Journal of Academic Research Vol. 30, Issu 02, 2026 "210-222"

Multiple Linear Regression Models for Crude Oil Forecasting in the A-NC115 Field, Libya

Abdulkader Fauzi Aissa Alagili
Akakus Oil Operations Company, Libya

DOI: https://doi.org/10.65540/wfrjkt63

Article information

Key words
El-sharara A-NC115 Libyan oilfield, Multiple Linear Regression, Oil Production Forecasting Model.

Received 10 05 2026,
Accepted 03 09 2026,
Available online 05 09 2026

Abstract

Accurate multi-horizon production forecasting remains a critical challenge in mature Libyan oilfields. In this paper, a Multiple Linear Regression (MLR) approach is proposed to forecast crude oil production in the A-NC115 oilfield, Libya, by identifying the main influencing variables. Unlike most previous studies conducted in Nigerian and other international fields, this study applies MLR forecasting to a mature Libyan oilfield, addressing an important regional research gap. The A-NC115 field is a mature water-supported reservoir where accurate forecasting of oil production is of essential importance. The trends of production, pressure, and water injection directly impact the operational and development decisions. the three different models were developed and evaluated: a short-term model based on daily data, a mid-term model based on monthly data and a long-term model based on yearly data. Historical production data from February 2018 to December 2023 was collected, pre-processed and analyzed using Minitab software. Model performance was evaluated on a hold-out testing dataset (Jan 2024 - Apr 2024) using statistical metrics such as R-squared (R^2), Adjusted R^2, Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The best accuracy was achieved for the daily model (R^2= 99.99%, MAPE = 7.20%) followed by the yearly model (R^2= 99.84%, MAPE = 8.67%). The monthly model is still quite explanatory (R^2= 99.96%) but with a higher MAPE (17.87%) showing a higher variance of prediction at this specific temporal scale. The results suggest that data-driven models provide accurate forecasts that support improved resource management, maintenance planning, and more efficient operational and strategic decision-making, along with a better understanding of future needs such as equipment upgrades, associated water management, and production optimization.

I. Introduction

Libya depends on its hydrocarbon industry for most of its economic activity because oil exports provide about 95 percent of government revenues. The national economy needs accurate oil production forecasts because it serves as an essential operational requirement for the country. The field operations rely on accurate production forecasts which help operators to distribute their resources efficiently and plan their maintenance activities and water injection systems and make important choices about their well operations and drilling projects [1,2].

abdulkader.alagili@akakusoil.com

The A-NC115 field which Akakus Oil Operations Company (AOO) located in the Murzuq Basin stands as the largest field within the El Sharara complex because it achieves daily production rates that exceed 70000 barrels [3,4,11]. The development of precise forecasting models for A-NC115 holds vital importance because the location has significant economic value. The available forecasting methods include decline curve analysis and numerical simulation and advanced machine learning techniques [5]. The multiple linear regression (MLR) method enables researcher to achieve three essential goals because it maintains system transparency while processing quickly and delivering accurate results when production factors show linear behavior [6].

While MLR has been widely used for production forecasting in other regions, particularly Nigerian oilfields, its application in mature Libyan reservoirs remains limited. Therefore, this study addresses this gap by developing and validating multi-horizon MLR forecasting models for the A-NC115 field using real field data.

Whereas the primary objective of this research is to develop a practical, reliable, and user-friendly approach for oil production forecasting that can be readily applied by reservoir and production engineers without the need for complex workflows, advanced computational resources, or expensive specialized software.

Production forecasting is a critical component of reservoir management and field development planning. Over the years, numerous forecasting techniques have been developed, including Decline Curve Analysis (DCA), numerical simulation models, machine learning algorithms, and artificial intelligence-based approaches. Although these methods can provide accurate predictions under certain conditions, they often require extensive data preparation, specialized expertise, significant computational effort, and access to costly software platforms. These limitations may restrict their practical application, particularly in environments where resources and software accessibility are limited.

A review of the existing literature indicates that Multiple Linear Regression (MLR) remains a promising alternative for production forecasting due to its simplicity, transparency, and ease of implementation. When supported by appropriate data preprocessing, rigorous variable selection, and systematic model development procedures, MLR can achieve a high level of forecasting accuracy and reliability comparable to more sophisticated forecasting techniques.

Based on the previous studies, it may be seen that there are similarities between contemporary studies and these studies, including the use of more than one MLR model and the identification of variables affecting oilfield output [1,2,6,9].

However, the principal distinction between these preceding studies and this study, the focus on the Akakus oil field, in particular on developing prediction models to forecast oil production in the short, medium, and long term.

As for the one-of-a-kind capabilities of the current study, it would take MLR model to forecast oil production in Akakus oil area A NC115. The observe targets to develop correct and dependable prediction models that could help in decision-making concerning oil production and meet the increasing local and global demand. The outcome of this study might also have implications for policymakers, engineers, and other experts operating within the Libyan energy industry.

Also, most of previous studies have utilized Multiple Linear Regression for oil production forecasting, the present study adopts a different perspective. Rather than focusing solely on developing a prediction model, this research proposes a structured and practical modeling framework that guides engineers through data preparation, variable screening, model development, validation, and performance evaluation. The proposed methodology aims to establish a reliable and reproducible workflow for building accurate forecasting models that are easy to use, practical to implement, and suitable for real fields applications [6,9].

Therefore, this study seeks to demonstrate that a carefully designed Multiple Linear Regression approach, combined with a systematic modeling procedure, can provide an effective, accessible, and highly reliable forecasting tool that helps engineers easily develop and apply forecasting models through a simple, practical, and fit-for-purpose methodology

The main study question is "Can Multiple Linear Regression models provide accurate and interpretable production forecasts for the A-NC115 field across daily, monthly, and annual time horizons? "

This study addresses a gap in the existing literature by developing and evaluating multi-horizon forecasting models for a Libyan oilfield, a context that has been largely underrepresented in MLR-based production studies. The results demonstrate that previous studies have applied multiple linear regression MLR methods to oil fields located in Nigeria [1,4,14] and worldwide [16] while only one study has created matching short-term and mid-term and long-term forecasting models. The primary objective of this study is to develop, validate, and compare three MLR models for forecasting oil production in the A-NC115 field across different time horizons (daily, monthly, yearly). The project aims to accomplish specific objectives which include:

1. Identifying the most significant variables that impact on production for each time horizon.
2. Building and test MLR models for short-term, mid-term, and long-term forecasts.
3. Evaluating and comparing the accuracy of the three models using standard statistical metrics.
4. Provide valuable insights for AOO to enhance field management and decision-making processes.

A. Objectives

The main objective of the study as listed below:

1. To develop an accurate forecasting model for short-term, mid-term, and long-term production prediction in the A-NC115 Oil Field.
2. To identify the key variables that significantly influence oil production in each forecasting model through MLR method using Minitab software.
3. To evaluate the performance of the developed forecasting models by comparing the forecasted production values with the actual production data.
4. To provide insights and recommendations for improving production efficiency, optimising resource allocation, and enhancing decision-making processes in the A-NC115 Oil Field using the developed forecasting models.
5. To contribute to the understanding of production patterns and trends in the A-NC115 Oil field through the application of forecasting techniques.
6. To contribute to the body of knowledge in oilfield production forecasting and provide a valuable resource for oilfield operators and decision-makers.

B. Hypotheses

The main hypotheses of the study as listed below:

1. The developed forecasting models using MLR will accurately predict production levels in the A-NC115 Oil Field across different time horizons (short-term, mid-term, and long-term).
2. There are significant variables that influence oil production in the A-NC115 Oil Field.
3. The forecasted production values from the developed models will exhibit a high degree of accuracy and reliability when compared to the actual production data in the A-NC115 Oil Field.
4. The utilization of the developed forecasting models will lead to optimised resource allocation and production strategies in the A-NC115 Oil Field, resulting in improved production performance.

These hypotheses focus on the accuracy of the forecasting models, identification of significant variables, performance evaluation, and the impact of the models on production efficiency and decision-making processes.

II. Methodology

A. study Area and Data Collection

The study focuses on the A-NC115 field. AOO's databases provided access to Historical production and operational data from February 2018 until April 2024 which illustrated in fig. 1. This dataset includes daily, monthly, and yearly records of oil production, water production, water injection rates, gas-oil ratios (GOR), bottom-hole pressures (SBHP and FBHP), the number of active oil and injection wells, and the Basic Sediment & Water (BS&W) content [7].

The objective of the study was to develop forecasting models representative of the current reservoir and operational conditions. Therefore, the period from 2018 to 2024 was selected because it reflects the most recent production strategy, mature waterflooding performance, and current field operating practices. Earlier production periods were characterized by different development stages, changing operational strategies, and reservoir conditions that are not representative of the field's current behavior. Using recent data improved model relevance and applicability to present-day forecasting and decision-making.

The contribution of this study is not limited to developing forecasting models for a single oilfield. Rather, the main contribution lies in establishing a systematic, practical, and easy-to-use forecasting methodology that can guide reservoir and production engineers through data preparation, variable selection, model development, validation, and performance evaluation. Although the developed models were calibrated specifically for the A-NC115 field, the methodology itself can be applied to other oilfields to obtain reliable forecasting results after field-specific calibration. Furthermore, the objective was not only to achieve accurate forecasts, but also to provide a transparent and efficient workflow that reduces time, effort, and computational requirements compared with more complex forecasting approaches while maintaining a high level of prediction accuracy.

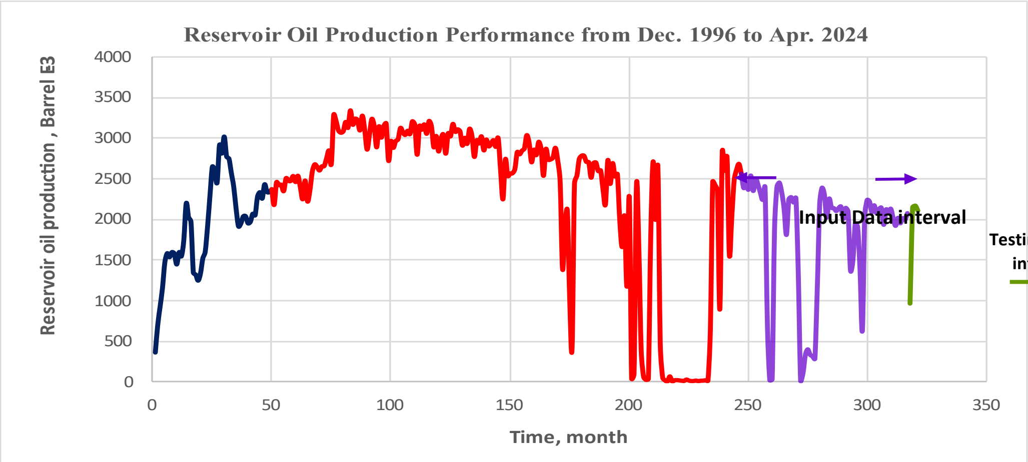

Figure (1): Reservoir oil production performance from Dec. 1996 to Apr. 2024.

B. Data Pre-processing

The data required extensive preprocessing efforts prior to its usage for modeling purposes:

- Data Cleaning the outliers and erroneous data points had to be identified and deleted.[8]
- Missing Data Handling The missing time periods in the data had to be addressed using either interpolation techniques or by excluding those gaps based on certain considerations.
- Feature Engineering - In order to capture the temporal trend in the data, lagged features were generated along with a set of moving average features using 2, 3, and 4 periods' duration to track the key measures, including the oil production and water injection.[9]
- Normalization Transformation—The Box-Cox transformation was conducted on the dependent variable (oil production) to make sure that the residuals in the regression model met the normality condition [10], which is crucial for proper statistical analysis as shown in fig. 3 and fig. 4.

The model was created under stable operational conditions. The time periods when the system experienced extreme operational issues such as the force majeure events and initial water injection were not considered during creating. The data from January 2018 till December 2023 was utilized to create the model. The data between January and April 2024 served for out-of-sample testing as explained in Fig. 2. So, the purpose of the developed models is to forecast production under normal operating conditions rather than predict exceptional events. Force majeure periods, prolonged shutdowns, and the initial water injection phase introduce abrupt operational changes that do not reflect the reservoir's typical performance behavior. Including such events may distort the statistical relationships between production and predictor variables and reduce forecasting accuracy during routine operations. Nevertheless, future studies may investigate dedicated models for abnormal operating conditions and production disruptions to further improve forecasting capabilities under non-routine scenarios.

C. Model Development

MLR models were developed using three different time intervals:

- Short Term: Using daily production information. It began with 23 candidate variables that are generated from daily production data.
- Mid-Term: Using monthly production information. It began with 17 candidate variables that are generated from daily production data.
- Long-Term: Using annual production information. It began with 14 candidate variables that are generated from daily production data.

The equation for MLR is expressed by:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \quad (1)$$

Where `$Y$` is the oil production (dependent variable), `$X_1 \ldots X_n$` are the independent predictor variables, `$\beta_0 \ldots \beta_n$` are the regression coefficients, and `$\varepsilon$` is the error term [11].

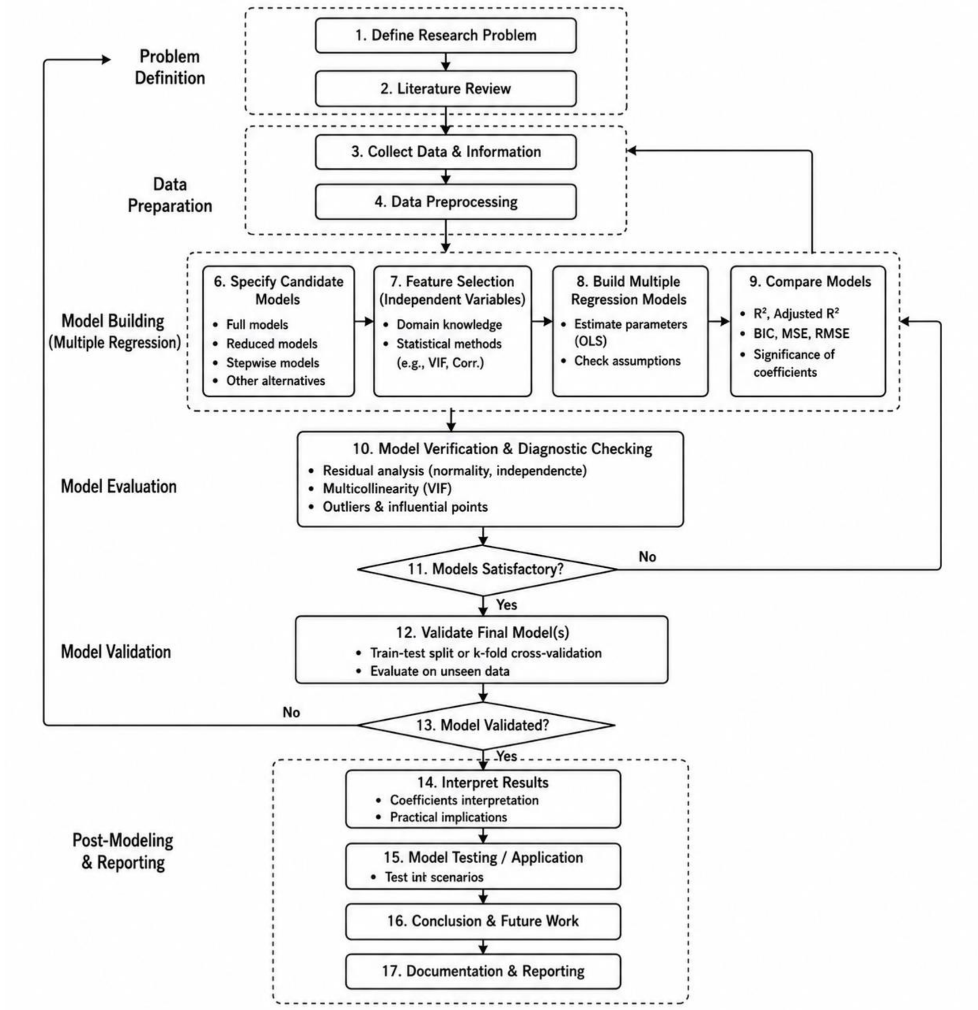

Figure (2): Modeling steps

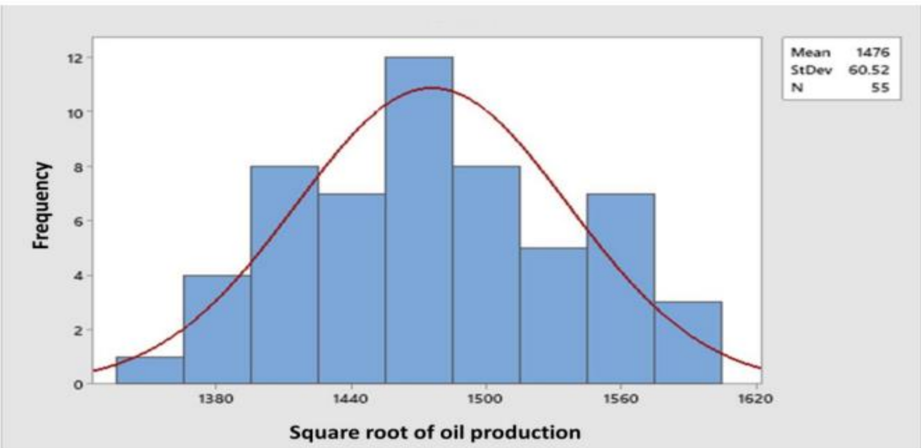

Figure (3): Histogram of the Square root oil production

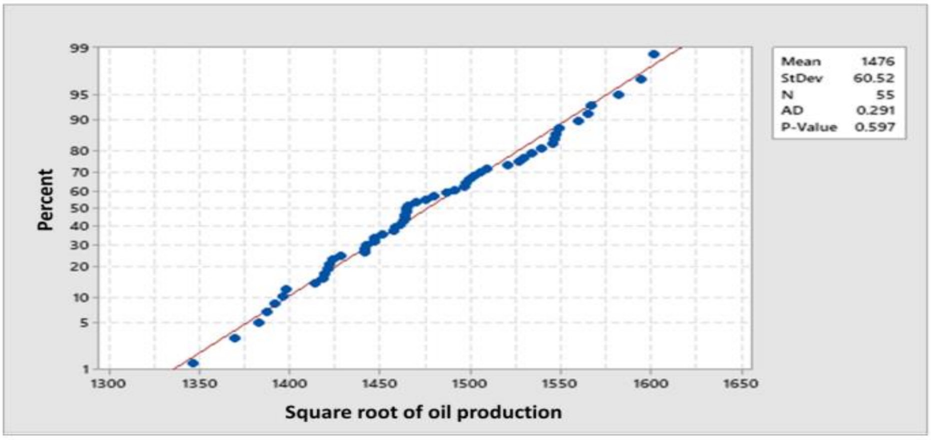

Figure (4): Normality Test

D. Variable Selection and Model Refinement

Stepwise Regression is applied by Minitab software to determine the statistically significant variables that need to be considered in every model. It is done automatically by the system were depending on their p-values' results, predictors are removed or added to the model considering a level of significance of α = 0.05. In such a way, only those variables having statistically significant influence on production are included in the model thereby making the model more parsimonious and generalizable [12,13,14]. Also, Stepwise Regression was selected because the study initially involved a relatively large number of potential predictor variables and aimed to identify the most influential parameters affecting oil production at different forecasting horizons. The method provides a practical and transparent approach for variable screening and model simplification, which is particularly valuable for operational applications. To reduce the risk of overfitting, only statistically significant variables (p-value < 0.05) were retained, and model performance was validated using an independent testing dataset. In addition, multiple performance indicators (R^2, Adjusted R^2, MAE, RMSE, and MAPE) were used to evaluate model reliability.

To further verify the statistical validity of the developed models, Analysis of Variance (ANOVA) was performed during the regression analysis process. Variables were selected based on their statistical significance (p-value < 0.05), while model adequacy was assessed using ANOVA results, R^2, Adjusted R^2, and residual analysis. This procedure ensured that only statistically significant predictors were retained in the final forecasting models.

In addition to the stepwise regression procedure, multicollinearity among the candidate predictor variables was evaluated using the Variance Inflation Factor (VIF). Variables with high multicollinearity were excluded to improve model stability and interpretability. Only statistically significant predictors (p-value < 0.05) with acceptable VIF values were retained in the final models. This procedure is just to ensure that the selected variables contributed independently to the prediction of oil production and to ensure model adequacy and reliability.

E. Model Validation and Performance

The test set, gathered between January and April 2024, was the evaluation technique for all final models. The following measures of performance were computed to measure accuracy and reliability:

- Coefficient of Determination (R^2) and Adjusted R^2: Used to gauge how well the model predicts the variability of the dependent variable [15]. The adjusted R^2 punishes the inclusion of irrelevant variables[21] .
- Mean Absolute Error (MAE): This is the average absolute error.
- Root Mean Squared Error (RMSE): This is the square root of the average squared error. RMSE is more sensitive to outliers than MAE [16,17].
- Mean Absolute Percentage Error (MAPE): This is the average percentage error. It gauges prediction accuracy, and its categories include Table 1 [18]:

Table (1): The Criteria of MAPE for Model Evaluation Based on Lewis [18].

MAPE, %	Evaluation
MAPE ≤10%	High accuracy prediction
10% < MAPE ≤ 20%	Good prediction
20% < MAPE ≤ 50%	Reasonable prediction
MAPE >50%	Inaccurate prediction

III. Results and Discussion

Final Model Equations and Significant Variables are selected Although stepwise regression has known limitations, it was selected due to its interpretability and suitability for operational field data where explain ability is crucial [19]. The stepwise regression process yielded the following final models, each with a reduced set of highly significant predictors (p-value < 0.05)[20] [22] .

- **Short-Term (Daily) Model:**

$$Y = 135.784 - 0.000009X_5 + 0.003708X_9 - 0.001851X_{13} + 0.000002X_{16} - 0.001352X_{19} + 0.03005X_{20} + 0.00584X_{21} - 0.00931X_{22} + 0.000082X_{23} \quad (2)$$

Where:
X_5: Moving average of Gross Oil produced (bpd) (4 periods).
X_9: Moving average of Net oil production rate (2 periods).
X_{13}: Net oil production rate of last period.
X_{16}: Moving average of Water injection (4 periods).
X_{19}: Last period's Static Bottom Hole Pressure (SBHP).
X_{20}: Number of oil wells.
X_{21}: Last period's BS&W.
X_{22}: Recovery percent (RF%) in the previous period.
X_{23}: Gross Gas Production in the previous period.

- **Mid-Term (Monthly) Model:**

$$Y = 717.43 - 0.000341X_3 + 0.2943X_7 + 0.000682X_{10} \quad (3)$$

Where:
X_3: Monthly oil production rate in the previous time.
X_7: Recovery rate percent (RF%) in the previous time.
X_{10}: Moving average of monthly oil production rate (2 periods).

- **Long-Term (Yearly) Model:**

$$Y = 2325.8 + 0.000211X_1 - 0.000104X_3 + 0.000020X_4 - 0.000006X_7 \quad (4)$$

Where:
X_1: Moving average of yearly oil production rate (2 periods).
X_3: Yearly oil production rate in the previous time.
X_4: Moving average of yearly water production rate (2 periods).
X_7: Moving average of yearly water injection (2 periods).

The difference in the number of selected variables between the daily and monthly models is primarily related to data resolution and variability. The daily dataset contains a significantly larger number of observations and captures short-term operational fluctuations, pressure changes, injection variations, well status changes, and other field dynamics. Consequently, more variables remain statistically significant and contribute to forecasting performance. In contrast, monthly aggregation smooths much of this operational variability, reducing the amount of information contained in the dataset and resulting in fewer variables being identified as significant predictors through the Stepwise Regression process. Therefore, only three variables were sufficient to explain most of the monthly production behavior.

The results show that main production drivers change according to different time measurement periods. The short-term operational changes of the system show strong dependency on current operational parameters which include well count and pressure and recent production and injection patterns [20,21,22]. The field's complete production record together with its recovery factor determines the mid-term development of its production activities. Long-term projections depend on production and water management which have been tracked over multiple years [23,24,25].

The Significant changes in operating conditions, well configuration, or water injection strategy may affect the validity of the developed models. In such cases, model updating or recalibration using recent data is recommended to maintain forecasting accuracy. This does not imply that the methodology becomes invalid, but rather that the forecasting equations should be periodically updated to reflect changing field conditions.

The regression equations developed in this study are specific to the A-NC115 field and should not be directly applied to other reservoirs. However, the proposed methodology can be implemented in other Libyan oilfields after recalibration using field-specific data. Similar levels of forecasting accuracy may be achieved if the same systematic workflow is followed and the models are properly adapted to the reservoir and operational conditions of the target field.

3.1 Model Performance Evaluation

The performance metrics for the three models on the testing dataset are summarized inTable 2.

Table (2): Overall comparison between the models.

Statistical results of the regression analysis

Overall measures error values of different models for crude oil production prediction

Model	Error Type	Error	**Evaluation**
Short-term model daily	MAD	3820	High accuracy prediction
	MSE	120727612	
	RMSE	10988	
	MAPE (%)	7.20	
	R^2	99.99%	
	Adjusted R^2	99.99%	
Mid-term model monthly	MAD	551265	Good prediction
	MSE	548618500975	
	RMSE	740688	
	MAPE (%)	17.869	
	R^2	99.96%	
	Adjusted R^2	99.95%	
Long-term model yearly	MAD	2230243	High accuracy prediction
	MSE	6049288300525	
	RMSE	2459530	
	MAPE (%)	8.666	
	R^2	99.84%	
	Adjusted R^2	99.79%	

- **Short-Term Forecasting Model:** This model produced excellent results with near-perfect R^2 values and an error of 7.20% MAPE, making it a very accurate predictive model. The production of A-NC115 on a daily basis is predictable with great accuracy using historical data and present-day situations as shown in the fig. 5.
- **Mid-Term Forecasting Model:** Even though the model achieved a very high R^2 value (99.96%), indicating that it explains nearly all the variability in the monthly production data, its MAPE value (17.87%) was considerably higher than those of the daily and yearly models. This is not contradictory, as R^2 and MAPE assess different aspects of model performance. While R^2 measures the model's ability to explain data variability, MAPE evaluates the magnitude of prediction errors. Monthly data aggregation tends to smooth short-term production fluctuations, allowing the model to capture the overall production trend effectively. However, monthly production remains sensitive to operational interruptions, maintenance activities, production adjustments, and temporary field events, which can increase forecasting uncertainty and prediction errors. Consequently, the model demonstrates excellent explanatory power while exhibiting a higher forecasting error relative to the daily and yearly models.

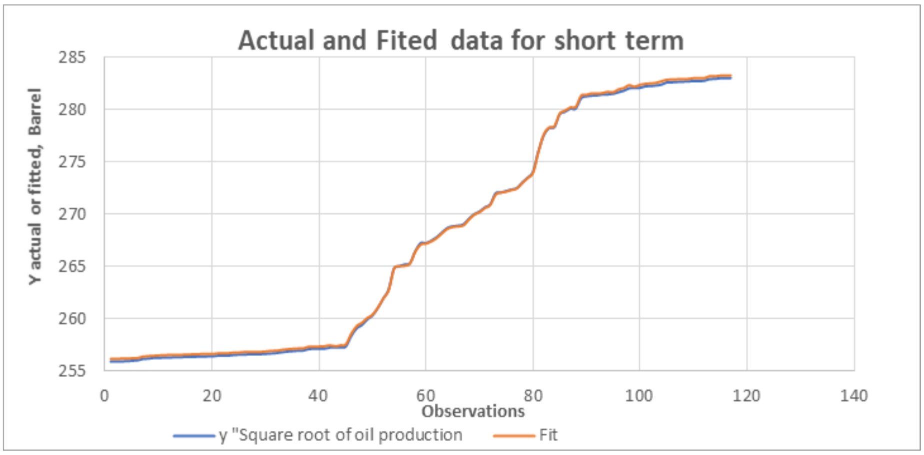

Figure (5): Fits and actual for unusual observations for short-term model

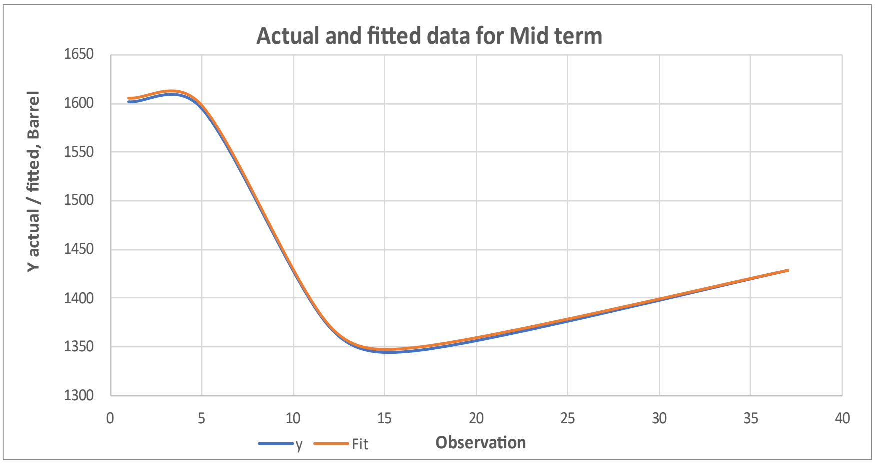

Figure (6): Fits and actual for unusual observations for Mid-term model.

The difference arises because even though the model has accurately reflected past trends, the variance in its monthly predictions is considerably larger compared to the other two models, and hence it is classified as a "Good" prediction model as shown in the fig. 6.

- **Long-Term Forecasting Model:** As discussed above, this model is considered to have "High Accuracy" with a MAPE of 8.67%. The high R^2 and MAPE values of this model mean that annual production trends are largely dictated by production and water injection history over the multi-year period as shown in the fig. 7.

From an operational perspective, the short-term model can support day-to-day production monitoring and operational optimization, while the yearly model can assist in long-term planning, reserve management, budgeting, and strategic decision-making. The monthly model may be used as a supplementary planning tool and should be interpreted alongside short-term forecasts. In overall the above comparison indicates that for the A-NC115 field, MLR performs exceedingly well in daily and annual forecasting. The monthly forecasting model, although statistically sound, should be interpreted cautiously.

The high R^2 values obtained in this study can be partly explained by the characteristics of the A-NC115 reservoir, which is a relatively homogeneous reservoir with permeability reaching approximately 1 to 5 Darcy and a stable production history. In addition, the selected predictor variables are physically related to oil production behavior, including production, pressure, and water injection parameters. To minimize the risk of data leakage and overfitting, the models were trained using historical data from 2018 to 2023 and validated on an independent testing dataset covering January to April 2024. The selected modeling period was intentionally chosen to represent the most recent and stable operating conditions of the field while excluding periods affected by force majeure events, prolonged shutdowns, and major operational disruptions that could distort the production trends. Consequently, the developed models were calibrated using representative operating data and validated using unseen data, supporting the reliability of the reported forecasting performance.

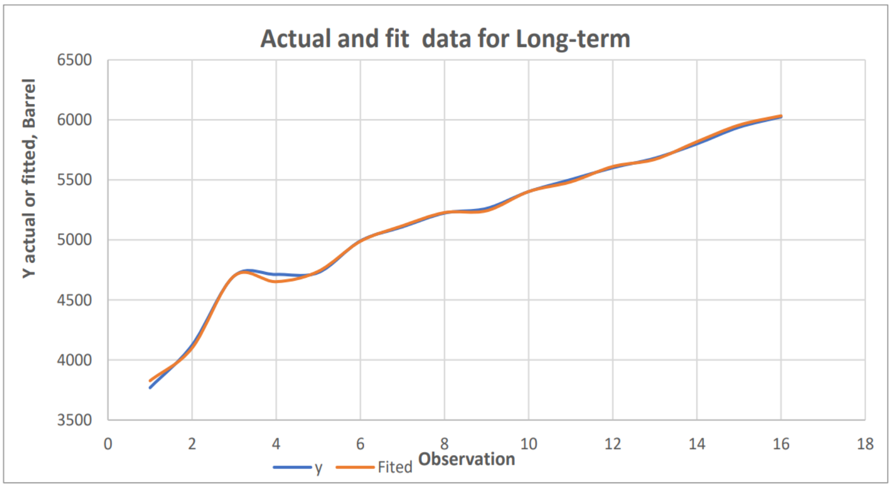

Figure (7): Fits and actual for unusual observations for Long-term model.

The four-month testing period was used as an independent out-of-sample dataset to evaluate model performance under unseen operating conditions. While the results demonstrated strong predictive capability, a longer validation period could further strengthen confidence in the model's robustness and generalization. This limitation has been acknowledged, and future studies are recommended to validate the models using longer testing intervals and extended historical datasets.

The contribution of this study is not the development of a new statistical algorithm, but the application and validation of a systematic multi-horizon forecasting framework in the A-NC115 field, a mature and relatively homogeneous Libyan reservoir that has been underrepresented in the literature. The study demonstrates that a carefully developed MLR methodology can provide highly accurate daily, monthly, and yearly forecasts using real field data. In addition, it addresses a regional research gap and provides a practical workflow that can be directly applied by reservoir and production engineers for operational, planning, and strategic decision-making.

IV. Conclusion and Recommendations

Three models based on MLR have been successfully developed and validated in this study for estimating the production of crude oil produced in the A-NC115 field within different periods of time. Also, the results confirmed the overall statistical significance of the developed regression models, supporting the reliability of the selected predictor variables and the forecasting equations. As result It has been shown that MLR can be used as an effective method for production forecasting when using the right methodology, as shown on modeling steps in fig 2.

The forecasting accuracy achieved by the developed MLR models is consistent with previous studies that successfully applied regression-based approaches to oil production forecasting. Furthermore, advanced forecasting techniques such as Artificial Neural Networks (ANN) have also shown promising results in petroleum applications. For example, Elmabrouk et al. (2014) demonstrated the applicability of ANN models for oil production prediction in Libyan oilfields. While ANN models may capture complex non-linear relationships, the MLR approach adopted in this study offers the advantages of simplicity, transparency, and ease of implementation, making it particularly suitable for operational field applications [3].

1. High accuracy was reached due to results received from both short-term daily and yearly long-term models as their accuracy level was estimated using MAPE with values lower than 10%. These models can be considered vital tools for decision-making.
2. Factors that influence prediction accuracy. Short-term predictions are based on detailed operational data, and long-term predictions are based on trends from previous periods.

3. Model package includes various modeling opportunities due to several models provided by this package. Daily model can be used for operational adjustment; at the same time, annual model should be used in cases of budget planning.
4. The daily production management model used in the field needs to be integrated with the short-term daily model, which will provide forecasts and alert about possible deviations in operations.
5. Annual model can be used for budgeting purposes and updating reserves estimation.
6. The organization needs to implement an ongoing system which will validate and clean data to maintain accurate and trustworthy model inputs.
7. The organization needs to conduct annual model re-calibration which requires updating all models with current data to reflect changes in reservoir behavior and operational practices.
8. The mid-term model's trend analysis should be used together with the more accurate short-term model's rolling forecasts for monthly planning because this combination will help reduce MAPE which results from using the mid-term model.
9. Future studies should investigate hybrid models that combine MLR with machine learning techniques which include Artificial Neural Networks and Long Short-Term Memory networks because this combination will enable the detection of complex non-linear relationships which will result in better mid-term forecasting results.
10. It is recommended that future studies incorporate longer historical datasets and external factors, such as oil prices and geopolitical conditions, to improve the generalizability and long-term applicability of the forecasting models.
11. Future studies are recommended to compare the developed MLR models with advanced machine learning techniques, such as Artificial Neural Networks (ANN) and other ML-based approaches, to assess potential improvements in forecasting accuracy, robustness, and model generalization.
12. The study requires dynamic geological and geophysical data to create improved models which will account for underground material variations.
13. This study must be carried out to assess how external variables like market price and geopolitical situations affect production forecasts based on their long-term nature.
14. External Factors Analysis: study the effect of external factors like market price or geopolitical situations on the production forecast, especially the long-term ones.

Conflict of interest statement: "The authors declare no conflicts of interest".
Author Contributions: "All authors have made a substantial, direct, and intellectual contribution to the work and approved it for publication."
Funding: "This research received no external funding."
Data Availability Statement: "No data were used to support this study."

V. References

[1] Oladeinde, M. H., Ohwo, A. O., & Oladeinde, C. A. (2015). A mathematical model for predicting output in an oilfield in the Niger Delta Area of Nigeria. Nigerian Journal of Technology,34(4), pp768-772. https:// doi.org/10.4314/njt.v34i4.14
[2] Mustafar, I. (2011). A research on prediction of output in oilfield using multiple linear regression. Perak, Malaysia: International Journal of Applied Science and Technology Vol. 1 No.4 .
[3] Madhukumar, M., Sebastian, A., Liang, X., Jamil, M., & Shabbir, M. N. S. K. (2022). Regression model-based short-term load forecasting for university campus load. IEEE Access, 10, pp. 8891-8905. https:// doi.org/10.1109/access.2022.3144206
[4] AlRassas, A. M., Al-qaness, M. A., Ewees, A. A., Ren, S., Abd Elaziz, M., Damaševičius, R., & Krilavičius, T. (2021). Optimized ANFIS model using Aquila Optimizer for oil production forecasting. Processes, 9(7), p. 1194. https://doi.org/10.3390/pr9071194
[5] Zhu, C., & Zhao, X. (2009). Application of Artificial Neural Network in the Prediction of Output in Oilfield. International Joint Conference on Artificial Intelligence. https:// doi.org/10.1109/jcai.2009.93
[6] Study by Guo, L., & Deng, X. (2009) entitled "Application of improved multiple linear regression method in oilfield output forecasting". In 2009 International conference on information management, innovation management and industrial engineering Vol. 1, pp. 133-136. https://doi.org/10.1109/iciii.2009.39
[7] Akukas Oil Operations Company. (2023). All Field data and maps related to El Sharara oil field, Al-Tahadi Street, Airport Road, Tripoli, Libya .

[8] Izni, G.M. and Radzuan, R. (2011). A Study on Prediction of Output in Oilfield Using Multiple Linear Regressions. International Journal of Applied Science and Technology, vol. 1, no. 4, pp. 21-27.
[9] Yale University. (1997). Multiple Linear Regression. Retrieved September 19, 2010, Available: https://www.stat.yale.edu/Courses/1997-98/101/anovareg.htm#mult
[10] Sun, Y., Wang, X., Zhang, C., & Zuo, M. (2023). Multiple Regression: Methodology and Applications. Highlights in Science, Engineering and Technology, 49, pp. 542-548. https://doi.org/10.54097/hset.v49i.8611
[11] Alin, A. (2010). Minitab. Wiley interdisciplinary reviews: computational statistics, 2(6), pp. 723-727. https://doi.org/10.1002/wics.113
[12] Hosseini, S., & Akilan, T. (2023) 'Advance deep regression models for forecasting time series oil production' arXiv preprint arXiv:2308.16105. https://doi.org/10.48550/arXiv.2308.16105
[13] AlRassas, A. M., Al-Qaness, M. A., Ewees, A. A., Ren, S., Sun, R., Pan, L., & Abd Elaziz, M. (2022) entitled 'Advance artificial time series forecasting model for oil production using neuro fuzzy-based slime mould algorithm', Journal of Petroleum Exploration and Production Technology, pp. 1-13. https://doi.org/10.1007/s13202-021-01405-w
[14] Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). Applied linear regression models (4th ed.). New York: McGraw-Hill/Irwin.
[15] Ifowodo, H. F., Chinwuko, E. C., & Mgbemena, C. O. (2020) "Forecast of Crude Oil Production Output in An Oil Field in the Niger Delta Region of Nigeria". International Journal of Industrial Engineering, 31(1), pp.161-170. https://doi.org/10.22068/IJIEPR.31.1.161
[16] Study by Ruiz, M., Espinosa, G. A., Obando, A., & Alvarez, H. (2019) "Combined artificial intelligence modeling for production forecast in an oil field" CT&F-Ciencia, tecnología y futuro,9(1)، pp.27-35. https://doi.org/10.29047/01225383.149
[17]Study by Aydin, G.Ö.K.H.A.N. (2015) Regression models for forecasting global oil production", Petroleum Science and Technology, 33(21-22), pp. 1822-1828.https://doi.org/10.1080/10916466.2015.1101474
[18] Lewis, C. D. (1982). Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting. Butterworth Scientific. https://doi.org/10.1002/for.3980020210
[19] U.S. Energy Information Administration (2024). Country Analysis Brief: Libya, www.eia.gov U.S. Department of Energy Washington, DC 20585.
[20] Pärt-Enander, E., & Sjöberg, A. (1999). The MATLAB 5 handbook (2nd ed.). Reading, MA: Addison-Wesley Longman.
[21] Wang, J. (1993). A theoretical discussion of production decline curve of oil and gas reservoirs. Petroleum Exploration and Development, China.
[22] Wang, T., Chen, X. G., & Li, Y. F. (2006). Optimization of multivariate model in oilfield output prediction. Computer Simulation, New York.
[23] Sokolov, S., (2016). Mathematical prediction model for base oil production with account of uncertainties based on the method of displacement characteristics. Tyumen State Univ. Herald. Phys. Math. Model. Oil Gas Energy 2 (1), 82–91. https://doi.org/ 10.21684/2411-7978-2016-2-1-82-91
[24] Uma, M. J., Yun, H., Jeong, C. S., and Heo, J. H. (2011). Factor analysis and multiple regression between topography and precipitation on Jeju Island, Korea. J. Hydrol. 410:189–203. https://doi.org/10.1016/j.jhydrol.2011.09.016
[25] Aydin, G. (2014). Production modeling in the oil and natural gas industry: An application of trend analysis. Pet. Sci. Technol 32:555–564. https://doi.org/10.1080/10916466.2013.825271

نماذج الانحدار الخطي المتعدد للتنبؤ بإنتاج النفط الخام في حقل A-NC115 النفطي، ليبيا

عبدالقادر فوزي عيسي العجيلي
شركة أكاكوس للعمليات النفطية، ليبيا

استلمت الورقة بتاريخ 2026/05/10، وقبلت بتاريخ 2026/09/03، ونشرت بتاريخ 2026/09/05

الكلمات المفتاحية: حقل الشرارة A-NC115 – الحقول النفطية الليبية، الانحدار الخطي المتعدد، نموذج التنبؤ بإنتاج النفط.

الملخص

لا يزال التنبؤ الدقيق بالإنتاج متعدد الآفاق الزمنية يمثل تحديًا كبيرا في الحقول النفطية الليبية الناضجة. في هذه الورقة، تم اقتراح منهجية الانحدار الخطي المتعدد (**MLR**) للتنبؤ بإنتاج النفط الخام في حقل **A-NC115** في ليبيا، من خلال تحديد المتغيرات الرئيسية المؤثرة وقياس علاقاتها الخطية، حيث يُعد التنبؤ الدقيق بإنتاج النفط أمرًا بالغ الأهمية. إن اتجاهات الإنتاج والضغط وحقن المياه تؤثر بشكل مباشر على قرارات التشغيل والتطوير. وفي هذه الدراسة قد تم تطوير نموذج قصير المدى يعتمد على البيانات اليومية، ونموذج متوسط المدى يعتمد على البيانات الشهرية، ونموذج طويل المدى يعتمد على البيانات السنوية. تم جمع بيانات الإنتاج التاريخية من فبراير 2018 إلى ديسمبر 2023، تم معالجتها مسبقًا وتحليلها باستخدام برنامج **Minitab**. كما تم تقييم أداء النماذج باستخدام بيانات اختبار مستقلة للفترة من يناير 2024 إلى أبريل 2024، وذلك باستخدام مقاييس إحصائية مثل معامل التحديد (**R^2**)، معامل التحديد المعدل (**Adjusted R^2**)، متوسط الخطأ النسبي المطلق (**MAPE**)، متوسط الخطأ المطلق (**MAE**)، وجذر متوسط مربع الخطأ (**RMSE**). أظهرت النتائج أن أفضل دقة كانت للنموذج اليومي (**R^2 = 99.99%**، **MAPE = 7.20%**)، يليه النموذج السنوي (**R^2 = 99.84%**، **MAPE = 8.67%**). أما النموذج الشهري فقد كان ايضا ذا قدرة تفسيرية عالية (**R^2 = 99.96%**) لكنه سجل قيمة أعلى لـ **MAPE** (17.87%) مما يشير إلى تباين أكبر في دقة التنبؤ على هذا المقياس الزمني. تشير النتائج إلى أن النماذج المبنية على البيانات توفر تنبؤات دقيقة تدعم تحسين إدارة الموارد، والتخطيط للصيانة، واتخاذ قرارات تشغيلية واستراتيجية أكثر كفاءة، مع فهم أفضل للاحتياجات المستقبلية مثل تطوير المعدات والتعامل مع المياه المصاحبة وتحسين الإنتاج.