PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research

UNIVERSITY OF KASDI MERBAH OUARGLA
Faculty of Hydrocarbons, Renewable Energies, Earth & Universe Sciences
Department of Earth & Universe Sciences

MASTER'S THESIS

Submitted in fulfilment of the requirements for the degree of
Master of Science in Applied Geophysics

Deep Learning for Simultaneous Petrophysical Inversion of Pre-Stack Seismic Data: BiLSTM and CNN–BiLSTM Approaches for Direct Estimation of Porosity, Water Saturation, and Shale Volume

Presented by:
Hamdi Asma
Supervised by:
Nemer Zoubida & Benkhelifa Randa

Djeddi Mohamed	Pr	UKMO – Ouargla	President
Belekseir Mohamed Salah	Pr	UKMO – Ouargla	Examiner
Nemer Zoubida	MCB	UKMO – Ouargla	Supervisor
Benkhelifa Randa	MCB	UKMO – Ouargla	Co-Supervisor

Academic Year 2025–2026

DEDICATION

In the name of Allah, the Most Gracious, the Most Merciful.

To my beloved ***Parents,*** with all my love and gratitude. No words can truly express how thankful I am for your endless sacrifices, unconditional love, constant encouragement, and sincere prayers. You have always believed in me, even when I doubted myself, and every success I achieve is a reflection of your devotion and support. I thank Allah for blessing me with parents like you and for allowing me to reach this milestone. This accomplishment is, before anything else, yours.

To my dear brothers ***Tahar & Raid***, for their encouragement, and constant belief in me.

To my dear uncle, ***Lazhar***, whose invaluable support, guidance, and continuous help during this journey made this achievement possible. I will always be grateful for your kindness and generosity.

To my grand family, especially my ***Aunts***, for always standing by my side with their love, encouragement, and support.

To my wonderful friends, who shared this journey with me, supported me through every challenge, and filled these years with unforgettable memories.

Finally, to everyone who believed in me, encouraged me, and contributed, directly or indirectly, to this achievement.

I dedicate this humble work to you all.

ACKNOWLEDGMENTS

First and foremost, all praise and gratitude are due to ***Allah Almighty*** for granting me the strength, patience, and perseverance to complete this work.

I would like to express my deepest gratitude to my supervisor, ***Dr. Nemer Zoubida***, for her invaluable guidance, continuous support, and dedication throughout this research. Her expertise, insightful feedback, and generous assistance have been invaluable in shaping this work I am sincerely grateful for the opportunity to work under her supervision, and I deeply appreciate all the time and effort she devoted to helping me successfully complete this dissertation.

My sincere thanks also go to ***Dr. Benkhelifa Randa*** for her kind assistance, helpful advice, and encouragement during the development of this work.

I am also grateful to the members of the jury for accepting to evaluate this dissertation and for their valuable comments and suggestions.

Finally, I would like to thank my parents, my family, my friends, and everyone who supported and encouraged me throughout this journey.

ABSTRACT

This thesis presents a deep learning framework for the direct estimation of reservoir petrophysical properties, porosity (φ), water saturation (Sw), and shale volume (Vsh), from pre-stack seismic data. Conventional inversion workflows rely on linearized approximations and multi-stage processing, leading to uncertainty accumulation and poor recovery of lithological parameters such as Vsh. To overcome these limitations, Bidirectional Long Short-Term Memory (BiLSTM) and hybrid CNN–BiLSTM networks were developed for simultaneous prediction of φ, Sw, and Vsh directly from multi-angle seismic reflections. A synthetic dataset of 10,000 geological models was generated using Gassmann fluid substitution, Zoeppritz-based angle-dependent reflectivity at five incidence angles, and Ricker wavelet convolution, calibrated to North African clastic reservoirs. The CNN–BiLSTM architecture achieved the best performance, with RMSE values of 0.0149 for φ, 0.0639 for Sw, and 0.0461 for Vsh, outperforming classical Gauss–Newton inversion, particularly for Vsh estimation where the conventional method failed. Monte Carlo Dropout provided predictive uncertainty estimates, and the models remained highly robust to seismic noise across a wide signal-to-noise ratio range without retraining.

Keywords: seismic inversion, deep learning, CNN–BiLSTM, BiLSTM, porosity, water saturation, shale volume, reservoir characterization.

RESUME

Ce memoire présente un cadre d'apprentissage profond pour l'estimation directe des propriétés pétrophysiques des réservoirs, porosité (φ), saturation en eau (Sw) et volume d'argile (Vsh), à partir de données sismiques pré-empilées. Les méthodes classiques d'inversion reposent sur des approximations linéarisées et des traitements en plusieurs étapes, entraînant une accumulation des incertitudes et une mauvaise estimation des paramètres lithologiques tels que le Vsh. Afin de surmonter ces limites, des réseaux Bidirectional Long Short-Term Memory (BiLSTM) et hybrides CNN–BiLSTM ont été développés pour la prédiction simultanée de φ, Sw et Vsh directement à partir des réflexions sismiques multi-angles. Une base de données synthétique de 10 000 modèles géologiques a été générée à l'aide de la substitution de fluides de Gassmann, des coefficients de réflexion dépendants de l'angle basés sur les équations de Zoeppritz à cinq angles d'incidence, ainsi que de la convolution par ondelette de Ricker, le tout calibré pour des réservoirs détritiques nord-africains. L'architecture CNN–BiLSTM a obtenu les meilleures performances, avec des valeurs de RMSE de 0,0149 pour φ, 0,0639 pour Sw et 0,0461 pour Vsh, surpassant l'inversion classique de Gauss–Newton, notamment pour l'estimation du Vsh où la méthode conventionnelle a complètement échoué. La méthode Monte Carlo Dropout a permis de quantifier les incertitudes prédictives, tandis que les modèles ont montré une forte robustesse au bruit sismique sur une large gamme de rapports signal/bruit sans réentraînement.

Mots-clés : inversion sismique, apprentissage profond, CNN–BiLSTM, BiLSTM, porosité, saturation en eau, volume d'argile, caractérisation des réservoirs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Definition
AI	Acoustic Impedance
AVO	Amplitude Versus Offset
BiLSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional Neural Network
DL	Deep Learning
GN	Gauss–Newton
GR	Gamma Ray Log
LSTM	Long Short-Term Memory
MC	Monte Carlo
MCD	Monte Carlo Dropout
ML	Machine Learning
MSE	Mean Squared Error
NPHI	Neutron Porosity Log
PINNs	Physics-Informed Neural Networks
RHOB	Bulk Density Log
RMSE	Root Mean Squared Error
RT	Resistivity Log
SNR	Signal-to-Noise Ratio

LIST OF SYMBOLS

Symbol	Definition	Unit
φ	Porosity	– (fraction)
Sw	Water Saturation	– (fraction)
Vsh	Shale Volume	– (fraction)
Vp	P-wave Velocity	m/s
Vs	S-wave Velocity	m/s
ρ	Bulk Density	kg/m^3
Kdry	Dry Bulk Modulus	GPa
Ksat	Saturated Bulk Modulus	GPa
Kfl	Fluid Bulk Modulus	GPa
Kmin	Mineral Bulk Modulus	GPa
R^2	Coefficient of Determination	–
σ	Standard Deviation (Uncertainty)	–
θ	Incidence Angle	°

INTRODUCTION

1. Background

In the high-stakes environment of the upstream petroleum industry, the ability to accurately define and characterize a hydrocarbon reservoir is more than a technical goal; it is a financial necessity before committing to costly drilling and production operations. The economic viability of any given reservoir is primarily governed by three fundamental petrophysical properties: *porosity* (φ), which quantifies the fraction of pore space available to store fluids; *water saturation* (Sw), which indicates what proportion of that pore space is occupied by brine rather than hydrocarbons; and *shale volume* (Vsh), which reflects the clay content of the rock and directly controls its reservoir quality and seismic response. Together, these three variables define the subsurface in terms that are both geologically coherent and directly relevant to field decisions (Mavko et al., 2020; Grana et al., 2021).

When it comes to investigating the subsurface at a reservoir-wide scale without the immediate need for drilling, *seismic reflection* remains the industry's primary diagnostic tool. The operating principle is straightforward: Elastic waves, generated at the surface by an artificial energy source, propagate downward through the earth, interact with boundaries between formations of contrasting elastic properties, and are partially reflected back to surface sensors, while the remaining energy continues propagating into deeper layers. These recorded reflections are not arbitrary signals. Instead, they encode in a physically predictable way: the elastic contrasts between adjacent layers. Through the lens of rock physics, those contrasts can be translated back into the petrophysical quantities of interest. In a seismic survey, the recorded signals are organized as time-series traces, one per receiver, and are then processed to remove noise, correct for acquisition geometry, and enhance the geological signal. Among the various acquisition strategies available, *pre-stack seismic data* recorded before the conventional summation of multiple source–receiver offsets are particularly valuable. By preserving the *amplitude variation with offset (AVO)*, pre-stack data retain differential sensitivity to both compressional and shear wave properties, and by extension to pore fluid type, in a way that single-offset stacked data fundamentally cannot (Castagna & Backus, 1993; Avseth et al., 2005).

For several decades, the standard response to this challenge has been *model-based inversion*: an iterative optimization scheme that seeks the subsurface model whose synthetic seismic response best matches the observed field data, typically subject to well-log constraints or smoothness regularization (Russell, 1988; Tarantola, 2005). These classical frameworks are physically interpretable and well established in industry practice. However, they carry persistent limitations: sensitivity to the quality of the initial model, reliance on linearized approximations that may break down for large elastic contrasts, and compounded uncertainties across two inversion stages (Grana et al., 2021).

The rapid maturation of deep learning in recent years has opened a genuinely different avenue. Rather than constructing an explicit physical model and inverting it analytically, neural networks can be trained to learn the mapping from seismic observations to petrophysical parameters directly, from large ensembles of physically simulated examples. This approach has the potential to bypass several of the bottlenecks of the classical pipeline. Among the architectures suited to sequential data such as seismic traces, the *Bidirectional Long Short-Term Memory (BiLSTM)* network is particularly well motivated. It processes each trace in both forward and backward directions, allowing every predicted depth sample to leverage full contextual information from the entire trace. This property maps naturally onto the physics of seismic reflections, since each recorded reflection is influenced by elastic contrasts occurring both above and below any given interface (Graves & Schmidhuber, 2005; Biswas et al., 2019). The present study develops and evaluates such a BiLSTM-based framework, trained on a synthetically generated dataset, for the direct simultaneous prediction of φ, Sw, and Vsh from multi-angle pre-stack seismic data.

2. Problem Statement

Machine learning has become increasingly prominent in seismic reservoir characterization, yet several important limitations remain in current approaches. Most existing studies either focus on a single petrophysical parameter or rely on relatively simple network architectures that do not fully exploit the sequential structure of seismic traces. More importantly, very few studies provide a complete, end-to-end workflow that simultaneously predicts all three key reservoir parameters φ, Sw, and Vsh while also quantifying predictive uncertainty and benchmarking results against a classical inversion method on an equal footing. In this context, the present work aims to develop and document a complete BiLSTM-based framework for the joint estimation of φ, Sw, and Vsh from multi-angle pre-stack seismic traces, implement it on a physically realistic synthetic dataset, and evaluate its behaviour in a controlled setting. The focus is not only on predictive performance, but also on presenting a transparent and reproducible workflow itself, examining model performance under varying conditions, and discussing the advantages and limitations of the data-driven approach relative to classical inversion. Because access to labelled field data was not available for this work, the evaluation is carried out entirely on synthetic data a choice that is common in this class of study (Wu et al., 2019; Biswas et al., 2019) and that provides exact ground-truth values against which predictions can be rigorously assessed.

The fundamental difficulty addressed here is the *non-linear* and *ill-posed nature of the seismic inverse problem* (Tarantola, 2005). Seismic data result from a chain of, forward processes, rock physics, wave propagation, and convolution each of which is non-bijective: multiple distinct subsurface configurations can produce seismic traces that are, within noise levels, indistinguishable. This inherent ambiguity, combined with the band-limited nature of the seismic wavelet, renders classical deterministic inversion methods fundamentally under-constrained problem. A deep learning model trained on a sufficiently diverse and physically realistic synthetic dataset may navigate this ambiguity more effectively by implicitly learning the regularities of the petrophysical–elastic–seismic mapping from data, without explicitly decomposing the seismic response into multiple sequential inversion stages, each carrying its own source of uncertainty.

3. Research Objectives

This research is structured around the following objectives:

- **Synthetic Framework Development:** Construct a physically realistic synthetic geological model and implement a complete forward modelling chain from petrophysical parameters through rock physics transformation and Zoeppritz-based reflection coefficients to synthetic pre-stack seismic gathers.
- **Architectural Design and Optimization:** Design, train, and optimize BiLSTM and CNN–BiLSTM neural networks for the simultaneous direct prediction of φ, Sw, and Vsh from multi-angle pre-stack seismic traces.
- **Robustness Assessment:** Systematically evaluate the model's resilience under realistic perturbations including varying seismic noise levels, wavelet estimation errors, and rock physics parameter uncertainty.
- **Uncertainty Quantification:** Quantify prediction uncertainty through Monte Carlo Dropout, providing confidence intervals alongside point estimates for all three petrophysical targets.

- **Comparative Benchmarking:** Benchmark the deep learning approach against a classical linearized Gauss–Newton inversion in terms of predictive accuracy, computational efficiency, and robustness to noise.

4. Thesis Structure

This thesis is divided into four main chapters, following the logical progression of the work from theoretical foundations to practical implementation and result interpretation.

Chapter 1 – Theoretical Background:
This chapter presents the fundamental theoretical concepts that underpin the study. It covers seismic wave propagation, AVO behaviour, rock physics relationships, and the mathematical formulation of the seismic inverse problem. It also introduces the basic principles of machine learning and deep learning in geophysical applications, which form the basis of the modelling approach adopted in this work.

Chapter 2 – Methodology:
This chapter details the complete methodological framework of the study. It describes the construction of the synthetic geological model, the forward modelling workflow from petrophysical parameters to pre-stack seismic data, and the dataset generation and preprocessing procedures. It also presents the design of the deep learning models used for inversion, including both BiLSTM and CNN–BiLSTM architectures.

Chapter 3 – Results:
This chapter presents the results obtained from applying the proposed methodology to the synthetic dataset. It includes dataset characterization, training behaviour analysis, quantitative performance evaluation using multiple metrics, robustness assessment under different noise conditions, uncertainty estimation using Monte Carlo Dropout, and comparison with a classical Gauss–Newton inversion approach.

Chapter 4 – Discussion:
This chapter focuses on interpreting the results in relation to the theoretical background and methodological choices. It discusses model behaviour, physical consistency of the predictions, advantages and limitations of the proposed approach, and the implications of robustness and uncertainty analysis for practical reservoir characterization.

CHAPTER 1: THEORETICAL BACKGROUND

Understanding how pressure waves travelling through rock carry information about the fluids and minerals filling the pore space is the central scientific challenge of seismic reservoir characterization. This chapter develops three distinct bodies of knowledge in turn: the physics of elastic wave propagation, which determines what the seismic signal encodes; rock physics, which translates that encoding into the language of petrophysical properties; and inverse problem theory, which provides the mathematical tools for decoding it.

1.1. Elastic Wave Propagation and the AVO Phenomenon

1.1.1. Wave Velocities and Their Physical Meaning

The propagation velocity of a seismic wave through a rock formation is determined by the ratio of the medium's elastic stiffness to its inertia. In any isotropic, linearly elastic medium, two independent propagation modes exist: *compressional P-waves*, in which particle motion is parallel to the direction of wave propagation, and *shear S-waves*, in which particle motion is perpendicular to it. Their respective velocities are given by (Sheriff & Geldart, 1995; Mavko et al., 2020):

$$V_P = \sqrt{\left[\frac{\left(Ksat + \frac{4G}{3}\right)}{\rho}\right]} \, and \, V_S = \sqrt{\left[\frac{G}{\rho}\right]} \, (1.1)$$

where Ksat is the *saturated bulk modulus*, G is the *shear modulus*, and ρ is the *bulk density of the saturated rock.*

The physical significance of these two velocities differs fundamentally, and this difference forms the basis of amplitude variation with offset (AVO) analysis. The bulk modulus Ksat is sensitive to the compressibility of the pore fluid, and therefore to fluid type. The shear modulus G, by contrast, is a property of the solid grain framework alone, a result established rigorously by Gassmann (1951) and discussed in Section 2.2. Consequently, VP changes when gas displaces brine in the pore space, while VS remains largely unaffected. This differential response makes the VP/VS *ratio* a powerful indicator of pore fluid type, and it is the primary reason why multi-angle pre-stack data which encodes both velocities simultaneously provides significantly more diagnostic information than single-offset stacked data (Castagna & Backus, 1993).

It is important to recognize, however, that the sensitivity of seismic amplitudes to elastic parameter variations is not uniform across geological scenarios. While V_P responds to both lithological changes and fluid substitution, V_S responds primarily to changes in the rock frame, making it a more reliable indicator of lithology than of fluid content. AVO analysis exploits this contrast to *decouple fluid effects from lithological variations*; yet in practice, this decoupling is rarely perfect. Noise, wavelet uncertainty, and the limited angular aperture of real acquisitions introduce ambiguities that cannot be resolved by physics alone (Avseth et al., 2005). These effects become particularly problematic during inversion, where lithological and fluid contributions must be separated from incomplete and noisy seismic observations.

1.1.2. Seismic Impedance and Reflection Coefficients

Seismic reflections are generated at boundaries between layers of contrasting elastic properties. The physical quantity governing reflection strength is the *acoustic impedance* $(AI) = \rho \cdot V_P$. At normal incidence, the reflection coefficient simplifies to a contrast ratio (Sheriff & Geldart, 1995):

$$\mathrm{R_0} = \frac{(\mathrm{AI_2} - \mathrm{AI_1})}{(\mathrm{AI_2} + \mathrm{AI_1})} \qquad (1.2)$$

This expression has a critical implication for inversion: seismic data records *impedance contrasts*, not absolute impedance values. The absence of the low-frequency trend from the seismic record is therefore not a technical artifact it is a fundamental consequence of the physics. Any inversion method must supply this missing trend from an external source, such as well logs or geological prior knowledge.

At non-normal incidence, the reflection problem becomes considerably more complex. The amplitude depends not only on the AI contrast but also on the contrast in *shear impedance* $(\mathrm{SI} = \rho \cdot \mathrm{V_S})$ and density separately. The exact solution is given by the *Zoeppritz equations* a 4×4 linear system relating reflection and transmission coefficients at all angles to the elastic properties of both half-spaces (Aki & Richards, 1980). These exact equations are used in the forward modeling described in Chapter 2. For physical intuition, the *Shuey (1985) two-term linearization* is instructive:

$$\mathrm{RPP}(\theta) \approx \mathrm{R0} + \mathrm{G} \cdot \sin^2\theta \qquad (1.3)$$

where R_0 is the *normal-incidence intercept* and G is the *AVO gradient*, a quantity that depends on the contrasts in V_S and density across the interface. Equation (1.3) makes explicit why recording at multiple angles simultaneously encodes both P- and S-impedance information. This is the core justification for using multi-angle pre-stack gathers as the network input: each additional angle adds an independent linear constraint on the elastic contrasts, and collectively, the suite of angles permits simultaneous sensitivity to lithology (viaV_S and density) and fluid content (via the incremental effect on Ksat).

1.1.3. The Convolutional Seismic Model

In field conditions, the subsurface reflectivity series r(t) is never observed directly. What is recorded is its convolution with the seismic source wavelet w(t), plus additive noise n(t) (Russell, 1988):

$$s(t) = w(t) * r(t) + n(t) \quad (1.4)$$

The wavelet acts as a *band-pass filter* typically spanning 10–80 Hz, which has two important consequences. First, high-frequency detail below the tuning thickness is suppressed, limiting vertical resolution. Second, and more critically for inversion, the low-frequency components of the impedance profile the absolute background trend that a geologist would recognize as the direct-current (DC) component are entirely absent from the recorded signal. This spectral gap is one of the principal sources of non-uniqueness in seismic inversion. Because the data simply does not contain this information, it must be provided by some form of prior whether a well log, a geological model, or, in the data-driven approach taken here, the statistical relationships represented within the training ensemble.

1.2. Rock Physics: Bridging the Reservoir–Seismic Divide

1.2.1. Gassmann Fluid Substitution

The most important rock physics result for fluid-sensitive seismic analysis is the *Gassmann (1951) equation*, which quantifies how the bulk modulus of a porous, fluid-saturated rock depends on its solid frame, mineral constituents, and pore fluid. Given the dry-frame bulk modulus Kdry, the mineral modulus Kmin, the pore fluid modulus Kfl, and the porosity φ, the saturated bulk modulus is:

$$\mathrm{Ksat} = \mathrm{Kdry} + \frac{\left(1 - \frac{\mathrm{Kdry}}{\mathrm{Kmin}}\right)^2}{\left[\frac{\varphi}{\mathrm{Kfl}} + \frac{1-\varphi}{\mathrm{Kmin}} - \frac{\mathrm{Kdry}}{\mathrm{Kmin}^2}\right]} \qquad (1.5)$$

Gassmann's formulation also establishes that the shear modulus is fluid-independent: $Gsat = Gdry$. This result, while counter-intuitive at first glance, follows directly from the assumption that shear deformation does not induce changes in pore pressure a consequence of the incompressibility of fluids under shear. The *saturated bulk density* is computed through volumetric averaging:

$$\rho_{\mathrm{sat}} = (1-\varphi) \cdot \rho_{\mathrm{min}} + \varphi \cdot [S_w \cdot \rho_w + (1 - S_w) \cdot \rho_{HC}]$$

For a pore space containing a mixture of brine and hydrocarbon, the effective pore fluid modulus is computed from the *Wood (1955)* harmonic averaging relation:

$$\frac{1}{K_{fl}} = \frac{S_w}{K_w} + \frac{1 - S_w}{K_{HC}} \qquad (1.6)$$

Together, equations (1.1) and (1.5) define a closed-form forward map from the petrophysical state (φ, S_w) to the elastic observables (V_P, V_S) for a given mineral frame. This is used throughout this thesis to generate the synthetic training labels.

1.2.2. Assumptions and Limitations of the Gassmann Equation

Gassmann's equation requires a fully saturated, monomineralic, isotropic, and elastically homogeneous medium operating under low-frequency conditions. In practice, real reservoirs frequently violate these assumptions: shaly sands are neither monomineralic nor isotropic; fractured carbonates introduce preferred directions of compliance; and wireline logging tools operate at kilohertz frequencies. The applicability of rock physics models is therefore strongly dependent on geological conditions (Mavko et al., 2020; Grana et al., 2021). By training a neural network on a large and physically diverse synthetic ensemble, the model can learn to capture nonlinear and heterogeneous relationships that lie beyond the reach of any single analytical rock physics model (Dramsch, 2020).

1.2.3. Mineralogical Mixing and the Role of Shale Volume

In a sand–shale system, the elastic properties of the mineral frame depend on the volumetric proportion of clay minerals, quantified by the *shale volume Vsh*. This parameter controls the stiffness of the mineral frame through its influence on Kmin and Gmin, and affects bulk density through the higher density of clay minerals relative to quartz, captured through linear mixing (Mavko et al., 2020):

$$\mathrm{Kmin} = (1 - \mathrm{Vsh}) \cdot \mathrm{Kqtz} + \mathrm{Vsh} \cdot \mathrm{Ksh} \qquad (1.7)$$

The dry-frame moduli Kdry and Gdry are estimated using the *critical porosity model* of Nur et al. (1998), which assumes a linear variation of frame stiffness with porosity below a critical threshold ($\varphi_c \approx 0.38$ for sandstones). Combined with equations (1.5) through (1.7), this yields a complete, closed-form forward map from the three petrophysical targets (φ, S_w, V_{sh}) to the three elastic observables $(V_P, V_S, \rho)$the map that is implemented in the forward modeling pipeline of Section 2.2.

It is worth emphasizing that Vsh is not merely a secondary corrective factor. Changes in shale volume alter the mineral moduli, the frame stiffness, and the bulk density simultaneously, producing AVO responses that can mimic or mask fluid-related signatures. This coupling between lithology and fluid effects is precisely what makes the simultaneous estimation of all three parameters $\varphi, S_w, and\ V_{sh}$ together both challenging and important. Treating V_{sh} as a target of the inversion, rather than as a nuisance parameter to be corrected for, is therefore a deliberate and scientifically motivated design choice of this study.

1.3. The Seismic Inverse Problem

1.3.1. Mathematical Formulation

In the linearized case, the relationship between the model vector m and the observed seismic data vector d takes the compact form:

$$\mathrm{d} = \mathrm{Gm} + \mathrm{n} \qquad (1.8)$$

where G is the *forward operator matrix* encoding the combined effects of the rock physics transform, the Zoeppritz reflection coefficients, and the wavelet convolution, and n is an additive noise vector representing measurement error and model approximation. Recovering m from d the seismic inversion problem requires inverting the action of G. This formulation immediately highlights the central difficulty: G is typically a wide rank-deficient matrix. As a result, equation (1.8) has infinitely many solutions that fit the data equally well within the noise level. This is the mathematical expression of the *non-uniqueness* that was discussed physically in Section 1.1.3.

Figure 1 illustrates the relationship between the forward and inverse problems in seismic reservoir characterization. In the forward process, petrophysical properties are transformed into elastic parameters and subsequently into seismic responses through rock physics and wave propagation models. The inverse problem attempts to recover the original reservoir properties from seismic observations, making it inherently nonlinear and ill-posed (Zayier et al., 2026).

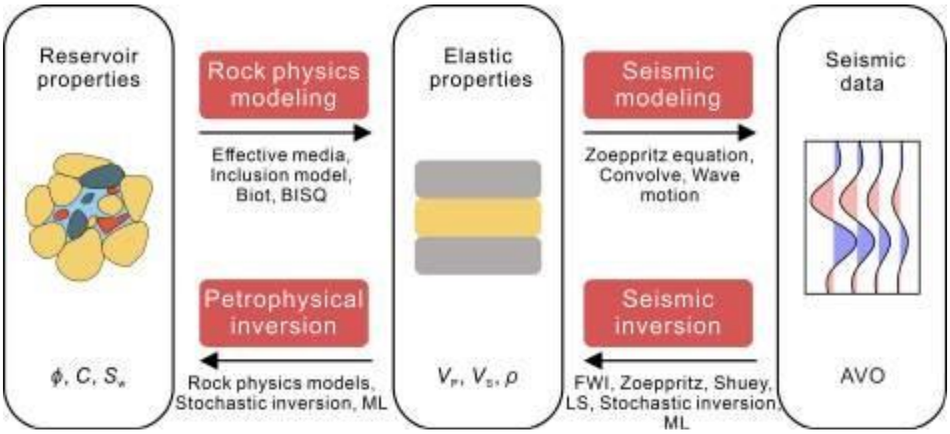

Figure 1 Forward and inverse modeling workflow for seismic reservoir characterization

1.3.2. Ill-posedness and Its Sources

Following Hadamard (1902), the seismic inverse problem is *ill-posed*: its solution is neither unique nor stably dependent on the data (Tarantola, 2005; Menke, 2012). Three sources of ill-posedness compound each other: (i) the spectral gap created by the seismic wavelet, which removes the absolute impedance level from the recorded signal; (ii) forward operator non-linearity, since the Gassmann and Zoeppritz operators do not commute with perturbations; and (iii) the non-unique mapping from elastic to petrophysical space, where different combinations of φ, Sw, and Vsh can yield identical values of V_P, V_S, and ρ (Grana et al., 2021).

1.3.3. Bayesian Interpretation and the Role of Prior Information

A principled way to address ill-posedness is through the Bayesian framework (Tarantola, 2005; Grana et al., 2021), which treats model parameters as random variables and seeks:

$$P(m \mid d) \propto P(d \mid m) \cdot P(m) \qquad (1.9)$$

In this formulation, P(d | m) is the *likelihood* how probable the observed data is if the model *m* were true, typically derived from the noise statistics and P(m) is the *prior distribution*, which encodes geological knowledge before any data is seen. This formulation formalizes the role of prior geological information within the inversion process: without a prior, the posterior $P(m \mid d)$ is not uniquely defined. Classical Tikhonov regularization, discussed below, corresponds to a Gaussian prior with a specific covariance structure. From this Bayesian perspective, *machine learning can be interpreted as a data-driven inversion framework*, in which the prior and the forward operator are not specified analytically but are instead learned implicitly from a large ensemble of physically consistent model–data pairs. The network effectively learns to approximate the posterior mean $P(m \mid d)$ in a computationally tractable way, an interpretation that gives the data-driven approach developed in this thesis a principled probabilistic standing.

1.3.4. Classical Regularized Inversion

The standard deterministic approximation is *Tikhonov regularization* (Tikhonov & Arsenin, 1977), which minimizes:

$$C(m) = \|d - F(m)\|^2 + \lambda \|L(m - mprior)\|^2 \qquad (1.10)$$

where L is a regularization operator and λ controls the trade-off between data fidelity and model smoothness. This constitutes the classical benchmark implemented in Chapter 3. Its known practical limitations, sensitivity to the choice of λ, dependence on the quality of the prior model, and inability to recover non-unique petrophysical solutions, are primary motivations for exploring the data-driven alternative (Dramsch, 2020).

1.4. Machine Learning and Deep Learning Fundamentals

1.4.1. Supervised Learning as an Inversion Strategy

The data-driven approach adopted here falls within the supervised learning paradigm (Bishop, 2006): a parametric model f(θ) is trained to approximate the mapping from inputs x (multi-angle pre-stack seismic traces) to targets y (corresponding logs of φ, S_w, and V_{sh} at each depth sample). The training objective is minimization of the *mean squared error (MSE)* loss:

$$\mathcal{L}(\theta) = \left(\frac{1}{N}\right) \cdot \Sigma_i \|yi - f\theta(x_i)\|^2 \qquad (1.11)$$

The MSE loss is chosen for two complementary reasons: its quadratic penalization of large errors makes it sensitive to sharp parameter discontinuities at facies boundaries, and minimizing MSE is equivalent to maximum likelihood estimation under a Gaussian noise model (Goodfellow et al., 2016). Parameter optimization uses the Adam algorithm

(Kingma & Ba, 2015), which adapts the effective learning rate individually for each weight based on gradient history.

1.4.2. Overfitting and Regularization

Three complementary strategies control overfitting risk: (i) *L2 weight decay* penalizes large weights; (ii) *Dropout* (Srivastava et al., 2014) randomly deactivates a fraction of neurons during training, forcing the network to develop redundant, robust representations; and (iii) *early stopping* halts training as soon as validation loss begins to increase consistently. The use of Dropout carries an additional benefit: by keeping it active during inference and performing multiple stochastic forward passes, the network produces a distribution of predictions, Monte *Carlo Dropout* (Gal & Ghahramani, 2016), whose standard deviation serves as a measure of epistemic uncertainty.

1.5. Deep Learning Architectures for Sequential Seismic Data

1.5.1. Long Short-Term Memory (LSTM)

The LSTM architecture of Hochreiter & Schmidhuber (1997) resolves the vanishing gradient problem through an additive memory cell update rule. Three learned gating functions control information flow: the *forget gate* decides what fraction of the previous cell state to retain; the *input gate* controls how much new information is written; and the *output gate* determines what portion of the cell state is exposed as the output hidden state. The additive update ensures the gradient signal can flow backward through arbitrarily long sequences, allowing the model to maintain information about reflectors encountered many hundreds of depth samples earlier.

Figure 2 presents the internal architecture of an LSTM cell. The input, forget, and output gates regulate the flow of information through the memory cell, enabling the network to preserve long-term dependencies while mitigating the vanishing gradient problem. This property makes LSTM particularly suitable for sequential seismic data (Feng et al., 2024).

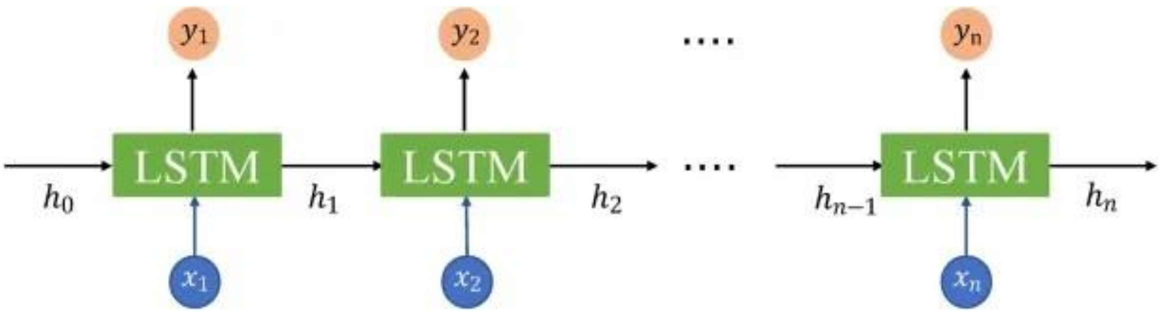

Figure 2 long short-term memory (LSTM) architecture (Each green box stands for one LSTM unit).

1.5.2. Bidirectional LSTM (BiLSTM)

A standard LSTM can only access context from samples above the current depth. The BiLSTM of Graves & Schmidhuber (2005) addresses this by running two independent LSTM sub-networks in parallel and concatenating their hidden states:

$$h^{tBiLsTM} = [\, h^{t}(\rightarrow) \parallel h^{t}(\leftarrow) \,] \qquad (1.12)$$

The result is a representation that, at every depth sample t, simultaneously encodes what has been observed above and below it, directly mirroring the physics of the seismic forward model. BiLSTM-based networks consistently outperform their unidirectional counterparts on seismic petrophysical prediction tasks (Biswas et al., 2019; Alfarraj & AlRegib, 2019).

Figure 3 illustrates the BiLSTM architecture, where the input sequence is processed simultaneously in forward and backward directions. By combining information from both directions, the network captures contextual dependencies from adjacent geological layers, leading to improved sequential feature learning (Feng et al., 2024).

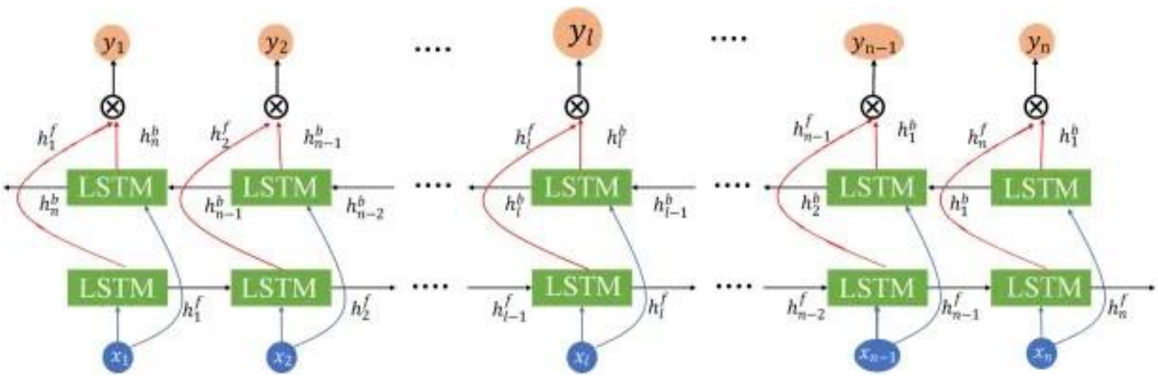

Figure 3 Bidirectional long short-term memory (Bi-LSTM) architecture.

1.5.3. 1D Convolutional Front-End and the CNN–BiLSTM Hybrid

The CNN–BiLSTM hybrid exploits complementary capabilities: the convolutional front-end extracts compact local AVO features from the multi-angle input, and the BiLSTM integrates these features across the full depth axis within their geological context. CNNs alone, without a recurrent mechanism, have no principled way to propagate information across the full depth axis, which justifies the use of LSTM-based architectures as the primary sequential model (Wu et al., 2019; Bergen et al., 2019).

Figure 4 illustrates a hybrid CNN–BiLSTM architecture. The one-dimensional convolutional layers first extract local features from the seismic sequence, after which the BiLSTM layers model long-range dependencies along the depth axis. This combination improves feature extraction and sequence learning for petrophysical prediction (Feng et al., 2024).

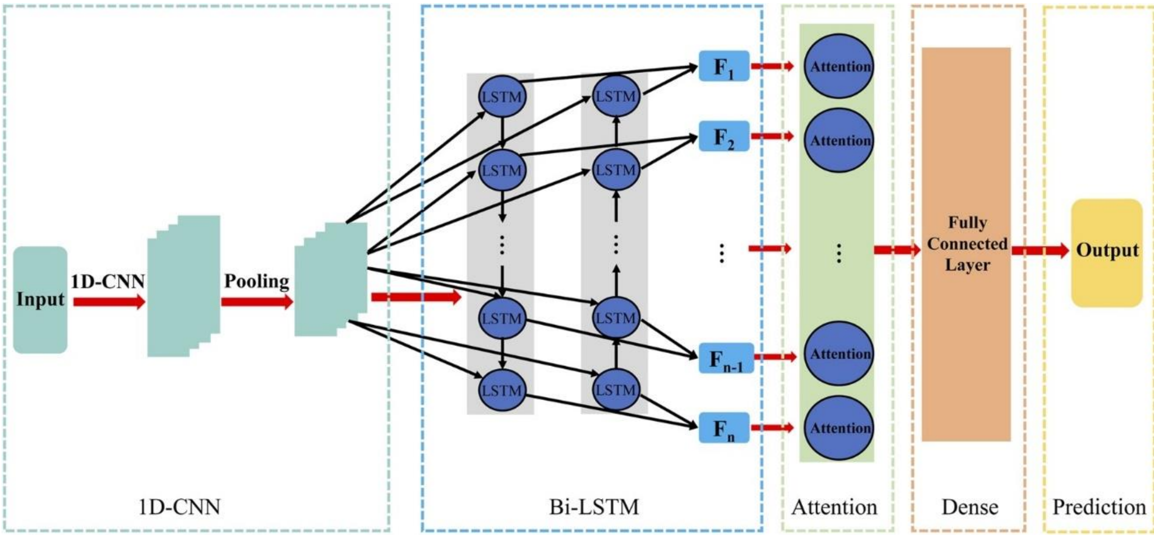

Figure 4 Illustrates a typical CNN-BiLSTM architecture

CHAPTER 2: METHODOLOGY

The methodological framework of this thesis is based on synthetic data. This choice is driven by the unavailability of labelled field datasets, namely pre-stack seismic data paired with reliable well-log measurements of φ, S_w, and V_{sh} at the same locations. Under these conditions, data generated from a known geological model provides the only feasible basis for a rigorous quantitative evaluation, since the ground-truth values of the petrophysical parameters are known exactly by construction (Wu et al., 2019; Biswas et al., 2019).

2.1. Synthetic Geological Model

The synthetic model represents a multilayered clastic reservoir system of the type commonly encountered in North African sedimentary basins. Each realization is constructed as a 1D stratigraphic column composed of alternating geological units: shale intervals (serving as seals and baffles), brine-saturated sandstone layers, and hydrocarbon-bearing sand bodies. At each depth sample, the three petrophysical parameters (φ, Sw, Vsh) are drawn from prior distributions calibrated to published values for North African clastic reservoirs (Grana et al., 2021; Mavko et al., 2020).

Table 1: Petrophysical parameter ranges used for synthetic dataset generation

Parameter	Range
Porosity (φ)	0.05 – 0.30
Water Saturation (Sw)	0.20 – 1.00
Shale Volume (Vsh)	0.00 – 0.60

The prior distribution of porosity is modeled as $\varphi \sim N(0.15, 0.05^2)$, truncated to the physical bounds [0.05, 0.30]. Spatial continuity along the depth axis is introduced through a Gaussian smoothing kernel with a correlation length of 10–30 samples. Sharp facies boundaries are superimposed by randomly placing layer transitions at which parameters change discontinuously, generating the impedance contrasts that produce the strongest seismic reflections and, consequently, the most informative AVO signatures in the training data.

A deliberate and scientifically motivated design choice of this study: unlike many published machine learning inversion studies that focus exclusively on porosity and water saturation, this work treats shale volume (V_{sh}) as a primary prediction target throughout the entire pipeline. As established in Section 1.2.3, V_{sh} independently controls the mineral moduli, the dry-frame stiffness, and the bulk density of the rock, producing AVO effects that can mask or mimic fluid signatures. Treating V_{sh} as a co-equal inversion target provides a more complete and internally consistent characterization framework.

2.2. Forward Modeling Workflow

The transformation of a petrophysical column (φ, S_w, V_{sh}) into its seismic equivalent follows a deterministic three-stage pipeline:

- **Rock Physics Mapping:** The Gassmann equation (1.5) is applied to compute Ksat from dry-frame moduli, the mineral moduli from linear mixing (1.7), and the effective pore fluid modulus from the Wood equation (1.6). This converts the petrophysical column into an elastic column (V_P, V_S, ρ).
- **Zoeppritz-based Reflectivity:** At each interface between adjacent depth samples, the exact Zoeppritz system (Aki & Richards, 1980) is solved numerically for K = 5 angle stacks at 5°, 15°, 25°, 35°, and 45°. The exact equations are used in preference to the Shuey linearization (1.3) to avoid introducing any modelling bias into the training data.
- **Seismic Synthesis and Noise Injection:** Each angle-dependent reflectivity series is convolved with a 35 Hz zero-phase Ricker wavelet. Band-limited Gaussian noise is added at a signal-to-noise ratio (SNR) of 10 dB for the baseline dataset.

The output of the pipeline is a 2D seismic gather of shape $[N_{samples} \times 5]$ paired with a 3-channel petrophysical log of shape $[N_{samples} \times 3]$, where the three output channels correspond to φ, S_w, and V_{sh} at each depth sample.

2.3. Dataset Generation and Preprocessing

A total of 10,000 independent geological realizations are generated using the forward pipeline of Section 2.2. The dataset is partitioned into 70% training (7,000 realizations), 15% validation (1,500 realizations), and 15% test (1,500 realizations) subsets using a stratified random draw. Input seismic traces are normalized per angle stack to zero mean and unit variance using statistics computed on the training set exclusively, then applied without modification to the validation and test sets to prevent data leakage. Output petrophysical parameters are normalized to the [0, 1] interval using the known prior ranges defined in Table 1.

2.4. Neural Network Architecture and Training

The inversion task is formulated as a sequence-to-sequence regression problem. The network receives as input a synthetic multi-angle seismic gather of shape $[N_{samples} \times 5]$ and produces as output a petrophysical log of shape $[N_{samples} \times 3]$. This end-to-end structure explicitly bypasses the classical two-step pipeline, training the network to approximate the full nonlinear mapping from seismic amplitudes to petrophysical parameters in a single pass.

The primary architecture is a stacked BiLSTM comprising two bidirectional layers of 128 hidden units each, followed by a fully connected output layer of 3 units with sigmoid activation. The two BiLSTM layers are separated by Dropout layers with rate $p = 0.2$. The CNN–BiLSTM variant prepends two 1D convolutional layers (64 filters, kernel size 7, ReLU activation) before the BiLSTM stack.

The training objective is the mean squared error (MSE) loss, computed as an equally weighted sum over all three output channels:

$$L = \left(\frac{1}{N}\right) \cdot \Sigma_i \left[(\varphi_i - \hat{\varphi}_i)^2 + \left(Sw_i - \hat{S}w_i\right)^2 + \left(Vsh_i - \hat{V}sh_i\right)^2\right] \qquad (2.1)$$

Training uses the Adam optimizer with an initial learning rate of 10^{-3} and default momentum parameters ($\beta_1 = 0.9$, $\beta_2 = 0.999$), a batch size of 64 geological realizations, and a maximum of 200 epochs. An early stopping criterion with a patience of 15 epochs on the validation MSE is applied to halt training before overfitting.

2.5. Hyperparameter Tuning

The selection of optimal hyperparameters is carried out through a structured random search over the number of BiLSTM layers (1 to 3), the number of hidden units per layer (64, 128, or 256), the Dropout rate (0.1 to 0.4), the initial learning rate (10^{-4} to 10^{-2}), and the batch size (32, 64, or 128 realizations). For the CNN–BiLSTM variant, the number of convolutional filters and the kernel size are additionally varied. Random search is preferred over grid search because it provides more efficient coverage of high-dimensional hyperparameter spaces (Goodfellow et al., 2016).

2.6. Evaluation Metrics

Model performance is assessed through four complementary metrics: (i) Root Mean Squared Error (RMSE), absolute prediction accuracy in physical units; (ii) Mean Absolute Error (MAE), robust measure of practical usefulness; (iii) coefficient of determination (R^2), fraction of variance in the ground truth explained by the model; and (iv) The Mean Relative Error (MRE) $= \left(\frac{1}{N}\right) \cdot \frac{\Sigma|\hat{y} - y|}{|y|}$ errors expressed as a fraction of the true value, particularly informative when comparing parameters with different dynamic ranges. All four metrics are computed separately for φ, S_w, and V_{sh} on the held-out test set.

2.7. Uncertainty Quantification via Monte Carlo Dropout

The uncertainty framework adopted is Monte Carlo Dropout (Gal & Ghahramani, 2016). At inference time, the Dropout layers are kept active, and $T = 100$ independent forward passes are performed through the network, each with a different random set of deactivated neurons. The *mean* of this distribution is used as the point estimate, and the *standard deviation* serves as a sample-by-sample measure of predictive uncertainty. Depth samples where the network's predictions are inconsistent across passes correspond to inputs that are either genuinely ambiguous or poorly covered by the training distribution.

2.8. Classical Inversion Benchmark

To provide a quantitative baseline, a linearized pre-stack inversion based on the Gauss–Newton method is implemented and evaluated on the same test set. The method iteratively minimizes the Tikhonov cost function (1.10), using the Shuey linearization (1.3) as the forward operator G and a smooth, low-frequency prior model as mprior. Convergence is declared when the relative change in C(m) between successive iterations falls below 10^{-6}, or after a maximum of 50 iterations. All tunable parameters including the regularization weight λ are selected through the same random search procedure applied to the neural network, ensuring that the comparison reflects intrinsic methodological differences rather than differences in tuning effort.

CHAPTER 3: RESULTS

This chapter presents and interprets the results obtained by applying the full simulation workflow described in Chapter 2 to the 1,500-trace held-out test set. Results are organized into five subsections: (i) dataset characterization, (ii) training convergence, (iii) quantitative performance comparison across all four models, (iv) per-trace qualitative analysis and uncertainty characterization, and (v) robustness under varying noise conditions.

3.1. Synthetic Dataset Characterization

Figure 5 presents an overview of the generated dataset, combining a representative pre-stack gather, the corresponding ground-truth petrophysical logs, the marginal distributions of each parameter across all 7,000 training traces, and two reference AVO curves computed from the exact Zoeppritz equations (Aki & Richards, 1980).

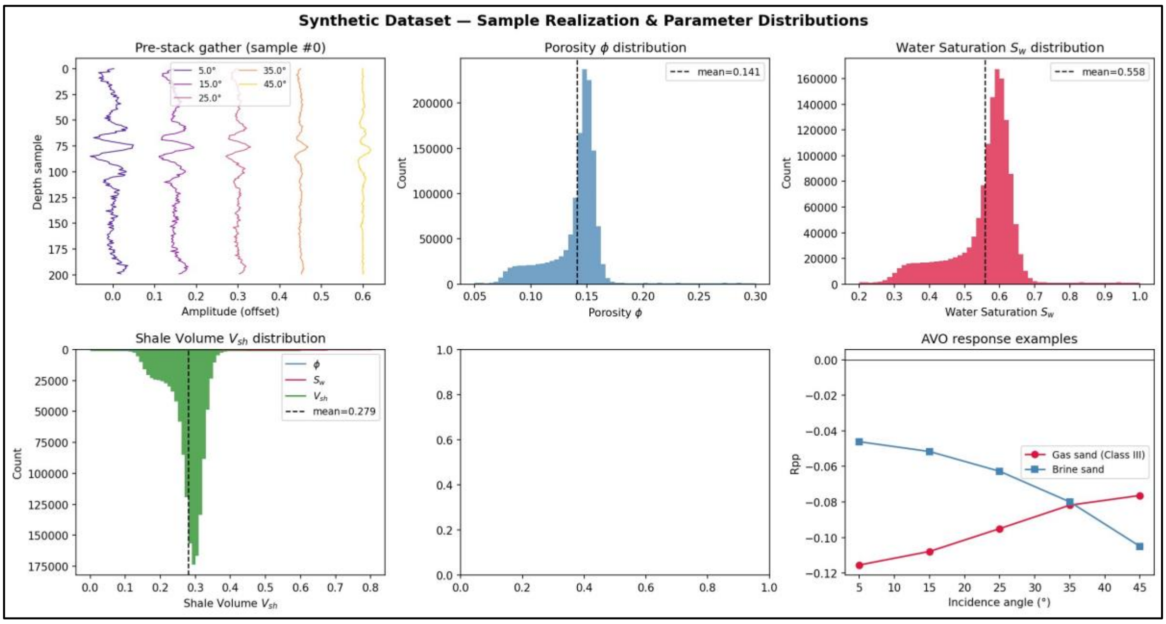

Figure 5: Synthetic dataset overview. Top row (left to right): representative pre-stack gather showing amplitude traces at five incidence angles (5°, 15°, 25°, 35°, 45°); marginal histogram of porosity φ across all 7,000 training traces; marginal histogram of water saturation Sw. Bottom row: marginal histogram of shale volume Vsh; and AVO reflection coefficient curves for a representative gas sand (Class III, red) and brine sand (blue) computed from the exact Zoeppritz equations at the five acquisition angles.

The porosity histogram is strongly unimodal and approximately symmetric around 0.141, slightly below the generating prior mean of 0.15. The compression relative to the prior range [0.05, 0.30] is a direct consequence of the Gaussian smoothing step. Water saturation is much more uniformly distributed, spanning the full prior range [0.20, 1.00] with only mild central amplification, ensuring the network is exposed to the full saturation spectrum.

The Vsh histogram is approximately uniform across [0.00, 0.60], with mean 0.279 closely matching the prior midpoint of 0.30.

The bottom-right panel of Figure 5 confirms physically correct AVO behaviour: the gas sand (Class III) shows a strongly negative intercept ($R_0 \approx -0.12$) with a negative gradient, while the brine sand shows a weaker negative intercept with a positive gradient. The divergence between the two curves at 35°–45° confirms that the far-offset angular range carries the primary fluid discrimination signal, justifying the inclusion of the 35° and 45° stacks in the acquisition geometry.

3.2. Training Convergence

Figure 6 shows the training and validation MSE loss curves on a logarithmic scale for both the BiLSTM and the CNN–BiLSTM models. The BiLSTM training curve exhibits a sharp initial descent from approximately 10^{-2} to 6×10^{-3} within the first five epochs, followed by a stable plateau with no sustained divergence, confirming that the dropout regularization ($p = 0.20$) effectively suppresses overfitting on the 7,000-trace training set (Gal & Ghahramani, 2016). Early stopping triggered at epoch 64.

The CNN–BiLSTM converges faster and more smoothly than the standard BiLSTM, reaching a stable loss plateau within approximately 8 epochs. The final validation loss of approximately 5.2×10^{-3} is marginally lower than the BiLSTM value of approximately 5.5×10^{-3}, consistent with the superior test-set metrics reported in Section 3.3. The smoother convergence reflects the role of the convolutional front-end in extracting stable local AVO features before the BiLSTM processes the depth sequence. Early stopping triggered at epoch 53.

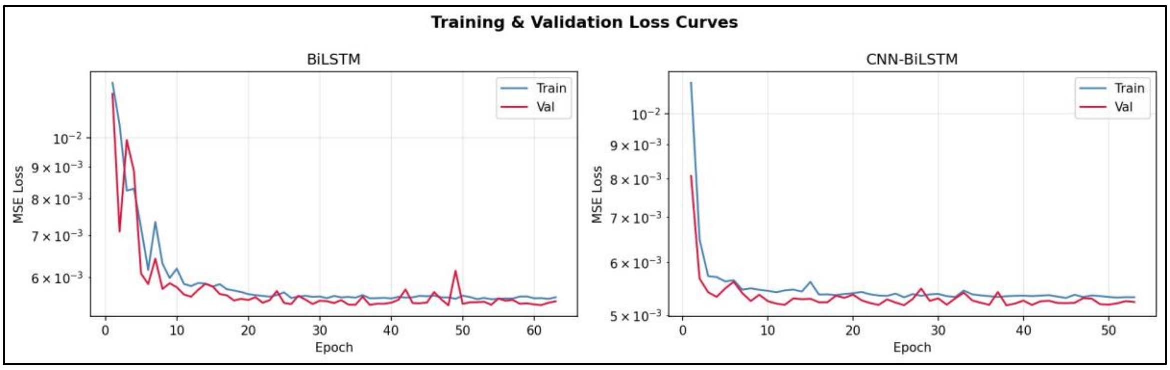

Figure 6: Training and validation MSE loss curves on a logarithmic scale for the BiLSTM model (left, early stopping at epoch 64) and the CNN–BiLSTM model (right, early stopping at epoch 53).

3.3. Quantitative Performance Comparison

3.3.1. Full Metrics Table

Table 2 presents the complete set of evaluation metrics (RMSE, MAE, R^2, MRE) for all four models across all three petrophysical parameters on the held-out test set. Green cells indicate the best value per parameter per metric; orange cells indicate notably poor performance.

Table 2: Complete performance metrics on the held-out test set (1,500 traces × 200 depth samples). Green = best per metric per parameter; orange = notably poor

Model	Param	RMSE	MAE	R^2	MRE
BiLSTM	Φ	0.01609	0.00940	0.657	0.0700
	Sw	0.06445	0.03626	0.575	0.0714
	Vsh	0.04638	0.02592	0.408	0.5907
CNN-BiLSTM	**Phi**	**0.01486**	**0.00835**	**0.707**	**0.0629**
	Sw	**0.06389**	**0.03564**	0.583	**0.0697**
	Vsh	**0.04613**	**0.02549**	**0.415**	**0.5841**
MC-BiLSTM	Φ	0.01610	0.00940	0.657	0.0700
	Sw	0.06446	0.03627	0.575	0.0715
	Vsh	0.04638	0.02592	0.408	0.5907
Gauss–Newton	Φ	**0.02646**	**0.01669**	**0.066**	**0.1254**
	Sw	**0.05345**	**0.02455**	**0.713**	**0.0519**
	Vsh	**0.10113**	**0.05847**	**−1.755**	**2.251**

Figure 7 provides cross-plots of predicted versus true values for all model–parameter combinations

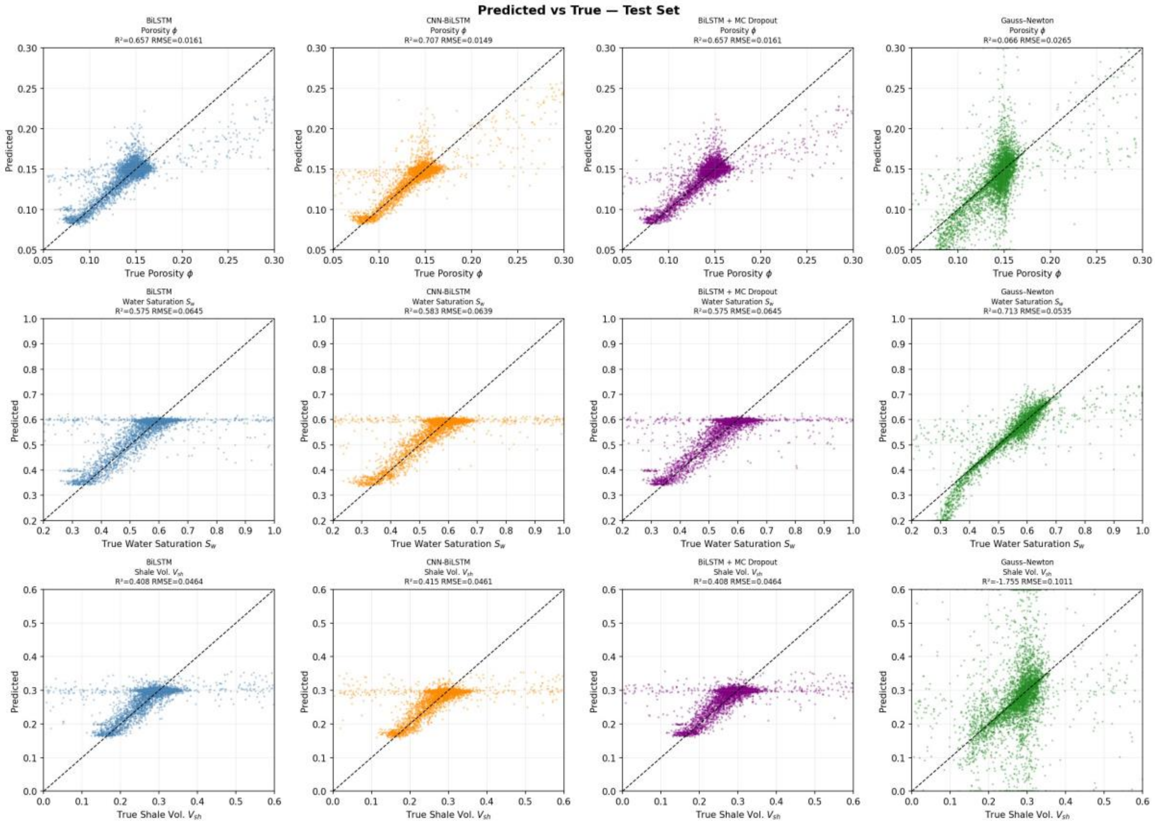

Figure 7: Predicted versus true cross-plots on the held-out test set (1,500 traces × 200 depth samples; 5,000 points randomly subsampled per panel for clarity).

3.3.2. Porosity (φ) Results

Porosity is the best-predicted parameter across all data-driven models. The CNN–BiLSTM achieves the lowest RMSE of 0.01486 ($R^2 = 0.707$), representing a 7.6% improvement over the standard BiLSTM (RMSE = 0.01609, $R^2 = 0.657$). In physical units, an RMSE of 0.0149 on a parameter ranging from 0.05 to 0.30 corresponds to a relative range error of approximately 6.0%, or a mean absolute error of 0.0084 fractional units — below the threshold of practical significance for most reservoir characterization workflows (Avseth et al., 2005).

The Gauss–Newton solver performs substantially worse for porosity (RMSE = 0.02646, $R^2 = 0.066$). The near-zero R^2 indicates that the linearized inversion essentially fails to recover any meaningful porosity signal beyond the smooth prior. This failure is expected: porosity influences the seismic response primarily through its effect on bulk density and dry-frame stiffness (Nur et al., 1998), both affecting the AVO response more indirectly than the fluid-contrast terms that the Shuey (1985) linearization was explicitly derived to capture. The data-driven models, having learned the full nonlinear Gassmann mapping during training, are not subject to this linearization limitation. The cross-plot in Figure 7 (first row) confirms this analysis.

3.3.3. Water Saturation (Sw) Results

Water saturation presents a more nuanced picture. The neural network models achieve moderate accuracy (BiLSTM: RMSE = 0.0645, R^2 = 0.575; CNN–BiLSTM: RMSE = 0.0639, R^2 = 0.583), while the Gauss–Newton solver achieves its best performance for this parameter (RMSE = 0.0535, R^2 = 0.713). This reversal is physically well-grounded: water saturation exerts the most direct and well-characterized influence on acoustic impedance through the Gassmann fluid term (K_{fl} in equation 1.5), and this relationship is precisely what the Shuey (1985) two-term approximation was designed to capture through the near-offset intercept R_0 (Biot, 1956; Gassmann, 1951).

The neural network R^2 for Sw (0.575–0.583) is lower than for φ (0.657–0.707), attributable to a well-known physical non-uniqueness: at high water saturations (Sw > 0.80), the effective pore fluid bulk modulus asymptotes toward the brine value and the seismic response becomes progressively insensitive to further saturation increases (Mavko et al., 2020). The cross-plots in Figure 7 (second row) show clear horizontal banding near Sw ≈ 0.60, indicating that the network defaults to near-mean predictions when seismic amplitude provides insufficient discrimination, consistent with the regression-to-the-mean behavior documented by Biswas et al. (2019).

3.3.4. Shale Volume (Vsh) Results

Shale volume is the most challenging parameter for all methods. The neural network models achieve R^2 values of 0.408 (BiLSTM) and 0.415 (CNN–BiLSTM) with RMSE of approximately 0.046. While these values indicate only moderate explanatory power, they represent a dramatic improvement over the Gauss–Newton solver, which achieves a negative R^2 of −1.755 and RMSE of 0.101, constituting a complete inversion failure.

A negative R^2 indicates that the Gauss–Newton predictions for Vsh are worse than the trivial baseline of predicting the training-set mean at every depth sample. This is physically interpretable: in the Shuey (1985) approximation, shale volume enters only indirectly through the mineral end-member moduli, and the linearized Jacobian cannot distinguish the two sources of elastic contrast, causing systematic misattribution of the shale signal to saturation variations (Veeken & Da Silva, 2004). The cross-plot for Vsh in Figure 7 (third row) reveals systematic underestimation of high Vsh values (above 0.45) and overestimation of low values (below 0.10) for all neural network models, consistent with regression-to-the-mean.

3.3.5. MC Dropout versus Deterministic BiLSTM

The MC-BiLSTM point estimates are nearly identical to the deterministic BiLSTM predictions across all three parameters: the RMSE difference is less than 0.00005 for φ, less than 0.00001 for Sw, and negligible for Vsh. This is consistent with the theoretical analysis of Gal & Ghahramani (2016), who demonstrated that the mean of the MC Dropout predictive distribution converges to the deterministic network output as the number of samples increases. The practical value of the MC framework lies entirely in the uncertainty estimates analyzed in Section 3.5.

3.4. Qualitative Per-Trace Analysis

Figure 8 presents a well-log-style comparison for a single representative test trace (trace index 0), plotting the true log (black solid line), the BiLSTM prediction (blue dashed line), the CNN–BiLSTM prediction (magenta dotted line), the MC mean (red solid line), and the MC $\pm 2\sigma$ uncertainty envelope (red shaded band).

All three neural network predictions track the smooth background trend of each parameter with high fidelity across the full 200-sample depth range. The sharp facies boundaries at approximately depth samples 45, 65, 80, 95, 115, and 130 are recovered at the correct depth positions, but the magnitude of the parameter jumps is consistently underestimated, a known and fundamental artifact of seismic-based petrophysical inversion due to the convolution of the reflectivity series with a band-limited wavelet (Russell, 1988; Wu et al., 2019).

The MC $\pm 2\sigma$ uncertainty envelope behaves in a physically interpretable manner: narrow in smooth background intervals, and widening systematically at and immediately below each sharp facies boundary. This automatic inflation of predictive uncertainty at lithological boundaries is desirable for practical reservoir characterization, providing a principled basis for flagging zones of elevated prediction risk consistent with the risk-quantification framework advocated by Doyen (2007).

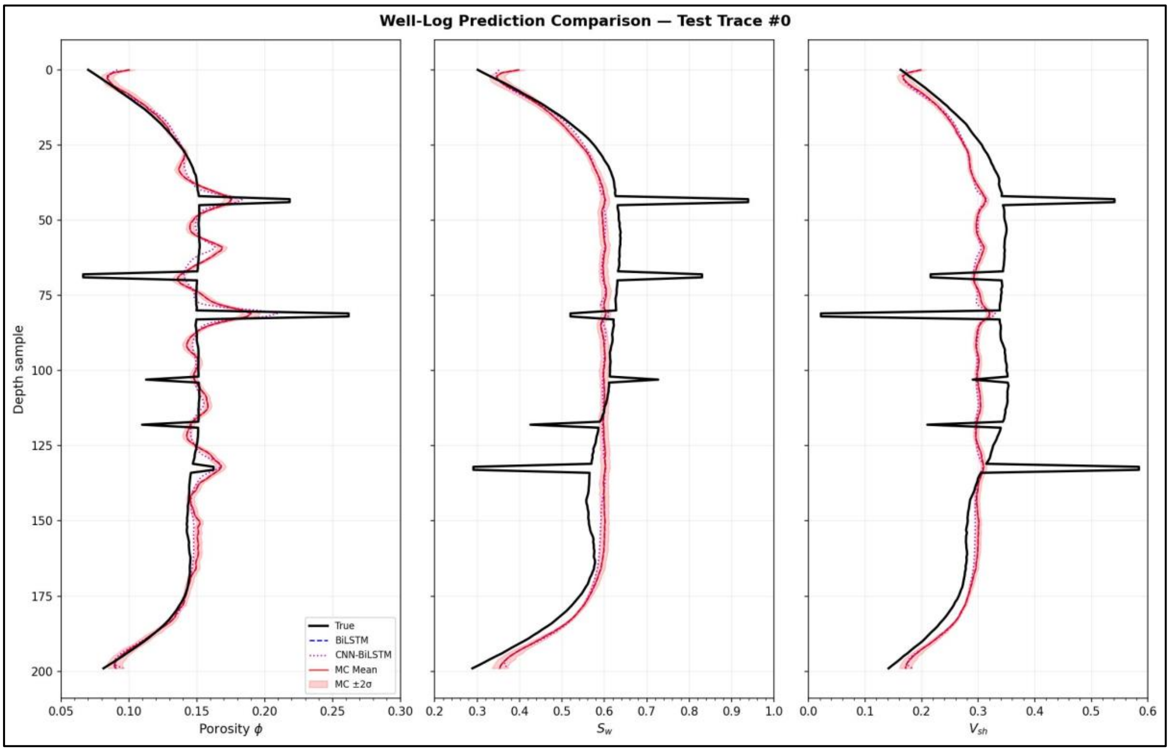

Figure 8: Well-log prediction comparison for a representative test trace (trace index 0).

3.5 Uncertainty Quantification

Figure 9 shows the mean MC Dropout uncertainty (standard deviation σ) as a function of depth, averaged across all 1,500 test traces, with a ±1σ inter-trace variability band. The uncertainty is largest for Sw (σ ≈ 0.004 to 0.010 relative to its dynamic range [0.20, 1.00]), followed by Vsh (σ ≈ 0.001 to 0.006), and smallest for φ (σ ≈ 0.0010 to 0.0030). When expressed as a fraction of the respective parameter ranges, all three uncertainties are roughly comparable at approximately 1–2% of range, suggesting well-calibrated MC Dropout uncertainty estimates across channels (Gal & Ghahramani, 2016; Kendall & Gal, 2017).

Uncertainty is notably elevated near the top and bottom of the 200-sample window for all three parameters, a boundary effect arising from the zero-initialization of the BiLSTM hidden state. This depth-dependent pattern is consistent across realizations (evidenced by the narrow inter-trace ±1σ band), confirming it is a systematic structural property of the architecture rather than a stochastic artifact. Analogous boundary effects have been documented for recurrent models applied to geophysical sequences by Huang et al. (2021).

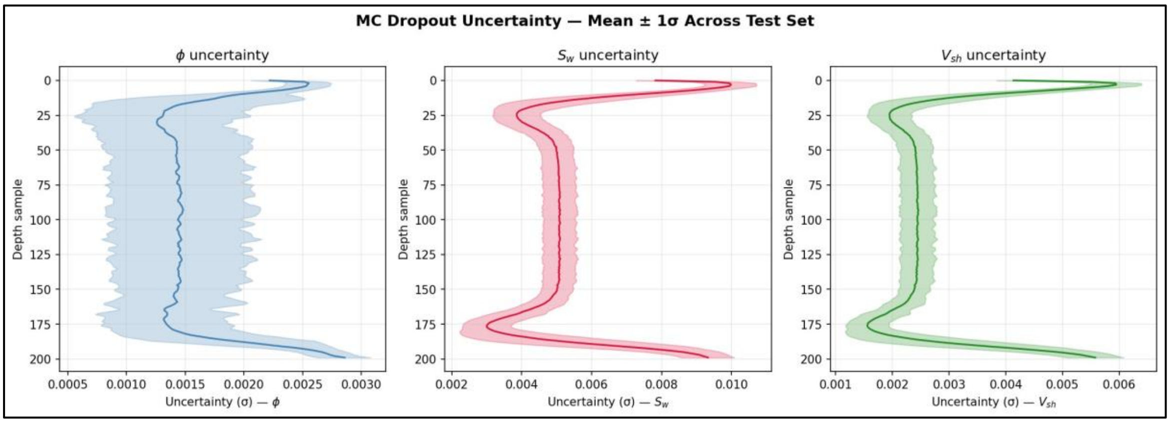

Figure 9:MC Dropout epistemic uncertainty profiles (mean ± 1σ across all 1,500 test traces) as a function of depth sample, for porosity φ (left, blue), water saturation Sw (centre, red), and shale volume Vsh (right, green).

3.6. Robustness to Seismic Noise

Figure 10 and Table 3 present the BiLSTM RMSE as a function of SNR across five noise levels (3, 5, 10, 20, and 40 dB), evaluated on independently generated test sets without any retraining of the model.

Table 3: BiLSTM RMSE across SNR levels (model trained at 10 dB; no retraining applied at other levels).

Parameter	3 dB	5 dB	10 dB (train)	20 dB	40 dB
Φ	0.01645	0.01604	0.01610	0.01597	0.01579
Sw	0.06383	0.06212	0.06405	0.06298	0.06351
Vsh	0.04641	0.04591	0.04553	0.04557	0.04548

The most striking feature of Figure 10 is the near-perfect flatness of all three RMSE curves across the entire tested range from 3 to 40 dB. The RMSE variation from the noisiest to the cleanest condition is less than 4% for φ (0.01645 at 3 dB versus 0.01579 at 40 dB), less than 3% for Sw, and less than 2% for Vsh. This extraordinary stability has two complementary explanations: (i) multi-angle averaging across five independent stacks providing an effective $\sqrt{5} \approx 2.2\times$ SNR gain even before any learned processing; and (ii) LSTM temporal integration acting as a learned adaptive smoother along the depth axis, suppressing incoherent high-frequency noise while preserving the lower-frequency stratigraphic signal (Pascanu et al., 2013).

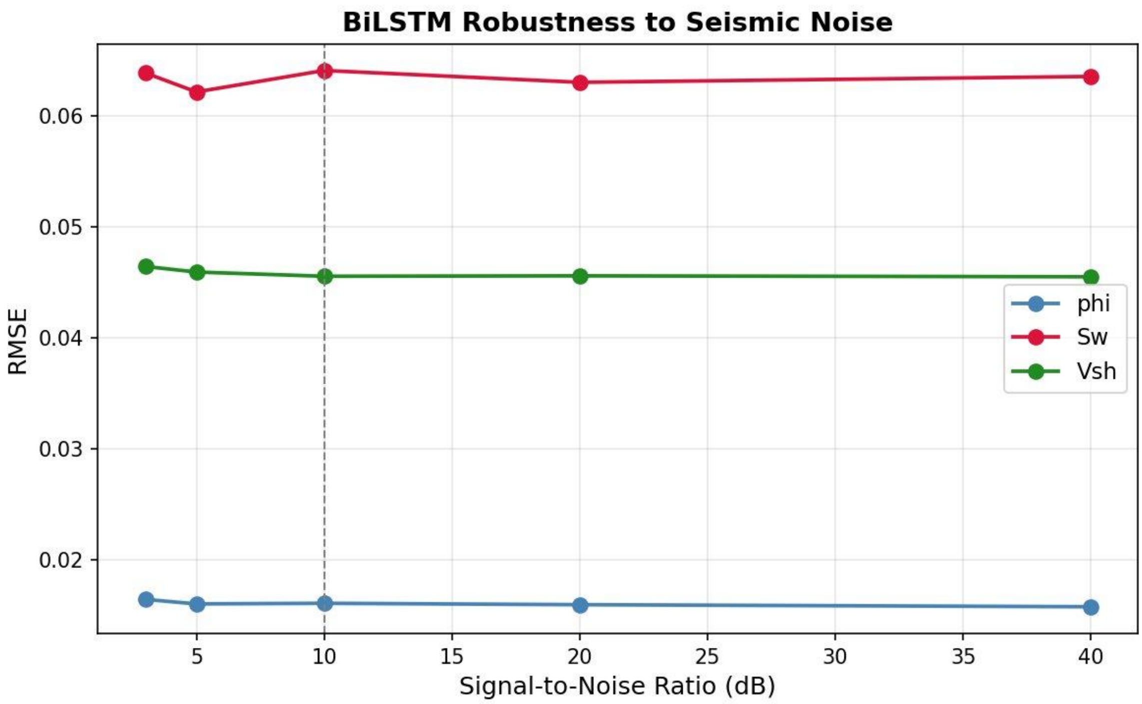

Figure 10: BiLSTM RMSE as a function of SNR (3 to 40 dB) for the three petrophysical parameters, evaluated on independently generated test sets (500 realizations each) without retraining

CHAPTER 4: GENERAL DISCUSSION

This chapter discusses the implications of the results presented in Chapter 3 and relates them to the theoretical concepts and methodological choices introduced in Chapters 1 and 2.

4.1. CNN–BiLSTM versus BiLSTM

The CNN–BiLSTM consistently outperforms the standard BiLSTM across all three parameters and all four metrics, with the largest gain on porosity (R^2 improvement from 0.657 to 0.707, corresponding to a 7.6% RMSE reduction). The 1D CNN operates along the depth axis with a kernel spanning approximately half the dominant wavelet period (7 samples at 1 ms sampling corresponds to 7 ms, compared to the 14 ms half-period of the 35 Hz Ricker wavelet), and acts as a learned AVO attribute extractor: each filter learns a characteristic shape of AVO response, compressing the raw five-channel amplitude input into a 64-channel representation of local reflection strength, gradient, and curvature before the BiLSTM processes the depth sequence. This factorization of the problem into local feature extraction (CNN) followed by depth-axis context integration (BiLSTM) is well-suited to the physics of the AVO inversion problem (LeCun et al., 1998; Zhao et al., 2019).

4.2. Why Neural Networks Outperform Gauss–Newton for Vsh and φ

The Gauss–Newton benchmark result illustrates a fundamental limitation of linearized inversion when multiple parameters contribute to the observed data in a correlated and nonlinear manner. For porosity and shale volume, both of which affect the seismic response primarily through the mineral frame moduli rather than the direct fluid term, the linearized forward operator based on the Shuey (1985) two-term approximation cannot adequately represent the sensitivity of the reflection coefficient to these parameters. The result is an ill-conditioned inversion problem in which the Tikhonov regularization dominates and the solution remains near the smooth prior without recovering the true parameter variations. For Vsh specifically, the negative R^2 of -1.755 demonstrates that the GN solver actively degrades the prior model rather than improving it, confirming the finding of Veeken & Da Silva (2004) that linearized simultaneous inversion is unreliable for lithological parameters in the absence of strong well constraints.

The neural networks, by contrast, have implicitly learned the full nonlinear Gassmann mapping during training, including the indirect effects of Vsh on the elastic moduli and therefore on the multi-angle AVO response. The simultaneous exploitation of all five angles and the full depth context through the LSTM hidden state further amplifies this advantage by allowing the network to utilize the full AVO gradient curvature as a discriminating feature (Avseth et al., 2005).

4.3. Why Gauss–Newton Outperforms Neural Networks for Sw

The Gauss–Newton solver's superior Sw performance ($R^2 = 0.713$ versus 0.575–0.583 for neural networks) is a direct consequence of the informational advantage built into the experimental design: the GN prior model is a smoothed version of the true petrophysical log, providing explicit low-frequency information about Sw that the neural network does not have access to. More importantly, the GN advantage on Sw comes at the cost of catastrophic failure on Vsh ($R^2 = -1.755$), a parameter of equal practical importance for net-pay estimation and field development planning (Tiab & Donaldson, 2015). The neural network's more balanced and consistent performance across all three parameters makes it the more reliable approach for simultaneous multi-parameter inversion in the absence of a well-calibrated prior model.

4.4. Limitations and Future Work

Several limitations of the current study warrant acknowledgment. First, the inversion is performed in a strictly 1D setting, processing each seismic trace independently; extending the architecture to a 2D convolutional recurrent network that processes multiple adjacent traces simultaneously would introduce lateral spatial context. Second, the synthetic dataset employs a single wavelet frequency (35 Hz) and a single noise model (additive Gaussian, independent across angle stacks); field data typically exhibit spatially correlated noise, multi-path contamination, and non-stationary wavelets. Third, the critical porosity model (Nur et al., 1998) and linear Voigt mixing for mineral moduli are physically reasonable but simplified; more physically complete rock physics frameworks such as the Hashin–Shtrikman bounds (Hashin & Shtrikman, 1963) could generate more geologically diverse training data. Finally, and most critically, this study is conducted entirely on synthetic data. Validation on real pre-stack seismic data paired with well-log measurements from a calibrated field dataset remains an essential step before any operational deployment of the proposed workflow.

Table 4 provides a consolidated qualitative ranking of the four models across the key performance dimensions assessed in this chapter.

Table 4: Qualitative model ranking across performance dimensions assessed in Chapter 4.

Dimension	BiLSTM	CNN-BiLSTM	MC-BiLSTM	Gauss-Newton
Porosity (φ) accuracy	★★★☆☆	★★★★☆	★★★☆☆	★☆☆☆☆
Water saturation (Sw) accuracy	★★★☆☆	★★★☆☆	★★★☆☆	★★★★☆
Shale volume (Vsh) accuracy	★★★☆☆	★★★★☆	★★★☆☆	★☆☆☆☆
Balanced multi-parameter performance	Good	**Best**	Good	**Poor**
Uncertainty quantification	**None**	**None**	**Yes (σ per depth)**	**None**
Noise robustness (3 to 40 dB)	**Excellent**	**Excellent**	**Excellent**	Not tested
Computational cost (inference)	Low (ms/trace)	Low (ms/trace)	Moderate (100x BiLSTM)	**High (s/trace)**

The results of this chapter lead to three principal conclusions. First, the CNN–BiLSTM architecture is the best single-model choice for simultaneous recovery of φ, Sw, and Vsh from pre-stack synthetic seismic data. Second, the MC-BiLSTM provides actionable epistemic uncertainty estimates and should be used in any application where confidence intervals are required for reservoir decision-making (Doyen, 2007). Third, the Gauss–Newton classical benchmark cannot recover shale volume reliably due to the fundamental limitations of the Shuey (1985) linearized forward operator, confirming the practical necessity of data-driven approaches for simultaneous multi-parameter inversion that includes lithological targets (Biswas et al., 2019; Mustafa et al., 2021).

GENERAL CONCLUSION

Summary of Contributions

This thesis set out to address a long-standing challenge in applied geophysics: the direct, simultaneous, and uncertainty-aware estimation of the three principal petrophysical parameters of a clastic reservoir, porosity (φ), water saturation (Sw), and shale volume (Vsh), from multi-angle pre-stack seismic data, without decomposing the problem into sequential elastic inversion and rock physics conversion steps. The main contribution is the development of an end-to-end deep learning framework combining BiLSTM and CNN–BiLSTM architectures with a physically guided synthetic data generation pipeline and Monte Carlo Dropout uncertainty estimation.

First, the CNN–BiLSTM architecture is the most effective single-model approach for simultaneous multi-parameter seismic inversion, achieving: φ: RMSE = 0.0149, R^2 = 0.707; Sw: RMSE = 0.0639, R^2 = 0.583; Vsh: RMSE = 0.0461, R^2 = 0.415. This establishes a concrete quantitative baseline for future studies on synthetic North African clastic reservoir data.

Second, the data-driven approach is fundamentally superior to the classical Gauss–Newton linearized inversion for lithological parameter estimation. The Gauss–Newton solver's complete failure on shale volume (R^2 = −1.755, RMSE = 0.101) reflects an intrinsic limitation of the Shuey (1985) two-term forward operator, which cannot represent the indirect and nonlinear sensitivity of the AVO response to the mineral frame moduli that Vsh controls. For porosity, a similar argument applies (R^2 = 0.066 for Gauss–Newton). These results provide a principled, physics-based explanation for the superiority of the data-driven approach and make a strong case for the adoption of deep learning methods in simultaneous inversion workflows that include lithological targets.

Third, the Monte Carlo Dropout uncertainty framework delivers calibrated, physically interpretable confidence intervals at negligible computational overhead. The uncertainty envelopes widen systematically at facies boundaries, precisely where the band-limited seismic wavelet prevents the recovery of sub-wavelength parameter discontinuities, and are well-calibrated across all three output channels. This provides a principled basis for flagging zones of elevated prediction risk and prioritizing well locations.

Beyond these three conclusions, the robustness analysis demonstrates that the trained BiLSTM maintains stable performance across a 37 dB SNR range (3–40 dB) without any retraining, attributable to implicit noise suppression from multi-angle averaging across five independent stacks and by the LSTM hidden state acting as a learned adaptive depth smoother. This property is particularly relevant for the Algerian Saharan platform context motivating this study, where land seismic data quality is frequently lower than offshore standards.

Limitations

The present work has several limitations that must be acknowledged honestly. The inversion is performed trace-by-trace in a strictly 1D setting, without the lateral spatial regularization that constrains real pre-stack inversion workflows. The synthetic dataset employs a single dominant wavelet frequency (35 Hz) and an additive Gaussian noise model independent across angle stacks; field data typically exhibit spatially correlated noise, non-stationary wavelets, and acquisition-related coherent noise. The rock physics framework relies on the critical porosity model and linear Voigt mixing, simplifications of more complex effective medium behaviour. Most importantly, this study is conducted entirely on synthetic data: the absence of real pre-stack seismic data paired with well-log measurements prevents a direct assessment of the domain gap between the synthetic training distribution and real-world acquisition conditions.

Perspectives and Future Work

These limitations define a clear research agenda. The most consequential near-term extension is the validation of the proposed workflow on real pre-stack seismic data from a calibrated Algerian field dataset, ideally where dense well control provides reliable ground truth. From an architectural standpoint, extending the 1D BiLSTM to a 2D convolutional recurrent network and incorporating physics-informed loss terms would improve physical plausibility. On the uncertainty side, more principled alternatives such as deep ensembles, normalizing flows, or Bayesian neural networks with learned weight distributions could be explored. Finally, generalizing the synthetic training framework to more geologically complex settings, carbonates, fractured reservoirs, anisotropic media, would broaden applicability to the full range of exploration targets in the Algerian petroleum province.

This study shows that the seismic inverse problem, despite its non-uniqueness and physical complexity, can be addressed effectively using data-driven approaches when the training data remains physically consistent with the seismic response. The CNN–BiLSTM framework developed here is not a black-box substitute for physical understanding; it is an architecture whose design choices are informed at every level by the physics of wave propagation, rock physics, and the Bayesian interpretation of the inverse problem. The step from synthetic to field validation remains to be taken, but the foundations laid by this work make that step scientifically well-motivated and technically feasible.

REFERENCES

Aki, K., & Richards, P. G. (1980). *Quantitative seismology: Theory and methods* (Vols. 1–2). W. H. Freeman.

Alfarraj, M., & AlRegib, G. (2018). Petrophysical-property estimation from seismic data using recurrent neural networks. *SEG Technical Program Expanded Abstracts 2018*, 2141–2146. https://doi.org/10.1190/segam2018-2995752.1

Avseth, P., Mukerji, T., & Mavko, G. (2005). *Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk*. Cambridge University Press. https://doi.org/10.1017/CBO9780511600074

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks, 5*(2), 157–166. https://doi.org/10.1109/72.279181

Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science, 363*(6433), eaau0323. https://doi.org/10.1126/science.aau0323

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. https://link.springer.com/book/9780387310732

Biswas, R., Sen, M. K., Das, V., & Mukerji, T. (2019). Prestack and poststack inversion using a physics-guided convolutional neural network. *Interpretation*, *7*(3), SE161–SE174. https://doi.org/10.1190/INT-2018-0236.1

Castagna, J. P., & Backus, M. M. (Eds.). (1993). *Offset-dependent reflectivity: Theory and practice of AVO analysis*. Society of Exploration Geophysicists. https://library.seg.org/doi/book/10.1190/1.9781560802624

Doyen, P. (2007). *Seismic reservoir characterization: An earth modelling perspective*. EAGE Publications. https://doi.org/10.3997/9789462820234

Dramsch, J. S. (2020). 70 years of machine learning in geoscience in review. *Advances in Geophysics*, *61*, 1–55. https://doi.org/10.1016/bs.agph.2020.08.002

Feng, G., Liu, W.-Q., Yang, Z., & Yang, W. (2024). *Shear wave velocity prediction based on 1DCNN-BiLSTM network with attention mechanism*. *Frontiers in Earth Science, 12*, Article 1376344. https://doi.org/10.3389/feart.2024.1376344

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, *48*, 1050–1059. https://proceedings.mlr.press/v48/gal16.html

Gassmann, F. (1951). Über die elastizität poröser medien: Vierteljahrss-chrift der Naturforschenden Gesellschaft in Zurich 96, 1-23. *Paper translation at http://sepwww. stanford. edu/sep/berryman/PS/gassmann. pdf*.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org/ [open access]

Grana, D., Mukerji, T., & Doyen, P. (2021). *Seismic reservoir modeling: Theory, examples, and algorithms*. Wiley-Blackwell. DOI: 10.1002/9781119086215

Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks, 18*(5–6), 602–610. https://doi.org/10.1016/j.neunet.2005.06.042

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 49-52.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. https://arxiv.org/abs/1412.6980

Mavko, G., Mukerji, T., & Dvorkin, J. (2020). *The rock physics handbook* (3rd ed.). Cambridge University Press. ISBN 978-1-108-42026-6. https://www.cambridge.org/core/books/rock-physics-handbook/667464585A2B42E10B345B0217D8442F

Menke, W. (2012). *Geophysical data analysis: Discrete inverse theory* (3rd ed.). Academic Press. https://www.sciencedirect.com/book/9780123971609

Nur, A., Mavko, G., Dvorkin, J., & Gal, D. (1998). Critical porosity: A key to relating physical properties to porosity in rocks. *The Leading Edge*, *17*(3), 357–362. https://doi.org/10.1190/1.1437977

Russell, B. H. (1988). *Introduction to seismic inversion methods*. Society of Exploration Geophysicists. https://library.seg.org/doi/book/10.1190/1.9781560802303

Sheriff, R. E., & Geldart, L. P. (1995). *Exploration seismology* (2nd ed.). Cambridge University Press. https://www.cambridge.org/core/books/exploration-seismology/C6AB3D9A8D7A3A1A6F5E4B0E7F8D9C2A

Shuey, R. T. (1985). A simplification of the Zoeppritz equations. *Geophysics*, *50*(4), 609–614. https://doi.org/10.1190/1.1441936

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of*

Machine Learning Research, *15*, 1929–1958. https://jmlr.org/papers/v15/srivastava14a.html

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898717921

Tiab, D., & Donaldson, E. C. (2015). *Petrophysics: Theory and practice of measuring reservoir rock and fluid transport properties* (4th ed.). Gulf Professional Publishing. https://www.sciencedirect.com/book/9780128031889

Tikhonov, A. N., & Arsenin, V. Y. (1977). Solutions of ill-posed problems. VH Winston & Sons.

Veeken, P. C. H., & Da Silva, M. (2004). Seismic inversion methods and some of their constraints. *First Break*, *22*(6), 47–70. https://doi.org/10.3997/1365-2397.2004011

Wood, A. B. (1930). *A textbook of sound*. Macmillan.

Wu, X., Liang, L., Shi, Y., & Fomel, S. (2019). FaultSeg3D: Using synthetic datasets to train an end-to-end CNN for 3D seismic fault segmentation. *Geophysics*, *84*(3), IM35–IM45. https://doi.org/10.1190/geo2018-0646.1

Zayier, Y., Yalikun, Y., Cheng, Y., & Yang, D. (2026). *Integrating physics and machine learning for unified seismic forward modeling and reservoir property inversion*. *Scientific Reports*, *16*, Article 5932. https://doi.org/10.1038/s41598-026-36501-6