# IMPROVING SHORT-TERM DROUGHT PREDICTION USING LSTM WITH LAGGED SPEI FEATURES: A COMPARATIVE STUDY WITH ARIMA

Ninoslava Tihi[1]*,
[0009-0004-8009-8120]

Srđan Popov[2],
[0000-0003-1215-3111]

Sanja Stankov[1],
[0000-0002-0537-6782]

Igor Vecštejn[3]
[0009-0009-9940-4411]

[1]The Higher Education Technical School of Professional Studies in Novi Sad,
Novi Sad, Serbia

[2]Faculty of Technical Science,
Novi Sad, Serbia

[3]Technical faculty “Mihajlo Pupin”,
Zrenjanin, Serbia

Correspondence:
Ninoslava Tihi

e-mail:
tihi@vtsns.edu.rs

Abstract:

Accurate short-term drought forecasting is essential for the implementation of early warning systems and managing water resources. This study investigates the applicability of data-driven models for one-month-ahead drought forecasting by comparing a traditional ARIMA method with two configurations of Long Short-Term Memory (LSTM) neural networks. The Standardized Precipitation Evapotranspiration Index (SPEI) at a one-month scale is computed using precipitation and temperature data.

Two LSTM models were developed. The first LSTM model uses only meteorological inputs, while the second uses lagged SPEI values in order to capture temporal persistence in drought conditions. The performance of both models is evaluated using MAE, RMSE, and MAPE on an independent test dataset. The results show that ARIMA achieves the lowest MAE and RMSE, confirming that short-term SPEI dynamics have strong autoregressive characteristics. The LSTM model enhanced with lagged SPEI information outperforms the standard LSTM, showing increased stability and predictive accuracy.

The results highlight the additional role of deep learning in drought prediction and propose that integrating domain-specific lagged indicators can enhance model effectiveness, enabling the development of data-driven decision support systems.

Keywords:

Drought Prediction, Lstm, Arima, Spei, Time Series Forecasting.

## INTRODUCTION

Predicting drought is a challenging task due to the complex, non-linear and non-stationary characteristics of meteorological time series. Accurate short-term drought forecasting is especially crucial for environmental monitoring and managing water resources, because it needs to be fast and dependable [1], [2]. As long-term weather data becomes more widely available, data-driven methods have become an important addition to classic statistical methods.

Classical time series models, such as the AutoRegressive Integrated Moving Average (ARIMA) model, have been widely used for forecasting drought due to their mathematical simplicity and interpretability [3], [4]. However, these models are often limited in their ability to capture non-linear relationships that are commonly present in meteorological data.

DOI: 10.15308/Sinteza-2026-213-219

Recently, deep learning techniques, especially LSTM networks, have been very popular for time series forecasting. LSTM models are made to learn long-term dependencies in sequential data and have shown strong performance across a wide range of applications [5]. Several studies have investigated the use of deep learning in meteorological and hydrological forecasting and found that they overperform traditional methods [6], [7]. However, their performance depends on different factors such as data characteristics, forecasting time frame and model design.

For this reason, a systematic comparison between classical statistical models and deep learning approaches is necessary to better understand their advantages and disadvantages for forecasting drought in the real world [8]. The role of domain-specific characteristics, such as lagged drought indices, is still an essential factor that can influence model performance.

This paper presents a comparative analysis of an ARIMA model and two LSTM-based models for short-term drought predictions using monthly aggregated data. Drought conditions are represented using the Standardized Precipitation Evapotranspiration Index (SPEI), which is based on precipitation and temperature data. In addition to a baseline LSTM model using only meteorological data, an extended LSTM model that uses lagged SPEI values is also proposed to capture drought persistence.

The models are evaluated using standard error metrics, such as MAE, RMSE and MAPE, on an independent test dataset. The results show predictive performances, stability and practical applicability of different approaches. They also highlight the potential of using hybrid feature design in improving data-driven drought forecasting systems.

## 2. RELATED WORK

Drought prediction has been extensively researched using both classical statistical approaches and, recently, machine learning techniques. Traditional models, especially ARIMA, have been widely used in hydrology and climate-related forecasting due to their solid theoretical base and easy understanding. Early research has shown that ARIMA models can accurately capture linear temporal relationships in time series related to drought, making them a good basis for short-term forecasting [3], [9].

However, climatic processes are naturally nonlinear and affected by complex interactions between various variables. As a result, classical models often struggle to accurately show such dynamics, especially when dealing with non-stationary and highly variable climate data. In recent years, machine learning and deep learning approaches have been increasingly studied to reduce these limits.

Artificial neural networks (ANNs) were one of the first data-driven methods used for time series forecasting, showing very good results at finding nonlinear trends [8]. Long Short-Term Memory (LSTM) networks have recently attracted a lot of attention for their ability to model long-term dependencies in sequential data. Sepp Hochreiter and Jürgen Schmidhuber's work on LSTM was a big step forward for recurrent neural network architectures. It made training more reliable and made time series tasks more effective [5].

In hydrology, several studies have successfully used LSTM models to model rainfall and runoff and other related environmental predictions, typically overperforming previous methods [6], [7]. These models are especially effective with large datasets and when time correlations are very important. However, their performance isn't always better, and it may change depending on data quality, the prediction length, and feature design.

Recent studies have shown how important it is to include domain-specific information in models that are based on data. Standardized indicators, such the Standardized Precipitation Evapotranspiration Index (SPEI), are commonly used to show dry conditions when predicting droughts [10], [11]. Most studies focus on meteorological data as inputs, but fewer methods use lagged drought indices to address the long-term effects of drought events.

Therefore, there is a need for systematic studies that compare both classical statistical models with deep learning techniques, while also evaluating the impact of domain-specific characteristics on model performance. This study contributes to the field by evaluating ARIMA and LSTM models in a single experimental setting, highlighting the importance of lagged SPEI features in predicting short-term droughts.

## 3. DATA AND PREPROCESSING

The research employs meteorological data from a particular region characterized by significant seasonal fluctuations and recurrent droughts. The region has a moderate continental climate, which implies that droughts highly affect the ecology, agriculture, and water resources. These characteristics represent an excellent case study for evaluating data-driven methodologies for drought prediction.

The datasets of daily meteorological observations, collected from January 2014 until December 2020, from a weather station near Novi Sad, Serbia. Environmental sensors measure six parameters, such as air temperature (AT1), air humidity (AH1), wind speed (WS1), wind direction (WD1), precipitation (PP1), and surface soil moisture (SM1). Every hour, measurements are recorded, which provides high-resolution temporal data which can be used for time series analysis.

The preprocessing processes included handling missing values, verification of data consistency, aggregation of data, and normalizing input variables in order to ensure high quality data. These measures are essential for improving the stability and accuracy of time series forecasting models [9].

The Standardized Precipitation Evapotranspiration Index (SPEI) serves as the target variable in this study. SPEI is a well-known multiscalar drought index that combines precipitation and atmospheric evaporative demand, making it especially suitable for drought assessment under changing climate conditions[10], [11]. The index was computed from the available meteorological data and used for both model training and evaluation.

To preserve the temporal structure of the data and prevent information leaking [12], the dataset was divided into training and testing subsets using a chronological split. This method makes sure that the evaluation of model performance is fair and consistent. The resulting dataset and experimental setup provide a reliable basis for comparing deep learning with classical statistical methods for predicting short-term droughts.

## 4. METHODOLOGY

This study uses a standardized drought index based on precipitation and air temperature data to represent drought conditions. This index combines data on water availability and atmospheric evaporative demand, allowing for a more comprehensive description of drought dynamics. The resulting drought index time series is used as the target variable.

The forecasting task is defined as a one-month-ahead prediction problem. In other words, the goal is to use the data up to time step t to predict the value of the drought index at time t+1. To ensure consistency with the SPEI formulation and the forecasting setup, daily meteorological measurements are aggregated to monthly averages before model development.

To prepare the data for modeling, the time series is transformed into a fixed-length look-back window. This means that the model uses a set number of previous time steps as input to learn patterns and predict the next value in the time series. This method enables the models to learn from past observations, which helps them find patterns over time. The models use historical data from the defined window to estimate the next value of the drought index.

All models are evaluated under the same conditions to ensure a fair and consistent comparison of their performance.

### 4.1. ARIMA MODEL

The AutoRegressive Integrated Moving Average (ARIMA) model was used as a baseline statistical method for drought prediction. ARIMA is defined by three parameters: p, d, and q. These represent the autoregressive order, the degree of differencing, and the moving average order, respectively [13].

Standard information criteria [14] were used to automatically identify the model and select the parameters. The selected ARIMA configuration was then trained on the training subset of the drought index time series and used to provide one-month-ahead forecasts for the testing period.

### 4.2. LSTM MODEL

Along with the ARIMA model, a Long Short-Term Memory (LSTM) neural network is used to find nonlinear temporal correlations in the data. LSTM networks are very suitable for predicting time series as they can learn patterns from sequential inputs and model dependencies over time. In this study, the LSTM model takes sequences of historical drought index values and meteorological variables as input and produces a one-step-ahead forecast.

The architecture of the model consists of a single LSTM layer followed by a fully connected output layer. Backpropagation through time is used to train the model, and the loss function is the mean squared error. All input variables are normalized to make the training more stable and generalizable before training.

Drought indices like SPEI are derived from precipitation and temperature, and their past values provide a valuable representation of historical drought conditions. ARIMA models temporal dependence directly through its autoregressive structure, but LSTM works inside a supervised learning framework and depends on the provided input features. Therefore, lagged values of the drought index are included as extra inputs. This helps the mode to better capture drought length and temporal dynamics over time. This design also ensures consistent comparison between the statistical and deep learning methods.

### 4.3. MODEL TRAINING AND EVALUATION

The dataset is divided into training and testing sets based on the order of the data. Performance of all models was evaluated using standard accuracy metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE).

MAPE is often used in prediction tasks, but it is not suitable for standardized drought indices like SPEI due to the possible presence of values close to zero. Therefore, MAE and RMSE are used as the primary measures for predictive accuracy in this study.

MAE measures the average prediction error, providing a simple and comprehensive assessment of model accuracy. RMSE gives more weight to larger errors, which makes it very effective for finding significant differences between actual and anticipated values, especially when conditions are more extreme. MAPE shows the prediction error in relative terms, allowing comparisons of different datasets and time periods.

All evaluation metrics were calculated only on the test dataset to ensure an objective assessment of model generalization efficacy. Using multiple metrics gives a more comprehensive comparison between ARIMA and LSTM models.

All experiments were conducted under identical settings to provide a fair comparison, including the forecasting horizon, evaluation measures, and testing period. The effectiveness of ARIMA and LSTM models in short-term drought forecasting was analyzed using monthly aggregated meteorological time series data.

### 4.4. SOFTWARE AND IMPLEMENTATION

The R environment was used for all of the tests. The *forecast* package was used to implement the ARIMA baseline, while *keras* and *tensorflow* interfaces for R were used to create the LSTM model. The *dplyr, lubridate*, and *zoo* packages were used for data cleaning and creating new features. Model evaluation was performed using standard error metrics and the *ggplot2* package was used for visualization of the results.

## 5. RESULTS AND DISCUSSION

This section presents the results of a comparison between the ARIMA model and two LSTM configurations, one based only on meteorological data and the other that us extended with lagged SPEI values, for one-month-ahead drought forecasting. All models were evaluated using the same training and testing datasets, forecasting horizon, and evaluation metrics. First, the quantitative performance of the models is assessed using standard error metrics. Then, the analysis of qualitative performance is performed by comparing the observed and anticipated drought index values in graphs. The results are discussed in terms of each model's ability to capture temporal dependencies and variability in the drought time series across time.

### 5.1. EXPERIMENTAL SETUP

The forecasting experiments used a one-month-ahead prediction window. The dataset was partitioned into training and testing subsets by a chronological split to keep the temporal dependencies. The performance of the models was evaluated by MAE, RMSE, and MAPE to ensure comparison between the ARIMA and LSTM models in a consistent way.

### 5.2. QUANTITATIVE RESULTS

Table I presents the performance results of the ARIMA baseline and the two LSTM models for one-month-ahead SPEI prediction. The ARIMA model had the lowest MAE and RMSE, which means that short-term drought dynamics are very stable and autoregressive.

The LSTM model with only meteorological data showed weaker performance abilities, suggesting that meteorological inputs are not sufficient to capture short-term SPEI variability. The LSTM model with lagged SPEI values showed improved accuracy compared to the model when only meterological data were used. This shows how important drought persistence and system memory are.

Even though the LSTM model with SPEI lags did not perform better than ARIMA baseline, it showed greater stability in predictions and better consistency in capturing drought evolution. High MAPE values for all models are due to the near-zero SPEI values during transition periods. This shows a known problem with using percentage-based error metrics for standardized drought indices. Therefore, MAE and RMSE provide a more reliable evaluation of model performance.

The improvement made by LSTM is in accordance with its ability to model nonlinear temporal dependencies and complex interactions between predictors. ARIMA, on the other hand, mostly captures linear autocorrelation patterns in the target series. This can be a problem when climatic conditions are changing fast. Adding lagged SPEI values made the results more accurate, which showed that the effects of drought are long-lasting.

### 5.3. QUALITATIVE RESULTS

Figure 1 shows the real and predicted values of the drought index during the testing period. The LSTM model more precisely follows the observed dynamics, especially during periods of higher variability. The ARIMA model shows the general seasonal trend, although it has greater differences in months when there are sudden changes.

Figure 2 shows the absolute forecast errors over time for all models. The LSTM model performs better throughout the test period, with smaller and more evenly spread errors. On the other hand, the ARIMA model shows larger error fluctuations, especially when the observed SPEI series is more variable.

**Table 1.** Forecasting performance of all three models on the test set

| Models | MAE | RMSE | MAPE |
|---|---|---|---|
| ARIMA (baseline) | 0.797 | 0.935 | 126.842 |
| LSTM (mete only) | 0.918 | 1.019 | 138.417 |
| LSTM (mete + SPEI lags) | 0.868 | 1.011 | 142.900 |

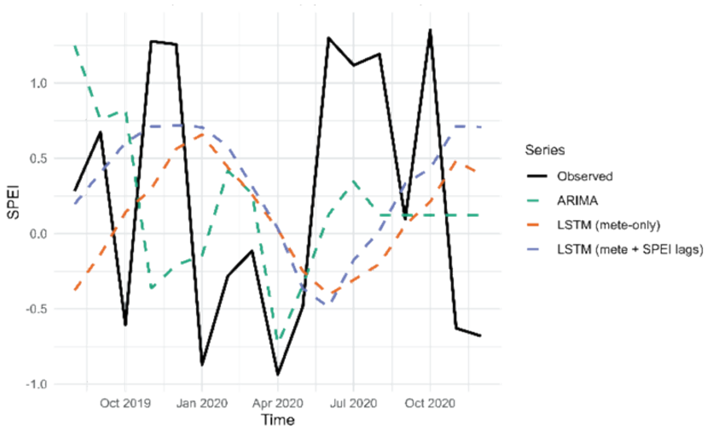

**Figure 1.** Comparison of observed and predicted SPEI (1) values for the test period

**Figure 2.** Absolute forecast errors of ARIMA and LSTM models over the test period

These results show that the LSTM model makes more stable predictions under dynamically changing drought conditions, eventhough ARIMA performs well on average.

### 5.4. DISCUSSION

The key features that make the LSTM model work better are drought phenomena dependency on different time scales and its persistence. Adding lagged values of the drought index as input features helps the LSTM model understand how drought memory works, while meteorological predictors cover the basic physical elements. This combined representation enables the model to make more consistent one-step-ahead predictions, which can be seen by slower error evolution shown in Figure 2.

By looking at the absolute forecast errors, the performance  of the models is even clearer. ARIMA has lower average MAE and RMSE values, but their errors tend to get greater when the observed SPEI series is more variable. The LSTM model, on the other hand, has more consistent error sizes during the test period, which means it is less likely to change in a short term. This means that the LSTM framework, when combined with historical drought data, may be better at adjusting to changing conditions, even though its overall accuracy is slightly lower than the ARIMA baseline.

Even with these differences, ARIMA is still a good modeling choice because it is simple, easy to understand, and cheap to run. This makes it a good choice for short-term operational drought forecasting. In general, the results show that deep learning models should be seen as an addition to, not a replacement for, traditional methods, especially when the stability of predictions and the length of drought are the most important factors.

## 6. CONCLUSION

This study investigated the effectiveness of deep learning models for short-term drought forecasting by comparing a traditional ARIMA approach with two LSTM-based configurations for one-month-ahead SPEI prediction. The results show that drought dynamics at short forecasting horizons are significantly influenced by temporal persistence, which is accurately represented by autoregressive models.

The ARIMA model has the lowest MAE and RMSE values, which shows its strong performance in short-term SPEI forecasting. This research highlights that, despite recent advances in deep learning, traditional statistical models are still quite competitive when the target variable has strong temporal autocorrelation and relatively stable seasonal patterns.

The LSTM model based only on meteorological data showed weaker performance, which suggests that meteorological data alone are not sufficient to fully reflect short-term drought variability. However, when lagged SPEI values were added as additional data, the LSTM model showed improved accuracy and more stable predictions. This finding emphasizes the importance of drought persistence and system memory, which should be explicitly integrated into data-driven models.

Although the LSTM model was improved with lagged SPEI values, it did not perform better than the ARIMA baseline; it gives a flexible nonlinear framework capable of taking different data and finding more complex relationships between meteorological conditions and past drought states. This feature may become especially valuable for longer forecasting periods or under unstable climatic conditions, where traditional linear models may be less effective.

In general, the results show that deep learning approaches should be viewed as complementary to classical time series methods and not as replacements, especially when historical characteristics specific to the domain are included. This study is limited to one-month-ahead forecasting and a single drought index. Future work will focus on extending the forecasting time period, exploring hybrid modeling approaches, and adding more variables to further improve drought prediction under changing climatic conditions.

## REFERENCES

[1] A. K. Mishra and V. P. Singh, "A review of drought concepts," *J. Hydrol.*, vol. 391, no. 1-2, pp. 202-216, Sep. 2010, doi: 10.1016/j.jhydrol.2010.07.012.

[2] A. Dai, "Drought under global warming: a review," *Wiley Interdiscip. Rev. Climate Change*, vol. 2, no. 1, pp. 45-65, Jan. 2011, doi:10.1002/wcc.81.

[3] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ, USA: Wiley, 2015, doi: 10.1111/jtsa.12194.

[4] K. W. Hipel and A. I. McLeod, Time Series Modelling of Water Resources and Environmental Systems. 1st ed. Amsterdam, The Netherlands: Elsevier, 1994.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[6] F. Kratzert, D. Klotz, C. Brenner, S. Karsten and M. Herrnegger, “Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks,” *Hydrol. Earth Syst. Sci.*, vol. 22, pp. 6005–6022, 2018, doi: 10.31223/osf.io/qv5jz.

[7] C. Shen, “A transdisciplinary review of deep learning research in hydrology,” *Water Resour. Res.*, vol. 54, no. 11, pp. 8558–8593, 2018, doi: 10.1029/2018WR022643.

[8] P. G. Zhang, E. Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: the state of the art,” *Int. J. Forecasting*, vol. 14, no. 1, pp. 35–62, March 1998, doi: 10.1016/S0169-2070(97)00044-7.

[9] D. S. Wilks, Statistical Methods in the Atmospheric Sciences, 3rd ed. Oxford, UK: Academic Press, 2011.

[10] S. M. Vicente-Serrano, S. Beguería, and J. I. López-Moreno, “A multiscalar drought index sensitive to global warming: *The SPEI,” J. Climate*, vol. 23, no. 7, pp. 1696–1718, April 2010, doi: 10.1175/2009 JCLI2909.1.

[11] S. Beguería, S. M. Vicente-Serrano, F. R. Gracia, and B. Latorre, “Standardized precipitation evapotranspiration index (SPEI) revisited,” *Int. J. Climatol.*, vol. 34, no. 10, pp. 3001–3023, Aug. 2014, doi: 10.1002/joc.3887.

[12] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021.

[13] N. Tihi and S. Popov, “Selection of the best ARIMA models for urban drought prediction”, Fresenius Environmental Bulletin (FEB), vol.32, no.06, pp.2564-2572, 2023.

[14] N. Tihi and S. Popov, “A comparison of ARIMA and Random Forest time series models for urban drought prediction”, in *Proc. Int. Sci. Conf. Information Technology, Computer Science and Data Science (Sinteza)*, Belgrade, Serbia, 2024, ISBN: 978-86-7912-821-8.