Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20250402038
Volume 4, Issue2 (May-August 2025), PP: 290-294.
www.indjcst.com
ISSN No: 2583-5300

A Drought Forecasting Model Using the Prophet Time Series Analysis Technique

Erick Katumo[1], Dr. Charles Katila[2], Dr. Richard Omolo[3], Dr. Hadullo Ken[4]
[1,2,3,4] *Department of Computer Science and Information Technology at the Cooperative University of Kenya.*

To Cite this Article: *Erick Katumo[1], Dr. Charles Katila[2], Dr. Richard Omolo[3], Dr. Hadullo Ken[4], "A Drought Forecasting Model Using the Prophet Time Series Analysis Technique", Indian Journal of Computer Science and Technology, Volume 04, Issue 02 (May-August 2025), PP: 290-294.*

Abstract: *Drought remains one of the most devastating environmental hazards, significantly affecting agricultural productivity, water availability, and socio-economic stability, particularly in semi-arid regions like Machakos County, Kenya. This study proposes a hybrid drought forecasting model that integrates Facebook's Prophet Time series algorithm with Long Short-Term Memory (LSTM) neural networks to enhance the accuracy and reliability of drought prediction. Using historical climate data — including rainfall, temperature, humidity, and soil moisture — sourced from the Visual Crossing Weather Data platform, the model captures complex temporal patterns and non-linear dependencies. The research demonstrates the limitations of traditional models such as ARIMA and SARIMA and highlights the advantages of combining Prophet's seasonality modeling with the temporal depth of LSTM networks. Evaluation metrics such as RMSE, MAE, and R² are used to validate the model's performance. This approach contributes to early warning systems and decision-making processes for drought management in Kenya and similar semi-arid regions.*

Key Words: *Drought forecasting, Prophet Model, LSTM, time series analysis, machine learning, Machakos County, climate change adaptation*

I.INTRODUCTION

Drought is a common natural disaster that seriously upsets both human livelihoods and ecological balance, especially in areas where communities mainly depend on rain-fed agriculture. Economic stagnation, crop failures, and water shortages have resulted from frequent drought episodes in Machakos County, Kenya (Mdemu, 2021; Syomiti, n. "d.".). As climate change increases the frequency and severity of extreme weather events, these problems are made worse (Alamgir et al. 2020; Elneel & Co. (2024).

Traditional drought forecasting models, such as ARIMA and SARIMA, have long been employed to predict drought trends. However, they often fall short in capturing the non-linear relationships and complex seasonality inherent in climatic data (Rezaiy & Shabri, 2023). This limitation underscores the need for more advanced, data-driven models capable of integrating various climatic indicators.

This study aims to develop and validate a drought forecasting model using the Prophet time series technique, enhanced by LSTM neural networks, to improve the prediction accuracy of drought onset, severity, and duration in Machakos County. The study will provide critical insights for policy-making, agricultural planning, and sustainable water resource management.

Problem Statement

The cumulative effects of climate change and environmental degradation have made drought events in Kenya more frequent, severe, and unpredictable, especially in arid and semi-arid areas like Machakos County. These droughts have a major negative impact on rural livelihoods, food insecurity, water scarcity, and agricultural output. The current early warning systems are not very good at making timely and accurate predictions, even with their devastating effects. Climate variables like rainfall, temperature, and soil moisture exhibit complex, non-linear behavior that is incompatible with the linearity and stationarity assumptions of the majority of conventional models, including ARIMA and SARIMA.

Furthermore, these models often fail to incorporate macro-climatic indices like the El Niño-Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD), which are known to influence drought dynamics in Eastern Africa. This results in delayed responses by farmers, policymakers, and disaster preparedness agencies, exacerbating the socio-economic consequences of droughts.

There is thus a critical need for an adaptive, data-driven, and context-specific forecasting model that can accurately predict drought onset, duration, and intensity by integrating diverse climate variables. Addressing this gap is essential for enhancing local resilience, improving resource planning, and supporting proactive interventions in drought-prone regions like Machakos County.

Objective of the study

To develop a Prophet-based drought prediction model that enhances planning and resource management in Machakos County.

II.LITERATURE REVIEW

Traditional Statistical Models for Drought Forecasting

Statistical models like Autoregressive Integrated Moving Average (ARIMA) and its seasonal variant SARIMA were a major part of early drought forecasting efforts. These models require stationarity and assume linear relationships within time series data, which are frequently broken in climate datasets because of erratic rainfall patterns and abrupt drought events (Rezaiy and Shabri, 2023). Although SARIMA is more adept at managing seasonality, it is less adaptable when it comes to handling intricate multifaceted climate influences, like the interplay of temperature, precipitation, and soil moisture (Karimi et al. (2019).

The SARIMAX model, an extension incorporating exogenous variables, was introduced to address some of these limitations. However, its performance still declines when faced with missing data, outliers, or non-linear responses, which are common in drought-prone regions (Manigandan et al., 2021; Nandgude et al., 2023).

Emergence of Machine Learning Models

The limitations of linear statistical methods led to increased interest in Machine Learning (ML) models, which can capture complex, non-linear relationships between environmental variables. Algorithms such as Random Forests (RF), Support Vector Machines (SVM), and Artificial Neural Networks (ANNs) have shown promise in this regard (Sundararajan et al., 2021).

For instance, RF models are well-suited for drought prediction because they can handle large feature sets without overfitting and offer variable importance rankings (Probst et al., 2019; Dikshit et al., 2020). SVMs, on the other hand, perform well in high-dimensional spaces and have been successfully used to predict soil moisture and classify drought severity (Kolachian & Saghafian, 2021).

Despite their strengths, many ML models lack interpretability and often require extensive preprocessing, large training datasets, and domain-specific feature engineering, which can limit their adoption in data-scarce regions (Mustafa & Abdulazeez, 2021).

The Prophet Time Series Model

Developed by Facebook, the Prophet model was designed to provide accurate forecasts in business and scientific domains where time series data contain outliers, missing points, and strong seasonality (Taylor & Letham, 2021). Unlike ARIMA, Prophet is an additive model that decomposes time series into trend, seasonality, and holiday effects, and automatically handles change points.

Its ability to incorporate external regressors (such as soil moisture or ENSO indices) makes it well-suited for drought forecasting (Basak et al., 2022). Moreover, Prophet models have shown high robustness in handling noisy and incomplete climate data, which is critical for developing countries where monitoring infrastructure may be limited (Menculini et al., 2021).

Applications in climate studies include rainfall prediction, hydrological modeling, and particulate matter forecasting (J. Zhao et al., 2018; Topping et al., 2020). However, Prophet struggles to capture non-linear temporal dependencies, particularly abrupt drought onset caused by climate shifts — a limitation addressed by hybrid modeling approaches.

Deep Learning in Climate Forecasting

Deep learning approaches, particularly Long Short-Term Memory (LSTM) networks, are a subclass of recurrent neural networks (RNNs) designed to handle long-term dependencies in sequential data. LSTMs have demonstrated exceptional performance in climate modeling because of their capacity to learn from past temporal sequences and forecast into the future (Chien et al., 2021; Hu & Zheng, 2019).

Studies show LSTM-based models outperform conventional models in capturing complex interactions between rainfall, temperature, and soil moisture. For instance, Dikshit et al. (2021) used LSTM to enhance the Standardized Precipitation Evapotranspiration Index (SPEI), achieving more accurate short- and long-term drought forecasts across Australia.

While powerful, LSTMs are considered black box models, offering limited interpretability, which is a challenge when results must inform policy or risk communication.

Hybrid Models: Combining Prophet and LSTM

Hybrid models combining Prophet and LSTM have gained traction in recent years for improving forecasting performance. Prophet handles trend and seasonality effectively, while LSTM captures residual non-linear relationships in the data. This synergy addresses the limitations of each model when used independently.

Shohan et al. (2022) demonstrated that a Prophet-LSTM hybrid achieved superior results in load forecasting. Similarly, Balti et al. (2021) applied this hybrid in drought forecasting across China's Jiangsu Province, outperforming standalone ARIMA and LSTM models across all weather stations. Such models are especially valuable in semi-arid regions where both long-term trends and short-term anomalies are crucial to capture.

However, research applying Prophet-LSTM hybrids to drought forecasting in Sub-Saharan Africa remains sparse, representing a significant knowledge gap this study addresses.

Identified Gaps in Existing Literature

The literature clearly indicates that:

- Most traditional models struggle with non-linearity and irregular patterns
- ML and DL models improve accuracy but suffer from interpretability and resource dependency
- Prophet excels in seasonality modeling but lacks sensitivity to abrupt climate changes
- Hybrid approaches have strong potential but are underutilized in drought forecasting within local Kenyan contexts

III.METHODOLOGY

Research Design

This study uses a quantitative research design in accordance with positivism, which prioritizes empirical analysis and objective measurement of past climate data. In order to predict drought events in Machakos County, the design combines deep learning (LSTM) with time series modeling (Prophet). This hybrid strategy makes use of both methods' advantages: LSTM's ability to learn non-linear dependencies over time and Prophet's seasonality detection.

Data Sources

Visual Crossing Weather Data, which provides high-resolution data on precipitation, temperature, humidity, and wind speed, was the source of historical climate data from 2014 to 2024. These factors were chosen because they have a known relationship to drought conditions (Bosire et al. 2019; Ojo and associates. 2021).

Additionally, macro-climatic indices such as the El Niño-Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) were incorporated as external regressors to capture regional climatic variability (AghaKouchak et al., 2022).

Sampling and Preprocessing

A systematic sampling technique was employed to evenly distribute the data across the 10-year period. Preprocessing steps included:

KNN Imputation for handling missing values

Z-score normalization to standardize feature scales

Outlier detection using the IQR method

Feature selection based on correlation and feature importance from RF models

Model Architecture

Prophet Model:

The additive model used in Prophet can be defined as: $y(t)=g(t)+s(t)+h(t)+\varepsilon t$

Where:

- $g(t)$ is the trend
- $s(t)$ is the seasonal component
- $h(t)$ represents holiday effects
- ε_t is the error term

LSTM Model:

LSTM layers were stacked with dropout regularization to prevent overfitting. The hybrid approach involved:

1. Fitting Prophet to capture seasonality and trends
2. Passing Prophet residuals to the LSTM network for learning non-linear patterns

Model Evaluation

Models were evaluated using the following metrics:

- MAE (Mean Absolute Error)
- MSE (Mean Squared Error)
- RMSE (Root Mean Squared Error)
- R^2 (Coefficient of Determination)
- Precision, Recall, and F1-Score for drought classification

Findings and discussions

Historical Climate Trends in Machakos County

The analysis of climate data from 2014 to 2024 revealed a clear decline in average seasonal rainfall, particularly during the long rains (March–May) and short rains (October–December). At the same time, mean annual temperatures showed an increasing trend, rising by an estimated 0.6°C, over the decade. These observations are consistent withearlier reports highlighting the vulnerability of Eastern Kenya to climate variability (Bosire et al., 2019; Gebrechorkos et al., 2019).

Soil moisture and humidity patterns aligned closely with rainfall variability, confirming that periods of low precipitation coincide with significant reductions in soil moisture — a key trigger for agricultural drought (Wainwright et al., 2021). The study also noted intensified inter-annual variability, with alternating years of mild and severe droughts, which are likely linked to broader climate cycles such as ENSO and IOD.

Prophet Model Forecast Performance

The standalone Prophet model successfully identified seasonal drought cycles, especially the recurrent dry periods in July–September and January–February. The model achieved an RMSE of 7.3 mm, MAE of 5.2 mm, and R^2 of 0.67, indicating moderate predictive capability. Prophet's strength lay in its ability to model long-term trends and seasonal effects, such as rainfall deficits during known drought years (e.g., 2016, 2019, 2022).

Moreover, the model proved highly resilient to data anomalies and missing values — outperforming SARIMA and ARIMA models, which struggled under such irregular conditions. However, Prophet alone had limitations in forecasting non-linear and abrupt shifts in climate patterns.

Improved Accuracy with the Hybrid Prophet-LSTM Model

To overcome Prophet's limitations, residuals from the Prophet model were passed to a Long Short-Term Memory (LSTM) neural network. This hybrid approach substantially improved forecasting accuracy:

- RMSE decreased to 4.5 mm
- MAE improved to 3.1 mm
- R^2 increased to 0.88, indicating a high proportion of variance explained
- F1-score for drought classification reached 0.84, compared to 0.72 using Prophet alone

The LSTM component captured non-linear dependencies, especially during periods with irregular rainfall influenced by external climate phenomena. The hybrid model demonstrated better sensitivity to sudden drought onset and captured subtle signals from macro-climatic indices (ENSO, IOD), which improved its early warning capability.

Temporal and Spatial Forecasting Accuracy

Temporally, the model predicted drought onset within a ±1-month margin, which is highly beneficial for early preparedness in agriculture and water resource allocation. The ability to anticipate drought duration and severity provided actionable insight for planning irrigation and emergency interventions.

Spatially, while the study focused primarily on Machakos County, preliminary application to neighboring counties (e.g., Makueni and Kitui) showed that the model retained predictive robustness. However, spatial resolution was limited due to insufficient localized soil moisture data. Future work with higher-resolution satellite data could enhance spatial granularity.

4.5 Practical Implications and Decision Support

The hybrid model has broad applications:

- For agricultural planning, the model helps farmers align planting cycles with rainfall forecasts.
- In water management, local governments can use predictions to plan reservoir releases and community water supply.
- For disaster response, stakeholders like the National Drought Management Authority (NDMA) can activate contingency plans based on forecasts.

IV.CONCLUSION AND RECOMMENDATIONS

Conclusion

This study developed and validated a hybrid drought forecasting model integrating the Prophet time series technique with Long Short-Term Memory (LSTM) neural networks, using climate data from Machakos County, Kenya. The model successfully addressed the limitations of traditional statistical approaches — such as ARIMA and SARIMA — which often fall short in modeling the complex, non-linear patterns typical of climatic systems.

By combining the strengths of both Prophet and LSTM, the model demonstrated significant improvements in forecast accuracy, particularly in capturing drought onset, duration, and severity. Prophet effectively modeled long-term trends and seasonality, while LSTM accounted for residual non-linear dependencies. The inclusion of macro-climatic indicators such as ENSO and IOD further enhanced the model's capacity to capture regional climate variability.

The validated model achieved an R^2 of 0.88 and an F1-score of 0.84, confirming its reliability for operational drought forecasting. These results suggest that hybrid models can serve as effective early warning tools, enabling local governments, farmers, and disaster risk management agencies to make data-driven decisions for resource planning and climate resilience.

In addition to advancing the academic understanding of drought prediction, this study contributes a scalable and adaptable framework that can be replicated in other semi-arid or drought-prone regions across Sub-Saharan Africa. It bridges the gap between cutting-edge AI techniques and practical, localized drought management applications.

Recommendations

Operational Deployment: The developed model should be deployed as a web-based tool or API integrated with early warning systems operated by county governments, the Kenya Meteorological Department (KMD), and the National Drought Management Authority (NDMA).

Enhance Spatial Resolution: Incorporate higher-resolution, georeferenced data (e.g., from SMAP or Sentinel satellites) to improve the spatial granularity of predictions at the ward or village level.

Community Training: Empower local users (farmers, extension officers, NGOs) through capacity-building workshops on how to interpret and act on forecast outputs.

Extend Variable Inputs: Future research should integrate socio-economic indicators (e.g., food prices, livestock health) and hydrological data (e.g., river flows) for a more comprehensive drought impact assessment.

Explore Advanced Models: Future studies could evaluate the performance of Transformer-based architectures or Generative AI models in multi-horizon climate forecasting to compare with the Prophet-LSTM framework.

References

1. Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., Bashir, M., & Khan, S. U. (2021). Predicting at-risk students at different percentages of course length for early intervention using machine learning models. *IEEE Access*, *9*, 7519–7539. https://doi.org/10.1109/ACCESS.2021.3049446

2. AghaKouchak, A., Pan, B., Mazdiyasni, O., Sadegh, M., Jiwa, S., Zhang, W., Love, C. A., Madadgar, S., Papalexiou, S. M., Davis, S. J., Hsu, K., & Sorooshian, S. (2022). Status and prospects for drought forecasting: Opportunities in artificial intelligence and hybrid physical–statistical forecasting. *Philosophical Transactions of the Royal Society A*, *380*(2238), 20210288. https://doi.org/10.1098/rsta.2021.0288
3. Al Mamun, M. A., Sarker, M. R., Sarkar, M. A. R., Roy, S. K., Nihad, S. A. I., McKenzie, A. M., Hossain, M. I., & Kabir, M. S. (2024). Identification of influential weather parameters and seasonal drought prediction in Bangladesh using machine learning algorithm. *Scientific Reports, 14*(1), 566. https://doi.org/10.1038/s41598-023-51111-2
4. Alamgir, M., Khan, N., Shahid, S., Yaseen, Z. M., Dewan, A., Hassan, Q., & Rasheed, B. (2020). Evaluating severity–area–frequency of seasonal droughts in Bangladesh under climate change scenarios. *Stochastic Environmental Research and Risk Assessment, 34*(2), 447–464. https://doi.org/10.1007/s00477-020-01768-2
5. Basak, A., Rahman, A. T. M. S., Das, J., Hosono, T., & Kisi, O. (2022). Drought forecasting using the Prophet model in a semi-arid climate region of western India. *Hydrological Sciences Journal*, *67*(9), 1397–1417. https://doi.org/10.1080/02626667.2022.2082876
6. Bosire, E., Gitau, W., Karanja, F., & Ouma, G. (2019). Analysis of climate variability, trends and expected implications on crop production in a semi-arid environment of Machakos County, Kenya. *International Journal of Scientific Research and Innovative Technology*, *6*(4), 30–39.
7. Chien, H.-Y. S., Turek, J. S., Beckage, B., & Eppstein, M. J. (2021). Modeling drought conditions using deep learning techniques. *Environmental Modelling & Software, 144*, 105132. https://doi.org/10.1016/j.envsoft.2021.105132
8. Dikshit, A., Pradhan, B., & Alamri, A. M. (2020). Short-term spatio-temporal drought forecasting using random forests model. *Applied Sciences*, *10*(13), 4516. https://doi.org/10.3390/app10134516
9. Dikshit, A., Pradhan, B., Shukla, A. K., Alamri, A. M., & Alam, S. (2021). Improved drought forecasting using LSTM deep learning model. *Journal of Hydrology, 597,* 126190. https://doi.org/10.1016/j.jhydrol.2021.126190
10. Gebrechorkos, S. H., Hülsmann, S., & Bernhofer, C. (2019). Changes in temperature and precipitation extremes in Ethiopia, Kenya, and Tanzania. *International Journal of Climatology, 39*(1), 18–30. https://doi.org/10.1002/joc.5777
11. Hu, Y., & Zheng, H. (2019). Time series drought forecasting using LSTM networks. *Advances in Meteorology, 2019*, 1–10. https://doi.org/10.1155/2019/3527341
12. Kimaro, J., Massawe, B., & Kimaro, E. (2022). Effects of deforestation and poor agricultural practices on soil moisture content in East Africa. *Environmental Sustainability, 5*(1), 38–48.
13. Mahmoud, Y. M., & Mohammed, N. B. (2024). Deep learning for long-range climate forecasting: A comparative study of LSTM and GRU. *Journal of Climate Analytics*, *12*(1), 55–67.
14. Menculini, L., Pizzi, S., Mancini, M., & Vantaggiato, S. (2021). Comparing Prophet and deep learning to ARIMA in forecasting food prices. *Agricultural Economics*, *52*(3), 331–342. https://doi.org/10.1111/agec.12616
15. Mustafa Abdullah, M., & Mohsin Abdulazeez, A. (2021). Applications of support vector machines in weather and drought forecasting: A survey. *Journal of Soft Computing and Decision Support Systems, 8*(1), 1–11.
16. Nandgude, S., Gupta, A., & Nandgude, P. (2023). Evaluation of SARIMA and SARIMAX in drought forecasting: A multivariate analysis. *Water Resources Management, 37*(2), 421–437.
17. Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(3), e1301.
18. Rezaiy, M., & Shabri, A. (2023). Using ARIMA/SARIMA for Afghanistan's drought forecasting based on SPI. *Climate Risk Management, 39*, 100510. https://doi.org/10.1016/j.crm.2023.100510
19. Shohan, M. A. I., Hossain, M. S., & Chowdhury, M. T. (2022). Prophet-LSTM hybrid model for electric load forecasting. *Smart Grid and Sustainable Energy, 13*(1), 101–112.
20. Spinoni, J., Vogt, J., Naumann, G., Barbosa, P., & Dosio, A. (2019). Will drought events become more frequent and severe in Europe? *International Journal of Climatology, 39*(4), 1445–1462. https://doi.org/10.1002/joc.5858
21. Sundararajan, S., Venkatesan, R., & Arunachalam, D. (2021). Machine learning approaches for drought forecasting: A review. *Computers and Electronics in Agriculture, 187*, 106285. https://doi.org/10.1016/j.compag.2021.106285
22. Taylor, S. J., & Letham, B. (2021). Forecasting at scale. *The American Statistician, 72*(1), 37–45. https://doi.org/10.1080/00031305.2017.1380080
23. Topping, J. C., Smith, L. M., & White, A. B. (2020). Forecasting air quality using the Prophet time series model. *Atmospheric Environment, 242*, 117823. https://doi.org/10.1016/j.atmosenv.2020.117823
24. Wainwright, C. M., Finney, D. L., & Funk, C. (2021). Understanding East African soil moisture variability using seasonal rainfall and land cover data. *Climate Dynamics, 57*, 987–1002. https://doi.org/10.1007/s00382-020-05629-5
25. Zhao, Y., Li, Z., & Zhang, R. (2022). Regional drought prediction with integrated satellite-based data using Prophet. *Remote Sensing, 14*(12), 2721. https://doi.org/10.3390/rs14122721