Journal of Information Systems Engineering and Management
2025, 10(20s)

e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Comparative Study of SARIMAX and Artificial Neural Networks for Drought Forecasting in North Karnataka

Deepika Prabhu[1], Dr. Suma A P[2]
[1]Research Scholar, JAIN (Deemed-to-be University), Bengaluru, India.
ORCID: ORCID: 0009-0001-2154-1944,
[2]Research Guide, JAIN (Deemed-to-be University), Bengaluru, India
ORCID: 0009-0004-7280-6350

ARTICLE INFO

Received: 21 Dec 2024

Revised: 27 Jan 2025

Accepted: 15 Feb 2025

ABSTRACT

Drought forecasting is crucial for managing its effects on crop production and water supply. This study evaluates the effectiveness of the Seasonal Auto-Regressive Integrated Moving Average model with external predictors (SARIMAX) and Artificial Neural Networks (ANN) in predicting the Standardized Precipitation Index (SPI), a crucial indicator of drought conditions in Karnataka. Using data from 2000 to 2023, both models incorporate variables such as rainfall, temperature, NDWI (Normalized Difference Water Index), and NDVI (Normalized Difference Vegetation Index). The results indicate that the SARIMAX model significantly outperforms the ANN model, with a lower RMSE of 0.2699 and MAE of 0.2114, highlighting its superior accuracy and predictive reliability. The SARIMAX model also demonstrates minimal bias (ME = 5.89e-14) and uncorrelated residuals (ACF1=-0.0750), confirming its robustness in capturing the underlying trends. In contrast, the ANN model exhibited higher errors and lower predictive performance in extreme drought conditions. Based on these findings, the SARIMAX model is recommended as the more effective tool for SPI forecasting in North Karnataka, offering a reliable approach to enhancing agricultural resilience in drought-prone areas.

Keywords: SARIMAX, ANN, drought, SPI forecasting, R software.

1. INTRODUCTION

Droughts are among the most severe and destructive environmental catastrophes, posing severe threats to agriculture, water resources, and livelihoods, especially in regions with predominantly agricultural economies. North Karnataka, a semi-arid region in southern India, frequently experiences droughts that disrupt agricultural productivity and socio-economic stability. The growing unpredictability of rainfall patterns due to climate change underscores the necessity of robust drought forecasting models to enhance resilience and mitigate adverse impacts.

Accurate forecasting of drought events is critical for informed decision-making in agricultural planning, water resource management, and disaster preparedness. Traditional drought indexes, such as the Standardized Precipitation Index (SPI), have been widely used to assess drought severity. However, the challenge lies in accurately predicting SPI values, which depend on a combination of climatic variables and their complex interrelationships. Vyom Shah et al., 2024, demonstrated the potential of SARIMAX models in capturing seasonal patterns and external climatic variables, proving their efficacy in agricultural planning. Similarly, Kumar et al., 2021 highlighted the capabilities of ANN models in handling non-linear climatic interactions, showcasing their advantage in regions with complex weather patterns.

This article provides two advanced modeling techniques: the Seasonal Auto-Regressive Integrated Moving Average with Exogenous Variables (SARIMAX) and Artificial Neural Networks (ANN). SARIMAX, a statistical model, captures seasonality and external influences, making it suitable for time-series data with predictable patterns. On the other hand, ANN, a machine learning-based model, excels in handling non-linear relationships and complex data structures, offering a data-driven approach to drought forecasting.

This article seeks to assess the effectiveness of these two models in forecasting SPI for North Karnataka. The study leverages 23 years of historical data (2000-2023) encompassing key Meteorological factors like temperature, rainfall, and vegetation indices. By comparing the models' accuracy using metrics like Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), this study aims to identify the most reliable and efficient model for drought prediction.

2. LITERATURE REVIEW

Amr Mossad and Abdulrahman AliAlazba (2015) highlight the significance of drought forecasting in facilitating effective drought mitigation strategies. Their study explores the applicability of linear stochastic models, specifically the autoregressive integrated moving average (ARIMA) model, for predicting drought conditions. Using the Standardized Precipitation Evapotranspiration Index (SPEI) in a hyper-arid climate, multiple ARIMA models were developed to assess their forecasting capabilities. The findings indicate that these models can effectively predict drought across various time scales. Consequently, the study suggests that ARIMA models are valuable tools for drought forecasting, enabling water resource managers and planners to implement proactive measures based on anticipated drought severity.

A.K. Mishra, V.R. Desai, and V.P. Singh (2007) present key findings from their study on drought forecasting using the Standardized Precipitation Index (SPI). Their analysis indicates that a hybrid model outperforms other approaches for predicting drought in the Kansabati basin, India. This hybrid model integrates the strengths of both stochastic and neural network models. A comparison among various models—including ARIMA/SARIMA (stochastic models), recursive multistep neural network models (RMSNN), direct multistep neural network models (DMSNN), and hybrid stochastic neural networks with recursive (HSNNRA) and direct (HSNNDA) approaches—reveals that HSNNRA provides the best performance for a one-month lead time. However, as the lead time increases, the accuracy of recursive-based models declines due to error accumulation. In contrast, HSNNDA demonstrates superior performance when longer lead-time forecasts are required.

2.1 Study Area: The Drought Crisis in North Interior Karnataka

North Interior Karnataka extends across latitudes 14°N to 18°N and longitudes 74°E to 77°E, encompassing districts such as Bagalkote, Ballari, Belagavi, Bidar, Dharwad, Gadaga, Haveri, Kalaburagi, Koppala, Raichur, Vijayapura, Yadgir, and Vijayanagar. The region experiences semi-arid climatic conditions, withyearly precipitation ranging from 500 mm to 850 mm, predominantly occurring during the southwest monsoon(June to September). Summers are typically hot, with temperatures often exceeding 35°C, while winters are mild.

Drought is a recurring issue in this area, driven by erratic and insufficient rainfall. These dry spells lead to reduced crop yields and economic hardship for farmers reliant on rain-fed agriculture. The unpredictability of rainfall, exacerbated by climate change, has strained water resources and increased the risk of desertification and soil degradation. This study seeks to analyze these seasonal drought patterns and use advanced forecasting models to support sustainable agricultural practices and mitigate the drought crisis.

Maltare et al. (2023) explored the rainfall pattern and groundwater level of the Banaskantha district of Gujarat and predicted a rise in the groundwater level using Artificial Intellegent such as SARIMA, multi-variable regression, ridge regression, and KNN regression.

The literature review reveals limited efforts have been made to develop a statistical model for reliably predicting drought occurrences in Karnataka. Additionally, the impact of hydrological, meteorologicalparameters, and remote sensing profiles on drought prediction has not been thoroughly explored. This study aims to address these gaps by leveraging data and advanced technological approaches to enhance drought forecasting accuracy.

3. METHODOLOGY

This study utilizes historical data (2000-2023) collected from official sources such as the Karnataka State Natural Disaster Monitoring Centre (KSNDMC), India Meteorological Department (IMD), United States Geological Survey (USGS), BHUVAN, and other official sources. The variables considered for drought forecasting include rainfall, temperature, NDVI, NDVI lag, NDWI, and SPI. These parameters are crucial climatic and vegetation health indicators, directly influencing drought conditions.

The methodology for this study is summarized in the flowchart below (Figure 1). It provides a detailed, sequential explanation of the process., starting from raw data collection to drought forecasting. The raw data, comprising Rainfall, Temperature, NDVI, NDVI lag, and NDWI, was processed to generate the SPI values. These processed data were then used for drought modeling through two approaches: Time Series Modeling (SARIMAX) and Artificial Intelligence Modeling (ANN). The model's accuracy and efficiency were assessed using statistical measures, including RMSE and MAE for error estimation, while AIC and BIC were employed to determine model fit and complexity. The final step involved drought forecasting in terms of duration and severity. The analysis is conducted using R software to ensure precise modeling and forecasting.

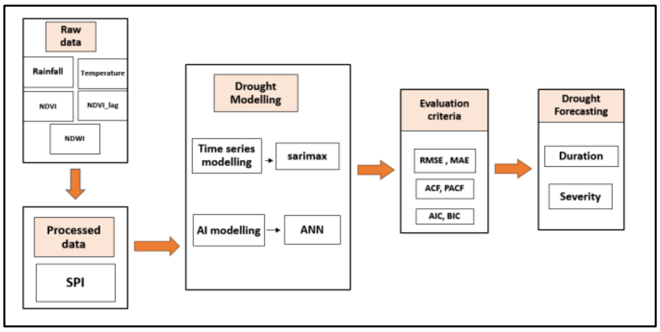

Figure 1: Flowchart depicting the methodology for drought forecasting

3.1 Normalized Difference Vegetation Index (NDVI)

NDVI is a remote sensing index utilized to evaluate vegetation density and health. It is derived from the difference in reflectance between near-infrared (NIR) and red (RED) light, as healthy vegetation strongly reflects NIR while absorbing more RED light.

$$\textbf{NDVI}= \frac{(NIR-RED)}{(NIR+RED)}$$

Where:

NIR: Reflection in the near-infrared wavelength range

RED: Reflectance in the red wavelength range

The NDVI scale spans from -1 to +1, where elevated values reflect healthier and more abundant vegetation. Values near zero suggest barren or sparsely vegetated areas, while negative values often correspond to water bodies or clouds.

3.2 NDVI Lag

NDVI lag represents the delayed influence of vegetation health on drought conditions. Vegetation changes, as captured by NDVI, may impact subsequent month's or season's climatic and soil conditions. Incorporating lagged NDVI values helps capture this delayed effect, enhancing the predictive capacity of drought forecasting models.

$$\text{NDVI lag}(t) = \text{NDVI}(t-k)$$

Where:

- NDVI lag(t): Lagged NDVI value at time t
- t: Current time period
- k: Lag interval

3.3 Normalized Difference Water Index (NDWI)

NDWI is a satellite-derived index that assesses water content in vegetation and soil by analyzing the difference in reflectance between the green and near-infrared bands, it highlights water features and minimizes interference from vegetation and soil. NDWI values typically range between -1 and 1. Positive values indicate water bodies or areas with high water content. Negative or low values indicate dry or non-water regions. This index is widely used in drought monitoring, agricultural studies, and hydrological assessments, as it helps identify areas of water stress or availability.

$$NDWI = \frac{(GREEN-NIR)}{(GREEN+NIR)}$$

Where:

- Green Band: Reflectance value from the green wavelength of the electromagnetic spectrum.
- NIR (Near-Infrared Band): Reflectance value from the near-infrared wavelength of the spectrum.

3.4 SPI

The Standardized Precipitation Index (SPI) serves as an essential tool for assessing and quantifying drought severity. It measures deviations from normal precipitation over a specific period, ranging from -2 to +2. SPI values closer to -2 indicates extreme drought, with -1 indicating moderate drought conditions. Conversely, values above +1 represent wetter-than-normal conditions, with +2 indicating unusually high precipitation. This index aids in determining the intensity and duration of droughts, with lower values indicating more severe water scarcity.

3.5 SARIMAX (Seasonal Auto-Regressive Integrated Moving Average with exogenous variables)

SARIMAX is a time series forecasting model based on statistical methods. It builds upon the SARIMA model by incorporating exogenous variables - external factors influencing the dependent variable. SARIMAX effectively accounts for:

- Seasonality: Repeating patterns or cycles in the data.
- Trend: Long-term directional movement.
- Autocorrelation: Relationships between current and past values.
- External predictors: Key factors, including climatic conditions and economic metrics.

SARIMAX Model

The Seasonal Auto-Regressive Integrated Moving Average with Exogenous Variables (SARIMAX) model was employed to forecast the SPI. The model incorporated SPI as the dependent variable and used NDVI, NDVI lag, NDWI, rainfall, and temperature as exogenous variables.

The SARIMAX model is represented as:

SARIMAX (p, d, q) × (P, D, Q) s

where p, d, and q represent the non-seasonal components, corresponding to the autoregressive (AR) terms, differencing, and moving average (MA) terms, respectively. Similarly, P, D, and Q denote the seasonal counterparts of these components. The parameter s specifies the length of the seasonal cycle.

The SARIMAX equation incorporates exogenous predictors (X_t) as:

$$\boldsymbol{\phi_p(B)\, \Phi_P(B^s)\, (1-B)^d (1-B^s)^D Y_t = \theta_q(B)\, \Theta_Q(B^s)\, \varepsilon_t + \beta X_t}$$

In the SARIMAX model, Y_t represents the dependent variable, the time series being analyzed. The term $\phi_p(B)$ denotes the non-seasonal autoregressive (AR) component of order p, while $\Phi_P(B^s)$ represents the seasonal autoregressive (SAR) component of order P with a defined seasonality s. To achieve stationarity, the model applies differencing, where $(1-B)^d$ accounts for non-seasonal differencing of order d, and $(1-B^s)^D$ manages seasonal differencing of order D to remove seasonal trends. The moving average (MA) components are captured through $\theta_q(B)$ which represents the non-seasonal MA component of order q, and $\Theta_Q(B^s)$, which accounts for the seasonal MA component of order Q with seasonality s. Additionally, the model includes a linear combination of exogenous predictors, denoted as βX_t, where X_t represents external influencing factors, and β is their corresponding coefficients. Finally, ε_t the error term represents white noise, which is assumed to have a mean of zero and constant variance, ensuring that the model captures only the systematic patterns in the data.

3.6 Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) are machine learning models designed to emulate the functioning of the human brain in processing information and identifying patterns. They are composed of multiple layers of interconnected neurons that analyze input data and identify patterns for making predictions. ANNs are widely used for nonlinear and complex data analysis, including time-series forecasting.

Artificial Neural Network (ANN) Model

An Artificial Neural Network (ANN) model was developed to forecast SPI, leveraging its capacity to capture complex nonlinear relationships. The input layer consisted of NDVI, NDVI lag, NDWI, rainfall, and temperature as input features. The network architecture was configured with:

- Input layer: 5 neurons (one for each feature)
- Hidden layers: Two layers with 5 and 3 neurons respectively, activated using the ReLU function
- Output layer: 1 neuron for SPI prediction

The input layer, consisting of nodes representing these features, served as the entry point for the data. These input nodes transmit the data to hidden layers, processing the information through weighted connections. The hidden layers applied activation functions to transform the input data into a format suitable for making predictions. The weights assigned to each feature highlighted their relative influence on the model's output, offering insights into the importance of each variable in predicting SPI. The final SPI predictions were generated through a single-node output layer.

4. RESULTS AND FINDINGS

4.1 SARIMAX Model Performance and Error Analysis for SPI Forecasting

Table 1: Error and Accuracy Metrics for Training Set Predictions

	ME	RMSE	MAE	ACF1
Training set	5.88e-14	0.2698	0.2113	-0.0750

To build and validate the forecasting model, the data is divided into separate training and testing subsetsin an 80:20 proportion. The training set is used to build and fine-tune the SARIMA model, ensuring optimal parameter selection and model accuracy. Meanwhile, the testing set is set aside. To assess the model's forecasting accuracy and determine its ability to generalize to new data.

The training set error measures demonstrate that the model performs well in forecasting SPI. The Mean Error (ME) is nearly zero, indicating that the model is unbiased and correctly aligned with observed values. The Root Mean Squared Error (RMSE) of 0.2698 shows a relatively small average error in predictions, suggesting the model's accuracy. Additionally, the Autocorrelation of Residuals (ACF1) value of 0.0138 indicates minimal correlation among residuals, signifying that the model effectively captures the data’s underlying patterns. Overall, the model demonstrates strong forecasting ability and accurately predicts SPI values.

4.2 Comparison between the SARIMAX and ANN models

Table 2: Error Metrics for SARIMAX and ANN Models in SPI Forecasting

Model	RMSE	MAE	AIC	BIC
SARIMAX	0.2698	0.2113	26.4381	40.7395
ANN	0.6035	0.4860	-29.7657	43.7842

The evaluation of SARIMAX and ANN models for SPI forecasting reveals that SARIMAX outperforms ANN in both accuracy and model fit. With a lower RMSE of 0.2698 andMAE of 0.2113, the SARIMAX model demonstrates greater predictive precision compared to the ANN model, which recorded an RMSE of 0.6035 and MAE of 0.4860.

The comparison of model performance revealed that the AIC value for the ANN model (-29.7657) was lower than that of the SARIMAX model (26.4381). However, the BIC value for the SARIMAX model (40.7395) was slightly lower than that of the ANN model (43.7842). Considering both criteria and the overall performance metrics, the results favor the SARIMAX model for its superior predictive accuracy and reliability.

These findings suggest that the SARIMAX model is more reliable for SPI forecasting, making it a preferable choice for this study's drought prediction objectives.

4.3 ACF and PACF plots of the residuals for Model Comparison

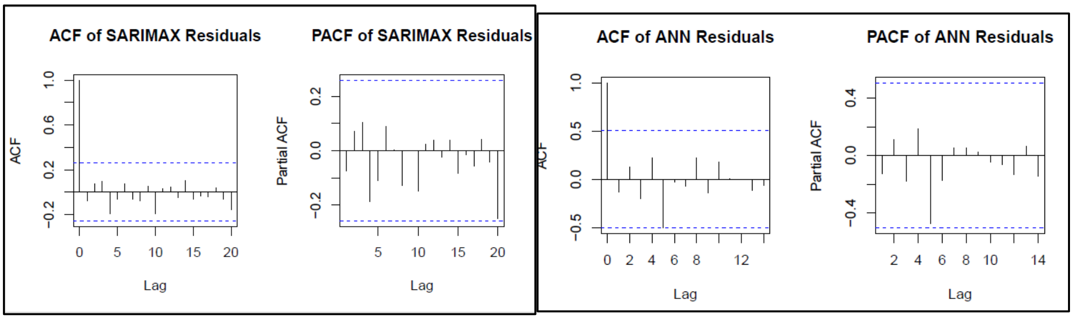

Figure 2: ACF and PACF plots of the residuals of SARIMAX and ANN models

The Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) plots of the residuals were analyzed to assess the adequacy of the SARIMAX and ANN models.

SARIMAX Residuals:

The ACF and PACF plots of SARIMAX residuals exhibit no significant lagsexceeding the confidence intervals, indicating the absence of correlation. This confirms that the SARIMAX model effectively captured the data's underlying pattern, leaving minimal autocorrelation in the residuals.

ANN Residuals:

The ACF and PACF plots of ANN residuals display some lags exceeding the confidence intervals, suggesting the remaining autocorrelation. This indicates that the ANN model may not have fully captured the temporal dependencies in the data.

4.4 Q-Q plot for SARIMAX Residuals and ANN Residuals for normality check

Shapiro-Wilk normality test

- **SARIMAX Residuals**

  W= 0.9795, p-value=0.2978

- **ANN Residuals**

W=0.98711, p-value=0.8032

The Shapiro-Wilk test results indicate that the residuals of both the SARIMAX and ANN models follow a normal distribution, as their p-values exceed 0.05. This is a good sign that the assumptions for normal residuals are met in both models.

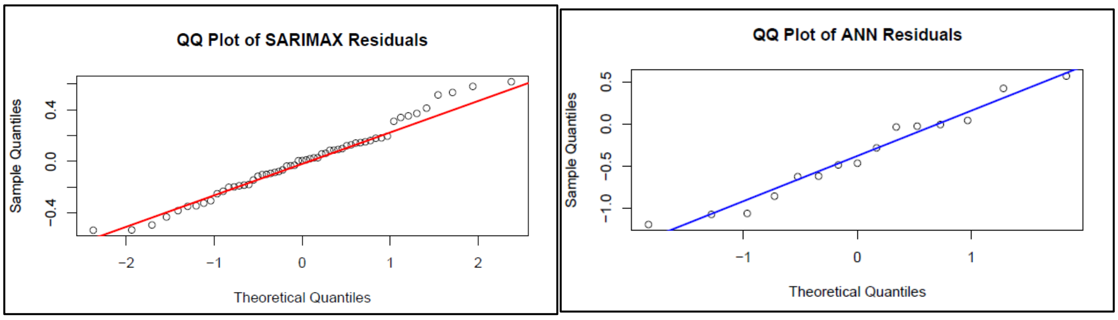

Figure 3: Q-Q Plot for SARIMAX Residuals and ANN Residuals

The Quantile-Quantile (Q-Q) plot for the SARIMAX residuals shows that they align well with the expected quantiles of a normal distribution. The Quantile-Quantile (Q-Q) plot for the SARIMAX residuals shows that they align well with the expected quantiles of a normal distribution. Most observations align with the reference trend line, suggesting that the residuals follow a normal distribution.This result supports the adequacy of the SARIMAX model in capturing the basic data patterns. It confirms that its residuals meet the assumption of normality, which is essential for reliable forecasting. This result supports the adequacy of the SARIMAX model in capturing the basic data patterns. It confirms that its residuals meet the assumption of normality, which is essential for reliable forecasting.

The QQ plot of the residuals from the Artificial Neural Network (ANN) model assesses the normality of the prediction errors. Most residuals align well with the diagonal line, indicating that they follow normal distribution, which is a desirable property for validating model assumptions.

4.5 Artificial Neural Network Architecture for SPI Forecasting

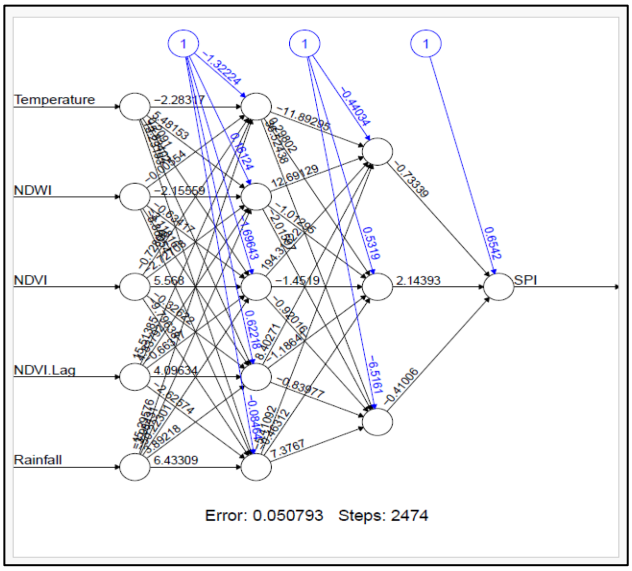

Figure 4: Artificial Neural Network Architecture for SPI Forecasting

The diagram illustrates the ANN model's architecture for forecasting SPI. It consists of five input nodes representing Temperature, NDWI, NDVI, NDVI Lag, and Rainfall, which feed into multiple hidden layers of interconnected neurons. The connections between nodes are weighted, indicating the relative importance of each input and hidden layer neuron. The output node represents the predicted SPI value. The model's prediction error is 0.050793, indicating a high level of accuracy. The document also lists the weights associated with each feature, highlighting their influence on the model's predictions. This ANN model can be utilized for accurate environmental monitoring and forecasting.

4.6 Actual and Predicted values of SPI using SARIMAX model (2000-2023)

The SARIMAX model effectively predicted the SPI values, closely aligning with the observed (actual) values (refer to Table 3). The residuals, representing the discrepancies between the observed and predicted SPI values, were mostly minimal, suggesting that the model fits the data well.

Table 3: Actual vs. predicted and Residual values of SPI using the SARIMAX model (2000-2023)

Year	Season	Actual SPI	Predicted SPI	Residual
2000	Kharif	-0.4231	-0.4795	0.0564
2000	Rabi	-0.3484	-0.3715	0.0231
2000	Zaid	-0.2793	-0.3895	0.1103
2001	Kharif	-0.1533	-0.1312	-0.0221
2001	Rabi	0.0979	-0.0963	0.1942
2001	Zaid	-0.7195	-0.7318	0.0122
2002	Kharif	-0.5122	-0.4958	-0.0164
2002	Rabi	0.4764	0.4907	-0.0143
2002	Zaid	0.2624	0.3749	-0.1125
2003	Kharif	-0.7598	-0.8449	0.0851
2003	Rabi	0.1454	0.0782	0.0672
2003	Zaid	-0.3009	-0.2433	-0.0576
2004	Kharif	-0.3907	-0.3125	-0.0782
2004	Rabi	-0.0130	0.0499	-0.0629
2004	Zaid	-0.1515	0.0458	-0.1974
2005	Kharif	-0.7202	-0.7013	-0.0189
2005	Rabi	0.0954	-0.0446	0.1401
2005	Zaid	-0.1393	-0.0401	-0.0991
2006	Kharif	-0.9973	-0.9833	-0.0140
2006	Rabi	0.2467	0.2889	-0.0422
2006	Zaid	0.5180	0.5191	-0.0011
2007	Kharif	0.1248	0.1041	0.0207
2007	Rabi	-0.2611	-0.4965	0.2354
2007	Zaid	0.2020	0.1201	0.0819
2008	Kharif	-0.6218	-0.6839	0.0621
2008	Rabi	0.2020	0.1201	0.0819
2008	Zaid	-0.1282	-0.0688	-0.0594
2009	Kharif	-0.8698	-0.8199	-0.0499
2009	Rabi	-0.5494	-0.1856	0.1356
2009	Zaid	0.3272	0.2402	0.0870
2010	Kharif	-0.0621	-0.0771	0.0150
2010	Rabi	-0.5352	-0.6216	0.0864
2010	Zaid	-0.5594	-0.5493	-0.0101
2011	Kharif	-0.9926	-0.8972	-0.0954

2011	Rabi	-0.2214	0.1845	-0.4059
2011	Zaid	0.2295	0.1862	0.0433
2012	Kharif	-0.4449	-0.3560	-0.0889
2012	Rabi	0.4764	0.4907	-0.0143
2012	Zaid	0.7243	0.7013	0.0230
2013	Kharif	0.2446	0.2924	-0.0478
2013	Rabi	0.1454	0.0782	0.0672
2013	Zaid	-0.0439	0.0281	-0.0720
2014	Kharif	-0.7622	-0.7849	0.0227
2014	Rabi	-0.5398	-0.4787	-0.0610
2014	Zaid	0.4873	0.5308	-0.0435
2015	Kharif	-0.6147	-0.4992	-0.1155
2015	Rabi	0.4641	0.5069	-0.0428
2015	Zaid	-0.0530	-0.0978	0.0447
2016	Kharif	0.2226	0.1992	0.0234
2016	Rabi	-0.6921	-0.7792	0.0871
2016	Zaid	-0.2551	-0.1340	-0.1211
2017	Kharif	0.0356	0.1086	-0.0730
2017	Rabi	-0.7265	-0.6866	-0.0398
2017	Zaid	-0.0396	0.0042	-0.0439
2018	Kharif	0.0181	0.0213	-0.0033
2018	Rabi	-0.3529	-0.3715	0.0186
2018	Zaid	0.1267	0.2123	-0.0856
2019	Kharif	-0.5012	-0.4747	-0.0266
2019	Rabi	-0.4925	-0.4798	-0.0127
2019	Zaid	0.4286	0.4105	0.0181
2020	Kharif	-0.2398	-0.2688	0.0291
2020	Rabi	0.1178	0.0635	0.0543
2020	Zaid	-0.2009	-0.1903	-0.0106
2021	Kharif	-0.2658	-0.2381	-0.0277
2021	Rabi	-0.0287	-0.0624	0.0337
2021	Zaid	0.0892	0.1277	-0.0385
2022	Kharif	-0.6545	-0.6762	0.0217
2022	Rabi	-0.0572	-0.0604	0.0032
2022	Zaid	0.0912	0.1002	-0.0090
2023	Kharif	0.4362	0.3991	0.0372
2023	Rabi	0.1975	0.2619	-0.0644
2023	Zaid	-0.3691	-0.3337	-0.0354

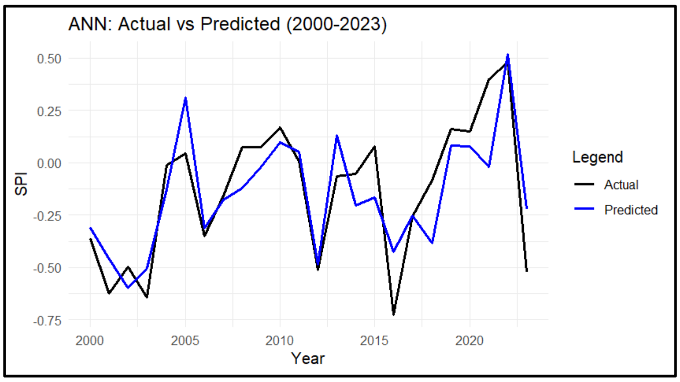

Figure 5: Actual and Predicted SPI values using ANN Model

4.7 Forecasted SPI values using SARIMAX

Since SARIMAX outperforms the ANN model, the forecasted SPI values for the period from 2024 to 2030 are presented season-wise. Refer to Table 5 and the plot (Figure 6) for a detailed view of Actual, Predicted, and forecasted SPI values.

Table 5: Forecasted SPI Values Using SARIMAX (2024-2030)

Year	Season	Forecasted SPI
2024	Kharif	-0.1188
2024	Rabi	0.1093
2024	Zaid	-0.2576
2025	Kharif	-0.4668
2025	Rabi	0.0020
2025	Zaid	-0.2804
2026	Kharif	-0.7106
2026	Rabi	0.0726
2026	Zaid	-0.3094
2027	Kharif	-0.2969
2027	Rabi	-0.1077
2027	Zaid	-0.376
2028	Kharif	-0.0933
2028	Rabi	-0.1634
2028	Zaid	0.60383
2029	Kharif	1.1010
2029	Rabi	0.1923
2029	Zaid	0.3136
2030	Kharif	-0.2284
2030	Rabi	0.1458
2030	Zaid	-0.0593

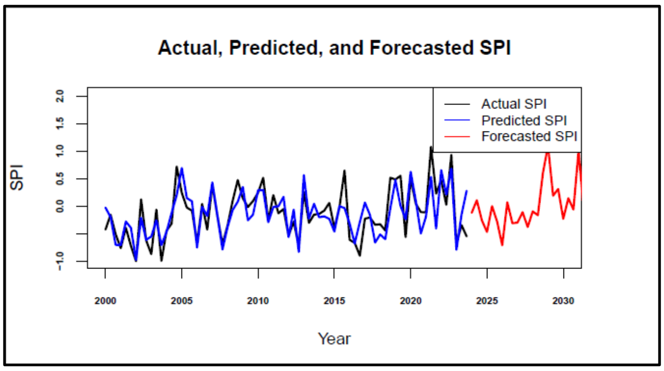

Figure 6: Actual, Predicted, and Forecasted SPI values using SARIMAX Model

5. CONCLUSION

This study compares the forecasting performance of two models SARIMAXand ANN for predicting SPI (Standardized Precipitation Index) in North Karnataka. The findings indicate that the SARIMAX model demonstrates greater predictive accuracy compared to the ANN model, as reflected in its lower RMSE (0.2698) and MAE (0.2114). demonstrating its ability to predict drought conditions reliably. The SARIMAX model also exhibited minimal bias (ME = 5.89e-14) and uncorrelated residuals (ACF1 = -0.0750), making it a robust tool for drought forecasting.
However, the ANN model still showed considerable promise, particularly in capturing complex, non-linear relationships between the predictors and SPI. While the ANN had slightly higher RMSE and MAE values, it was still effective in modeling the complex patterns in the data. It can be a valuable tool in scenarios where non-linear interactions are critical. The flexibility of ANN models makes them suitable for a wide range of forecasting tasks, especially when data patterns are highly complex.
In conclusion, while the SARIMAX model is recommended for its superior forecasting accuracy and reliability, the ANN model should not be dismissed. It can complement SARIMAX in areas where capturing non-linear relationships is essential. Both models provide valuable insights into SPI forecasting, and the selection between the two should be based on the specific requirements of the forecastingobjectives and the characteristics of the data.

IMPLICATIONS and FUTURE RESEARCH SCOPE

The results of this study have significant implications for agricultural planning, water resource management, and drought mitigation strategies in North Karnataka. A more reliable drought forecasting model can help policymakers and farmers make informed decisions regarding crop planning, irrigation scheduling, and disaster preparedness. By incorporating multiple meteorological and environmental variables, SARIMAX-based forecasting can enhance resilience against drought-induced agricultural losses.

Future studies could extend this research by evaluating the performance of hybrid models, such as SARIMAX-ANN ensembles, for improved drought prediction. Incorporating high-resolution remote sensing data and machine learning techniques could further enhance model accuracy. Additionally,taluk-wise drought forecasts could provide more localized insights, enabling better resource allocation and targeted drought management strategies. Expanding the study to include longer forecast horizons and climate change scenarios could also provide valuable insights for long-term drought risk assessment.

REFERENCES

[1] A.K. Mishra, V. R. Desai, and V. P. Singh (November 2007), Drought Forecasting Using a Hybrid Stochastic and Neural Network Model, Journal of Hydrologic Engineering, 10.1061 /ASCE1084-0699 12:6, 626-638
[2] Abdelhamid El Ibrahimi, AbdennasserBaali (2018), Application of Several Artificial Intelligence Models for Forecasting Meteorological Drought Using the Standardized Precipitation Index in the Saïss Plain (Northern Morocco), International Journal of Intelligent Engineering and Systems, Vol.11, No.1, 267-275
[3] Abrar Islam, et al.,(March 2015), "Characterization of RAP Signal Patterns, Temporal Relationships, and Artifact Profiles Derived from Intracranial Pressure Sensors in Acute Traumatic Neural Injury", Sensors, 2025

[4] Amr Mossad and Abdulrahman Ali Alazba (Drought Forecasting Using Stochastic Models in a Hyper-Arid Climate, Atmosphere 2015, 6, 410-430.

[5] Ashok K. Mishra, Vijay P. Singh (March 2011), Drought modeling – A review, Journal of Hydrology 403,157–175.

[6] Dipanwita Dutta (April 2015), Assessment of agricultural drought in Rajasthan (India) using remote sensing derived Vegetation Condition Index (VCI) and Standardized Precipitation Index (SPI), The Egyptian Journal of Remote Sensing and Space Sciences, 53-63.

[7] K. F. Fung et.al (2020), Drought forecasting: A review of modelling approaches 2007–2017, Journal of Water and Climate Change , 11.3,771-799.

[8] Kumar, P., Sharma, R., & Verma, S. (2021). *Drought prediction using statistical and machine learning models: A case study in Karnataka, India*. Journal of Climate Research, 45(2), 134-150.

[9] KSNDMC. (2017). *Drought vulnerable assessment in Karnataka*. Karnataka State Natural Disaster Monitoring Centre.

[10] Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal ofHydrology*, *391*(1–2), 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012

[11] Maltare, N. N., Sharma, D. & Patel, S. (2023). An Exploration and Prediction of Rainfall and Groundwater Level for the District of Banaskantha, Gujrat, India. International Journal of Environmental Sciences, 9(1), 1-17. https://www.theaspd.com/resources/v9-1-1-Nilesh%20N.%20Maltare.pdf

[12] Suma A P et al. (2015), Forecasting Anthrax in Livestock in Karnataka State using Remote Sensing and Climatic Variables, International Journal of Science and Research (IJSR), Volume 6, 1891-1897.

[13] Vyom Shah et al., (2024), Forecasting Temperature Trends Using SARIMAX: A Case Study in Ahmedabad City, India, preprints.org doi: 10.20944/preprints202407.1521.v1.

Authors

First Author: Deepika Prabhu is a research scholar in the Department of Statistics at Jain (Deemed-to-be) University, Bengaluru, specializing in drought forecasting using statistical and machine learning models. Her research focuses on developing predictive models to enhance drought resilience in Karnataka, with a particular emphasis on drought-prone districts. She has published three research articles in reputed journals and actively contributes to advancements in climate modelling and agricultural sustainability.
She was honoured with the **Best Research Scholar Award** for her article *"Drought Resilience and Soil Degradation in Drought-Hit Districts of Karnataka: The Relevance of Regenerative Farming"*, presented at the 7th International Conference on Advances in Agriculture Technology and Allied Sciences (ICAATAS 2024).

Second Author: Dr. Suma A P is a distinguished academic and research guide in the Department of Statistics at Jain (Deemed-to-be) University,Bengaluru. With extensive expertise in statistical modeling and data analysis, she has mentored numerous research scholars in the fields of climate science, machine learning, and applied statistics. As a research supervisor, she has played a pivotal role in guiding studies on drought forecasting, environmental statistics, and agricultural resilience. Her contributions to academia include several research publications in reputed journals and active participation in national and international conferences.