Turkish Journal of Agricultural and Natural Sciences 12 (3): 733–742, 2025
https://doi.org/10.30910/turkjans.1699513

TÜRK
TARIM ve DOĞA BİLİMLERİ
DERGİSİ

TURKISH
JOURNAL of AGRICULTURAL
and NATURAL SCIENCES

www.dergipark.gov.tr/turkjans

Original Article

Comparative Analysis of Trend Models for Standardized Precipitation Index (SPI) Data for Çanakkale

Nazlı Çiçek Söylemez [1] , Ünal Kızıl [2]✉

[1]Devlet Su İşleri Balıkesir 25. Bölge Müdürlüğü, Altıeylül, Balıkesir, Türkiye, [2]Çanakkale Onsekiz Mart Üniversitesi, Ziraat Fakültesi, Tarımsal Yapılar ve Sulama Bölümü, Çanakkale, Türkiye

[1]https://orcid.org/0009-0003-8660-9756, [2]https://orcid.org/0000-0002-8512-3899

✉: unal@comu.edu.tr

ABSTRACT

Understanding precipitation trends is critical for assessing climate change and its impacts on water resources, and disaster preparedness. In this study it was aimed to analyzes the long-term trends of the Standardized Precipitation Index (SPI) for Çanakkale. The precipitation data from a period of 1929 to 2023 was used. Three distinct models—Linear Regression, Autoregressive Integrated Moving Average (ARIMA), and Long Short-Term Memory (LSTM) networks—were employed to evaluate SPI trends. The linear regression model indicated significant short-term fluctuations in SPI values but did not reveal a clear long-term trend toward wetter or drier conditions. The ARIMA model, optimized for stationarity, also suggested relatively stable precipitation patterns, with no pronounced directional trend over the study period. The LSTM model, designed for sequential data analysis, captured complex temporal dependencies in SPI values but did not indicate a persistent long-term trend. Instead, the results highlighted substantial interannual variability in precipitation. These findings underscore the complexity of climate patterns in Çanakkale Province and emphasize the need for diverse modeling approaches to accurately assess precipitation trends. The lack of a clear directional trend suggests that short-term climate variability plays a more significant role than long-term changes in precipitation patterns in the region. This study provides a foundation for further research into advanced modeling techniques to enhance climate prediction capabilities. Future studies should explore hybrid and ensemble methods to improve accuracy, which is crucial for climate adaptation strategies and water resource management.

Key words: ARIMA, Linear Regression, Long Short-Term Memory, Climate Variability, Precipitation Trends.

Çanakkale için Standart Yağış İndeksi (SPI) Verileri için Trend Modellerinin Karşılaştırmalı Analizi

ÖZ

Yağış eğilimlerini anlamak, iklim değişikliğini ve su kaynakları ile afetlere karşı hazırlığı etkileyen faktörleri değerlendirmek için kritik öneme sahiptir. Bu çalışmada Çanakkale için Standart Yağış İndeksi'nin (SPI) uzun vadeli eğilimlerini analiz etmek amaçlanmıştır. 1929-2023 dönemine ait yağış verileri kullanılmıştır. SPI eğilimlerini değerlendirmek için Doğrusal Regresyon, Otoregresif Entegre Hareketli Ortalama (ARIMA) ve Uzun Kısa Vadeli Bellek (LSTM) ağları olmak üzere üç ayrı model kullanılmıştır. Doğrusal regresyon modeli, SPI değerlerinde kısa vadeli önemli dalgalanmalar göstermiştir ancak daha yağışlı veya daha kuru koşullara doğru net bir uzun vadeli eğilim ortaya koymamıştır. Durağanlık için optimize edilmiş ARIMA modeli, çalışma dönemi boyunca belirgin bir yön eğilimi olmaksızın nispeten istikrarlı yağış desenleri de önermiştir. Sıralı veri analizi için tasarlanan LSTM modeli, SPI değerlerinde karmaşık zamansal bağımlılıkları yakalamış ancak kalıcı uzun vadeli bir eğilim ortaya koymamıştır. Sonuçlar yağışta önemli yıllık değişkenliği ortaya koymaktadır. Bu bulgular, Çanakkale İli'ndeki iklim modellerinin karmaşıklığını vurgulamakta ve yağış eğilimlerini doğru bir şekilde değerlendirmek için çeşitli modelleme yaklaşımlarına olan ihtiyacı ortaya koymaktadır. Net bir yön eğiliminin olmaması, kısa vadeli iklim değişkenliğinin bölgedeki yağış modellerindeki uzun vadeli değişikliklerden daha önemli bir rol oynadığını göstermektedir. Bu çalışma, iklim tahmin yeteneklerini geliştirmek için gelişmiş modelleme tekniklerine yönelik daha fazla araştırma için bir temel sağlamıştır. Gelecekteki çalışmalar, iklim adaptasyon stratejileri ve su kaynakları yönetimi için çok önemli olan doğruluğu artırmak için hibrit ve topluluk yöntemlerini araştırmalıdır.

Anahtar kelimeler: ARIMA, Doğrusal Regresyon, Uzun Kısa Dönemli Bellek, İklim Değişkenliği, Yağış Eğilimleri.

INTRODUCTION

Studying precipitation trends is vital in climate science due to its significant impact on various environmental and societal aspects. Precipitation patterns influence water resources, agriculture, ecosystems, and urban planning, making it crucial to understand and predict these trends accurately. Analyzing spatial patterns of precipitation trends helps in understanding regional climatology and preparing for climatic changes. For example, research on the continental United States identified hotspots of increasing and decreasing precipitation trends, which can guide regional planning and adaptation strategies (O'Brien, 2018). Accurate prediction of precipitation trends is crucial for economic development as well. Studies in the Swat River basin, Pakistan, showed variability in precipitation trends across different time scales, highlighting the importance of precise predictions for water resource management and agricultural planning (Ahmad et al., 2015). On a global scale, mapping minimal detectable trends in annual precipitation is highly important for assessing the impact of climate change. This helps in identifying regions at risk and formulating global adaptation strategies (Morin, 2011). Detecting overlooked trends in precipitation, especially in extreme events, is critical for risk management in agriculture, water supply, and ecosystems. Quantile regression analyses have revealed significant trends in extreme precipitation that traditional methods might miss (Lausier & Jain, 2018). Precipitation trends directly affect the hydrological cycle and water management policies. Studies in Turkey demonstrated that understanding the spatial and temporal variations of precipitations is critical for effective management of water resources (Yavuz & Erdogan, 2012). Studying trends in precipitation alongside temperature helps predict future climate scenarios. Joint modeling of these parameters provides more accurate predictions of climate impacts, which are essential for planning and adaptation (Mesbahzadeh et al., 2019).

In this study it was aimed to monitor Standardized Precipitation Index (SPI) trends for further understanding and mitigating the impact of climate change. To achieve this objective three different approaches in modeling and predicting SPI data are employed using long term (1929 – 2023) precipitation data.

MATERIALS AND METHODS

Study Area and Dataset

Çanakkale province is located at of 39°27' north latitude and 26°16' east longitude (Figure 1). These coordinates show that the province is in the northwest of Turkey, bordering the Aegean and the Marmara Seas (Ünsal, 2015). Çanakkale has a climate that is transitional between the Mediterranean climate and the Black Sea climate. A warm and temperate climate prevails throughout the province. The average annual temperature is around 15°C. Summer months are hot and dry. In July and August, temperatures often rise above 30°C. Winters are warm and rainy. The average temperature in January varies between 5-10°C. Snowfall is rare, it usually rains. The mean precipitation varies between 600-700 mm. Precipitation generally occurred in fall and winter (Kale, 2017).

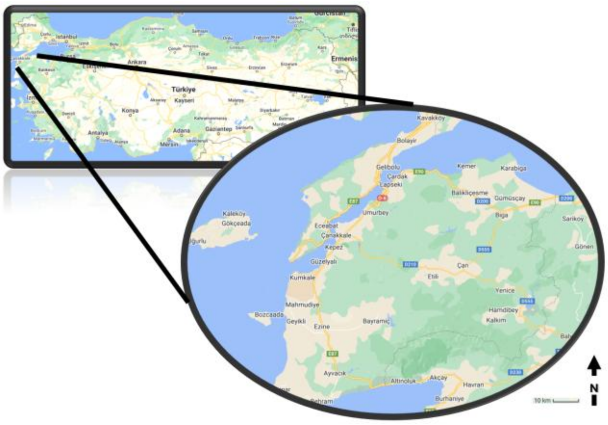

Figure 1. Study area

Montly precipitation data for the years 1929 - 2023 were obtained from the Turkish State Meteorological Service. The file containing precipitation data also includes data from all stations in the country. Therefore, first of all, only information about Çanakkale meteorology station was extracted from the txt file. These data were then converted into an Excel file and made ready for statistical analysis. Statistical calculations were made in Python programming language. Libraries such as Keras, Tensorflow, Sklearn, NumPy and Panda were used in the analyses. Necessary coding for trend analysis was done on Google Colab.

Standardized Precipitation Index (SPI)

The SPI is a commonly used technique designed to quantify precipitation deficits within different time scales. It provides a probabilistic measure of precipitation anomalies. Its calculation requires a historical data that fits probability distribution followed by transformation into a normal distribution. This allows the SPI to express the precipitation deficit or surplus as standard deviations from the mean. The commonly used probability distribution for SPI calculation is the gamma distribution, but other distributions such as normal, log-normal, and Weibull can also be employed based on the region and time scale (Angelidis et al., 2012). It can be calculated for different time scales, mainly from 1 to 48 months. Time scale of 1-6 months are useful for agricultural droughts, time scales of 12-48 months are better for hydrological droughts (Guenang & Kamga, 2014).

In this study it was aimed to monitor hydrological drought. Therefore, 12-month time scale was used for SPI calculations. The 12-month SPI integrates seasonal and interannual variability, providing a more stable and smoothed indicator of precipitation anomalies that directly impact water resource availability, reservoir levels, and watershed-scale hydrology. Additionally, focusing on a single, longer time scale reduced the dimensionality and complexity of model comparisons, allowing a clearer evaluation of the models' performance in capturing meaningful long-term climate signals. To fit a probability distribution, two-parameter gamma equation was used due to its flexibility and suitability for precipitation data (Equation 1) (Blain & Meschiatti, 2015).

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \tag{1}$$

Where; x is the variable of interest, which is the value at which the probability density function (PDF) is evaluated (precipitation amount), α is the shape parameter, β is the scale parameter, Γ is the gamma function, and e is the base of the natural logarithm.

Then, cumulative probabilities of each precipitation value were computed by using Equation (2).

$$G(x) = \int_0^x \frac{t^{\alpha-1}e^{-t/\beta}}{\beta^{\alpha}\Gamma(\alpha)}dt \tag{2}$$

Where; G(x) is the cumulative distribution function (CDF), t is variable of integration, which represents all possible values up to x, α is the shape parameter.

Finally, the cumulative probability transformed into standard normal distribution. This is done by the inverse normal (Gaussian) function (Equation 3).

$$SPI = \Phi\text{-}1(G(x)) \tag{3}$$

Where; Φ−1 represents the inverse function of the standard normal cumulative distribution.

Trend Analysis

Linear regression model

Linear regression is a fundamental statistical tool used to analyze and interpret trends in different fields. Linear regression models are employed to define relationships between an outcome (dependent) variable and one or more input (independent) variables. The fundamental equation governing this relationship is presented in Equation (4).

$$Y=\beta 0+\beta 1X+\epsilon \tag{4}$$

Where; Y represents the dependent variable while X represents the independent variable. β0, β1, and ϵ are the intercept, slope and error term, respectively.

Autoregressive integrated moving average (ARIMA)

In environmental sciences, ARIMA models are used to analyze trends in climatic data, such as precipitation and temperature patterns, which are crucial for understanding climate change impacts and planning mitigation strategies (Dimri et al., 2020). These models are characterized by three key parameters: p (autoregressive order), d (differencing order), and q (moving average order), and are generally represented as ARIMA(p,d,q). The Autoregressive (p) component signifies that the variable of interest is predicted based on its past values. The integrated (d) component accounts for differencing the raw data to achieve stationarity in the time series. Lastly, the moving average (q) component identifies the relationship between an observed value and the residual errors. The general equation for the ARIMA model is presented in Equation (5).

$$Y_t=c+\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+...+\phi_p Y_{t-p}+\theta_1 \epsilon_{t-1}+\theta_2 \epsilon_{t-2}+...+\theta_q \epsilon_{t-q}+\epsilon_t \quad (5)$$

Where; Y_t represents the value at time t, c is a constant, ϕ_i are the coefficients for the autoregressive terms, θ_j are the coefficients for the moving average terms, and ϵ_t is the error term at time t (Green & Noles, 1977).

The best model parameters are determined using the AIC (Akaike Information Criterion), and the combination of (p, d, q) which yields the lowest value is defined as the optimal model. Zhang & Meng, (2023) demonstrates the application of AIC in the selection of ARIMA model parameters, balancing model complexity and accuracy. The AIC criterion aids in selecting the optimal model by minimizing its value, ensuring the best combination of (p, d, q) is chosen.

Using the ADF test's p-value and ADF statistics it was evaluated that if the data is stationary. If this value is less than a threshold value (commonly 0.05), this means that the data is likely stationary. Additionally, if the ADF test statistic is more negative than the critical values at different significance levels (1%, 5%, 10%), it indicates that the data is stationary. In summary, a low p-value (typically less than 0.05) or an ADF statistic more negative than the critical value suggests that both data are stationary.

Long Short-Term Memory Networks (LSTM)

LSTMs are a type of recurrent neural networks (RNNs). This advanced type is developed to handle the shortcomings of traditional RNNs. They perform well in managing long-term dependencies. They are highly effective for sequential data tasks, such as time series analysis (Oruh et al., 2022).

LSTM networks consist of multiple units, each incorporating a cell, an input gate, an output gate, and a forget gate. The traffic of information within the network is controlled by these gates, enabling LSTMs to efficiently retain and update long-term dependencies. They operate by iteratively processing input sequences through the gates and cell states. The primary equations governing an LSTM unit at time step t are given in following Equations (6, 7, 8, 9, 10, 11).

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f) \quad (6)$$

Where; f_t represents the forget gate activation at time step t, σ denotes the Sigmoid activation function, W_f is the weight matrix associated with the forget gate, h_{t-1} is hidden state from the previous time step $t-1$, x_t is input at the current time step t, and b_f is bias for the forget gate.

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i) \quad (7)$$
$$\acute{C}_t=\tanh(W_C\cdot[h_{t-1},x_t]+b_C) \quad (8)$$

Where; i_t is input gate activation at time step t, W_i is weight matrix for the input gate, b_i is bias for the input gate, $\acute{C}_t$ is candidate cell state at time step t, tanh is hyperbolic tangent activation function, W_C is weight matrix for the candidate cell state, and b_C is bias for the candidate cell state.

$$C_t=f_t\cdot C_{t-1}+i_t\cdot \acute{C}_t \quad (9)$$

Where; C_t is updated cell state at time step t, f_t is forget gate activation at time step t, C_{t-1} is cell state from the previous time step $t-1$.

$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o) \quad (10)$$
$$h_t=o_t\cdot\tanh(C_t) \quad (11)$$

Where; ot is output gate activation at time step t, Wo is weight matrix for the output gate, ht−1 is hidden state from the previous time step t−1, xt is input at the current time step t, bo is bias for the output gate, ht is hidden state or output of the LSTM unit at time step t, Ct is cell state at time step t.

The statistical parameters of LSTM networks, such as weights, biases, activation functions, learning rate, number of units, batch size, and dropout rate, play a highly important role in their performance. Understanding and optimizing these parameters is essential for leveraging the full potential of LSTM networks in various applications (Merity et al., 2017).

RESULTS AND DISCUSSION

Linear Regression Model

12-month SPI and trend line drawn via linear regression model are represented in Figure 2. The SPI values exhibit significant fluctuations, indicating variability in annual precipitation with periods of both wetter and drier conditions. On the other hand, the nearly horizontal red trend line indicates that, on average, there is no significant increase or decrease in annual SPI values over the observed period. This implies that there is no strong indication of wetter or drier conditions for the long term.

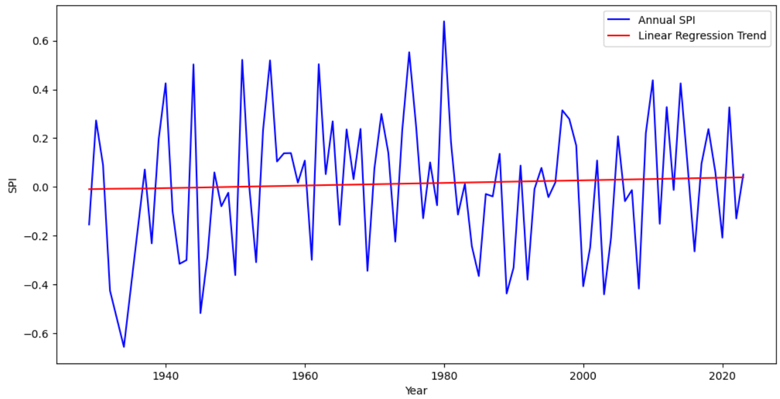

Figure 2. Annual SPI trend analysis with linear regression

The statistics for the linear regression model are listed in Table 1. Given in Table 1, a highly small R2 value suggesting that the linear trend does not capture much of the variability in the data. Additionally, a relatively low MSE value indicates that the model's predictions closely align with the actual values on average. Another parameter is the slope of the linear regression model. The small positive slope indicates a slight upward trend in SPI values over time, but the trend is very minimal.

Table 1. Linear regression model statistics

Parameter	Value
R^2	0.00297
Mean Squared Error (MSE)	0.07507
Slope	0.00056
Intercept	-1.08741

The linear regression analysis reveals no substantial long-term trend in the annual SPI values. The natural variability in annual precipitation remains the dominant feature of the SPI data.

ARIMA Analysis

After calculating the annual SPI values (Figure 2), an ARIMA trend analysis was conducted. Important model parameters in this analysis are defined by the letters (p, d, q) as explained above. Using Python’s itertools library, the optimal parameters that yield lowest AIC value were attempted to be determined by trial-error method. However, the iteration resulted in (0,0,0) values, indicating that the data is non-stationary. Stationarity is highly important for ARIMA, because these models assume that the statistical properties of the series are not changed. To convert the data from non-stationary to stationary form, differencing was applied. After differencing, the data appeared as given in Figure 3.

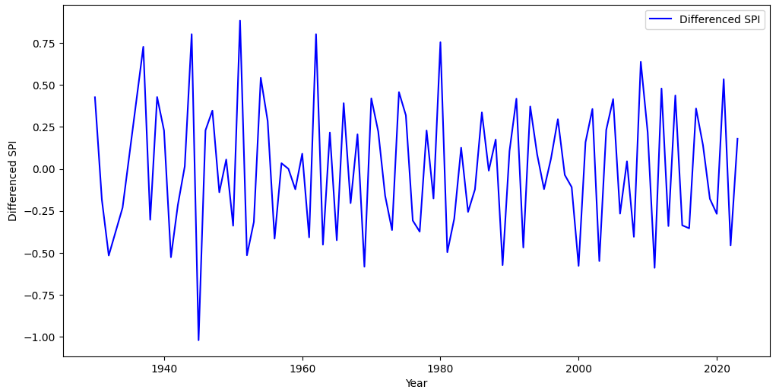

Figure 3. Differenced annual SPI data

Stationarity of original and differenced data was evaluated using the Augmented Dickey-Fuller (ADF) test. The results are listed in Table 2.

Table 2. ADF test statistics

	Value	
Parameter	**Original Data**	**Differenced Data**
ADF Statistic:	-9.59	-6.055
p-value:	2.04×10^{-16}	1.25×10^{-7}
Critical Values:		
1%:	-3.50	-3.51
5%:	-2.89	-2.90
10%:	-2.58	-2.59

Based on the test statistics an ARIMA model without differencing was used as the data does not require differencing.While the test results indicate that the original data is stationary, the inability to fit ARIMA models effectively on the original data suggests that there might be other factors at play, such as higher-order dependencies or non-linearities that are not captured by simple ARIMA models without differencing. As a result, differenced data is used to run ARIMA model. Once the iteration process was applied (1,1,1) model fitted well on the differenced data. Therefore ARIMA (1, 1, 1) model was applied. The ARIMA model output and differenced SPI data plot are given in Figure 4.

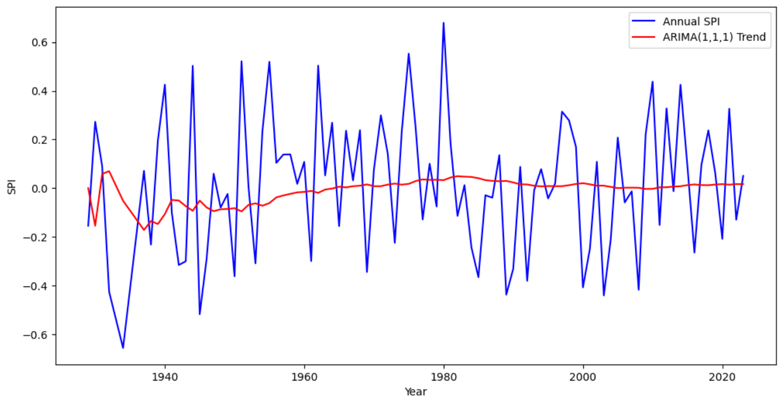

Figure 4. Annual SPI trend analysis with selected ARIMA model

Based on ARIMA, the trend line appears relatively flat with minor fluctuations over the years. This flat trend line suggests that there is no significant long-term trend in the SPI values, indicating either wetter or drier conditions over the entire period. In early years (Pre-1950s) the SPI values exhibit high variability with noticeable peaks and troughs. The red trend line shows a slight downward trend initially, indicating drier conditions during this period.

However, during mid to late period (Post-1950s) the SPI values continue to show variability, but the fluctuations become more moderate. The red trend line flattens out and remains close to zero, indicating that there is no strong indication of wetter or drier conditions for this period. The ARIMA (1, 1, 1) model indicates that the annual SPI values have not shown a consistent long-term trend towards wetter or drier conditions over the years. The conditions appear relatively stable, with fluctuations that do not point towards a significant change in either direction.

Long Short-Term Memory (LSTM) Networks Analysis

Used as a version of recurrent neural networks that is particularly good at learning from sequences and temporal data, making it applicable for time series analysis. Improving the performance of an LSTM model can involve several strategies, including tuning hyperparameters, increasing model complexity, adding regularization, and improving data preprocessing. In order to find the select best LSTM parameters, Keras' optimizer library was employed. Based on the optimization (tuning hyperparameters) following model was developed representing the SPI data trend (Table 3).

Table 3. LSTM model parameters

Parameter	Value
Validation loss	0.0272
Units in the first layer	64
Units in the second layer	160
Dropout rate for first layer	0.2
Dropout rate for second layer	0.3
Number of layers	3
Learning rate	0.0025

Based on Table 3, A lower validation loss indicates better generalization and model performance. The number of units (neurons) in the first LSTM layer having 64 units means that the layer has 64 LSTM cells, each with its own set of weights and biases to learn from the data. Similar to the first LSTM layer, the number of units in the second LSTM layer having 160 units suggests that this layer is larger and might be capturing more detailed temporal patterns from the data. A dropout rate of 0.2 means that 20% of the neurons in the first LSTM layer are randomly dropped during each epoch, forcing the model to learn redundant representations and improve its generalization ability. Similarly, this dropout rate is applied to the second LSTM layer. A rate of approximately 0.3 (30%) means that 30% of the neurons in this layer are randomly dropped during training. The number of layers indicates the total number of layers in the model. In this case, the optimal number is 3, which typically includes the two LSTM layers and one additional layer (often a Dense layer) for the final output. A learning rate of approximately 0.002465 means the optimizer (Adam) will update the model weights by this fraction of the gradient. Selecting an appropriate learning rate is essential for model performance. If the learning rate is too high, the model may converge too quickly to a suboptimal solution, whereas a low learning rate can lead to an overly slow training process. Utilizing this optimized model, the trend analysis obtained through LSTM is presented in Figure 5.

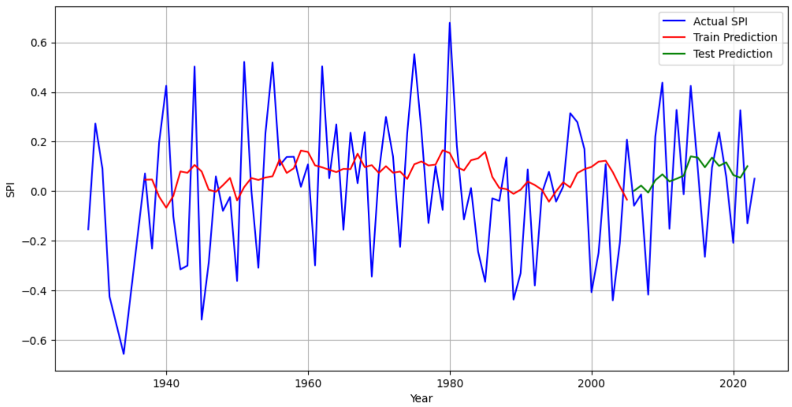

Figure 5. Improved annual SPI trend analysis with LSTM

As stated above the SPI values fluctuate significantly, indicating periods of both wet and dry conditions. The red line, on the other hand, represents the predicted values. The model reflects some general trends in the training data, but there are some deviations from the actual SPI values. It also seems to smooth out some of the higher peaks and lower troughs, indicating that it captures the overall trend but might miss some short-term fluctuations. The green line, representing the model's predictions on the test dataset, follows the general trend of the actual SPI values reasonably well, with some deviations. The performance of the model on the test set indicates it has generalized well to unseen data, though it still shows some smoothing of extreme values. The overall trend captured by the LSTM model does not show a clear and consistent long-term increase or decrease in SPI values, which would indicate a significant trend towards wetter or drier conditions. There are periods of negative SPI values, but these are interspersed with periods of positive SPI values, indicating a balance between wet and dry periods over time. As a result, based on the trend analysis, there is no clear long-term trend towards drought in the given SPI data. The SPI values fluctuate around zero, indicating alternating periods of wet and dry conditions without a significant downward trend.

Ahmad et al. (2015) determined high variability in SPI trends in the Swat River basin using non-parametric tests, stating that precipitation trends are better explained with flexible and dynamic modeling rather than static linear methods. Yavuz & Erdogan (2012) also determined strong spatial and temporal variability in Turkish precipitation records with no consistent trends between areas, which corroborated the theory that local influences have a tendency to overpower larger-scale climate signals.

At the global level, Morin (2011) highlighted that the few discernible trends in annual precipitation are overshadowed by huge interannual variability, a result mimicked in this study's outcome of SPI trends fluctuating around a zero mean with no directional trend. In addition, Lausier & Jain (2018) demonstrated how traditional statistical approaches can overlook subtle trends, especially in extremes, which further validates the manuscript's use of advanced models like LSTM in identifying subtle trends.

Model Performance Evaluation and Statistical Comparison

To enhance the comparative evaluation of the three models, additional performance metrics—Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2)—were calculated (Table 4). The Linear Regression model had the poorest performance, with RMSE of 1.867, MAE of 1.693, and a negative R^2 value of -2.508, indicating that it failed to capture the variation in SPI values. In contrast, both the ARIMA and LSTM models yielded substantially better metrics, with nearly identical RMSE (0.998 for ARIMA, 0.996 for LSTM) and MAE (0.791 for ARIMA, 0.790 for LSTM), though their R^2 values remained close to zero (−0.0031 and 0.0004 respectively), suggesting limited predictive strength in terms of explained variance.

Table 4. Model Performance Metrics

	RMSE	MAE	R^2
Linear Regression	1.8669	1.6923	-2.5089
ARIMA	0.9982	0.7912	-0.0031
LSTM	0.9964	0.7902	0.0004

To statistically assess the difference in forecast accuracy, the Diebold-Mariano (DM) test was applied (Table 5). The test showed statistically significant differences between Linear Regression and both ARIMA (DM = 27.83, $p < 0.001$) and LSTM (DM = 27.91, $p < 0.001$). However, the difference between ARIMA and LSTM was not statistically significant (DM = 0.61, p = 0.54). These results confirm that while ARIMA and LSTM models outperform Linear Regression, their relative predictive performance is statistically similar. Therefore, both ARIMA and LSTM offer more reliable modeling options for SPI trend analysis in Çanakkale.

Table 5. Diebold-Mariano Test Results

	DM Statistic	p-value
LR vs ARIMA	27.8322	0.0000
LR vs LSTM	27.9088	0.0000
ARIMA vs LSTM	0.6104	0.5417

CONCLUSION

This study meticulously analyzed the Standardized Precipitation Index (SPI) for Çanakkale using three distinct models—Linear Regression (LR), Autoregressive Integrated Moving Average (ARIMA), and Long Short-Term Memory (LSTM)—spanning the period from 1929 to 2023. The findings reveal that none of the models identified a strong, consistent long-term trend in precipitation, highlighting the complex and variable nature of climate patterns in the region.

Quantitative performance metrics further clarified the relative effectiveness of the models. Linear Regression exhibited the weakest performance, with the highest RMSE and MAE values and a strongly negative R^2, indicating its inability to capture the underlying variability in SPI. In contrast, both ARIMA and LSTM models achieved significantly lower error rates and nearly identical performance, though their R^2 values remained close to zero. The Diebold-Mariano statistical test confirmed that both ARIMA and LSTM significantly outperformed Linear Regression in forecasting accuracy ($p < 0.001$), while no statistically significant difference was found between ARIMA and LSTM ($p = 0.54$).

The analysis carried out here took into consideration trend modeling of SPI values based on local precipitation only, ignoring overall atmospheric or oceanic climate forcing factors. Thus, the lack of persistence in a long-term trend in SPI does not necessarily imply climatic stability, but is instead a reflection of the non-availability of modulating variables that effect precipitation variability on interannual to decadal timescales. Climate indices such as the North Atlantic Oscillation (NAO), the Arctic Oscillation (AO), and the El Niño-Southern Oscillation (ENSO) have been seen to influence regional precipitation traits over Europe and the eastern Mediterranean, including Turkey. For example, positive and negative phases of the NAO have been shown to be related to wetter or drier conditions in western Anatolia. Leaving these indices out may then minimize SPI trend interpretation to disregard such important teleconnections that could be primarily behind the interannual variability observed. Although no statistically significant long-term trend towards increasing or decreasing precipitation based on SPI-12 analysis was detected in this research, the fact of substantial interannual variability renders adaptive water management essential to Çanakkale. Policy makers need to put top priority on the development of variable drought preparedness policies capable of responding to short-term variations over the single strategy of long-term prediction. Investments should be made in climate-resilient infrastructure, such as multi-year water storage infrastructure, to serve as a cushion against unpredictable dry years.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Author Contributions

The authors declare that they have contributed equally to the article.

ORCID

Nazlı Çiçek SÖYLEMEZ **http://orcid.org/0009-0003-8660-9756**
Ünal KIZIL **http://orcid.org/0000-0002-8512-3899**

Article History	
Submission received:	14.05.2025
Revised:	06.07.2025
Accepted:	09.07.2025

REFERENCES

Ahmad, I., Tang, D., Wang, T., Wang, M., & Wagan, B. (2015). Precipitation Trends over Time Using Mann-Kendall and Spearman's rho Tests in Swat River Basin, Pakistan. *Advances in Meteorology*, 2015, 1-15. https://doi.org/10.1155/2015/431860.

Angelidis, P., Maris, F., Kotsovinos, N., & Hrissanthou, V. (2012). Computation of Drought Index SPI with Alternative Distribution Functions. *Water Resources Management, 26, 2453-2473*. https://doi.org/10.1007/s11269-012-0026-0.

Blain, G., & Meschiatti, M. (2015). Inadequacy of the gamma distribution to calculate the Standardized Precipitation Index. *Revista Brasileira de Engenharia Agricola e Ambiental*, 19, 1129-1135. https://doi.org/10.1590/1807-1929/AGRIAMBI.V19N12P1129-1135.

Dimri, T., Ahmad, S., & Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach. *Journal of Earth System Science*, 129, 1-16. https://doi.org/10.1007/s12040-020-01408-x.

Green, S., & Noles, K. (1977). A Computer Program for the Generation of ARIMA Data. *Educational and Psychological Measurement*, 37, 249 - 252. https://doi.org/10.1177/001316447703700132.

Guenang, G., & Kamga, F. (2014). Computation of the Standardized Precipitation Index (SPI) and Its Use to Assess Drought Occurrences in Cameroon over Recent Decades. *Journal of Applied Meteorology and Climatology*, 53, 2310-2324. https://doi.org/10.1175/JAMC-D-14-0032.1.

Kale, S. (2017). Climatic Trends in the Temperature of Çanakkale City, Turkey. *Natural and Engineering Sciences*, 2, 14-27. https://doi.org/10.28978/NESCIENCES.348449.

Lausier, A., & Jain, S. (2018). Overlooked Trends in Observed Global Annual Precipitation Reveal Underestimated Risks. *Scientific Reports*, 8. https://doi.org/10.1038/s41598-018-34993-5.

Merity, S., Keskar, N., & Socher, R. (2017). Regularizing and Optimizing LSTM Language Models. *ArXiv*, abs/1708.02182.

Mesbahzadeh, T., Miglietta, M., Mirakbari, M., Sardoo, F., & Abdolhoseini, M. (2019). Joint Modeling of Precipitation and Temperature Using Copula Theory for Current and Future Prediction under Climate Change Scenarios in Arid Lands (Case Study, Kerman Province, Iran). *Advances in Meteorology*. https://doi.org/10.1155/2019/6848049.

Morin, E. (2011). To know what we cannot know: Global mapping of minimal detectable absolute trends in annual precipitation. *Water Resources Research*, 47. https://doi.org/10.1029/2010WR009798.

O'Brien, S. (2018). Spatial Patterns of Precipitation Trends in the Continental United States, 1950-2016. (Masters Thesis) Fort Hays State University, USA.

Oruh, J., Viriri, S., & Adegun, A. (2022). Long Short-Term Memory Recurrent Neural Network for Automatic Speech Recognition. *IEEE Access*, 10, 30069-30079. https://doi.org/10.1109/ACCESS.2022.3159339.

Ünsal, V. (2015). Fırst Traces Of Settlement In Canakkale And Gallıpolı Penınsula. *The Journal of International Social Research*, 8, 340-340. https://doi.org/10.17719/JISR.20154013909.

Yavuz, H., & Erdogan, S. (2012). Spatial Analysis of Monthly and Annual Precipitation Trends in Turkey. *Water Resources Management*, 26, 609-621. https://doi.org/10.1007/s11269-011-9935-6.

Zhang, Y., & Meng, G. (2023). Simulation of an Adaptive Model Based on AIC and BIC ARIMA Predictions. *Journal of Physics: Conference Series*, 2449. https://doi.org/10.1088/1742-6596/2449/1/012027.