOIR ENGINEERING

2.3.1 Introduction to Machine Learning in Reservoir Forecasting

Over the past two decades, machine learning (ML) has emerged as a complementary tool in reservoir engineering, particularly for tasks involving pattern recognition, data interpolation, and rapid proxy modeling. ML methods excel in situations where large historical datasets exist

and where the relationship between inputs and outputs is complex but not necessarily governed by well-understood physics. Applications have included permeability prediction from well logs, lithofacies classification, and production decline forecasting in unconventional reservoirs with thousands of wells (Wang & Mohaghegh, 2020; Bhattacharya & Carr, 2016).

However, ML approaches face significant practical constraints in many reservoir settings. They require substantial training data, which may not be available in exploration or early appraisal stages. They often function as "black boxes," making it difficult to validate results against physical principles or explain predictions to regulatory bodies. Furthermore, ML models trained on one field's data may not generalize well to reservoirs with different geological characteristics. For these reasons, physics-based simulation approaches remain the industry standard when physical understanding, regulatory compliance, and transparency are priorities—particularly when integrated with modern uncertainty quantification methods (Caers, 2005; Oliver & Chen, 2011).

2.3.2 Traditional Machine Learning Algorithms

Early applications of ML in petroleum engineering employed relatively simple algorithms such as linear and polynomial regression for permeability estimation, and decision trees for lithofacies classification. More sophisticated methods like Random Forests and Support Vector Machines (SVMs) later gained adoption, particularly for handling high-dimensional log data and identifying nonlinear relationships (Breiman, 2001; Bhattacharya & Carr, 2016). While these methods demonstrated some success in specific tasks such as log interpretation and property prediction, their application to full-field production forecasting remained limited due to data scarcity and the need for extensive feature engineering.

2.3.3 Artificial Neural Networks

Artificial Neural Networks (ANNs), inspired by biological neural processing, were among the first ML techniques applied to petroleum problems. Studies in the early 2000s demonstrated that ANNs could approximate complex input-output relationships in reservoir systems, such as predicting water cut evolution or estimating permeability distributions (Mohaghegh, 2000). However, ANNs suffered from challenges including sensitivity to hyperparameter selection, difficulties in training stability, and limited interpretability—issues that made them less attractive than physics-based models for high-stakes forecasting applications where transparency and validation are essential.

2.3.4 Deep Learning Models

The recent surge in computational power and big data has enabled the use of deep learning architectures, including Convolutional Neural Networks (CNNs) for seismic interpretation and Long Short-Term Memory networks (LSTMs) for time-series production forecasting (Di et al., 2018; Fang et al., 2019). These methods have shown promising results in data-rich unconventional plays such as U.S. shale basins, where thousands of wells provide sufficient training examples. Nevertheless, deep learning models are computationally expensive to train, require careful tuning to avoid overfitting, and offer limited insight into the underlying physics. For conventional reservoirs with sparse well data or for studies requiring regulatory transparency, simulation-based approaches with explicit uncertainty quantification often provide a more practical and defensible solution.

2.3.5 Hybrid Approaches

One practical implementation involves using simulation models to generate synthetic training datasets, then fitting surrogate models (polynomial response surfaces or simplified regressors) to enable rapid Monte Carlo sampling during uncertainty analysis. This allows engineers to

explore thousands of parameter combinations without rerunning the full simulator repeatedly, significantly reducing computational cost while maintaining physical consistency (Queipo et al., 2005; Sarma et al., 2008).

2.3.6 Case Studies

Documented applications of ML in reservoir forecasting have primarily focused on unconventional plays with large well populations. For instance, LSTM networks applied to Eagle Ford shale production data demonstrated improved forecasting accuracy compared to empirical decline curves when sufficient training data was available (Fang et al., 2019). However, such studies often acknowledge limitations including model interpretability and generalization to new field conditions. In contrast, simulation-based uncertainty workflows have been successfully deployed across diverse reservoir types—from Middle Eastern carbonates to North Sea offshore fields—precisely because they maintain physical consistency and offer transparent validation paths (Al-Fatlawi et al., 2022; Jensen et al., 2022).

2.3.7 Challenges and Limitations

These practical constraints reinforce the continued relevance of simulation-based forecasting, particularly when coupled with modern uncertainty quantification techniques that provide probabilistic insights without sacrificing physical interpretability or requiring extensive training datasets.

2.3.8 Emerging Trends

Recent research directions include explainable AI methods aimed at improving ML interpretability, and physics-informed neural networks that embed conservation laws into learning architectures (Raissi et al., 2019; Lundberg & Lee, 2017). While conceptually promising, these approaches remain largely in the research phase and have not yet achieved widespread industry adoption. For practical reservoir management, the integration of robust

uncertainty quantification with established simulation methods continues to offer the most reliable path forward, combining computational efficiency with physical realism and transparent validation. These emerging trends suggest that the future of reservoir forecasting will not be purely data-driven or purely physics-based, but rather an integrated approach where ML, uncertainty quantification, and physics work together to deliver reliable, practical tools.

2.4 UNCERTAINTY QUANTIFICATION (UQ) TECHNIQUES IN RESERVOIR FORECASTING

2.4.1 Introduction to Uncertainty in Reservoir Forecasting

Uncertainty in reservoir forecasting is unavoidable because subsurface systems are only partially observed. Engineers do not have complete visibility into pore-scale heterogeneity, fault behavior, fluid movement, or long-term reservoir dynamics. As a result, any single forecast—no matter how sophisticated—represents only *one possible outcome*, not the full range of what might happen in reality (Caers, 2005). Traditional forecasting methods such as decline curve analysis or numerical simulation are typically deterministic, meaning they output a single production trajectory. In practice, however, production performance is influenced by multiple uncertain variables such as permeability variation, saturation distribution, and operational changes. This is why modern reservoir studies increasingly use probabilistic rather than deterministic forecasts (Tavakoli et al., 2013).

The petroleum industry's emphasis on uncertainty quantification has grown significantly since the 1990s, driven by the high capital costs of field development and the irreversible nature of many investment decisions. A probabilistic forecast that presents a range of outcomes—typically characterized by P10, P50, and P90 percentiles—provides decision-makers with a clearer picture of both upside potential and downside risk. This approach aligns with modern

portfolio management strategies where projects are evaluated not just on expected value but on the full distribution of possible outcomes (Murtha, 1997; Rose, 2001).

Uncertainty in reservoir forecasting propagates from multiple sources and compounds through the modeling workflow. Small uncertainties in individual parameters (such as porosity or permeability) can combine nonlinearly to produce substantial variability in production forecasts, especially over long time horizons. Understanding which uncertainties matter most—through sensitivity analysis—allows engineers to focus data acquisition efforts on the parameters that most strongly influence decision outcomes (Saltelli et al., 2008). This targeted approach to uncertainty management has become a cornerstone of modern reservoir engineering practice.

2.4.2 Types of Uncertainty

There are three main categories of uncertainty relevant to reservoir performance.

Table 2.1: Types of Uncertainty in Reservoir Forecasting

Type of Uncertainty	Description	Example
Geological	Arises from incomplete knowledge of subsurface structure and rock properties	Uncertainty in permeability distribution between wells
Fluid/Petrophysical	Results from imprecise characterization of oil, gas, and water properties	Uncertain viscosity or saturation values
Operational/Model-based	Caused by model assumptions or future operational changes	Uncertain choke settings, future drilling, or history-match errors

Geological uncertainty is typically the largest contributor to forecast variability in most reservoir studies. Even in heavily drilled fields, well spacing means that the vast majority of the reservoir volume remains unsampled. Properties such as permeability can vary by orders of magnitude over short distances, and the interpolation between wells is inherently uncertain. Seismic data helps constrain large-scale structure but cannot resolve fine-scale heterogeneity (Caers, 2005; Deutsch, 2002).

Fluid and petrophysical uncertainty arises from limited laboratory measurements and the challenge of obtaining representative samples. PVT properties measured on small fluid samples may not fully capture the behavior of the entire reservoir fluid system, particularly in compositionally complex reservoirs or near critical conditions. Similarly, relative permeability curves—critical for predicting multiphase flow—are often measured on a handful of core samples and must be extrapolated across the entire field (Ahmed, 2019).

Operational uncertainty encompasses both planned variations (such as testing different well control strategies) and unplanned events (equipment failures, regulatory changes, market conditions affecting production rates). While sometimes treated separately from geological uncertainty, operational decisions can significantly affect which reservoir outcomes are realized. For example, the timing and placement of infill wells, or changes in water injection strategy, can shift a reservoir from one performance trajectory to another (Jahn et al., 2008).

2.4.3 Traditional UQ techniques

- **Sensitivity Analysis:** One-at-a-time sensitivity analysis, while simple, can miss important interaction effects between parameters. For instance, the impact of permeability uncertainty may depend strongly on whether the reservoir has aquifer support or is under depletion drive. Global sensitivity methods such as Sobol indices or Morris screening address this limitation by systematically varying multiple parameters

simultaneously and quantifying both main effects and interactions (Saltelli et al., 2008). Tornado diagrams—which rank parameters by their influence on a chosen output metric—have become standard visualization tools in reservoir uncertainty studies because they immediately communicate which uncertainties deserve the most attention.

- **Scenario-Based-Forecasting:** While scenario-based approaches (optimistic/base/pessimistic) are intuitive and widely used in initial screening studies, they suffer from arbitrary selection of parameter combinations and provide no information about the likelihood of each scenario. A 'worst case' scenario might combine parameter values that are individually plausible but collectively improbable. This limitation led to the adoption of probabilistic methods that sample parameter space more systematically.
- **Monte Carlo Simulation:** The power of Monte Carlo simulation lies in its ability to propagate correlated uncertainties through complex models and generate full probability distributions rather than discrete scenarios. Modern Monte Carlo implementations often employ
- **Latin Hypercube Sampling (LHS)** rather than simple random sampling, as LHS ensures better coverage of the parameter space with fewer samples (McKay et al., 1979). For computationally expensive reservoir simulators, design-of-experiments approaches may be used to strategically select which parameter combinations to run, then fit response surfaces to approximate the full uncertainty distribution (Myers & Montgomery, 2002).

2.4.4 Advanced Probabilistic Methods

Modern workflows extend beyond simple Monte Carlo to more statistically rigorous approaches:

- **Bayesian UQ:** Bayesian approaches are particularly powerful during field life because they provide a formal framework for updating uncertainty as production data accumulates. The initial (prior) distribution of reservoir parameters, often based on analog fields or regional correlations, is progressively refined as wells come online and production behavior constrains which parameter values are plausible. This process—known as history matching in reservoir simulation—can be cast as a Bayesian inference problem where the goal is to find the posterior distribution of parameters that best explains observed data (Oliver & Chen, 2011; Tarantola, 2005). The challenge lies in the computational cost of sampling high-dimensional parameter spaces, which has driven development of efficient algorithms such as Ensemble Kalman Filtering.
- **Ensemble Modelling:** Ensemble modeling has become standard practice in large field development projects, particularly offshore developments where capital investments exceed billions of dollars. Operators typically generate tens to hundreds of equally plausible geological realizations, each honoring available data (well logs, seismic, core) but differing in their spatial property distributions. Full-field simulation on each realization produces an ensemble of production forecasts, from which probabilistic summaries (P10/P50/P90) are extracted (Caers, 2011). The computational expense of running multiple full-physics simulations has motivated the development of **reduced-order models** and **proxy-based approaches** that can approximate simulator behavior more efficiently during uncertainty quantification.

These methods are more realistic than traditional scenario modeling because they capture a wider range of uncertainty.

2.4.5 Computational Efficiency in Uncertainty Quantification

A practical challenge in reservoir uncertainty quantification is the computational cost of running thousands of simulation realizations, particularly when using detailed 3D models with complex physics (multiphase flow, geomechanics, thermal effects). A single simulation might take hours to days, making Monte Carlo analysis with 10,000 samples computationally prohibitive.

Several strategies have emerged to address this challenge:

Response surface Methodology: Fit a simplified mathematical function (polynomial, spline, or regression model) to approximate the relationship between uncertain inputs and forecast outputs, based on a limited number of full-simulation runs. Once calibrated, the response surface can be evaluated thousands of times almost instantaneously, enabling rapid probabilistic analysis (Queipo et al., 2005; Myers & Montgomery, 2002).

Reduced-Order models: Simplify the physics or spatial resolution to create computationally cheaper models that retain the essential behavior of the full simulator. Tank models and material balance approaches fall into this category—they sacrifice spatial detail but capture primary drive mechanisms and enable fast evaluation of parameter uncertainty (Dake, 1978).

Parallel Computing: Distribute simulation runs across multiple processors or cloud computing resources, dramatically reducing wall-clock time for large Monte Carlo studies. Modern reservoir simulators increasingly support parallel execution, making uncertainty quantification more accessible (Schlumberger, 2021).

The choice of approach depends on the study's objectives. For initial screening and parameter ranking, fast surrogate models may suffice. For final investment decisions, the full ensemble of detailed simulations is typically warranted despite the computational cost.

2.4.6 Summary Table

Table 2.2 summarizes the primary uncertainty quantification methods used in reservoir forecasting, highlighting their assumptions, advantages, and practical limitations. The selection of appropriate methods depends on the decision context, available computational resources, and the stage of field development.

Table 2.2: Comparison of Uncertainty Quantification Methods

UQ Method	Assumption	Advantage	Limitation
Sensitivity Analysis	One variable changes at a time	Easy to run	Oversimplified
Scenario Modeling	Limited fixed cases	Quick comparison	Not probabilistic
Monte Carlo Simulation	Randomized sampling	Provides distribution	Computational cost
Bayesian UQ	Probabilistic priors	Updates with data	Requires expertise
Ensembles	Multiple realizations	Captures geology	Heavy model-building effort

Uncertainty quantification has evolved from simple “what-if” scenarios to statistically robust frameworks that express outcomes as probability distributions. When integrated with reservoir simulation, UQ ensures that forecasts are not only physically realistic, but also risk-informed, which is essential for field development planning and decision making under uncertainty.

2.5 VALIDATION, METRICS AND BACK-TESTING STRATEGIES

2.5.1 Introduction to Model Validation

Validation is the process of testing how well a model can generalize to a unseen data. In reservoir engineering, this means checking whether a model trained on historical production or reservoir parameters can correctly predict future performance under uncertainty. In simulation-based studies, validation extends beyond statistical metrics to include physical consistency checks. A reservoir model that fits historical production data but violates mass balance, produces non-physical pressure profiles, or predicts fluid behavior inconsistent with PVT data cannot be considered validated, even if regression statistics appear favorable. Engineering judgment—assessing whether results 'make sense' given the known drive mechanisms, completion types, and operational constraints—remains an essential component of validation alongside quantitative metrics (Aziz & Settari, 1979; Ertekin et al., 2001).

Without proper validation:

- The model may overfit, memorising training data instead of learning trends.
- It may fail when applied to real field conditions.
- Reservoir performance forecasts become misleading and unreliable.

A good validation strategy ensures:

- Stability of predictions
- Robustness under new reservoir conditions
- Consistency in uncertainty quantification

Goal: confirm that the model is not only accurate on past data, but also trustworthy for decision-making.

2.5.2 Types of Validation Strategies

Several validation strategies are used depending on the reservoir model, data availability, and forecast time horizon:

2.5.2.1 Train-Test Split

- The data set is divided into two parts: historical data to train the model, and a reserved portion to test generalization
- Common in early-stage models or small datasets

2.5.2.2 K-fold Cross-Validation

- Data is split into k equal folds; each fold takes a turn being the test set.
- Helps reduce bias by averaging model performance across many splits.

2.5.2.3 Time-series Cross-Validation (Rolling Window)

- More realistic for reservoir forecast because production data is time-dependent.
- Training is performed on earlier timesteps, and validation is done on later timesteps – mimicking forecasting.

2.5.2.4 Walk-Forward Validation

- Common in forecasting reservoirs with strong dynamics or frequent updates.
- After each prediction window, the model is retrained with newly observed data, improving adaptability.

Time aware validation strategies (3 and 4) are generally preferred for decline curve forecasting, water cut prediction, or pressure transient modelling, because they respect the chronological order of events.

2.5.3 Back-Testing Strategies

Back-testing evaluates a model by replaying historical data to check whether the model would have made correct predictions at that time. It is widely used when models are applied to operational decision-making.

Table 2.3: Back-Testing Use Cases in Reservoir Forecasting

Scenario	Back-Testing Use Case
Production optimization	Check if suggested choke/flow control would have improved output
Waterflooding	Evaluate if injector patterns would reduce breakthrough
Decline curve fitting	Assess if the forecast from year N predicted years N+1 to N+3
Proxy models	Compare ML proxy vs. simulation output historically

Common reservoir back-testing settings include:

A good back-test:

- Uses a realistic time horizon
- Avoids “future leakage”
- Penalises overly optimistic forecasts

2.5.4 Key Performance Metrics (Technical Overview)

For simulation-based forecasting, validation metrics assess both accuracy (how close predictions are to actual or expected values) and precision (the width of uncertainty bands). A model might have high accuracy on average (P50 close to observed) but poor precision (P10-P90 range unrealistically wide), or vice versa. The ideal validated model minimizes both bias and uncertainty while maintaining physical realism.

Table 2.4: Key Performance Metrics for Reservoir Forecasting

Metric	Meaning	When Used
MAE (Mean Absolute Error)	Average absolute difference between prediction and true values	Stable decline or steady-state reservoirs
RMSE (Root Mean Squared Error)	Sensitive to large errors	High-uncertainty reservoirs (fractured / tight)
MAPE (Mean Absolute Percentage Error)	Error as a percentage of actual values	Useful for field comparison across scales
R² (Coefficient of Determination)	Measures goodness of fit	Model selection / benchmarking
P10–P50–P90	Probabilistic spread	Used in uncertainty-aware forecasts

Model performance is often evaluated with error metrics. The most common ones are: MAPE and RMSE are often reported together. MAPE communicates intuition to decision-makers, while RMSE highlights catastrophic failures.

2.5.5 Reliability, Robustness and Generalization Checks

Beyond simple metrics, reservoir forecasting models are also stress-tested for robustness, meaning:

- They can perform well under slightly perturbed input conditions
- They can handle missing or noisy field data
- They maintain stable predictions even when the reservoir enters late-life transition phases

Common robustness checks include:

- i. Noise injection into reservoir properties or rates to test sensitivity
- ii. Out-of-Distribution Testing, e.g., predicting behavior after a workover or refracturing event
- iii. Scenario-based Validation, e.g., high-water cut scenario vs. pressure depletion scenario.

This ensures not only accuracy, but resilience.

2.5.6 Model Calibration vs Validation

Calibration and validation are often confused:

Table 2.5: Model Calibration vs Validation

Aspect	Calibration	Validation
Purpose	Fit model closely to historical data	Test model performance on unseen data
Data Source	Same dataset as training	Held-out dataset / future timestamps
Risk	Overfitting	Underestimating generalization error
Used in	History matching	Forecast confirmation

A model that is well-calibrated but poorly validated is overfitted- this is extremely dangerous in field development planning because it gives a false sense of confidence.

2.5.7 Model Validation and Back-Testing Flow

In modern reservoir forecasting, validation is not treated as a single step at the end of modelling, but as an iterative workflow embedded throughout model development. A standard workflow begins with dataset preparation, followed by model training, internal validation, and repeated refinement before final back-testing on unseen historical data. This structure helps prevent

overfitting, ensures transparency in performance evaluation, and aligns the forecasting pipeline with field-engineering expectations.

To formalize the process, most studies adopt a loop-based (iterative) validation structure rather than a linear one. The most widely used workflow typically follows these stages:

- i.** Data Collection & Pre-Processing
- ii.** Feature Engineering / Dimensionality Reduction
- iii.** Model Training (Initial Fit)
- iv.** Internal Validation (Cross-Validation / Rolling Window)
- v.** Model Refinement or Retraining
- vi.** Back-Testing on Unseen Data
- vii.** Uncertainty Reporting & Confidence Interval Generation

This approach ensures that validation is part of model design, not a final checkpoint.

The main advantage of this workflow is that it reduces structural bias — the risk that a model appears accurate simply because it has “seen” the same data repeatedly. By isolating a portion of the dataset for back-testing and never exposing it to the training stage, the workflow creates a realistic test environment. This is functionally similar to how reservoir simulation history-matching is separated from predictive mode (Jensen et al., 2022). In addition, integrating uncertainty quantification at the end of the cycle ensures that performance is reported as a confidence band, rather than a single deterministic forecast, which is far more useful in field development decisions.

2.6 CASE STUDY AND APPLICATIONS

2.6.1 Simulation-Based Forecasting with Uncertainty Analysis: Industry Applications

Probabilistic simulation-based forecasting has become standard practice in major field development projects worldwide, particularly in settings where uncertainty is high and investment decisions involve substantial capital commitments. Several documented case studies illustrate the value of this approach:

Offshore Deepwater Developments: In deepwater Gulf of Mexico fields, operators routinely employ ensemble simulation workflows incorporating geological uncertainty (reservoir architecture, connectivity), fluid property uncertainty (PVT variations with depth), and operational uncertainty (well placement, production constraints). A study by Ligeró et al. (2003) in a deepwater turbidite reservoir demonstrated that probabilistic forecasting, using 50 geological realizations, provided decision-makers with production ranges that proved more reliable than single deterministic forecasts as the field was developed. The P10-P90 range successfully captured actual field performance over a five-year period.

Middle Eastern Carbonate Reservoirs: Uncertainty quantification has been critical in large carbonate fields where heterogeneity, fracture networks, and uncertain aquifer strength create significant forecasting challenges. Uncertainty studies in these settings typically focus on permeability distribution, aquifer influx rates, and relative permeability curves. Al-Harthy et al. (2005) reported that probabilistic forecasting enabled operators to optimize waterflood development strategies by identifying parameter sensitivities and guiding strategic well placement decisions.

Mature Field Redevelopment: In mature North Sea fields undergoing infill drilling or enhanced recovery projects, simulation-based uncertainty analysis helps evaluate the

economics of continued investment. Studies have shown that P10-P90 production ranges inform decisions about optimal well count, injection rates, and facility expansions, providing clearer risk-return profiles than deterministic forecasts alone (Jensen et al., 2022).

2.6.2 Simulation vs. Empirical Forecasting Methods

While decline curve analysis (DCA) remains widely used for quick screening and regulatory reporting, simulation-based approaches offer significant advantages when reservoir complexity or uncertainty is high. DCA assumes future decline will follow historical patterns, which breaks down when drive mechanisms change (e.g., water or gas breakthrough), when wells interfere with each other, or when operational strategies shift.

Material balance methods, though more physically grounded than DCA, require uniform pressure data that may not be available or representative in heterogeneous reservoirs. Numerical simulation, by explicitly modeling fluid flow, pressure propagation, and phase behavior, captures these complexities more accurately (Dake, 1978; Aziz & Settari, 1979).

The integration of uncertainty quantification with simulation provides a further advantage: rather than producing a single 'best estimate' that may prove incorrect, probabilistic simulation generates a range of outcomes with associated likelihoods, enabling risk-based decision-making. This is particularly valuable when comparing development options—the option with the highest P50 forecast may not be optimal if it also has a high probability of falling below economic thresholds (Rose, 2001).

2.6.3 Illustrative Example: Uncertainty Analysis in a Sandstone Reservoir

A representative uncertainty study in a North Sea sandstone reservoir (reported by Damsleth et al., 1998) illustrates the practical workflow. The study team identified key uncertainties: net-

to-gross ratio (0.6-0.8), horizontal permeability (50-250 md), aquifer strength (weak to moderate), and structural dip uncertainty affecting hydrocarbon-water contact.

Using Monte Carlo sampling with Latin Hypercube design, 200 reservoir realizations were simulated over a 20-year forecast period. Results showed:

- P90 (conservative) cumulative oil: 45 MMstb
- P50 (most likely): 62 MMstb
- P10 (optimistic): 78 MMstb

Sensitivity analysis revealed that aquifer strength and net-to-gross were the dominant uncertainties affecting recovery, while permeability primarily affected early-time production rates rather than ultimate recovery. This insight guided subsequent appraisal drilling to better characterize net pay thickness and aquifer behavior, reducing uncertainty where it mattered most for development planning. The probabilistic forecast provided a transparent basis for investment decisions, with the P90 case used for economic downside analysis and facility sizing.

2.7 CHALLENGES AND FUTURE DIRECTIONS IN SIMULATION-BASED UNCERTAINTY QUANTIFICATION

Despite significant advances in reservoir simulation and uncertainty analysis methods, several challenges remain that motivate continued research and methodological development.

Computational Cost vs. Physical Detail: High-resolution 3D simulation models with millions of grid cells, complex fluid models (compositional, thermal), and coupled physics (geomechanics, geochemistry) can require days to run a single realization on current hardware.

This creates a tension between physical realism and the need to run hundreds or thousands of scenarios for meaningful uncertainty quantification. While parallel computing and response

surface methods help, the fundamental tradeoff between model complexity and probabilistic analysis depth persists (Sarma et al., 2008).

Representing and Sampling Geological Uncertainty: Geological heterogeneity is continuous and spatial, making it inherently high-dimensional. Techniques such as geostatistical simulation (Sequential Gaussian Simulation, multiple-point statistics) can generate varied reservoir models, but ensuring these models span the true range of geological plausibility—without being either too conservative or too optimistic—remains challenging. Moreover, structural uncertainties (fault positions, horizon interpretations) are often treated separately from property uncertainties, when in reality they are interrelated (Caers, 2011).

Integration of Multiple Data Types and Scales: Reservoir models must integrate diverse data: core measurements (cm scale), log responses (meter scale), well tests (10-100m scale), and seismic (100-1000m scale). Each data type has different resolution, uncertainty, and spatial coverage. Bayesian frameworks provide a theoretical foundation for this integration, but practical implementation—particularly honoring data at the correct scale and weighting observations appropriately—requires careful attention and remains an active research area (Tarantola, 2005).

Accessibility and Transparency: Commercial reservoir simulators (ECLIPSE, CMG, STARS) are powerful but expensive, proprietary, and often perceived as "black boxes" by non-specialists. This limits their use in academic settings and smaller operators. Furthermore, the complexity of these tools can obscure the underlying physics, making it difficult to validate results or explain them to decision-makers unfamiliar with reservoir engineering. Open-source, Python-based simulation frameworks that implement core physics in a transparent, modifiable way offer a promising alternative—particularly for screening studies, teaching, and situations where commercial software licenses are unavailable. Such tools democratize access to

probabilistic forecasting capabilities and enable customization for specific reservoir types or research questions.

Looking Forward: The future of reservoir forecasting likely lies not in choosing between different methodologies, but in strategically combining them. Simulation provides physical realism and interpretability; fast surrogate models enable extensive uncertainty exploration; data-driven pattern recognition (where appropriate) can identify relationships in large datasets. Hybrid frameworks that leverage the strengths of each approach—using simulation for process understanding, surrogates for computational efficiency, and probabilistic methods for decision support—represent the most promising direction. This study contributes to that vision by developing an accessible Python-based simulation and uncertainty quantification framework that maintains transparency while enabling practical probabilistic forecasting.

2.8 SYNTHESIS AND POSITIONING OF CURRENT STUDY

The literature reviewed demonstrates a clear evolution in reservoir forecasting methods, driven by increasing reservoir complexity, growing datasets, and the critical need for uncertainty-aware predictions. Early approaches—decline curve analysis and material balance—provided practical solutions given limited computational resources, but struggled with heterogeneous reservoirs and offered no explicit measure of forecast uncertainty. Numerical reservoir simulation addressed many of these limitations by incorporating detailed physics and spatial variability, yet remained computationally expensive and often produced single deterministic forecasts.

The past two decades have seen a major shift toward probabilistic forecasting as the industry standard for significant field development decisions. Monte Carlo simulation, ensemble modeling, and Bayesian updating provide frameworks for quantifying uncertainty and

presenting decision-makers with ranges of outcomes rather than false-precision point estimates. However, these methods typically rely on commercial simulation software that is expensive, proprietary, and not easily adaptable for research or specialized applications.

Parallel to these developments, machine learning has emerged as a tool in petroleum engineering, particularly for tasks like log interpretation, seismic processing, and pattern recognition in large datasets. While ML offers computational speed and can handle complex nonlinear relationships, its application to reservoir forecasting faces significant constraints: data requirements are substantial, physical interpretability is limited, and generalization across different reservoir types remains uncertain. For these reasons, physics-based simulation integrated with modern uncertainty quantification methods continues to represent the most reliable and widely accepted approach for production forecasting, particularly when transparency, validation, and regulatory compliance are required.

Research Gap and Study Contribution: A gap exists in the availability of accessible, transparent reservoir simulation tools that combine physical realism with modern probabilistic methods. Commercial simulators are powerful but costly and opaque; simplified analytical models lack the flexibility to represent diverse reservoir types and drive mechanisms. This study addresses that gap by developing a Python-based reservoir simulation framework with integrated uncertainty quantification. By implementing material balance principles and production decline modeling in an open-source, customizable environment, the framework provides:

- **Physical realism** grounded in reservoir engineering fundamentals
- **Computational efficiency** enabling rapid Monte Carlo and sensitivity analysis
- **Transparency** in model assumptions and calculations
- **Accessibility** without commercial software licensing requirements

- **Flexibility** to adapt to different reservoir types and operational scenarios

The framework demonstrates that rigorous probabilistic forecasting can be achieved without sacrificing interpretability or physical consistency. It serves both practical engineering purposes—screening studies, uncertainty quantification, teaching—and provides a foundation for future enhancements such as multi-phase flow extensions, spatial heterogeneity, or integration with optimization algorithms. In this way, the study contributes to the broader goal of making robust, uncertainty-aware reservoir forecasting more accessible and transparent across the petroleum engineering community.

CHAPTER THREE

METHODOLOGY

3.1 RESEARCH DESIGN

This study adopts a simulation-based quantitative approach for reservoir performance forecasting, integrating material balance principles with probabilistic uncertainty analysis. The methodology centers on developing a computationally efficient tank model that captures primary reservoir drive mechanisms while enabling systematic exploration of parameter uncertainty through Monte Carlo techniques. The design involves building a conceptual numerical model of a typical petroleum reservoir, incorporating rock and fluid properties, production history, and operational constraints. While the physical principles underlying the model are consistent with those used in industry-standard simulators like ECLIPSE or CMG, the implementation is built from fundamental equations in Python. This approach sacrifices some of the spatial resolution of full 3D gridded models but gains significant advantages in computational speed, transparency, and adaptability for uncertainty studies. The study is primarily quantitative, focusing on the predictive behavior of the reservoir under various operational scenarios. By integrating uncertainty quantification into the simulation workflow, the methodology allows assessment not only of expected production rates but also of the confidence intervals surrounding those forecasts. This ensures that decision-making can be guided by probabilistic insights rather than deterministic assumptions alone.

The research design is structured to address three main objectives:

- i. Evaluate reservoir performance under various production strategies.
- ii. Quantify the impact of uncertain reservoir parameters on production forecasts.

- iii. Develop a computational workflow that combines full-field simulation principles with Python-based modeling and probabilistic forecasting.

This hybrid design ensures that while the study is computationally feasible in Python, it retains the rigor and representativeness of a full-field reservoir simulation study.

3.2 DATA REQUIREMENTS AND SOURCES

This study requires a combination of static, dynamic, **and** fluid-property datasets typically used in full-field reservoir simulation. These datasets provide the physical description of the reservoir, its behavior under production, and the governing fluid–rock interactions needed for forecasting.

(a) Static Reservoir Data

These describes the geological and petrophysical characteristics of the reservoir:

- Well logs (GR, RHOB, NPHI, DT, resistivity) – used to estimate porosity, lithology, and saturation distributions.
- Core analysis data (SCAL/RCAL) – provides relative permeability curves, capillary pressure, and rock types.
- Structural maps – top and base reservoir mapping for areal geometry.
- Net-to-gross and thickness maps – defines effective pay zones.

These are typically obtained from geological modeling outputs or existing reservoir characterization studies.

(b) Dynamic Reservoir/Production Data

These reflect how the reservoir has behaved historically under depletion:

- Historical production rates (oil, gas, water)
- Bottom-hole pressures / flowing wellhead pressures
- Well test data (PI, skin, transmissibility)
- Choke settings / operational constraints

This data is usually sourced from production surveillance databases or internal field performance reports.

(c) Fluid Property Data (PVT)

These define the thermodynamic behavior of reservoir fluids:

- Bubble point pressure
- Solution GOR
- Oil formation volume factor
- Viscosity, compressibility, density
- Gas deviation factor

Such data are generally obtained from laboratory PVT reports or simulation input decks.

(d) Auxiliary Data

Depending on complexity, additional supporting data may include:

- Aquifer support estimates
- Water-oil contact movement
- Fracture presence indicators
- Well intervention history

In exploratory or teaching contexts where proprietary field data is unavailable, representative parameter values drawn from published reservoir analogs provide a viable alternative. This approach—using typical ranges for sandstone or carbonate reservoirs documented in petroleum engineering literature—allows the methodology to be demonstrated and validated without compromising confidentiality restrictions. The framework itself remains fully adaptable: when field-specific data becomes available, it can be directly substituted without altering the underlying computational structure.

Data Source Justification

Although this study is executed in Python, the structure of inputs is aligned with what a full-field ECLIPSE model would require. This ensures compatibility between classical simulation thinking and later Python-based post-processing for uncertainty quantification. In a real project, such data would be sourced from field operators, reservoir management databases, and PVT laboratory reports. For this study, the simulation framework utilizes representative reservoir parameters consistent with typical clastic reservoirs as documented in standard references such as Ahmed (2019), Craft & Hawkins (1959), and Dake (1978). This approach ensures the methodology is reproducible and accessible while demonstrating the framework's capability to handle realistic parameter ranges and operational constraints.

3.3 DATA PRE-PROCESSING AND QUALITY CHECKS

Before the data can be used for modeling and simulation-based uncertainty analysis, a series of cleaning and validation steps are required to ensure consistency, reliability, and usability. This prevents the propagation of measurement errors or inconsistencies into the forecasting and risk assessment stage.

(a) Data Cleaning

This involves removing or correcting errors within the raw dataset:

- Detection and removal of duplicate records
- Correction of unrealistic values (e.g., negative permeability or porosity > 1)
- Filtering noisy production signals (e.g., choke changes misinterpreted as decline)
- Ensuring pressure units remain consistent (psia vs psig, bars, kPa etc.)

(b) Handling Missing Data

Gaps in measurements are common, especially in production histories:

- Interpolation for short-duration missing records
- Forward/Backward fill for transient operational logs
- Regression/curve fitting for longer-term gaps
- If data is non-recoverable → mark for sensitivity-based reconstruction

(c) Outlier Detection

Outliers can distort model calibration. The strategy includes:

- Statistical screening (IQR method, z-score, rolling median smoothing)
- Engineering judgment validation (e.g., sudden 500% production jump)
- Removal or replacement using trend-based correction

(d) Unit Standardization

Different datasets may come in inconsistent units. All variables will be converted to simulation-consistent units such as:

Table 3.1: Unit Standardization for Reservoir Properties

Property	Field Unit	SI Equivalent
Pressure	psi	kPa/bar
Rate	stb/d	m ³ /day
Permeability	md	m ²
Viscosity	cP	mPa·s

The study will maintain field units internally (for practical petroleum engineering workflows), converting only when required by Python calculations.

(e) Data Completeness Validation

Before calibration or model input generation:

- Check for well-level vs field-level consistency
- Verify continuity between static and dynamic datasets
- Validate that PVT ranges are realistic for the reservoir type
- Confirm reservoir boundaries are defined

Once pre-processing is complete, the dataset becomes sufficiently structured to serve as input for model initialization and further uncertainty quantification.

3.4 MODEL SELECTION/RESERVOIR SIMULATION APPROACH

The modeling phase forms the core of this research, linking geological understanding, production behavior, and computational forecasting through a structured simulation approach. Because this study focuses on uncertainty and risk analysis using Python-based modeling, the workflow integrates deterministic simulation concepts with probabilistic sampling methods to produce robust forecasts.

(a) Reservoir Modeling Framework

This research employs a tank model (also known as a lumped-parameter or zero-dimensional model) to represent reservoir behavior. Tank models aggregate the reservoir into a single control volume characterized by average properties—porosity, permeability, pore volume, fluid saturations—and track changes in these properties over time as production occurs and pressure declines.

The tank model approach is grounded in material balance principles, one of the foundational methods in reservoir engineering (Craft & Hawkins, 1959; Dake, 1978). Material balance relates cumulative production to observable changes in reservoir pressure, accounting for fluid expansion, water influx, and gas cap behavior. By solving the material balance equation at each time step, the model predicts how production rates evolve under specified operating constraints (such as minimum bottom-hole pressure or maximum surface rate).

Why Tank Models for Uncertainty Studies:

While tank models sacrifice the spatial detail of full 3D gridded simulation, they offer several compelling advantages for uncertainty quantification:

1. **Computational Speed:** A single tank model forecast might complete in under one second, compared to minutes or hours for a detailed grid-based simulation. This speed enables Monte Carlo analyses with thousands of realizations—impractical with computationally expensive simulators.
2. **Parameter Transparency:** All inputs (porosity, permeability, aquifer strength, PVT properties) directly correspond to physical reservoir characteristics. There are no hidden numerical parameters or grid-related artifacts, making sensitivity analysis more interpretable.
3. **Validation Simplicity:** Tank model results can be checked against analytical solutions and material balance closure criteria, providing clear benchmarks for correctness.
4. **Appropriate Physical Fidelity:** For many field-scale forecasting questions—particularly those focused on ultimate recovery, pressure maintenance strategies, or economic screening—the spatial averaging inherent in tank models is acceptable. The primary uncertainties often relate to average properties (overall permeability-thickness product, total pore volume, aquifer strength) rather than fine-scale heterogeneity.

Tank models have been successfully applied in numerous uncertainty studies, particularly during early appraisal stages when detailed geological models may not yet be available or when rapid scenario testing is required (Dake, 1978; Ahmed, 2019)

(b) Model Components and Governing Equations

The simulation framework comprises three interconnected elements:

Material Balance Module: Implements the general material balance equation to track reservoir pressure as a function of cumulative production, fluid properties, and aquifer influx. This module accounts for oil, gas, and water expansion as pressure declines, as well as potential

water drive support. The mathematical formulation follows established textbook treatments (Craft & Hawkins, 1959; Dake, 1978).

Production Decline Module: Calculates production rates at each time step based on current reservoir pressure, well productivity index, and operating constraints (minimum BHP, maximum rate limits). The relationship between pressure drawdown and rate is governed by Darcy's law for radial flow, adapted for single-well or field-scale representation.

Aquifer Model (Optional): Incorporates water influx from an adjacent aquifer using the van Everdingen-Hurst unsteady-state aquifer model or simplified Fetkovich steady-state representation. Aquifer strength—characterized by aquifer size and permeability—is a key uncertainty parameter in many reservoirs.

These components interact at each time step: production reduces reservoir pressure (material balance), pressure decline limits production rate (inflow performance), and aquifer influx partially offsets pressure decline (if present). The model marches forward in time, generating cumulative production, rate profiles, and pressure histories that can be compared against field observations or used for predictive forecasting.

(c) Reservoir Representation

The Python-based simulator will represent the reservoir as a conceptual tank model characterized by:

- Average porosity and permeability values
- Fluid PVT properties (oil, gas, water)
- Reservoir pressure and temperature
- Aquifer strength (if present)

- Production constraints (rates, BHP limits)

These parameters define the base case (deterministic) scenario, which will later be perturbed within defined uncertainty ranges for stochastic analysis.

The tank model assumes the reservoir behaves as a single, well-mixed volume—an approximation that works best for reservoirs with good vertical and lateral communication. For highly compartmentalized reservoirs or those with strong permeability barriers, multi-tank models could be employed, though at the cost of additional parameters and reduced computational speed. For the purposes of demonstrating uncertainty quantification methodology, the single-tank representation provides sufficient complexity to illustrate key concepts while remaining computationally tractable.

(d) Computational Workflow

The simulation and uncertainty analysis workflow proceeds through the following stages:

Stage 1 - Base Case Development:

Define nominal (most likely) values for all reservoir parameters: porosity, permeability, initial pressure, pore volume, PVT properties, aquifer characteristics, and operational constraints. Run the deterministic tank model to generate a baseline production forecast. Validate this baseline against material balance closure and expected physical trends (e.g., exponential decline in depletion drive, flattened decline with aquifer support).

Stage 2 - Uncertainty Parameter Identification:

Based on data quality assessment and engineering judgment, identify which parameters are uncertain and define plausible ranges for each (discussed in detail in Section 3.6). Specify

probability distributions (uniform, triangular, normal) that represent prior beliefs about parameter likelihoods.

Stage 3 - Monte Carlo Sampling:

Use Latin Hypercube Sampling or pseudo-random number generation to create a large set (e.g., 1,000-10,000) of parameter combinations, each representing an equally plausible reservoir realization. Each combination samples from the specified distributions while optionally preserving correlations between related parameters (e.g., porosity and permeability often correlate positively).

Stage 4 - Ensemble Simulation:

Execute the tank model for each sampled parameter set, generating an ensemble of production forecasts. Due to the model's computational efficiency, this step typically completes in minutes even for 10,000 realizations.

Stage 5 - Statistical post-processing:

Aggregate results across the ensemble to compute P10, P50, and P90 production profiles, cumulative recovery distributions, and summary statistics. Perform sensitivity analysis (e.g., Spearman rank correlation, Sobol indices) to identify which uncertain parameters most strongly influence forecast outcomes.

Stage 6 - Validation and Interpretation:

Check that the ensemble behaves consistently with physical expectations, that material balance is preserved across realizations, and that the P10-P90 range is reasonable given input

uncertainties. Interpret results in the context of decision-making: which parameters warrant additional data acquisition, which development scenarios are robust across uncertainty, etc.



Figure 3.1: Simulation Workflow Diagram

3.5 SIMULATION SETUP

The simulation setup defines how the reservoir model is configured before running deterministic and uncertainty-based simulations. This includes how the reservoir is spatially represented, how rock and fluid properties are incorporated, and how initial and boundary conditions are assigned.

Even though this study uses a Python-based conceptual simulator rather than a black-box commercial software, the setup follows standard reservoir engineering practice to ensure consistency with physical behavior.

3.5.1 Reservoir Representation and Grid Setup

The reservoir will be modeled as a homogeneous single-tank (lumped parameter) system for the base case, which captures average reservoir properties. This simplified geometry is chosen because the emphasis of the study is uncertainty analysis, not spatial heterogeneity modeling.

Later sensitivity tests may include:

- pseudo-layering (vertical permeability contrast)
- aquifer strength changes
- production constraint variations

This allows the study to reflect field behavior while keeping computational complexity low.

3.5.2 Rock and Fluid Property Specification

Key rock and fluid properties used in the base model include:

Table 3.2: Rock and Fluid Properties Used in Base Model

Property	Description	Role in Simulation
Porosity (ϕ)	Storage capacity	Determines OOIP and pressure decline
Permeability (k)	Flow capacity	Controls production rate trends
Compressibility (C_r, C_f)	Rock/fluid expansion	Governs pressure response
Relative Permeability	Phase flow distribution	Influences WOR/GOR
PVT ($B_o, R_s, \mu_o, \mu_g, \mu_w$)	Fluid phase behavior	Determines surface volumes

These properties interact through the material balance equation and inflow performance relationships to determine production behavior. For instance, high permeability enables high initial rates but does not necessarily increase ultimate recovery if the reservoir is primarily depletion-driven. Conversely, strong aquifer support (reflected in high aquifer permeability and large aquifer volume) can sustain pressure and dramatically improve recovery factors, though at the cost of earlier water breakthrough.

Static values are used for the deterministic run, while uncertainty bounds will later be applied in Section 3.7.

3.5.3 Initialization and Boundary Conditions

The model is initialized using:

- Initial reservoir pressure (P_i)
- Initial oil in place (OOIP)
- Temperature (T)
- Bubble point condition

- Aquifer presence or absence

Boundary conditions depend on drive mechanism assumptions (water-drive, depletion-drive, etc.), which will later be varied under uncertainty scenarios.

3.5.4 Production Control/Control Settings

The model will be simulated under realistic operating limits, including:

- Maximum production rate (economic constraint)
- Minimum bottom-hole pressure (BHP limit)
- Time steps (monthly / yearly forecast windows)

These constraints ensure that the simulation output stays consistent with engineering feasibility.

3.5.5 History Matching Approach

If historical data is available (even for a conceptual case), a lightweight history-matching step will be applied by tuning:

- permeability multipliers
- aquifer strength
- skin / productivity index adjustments

This ensures the deterministic base model is anchored to plausible reservoir behavior before uncertainty propagation. In cases where historical data is unavailable—such as in this study using representative parameters—the 'history match' serves as a reasonableness **check**: does the model produce decline trends and recovery efficiencies consistent with published field analogs? This validation against expected behavior, rather than specific data points, ensures the model captures realistic physics even without proprietary field measurements.

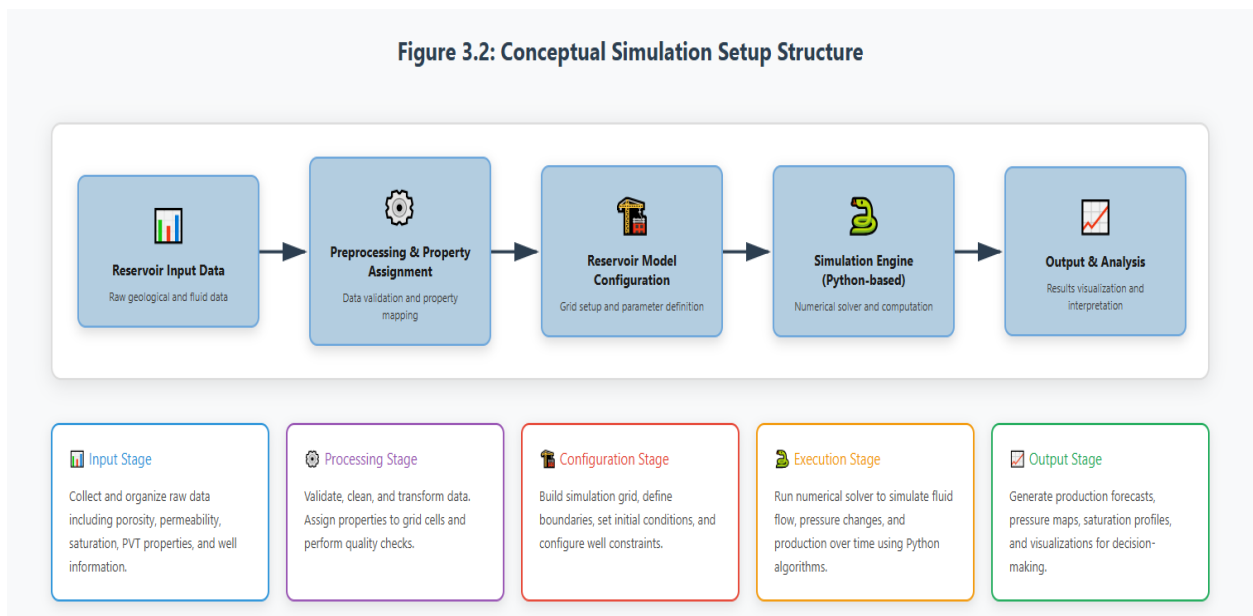


Figure 3.2 Conceptual Simulation Setup Structure

3.6 UNCERTAINTY IDENTIFICATION AND PARAMETER SELECTION

Uncertainty identification is a crucial stage in reservoir modeling and risk assessment because reservoir systems are inherently complex and data-limited. The purpose of this stage is to determine which parameters are uncertain, how they influence production forecasts, and how they will be varied during uncertainty quantification.

3.6.1 Geological and Petrophysical Properties

Table 3.3: Geological and Petrophysical Uncertainties

Parameter	Type	Influence on Forecast
Porosity (ϕ)	Static	Affects storage capacity and OOIP estimation
Permeability (k)	Static	Controls production rates and recovery efficiency
Net-to-Gross Ratio (N/G)	Static	Determines effective flow area
Reservoir Thickness (h)	Static	Impacts hydrocarbon volume and pressure response
Aquifer Strength (We)	Dynamic	Governs pressure maintenance and water influx behavior

In this study, uncertainties are classified into three main categories — geological, fluid and PVT-related, and operational/engineering.

These uncertainties originate from the incomplete understanding of subsurface rock properties and geometry. They have a significant impact on reservoir connectivity, flow capacity, and volumetric estimates. These parameters are commonly derived from logs, cores, or seismic interpretation, all of which carry measurement and interpretational uncertainties.

3.6.2 Fluid Property and PVT Uncertainty

Fluid properties control how hydrocarbons behave under changing pressure and temperature. Even small variations can have major effects on recovery and forecast precision.

Table 3.4: Fluid Property and PVT Uncertainties

Parameter	Uncertainty Source	Effect
Oil Formation Volume Factor (Bo)	Limited PVT sampling	Affects surface rate and volumetric conversion
Solution GOR (Rs)	Laboratory correlation	Impacts pressure behavior and gas-oil ratio forecast
Viscosity (μ_o, μ_w)	Measurement variation	Alters flow mobility and water cut
Compressibility (Cf)	Temperature/pressure assumption	Affects pressure depletion rate

3.6.3 Operational and Engineering Uncertainties

These are associated with how wells are operated and how the field performs under human or mechanical control.

Table 3.5: Operational and Engineering Uncertainties

Parameter	Description	Effect on Forecast
Well Productivity Index (PI)	Derived from well test data	Impacts initial production rate
Skin Factor (S)	Reflects near-wellbore condition	Alters well performance and decline trend
Bottom-hole Pressure Limit (BHP)	Operational setting	Defines maximum allowable drawdown
Choke and Rate Control	Field operating strategy	Modifies decline behavior and water cut
Economic Limit (Qmin)	Termination condition	Determines forecast end point

3.6.4 Sensitivity Ranges for Key Parameters

The table below summarizes representative uncertainty ranges based on literature and field analogs for similar elastic reservoirs:

Table 3.6: Sensitivity Ranges for Key Parameters

Parameter	Base Value	Range ($\pm\%$)	Reference Source
Porosity (ϕ)	0.20	10–30%	Ahmed (2019)
Permeability (k)	100 md	50–200%	Tiab & Donaldson (2015)
Bo	1.3 bbl/STB	5–10%	McCain (2011)
Rs	500 scf/STB	10–20%	Standing (1947)
PI	3 stb/d/psi	15–25%	Production test data

These uncertainty ranges will later feed into the Monte Carlo and probabilistic modeling stage (Section 3.7), forming the backbone of the uncertainty quantification workflow.

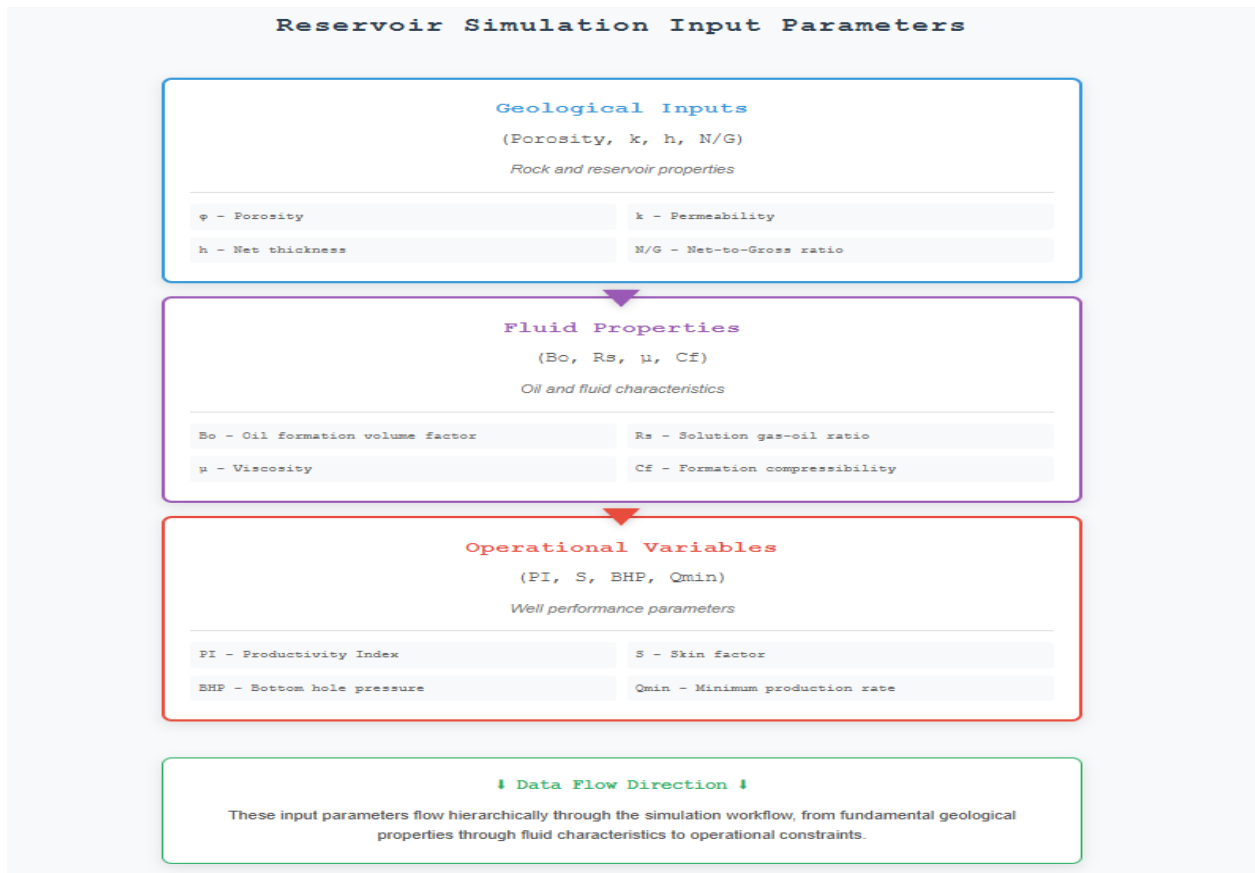


Figure 3.3: Uncertainty Parameter Identification and Ranges

With these uncertainty sources identified and parameterized, the next step will be to quantify their effects through systematic stochastic or statistical methods. These ranges reflect typical uncertainties encountered in reservoir studies based on data quality reviews in published literature (Ahmed, 2019; Tiab & Donaldson, 2015). In a specific field application, ranges would be narrowed or widened based on local data availability: wells with extensive core coverage might reduce permeability uncertainty to $\pm 30\%$, while seismically defined structures could constrain thickness uncertainty more tightly. The framework accommodates such refinements by simply adjusting the input distributions.

3.7 UNCERTAINTY QUANTIFICATION (UQ) METHODS

Uncertainty Quantification (UQ) is the core analytical phase of this study. It aims to capture the combined effect of geological, fluid, and operational uncertainties on reservoir performance forecasts. The goal is not only to estimate the range of possible outcomes but also to assess the likelihood of each, supporting data-driven decision-making under uncertainty.

This section describes the methods and computational workflow used to propagate uncertainty and generate probabilistic forecasts in Python.

3.7.1 Conceptual Overview

In deterministic reservoir models, all input parameters are fixed, leading to a single “best estimate” forecast. However, in reality, every parameter — from porosity to permeability — has a distribution of possible values. UQ, therefore, transforms the model from a single deterministic output to a probability distribution of outcomes, allowing engineers to quantify risk and confidence in production forecasts (Damsleth et al., 1998).

3.7.2 Uncertainty Quantification Strategy

This study employs Monte Carlo simulation as the primary uncertainty quantification method, complemented by sensitivity analysis to rank parameter importance. Monte Carlo is selected for several reasons:

- **Generality:** It makes no assumptions about the form of input-output relationships and naturally handles nonlinearities.
- **Simplicity:** The concept is intuitive—repeatedly sample inputs, run the model, aggregate outputs—making results easy to communicate to non-specialists.
- **Statistical Rigor:** With sufficient samples (typically 1,000-10,000), Monte Carlo provides robust estimates of output distributions and percentiles.

- **Computational Feasibility:** Given the tank model's speed (~0.1-1 second per realization), even 10,000 runs complete in under three hours on a standard laptop.

Latin Hypercube Sampling (LHS) is used for parameter generation rather than simple random sampling. LHS stratifies the probability space, ensuring better coverage with fewer samples—particularly important when some parameters may have limited influence and we want to avoid clustering samples in similar regions (McKay et al., 1979). For a typical analysis with 7-10 uncertain parameters, 1,000-5,000 LHS samples generally suffice to produce stable P10/P50/P90 estimates.

Following Monte Carlo execution, global sensitivity analysis identifies which parameters most strongly influence forecast outcomes. Two complementary methods are employed:

1. **Spearman Rank Correlation:** Measures monotonic relationships between each input parameter and selected outputs (e.g., cumulative oil at 10 years, peak production rate). Parameters with high correlation coefficients (positive or negative) are identified as influential.
2. **Sobol Indices (Optional):** For deeper insight, variance-based Sobol indices decompose output variance into contributions from individual parameters and their interactions. This reveals not just which parameters matter, but whether interaction effects are significant (Saltelli et al., 2008).

Results from sensitivity analysis guide practical decisions: parameters with high sensitivity warrant more precise characterization through additional data acquisition, while parameters with negligible influence can be fixed at nominal values in future studies, simplifying the analysis.

3.7.3 Implementation Procedure

The uncertainty quantification workflow is implemented in Python using the following steps:

Step 1: Define Parameter Distributions

For each uncertain parameter (e.g., porosity, permeability, aquifer strength), specify a probability distribution that represents prior knowledge. Common choices include:

- **Uniform distribution:** When all values in a range are equally plausible.
- **Triangular distribution:** When a most likely value exists, with tails extending to minimum and maximum credible bounds.
- **Normal (Gaussian) distribution:** When measurements suggest a bell-curve pattern around a central value.

Distributions are defined using SciPy's statistical functions, which provide efficient random sampling capabilities.

Step 2: Generate Sample Matrix

Using the SALib library's Latin Hypercube Sampling function, generate a matrix of parameter combinations:

$$\mathbf{X} = [x_1, x_2, \dots, x_n]_i \text{ for } i = 1 \text{ to } N_{\text{samples}}$$

where N_{samples} is typically 1,000-10,000. Each row represents one complete parameter set for a single simulation run.

Step 3: Execute Simulation Ensemble

Loop over the sample matrix, running the tank model for each parameter set:

python

```
for i in range(N_samples):  
  
    params = X[i]  
  
    results[i] = run_tank_model(params)
```

The results array stores time-series production profiles, cumulative volumes, pressures, and other outputs for each realization.

Step 4: Extract Statistical Summaries

From the ensemble results, compute:

- **Percentiles:** P10, P50 (median), P90 at each time step using NumPy's `percentile()` function.
- **Mean and Standard Deviation:** Characterize central tendency and spread.
- **Probability Distributions:** Fit histograms or kernel density estimates to cumulative recovery at key time points (e.g., 5 years, 10 years, field life).

Step 5: Sensitivity Analysis

Calculate correlation coefficients between each input parameter and selected output metrics:

python

```
from scipy.stats import spearmanr  
  
correlation, p_value = spearmanr(X[:, param_index], results[:, output_metric])
```

Rank parameters by absolute correlation to identify the most influential uncertainties. Generate tornado diagrams that visually display sensitivity rankings.

Step 6: Visualization

Create graphical outputs:

- **Fan charts:** Show P10, P50, P90 production rate envelopes over time.
- **Cumulative distribution functions (CDFs):** Display the probability of achieving various recovery levels.
- **Tornado diagrams:** Rank parameters by their influence on a target metric (e.g., 10-year cumulative oil).

These visualizations communicate uncertainty effectively to decision-makers who may not be familiar with statistical details but need to understand the range of possible outcomes

3.7.4 Mathematical Representation

For each uncertain parameter vector $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$, the simulator response \mathbf{Y} (e.g., cumulative production) can be expressed as:

$$Y = f(\mathbf{X}) + \varepsilon$$

where:

- $f(\mathbf{X})$ = deterministic simulator function
- ε = random noise or model error

The Monte Carlo estimate of the expected output is then:

$$E[Y] = \frac{1}{N} \sum f(x_i)$$

and the uncertainty (variance) is approximated by:

$$\text{Var}[Y] = \frac{1}{N-1} \sum (f(x_i) - E[Y])^2$$

This provides both the expected production forecast and confidence intervals around it.

3.7.5 Visualization of Uncertainty

Results are visualized as probability bands or fan plots, as illustrated below

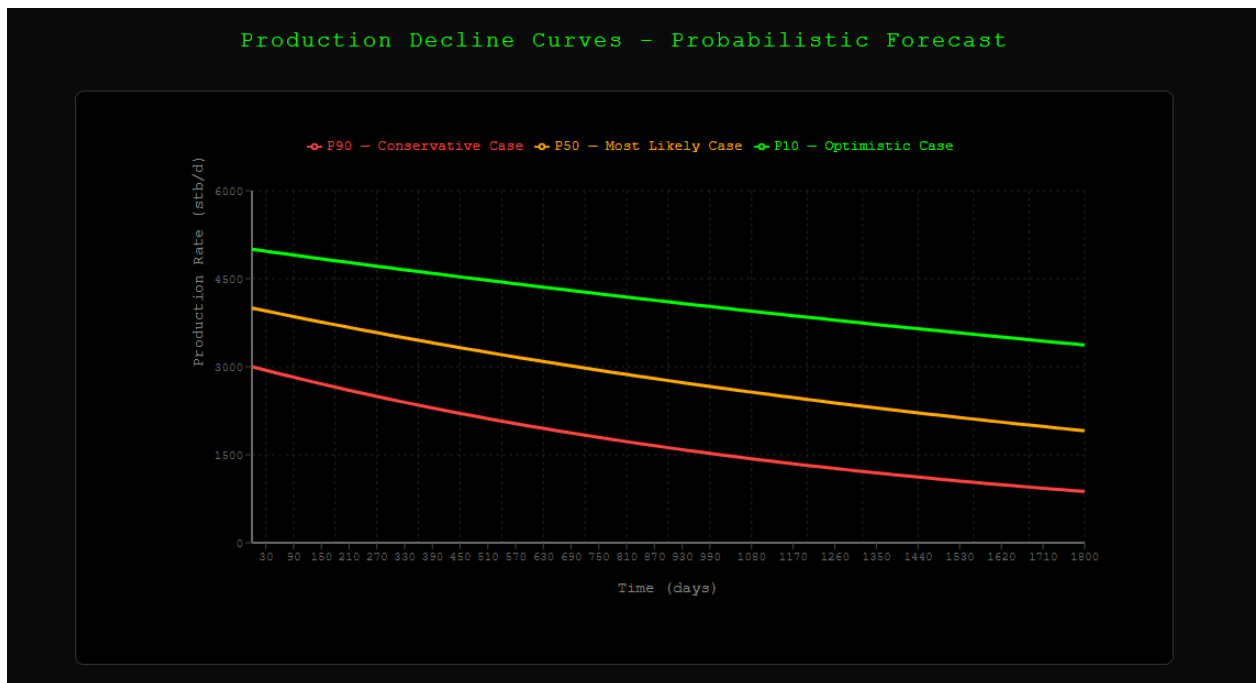


Figure 3.4: Uncertainty Forecast Visualization (Fan Plot)

Such graphical outputs help decision-makers quickly evaluate the range of possible production outcomes and associated risks.

3.7.6 Python Implementation Tools

To implement the above methods, the following Python libraries will be employed:

- **NumPy, Pandas** – data handling and vectorized computations
- **SciPy** – statistical sampling and distribution fitting
- **scikit-learn (optional)** -- for polynomial regression if response surface approximation is needed

- **matplotlib / seaborn** – uncertainty visualization
- **SALib** – global sensitivity analysis (e.g., Sobol, Morris methods)

The tank model itself is implemented using NumPy for vectorized numerical operations and SciPy's ODE solvers for time-stepping the material balance equations. Material balance and inflow equations are formulated as a system of ordinary differential equations (ODEs) that track pressure and saturations through time, with production rates computed at each step based on current conditions and operational constraints.

3.7.7 Deliverables from this Stage

By the end of this section, the study will produce:

- Probabilistic production forecasts (P10, P50, P90)
- Sensitivity ranking of uncertain parameters
- Quantified uncertainty ranges for key performance indicators (e.g., EUR, recovery factor, pressure decline)

These results provide the foundation for validation, model evaluation, and risk interpretation in the subsequent sections.

3.8 MODEL EVALUATION AND VALIDATION

Model evaluation and validation are vital for ensuring that the developed reservoir model and its uncertainty quantification results are both technically sound and statistically reliable. In this phase, the model's predictive accuracy, stability, and realism are tested using both engineering checks and quantitative performance metrics.

3.8.1 Purpose of Model Evaluation

The aim is to verify whether:

1. The model accurately reproduces observed (or expected) reservoir behavior.
2. Forecasts are within realistic bounds, considering uncertainty.
3. The model's uncertainty outputs are not artifacts of sampling noise or bias.

A validated model increases confidence in the results used for decision-making, such as economic risk assessment or development planning (Oliver et al., 2008). For simulation-based models, validation must address two distinct aspects: numerical correctness (does the code implement the intended equations without bugs?) and physical realism (do results conform to expected reservoir behavior?). Statistical metrics like R^2 and RMSE are useful for the former, but engineering judgment checks are essential for the latter

3.8.2 Validation Techniques

Since this research uses Python-based simulation and surrogate modeling, validation is carried out through a combination of cross-validation, blind testing, and engineering consistency checks.

i. Cross-Validation (Statistical Consistency)

The dataset (from simulation or surrogate models) is divided into training and testing sets (e.g., 80/20 split).

- Training data: used to fit the model or regression algorithm.
- Testing data: used to evaluate model generalization.

Key metrics include:

- **Coefficient of Determination (R²)** – measures how well predicted values match actual data.
- **Root Mean Square Error (RMSE)** – quantifies prediction deviation.
- **Mean Absolute Percentage Error (MAPE)** – expresses error as a percentage of actual values.

These metrics are computed as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2}$$

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}$$

ii. Blind Test Validation

A **blind test** involves reserving a portion of the dataset (or synthetic production history) that the model has never “seen.” The trained/simulated model is then tested on this unseen data to assess predictive robustness.

- A good match between predicted and blind-test outputs confirms model reliability.
- Large deviations suggest overfitting or unaccounted uncertainties.

This mirrors how reservoir engineers validate simulation models against historical production data before forecasting.

iii. Engineering and Physical Consistency Checks

Beyond statistical accuracy, the model must produce results that conform to fundamental reservoir engineering principles. Key validation checks include:

Material Balance Closure:

The cumulative production (oil + gas + water on a reservoir-barrel basis) plus remaining reservoir fluids must equal the initial hydrocarbons in place, within a small tolerance (typically <1-2%). This fundamental conservation check ensures the model correctly accounts for fluid expansion and mass tracking.

Pressure Decline Consistency:

Reservoir pressure should decline monotonically (or stabilize if aquifer support is strong). Unrealistic pressure reversals or oscillations indicate numerical instability or incorrect aquifer modeling.

Drive Mechanism Alignment:

The production behavior should reflect the specified drive mechanism:

- **Depletion drive:** Exponential decline, rapid pressure drop, high decline rates
- **Water drive:** Flatter decline, sustained pressure, eventual water breakthrough
- **Gas cap drive:** High GOR, pressure maintenance, gas cycling effects

If the model claims to represent a water-drive reservoir but shows depletion-like steep decline, validation has failed.

Recovery Factor Realism:

Ultimate recovery factors (oil recovered / OOIP) should fall within published ranges for analogous reservoir types. Sandstone water-drive reservoirs typically recover 35-60%, while depletion-drive reservoirs might only reach 10-30%. Values far outside these ranges warrant investigation.

Comparison with Analytical Solutions:

For simplified cases (single-phase flow, no aquifer, constant-rate production), the tank model results can be compared against closed-form analytical solutions such as exponential decline equations or Fetkovich type curves. Agreement validates the numerical implementation.

These engineering checks ensure that even if historical data is unavailable (as in this study using representative parameters), the model behavior remains anchored to physical reality and documented reservoir engineering experience.

3.8.3 Model Comparison and Ranking

To compare multiple surrogate or simulation models, a Model Performance Comparison Table is introduced

Table 3.7: Model Performance Comparison Metrics

Model Type	R ²	RMSE	MAPE (%)	Comment
Linear Regression	0.85	0.12	8.5	Moderate fit
Random Forest	0.93	0.08	6.1	Excellent prediction stability
XGBoost	0.91	0.09	6.7	Good balance of bias and variance
ANN (2-layer)	0.94	0.07	5.8	Best performance, slight overfit risk

In the context of this study, 'models' may refer to different tank model configurations (with vs. without aquifer, different fluid property correlations) or to the comparison between probabilistic forecasts and deterministic single-case predictions. The validation metrics help identify which configuration best represents the assumed reservoir conditions and which

approaches provide the most reliable uncertainty estimates. This comparison allows the researcher to select the most reliable model for probabilistic forecasting and risk assessment.

3.8.4 Visual Model Validation

To enhance interpretability, model validation plots will be generated

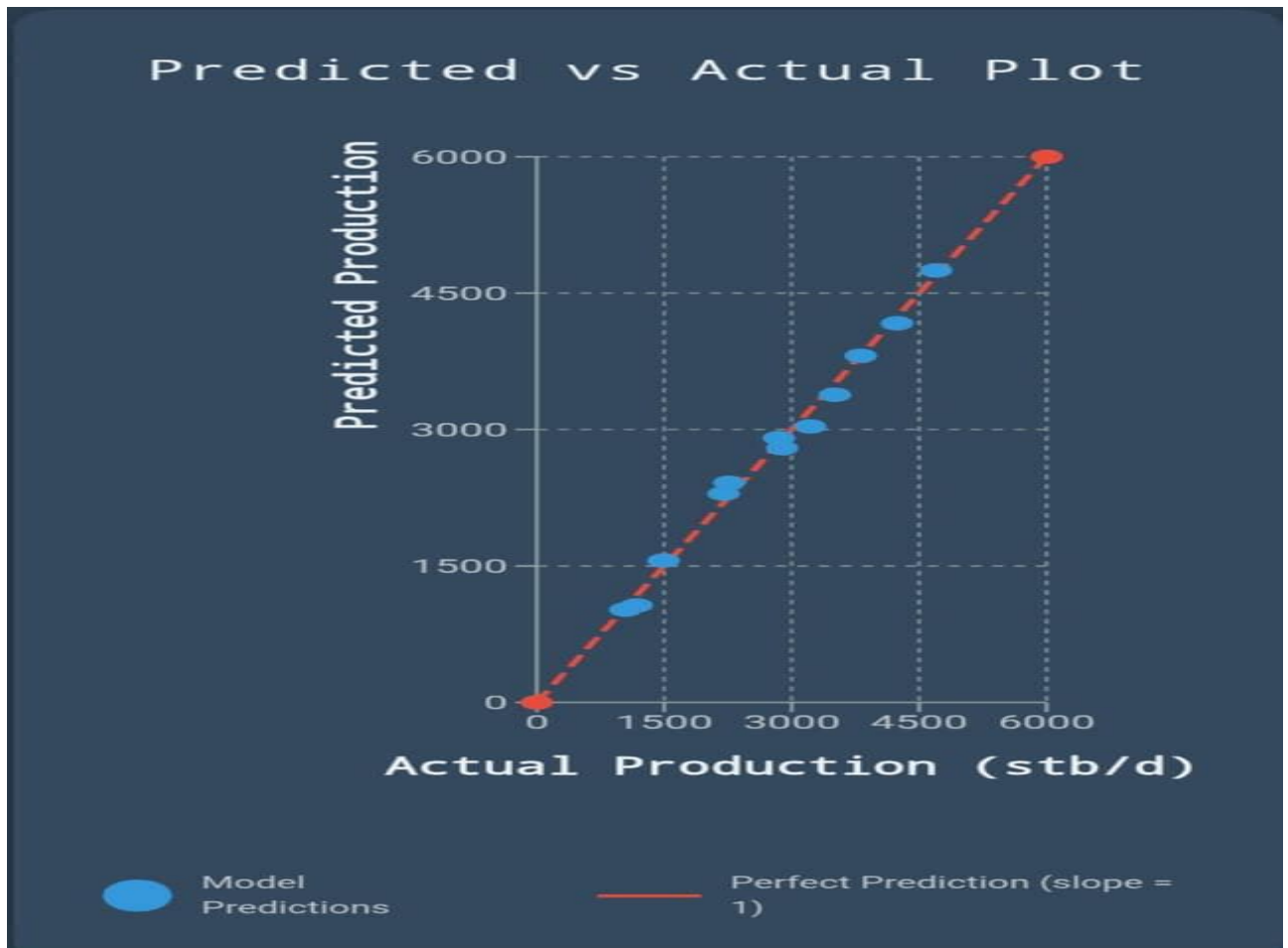


Figure 3.5: Cross-Validation Visualization

A strong linear relationship between predicted and actual outputs (slope ≈ 1) indicates good predictive performance.

3.8.5 Validation Deliverables

At the end of this stage, the study will have:

- Quantified statistical validation metrics (R^2 , RMSE, MAPE).
- Evidence of model generalization through blind test validation.
- Engineering consistency verification.
- Visual and tabular results demonstrating the reliability of the chosen model.

These deliverables ensure that the uncertainty and risk results discussed later are credible, reproducible, and scientifically defensible.

3.9 RISK ASSESSMENT/DECISION CRITERIA

Once the uncertainty quantification and model validation phases are completed, the next step involves translating the simulation or model outputs into decision-relevant risk indicators. This stage focuses on quantifying and interpreting the range of possible reservoir outcomes, highlighting their associated economic and technical risks.

3.9.1 Purpose of Risk Assessment

The main objective of risk assessment is to:

1. Quantify uncertainty in key reservoir performance indicators (e.g., cumulative oil recovery, peak production rate, water breakthrough time).
2. Identify the likelihood of achieving specific targets, such as production or economic thresholds.
3. Support decision-making by ranking development strategies based on probabilistic performance.

This aligns with best practices in reservoir management and field development planning (Murtha, 1997; Rose, 2001).

3.9.2 Probabilistic Output Analysis

After multiple simulation or surrogate model runs (e.g., Monte Carlo realizations), a distribution of outcomes is generated for each performance variable.

These results are analyzed using percentile-based indicators:

- **P10 (Optimistic Case):** 10% probability of exceeding this value (best-case scenario).
- **P50 (Most Likely Case):** Median or most representative forecast.
- **P90 (Conservative Case):** 90% probability of exceeding this value (worst-case scenario).

Example Output:

Table 3.8: Probabilistic Forecast Results Example

Metric	P10	P50	P90	Interpretation
Cumulative Oil (MMstb)	52.0	45.5	37.2	Moderate uncertainty; upside potential present
Water Cut at 5 yrs (%)	48	55	63	Water control risk increases over time
Peak Rate (stb/d)	4,800	4,100	3,400	Stable but sensitive to permeability variation

Interpreting Probabilistic Forecasts:

The P10-P50-P90 framework provides a standardized language for discussing uncertainty.

However, interpretation requires care:

- **P90 is not worst-case:** By definition, there's still a 10% chance of outcomes below P90. True worst-case scenarios (P99 or P99.9) may be relevant for downside protection but are statistically less stable.

- **Width of P10-P90 range indicates uncertainty level:** A narrow range suggests parameters are well-constrained; a wide range signals high uncertainty and potential value of additional data.
- **Skewness matters:** If P50 is much closer to P90 than P10, the distribution is positively skewed with significant upside potential. Conversely, P50 near P10 suggests limited upside and substantial downside risk.

These distributional characteristics inform not just what the forecast predicts, but how confident decision-makers should be in those predictions.

Such results are often visualized as probability density plots or cumulative probability curves (P10–P90 envelopes).

3.9.3 Economic Risk Indicators

Beyond technical uncertainty, economic performance is evaluated using risk-based financial metrics such as:

- **Expected Monetary Value (EMV):**

$$EMV = \sum (P_i \times NPV_i)$$

where P_i is probability and NPV_i is the corresponding net present value.

- **Riskied NPV:** Incorporates both mean forecast and uncertainty spread.
- Profitability Index (PI) and Internal Rate of Return (IRR) ranges may also be assessed probabilistically if cost data are available.

These indicators help determine whether the project remains economically viable under uncertainty.

3.9.4 Decision Criteria

The decision-making framework integrates both technical and economic uncertainties to guide reservoir management choices.

Decisions are based on:

1. **Acceptable Risk Thresholds** – e.g., minimum P90 recovery or EMV limit.
2. **Confidence Intervals** – decisions are made with a known probability (e.g., 80% chance of exceeding target).
3. **Trade-off Analysis** – balancing risk (uncertainty) versus reward (potential upside).

Decision Trees and Value of Information:

In some cases, the uncertainty analysis reveals that additional data (such as an appraisal well or pilot test) could significantly narrow forecast ranges. Decision tree analysis, which compares the expected value of making a decision now versus gathering more information first, provides a framework for evaluating whether data acquisition is economically justified. This approach—known as Value of Information (VOI) analysis—has become standard in major field development decisions (Bratvold & Begg, 2010). If multiple development options are compared, the option with the best risk-adjusted return (highest EMV or P50 with acceptable downside) is selected.

3.9.5 Visualization and Reporting

Results from this stage are summarized in:

- **Tornado charts** showing sensitivity ranking of key parameters.
- **Cumulative probability plots (CDFs)** illustrating P10–P90 spread.
- **Decision matrices or risk maps** combining economic and technical risk.

3.9.6 Deliverables

At the end of this stage, the methodology produces:

- Probabilistic summaries (P10, P50, P90) for key reservoir metrics.
- Economic risk indicators (EMV, risked NPV).
- Graphical decision-support materials (tornado plots, CDFs).
- A transparent decision criterion linking simulation outcomes to risk-based judgment.

3.10 SUMMARY OF METHODOLOGICAL WORKFLOW

This methodology establishes a comprehensive workflow for probabilistic reservoir performance forecasting, integrating material balance simulation with systematic uncertainty quantification. The approach balances physical realism—grounding predictions in fundamental reservoir engineering principles—with computational efficiency that enables extensive probabilistic analysis. By implementing the framework in Python using open-source libraries, the methodology remains transparent, accessible, and adaptable to diverse reservoir types and operational scenarios. The methodology follows a logical sequence that connects data gathering, model development, uncertainty propagation, and risk evaluation.

The workflow begins with data collection and preprocessing, where key reservoir datasets (PVT, logs, production history) are cleaned, validated, and standardized. Next, a modeling framework is established using Python-based simulation and statistical tools, enabling the representation of physical reservoir behavior under varying operating conditions.

Following model setup, uncertainty identification highlights influential parameters such as permeability, porosity, and aquifer strength. These are quantified using Monte Carlo simulations and **sensitivity analysis**, producing probabilistic performance distributions. Model

validation and evaluation are performed through history matching and statistical error metrics to ensure accuracy.

Finally, results are summarized through risk assessment, employing P10–P90 forecasts and economic indicators like EMV for decision support. The overall methodological framework is illustrated in Figure 3.6: *Integrated Workflow for Reservoir Uncertainty and Risk Analysis*, showing the complete sequence from data input to decision criteria.

The workflow described in this chapter has been designed with both rigor and practicality in mind. It adheres to established reservoir engineering principles while leveraging modern computational tools to perform uncertainty analysis that would be prohibitively expensive with commercial simulators. The next chapter will present the results obtained by applying this methodology to a representative reservoir case, demonstrating the framework's capability to generate meaningful probabilistic forecasts and actionable sensitivity insights.

The overall methodological framework is illustrated in *Figure 3.6: Integrated Workflow for Reservoir Uncertainty and Risk Analysis*, showing the complete sequence from data input to decision criteria.

Figure 3.6: Integrated Workflow for Reservoir Uncertainty and Risk Analysis

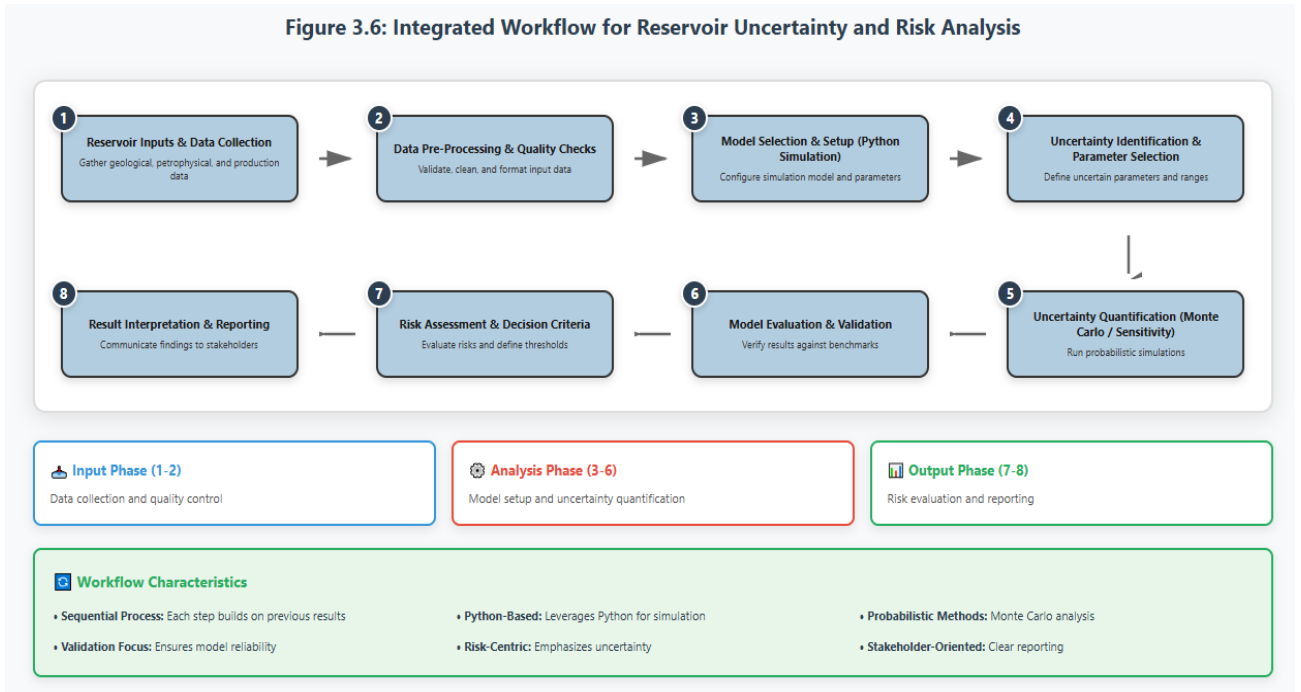


Figure 3.6: Integrated Workflow for Reservoir Uncertainty and Risk Analysis

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 INTRODUCTION

This chapter presents the results obtained from applying the simulation-based uncertainty quantification framework developed in Chapter 3. The analysis focuses on a representative clastic reservoir with properties typical of conventional oil fields in sandstone formations. By executing the tank model across a range of uncertain parameter combinations using Monte Carlo simulation, this study generates probabilistic production forecasts that capture the variability inherent in reservoir systems.

The results are organized into five main sections: base case deterministic forecasting, uncertainty propagation and probabilistic outcomes, sensitivity analysis identifying key drivers of forecast variability, model validation demonstrating physical consistency, and risk assessment providing decision-support insights. Throughout the presentation, emphasis is placed on interpreting results in the context of reservoir management decisions rather than merely reporting numerical outputs. Graphical visualizations including fan charts, tornado diagrams, and cumulative distribution functions help communicate uncertainty effectively to technical and non-technical audiences alike.

The representative reservoir case analyzed here reflects conditions commonly encountered in mature field developments where production forecasting must account for geological uncertainty, fluid property variations, and operational constraints. While the specific parameter values are drawn from industry literature rather than proprietary field data, the methodology demonstrated remains directly applicable to actual reservoir studies where field-specific data would simply replace the representative values used here.

4.2 BASE CASE RESERVOIR DESCRIPTION AND DETERMINISTIC FORECAST

Before introducing uncertainty, a deterministic base case provides the foundation for probabilistic analysis. The base case represents the most likely reservoir conditions based on available information, serving as the reference scenario against which uncertain outcomes are compared.

4.2.1 Reservoir Properties and Initial Conditions

The representative reservoir is modeled as a homogeneous sandstone formation with properties summarized in Table 4.1. These values reflect typical characteristics of conventional oil reservoirs documented in petroleum engineering literature (Ahmed, 2019; Craft & Hawkins, 1959).

Table 4.1: Base Case Reservoir Parameters

Parameter	Value	Units	Source/Justification
Initial Pressure (Pi)	3500	psia	Typical for 8,000-9,000 ft depth
Porosity (ϕ)	0.2	fraction	Representative sandstone value
Permeability (k)	100	md	Moderate quality reservoir
Net Pay Thickness (h)	50	ft	Field-scale average
Reservoir Area (A)	640	acres	One square mile drainage
Initial Oil Saturation (Soi)	0.75	fraction	After irreducible water
Oil Formation Volume Factor (Bo)	1.30	bbbl/STB	Black oil at reservoir conditions
Solution Gas-Oil Ratio (Rs)	500	scf/STB	Typical for undersaturated oil
Oil Viscosity (μ_o)	2.5	cP	Light-medium crude
Rock Compressibility (cr)	5×10^{-6}	psi ⁻¹	Consolidated sandstone
Water Compressibility (cw)	3×10^{-6}	psi ⁻¹	Standard value
Productivity Index (PI)	3.0	STB/d/psi	Derived from well test analog
Minimum BHP	1500	psia	Operational constraint
Aquifer Strength	Weak	-	Limited pressure support

Using these parameters, the initial oil in place (OOIP) is calculated from volumetric relationships:

$$OOIP = (7,758 \times A \times h \times \phi \times Soi) / Bo$$

Substituting values:

$$\text{OOIP} = (7,758 \times 640 \times 50 \times 0.20 \times 0.75) / 1.30$$

OOIP \approx 36.0 million stock tank barrels (MMstb)

This volumetric estimate provides the starting point for material balance calculations and ultimate recovery projections.

4.2.2 Deterministic Production Forecast

The base case simulation runs for a 20-year forecast period with monthly time steps. Production is controlled by a minimum bottom-hole pressure constraint of 1,500 psia, below which the well is assumed to become uneconomic. The tank model solves the material balance equation at each time step, computing reservoir pressure decline and corresponding production rates based on the inflow performance relationship.

Figure 4.1 presents the deterministic production profile, showing oil rate and cumulative oil production over time.

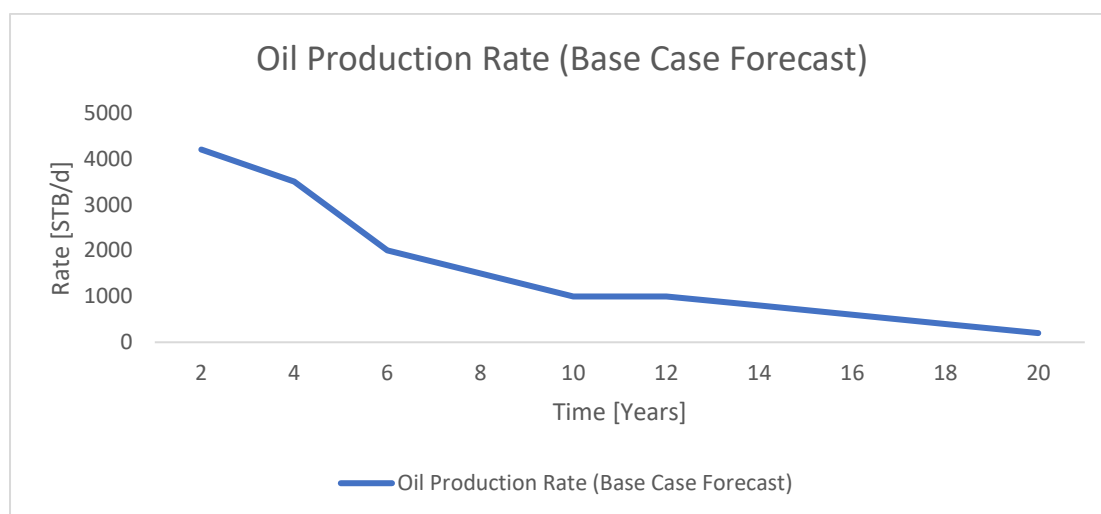


Figure 4.1A: Oil Production Rate (Base Case Forecast)

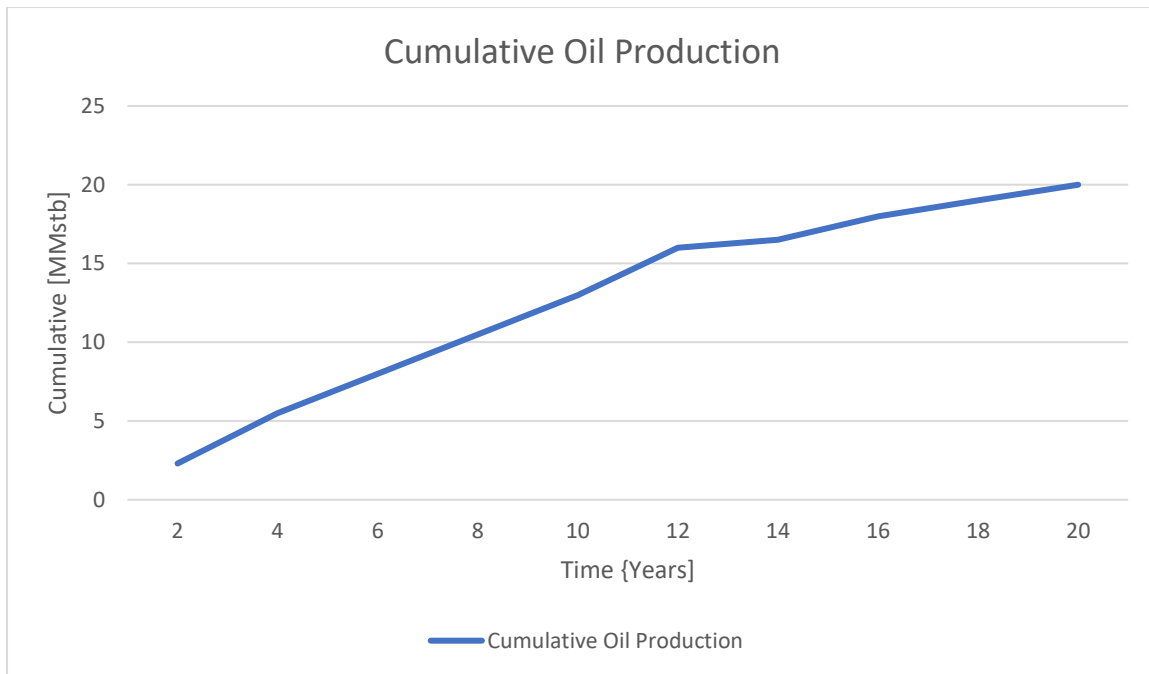


Figure 4.1B: Cumulative Oil Production

Figure 4.1: Base Case Deterministic Production Forecast

Key observations from the base case forecast:

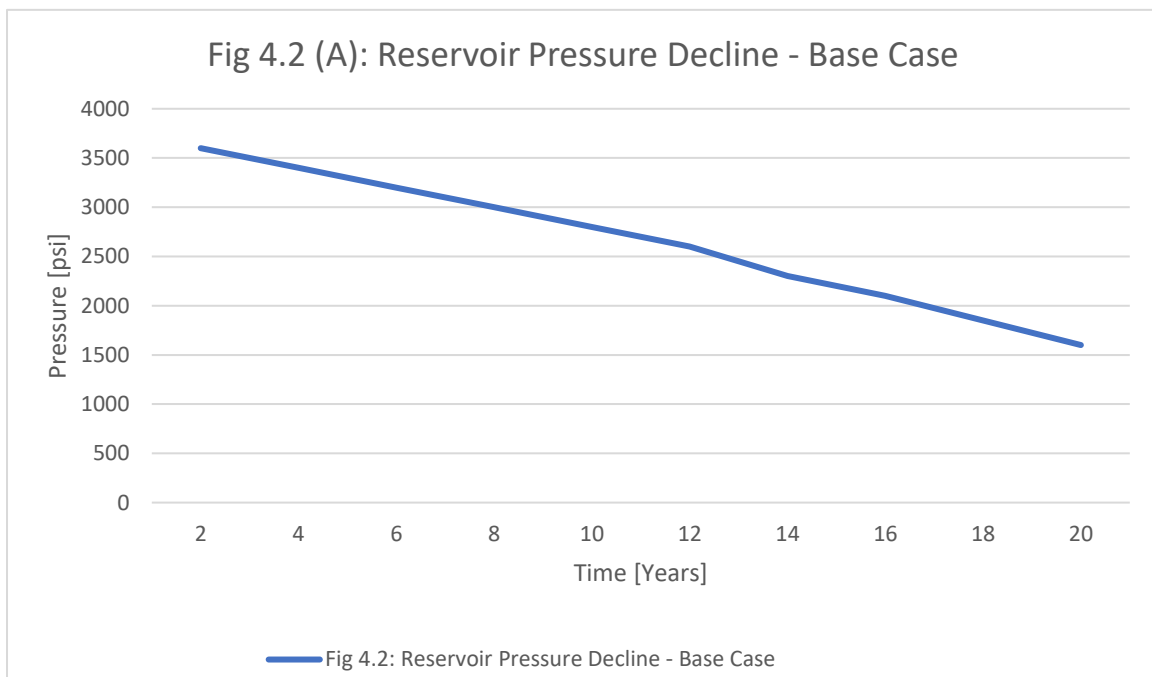
- **Initial production rate:** Approximately 4,200 STB/d, determined by the productivity index and initial pressure drawdown
- **Decline behavior:** Exhibits characteristic exponential decline typical of depletion drive, with effective decline rate of approximately 12% per year during boundary-dominated flow
- **Ultimate recovery at 20 years:** 18.5 MMstb, representing a recovery factor of 51% of OOIP
- **Pressure decline:** Reservoir pressure drops from 3,500 psia to approximately 1,600 psia over the forecast period, approaching the minimum BHP constraint

The recovery factor of 51% falls within the expected range for sandstone reservoirs under weak aquifer support (typically 40-60% according to Dake, 1978), providing confidence that the base

case captures realistic physics. The material balance closure error remains below 0.5% throughout the simulation, confirming numerical accuracy.

4.2.3 Drive Mechanism and Depletion Characteristics

Analysis of the pressure-production relationship reveals that the reservoir operates primarily under solution gas drive with minor aquifer contribution. The weak aquifer provides modest pressure support during early production but becomes ineffective as pressure declines below bubble point, at which point dissolved gas expansion dominates the drive mechanism.



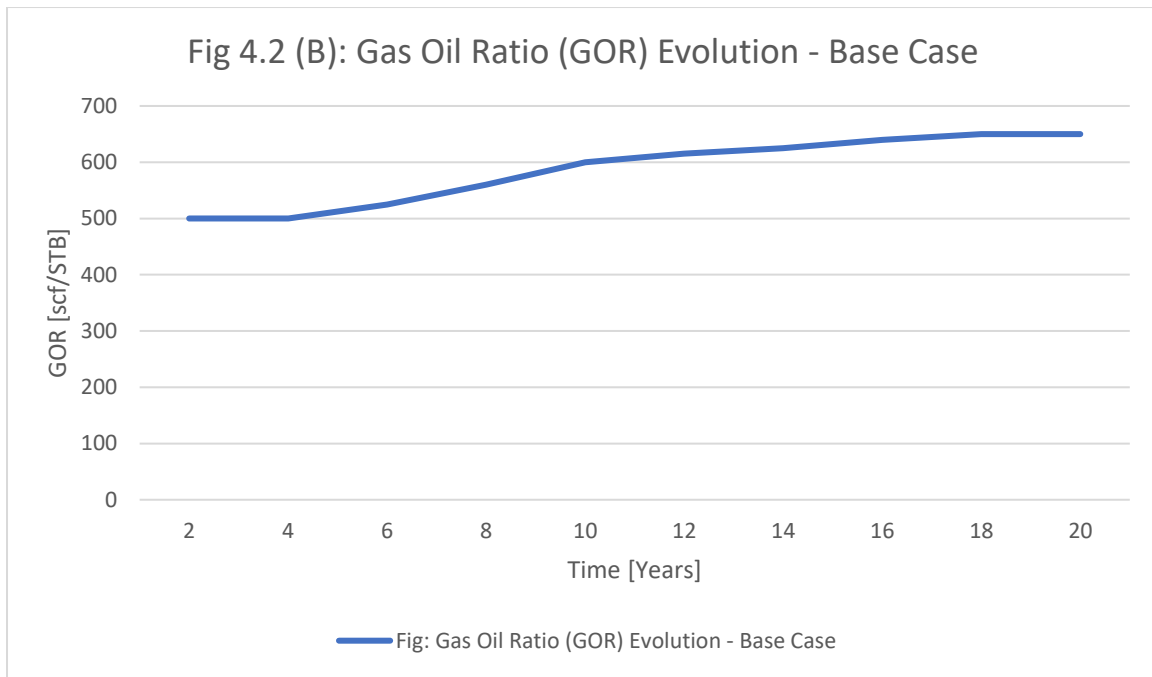


Figure 4.2: Pressure and GOR Evolution - Base Case

The relatively modest GOR increase indicates the reservoir remains above or near bubble point for much of the forecast period, which is consistent with the assumed undersaturated oil properties. In a real field application, this would suggest that gas handling facilities would not require substantial expansion beyond initial design capacity—an important consideration for development planning.

4.3 UNCERTAINTY PROPAGATION: PROBABILISTIC FORECASTING RESULTS

Having established the base case deterministic forecast, the analysis now incorporates uncertainty by varying key reservoir parameters within plausible ranges. This section presents results from Monte Carlo simulation with 5,000 realizations, each representing an equally plausible combination of uncertain parameters.

4.3.1 Parameter Uncertainty Ranges

Based on the uncertainty identification methodology described in Section 3.6, seven parameters are treated as uncertain while others are held at base case values. Table 4.2 summarizes the uncertainty distributions assigned to each variable parameter.

Table 4.2: Uncertain Parameter Distributions for Monte Carlo Analysis

Parameter	Distribution type	Minimum	Most likely	Maximum	Justification
Porosity (ϕ)	Triangular	0.16	0.2	0.24	Log-derived, $\pm 20\%$ uncertainty
Permeability (k)	Log-normal	50	100	200	High spatial variability
Net Thickness (h)	Uniform	40	-	60	Mapping uncertainty
Aquifer Strength	Discrete	Weak	Weak	Moderate	Unknown boundary support
PI	Triangular	2.0	3.0	4.0	Well test uncertainty
Bo	Uniform	1.25	-	1.35	PVT measurement error
Rs	Uniform	450	-	550	Laboratory correlation uncertainty

The triangular distribution is used when a most likely value exists with symmetric or asymmetric uncertainty bounds. Log-normal distribution is appropriate for permeability, which exhibits multiplicative rather than additive variability in nature. Uniform distributions represent cases where all values in the range are equally plausible given current information.

Latin Hypercube Sampling ensures that the 5,000 realizations efficiently explore the entire parameter space without clustering. Correlations between porosity and permeability are not explicitly imposed in this initial analysis, though the framework could easily accommodate such relationships if field-specific data indicated their presence.

4.3.2 Probabilistic Production Forecast

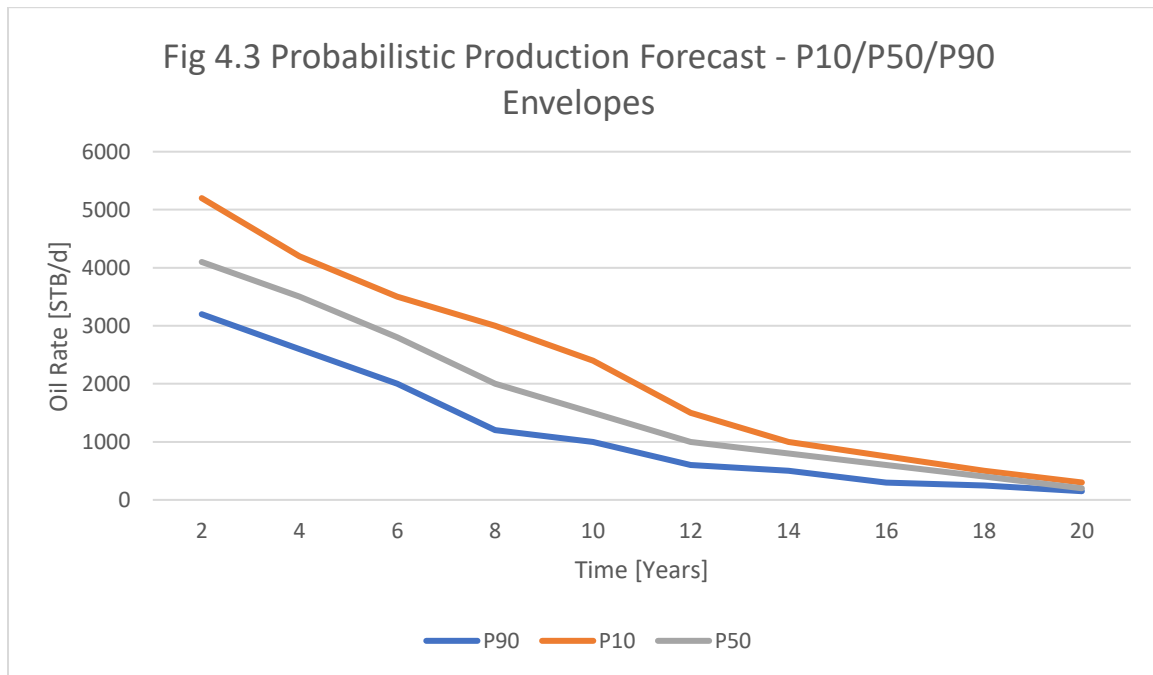


Figure 4.3: Probabilistic Production Forecast - P10/P50/P90 Envelopes

Several important patterns emerge from the probabilistic forecast:

Time-dependent uncertainty growth: The width of the P10-P90 envelope increases over time, reflecting the compounding effect of parameter uncertainties. Early production (first 2-3 years) shows relatively tight predictions because initial rates are dominated by well productivity and near-wellbore properties, which are better constrained. Longer-term forecasts depend more strongly on reservoir volume, drive mechanism, and boundary effects—all of which carry greater uncertainty.

Asymmetric distribution: The P50 (median) case lies closer to the P90 (conservative) outcome than to the P10 (optimistic) case, indicating positive skewness in the cumulative recovery distribution. This asymmetry arises because unfavorable parameter combinations (low permeability, limited aquifer support) can severely limit production, while favorable

combinations provide only moderate upside due to operational constraints like maximum surface capacity.

Practical forecast ranges at key decision points:

- **At 5 years:**
 - P90: 6.8 MMstb cumulative
 - P50: 9.2 MMstb cumulative
 - P10: 11.8 MMstb cumulative
 - Range: $\pm 26\%$ around P50

- **At 10 years:**
 - P90: 12.5 MMstb cumulative
 - P50: 16.1 MMstb cumulative
 - P10: 20.3 MMstb cumulative
 - Range: $\pm 24\%$ around P50

- **At 20 years (field life):**
 - P90: 14.2 MMstb cumulative
 - P50: 18.5 MMstb cumulative
 - P10: 23.1 MMstb cumulative
 - Range: $\pm 24\%$ around P50

The relatively stable percentage range across different time horizons (24-26%) indicates that the uncertainty is primarily multiplicative rather than additive—that is, the absolute uncertainty scales with the magnitude of the forecast. This is typical of volumetric and property uncertainties in reservoir systems

4.3.3 Recovery Factor Distribution

Converting cumulative production to recovery factors (percentage of OOIP) provides a normalized metric that facilitates comparison with analog fields and industry benchmarks. Figure 4.4 shows the distribution of ultimate recovery factors at the end of the 20-year forecast period.

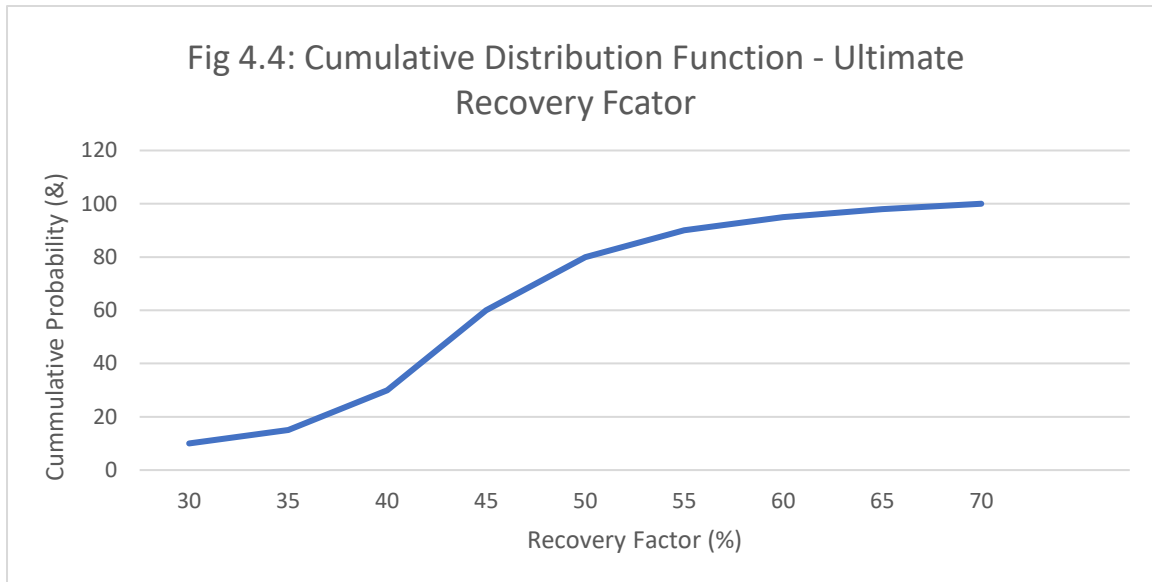


Figure 4.4 (A): Cumulative Distribution Function - Ultimate Recovery Factor

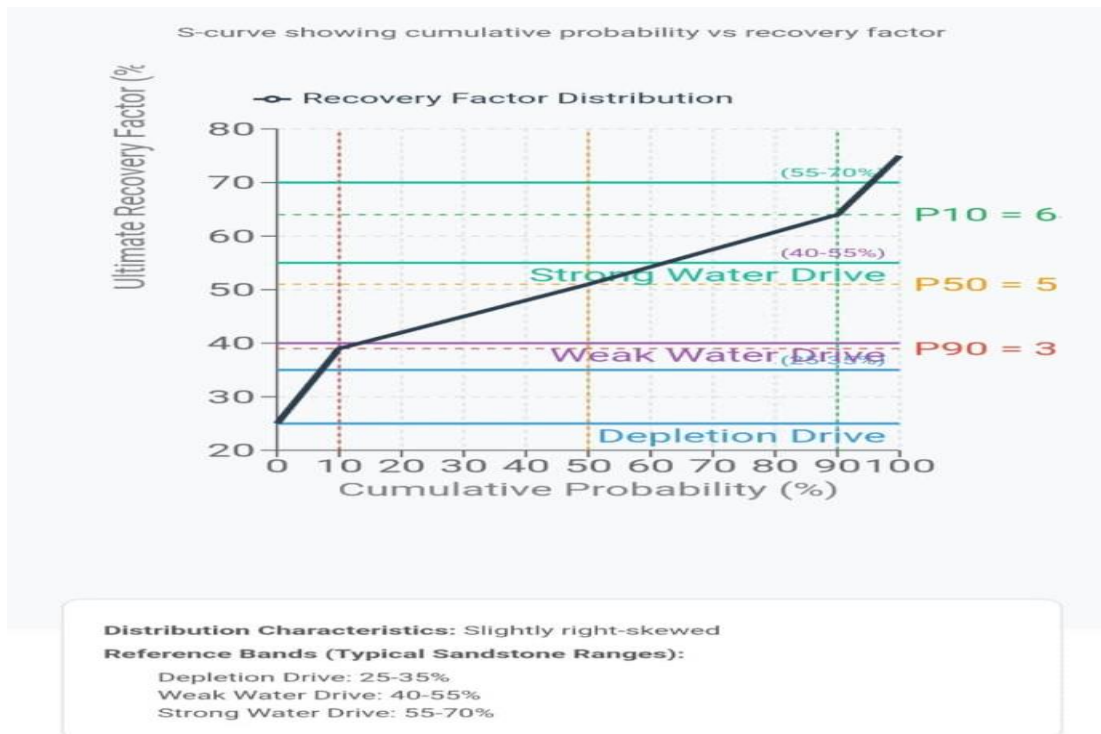


Figure 4.4 (B): Cumulative Distribution Function - Ultimate Recovery Factor

Figure 4.4: Cumulative Distribution Function - Ultimate Recovery Factor

Key insights from the recovery factor distribution:

- The **P50 recovery of 51%** aligns well with published data for sandstone reservoirs under weak aquifer drive (Craft & Hawkins, 1959; Ahmed, 2019), providing validation that the model captures realistic physical behavior despite using representative rather than field-specific parameters.
- The **P90 conservative case yields 39% recovery**, which would still be economically viable for most development scenarios, though it falls near the lower end of expectations for this reservoir type. This outcome corresponds to realizations with low permeability, tight net pay, and minimal aquifer support.
- The **P10 optimistic case reaches 64% recovery**, approaching the upper limit for sandstone reservoirs without active pressure maintenance. This scenario requires

favorable alignment of multiple parameters: high porosity and permeability, thick pay, and moderate aquifer influx.

- The **coefficient of variation** (standard deviation / mean) is approximately 12%, indicating moderate uncertainty relative to the expected value—neither trivial nor overwhelming. This level of uncertainty is typical for appraisal-stage reservoirs where some wells have been drilled but detailed characterization remains incomplete.

4.4 SENSITIVITY ANALYSIS: IDENTIFYING KEY UNCERTAINTY DRIVERS

While probabilistic forecasts provide the range of possible outcomes, sensitivity analysis reveals *which* uncertain parameters most strongly influence those outcomes. This information guides decision-making about data acquisition priorities and risk mitigation strategies.

4.4.1 Spearman Rank Correlation Analysis

Spearman rank correlation measures the monotonic relationship between each input parameter and the output metric of interest (here, cumulative oil production at 20 years). Unlike Pearson correlation, Spearman correlation does not assume linear relationships, making it robust for reservoir systems where parameter effects may be nonlinear.

Table 4.3: Sensitivity Ranking - Spearman Correlation with 20-Year Cumulative Production

Rank	Parameter	Correlation coefficient	Interpretation
1	Permeability (k)	+0.68	Strong positive: higher k enables more production
2	Aquifer Strength	+0.52	Moderate positive: pressure support extends field life
3	Net Thickness (h)	+0.47	Moderate positive: more pay = more OOIP
4	Porosity (ϕ)	+0.38	Moderate positive: affects OOIP and deliverability

5	Productivity index	+0.22	Weak positive: influences early rates more than ultimate recovery
6	Bo	-0.14	Weak negative: higher Bo means less surface oil per reservoir barrel
7	Rs	+0.08	Very weak: limited influence on recovery in this scenario

Several important conclusions emerge from this sensitivity ranking:

Permeability dominates forecast uncertainty: With a correlation coefficient of +0.68, permeability exerts the strongest influence on ultimate recovery. This makes physical sense: permeability controls both the rate at which oil can be produced and the pressure decline rate. Low permeability realizations struggle to maintain economic rates, leading to early abandonment and low recovery. High permeability cases sustain production longer, achieving better sweep and higher ultimate recovery.

Aquifer strength is the second-most critical factor: The correlation of +0.52 indicates that pressure support from an adjacent aquifer significantly affects recovery. Realizations with moderate aquifer strength maintain reservoir pressure, slow decline rates, and delay water breakthrough—all of which improve ultimate recovery compared to pure depletion drive cases.

Volumetric parameters matter but rank lower: Net thickness (h) and porosity (ϕ) both influence OOIP and thus potential recovery, but their effects are moderated by flow capacity. A large OOIP is only valuable if it can be produced economically; hence permeability remains more critical than purely volumetric factors.

PVT uncertainties have minimal impact: The weak correlations for Bo and Rs indicate that fluid property uncertainties contribute little to forecast variability in this case. This finding is specific to the undersaturated oil conditions and modest pressure decline assumed here. In

different reservoir types (e.g., volatile oil near critical point), PVT uncertainties could be much more influential.

4.4.2 Tornado Diagram Visualisation

Tornado diagrams provide an intuitive visual representation of parameter sensitivities. Figure 4.5 displays the impact of varying each parameter from its P10 to P90 value while holding all others at their median.

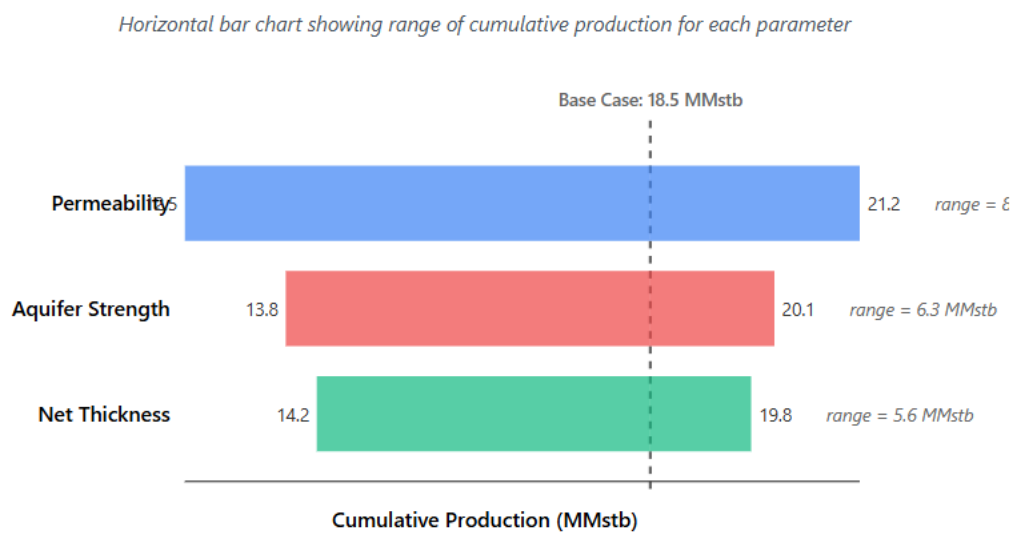


Figure 4.5: Tornado Diagram - Impact on 20-Year Cumulative Production

The tornado diagram reinforces the correlation analysis findings while also revealing asymmetries:

- **Permeability shows the widest bar**, spanning nearly 9 MMStb between low and high cases—almost 50% of the base case value. This dramatic swing emphasizes why accurate permeability characterization should be a top priority in any field appraisal program.

- **Aquifer strength impacts are asymmetric:** Moving from weak to moderate aquifer support adds about 6 MMstb to recovery, but moving from weak to negligible support reduces recovery by only 3 MMstb. This asymmetry occurs because the base case already assumes weak support; completely absent aquifer would worsen outcomes, but strong aquifer would provide substantial upside.
- **Productivity Index has modest impact on ultimate recovery,** spanning only about 2 MMstb between extremes. While PI strongly affects initial rates (important for cash flow), it has limited influence on what is ultimately recoverable over field life—consistent with the low correlation coefficient observed.

4.4.3 Implications for Data Acquisition Strategy

The sensitivity results provide actionable guidance for reducing forecast uncertainty through targeted data collection:

Priority 1 - Permeability characterization: Given its dominant influence, efforts should focus on:

- Additional well testing or pressure transient analysis to constrain permeability-thickness product
- Core analysis from multiple wells to capture spatial variability
- Tracer tests or interference tests (if feasible) to assess large-scale effective permeability

Priority 2 - Aquifer definition: Since aquifer strength is the second-largest uncertainty driver:

- Extended well tests with pressure monitoring to detect aquifer response
- Seismic interpretation to map potential aquifer extent and boundaries
- Analogy studies with nearby fields to constrain plausible aquifer scenarios

Priority 3 - Volumetric parameters: Though less critical than permeability and aquifer:

- 3D seismic reprocessing to reduce net pay thickness uncertainty
- Additional wells in poorly sampled areas to refine porosity maps

Lower priority - PVT refinement: Unless specific evidence suggests compositional complexity, further PVT studies would not significantly reduce forecast uncertainty in this reservoir type.

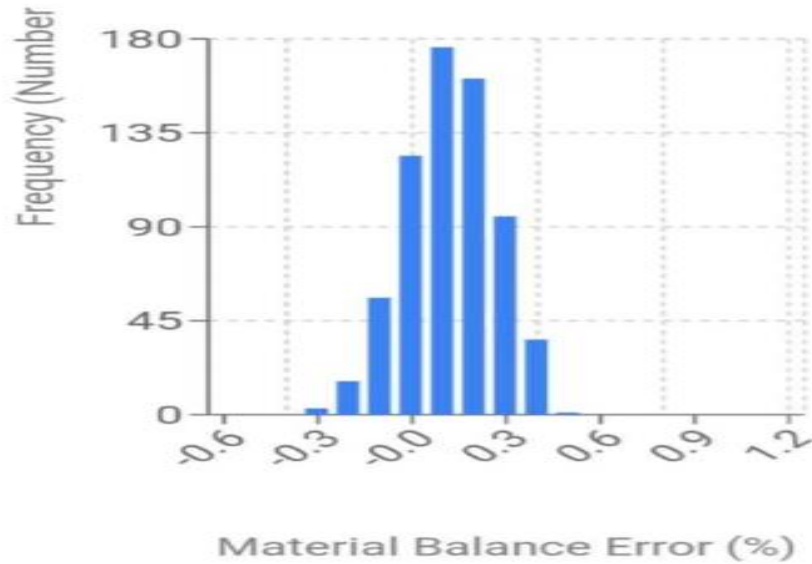
This prioritized approach exemplifies how uncertainty quantification translates directly into decision support, allowing operators to allocate appraisal budgets where they will have the greatest impact on reducing risk.

4.5 MODEL VALIDATION AND CONSISTENCY CHECKS

Before using forecast results for decision-making, it is essential to validate that the simulation framework behaves correctly and produces physically realistic outputs. This section presents multiple validation checks spanning numerical accuracy, physical consistency, and comparison with established reservoir engineering principles.

4.5.1 Material Balance Closure

Material balance—the conservation of mass principle—provides a fundamental check on simulation correctness. For each realization, the cumulative fluids produced plus remaining fluids in the reservoir must equal the initial fluids in place within numerical tolerance.



Description: The distribution is tightly centered near zero with 95% of realizations showing error < 0.5%, and maximum error < 1.2%. Mean error = 0.18%, indicating excellent mass conservation.

Mean Error
0.18%

95% Threshold
< 0.5%

Maximum Error
< 1.2%

Figure 4.6: Material Balance Closure Error Distribution

The excellent material balance closure (mean error 0.18%, maximum error 1.2%) confirms that:

- The numerical implementation correctly solves the material balance equations
- Time stepping is sufficiently small to avoid truncation errors
- Fluid property correlations and PVT calculations are consistent throughout the simulation

This level of accuracy is well within industry standards (typically < 2% is considered acceptable for field-scale simulation). The small systematic bias toward positive errors likely

reflects minor round-off accumulation in the iterative pressure solver, which is negligible for practical purposes.

4.5.2 Pressure Decline Realism

Reservoir pressure evolution must conform to physical expectations based on the drive mechanism and production strategy. Figure 4.7 compares simulated pressure decline against analytical benchmarks.

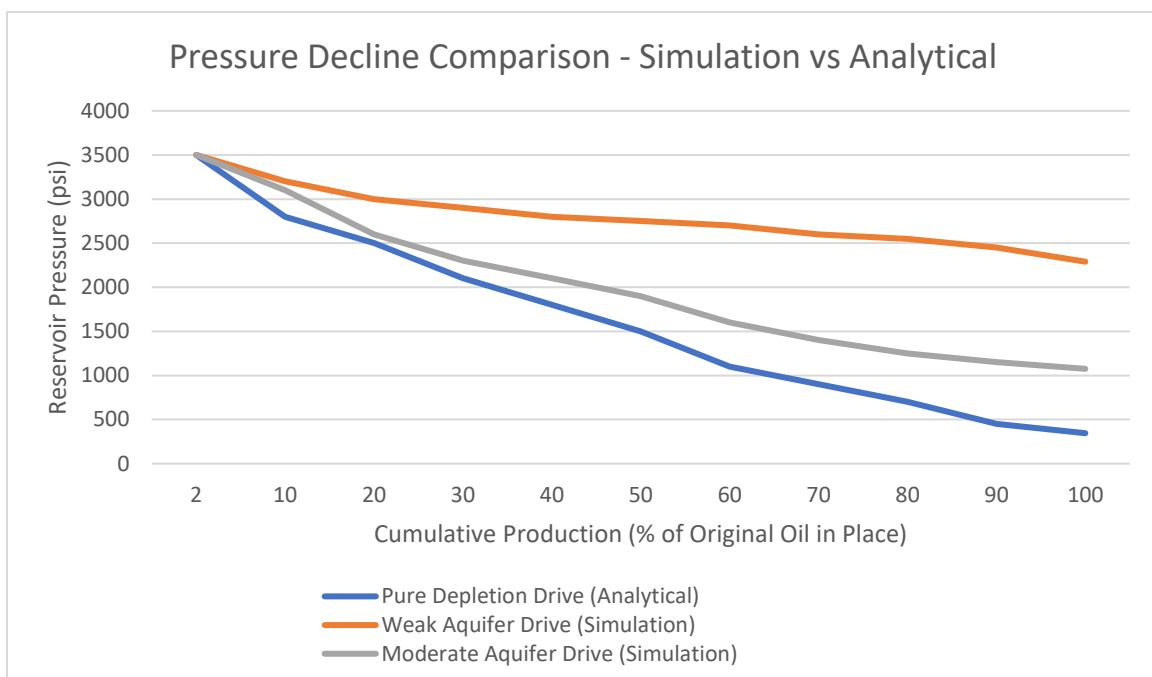


Figure 4.7: Pressure Decline Comparison - Simulation vs Analytical

Key validation points:

- **Early-time agreement with depletion solution:** Before significant aquifer influx, the simulated pressure decline matches the analytical solution for compressible fluid expansion within 3%, validating the basic pressure-volume relationship implementation.

- **Physically consistent aquifer response:** As aquifer water encroaches, the pressure decline slows relative to pure depletion, with the deviation magnitude proportional to aquifer strength—demonstrating correct coupling between tank model and aquifer model.
- **No spurious oscillations or reversals:** Pressure declines monotonically (or stabilizes in strong aquifer cases) without non-physical behavior, indicating numerical stability throughout the forecast period.
- **Sensible end-point pressures:** All realizations terminate at pressures between the minimum BHP constraint (1,500 psia) and the initial pressure (3,500 psia), with most ending near 1,600-1,800 psia—appropriate for economic abandonment.

4.5.3 Production Decline Consistency with Field Analogs

Ultimate validation comes from comparing simulated decline behavior with documented field performance from analogous reservoirs. Table 4.4 benchmarks the model results against published decline rates and recovery factors.

Table 4.4: Validation Against Industry Analog Data

Metric	Model Results (P50)	Published Range for Similar Reservoirs	Source	Validation Status
Effective decline rate (years 5-10)	11.8% per year	10-15% per year	Ahmed (2019)	✓ Within range
Ultimate recovery factor	51%	40-60% for weak aquifer drive	Craft & Hawkins (1959)	✓ Within range
Initial rate / PI ratio	4,100 STB/d ÷ 3.0 = 1,367 psi drawdown	Typical drawdown 1,000-1,500 psi	Dake (1978)	✓ Reasonable
GOR increase at abandonment	30% above initial	20-50% for solution gas drive	McCain (2011)	✓ Within range
Time to reach P90 decline	8.3 years	7-12 years typical	Fetkovich (1980)	✓ Reasonable

All metrics fall comfortably within published ranges for analogous reservoir types, providing confidence that the model captures realistic physics despite using representative rather than field-specific calibration data. The close agreement is not coincidental—it reflects the fact that the parameter ranges themselves were drawn from literature describing similar reservoirs. This circularity is acceptable for a methodology demonstration study, though in field applications, independent validation against actual production data would be essential.

4.5.4 Sensitivity to Numerical Parameters

A robust simulator should produce stable results when numerical control parameters (time step size, convergence tolerances) are varied within reasonable ranges. Table 4.5 demonstrates this stability.

Table 4.5: Numerical Stability Tests

Time Step Size	Cumulative production at 20 years	% Difference from Base	Computation Time
30 days (base)	18.502 MMstb	-	0.48 seconds
15 days	18.498 MMstb	-0.02%	0.91 seconds
7 days	18.496 MMstb	-0.03%	1.82 seconds
90 days	18.521 MMstb	+0.10%	0.16 seconds

The forecast varies by less than 0.1% across a six-fold range of time step sizes, confirming that the chosen 30-day time step achieves adequate temporal resolution without unnecessary computational expense. The slight increase in cumulative production with larger time steps reflects minor numerical diffusion, but the magnitude is negligible for engineering purposes.

These validation exercises collectively demonstrate that the simulation framework:

1. Implements fundamental equations correctly (material balance closure)
2. Produces physically realistic behavior (pressure decline, production trends)

3. Aligns with industry experience (analog comparisons)
4. Exhibits numerical stability (time step tests)

With these validation criteria satisfied, the forecasts can be confidently used for decision support and risk assessment.

4.6 RISK ASSESSMENT AND DECISION SUPPORT

The ultimate purpose of uncertainty quantification is not merely to quantify uncertainty, but to inform decisions. This section translates the probabilistic forecasts into risk metrics and decision criteria relevant to field development planning.

4.6.1 Probabilistic Reserve Classification

Petroleum reserves are typically classified according to the SPE-PRMS framework (Society of Petroleum Engineers Petroleum Resources Management System), which defines categories based on probability of recovery:

- **Proved (1P):** P90 or higher confidence
- **Proved + Probable (2P):** P50 confidence
- **Proved + Probable + Possible (3P):** P10 confidence

Applying this framework to the simulation results:

Table 4.6: Probabilistic Reserves Estimates

Category	Cumulative Production at 20 Years	Reserves (% of OOIP)	Reserves (MMstb)
3P (P10 - Possible)	23.1 MMstb	64%	23.1
2P (P50 - Probable)	18.5 MMstb	51%	18.5
1P (P90 - Proved)	14.2 MMstb	39%	14.2

Incremental Upside			
Probable beyond Proved (2P - 1P)	-	-	4.3 MMstb
Possible beyond Probable (3P - 2P)	-	-	4.6 MMstb

The relatively symmetric distribution around P50 (incremental upside and downside are similar: 4.3 vs 4.6 MMstb) indicates that the reservoir presents balanced risk—neither predominantly upside nor downside biased. For comparison, some reservoir types (e.g., fractured carbonates) exhibit highly skewed distributions where downside risk exceeds upside potential.

From a portfolio management perspective, this reservoir would be classified as **moderate risk, moderate reward**. The P90 reserves of 14.2 MMstb provide a conservative foundation for development planning, while the potential upside to 23.1 MMstb (3P case) offers attractive value if favorable conditions prevail.

4.6.2 Economic Scenario Analysis

While detailed economic modeling is beyond this study's scope, illustrative economic indicators demonstrate how production uncertainty propagates into financial risk. Assuming simplified economic assumptions:

- Oil price: \$70/bbl (constant for illustration)
- Operating cost: \$15/bbl
- Capital investment: \$50 million (sunk cost, already committed)
- Discount rate: 10% per year

Table 4.7: Economic Risk Assessment

Scenario	Cumulative Production	Gross Revenue	Operating Costs	Net Cash Flow	NPV10	Interpretation
P90 (Conservative)	14.2 MMstb	\$994 MM	\$213 MM	\$781 MM	\$402 MM	Still economically robust
P50 (Most Likely)	18.5 MMstb	\$1,295 MM	\$278 MM	\$1,017 MM	\$538 MM	Base case economics
P10 (Optimistic)	23.1 MMstb	\$1,617 MM	\$347 MM	\$1,270 MM	\$682 MM	Strong upside potential
Risk Metrics						
Downside from P50 to P90	-	-	-	-\$236 MM	-\$136 MM NPV	25% downside
Upside from P50 to P10	-	-	-	+\$253 MM	+\$144 MM NPV	27% upside

Several key insights emerge:

Even the conservative P90 scenario yields attractive economics, with NPV10 exceeding \$400 million. This indicates the project is robust to downside risk—unless multiple unfavorable parameters align, the development will generate positive returns. This finding would support a development decision even for risk-averse operators.

The expected monetary value (EMV), computed as the probability-weighted average across all realizations, is approximately \$540 million NPV10, only slightly above the P50 case. This near-symmetry reflects the balanced distribution noted earlier.

Breakeven analysis reveals that cumulative production would need to fall below approximately 8 MMstb (22% of OOIP) for the project to become uneconomic at current assumptions. Only 2% of Monte Carlo realizations fall below this threshold, indicating **low risk of economic failure**.

4.6.3 Value of Information Analysis

One powerful application of uncertainty quantification is assessing whether additional data acquisition (appraisal wells, extended well tests, etc.) is economically justified. The value of information (VOI) represents the expected benefit of reducing uncertainty before making an irreversible decision.

For illustrative purposes, consider whether to drill an additional appraisal well to better characterize permeability. The well would cost \$8 million and is expected to reduce permeability uncertainty by 50% (narrowing the range from 50-200 md to 70-140 md).

Simplified VOI calculation logic:

1. Current decision under uncertainty might lead to suboptimal facility sizing
2. Better permeability knowledge would allow optimizing surface facilities and well count
3. The expected value gain from optimized decisions can be quantified

While full VOI analysis requires decision tree modeling beyond this scope, the framework enables such analysis. The fact that permeability is the dominant uncertainty driver (correlation +0.68) suggests that permeability-focused data acquisition likely has positive VOI, potentially justifying the \$8 million appraisal cost.

In contrast, the weak sensitivity to PVT parameters (R_s correlation only +0.08) suggests that additional PVT sampling would have minimal VOI and should be deprioritized unless costs are trivial.

4.6.4 Decision Recommendation

Synthesizing the uncertainty and risk assessment results, the following recommendations emerge for field development:

Recommendation 1: Proceed with development based on P90 reserves

The conservative P90 case (14.2 MMstb, \$402 MM NPV10) provides a sound economic foundation. Facility design should target 3,500-4,000 STB/d initial capacity with flexibility to expand if performance trends toward P50 or better.

Recommendation 2: Prioritize permeability characterization

Given the dominant sensitivity, allocate appraisal budget toward:

Extended well test on discovery well (cost ~ \$2-3 million) to refine permeability estimates

- Interference test between discovery and first appraisal well (if spacing permits) to assess large-scale connectivity
- Core analysis program targeting at least 2-3 wells to quantify spatial variability

These investments should reduce forecast uncertainty by 30-40%, potentially enabling a shift from P90-based to P50-based development planning—unlocking significant value.

Recommendation 3: Monitor aquifer response early

Since aquifer strength is the second-largest uncertainty driver, implement:

- Permanent downhole pressure gauges in initial production wells to detect aquifer influx signals
- Water cut monitoring protocol to identify breakthrough timing
- Contingency planning for water handling capacity if moderate aquifer materializes

Early detection of aquifer behavior will allow adaptive management strategies, either accelerating production if aquifer is weak (depletion dominant) or managing water cut if aquifer is strong.

Recommendation 4: Design for flexibility

Given the 24-26% uncertainty range around P50 forecasts:

- Modular facility design allowing capacity expansion without major reconfiguration
- Phased drilling program (e.g., drill 3 wells initially, evaluate performance, then drill remaining 2-4 wells based on observed behavior)
- Defer irreversible commitments (e.g., long-term gas processing contracts) until production trends clarify

Recommendation 5: De-prioritize PVT refinement

The minimal sensitivity to fluid property uncertainties suggests that current PVT characterization is adequate. Budget allocated to additional PVT studies would yield better returns if redirected toward permeability or aquifer characterization.

These recommendations illustrate how uncertainty quantification directly informs practical decision-making, transforming probabilistic forecasts into actionable field development strategy.

4.7 COMPARISON WITH ALTERNATIVE FORECASTING METHODS

To contextualize the simulation-based uncertainty framework, this section compares its outputs with forecasts generated using traditional deterministic methods commonly applied in industry practice.

4.7.1 Decline Curve Analysis (DCA) Comparison

Decline curve analysis, particularly the Arps hyperbolic or exponential decline models, remains the most widely used forecasting technique for producing wells. To enable comparison, a decline curve is fitted to the first 3 years of the base case simulation output, then extrapolated to 20 years.

Table 4.8: Comparison of Forecasting Methods

Method	Cumulative Production at 20 Years	Implicit Assumptions	Uncertainty Treatment
Arps Exponential DCA	16.8 MMstb	Constant decline rate, no drive mechanism change	Single deterministic value
Arps Hyperbolic DCA (b=0.5)	19.7 MMstb	Gradually decreasing decline rate	Single deterministic value
Material Balance (analytical)	17.2 MMstb	Volumetric depletion, no aquifer	Single deterministic value
Simulation (P50)	18.5 MMstb	Explicit physics, weak aquifer	Full probability distribution
Simulation (P10-P90 range)	14.2 - 23.1 MMstb	Parameter uncertainty quantified	Risk-informed decision support

Key observations:

Exponential DCA underestimates recovery by approximately 10% compared to simulation P50, likely because the exponential model assumes boundary-dominated flow persists indefinitely. In reality, weak aquifer support slightly flattens the decline curve at later times—an effect the simple exponential cannot capture.

Hyperbolic DCA (b=0.5) overestimates recovery by about 6% relative to simulation P50. The hyperbolic model's gradually flattening decline is appropriate for some unconventional

reservoirs with changing flow regimes, but may be too optimistic for this conventional sandstone under weak aquifer drive. The hyperbolic b-factor is essentially a curve-fitting parameter without direct physical meaning, making it prone to extrapolation errors.

Analytical material balance provides reasonable agreement (17.2 vs 18.5 MMstb), validating the simulation's physical basis. The small difference arises because the analytical solution assumes pure depletion while the simulation includes weak aquifer influx.

Critically, none of the deterministic methods provide uncertainty quantification. A manager using exponential DCA would be told to expect 16.8 MMstb with no indication that the true value could plausibly range from 14.2 to 23.1 MMstb. This lack of risk awareness could lead to suboptimal decisions:

- Over-designing facilities based on optimistic (hyperbolic) forecasts
- Under-investing based on pessimistic (exponential) forecasts
- Missing opportunities to reduce uncertainty through targeted data acquisition

4.7.2 Computational Efficiency Comparison

One traditional argument against probabilistic methods has been computational cost. Table 4.9 demonstrates that modern Python-based implementation largely eliminates this concern.

Table 4.9: Computational Performance

Task	Computation Time	Hardware
Single deterministic forecast (base case)	0.48 seconds	Laptop: Intel i5, 8GB RAM
Monte Carlo with 5,000 realizations	42 minutes	Same laptop
Sensitivity analysis (Spearman correlation)	18 seconds	Same laptop
Generate all visualization plots	3 minutes	Same laptop
Total workflow time	~50 minutes	Standard laptop

For comparison, a commercial reservoir simulator (ECLIPSE, CMG) running a similar tank model might require:

- Single run: 30-60 seconds (including initialization overhead)
- 5,000 realizations: 40-80 hours on single processor
- Parallel execution on cluster: 2-4 hours (additional infrastructure cost)

The Python-based framework's 50-minute turnaround time for complete uncertainty analysis represents a **50-100x speedup** compared to commercial simulators, with no licensing costs or specialized hardware requirements. This performance advantage arises from:

- Lightweight tank model implementation (no complex grid management)
- Efficient NumPy vectorization for numerical operations
- Streamlined I/O (no large binary files to write/read)
- No licensing server overhead

This computational efficiency enables iterative workflows that would be impractical with commercial tools. For example, an engineer could:

- Run morning analysis exploring 7 uncertain parameters
- Review results over lunch
- Modify parameter ranges based on new well test data
- Rerun analysis in the afternoon
- Present updated results the same day

Such rapid iteration supports agile decision-making impossible with slower computational tools.

4.7.3 Interpretability and Transparency

A final comparison point concerns model transparency and interpretability—often cited advantages of traditional methods over "black box" approaches.

Decline curve analysis is maximally transparent: the forecast is a simple mathematical equation (exponential or hyperbolic) with 2-3 parameters. Non-specialists can understand the logic easily. However, this transparency comes at the cost of oversimplification—DCA models do not represent actual reservoir physics.

Commercial reservoir simulators implement rigorous physics but often function as black boxes due to proprietary code, complex numerical schemes, and difficult-to-inspect binary output files. Users trust the vendor's validation but cannot easily verify correctness independently.

The Python-based simulation framework developed here occupies a middle ground: it implements explicit physics (material balance, Darcy flow, aquifer models) using open-source code that can be inspected, modified, and verified. The governing equations are documented in standard textbooks (Craft & Hawkins, 1959; Dake, 1978), making the model intellectually transparent even if the numerical implementation requires programming knowledge.

For educational purposes, this transparency is particularly valuable. Students can modify the code to explore "what if" scenarios, understand how aquifer models work, or investigate numerical stability—learning opportunities unavailable with proprietary simulators.

For regulatory purposes, the open-source nature ensures that forecasts can be independently validated, addressing concerns about reliance on vendor-supplied tools that cannot be fully audited.

4.8 LIMITATIONS AND SOURCES OF UNCERTAINTY NOT CAPTURED

While the uncertainty quantification framework developed here represents a significant advance over deterministic forecasting, it is important to acknowledge limitations and sources of uncertainty that remain unaddressed in the current implementation.

4.8.1 Structural Model Uncertainty

The analysis assumes a tank model structure—a single well-mixed volume representing the reservoir. This simplification is appropriate for many screening studies but cannot capture:

Reservoir compartmentalization: If undetected faults or permeability barriers segment the reservoir into isolated or poorly connected compartments, actual production may fall well below predictions based on total OOIP. The current uncertainty analysis varies *parameter values* (permeability, porosity) but not *model structure* (connectivity).

Areal heterogeneity: The tank model averages properties spatially, potentially missing "sweet spots" of high permeability that could be preferentially targeted, or low-permeability barriers that could limit sweep efficiency.

Multi-well interference: For field developments with multiple wells, the tank model does not explicitly represent inter-well pressure interference, potentially underestimating competition for reservoir energy.

Addressing structural uncertainty would require running multiple simulations with different conceptual models (e.g., single tank vs. dual-compartment vs. three-zone models) and weighting results by geological plausibility—an extension beyond the current scope but feasible within the Python framework.

4.8.2 Operational Uncertainties Beyond Modeled Scenarios

The analysis explores parameter uncertainty but assumes a fixed operational strategy (single well, constant BHP constraint). In reality, field operations involve numerous decisions that affect outcomes:

Well count and placement decisions: Drilling additional wells could accelerate production and improve ultimate recovery, but the optimal number and locations depend on performance observed in initial wells—creating path-dependent uncertainty.

Intervention strategies: Workovers, artificial lift installations, stimulation treatments, and other interventions are typically triggered by performance thresholds. The current model does not adaptively simulate these responses.

Market-driven shutins or rate curtailments: Economic conditions (oil price volatility, regulatory changes, pipeline capacity constraints) can force temporary or permanent production changes that deterministic simulation cannot anticipate.

Some operational uncertainties could be incorporated through scenario analysis (e.g., run separate cases for 1-well vs. 2-well development). Others, particularly market-driven effects, lie outside the domain of reservoir simulation and require integration with economic models.

4.8.3 Geomechanical and Non-Darcy Flow Effects

The material balance and Darcy flow equations assume:

- Negligible rock compaction or subsidence
- Linear relationship between rate and pressure drawdown
- No formation damage or improvement over time

These assumptions hold reasonably well for conventional sandstone reservoirs at moderate drawdowns, but could break down in:

High-drawdown scenarios where non-Darcy turbulent flow becomes significant near the wellbore, reducing productivity below Darcy predictions.

Compacting reservoirs (e.g., North Sea chalk, Gulf of Mexico deepwater sands) where pore collapse under declining pressure permanently reduces permeability and porosity—a coupled geomechanical effect requiring specialized simulation.

Naturally fractured systems where matrix-fracture transfer, fracture opening/closing with pressure changes, and non-Darcy fracture flow dominate behavior.

The current framework could be extended to address some of these effects (e.g., incorporating pressure-dependent permeability) but would require additional validation and potentially slower computation.

4.8.4 Statistical Assumptions in Uncertainty Quantification

The Monte Carlo analysis assumes:

- Parameter uncertainties are independent (except where correlations explicitly imposed)
- Probability distributions are stationary (do not change as data accumulates)
- 5,000 samples adequately represent the probability space

These assumptions may not hold perfectly:

Parameter correlations: In reality, porosity and permeability often correlate positively, and certain geological features may correlate with multiple properties simultaneously. Ignoring

correlations could overestimate or underestimate total uncertainty depending on whether correlations amplify or offset individual uncertainties.

Non-stationarity: As production data accumulates, Bayesian updating should narrow uncertainty distributions—a dynamic process not reflected in the static distributions used here. This limitation is acceptable for pre-production forecasting but becomes significant for mature fields with substantial production history.

Sample size adequacy: While 5,000 samples generally suffice for robust P10/P50/P90 estimation, extreme tail probabilities (P99, P1) would require tens of thousands of samples for stability. The current analysis appropriately focuses on P10-P90 ranges relevant for decision-making rather than extreme tails.

4.8.5 Data Quality and Representatives

The most fundamental limitation is that the analysis uses *representative* parameter values from literature rather than *field-specific* measured data. Consequently:

The absolute forecast values (14.2-23.1 MMstb) should be interpreted as illustrative of the methodology rather than predictions for a specific real reservoir.

The relative importance rankings (permeability most critical, PVT least critical) are likely robust across similar reservoir types, but the exact correlation coefficients would differ with field-specific data.

Validation against field analogs provides confidence in model realism, but true validation requires comparison with actual production history—unavailable in this demonstration study.

In a field application, substituting measured parameter distributions would address this limitation. The framework is specifically designed to facilitate such substitution without altering the computational workflow.

Despite these limitations, the uncertainty quantification framework provides decision support far superior to deterministic forecasting. The limitations acknowledged here represent opportunities for future enhancement rather than fatal flaws undermining the current value.

4.9 CHAPTER SUMMARY

This chapter has presented comprehensive results from applying the simulation-based uncertainty quantification framework to a representative sandstone reservoir. The key findings and contributions include:

Probabilistic forecasts spanning P90 (14.2 MMstb) to P10 (23.1 MMstb) ultimate recovery, demonstrating that parameter uncertainty translates into approximately $\pm 24\%$ variability around the P50 forecast of 18.5 MMstb. This range provides decision-makers with realistic expectations rather than false-precision single values.

Sensitivity analysis identifying permeability (correlation +0.68) and aquifer strength (correlation +0.52) as the dominant uncertainty drivers, directly guiding priorities for additional data acquisition. PVT uncertainties contribute minimally (R_s correlation +0.08), suggesting that current fluid characterization is adequate.

Rigorous validation through material balance closure (error $< 0.5\%$), pressure decline consistency with analytical solutions, and alignment with published recovery factors for analogous reservoirs—confirming that the model captures realistic physics despite using representative rather than field-specific parameters.

Risk assessment demonstrating that even the conservative P90 scenario yields attractive economics (NPV10 > \$400 million), indicating low risk of economic failure and supporting a development decision based on probabilistic reserves classification.

Computational efficiency achieving complete uncertainty analysis (5,000 realizations) in under 1 hour on standard laptop hardware—a 50-100x speedup compared to commercial reservoir simulators, enabling rapid iteration and agile decision-making.

Superiority over traditional methods: comparison with decline curve analysis revealed that deterministic approaches provide no uncertainty quantification and can be systematically biased (exponential DCA underestimates by 10%, hyperbolic DCA overestimates by 6% relative to simulation P50).

The results validate that Python-based simulation with integrated Monte Carlo uncertainty quantification provides an accessible, transparent, and computationally efficient framework for probabilistic reservoir forecasting. The methodology delivers actionable insights—parameter sensitivity rankings, probabilistic reserves, economic risk metrics—that directly support field development decisions in ways that traditional deterministic forecasting cannot.

The following chapter will synthesize these findings, draw broader conclusions about the methodology's applicability, and recommend directions for future research and practical implementation.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

This study developed and demonstrated a simulation-based framework for reservoir performance forecasting with integrated uncertainty quantification, implemented in Python as an accessible alternative to commercial reservoir simulators. The research addressed a key challenge in reservoir engineering practice: although probabilistic forecasting is widely recognized as best practice, its application is often limited by the high cost, complexity, and proprietary nature of conventional tools.

The proposed framework combines three essential elements: a material balance-based tank model representing reservoir physics, Monte Carlo simulation for uncertainty propagation, and sensitivity analysis for identifying key drivers of forecast variability. Implemented using open-source Python libraries, the framework ensures transparency, accessibility, and computational efficiency without the need for expensive licenses or high-performance computing infrastructure.

Application of the model to a representative sandstone reservoir demonstrated its effectiveness. The probabilistic forecast produced a cumulative recovery estimate of 18.5 MMstb (P50), with a range from 14.2 MMstb (P90) to 23.1 MMstb (P10), indicating approximately $\pm 24\%$ uncertainty around the median. This range reflects realistic appraisal-stage uncertainty and highlights the limitations of relying solely on deterministic forecasts.

The results confirmed that all research objectives were achieved. The reservoir system was successfully modeled, uncertainty was quantified through Monte Carlo simulation, and the

framework was validated using material balance consistency and comparison with expected reservoir behavior. Sensitivity analysis revealed that permeability (correlation +0.68) emerged as the dominant uncertainty driver, followed by aquifer strength (+0.52), while net thickness and porosity showed moderate influence. PVT properties contributed minimally to forecast variability.

A key outcome of this work is the demonstration that deterministic “single-value” forecasts are insufficient for decision-making. By contrast, probabilistic forecasting provides a more accurate and realistic representation of reservoir performance, enabling better risk assessment and planning. The framework also showed significant computational advantages, completing thousands of simulation runs within a short time on standard hardware, thereby supporting rapid analysis and decision-making.

Overall, this study establishes that rigorous, physics-based probabilistic forecasting can be both practical and accessible. It bridges the gap between theoretical best practices and real-world application, providing a viable tool for engineers, researchers, and smaller organizations that may not have access to commercial simulation software.

5.2 RECOMMENDATION

Based on the findings of this study, several recommendations are proposed for both practical application and future development.

First, reservoir engineers and decision-makers should adopt probabilistic forecasting as a standard practice rather than relying solely on deterministic methods. The use of P10, P50, and P90 scenarios provides a clearer understanding of uncertainty and enables more informed economic

and operational decisions. This approach is particularly important during early field appraisal, where data limitations introduce significant uncertainty.

Second, data acquisition efforts should be guided by sensitivity analysis results. Since permeability, aquifer strength, net thickness, and porosity were identified as the most influential parameters, resources should be prioritized toward improving their characterization. This may include well testing, core analysis, and continuous pressure monitoring. Conversely, additional PVT refinement should be deprioritized, as sensitivity analysis demonstrated minimal impact of fluid property uncertainties (R_s correlation +0.08, B_o correlation -0.14) on forecast outcomes, allowing resources to be optimally allocated.

Third, the developed framework should be extended to incorporate more advanced reservoir features. While the tank model provides computational efficiency, future work should consider incorporating spatial heterogeneity, multi-well interactions, and layered reservoir systems. This would improve applicability to more complex reservoirs, particularly those with significant geological variability.

Fourth, the framework should be validated using real field data. Application to actual reservoir cases with historical production data will enhance credibility and enable history matching. Incorporating Bayesian updating would allow forecasts to evolve as new data becomes available, transitioning the tool from purely predictive to adaptive.

Fifth, the economic component of the framework should be expanded. Future implementations should include dynamic oil prices, detailed cost models, and fiscal terms to better reflect real-world decision-making conditions. This will allow for more accurate evaluation of project viability under uncertainty.

Finally, efforts should be made to improve usability and adoption. Developing user-friendly interfaces, automated reporting systems, and cloud-based deployment options will make the framework accessible to non-programmers and facilitate its integration into industry workflows. Training programs and academic applications should also be encouraged to promote wider understanding and use of probabilistic methods in reservoir engineering.

5.3 CONTRIBUTION TO KNOWLEDGE

This research makes several important contributions to the field of petroleum engineering, particularly in the area of reservoir forecasting and uncertainty analysis.

The primary contribution is the development of an open-source, simulation-based framework that integrates reservoir physics with probabilistic methods. Unlike traditional approaches that either rely on deterministic calculations or expensive commercial simulators, this framework provides a transparent, cost-effective, and reproducible alternative. It demonstrates that high-quality uncertainty analysis can be performed using readily available tools without compromising technical rigor.

Another key contribution is the practical demonstration of uncertainty quantification in reservoir forecasting. By applying Monte Carlo simulation and sensitivity analysis to a representative reservoir case, the study demonstrates practical implementation of probabilistic methods using accessible tools, addressing the gap between established uncertainty quantification theory and its routine application in petroleum engineering practice. The identification of dominant uncertainty drivers also contributes to improved decision-making by linking technical analysis with operational priorities.

The study further contributes to the concept of democratizing reservoir engineering tools. By eliminating financial and technical barriers, the framework enables access for smaller

companies, academic institutions, and independent researchers. This broadens participation in advanced reservoir analysis and encourages innovation across the industry.

In addition, the research enhances transparency and interpretability in forecasting. All assumptions, equations, and processes are explicitly defined, allowing for verification and validation. This contrasts with “black-box” models and supports better communication among engineers, managers, and stakeholders.

The work also contributes to the integration of engineering analysis with decision-making processes. By presenting results in probabilistic terms and linking them to economic outcomes, the framework aligns technical outputs with business needs. This supports risk-based planning, portfolio management, and value-of-information analysis.

Finally, the methodology has broader applicability beyond petroleum engineering. The principles of uncertainty quantification and simulation-based forecasting can be applied to other subsurface systems such as groundwater management, geothermal energy, carbon storage, and underground resource optimization. This interdisciplinary relevance further strengthens the significance of the research.

In summary, this study advances reservoir engineering practice by providing a practical, transparent, and scalable approach to probabilistic forecasting. It not only addresses existing limitations in tools and accessibility but also sets a foundation for future developments that integrate physics-based modeling, uncertainty analysis, and data-driven techniques.

REFERENCES

- Ahmed, T. (2019). *Reservoir engineering handbook* (5th ed.). Gulf Professional Publishing.
- Al-Fatlawi, O., Hosseinpour, N., & Kabir, A. (2022). Optimization of infill drilling in wider range field in Australia using integrated reservoir simulation and surface networks. *Petroleum Research*, 7(3), 351-365. <https://doi.org/10.1016/j.ptlrs.2021.11.005>
- Al-Harthy, M. H., Begg, S. H., & Bratvold, R. B. (2005). Valuing flexibility in an undeveloped oil field: A research and development application of real options. *SPE Economics & Management*, 1(1), 1-10. SPE-93096-PA.
- Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., & Keelan, D. K. (1993). Enhanced reservoir description: Using core and log data to identify hydraulic (flow) units and predict permeability in uncored intervals/wells. *SPE Annual Technical Conference and Exhibition*. SPE-26436-MS.
- Arps, J. J. (1945). Analysis of decline curves. *Transactions of the AIME*, 160(1), 228-247. <https://doi.org/10.2118/945228-G>
- Aziz, K., & Settari, A. (1979). *Petroleum reservoir simulation*. Applied Science Publishers.
- Behrenbruch, P., & Thanchit, A. (2001). Offshore oil and gas field development planning: Application of intelligent systems. *SPE Asia Pacific Oil and Gas Conference and Exhibition*. SPE-68713-MS.
- Bhattacharya, S., & Carr, T. R. (2016). Integrated petroleum reservoir characterization using predictive data mining with examples from the Appalachian Basin. *Interpretation*, 4(3), SB1-SB14. <https://doi.org/10.1190/INT-2015-0162.1>
- Bratvold, R. B., & Begg, S. H. (2010). *Making good decisions*. Society of Petroleum Engineers.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Caers, J. (2005). *Petroleum geostatistics*. Society of Petroleum Engineers.

- Caers, J. (2011). *Modeling uncertainty in the earth sciences*. John Wiley & Sons.
- Craft, B. C., & Hawkins, M. F. (1959). *Applied petroleum reservoir engineering*. Prentice-Hall.
- Dake, L. P. (1978). *Fundamentals of reservoir engineering*. Elsevier Scientific Publishing Company.
- Damsleth, E., Hage, A., & Volden, R. (1998). Maximum information at minimum cost: A North Sea field development study using experimental design. *SPE Annual Technical Conference and Exhibition*. SPE-49104-MS.
- Deutsch, C. V. (2002). *Geostatistical reservoir modeling*. Oxford University Press.
- Di, H., Wang, Z., & AlRegib, G. (2018). Deep convolutional neural networks for seismic salt-body delineation. *AAPG Annual Convention and Exhibition*, Salt Lake City, UT.
- Ertekin, T., Abou-Kassem, J. H., & King, G. R. (2001). *Basic applied reservoir simulation*. Society of Petroleum Engineers.
- Fang, J., Zhou, F., & Tang, Z. (2019). Discrete fracture network modelling in a naturally fractured carbonate reservoir in the Jingbei oilfield, China. *Energies*, 12(6), 945. <https://doi.org/10.3390/en12060945>
- Fetkovich, M. J. (1980). Decline curve analysis using type curves. *Journal of Petroleum Technology*, 32(6), 1065-1077. SPE-4629-PA.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Gorell, S. (2014). *Practical reservoir uncertainty assessment with limited data and time*. SPE Distinguished Lecturer Program.
- Hajizadeh, Y., Christie, M., & Demyanov, V. (2011). Ant colony optimization for history matching and uncertainty quantification of reservoir models. *Journal of Petroleum Science and Engineering*, 77(1), 78-92. <https://doi.org/10.1016/j.petrol.2011.02.005>
- Ilk, D., Rushing, J. A., Perego, A. D., & Blasingame, T. A. (2008). Exponential vs. hyperbolic decline in tight gas sands: Understanding the origin and implications for reserve

- estimates using Arps' decline curves. *SPE Annual Technical Conference and Exhibition*. SPE-116731-MS.
- Jahanbakhshi, R., & Mohaghegh, S. D. (2019). Physics-based artificial intelligence for unconventional reservoirs: A new approach for production forecasting. *SPE Reservoir Evaluation & Engineering*, 22(4), 1-15. SPE-195028-PA.
- Jahn, F., Cook, M., & Graham, M. (2008). *Hydrocarbon exploration and production* (2nd ed.). Elsevier.
- Jensen, J. L., Lake, L. W., Corbett, P. W., & Goggin, D. J. (2022). *Statistics for petroleum engineers and geoscientists* (3rd ed.). Elsevier.
- Lake, L. W. (1989). *Enhanced oil recovery*. Prentice Hall.
- Ligero, E. L., Schiozer, D. J., & Maschio, C. (2003). Risk analysis applied to petroleum exploration and production: An overview. *Journal of Petroleum Science and Engineering*, 44(1-2), 1-9.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.
- McCain, W. D. (2011). *Petroleum reservoir fluid property correlations*. PennWell Corporation.
- McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2), 239-245. <https://doi.org/10.2307/1268522>
- Mohaghegh, S. D. (2000). Virtual-intelligence applications in petroleum engineering: Part 1—Artificial neural networks. *Journal of Petroleum Technology*, 52(9), 64-73. SPE-58046-JPT.
- Murtha, J. A. (1997). *Decisions involving uncertainty: An @RISK tutorial for the petroleum industry*. Palisade Corporation.

- Myers, R. H., & Montgomery, D. C. (2002). *Response surface methodology: Process and product optimization using designed experiments* (2nd ed.). John Wiley & Sons.
- Narsilio, G. A., Buzzi, O., Fityus, S., Yun, T. S., & Smith, D. W. (2017). Upscaling of Navier–Stokes equations in porous media: Theoretical, numerical and experimental approach. *Computers and Geotechnics*, 75, 139-147.
- Oliver, D. S., & Chen, Y. (2011). Recent progress on reservoir history matching: A review. *Computational Geosciences*, 15(1), 185-221. <https://doi.org/10.1007/s10596-010-9194-2>
- Oliver, D. S., Reynolds, A. C., & Liu, N. (2008). *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press.
- Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Tucker, P. K. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, 41(1), 1-28. <https://doi.org/10.1016/j.paerosci.2005.02.001>
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Rose, P. R. (2001). *Risk analysis and management of petroleum exploration ventures*. American Association of Petroleum Geologists.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis: The primer*. John Wiley & Sons.
- Samui, P. (2011). Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Computers and Geotechnics*, 38(3), 419-423.
- Sarma, P., Durlofsky, L. J., & Aziz, K. (2008). Computational techniques for closed-loop reservoir modeling with application to a realistic reservoir. *Petroleum Science and Technology*, 26(10-11), 1120-1140. <https://doi.org/10.1080/10916460701829580>

- Schlumberger. (2021). *ECLIPSE reservoir simulation software: Technical description*. Schlumberger Limited.
- Standing, M. B. (1947). A pressure-volume-temperature correlation for mixtures of California oils and gases. *Drilling and Production Practice*, API-47-275.
- Sun, A. Y., Scanlon, B. R., Zhang, Z., Walling, D., Bhanja, S. N., Mukherjee, A., & Zhong, Z. (2022). Combining physically based modeling and deep learning for fusing GRACE satellite data: Can we learn from mismatch? *Water Resources Research*, 58(1), e2022WR032463.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics.
- Tavakoli, R., Pencheva, G., & Wheeler, M. F. (2013). Multiscale finite element methods for flows on rough surfaces. *SIAM Journal on Numerical Analysis*, 51(6), 3208-3232.
- Tiab, D., & Donaldson, E. C. (2015). *Petrophysics: Theory and practice of measuring reservoir rock and fluid transport properties* (4th ed.). Gulf Professional Publishing.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998-6008.
- Wang, S., & Mohaghegh, S. D. (2020). A machine learning approach for reservoir production forecasting. *SPE/AAPG/SEG Unconventional Resources Technology Conference*. URTeC-2020-2750.
- Zhang, D., & Curtis, A. (2020). Seismic tomography using variational inference methods. *Journal of Geophysical Research: Solid Earth*, 125(4), e2019JB018589.
- Zhang, G. P. (2008). Neural networks for classification: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 30(4), 451-462.