Imperial College London
Department of Earth Science and Engineering
MSc in Applied Computational Science and Engineering

Independent Research Project
Project Plan

Predicting LWD Compressional and Shear Sonic From Drilling Parameters. A Neural Network Approach.

by
Xiangyu Zhang

Email: xz2730@ic.ac.uk
GitHub username: acse-himzhang
Repository: https://github.com/acse-himzhang?tab=repositories

Supervisors:
Dr. Said Assous

June 2, 2024

Abstract

The project focuses on developing a synthetic well-log generator using machine learning to predict compressional (P-wave) and shear (S-wave) transit times from various drilling parameters. These parameters include torque, rate of penetration, rotations per minute, standpipe pressure, weight on bit, and gallons per minute. The goal is to derive accurate geomechanical characteristics such as Poisson's ratio and Young's modulus, essential for optimizing drilling efficiency and reducing risks. This machine learning-based tool aims to enhance model generalization across different drilling environments and mitigate biases due to human-operated drilling features, thereby improving the reliability and applicability of geomechanical assessments in real-world operations.

Introduction

Literature Review

Sonic logs are a crucial component of well logging, offering valuable data for the characterization of subsurface formations. However, their integration into drilling operations, particularly using Logging-While-Drilling (LWD) tools, is often limited. This literature review examines the reasons behind the phasing of sonic logs, the conventional and advanced methods for predicting sonic data, especially the role of machine learning in enhancing these predictions.

Sonic data can be obtained using LWD tools, but their use during drilling, especially in development wells, is not common due to economic and operational challenges. The harsh drilling environment presents significant limitations, making the integration of sonic tools less feasible. Instead, wireline logging tools are typically employed after drilling, but this approach delays the availability of sonic data, increasing the risk of encountering problems without the necessary information (Gamal, Hany, et al.).

In many wells, particularly in deep-water operations, the cost and time associated with recording shear sonic data are prohibitive. As a result, these logs are often excluded from well logging programs to reduce operational costs (Chakraborty, Shantanu, et al.). The industry frequently relies on empirical correlations to estimate shear slowness from available data, such as compressional velocity or density logs. However, these correlations have limitations and are not universally applicable, leading to potential inaccuracies in rock elastic properties estimation and subsequent financial implications (Chakraborty, Shantanu, et al.).

Traditional techniques for predicting shear slowness involve empirical relationships and rock physics modeling, which require extensive input data and carry multiple assumptions. These methods often involve the sonic log data or laboratory testing on core samples, both of which are costly and time-consuming. Studies have attempted to develop empirical correlations to determine shear wave velocity from compressional velocity, but these models have limited application ranges and are specific to certain lithologies (Alfaraj, Rima T., et al.).

The complexity and data requirements of rock physics models have led to the exploration of machine learning (ML) and deep learning (DL) methods for predicting shear data. DL-based techniques, such as Recurrent Neural Networks (RNN) (Zhong, Ruizhi, et al.), offer simplicity and independence from additional data requirements. These models can automatically analyze relevant data to predict shear slowness with high accuracy.

Several studies have demonstrated the effectiveness of DL in predicting compressional and shear wave slowness. For instance, a DL model using bulk density, gamma-ray, true porosity, and neutron porosity as inputs achieved a correlation coefficient of 0.96 (Gowida, Ahmad, et al.). Another study employed a RNN with nonlinear autoregressive exogenous inputs, utilizing neutron and gamma-ray logs to predict slowness, yielding high accuracy results (Onalo, David, et al.).

Recent research aims to develop novel ML models for predicting sonic data using only surface drilling parameters, which will further reduce the cost and save valuable time, such as weight on bit (WOB), rate of penetration (ROP), drill pipe rotations per minute (RPM), standpipe pressure (SPP), torque (T), and mud flowrate (Qm), etc... These models, including decision trees (DT) and random forests (RF), are designed to operate in real-time drilling environments, providing accurate predictions without the need for extensive rock measurements or laboratory analysis (Gamal, Hany, et al.).

Problem Description

Model generalization problem

While substantial progress has been made in applying machine learning to predict compressional (DTC) and shear (DTS), contemporary studies predominantly focus on specific oil fields. Typically, training data for these models are derived from wells within the same oil field and subsequently tested and validated on a different well in the same geographic area. Although this approach may yield seemingly ideal performance metrics, such as high RMSE and MAPE, it raises concerns about the model's generalization capability. Different oil fields exhibit significant geological variability, which may not be effectively captured by a model trained in such a localized context.

Evidence of poor generalization has been noted in the literature. For instance, studies by Gowida and Elkatatny, Hany Gamal et al. and Rima T. Alfara et al. examined the correlation matrix between various drilling parameters (GPM, WOB, SPP, RPM, ROP, T) and the target variables DTC and DTS across different oil fields. Observations of varying correlation factors indicate that models developed within these studies are likely to face challenges in generalizing to datasets from other oil fields.

A primary objective of this project is to develop a deep learning model that generalizes across various unseen oil fields. To achieve this, the dataset will encompass well logs from multiple oil fields, featuring diverse lithologies and cover more depth. This approach aims to enhance the robustness and applicability of the predictive model across a broader range of geological settings.

	GPM	WOB	SPP	RPM	ROP	Torque	GR
Rima T. Alfaraj et al.	0.07	-0.23	0.42	0.07	0.03	0.05	0.2
Gowida and Elkatatny	-0.1	0.08	0.12	0.23	0.21	0.07	N/A
Hany Gamal et al.	0.27	0.13	-0.06	0.29	0.04	0.04	N/A

	GPM	WOB	SPP	RPM	ROP	Torque	GR
Rima T. Alfaraj et al.	-0.09	-0.1	0.21	0.07	0.03	0.05	0.2
Gowida and Elkatatny	-0.12	-0.17	0.17	0.22	-0.08	-0.07	N/A
Hany Gamal et al.	0.51	0.37	0.16	0.32	0.29	0.31	N/A

Figure 1: Correlation factor for different features collected from three studies

Human influenced drilling feature

Drilling features that can be measured from the surface are the most convenient to be gathered, including Rate of Penetration (ROP), Rotations Per Minute (RPM), Weight on Bit (WOB), Standpipe Pressure (SPP), Gallons Per Minute (GPM), and Torque (T). These parameters, while directly related to rock properties and ultimately DTC and DTS, are significantly influenced by human operational choices. This variability means that data collected in the field for these parameters may not consistently reflect the intrinsic properties of the rock. Additionally, these parameters are sensitive to a wide range of drilling parameters and environmental factors that are not always controllable or consistently reported, potentially introducing significant noise and biases into your model. This could reduce the accuracy and generalizability of predictions. It is essential to critically assess and preprocess data for these parameters, possibly integrating additional features that capture operational and environmental conditions, to mitigate these issues and enhance the model's predictive performance.

Objective

The primary goal of this project is to construct a synthetic well-log generator using machine/deep learning techniques to predict compressional (DTC) and shear (DTS) transit times, known as slowness’s, from various drilling parameters. These parameters include torque (T), rate of penetration (ROP), rotations per minute (RPM), standpipe pressure (SPP), weight on bit (WOB), and gallons per minute (GPM), along with the target sonic data (DTC & DTS) obtained during logging while drilling (LWD). This project ultimately aims to accurately determine the geomechanical characteristics of subsurface formations—specifically Poisson’s ratio and Young’s modulus—derived from the DTC and DTS data. By developing geomechanical models, this project seeks to capture the in-situ stress states and the elastic behavior of the formations, which are essential for optimizing drilling efficiency and reducing associated risks. Additionally, this synthetic machine/deep learning based log tool aims to improve generalization across various drilling environments and compensate for human influences in drilling, enhancing the reliability and applicability of geomechanical assessments in real-world drilling operations. Moreover, in contrast to the wireline technology, the LWD sonic tool cannot measure the shear (DTS) in slow formation. So, we also aim to predict the shear in slow formation using the drilling parameters and the compressional DTC slowness.

Future plan

Data Collection and Preprocessing

- **Expand Dataset:** To enhance model generalization, collect well log data from multiple oil fields with diverse geological features to include a broader range of lithologies and depths.
- **Data Cleaning:** Assess and preprocess the drilling parameters to minimize noise and biases introduced by human-operated drilling features. This includes handling missing values and outlier detection.
- **Feature Engineering:** Develop new features that could improve model performance, such as interactions between parameters and derived statistical features.

2. Model Development

- **Model Selection:** Experiment with various machine learning models such as ensemble methods (Random Forests, Gradient Boosting Machines) and deep learning neural architectures (Convolutional Neural Networks, LSTM).
- **Cross-validation:** Implement k-fold cross-validation to assess model robustness and avoid overfitting.
- **Hyperparameter Tuning:** Use grid search or random search strategies to find the optimal model settings.

3. Model Evaluation

- **Performance Metrics:** Evaluate model performance using metrics like RMSE, MAPE, and correlation coefficients. Consider using a confusion matrix and ROC curve for classification tasks, if applicable.
- **Validation:** Validate the models on independent test sets from wells outside the training dataset to test for generalization across different geological settings.

References

Chakraborty, Shantanu, et al. “Using Rock Physics Analysis Driven Feature Engineering in ML-Based Shear Slowness Prediction Using Logs of Wells from Different Geological Setup.” Acta Geophysica, Springer Science and Business Media LLC, 5 Jan. 2024. Crossref, doi:10.1007/s11600-023-01266-3.

Rohit, et al. “A Machine Learning Approach to Predict Geomechanical Properties of Rocks from Well Logs.” International Journal of Data Science and Analytics, Springer Science and Business Media LLC, 21 Sept. 2023. Crossref, doi:10.1007/s41060-023-00451-3.

Sanei, Manouchehr, et al. “Applied Machine Learning-Based Models for Predicting the Geomechanical Parameters Using Logging Data.” Journal of Petroleum Exploration and Production Technology, vol. 13, no. 12, Springer Science and Business Media LLC, 23 Aug. 2023, pp. 2363–2385. Crossref, doi:10.1007/s13202-023-01687-2.

Alfaraj, Rima T., et al. “Application of Machine Learning for Real-Time Prediction of Sonic Well Logs Using Surface Drilling Parameters and Gamma Ray.” All Days, ARMA, 7 Nov. 2022. Crossref, doi:10.56952/igs-2022-101.

Zhong, Ruizhi, et al. “Machine Learning for Drilling Applications: A Review.” Journal of Natural Gas Science and Engineering, vol. 108, Elsevier BV, Dec. 2022, p. 104807. Crossref, doi:10.1016/j.jngse.2022.104807.

Gamal, Hany, et al. “Predicting the Rock Sonic Logs While Drilling by Random Forest and Decision Tree-Based Algorithms.” Journal of Energy Resources Technology, vol. 144, no. 4, ASME International, 16 July 2021. Crossref, doi:10.1115/1.4051670.

Onalo, David, et al. “Data Driven Model for Sonic Well Log Prediction.” Journal of Petroleum Science and Engineering, vol. 170, Elsevier BV, Nov. 2018, pp. 1022–1037. Crossref, doi:10.1016/j.petrol.2018.06.072.

Gowida, Ahmad, et al. “Prediction of Sonic Wave Transit Times From Drilling Parameters While Horizontal Drilling in Carbonate Rocks Using Neural Networks.” Petrophysics – The SPWLA Journal of Formation Evaluation and Reservoir Description, vol. 61, no. 5, Society of Petrophysicists and Well Log Analysts (SPWLA), 1 Oct. 2020, pp. 482–494. Crossref, doi:10.30632/pjv61n5-2020a6.

Glubokovskikh, Stanislav, et al. “Machine Learning Algorithms for Real-Time Prediction of the Sonic Logs Based on Drilling Parameters and Downhole Accelerometers.” SEG Technical Program Expanded Abstracts 2020, Society of Exploration Geophysicists, 25 Sept. 2020. Crossref, doi:10.1190/segam2020-3427085.1.