Demand Forecasting for Logging and Drilling Tools in the Oilfield Service Industry

by

Mingchao Wang
Master of Science in Oil & Gas Storage and Transportation Engineering, China University of Petroleum, 2013

and

Yiming Tang
Bachelor of Science in Mechanical Engineering, Shanghai Jiao Tong University, 2011

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2025

Signature of Author: ______________________________
Department of Supply Chain Management
May 9, 2025

Signature of Author: ______________________________
Department of Supply Chain Management
May 9, 2025

Certified by: ______________________________
Dr. Ilya Jackson
Research Scientist
MIT Center for Transportation & Logistics
Capstone Advisor

Accepted by: ______________________________
Prof. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Demand Forecasting for Logging and Drilling Tools in the Oilfield Service Industry

by
Mingchao Wang
and
Yiming Tang
Submitted to the Program in Supply Chain Management
on May 9, 2025 in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

Forecasting demand for logging and drilling tools is crucial in the oilfield service industry due to their highly volatile demand patterns and complex applications. These tools necessitate substantial capital investment and careful inventory management of materials and supplies (M&S). This capstone project develops a data-driven demand forecasting model to enhance the accuracy of capital expenditure (CAPEX) and M&S planning for the oilfield service industry. By leveraging historical demand data, the methodology combines advanced time series forecasting models, including temporal convolutional network (TCN) and exponential smoothing, to produce forecast results with an accuracy error of 7% over a 6-month horizon, compared to 25% of the current process used by the sponsor company. Adopting the forecast models will lead to more accurate CAPEX budgets and reduced M&S costs, enhancing operational efficiency and resource utilization for the sponsoring company.

Capstone Advisor: Dr. Ilya Jackson
Title: Research Scientist

TABLE OF CONTENTS

Contents

1. Introduction

1.1 Background

Oil and gas are extracted from reservoirs deep underneath the surface of the earth. Drilling of the wells to reach such reservoirs is known as well construction in the industry. To ensure the well trajectories hit the target reservoir and gather data for further evaluation, oil and gas companies (the clients of service companies) often request that directional drilling and formation evaluation services be provided while drilling. Rotary steerable systems (RSS) and measurements and logging while drilling (M/LWD) tools are used by service companies to provide such services to the client during the well construction phase. These RSS and M/LWD tools are owned and maintained by service companies. Compared to other services provided by service companies during well construction, RSS and M/LWD services are less common in general, making their demand highly volatile in terms of both time and geography. Since clients tend to request different logging and drilling services under different circumstances, service companies must maintain various tools for their RSS and M/LWD fleets. There typically are dozens of distinct types of RSS and M/LWD tools, which together require hundreds to thousands of stock-keeping units (SKUs) in terms of materials and supplies (M&S) for a service company. In addition, a substantial portion of M/LWD tools consist of technology or materials that are classified as dangerous or dual-use, making the movement of SKUs between countries more difficult and time-consuming.

All these factors make it very challenging for service companies to manage their supply chain for RSS and M/LWD tools. Therefore, it is essential to have a clear and accurate forecast of future M/LWD tool demand. Otherwise, the capital expenditure (CAPEX) plan and the M&S inventory management plan will be significantly impacted. This can lead to various problems, such as unpredictable tool availability and discrepancies between CAPEX and M&S inventory management strategies. Situations often arise where tools are available but lack the necessary M&S for maintenance~~,~~ or where there are no usable tools because of the under-forecasting of tool demand. This can result in either urgent SKUs movement (sometimes even by charter flight), or loss of revenue due to insufficient availability of resources.

The sponsor company, one of the largest energy technology companies globally, is motivated to enhance the accuracy of tool demand forecasting in its operations. By understanding future tool demand more clearly, the CAPEX planning for tools and the corresponding M&S inventory management plan will be more precise and resilient to operational fluctuations. This will enable the company to avoid potential resource waste and reduce operational costs, including inventory holding costs, CAPEX investments, and opportunity costs. Such improvement will ultimately lead to increased profitability.

1.2 Problem Statement

Currently, tool demand forecasting inputs are derived from quarterly sales and operations planning (S&OP) meetings and frequently adjusted based on daily operations. This process results in a gap between forecasted data and actual occurrences. Also, such a process leads to a relatively high forecast error. Current data indicates an average of 25% forecast error globally. To minimize the gap, the company needs to implement a more accurate and dynamic demand forecasting system. This system should leverage historical data analysis and advanced analytics methodologies to enhance the precision of tool demand forecasts, thereby reducing the need for frequent adjustments in daily operations.

To accomplish the objective, the following questions need to be answered:

1. What methodology should the sponsor company follow to achieve better tool demand forecasting?
2. Due to the unstable nature of the oil and gas industry, there are many exogenous factors that influence tool demand. What are the key exogenous factors to focus on?
3. How can these key exogenous factors be incorporated into the demand forecasting model?
4. How can the sponsor company translate these forecasts into actionable decisions for M&S inventory management and CAPEX investment planning?

1.3 Scope: Project Goals and Expected Outcomes

This project develops a demand forecasting model for a selected operational geographical unit (geo-unit) within the sponsor company. The chosen geo-unit is ideal for this pilot due to its traditional operational style and the presence of both offshore and land operations within one location. Additionally, the substantial operation volume in this geo-unit presents a significant opportunity for improvement. Enhanced demand forecasting in this geo-unit is expected to lead to considerable profit increases.

The project's overall goal is to deliver a demand forecasting model that enables the sponsor company to predict tool demand with 15% or lower planning error over a 6-month horizon. Following validation, the model will be extended to other geo-units as well as to the regional Operations Control Center and global headquarters.

We anticipate that a data-driven approach to determining future tool demand will significantly enhance the company's planning accuracy. By analyzing historical demand data, we aim to identify demand patterns and apply appropriate forecasting techniques for future predictions. Additionally, by examining discrepancies between historical demand forecasts and actual occurrence data, we will generate correction factors to improve the accuracy of future forecasts.

The deliverables to the company include:

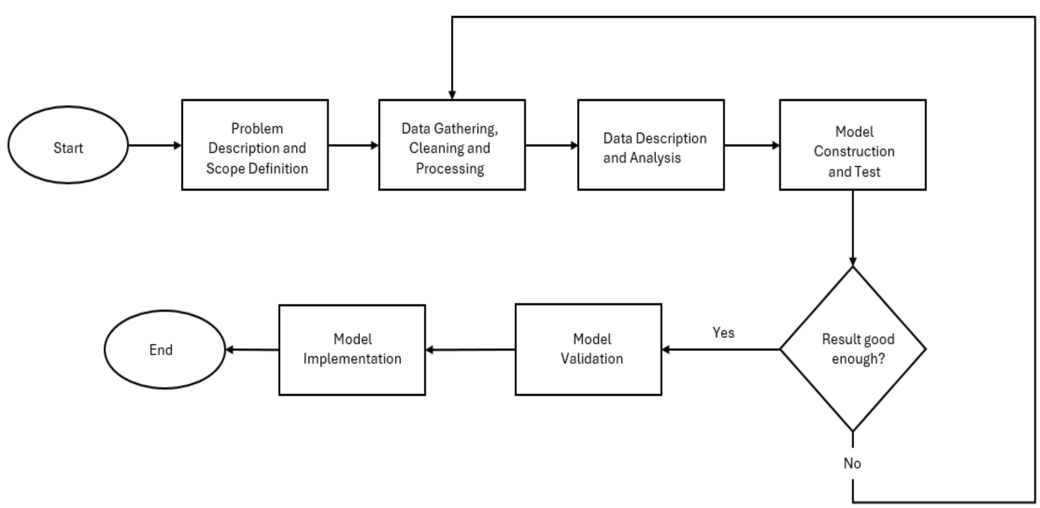
1. A model for determining future tool demand over a 6-month horizon.
2. A dashboard for communicating tool demand forecasts to the resources planning team.

By implementing the proposed model, the sponsor company is expected to enhance overall planning accuracy and develop a more resilient M&S inventory management plan and CAPEX investment strategy.

1.4 Plan of Work

Figure 1 illustrates the primary steps involved in this capstone. They include describing the problem and defining the project's scope; data gathering, cleaning, and processing; harmonizing and analyzing the data; utilizing that data to develop and test the demand forecasting model; and applying the model after validation.

Figure 1. Project Workflow



2. State of the Practice

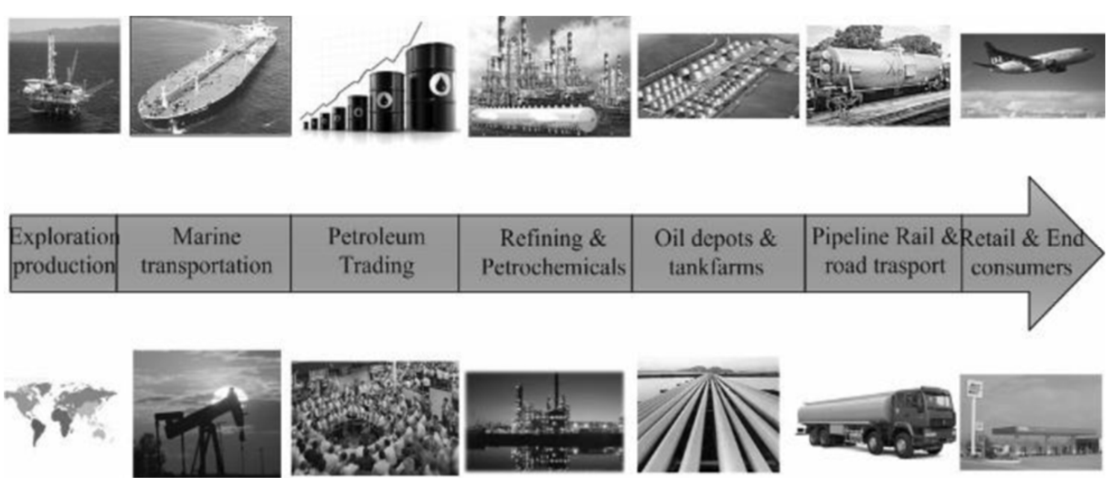
The purpose of this project is to develop a demand forecasting model that enables the sponsor company to better plan its capital expenditure (CAPEX) and M&S inventory management. Beyond the necessary dataset, a valid forecasting model must consider exogenous factors that critically influence the results. To identify these factors, we first need a comprehensive understanding of the oil and gas industry's supply chain, so we started by reviewing the literature on its characteristics.

Once we have an overall understanding and have identified the major influencing factors, the next step is to select appropriate forecasting models. This requires a deeper understanding of current forecasting techniques and theories. By combining these techniques with our specific case, we can select suitable models and implement them. Simultaneously, we must account for the impact of major factors when building the forecasting models to ensure their accuracy and reliability.

2.1 The Oil and Gas Supply Chain Overview

The oil and gas industry is distinguished from other industries in many ways. Projects in this sector tend to be bigger in scale and more technically complex than those in most other industries, requiring the cooperation of different entities with diverse expertise. Typically, multiple companies are involved in completing a single project in the oil and gas industry, either through joint ventures or subcontracting (Menhat, 2017). Consequently, the corresponding supply chain for the oil and gas industry becomes quite long and complicated. As shown in Figure 2, the oil and gas supply chain encompasses all stages from exploration to end users.

Figure 2. Oil and Gas Supply Chain (Fernandes et al. 2010)



Due to its complexity and scale, the oil and gas supply chain is normally divided into three sectors: Upstream, Midstream, and Downstream. Upstream, also known as exploration and production (E&P), focuses on allocating and extracting oil and gas from the reservoir underground. Midstream refers to the process of temporary storage and transportation from the production facility to the refinery. Oilfields across the world are located in some of the most remote landscapes on the planet, from the heart of the deserts to deep in the oceans, and from frozen soil in the Arctic to rain-forests on the Equator. Midstream activities involve gathering oil and gas from these fields and utilizing various means of transport—such as pipelines, oil tankers, liquefied natural gas (LNG) vessels, trucks, and trains—to deliver them to refineries. Downstream activities handle the refining and marketing processes, converting the oil and gas into various final products and delivering them to end users worldwide (Lisitsa et al., 2019).

"Exploration and production (E&P) is engineering-intensive and requires specialized technical expertise at every step along the way" (Jacoby & Gupta, 2021, p. 162). Therefore, oil and gas operators in the upstream segment typically rely on service companies and contractors to provide both the specialized equipment and technical personnel required. Drilling, one of the key steps in upstream operations, requires the oil and gas operators to organize a sophisticated supply chain network involving multiple service providers for such equipment and services as drilling rigs, drilling fluids, tubulars, directional drilling, formation evaluation, well completion, and transportation (Jacoby & Gupta, 2021). Additionally, to complete a well, the equipment and crew of each service provider must arrive at different stages of the operation and cooperate with each other at the drilling site, which is usually in a remote location. Menhat (2017) has summarized some of the distinctive characteristics of upstream operations as follows:

- Geographical location. The majority of exploration and production (E&P) sites today are in remote locations with limited or no infrastructure and are far from major settlements.
- High risks. High capital investment is required for equipment and facilities, compounded by the uncertainty inherent in the oil and gas reservoirs themselves.
- Complex engineering. Upstream activities involve multiple contractors and service providers.

The "bullwhip effect" describes the tendency of demand variability to be amplified through each level of the supply chain while moving further upstream. Ever since Jay Forrester defined the bullwhip effect in 1958, it has been identified in almost every industry. Given the fact of a multi-level supply chain network and the capital-intensive nature of the oil and gas industry, the bullwhip effect has not gone unnoticed within the industry, especially in the upstream sector. Azhar (2013) pointed out that smaller companies in the oil and gas industry tend to suffer more from the bullwhip effect compared to larger ones. This difference is because larger companies in the oil and gas industry are usually vertically integrated, providing them with easier or direct access to end users' demand data. Dujak et al. (2019) also urged the organizations at different levels of the oil and gas supply chain to collaborate on demand forecasts, using data on actual demand and variables collected from all members participating in the supply chain network.

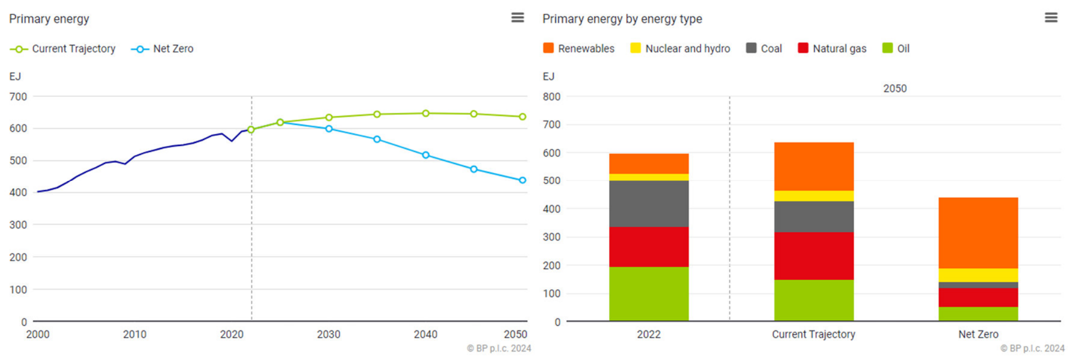
The oil and gas industry allocates substantial budgets for CAPEX and M&S annually, therefore, the cost of the bullwhip effect can be quite significant. In a system dynamics simulation model constructed by Jacoby and Gupta (2021), the cost of the bullwhip effect across the oil and gas supply chain accounts for almost 10% of the weighted average cost to produce a barrel of oil. This underscores the importance of demand forecasting in the oil and gas industry due to its substantial financial impacts across the entire supply chain.

2.2 Factors Influencing Oil and Gas Industry Forecasting

Oil price is perhaps the single most important factor for planning and forecasting in the oil and gas industry. Although different researchers utilize different models to describe the volatility of oil prices (Elder & Serletis, 2010; Rahman & Serletis, 2012; Serletis & Xu, 2016), they all agree that oil prices are highly volatile and extremely difficult to forecast, especially for a long-term horizon. Unlike an ideal free market with desired supply and demand equilibrium, oil and gas prices are generally influenced by factors beyond just supply and demand, causing them to fluctuate much more than expected (Cansado-Bravo & Rodríguez-Monroy 2018). As commodities, oil prices are influenced not only by supply and demand but also by factors such as geopolitical events and global economic performance. Therefore, all participants in the oil and gas industry must put in more effort in demand forecasting to counteract oil price volatility.

Another factor that concerns the oil and gas industry is the transition towards new energy sources. As of 2023, oil and gas accounted for 55% of the world's energy supply, while fossil fuels in total were responsible for almost 80% of the global energy supply (DNV, 2024). As the world continues to reduce carbon emissions, a considerable share of fossil fuels is expected to be substituted by renewable energy. However, the recently observed retreat confirms that this transition process is affected by many factors that are difficult to control and predict, including global economies, social events, and politics (Balcerzak et al. 2023). Almost everyone agrees that the share of coal in the global energy supply will decline in the coming decades, but when it comes to oil and gas, the outlook is less certain.

Figure 3. Global Primary Energy Demand and Share by Energy Type (BP, 2024)



As shown in Figure 3, BP (2024) provided two forecast results for global primary energy demand in the upcoming years. The forecast based on the current trajectory indicated a drop in oil's share, from 32% in 2022 to 23% by 2050, with a slight increase in gas's share, from 24% to 27%. Conversely, the forecast based on achieving net zero by 2050 shows a significant decrease in both oil's share, to 12%, and gas's share, to 15%, by 2050. These uncertainties complicate strategic demand forecasting for the oil and gas industry and suggest that companies should urgently consider the pace of global decarbonization in their planning.

2.3 Time Series Forecasting

The ultimate goal of forecasting is to predict the future based on the analysis of information about the past and factors that might influence future outcomes. Modern organizations require the development of a forecasting system that can generate outputs on different time scales by selecting and applying appropriate methods and evaluating the results (Hyndman & Athanasopoulos, 2021).

Forecasting methods can be divided into two categories: qualitative and quantitative. Qualitative forecasting, sometimes known as judgmental forecasting, is commonly used when historical data is unavailable (Hyndman & Athanasopoulos, 2021). In contrast, quantitative forecasting methods can be applied to deeply analyze past information and generate future forecasts.

Time series analysis is an effective method for uncovering patterns in historical data, including periodicity or seasonality, and using them to generate forecasts

(Sheremetov et al., 2013). In our case, the oil and gas industry has gathered and stored enormous amounts of data through years of operation. Most of the demand data is time-associated, which can be utilized to create a time series forecasting model.

Although multiple time series forecasting models are available, some stand out due to their flexibility and in-depth learning capabilities when dealing with time series data. Those models include the seasonal autoregressive integrated moving average (SARIMA) model, temporal convolutional network (TCN) model, neural basis expansion analysis (N-BEATS) model, and time-series dense encoder (TiDE) model.

The SARIMA model includes trend and seasonal patterns as covariates, allowing it to capture both the long-term progression and repeating cycles within the data. Additionally, it can handle the stochastic dependence of consecutive data points, meaning it accounts for the correlation between successive observations, which is common in time series data (Perez-Guerra et al., 2023). The TCN model is capable of detecting anomalies and performing deep learning on time series data (He & Zhao, 2019). N-BEATS, one of the latest recurrent neural network models, is flexible and effective in handling time series forecasting (Oreshkin et al., 2019). The TiDE model leverages a multi-layer perceptron (MLP)-based encoder-decoder architecture, effectively capturing the nuances of historical data (Das et al., 2023).

3. Methodology

Based on the literature review, we are convinced that a time series forecasting approach is the best method to solve the demand forecast problems our sponsor company has addressed. Monthly historical demand for the past couple of years is available from the sponsor company, and the historical crude oil price, which is the most important exogenous factor, can be obtained from the U.S. Energy Information Administration website. This makes it reasonable to apply some of the time series models.

Considering the numerous types of tools involved in this project, we need to forecast both the overall demand for all tools (to allocate CAPEX budget) and the specific demand for each tool type (to develop an effective inventory management plan for M&S). This means that we may need to identify multiple models to fulfill the forecasting requirements accurately.

The roadmap shown in the plan of work (Figure1) will be followed, and the key steps are as follows:

1. Scope Definition. During this step, the goal of the project and the specific forecasting requirements to achieve this goal are defined. We develop

multiple fit-for-purpose models to forecast the overall demand for all tools and the demand for each individual type of tool separately. Additionally, we need to define the forecast horizon in terms of both time and geography. As previously mentioned, this project will focus on a 6-month horizon and one selected operational geo-unit.

2. Data Gathering and Cleaning. In this step, the required data is collected and checked. Incomplete records and outliers will be removed, and as an output of this step, a data set is ready for processing.
3. Data Description and Analysis. This step is also referred to as "Preliminary (exploratory) analysis" (Hyndman & Athanasopoulos, 2021). The best way to understand a time series is by graphing the data. By doing so, we can uncover underlying trends, seasonality, or business cycles. Starting with a time plot is a good choice, as most trends and seasonality will be revealed. Following this, an autocorrelation function plot can be used to measure the linear relationship between lagged values in the time series. This helps determine the appropriate differencing level, which is an important parameter in most time series forecasting models.
4. Model Construction and Validation. We begin by building the baseline models, using "Naïve," "Cumulative," and "Naïve Seasonal" as our initial baselines. The complexity then increases as we apply other models, such as SARIMA, Exponential Smoothing, N-BEATS, TCN, and TiDE. By classifying historical data into training and validation datasets, we can perform a detailed comparison of each model's performance using metrics like mean square error (MSE) and mean absolute percentage error (MAPE). After this step, we will identify a few suitable models.
5. Model Implementation. Once the performance meets the criteria defined in the scope of work, the forecasting model will be implemented. Newly obtained actual demand data could be compared with the forecast generated to fine-tune the models for future use.

During each step, an overview of the project level is conducted to ensure the project proceeds per its original scope. This could also align all parties involved with the ultimate goals defined and ensure a smooth execution.

3.1 Scope Definition

We started by defining the scope of the project. The sponsor company operates in more than 120 countries globally. Its organization splits into a handful of GeoUnits, each covering operations in a few countries. Each GeoUnit has its own unique operational requirements. Therefore, the demand pattern also varies. Our project focuses on one of the largest and most important aspects for the sponsoring company, which has a full portfolio of products and services.

Strategic demand (1 to 3 year time horizon) for oilfield services is affected by many factors, such as geopolitical events and global economics. In the sponsor company, strategic demand planning is taken at a global level to better leverage the volatility in different countries. Since our project focuses only on one specific GeoUnit, we agree on a tactical forecast that suits the situations at the GeoUnit level. A 6-month forecast time horizon is then established after discussion with the sponsor.

The logging and drilling tools used by the sponsor company belong to a handful of different product families. It is quite important for the sponsor company to forecast the demand for each product family on top of the overall demand forecast. Therefore, the forecast model we develop, with some changes of parameters, provides a set of forecasts for both combined demand and each individual product family within the same time horizon. The overall forecast provides information for the sponsor company with regarding how they could plan their shared resources and overhead cost such as workforce and facilities. The segmented forecast for each tool family gives more insights on how to plan the CAPEX and M&S for the fleet.

3.2 Data Gathering and Cleaning

The sponsor provided PowerBI data starting from 2011, containing detailed information such as tools used, rig count, financial period, GeoUnit labels, and job count (the target variable). However, upon inspection, we observed that the data from 2011 to 2013 lacked some inputs and appeared incomplete.

Through interviews with experts in the oil and gas industry, we learned that 2015 marked a significant turning point due to major changes in oil prices. Discussions with the sponsor company further revealed that a strategic organizational restructuring took place in 2015 during the industry's downturn. Based on these insights, we decided to discard any data prior to 2016 to ensure consistency and reliability in our analysis.

The sponsor company has made substantial efforts to maintain data integrity, and we did not identify significant errors or missing data from 2016 onward. This suggests that a pre-cleaned dataset was supplied to us. The provided data is recorded at the job level across various locations within the GeoUnit. Our primary task involved preparing all data from 2016 onward and aggregating it to the GeoUnit level in batches. This aggregated dataset serves as the foundation for our demand forecasting models.

Moving forward, we classified the data based on the validation of model performance. One classification dimension involves grouping by tool type, while another considers the unique characteristics of different operational regions. We iteratively validated the models across these classifications to identify the most suitable approach for accurate demand forecasting.

3.3 Data Description and Analysis

The final dataset comprises historical demand data spanning from 2016 to June 2024. Each year's data is stored in a separate Excel spreadsheet, all of which follow the same structured format as detailed in the table below. Currently, the data has been aggregated at the geo-unit level, meaning that location and customer-specific details have been excluded. This aggregation aligns with the capital expenditure (CAPEX) planning framework, which is conducted at the geo-unit level. Additionally, forecasting demand at this level is crucial for informed decision-making and enhances accuracy by reducing data variability. Table 1 describes the detailed variables used in the dataset and their corresponding descriptions.

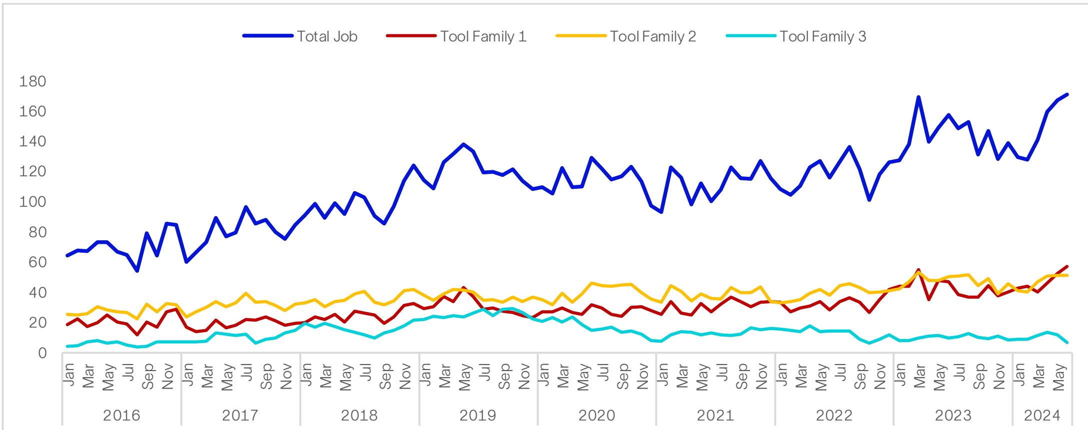
Table 1. Variables Descriptions

Variable	Description
Calender - Period	The time period in which the data was recorded, in months.
Basin	The higher level of geographical information about where the operations are taking place helps categorize different regions of drilling activity.
Geo Unit	A geographical unit representing a subdivision of a basin, used for tracking and reporting.
Mgt Country	The country responsible for managing the operations.
Location	The specific operation location that supports the job.
Rig Name	The name or identifier of the drilling rig used for operations.
Customer Name	The name of the customer.
# Job Number	A unique identifier assigned to each drilling or logging job for tracking purposes.
Tool Family Size	A classification of tools used in operations, indicating the type and size of tools deployed.
Job Volume	The job volume performed in the given month, representing operational activity levels.

Data visualization helps reveal more insights into the data, enabling better understanding and communication with stakeholders. It also indicates possible ways forward. We also started by plotting the dataset that was prepared in the previous step.

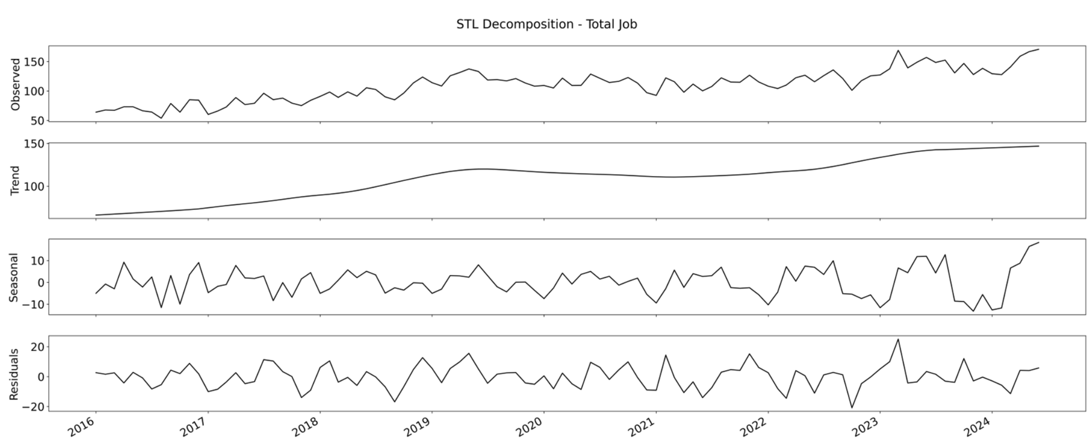
We first noticed that some of the tool families followed a similar trend as the overall job volume (see Figure 4). Tool family 1 significantly correlated with the overall job volume compared to the others.

Figure 4. Total Job Volume with Family-Level Segmentation



Next, we turned to the overall job volume. Although the data is volatile, there is an overall upward trend. We then applied decomposition of the total job volume, which gave more insights into the dataset. The decomposed data in Figure 5 confirmed the general upward trend, with the exception of the year 2020, which was likely due to the impact of COVID-19. It further suggested the existence of seasonality to a certain degree. The residuals were oscillating around 0, which was the ideal situation. However, the high absolute value of some data points indicated that there was more uncovered information about the total job volume.

Figure 5. Decomposition of Demand Data



3.4 Model Construction and Validation

Given the inherent volatility of the oil and gas industry, we set the forecasting horizon to six months. Initially, forecasting is applied to the total job volume and three tool families, representing the highest tool classification level in logging and drilling operations. Model performance is evaluated using key statistical metrics, as outlined in the table below. These metrics serve as a quantitative benchmark for comparing different forecasting approaches and ensuring accuracy in demand predictions.

Table 2. Forecasting Performance Metrics

Metric	Description	Interpretation
Mean Absolute Percentage Error (MAPE)	Measures the average percentage error between predicted and actual values.	Lower values indicate better accuracy, with 0% being a perfect forecast.
Mean Absolute Error (MAE)	Calculates the average absolute difference between predicted and actual values.	A lower MAE signifies more accurate predictions.
Mean Squared Error (MSE)	Computes the average squared difference between predicted and actual values, penalizing larger errors more heavily.	A lower MSE suggests better model performance, but it can be sensitive to large errors.
Root Mean Squared Error (RMSE)	The square root of MSE, providing an error metric in the same unit as the original data.	A lower RMSE indicates higher accuracy while reducing sensitivity to extreme errors.

3.4.1 Baseline Model Performance Analysis

Before implementing advanced time-series forecasting models, we developed a baseline model as a simple yet effective starting point. Baseline models are easy to apply, require no model training, and serve as a valuable reference for evaluating more sophisticated approaches. Additionally, they provide insights that can guide the development of advanced models.

Our baseline model includes the following approaches:

1. Last Value: Uses the most recent observed value as the forecast for all future periods.
2. Mean: Computes the historical mean and applies it as the forecast for future periods.
3. 6-Month Moving Average: Calculates the rolling average of the past six months and uses it as the forecast for the next period.

4. Naïve Seasonal: Since the dataset exhibits seasonality, we use the corresponding value from the previous seasonal cycle as the forecast. For example, next March's demand is equal to the demand from March of the prior year.

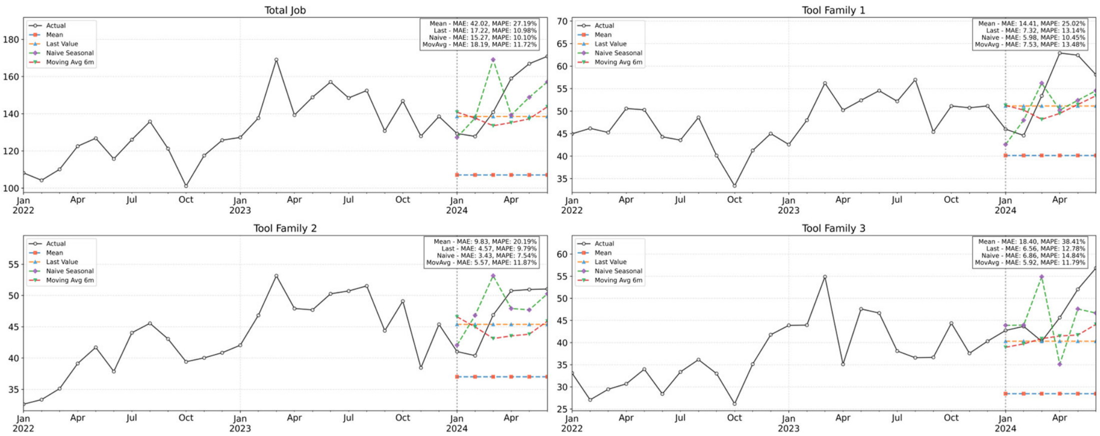
Table 3 shows the performance metrics for each baseline model we have constructed. The Naïve Seasonal model consistently outperforms the other baseline models across all categories, showing the lowest MAE and MAPE values. This confirms that seasonality is a significant factor in forecasting demand. The Last Value model performs reasonably well, but its accuracy varies depending on the fluctuations in the dataset. The Moving Average (6-month) model provides a smoothed forecast but struggles when demand exhibits sharp changes. The Mean model is the weakest, with the highest errors, as it does not capture trends or seasonality.

Table 3. Baseline Models Performance Comparison

Model	***Average_MAPE%***	***Average_MAE***	***Average_MSE***	***Average_RMSE***
Last Value	11.7	8.9	143.9	10.5
Mean	27.7	21.2	703.8	22.9
Moving Avg 6m	12.2	9.3	140.8	10.4
Naïve Seasonal	10.7	7.9	110.5	9.3

The strong performance of the Naïve Seasonal model suggests that future forecasting models should explicitly incorporate seasonality, such as Exponential Smoothing (Holt-Winters), SARIMA, or machine learning models with seasonal features. However, as illustrated in Figure 6, for certain tool families, the poor performance of baseline models urges the application of better models to achieve a reliable forecast. The baseline models provide helpful references, and their results assist in guiding the selection and fine-tuning of more advanced forecasting techniques.

Figure 6. Baseline Model Forecasts for Total Job and Selected Tool Families



3.4.2 Model Selection and Performance Analysis

The performance of the six selected forecasting models—Exponential Smoothing, FourTheta, N-BEATS, SARIMA, TCN, and TiDE—has been evaluated using MAPE, MAE, MSE, and RMSE metrics. Following the same logic as for the baseline model, the selected models were used to forecast demand for the next six months and validated against actual demand to evaluate their performance. Additionally, backtesting was performed to confirm the performance.

Backtesting is the process of evaluating a forecasting model's performance on historical data to assess how well it would have performed in the past. Instead of training a model on the full dataset and predicting future values, backtesting involves simulating a real forecasting scenario by using only past data at each step and assessing how well the model predicts the following points. In our practice, the backtesting process begins at 70% of the data set and trains the model on all the data up to that point. It then generates a 6-month forecast, as mentioned earlier, and moves forward one time step, repeating the process until the end of the data.

4. Results and Discussion

This chapter presents the findings from both the overall model evaluations and the backtesting results analysis. The results metrics of the six selected models (Exponential Smoothing, FourTheta, N-BEATS, SARIMA, TCN and TiDE) are presented and the best model is identified.

4.1 Model Forecast Results

Tables 4 and 5 present the performance metrics for the overall demand forecast of the six selected models, along with the backtesting results for each model. All models perform better compared to the baseline models. However, the TCN model stands out with the lowest forecast error, which is also validated by backtesting. This section will review the detailed results for each model.

Table 4. Model Performance Metrics

Model	Average MAPE%	Average MAE	Average MSE	Average RMSE
Exponential Smoothing	8.6	6.3	59.0	6.9
FourTheta	8.6	6.3	61.8	7.0
N-BEATS	10.0	6.2	72.4	8.1
SARIMA	9.6	7.5	96.4	8.3
TCN	6.6	4.9	46.7	6.0
TiDE	9.5	6.2	64.5	7.8

Table 5. Backtesting Results for Models

Model	Average MAPE%	Average MAE	Average MSE	Average RMSE
Exponential Smoothing	11.1	6.8	86.8	9.3
FourTheta	11.6	7.5	100.5	10.0
N-BEATS	14.5	8.2	124.3	11.2
SARIMA	10.9	6.7	86.4	9.3
TCN	10.5	7.1	109.8	10.5
TiDE	15.8	9.6	151.9	12.3

Exponential Smoothing is a lightweight forecasting method that applies weighted averages to past observations, with newer data points given greater importance. SARIMA, on the other hand, extends autoregressive integrated moving average (ARIMA) by explicitly modeling seasonal patterns, making it well-suited for time series with recurring trends. Both models demonstrate consistent and reliable performance, achieving:

- Low error rates: MAPE between 8.6–11.1% and MAE in the range of 6.3–7.5 across evaluations.
- Moderate MSE & RMSE values, indicating robust handling of demand fluctuations without excessive sensitivity to outliers.

Figures 7 and 8 plot the six-month forecasts of Exponential Smoothing and SARIMA models. The forecasts from both models follow the actual demand quite well for the overall job volume and all tool families. The small difference between the forecast and backtesting results proves the robustness of these models. Such stability suggests that Exponential Smoothing and SARIMA effectively balance accuracy and generalization, making them strong candidates for baseline forecasting in volatile demand environments.

Figure 7. Exponential Smoothing Model Result

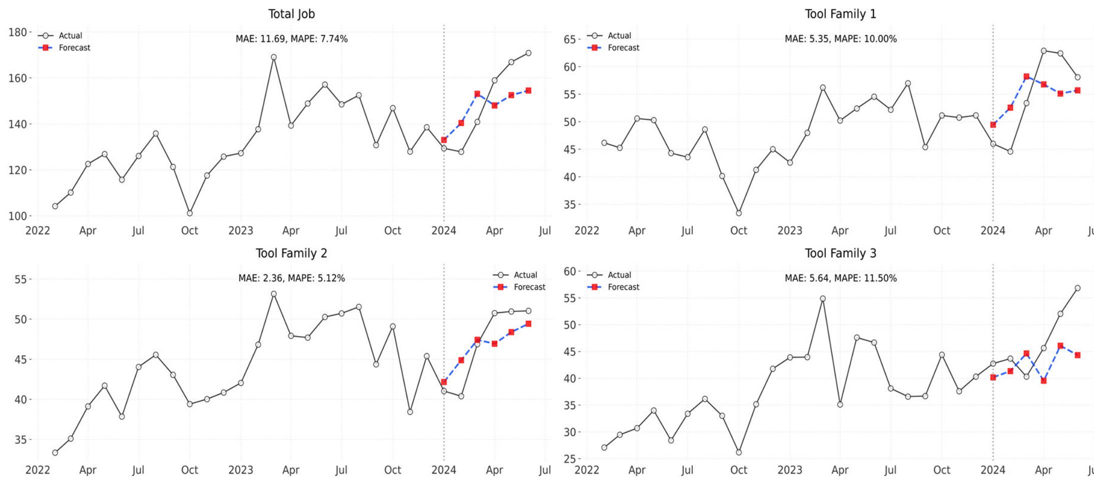
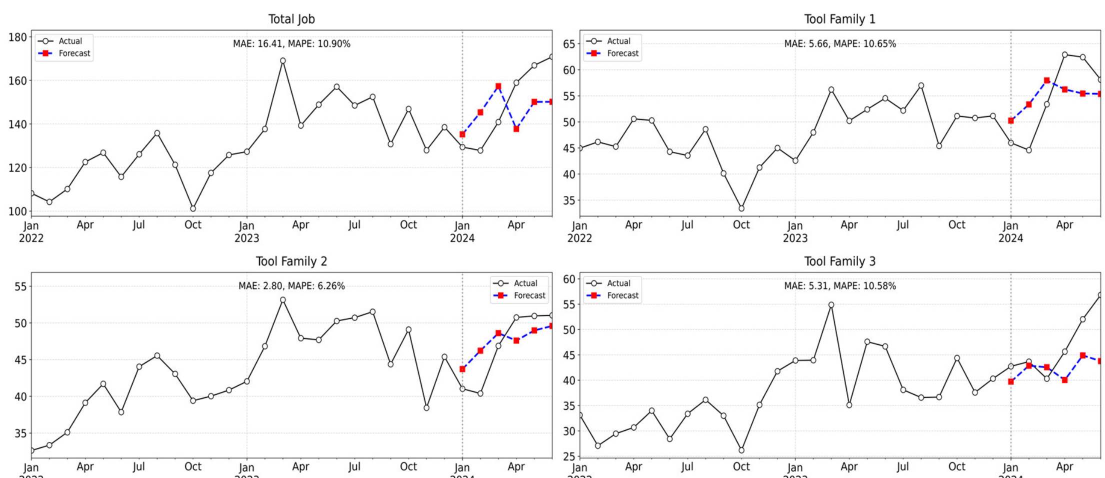


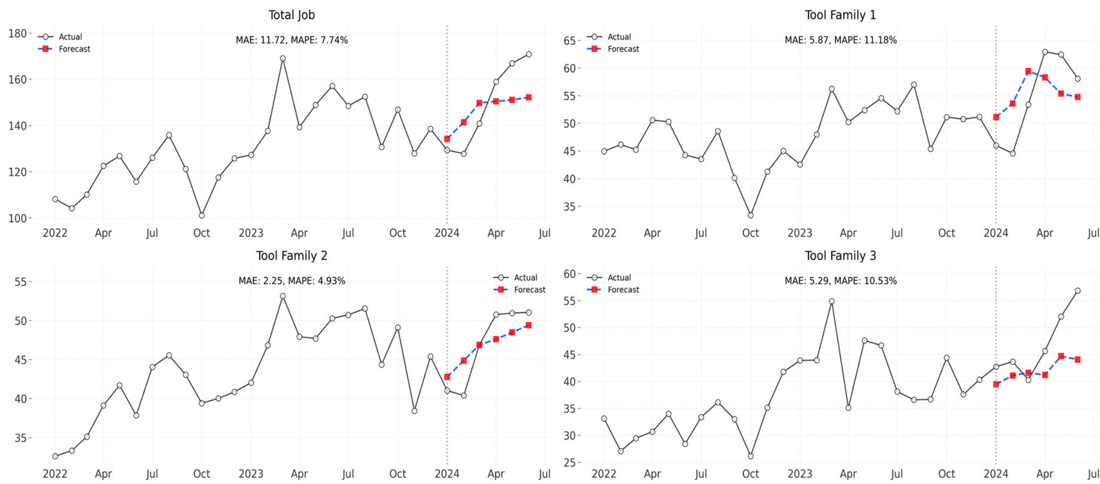
Figure 8. SARIMA Model Result



FourTheta is a univariate forecasting method that combines four theta decomposition lines to model different aspects of time series trends. In our result plots, as shown in Figure 9, it initially matched the performance of Exponential Smoothing but showed notable degradation during backtesting, with a higher MAE (7.5) and MSE

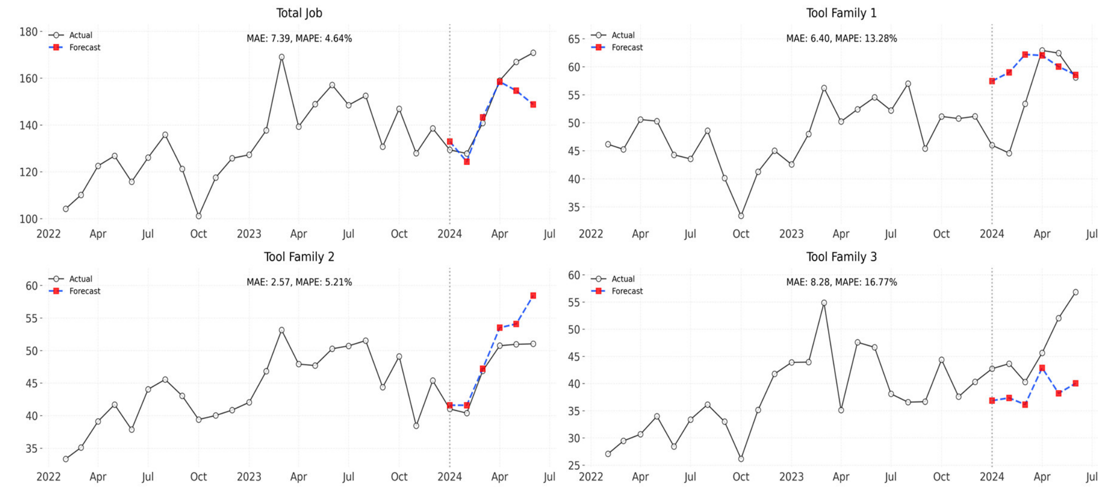
(100.5). This divergence suggests limited adaptability to unseen data patterns, potentially due to its fixed decomposition approach struggling with demand shifts.

Figure 9. FourTheta Model Result



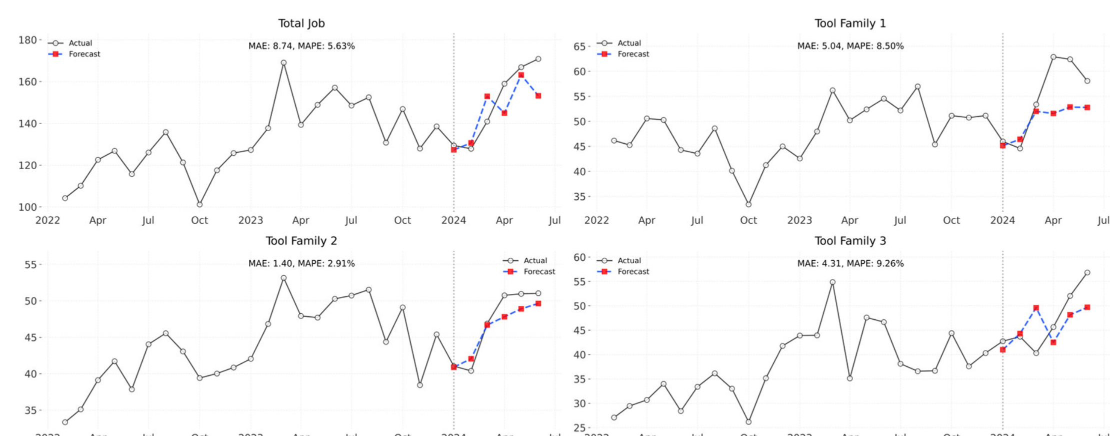
N-BEATS (Neural Basis Expansion Analysis) is a deep learning architecture designed for interpretable time series forecasting through basis function decomposition. While theoretically robust, our implementation produced higher errors (MAPE: 10.0% → 14.5%, MSE: 72.4 →124.3) in the backtesting compared to the evaluation, with an RMSE of 11.2 in the backtesting reflecting significant prediction swings. As demonstrated in Figure 10, the forecasted total job demand follows perfectly for the first four months~~,~~ but then follows the opposite trend compared to the actual demand data. Similar behavior is observed for Tool Family 2 as well. This behavior suggests over-reliance on training patterns or inadequate regularization for real-world volatility.

Figure 10. N-BEATS Model Result



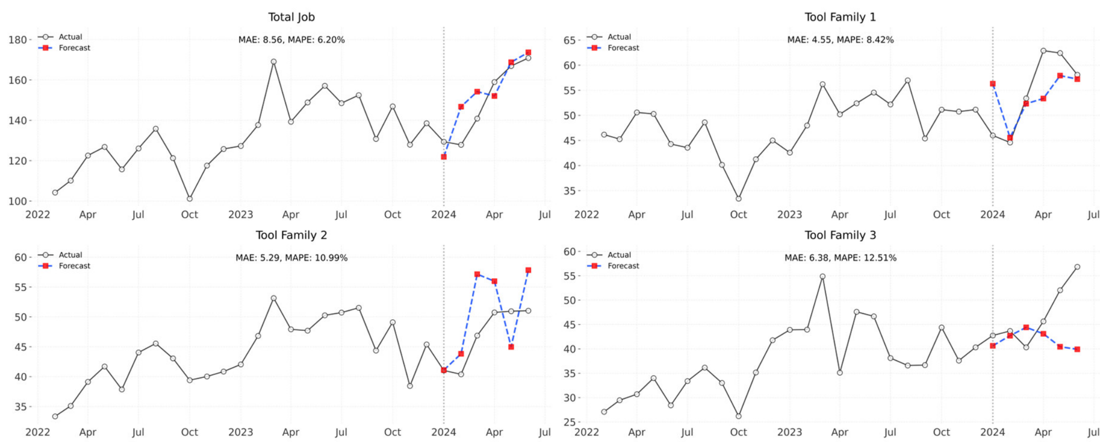
TCN leverages dilated causal convolutions to capture long-range temporal dependencies while maintaining computational efficiency. It emerged as the top performer in the initial evaluation, achieving exceptional accuracy with MAPE (6.6%) and MAE (4.9%), the lowest among all models. Figure 11 plots the six-month demand forecast against the actual value. Although the forecast does not match very well for Tool Family 1 and 3, the average results in terms of all metrics are superior among the six candidate models. However, the backtesting revealed moderate performance degradation (MAPE: 10.5%, MSE: 109.8 vs 46.7), suggesting that while its architecture is fundamentally strong, the model may benefit from additional regularization or training adjustments to improve robustness against dataset variations.

Figure 11. TCN Model Result



TiDE was developed “for long-term time-series forecasting that enjoys the simplicity and speed of linear models while also being able to handle covariates and non-linear dependencies” (Das et al., 2023). We tested this TiDE as a candidate model, and the results of the six-month forecast are shown in Figure 12. Unfortunately, it delivered the weakest performance in our tests, with MAPE worsening from 9.5% to 15.8% during backtesting and the highest MSE (151.9) among all models. The substantial RMSE (12.3) indicates larger errors and greater inconsistency in predictions. These results suggest the model may either be overly sensitive to noise or requires significantly more training data to stabilize its latent dynamics learning.

Figure 12. TiDE Model Result



4.2 Results Discussion

TCN appears to be the most promising model, based on its strong initial performance; however, its backtesting results suggest potential overfitting or instability. Further tuning may be required.

Exponential Smoothing and SARIMA provide stable performance, establishing them as robust baseline models for demand forecasting. FourTheta exhibits declining performance in backtesting, indicating it may not generalize well and requires further parameter fine-tuning to enhance its effectiveness. TiDE and N-BEATS also show diminishing performance in backtesting, rendering them the least suitable models for this task.

Based on insights gained during the model construction phase, the TCN model should undergo further fine-tuning to maintain its high accuracy while mitigating risks of overfitting. To establish robust baselines, we will also incorporate the Exponential

Smoothing and FourTheta models, which are known for their reliability in time series forecasting.

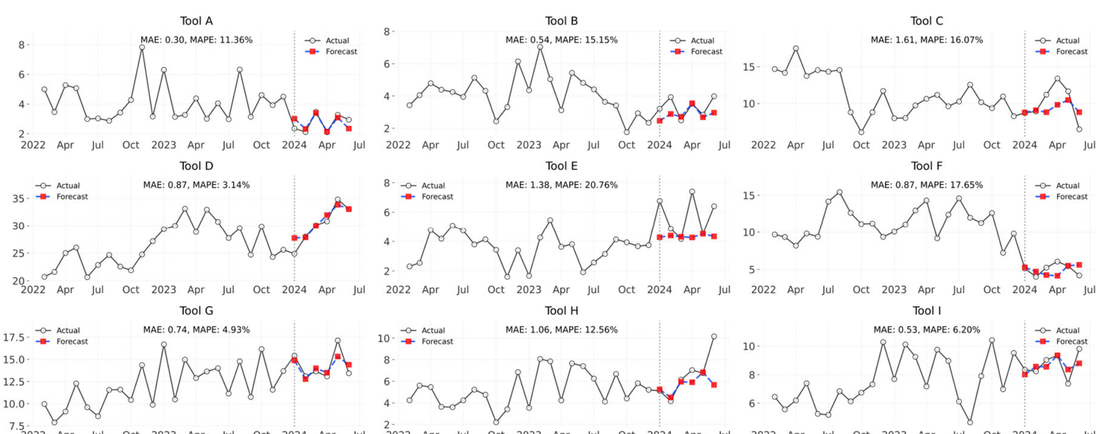
The selected three models were evaluated across nine distinct single-tool categories, strategically selected to represent the full spectrum of job volumes, from fewer than 10 to over 30. This comprehensive coverage ensures that model performance is tested across varying demand levels, providing a more reliable assessment of scalability and generalization. Table 6 presents the average performance metrics of the nine selected single tool categories.

Table 6. Selected Model Performance Comparison Summary

Model	***Average MAPE%***	***Average MAE***	***Average MSE***	***Average RMSE***
Exponential Smoothing	30.8	1.9	7.6	2.2
FourTheta	35.1	2.1	8.5	2.5
TCN	12.0	0.9	1.8	1.2

TCN emerged as the clear top performer, demonstrating superior accuracy across all metrics. With an average MAPE of 12.0% (60% lower than traditional methods), MAE of 0.9, and MSE of 1.8, it significantly outperformed both Exponential Smoothing and FourTheta. The model's consistently low error rates (RMSE = 1.2) indicate robust handling of demand variations across all job volume ranges, particularly excelling in low-volume scenarios where absolute precision (MAE = 0.9) matters most. Figure 13 plots the six-month forecast for the nine selected single tools. This strong performance suggests that TCN's temporal convolution architecture effectively captures complex patterns while maintaining stability.

Figure 13. TCN Performance for Single Tools



Traditional models showed notable limitations, with both Exponential Smoothing (MAPE = 30.8%) shown in Figure 14 and FourTheta (MAPE = 35.1%) shown in Figure 15 struggling with systematic errors. Their higher MSE/RMSE values (7.6-8.5/2.2-2.5 vs. TCN's 1.8/1.2) reveal sensitivity to outliers and volatility, especially in low-volume cases. FourTheta's consistent underperformance relative to Exponential Smoothing suggests its theta decomposition approach may be poorly suited to this dataset's characteristics.

Figure 14. Exponential Smoothing Performance for Single Tools

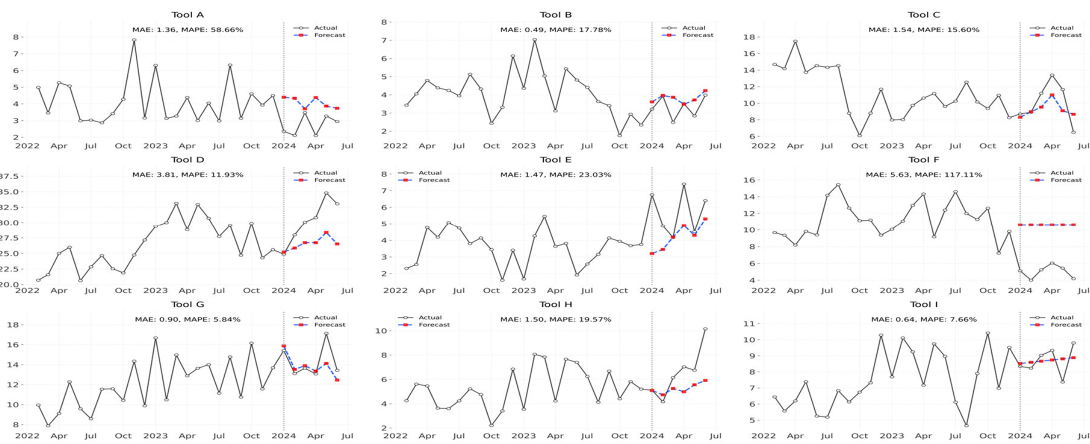
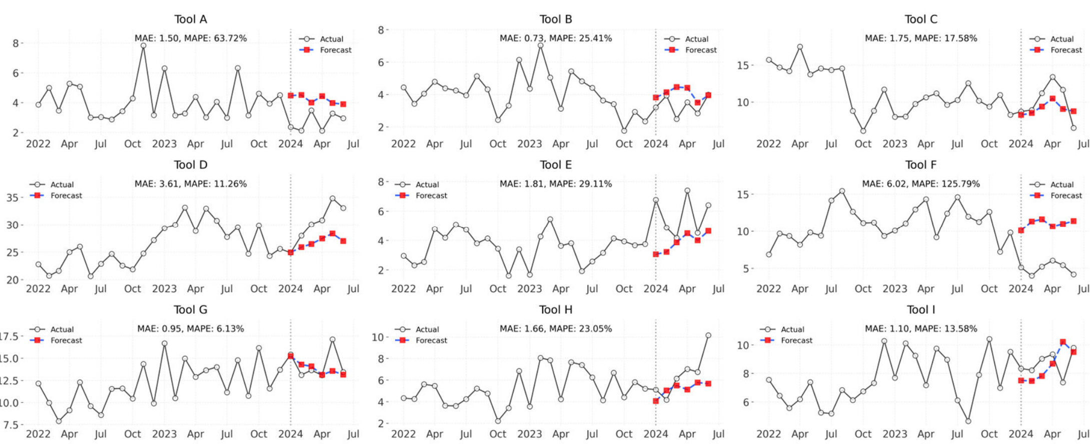


Figure 15. FourTheta Performance for Single Tools



For practical deployment, TCN is the recommended choice, although its computational complexity warrants evaluation against implementation constraints. Exponential Smoothing remains a viable baseline for simpler use cases, while FourTheta may require architectural reevaluation. Future work should segment the analysis by volume tiers to further validate TCN's robustness across demand levels.

5. Conclusion

The objective of this capstone project is to find a more accurate and dynamic demand forecasting system that enhances the current S&OP process for the sponsoring company. We identify machine learning forecast models to improve demand forecast accuracy for logging and drilling tools.

The comprehensive evaluation of forecasting models for logging and drilling tool demand yielded critical insights into model performance, robustness, and suitability for deployment. Below is a summary of key findings and recommendations:

- TCN (Temporal Convolutional Network) demonstrated the strongest overall performance, achieving the lowest errors during initial evaluation (MAPE: 6.6%, MAE: 4.9) and maintaining competitive accuracy in backtesting (MAPE: 10.5%). Its ability to capture complex temporal patterns makes it the top candidate for application. Further parameter fine-tuning could enhance its stability on unseen data.
- Exponential Smoothing and SARIMA provided reliable baseline performance with consistent metrics (MAPE: 8.6 – 11.1%) and handled volatility effectively. Their simplicity and interpretability make them ideal for scenarios where TCN's complexity is unnecessary.
- FourTheta displayed limited generalizability, with significant performance drops in backtesting (MAE: 7.5 vs. 6.3 in evaluation), suggesting it may not adapt well to demand shifts without recalibration.
- Deep Learning Models (TiDE, N-BEATS) underperformed, particularly in backtesting (TiDE MAPE: 15.8%, N-BEATS: 14.5%), indicating potential overfitting or insufficient training data. These models are not recommended for this use case without substantial architectural modifications.

Therefore, we recommend that the sponsoring company adopt TCN as the primary model, with Exponential Smoothing as the fallback. This strategy will maximize TCN's advanced capabilities for high-stakes forecasts while maintaining traditional methods for stable, low-resource scenarios, thereby obtaining a notable improvement in forecast accuracy. Currently, the demand forecasting inputs of the sponsor company are derived from quarterly sales and operations planning (S&OP) meetings and are frequently adjusted based on daily operations. We recommend implementing the model first as a baseline reference for the existing S&OP process instead of a total replacement. This approach will reduce the amount of work required to prepare for the S&OP meeting while allowing the team to transition smoothly from the current manual process to automated demand forecasting. Moreover, having input from the team during the S&OP meeting will help clarify any doubts and identify potential ways to further improve the model to best fit the sponsor company.

As our model is built on data from one GeoUnit of the sponsoring company, the immediate next step is to expand the application to a global scale. Additionally, incorporating historical and future S&OP records as covariates into the model merits further study regarding its effects on forecast accuracy and robustness.

In summary, this capstone project offers a valuable approach to demand forecasting for the logging and drilling tools of the sponsor company with improved accuracy and efficiency. Enhanced forecast accuracy will lead to reductions in operational and inventory costs. It will also help optimize the CAPEX plan and resource allocation across the company. Moreover, superior efficiency will reduce the time and effort required in the demand forecasting process. All combined, these improvements will enhance operational efficiency and resource utilization for the sponsoring company. The adoption of this technology in supply chain management will contribute to the competitive advantage of the sponsoring company, reinforcing its leading position in the oilfield service industry.

References

Azhar, M. (2013). The study of the bullwhip effect in the oil and gas industry. *Oklahoma State University ProQuest Dissertations & Theses.* https://www.proquest.com/dissertations-theses/study-bullwhip-effect-oil-gas-industry/docview/1517983954/se-2.

Balcerzak, A. P., Uddin, G. S., Igliński, B., & Pietrzak, M. B. (2023). Global energy transition: From the main determinants to economic challenges. Equilibrium. *Quarterly Journal of Economics and Economic Policy*, 18(3), 597–608. https://doi.org/10.24136/eq.2023.018.

BP (2024, July 20). Energy Demand. https://www.bp.com/en/global/corporate/energy-economics/energy-outlook/energy-demand.html.

Cansado-Bravo, P., & Rodríguez-Monroy, C. (2018). Persistence of Oil Prices in Gas Import Prices and the Resilience of the Oil-Indexation Mechanism. The Case of Spanish Gas Import Prices. *Energies*, vol. 11, no. 12, pp. 3486. https://doi.org/10.3390/en11123486.

DNV (2024). Energy Transition Outlook 2024. https://www.dnv.com/energy-transition-outlook/download/.

Das, A., Kong, W., Leach, A., Mathur, S., Sen, R., & Yu, R. (2023). Long-Term Forecasting with TiDE: Time-Series Dense Encoder. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.08424.

Dujak, D., Šebalj, D., & Koliński, A. (2019). Towards Exploring Bullwhip Effects in Natural Gas Supply Chain. *LogForum*, Vol. 15. http://doi.org/10.17270/J.LOG.2019.369.

Elder, J., Serletis, A., (2010). Oil price uncertainty. *J. Money Credit Bank*. 42, 1137–1159. https://doi.org/10.1111/j.1538-4616.2010.00323.x.

Fernandes, L.J., Barbosa-Póvoa, A. P., & Relvas, S. (2010). Risk Management Framework for the Petroleum Supply Chain, *Risk Management*, 28, pp. 157–158.

He, Y., & Zhao, J. (2019). Temporal Convolutional Networks for Anomaly Detection in Time Series. Journal of Physics. Conference Series, vol. 1213, no. 4, pp. 42050, https://doi.org/10.1088/1742-6596/1213/4/042050.

Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: principles and practice*, 3rd edition, OTexts: Melbourne, Australia. https://otexts.com/fpp3/.

Jacoby, D.S., & Gupta, A.R. (2021). *Reinventing the Energy Value Chain - Supply Chain Roadmaps for Digital Oilfields through Hydrogen Fuel Cells*. PennWell Corporation. Chapter 8 & Appendix 1.

Lisitsa, S., Levina, A., & Lepekhin, A. (2019). Supply-chain management in the oil industry. *E3S web of conferences*, 2019, Vol.110. https://doi.org/10.1051/e3sconf/201911002061.

Menhat, M.N.S. (2017). Performance measurement framework for the oil and gas supply chain. University of Central Lancashire (United Kingdom) ProQuest Dissertations & Theses. https://www.proquest.com/dissertations-theses/performance-measurement-framework-oil-gas-supply/docview/2164133874/se-2.

Oreshkin, B.N., Carpov, D., Chapodos, N., & Bengio, Y. (2019) N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1905.10437.

Perez-Guerra, U., Macedo, R., Manrique, Y. P., Condori, E. A., Gonzáles, H. I., Fernández, E., Luque N., Pérez-Durand, M. G., & García-Herreros, M. (2023). Seasonal autoregressive integrated moving average (SARIMA) time-series model for milk production forecasting in pasture-based dairy cows in the Andean highlands. *PLoS One*, 18(11). https://doi.org/10.1371/journal.pone.0288849.

Rahman, S., Serletis, A., (2012). Oil price uncertainty and the Canadian economy: evidence from a VARMA, GARCH-in-mean, asymmetric BEKK model. *Energy Economics*. 34, 603–610. https://doi.org/10.1016/j.eneco.2011.08.014.

Serletis, A., & Xu, L. (2016). Volatility and a Century of Energy Markets Dynamics. *Energy Economics*, vol. 55, pp. 1–9. https://doi.org/10.1016/j.eneco.2016.01.007.

Sheremetov, Leonid B., González-Sánchez, A., López-Yáñez, I., & Ponomarev, A.V. (2013). Time Series Forecasting: Applications to the Upstream Oil and Gas Supply Chain. *IFAC Proceedings Volumes*, vol. 46, no. 9, pp. 957–62, https://doi.org/10.3182/20130619-3-RU-3018.00526.