 ,

SEASONAL DROUGHT FORECASTING IN NORTH KARNATAKA: A SARIMA APPROACH USING SPI FOR ENHANCED AGRICULTURAL RESILIENCE

Deepika Prabhu[1] and Dr. Suma A P[2]

[1]Research Scholar, Department of Statistics, JAIN (Deemed-to-be University), Bengaluru, India.

[2]Research Guide, JAIN (Deemed-to-be University), Bengaluru, India.

ABSTRACT

Drought is an extended period of significantly reduced rainfall, resulting in water scarcity that negatively impacts agriculture, water resources, and ecosystems. As one of the most devastating natural hazards, drought disrupts socio-economic stability, especially in arid and semi-arid regions like North Karnataka. Accurate drought forecasting is essential for proactive management and mitigation, helping communities prepare for and reduce the impacts of these dry spells. This study uses the Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast drought in North Karnataka's interior districts.

Using historical rainfall, soil moisture, and drought occurrence data from 2010 to 2023, the study assesses drought severity using the Standardized Precipitation Index (SPI), a robust indicator that standardizes rainfall data over specific time frames to classify drought intensity. The SARIMA model, known for its effectiveness in capturing seasonal trends in time series data, was employed to predict short-term drought occurrences. The model was validated using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), with its components—Autoregressive (AR), Integrated (I), Moving Average (MA), and Seasonal (S) factors optimized for each district based on these analyses.

Results indicate that the SARIMA model, combined with SPI-based drought assessment, reliably predicts drought occurrences and provides valuable insights into drought onset and intensity across the region. This forecasting framework aids policymakers, farmers, and water resource managers in implementing timely interventions, offering a practical tool to build drought resilience in North Karnataka's drought-prone districts. The study highlights the importance of statistical forecasting models and drought indices in enhancing climate resilience, with significant implications for regional agricultural productivity and water conservation strategies.

Key Words: Drought, SARIMA, SPI, North Interior Karnataka, drought forecasting.

INTRODUCTION

Agriculture in India is significantly impacted by two major hydro-meteorological events: drought and flood. Given its heavy reliance on the monsoon for water, Indian agriculture remains vulnerable to these challenges. Nearly one-sixth of the country's geographical area, accounting for 12% of the population, is drought-prone (N.P. Kurade et al., 2021). Additionally, over 68-70% of India's cultivated land is susceptible to drought (NIDM, 2018).

Study Area - The Drought Crisis in North Interior Karnataka

Drought has become a recurrent and severe issue in North Interior Karnataka, affecting a significant portion of the region's districts yearly. This area includes the districts of **Bagalkote, Ballari, Belagavi, Bidar, Dharwad, Gadag, Haveri, Kalaburagi, Koppala, Raichur, Vijayapura, Yadgir, and Vijayanagar**, which face extreme water scarcity due to irregular and insufficient rainfall patterns. These prolonged dry periods drastically reduce crop yields, heightening the economic vulnerability of farmers, many of whom rely on rain-fed agriculture and are often pushed into debt as a result.

The region's dependency on the monsoon, coupled with the unpredictability of rainfall exacerbated by climate change, has strained its water resources and irrigation systems. With agriculture being a primary livelihood in North Interior Karnataka, these recurring droughts pose a significant risk to food security and socio-economic stability. This study focuses on the seasonal patterns of drought in North Interior Karnataka, aiming to forecast drought conditions to aid in strategic planning and to support sustainable agricultural practices, thus helping to alleviate the drought crisis in this vulnerable region.

Methodology

In Karnataka, several districts are particularly vulnerable to the interconnected challenges of drought, soil degradation, and climate variability, making regenerative agriculture practices vital. Drought, soil degradation, temperature, and rainfall are interrelated parameters that significantly impact agriculture, particularly in regions like Karnataka.

Data Collection

Climatic data, including rainfall and temperature, were collected from the India Meteorological Department (IMD), the Karnataka State Natural Disaster Monitoring Centre (KSNDMC), and other official sources for the years 2010 to 2023 to assess the parameters. The next step involved creating a drought database. This database categorizes data by district and agricultural season. It includes detailed, district-wise information on climatic conditions and, the Standardized Precipitation Index (SPI) for each year.

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.

SARIMA

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is a powerful statistical tool widely used for time series forecasting, especially for data with seasonal patterns. SARIMA extends the ARIMA model by incorporating seasonal components to address periodic variations within the data. It uses a combination of autoregressive (AR), moving average (MA), and differencing terms to model, both seasonal and non-seasonal patterns. This makes it suitable for applications where accurate forecasting of cyclic phenomena is critical.

The SARIMA model is expressed mathematically as:

SARIMA (p, d, q) (P, D, Q) s:

- p, d, q: Non-seasonal components representing AR (Autoregressive), differencing, and MA (Moving Average) terms, respectively.
- P, D, Q: Seasonal components representing AR, differencing, and MA terms, respectively.
- s: The duration of the seasonal cycle.

The mathematical representation of SARIMA is as follows:

$$(1-\varphi_1 B)(1-\varphi_1 B^s)(1-B)(1-B^s)Y_t = (1+\theta_1 B)(1+\Theta 1 B^s)\mathcal{E}_t$$

Where,

- Y_t denotes the observed time series at timet,
- B is the backward shift operator (lag operator)(i.e., $BY_t = Y_{t-1}$)
- φ_1 represents the non-seasonal autoregressive coefficient,
- $\Phi 1$ is the seasonal autoregressive coefficient,
- $\theta 1$ stands for the non-seasonal moving average coefficient,
- $\Theta 1$is the seasonal moving average coefficient,
- s indicates the seasonal period,
- $\mathcal{E}_t$is the white noise error term at time t.

The fitting of SARIMA model

SARIMA models have been integrated with drought indices such as the Standardized Precipitation Index (SPI) to enhance prediction accuracy. For instance, Patil et al. (2021) used SARIMA with SPI data to predict drought severity in Maharashtra, India, successfully identifying critical drought periods. Shumet and Mengesha **(2019)** applied SARIMA models to forecast rainfall patterns critical for agricultural planning in Ethiopia. Their findings emphasized that accurate seasonal forecasting using SARIMA can mitigate drought's impact on crop yields by guiding sowing and irrigation schedules. SARIMA has proven to be a reliable and interpretable model for drought forecasting, especially when seasonal patterns are prominent. However, it is essential to carefully tune the parameters and account for limitations in capturing non-linear relationships to maximize their effectiveness.

Step 1: Data Segregation and Seasonal Structure

The data used for forecasting is segregated into three agricultural seasons based on their relevance to the cropping cycles:

1. Kharif Season (June–October): The primary monsoon period, essential for rain-dependent agriculture.
2. Rabi Season (November–March): The winter season, depends on stored soil moisture or irrigation.
3. Zaid Season (April–June): A short summer cropping season between Rabi and Kharif.

To capture the seasonal nature of drought conditions, the Standardized Precipitation Index (SPI) is chosen as the variable for analysis. SPI is widely used for drought monitoring as it reflects rainfall anomalies across different timescales and is sensitive to short-term and seasonal drought conditions. The data is structured into a time series format with a seasonal frequency of 3, indicating three observations (Kharif, Rabi, Zaid) per year. This structure ensures the model can account for periodic fluctuations in drought conditions specific to agricultural seasons.

Step 2: Statistical Test for Stationarity: The Augmented Dickey-Fuller (ADF) Test

The Augmented Dickey-Fuller (ADF) test was used to evaluate the stationarity of the Standardized Precipitation Index (SPI) time series data, which is an essential step for applying the SARIMA model in seasonal forecasting. The results indicated that the data was non-stationary, guiding the necessary adjustments for effective modeling.

Step 3: Differencing to Achieve Stationarity

Differencing has been performed on non-stationary data to eliminate trends and seasonal patterns.

- **First-Differencing**: First-order differencing was applied to remove trends, and the series was checked for stationarity.
- **Recheck Stationarity**: After differencing, the ADF test was repeated to ensure the time series' stationarity.

Step 4: Fitting the SARIMA Model

The Seasonal ARIMA (SARIMA) model fitted on the differenced SPI (Standardized Precipitation Index) time series has the following specifications:

ARIMA (0,0,2)(0,0,2)[3] with zero mean

- Non-seasonal part: MA (2)
- Seasonal part: Seasonal MA (2) with a periodicity of 3 (seasonally adjusted).

Step 5: Diagnosing the Model (Residual Analysis)

After fitting the model, diagnostic checks were performed on the residuals to ensure they exhibited white noise characteristics. Key evaluations included assessing normality, verifying a zero mean, and confirming the absence of autocorrelation, essential criteria for validating model adequacy.

- **Residual Plots**: Visualizing the residuals by plotting them to check if they are randomly distributed around zero; any discernible patterns may indicate model misspecification.
- **ACF and PACF Plots**: After fitting a good model, the residuals should show no significant autocorrelations (i.e., the ACF plot should show only small spikes). If there are significant correlations, this indicates that the model could be improved.
- **Normality Check (Q-Q Plot):** The Q-Q plot compares the quantiles of the residuals with those of a normal distribution. If the residuals are normally distributed, the points

will closely follow a straight line. Significant departures from the line indicate that the residuals do not follow a normal distribution.

- **Shapiro-Wilk Test**: The Shapiro-Wilk test evaluates the null hypothesis that the data follow a normal distribution. A **p-value > 0.05** indicates that the residuals are likely normally distributed, while a **p-value < 0.05** suggests they are not normally distributed.

Step 6: Forecast using SARIMA Model

The forecasted SPI values for the seasons from 2024 to 2026 indicate variations in rainfall patterns across Kharif, Rabi, and Zaid.

Results and Findings

Statistical Test for Stationarity

The stationarity of the SPI time series was assessed by testing the following hypotheses:
Null Hypothesis (H0): The time series is non-stationary.
Alternative Hypothesis (H1): The time series is stationary.
Based on the Augmented Dickey-Fuller (ADF) test results, the stationarity assessment yielded a p-value of 0.09198, which exceeds the significance threshold of 0.05. This suggests that the null hypothesis cannot be rejected, implying that the SPI time series is non-stationary in its current state. Figure 1 further illustrates the SPI index pattern, emphasizing the need for transformation or differencing to achieve stationarity, a prerequisite for accurate time series modeling.

ADF Test Results
Test Statistic: −3.2552
Lag Order: 3 (indicating the number of lagged terms included in the test).
p-value: 0.09198.

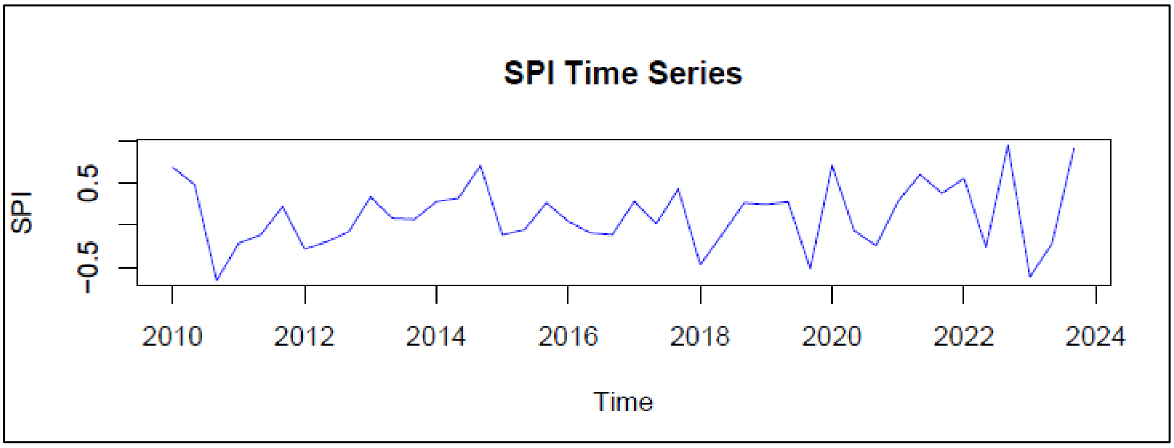

Figure 1: SPI index pattern (2010-2023)

Differencing to Achieve Stationarity

Hypothesis:
Null Hypothesis (H_0): The time series is non-stationary.

Alternative Hypothesis (H_1): The time series is stationary.
The stationarity test after first-order differencing, as shown in the section, resulted in a p-value of 0.01, which is below the 0.05 significance level. As a result, the null hypothesis is rejected, confirming that the differenced SPI time series is now stationary. With stationarity achieved, the time series is prepared for SARIMA modeling. The first-order differencing addresses long-term trends, while the seasonal components will be handled by incorporating appropriate seasonal parameters (D) in the SARIMA model. This transformation ensures that the model can effectively capture and forecast seasonal drought patterns, as illustrated in Figure 2.

ADF Test Results

Test Statistic: −6.4232

Lag Order: 3

p-value: 0.01

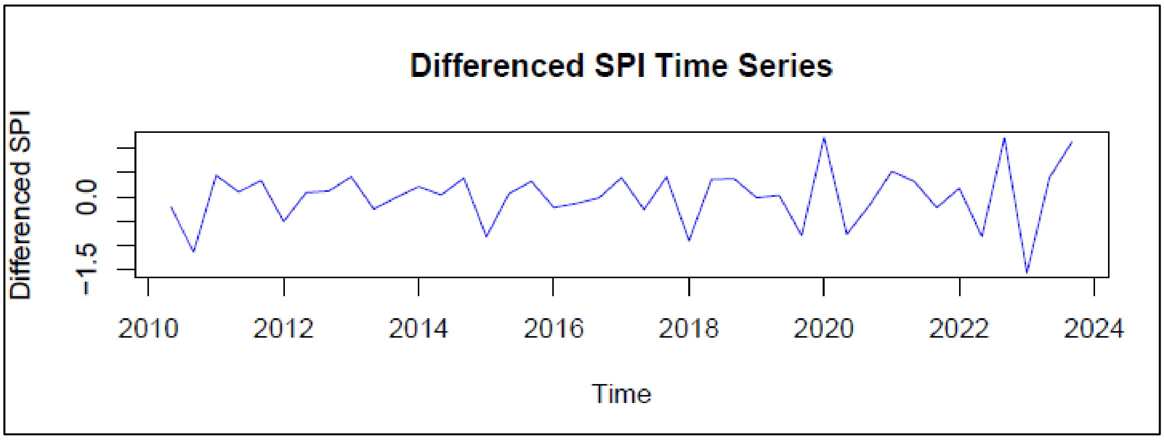

Figure 2: Differenced SPI pattern (2010-2023)

Fitting the SARIMA Model

The estimated coefficients and their standard errors are as follows:

Table 1: Estimated coefficients and S.E.

	MA1	MA 2	SMA1	SMA2
Estimated coefficients	-1.2142	0.3123	0.0087	-0.6856
S.E.	0.1663	0.1717	0.2104	0.2219

σ^2= 0.1317

The training set error measures = 0.1317

Table 2: Error and Accuracy Metrics for Training Set Predictions

	ME	RMSE	MAE	MASE	ACF1
Training set	0.02915	0.34478	0.27684	0.44603	0.069192

The SARIMA (0,0,2)(0,0,2)[3] model effectively captures the seasonal and non-seasonal dynamics in the SPI time series. Key observations include:

- **Model Fit**: The model demonstrates a strong fit with minimal bias, as indicated by a near-zero Mean Error (ME = 0.029).
- **Forecast Accuracy**: Measures like RMSE (0.34478) and MAE (0.27684) indicate good predictive performance, and suggest challenges in predicting extreme variations in SPI.
- **Residual Analysis**: The low autocorrelation in residuals (ACF1 = 0.069) confirms that the model adequately accounts for dependencies, with a residual variance (σ^2) of 0.1317 indicating reasonable error spread.

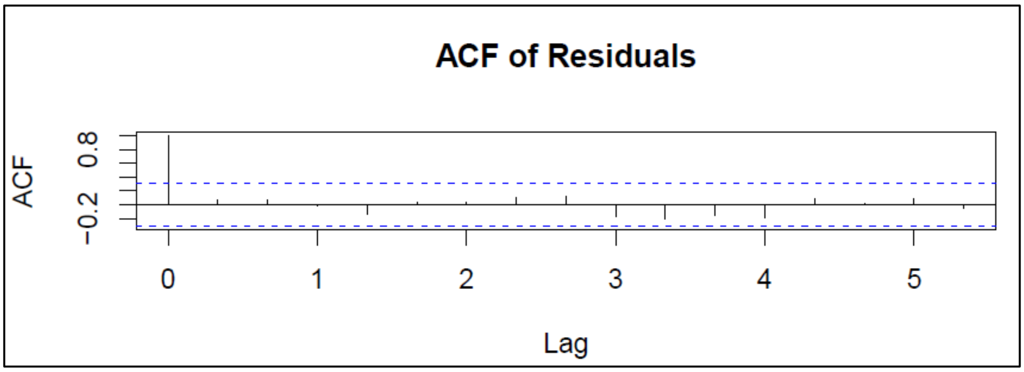

Figure 3: Autocorrelation Function (PACF) of Model Residuals

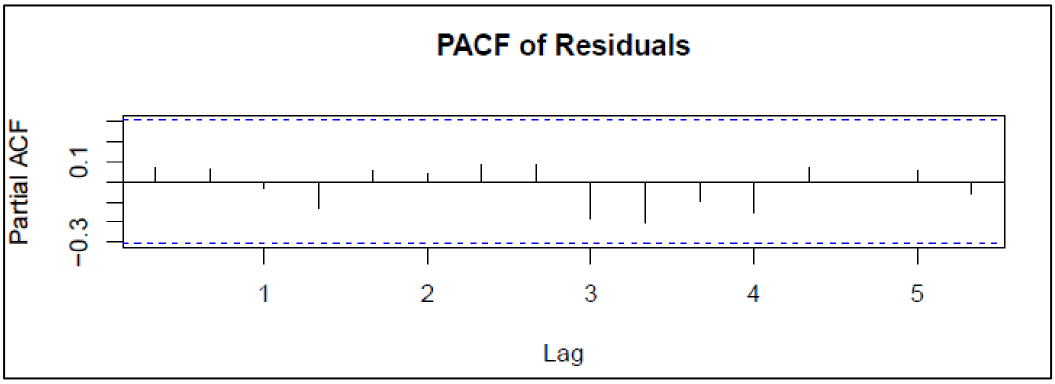

Figure 4: Partial Autocorrelation Function (PACF) of Model Residuals

Diagnosing the Model

Shapiro-Wilk normality test
W= 0.98969, p-value=0.968

The Shapiro-Wilk test assesses the null hypothesis that the data follow a normal distribution. A p-value > 0.05 suggests that the residuals are likely normally distributed.

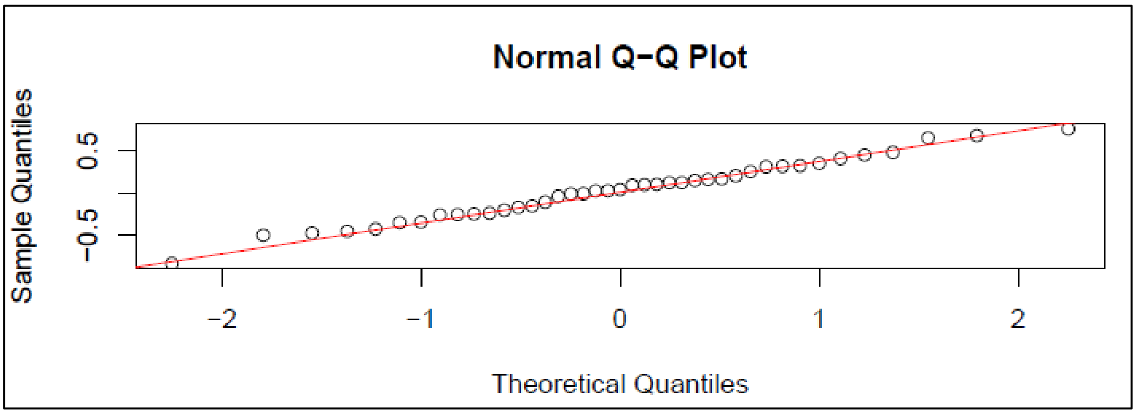

Figure 5: Normal Q -Q plot

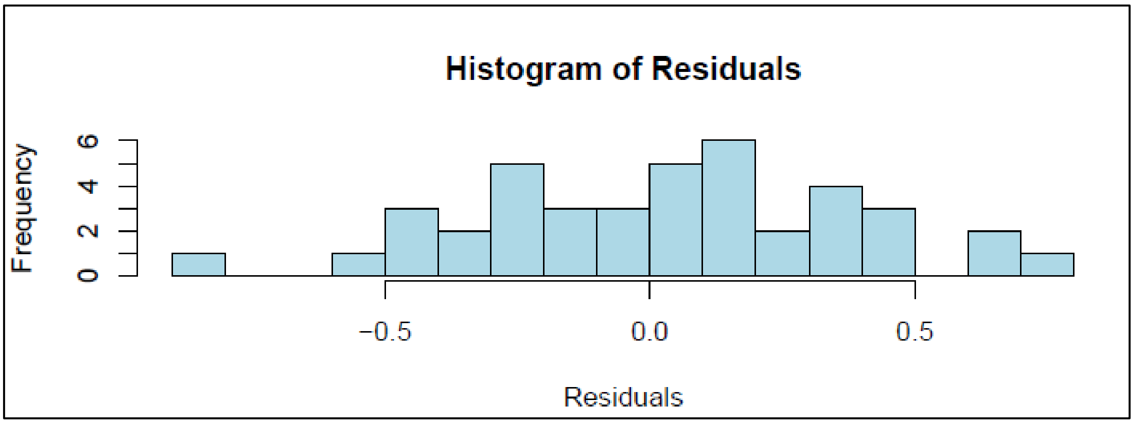

Figure 6: Histogram of Residuals

Forecasted values using SARIMA Model

The point forecasts show expected fluctuations in precipitation, while the 80% and 95% confidence intervals provide insight into the range of potential outcomes. The 80% confidence intervals suggest that there is an 80% probability that the actual SPI values will fall within the specified range, offering a narrower forecast window with moderate confidence. On the other hand, the 95% confidence intervals provide a wider range, reflecting a higher level of certainty (95%) that the actual values will lie within this broader interval. These intervals help quantify the uncertainty in the forecast, with the broader 95% range becoming more prominent as the forecast horizon extends. Overall, these projections, along with their associated confidence intervals, are valuable tools for decision-making, particularly for agricultural planning and drought preparedness.

Table 3: Forecasted Seasonal SPI Values with Prediction Intervals for Kharif, Rabi, and Zaid (2024–2026)

Year	Season	Point Forecast (SPI)	Lo 80	Hi 80	Lo 95	Hi 95

2024	Kharif	-1.359	-1.825	-0.893	-2.072	-0.647
2024	Rabi	1.071	0.339	1.803	-0.048	2.191
2024	Zaid	-0.876	-1.623	-0.130	-2.018	0.264
2025	Kharif	0.807	0.061	1.553	-0.334	1.949
2025	Rabi	-0.371	-1.117	0.375	-1.512	0.770
2025	Zaid	-0.550	-1.296	0.196	-1.691	0.591
2026	Kharif	0.662	-0.148	1.474	-0.578	1.903
2026	Rabi	-0.161	-1.060	0.737	-1.536	1.212

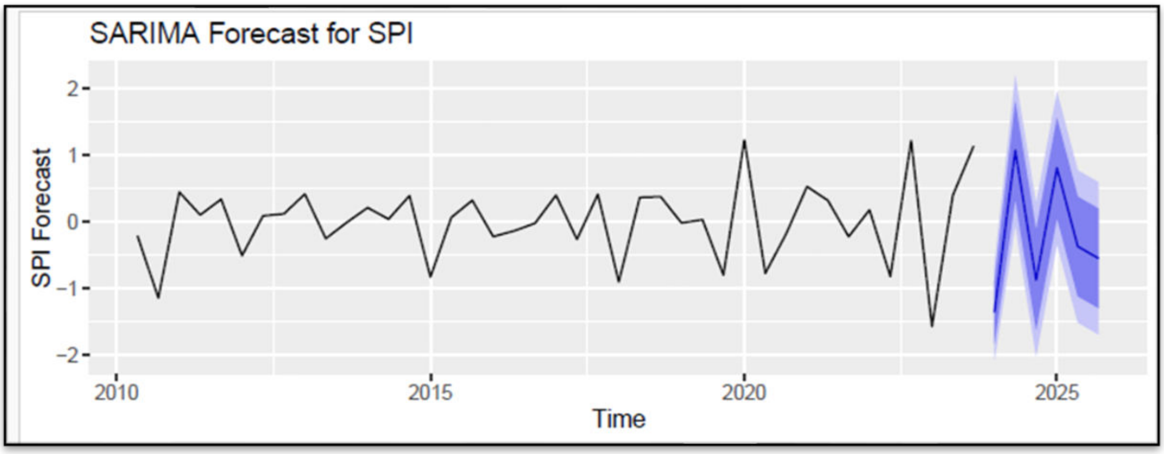

Figure 7: Forecasted SPI using SARIMA

Conclusion:

This study highlights the effectiveness of the SARIMA model in forecasting the Standardized Precipitation Index (SPI) to predict drought conditions. By addressing the non-stationarity in the time series through first-order differencing and selecting an ARIMA(0,0,2)(0,0,2)[3] configuration, The model showed a good capability to capture the underlying patterns in the data. Residual diagnostics, including the Shapiro-Wilk test and autocorrelation analysis, confirm the validity of the model's assumptions.

In conclusion, the SARIMA model provides a solid foundation for seasonal drought prediction and serves as a valuable tool for resource management and decision-making. Enhancing the model with more sophisticated techniques and broader datasets can further strengthen its applicability in drought resilience planning. This study underscores the potential of time-series forecasting in addressing critical environmental challenges like drought management.

References

- A D Diwan et al, 2021. Regenerative Agriculture, Indian Farming 71(12): 03–08.
- Amr Mossad and Abdulrahman Ali Alzaba, 2015. Drought Forecasting Using Stochastic Models in a Hyper-Arid Climate, Atmosphere 2015, 6, 410-430.
- Ashok K Mishra and Vijay P. Singh, 2011. Drought modeling – A review, Journal of Hydrology 403 (2011) 157–175.
- Dai, A. 2011, Drought under global warming: a review. Wiley Interdisciplinary Reviews: Climate Change 2, 45–65. doi:10.1002/wcc.81.

- David R. Montgomery, 2021, 11 - Soil health and the revolutionary potential of Conservation Agriculture, Woodhead Publishing Series in Food Science, Technology and Nutrition, Pages 219-229.
- Dr. Suma A P, 2015, Forecasting Anthrax in Livestock in Karnataka State using Remote Sensing and Climatic Variables, International Journal of Science and Research, 2319-7064.
- Eugen Popescu et al., 29 April 2022, Reducing the Effects of Drought and Degradation of Agricultural Soils, in the Context of Climate Change, through the Application of Regenerative Ecological Technologies, Chapter metrics overview, DOI: 10.5772/intechopen.104446.
- Jacob A. Miller-Klugesherz and Matthew R. Sanderson, 2023. Good for the soil, but good for the farmer? Addiction and recovery in transitions to regenerative agriculture, Volume 103, https://doi.org/10.1016/j.jrurstud.2023.103123
- Kasuni E. Adikari et al., 2021. Evaluation of Artificial Intelligence Models for Flood and Drought forecasting in Arid and Tropical regions, Elsevier
- L.Schreefel et al., September 2020, Regenerative agriculture – the soil is the base, Global food security, Volume 26, https://doi.org/10.1016/j.gfs.2020.100404.
- Mishra, A., Desai, V. and Singh V. 2007, Drought forecasting using a hybrid stochastic and neural network model. Journal of Hydrologic Engineering 12, 626–638. doi:10.1061/ ASCE1084-0699200712:6626.
- Rahul Patil, 2020. Drought Modelling and Forecasting using Arima and Neural Networks for Ballari District, Journal of the Indian Society of Agricultural Statistics 74(2) 149–157.
- Ravjit Khangura et al.,2023, Regenerative Agriculture—A Literature Review on the Practices and Mechanisms Used to Improve Soil Health.
- Ravjit Khangura et al.2023, Regenerative Agriculture—A Literature Review on the Practices and Mechanisms Used to Improve Soil Health, 2338; **https://doi.org/10.3390/su15032338**.
- Reibsame et al. 1991. Drought and Natural resource management in the United States :Impacts and Implications of the 1987-89 drought, Westview Press, Boulder, CO, p. 174.
- Stephen Sherwood and Norman Uphoff, August 2000. Soil health: research, practice and policy for a more regenerative agriculture, Applied Soil Ecology, Volume 15, Issue 1, Pages 85-97.