



Drought index time series forecasting via three-in-one machine learning concept for the Euphrates basin

Levent Latifoğlu¹ · Savaş Bayram¹ · Gaye Aktürk² · Hatice Citakoglu¹

Received: 9 July 2024 / Accepted: 22 August 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Droughts are among the most hazardous and costly natural disasters and are hard to quantify and characterize. Accurate drought forecasting reduces droughts' devastating economic effects on ecosystems and people. Eastern Anatolia is the largest and coldest geographical region of Türkiye. Previous studies lack drought forecasting in the Eastern Anatolia (Upper Mesopotamia) Region, where agriculture is limited due to being under snow most of the year. This study focuses on the Euphrates basin, specifically the Tercan and the Tunceli meteorological stations of the Karasu River sub-basin, a vital Eastern Anatolia Region water resource. In this context, time series of 1-, 3-, 6-, 9-, and 12-month Standardized Precipitation Index (SPI) and Standardized Precipitation Evapotranspiration Index (SPEI) values were created. The Tuned Q-factor Wavelet Transform (TQWT) method and Univariate Feature Ranking Using F-Tests (FSRFtest) were used for pre-processing and feature selection. Several models were created, such as stand-alone, hybrid, and tribrid. Machine Learning (ML) methods such as Artificial Neural Networks (ANN), Gaussian Process Regression (GPR), and Support Vector Machine (SVM) were conducted for the time series analyses. The GPR approach was concluded to perform better than the ANN and SVM at the Tercan station. In other words, GPR performs better in 80% of cases than SVM and ANN models. At the Tunceli station for the SPI output, SVM, which had a superior performance in 60% of the cases, demonstrated a performance comparable to GPR. At the same time, ANN once again exhibited an inferior performance. Similarly, for the SPEI output at the Tunceli station, no clear superiority was observed between the GPR and ANN methods. Because both methods were successful in 40% of cases. This study contributes by introducing a third concept to the stand-alone and hybrid model comparison of drought forecasting, adding tribrid models. It has been detected that the Hybrid and Tribrid ML methods lead to a 91% and 64% decrease relative root mean square error percentage compared stand-alone ML methods for SPEI and SPI in two stations. While the hybrid model at Tercan station was more successful in 80% of the cases, the hybrid model at Tercan station was more successful in 90% of the cases. While hybrid models were observed to be superior, tribrid models not only demonstrated performance close to the hybrid models but also provided advantages such as reducing computational load and shortening calculation time.

Keywords Climate Change · Drought · Forecasting · Standardized Precipitation Index · Standardized Precipitation Evapotranspiration Index · Machine Learning

Communicated by: Hassan Babaie

✉ Hatice Citakoglu
hcitakoglu@erciyes.edu.tr

Levent Latifoğlu
latifoglu@erciyes.edu.tr

Savaş Bayram
sbayram@erciyes.edu.tr

Gaye Aktürk
gayeoguzturk@kku.edu.tr

¹ Department of Civil Engineering, Erciyes University, Kayseri, Türkiye

² Department of Civil Engineering, Kırıkkale University, Kırıkkale, Türkiye

Introduction

The environment is constantly affected by the rapidly increasing population rate and technological progress (Kilinc and Yurtsever 2022). Climate change, which directly affects all living things in both the short and long term, has emerged as one of the foremost environmental challenges of our age (Katip 2018), especially gaining importance in recent years with the increasing awareness of environmental issues. Climate change is a statistically significant change in the average state or variability of the climate over an extended period, which can occur both due to changes in natural climate dynamics and external factors resulting from human actions. Various potential future scenarios, such as changing rainfall patterns, higher temperatures, and rising sea levels, could result from increasing climate change influences, the consequences of which could play a significant role in ecosystems, societies, and economies (IPCC 2014). Increasing temperatures, irregularities in precipitation, and changes in the frequency of extreme events change the total and seasonal water supply, and when these changes are combined with land use, they significantly affect hydrological processes at the basin level (Wang et al. 2019; Zhang et al. 2019; Alivi et al. 2021). This situation may stand out as considerable evidence about the effects of climate change. Therefore, it is vital to determine and implement management strategies compatible with climate change for the sustainable use of water resources today.

The Intergovernmental Panel on Climate Change (IPCC) Report on extreme events, which supports the targets of the United Nations Framework Convention on Climate Change, has acknowledged drought as a significant extreme climatic event requiring mitigation to minimize its adverse impacts (Field 2012; Deo et al. 2017b). Drought is a persistent and recurrent natural disaster on a global scale, often linked to climate change (Mohammed et al. 2018; Naumann et al. 2018). Drought forecasting is critical to combating drought natural disasters and is vital for risk management and mitigating the drought management process (Mishra and Singh 2010, 2011; Belayneh et al. 2016; Deo et al. 2017b). Additionally, drought forecasting accurately helps to lessen their catastrophic economic effects on ecosystems and people (Mo et al. 2009). It is challenging to forecast when a drought will start because it might appear out of nowhere, move swiftly, and have several outcomes (Wilhite 2000). Droughts are agricultural, hydrological, meteorological, and socioeconomic (Katip 2018; Evkaya and Kurnaz 2021). Most other natural disasters are not like droughts in many aspects, especially when it comes to the degree of difficulty in forecasting the drought's start, end, and severity (McKee et al. 1993).

Türkiye, situated in a semi-arid environment, is vulnerable to the disastrous effects of droughts (Aibaidula

et al. 2022). Therefore, forecasting the possible impacts of future droughts is vital. This study, specifically focusing on the Euphrates basin, aims to develop a more accurate and efficient method for drought forecasting. However, traditional stochastic models have limitations in forecasting nonlinear data.

Drought indices are crucial for continuously monitoring drought events in terms of their temporal and spatial extent, assessment of their severity and spatial dimension quantitatively, and early identification, i.e., predicting drought and enabling the development of management strategies under current climate conditions. Drought assessment plays a significant role in the planning and management of water resources. Among these drought indices, the Standardized Precipitation Index (SPI) and the Standardized Precipitation Evapotranspiration Index (SPEI) are widely used in the literature due to their ability to be utilized at multiple time scales, represent various types of droughts, and better reflect changes in drought characteristics (severity, duration, frequency, and spatial extent) (Mishra and Singh 2010). SPI and SPEI are commonly used in global research and calculated at various time scales such as 1-, 3-, 6-, 12-, 24-, and 48-months to represent droughts. The frequency of dry periods is exceptionally high at brief time scales, but as the time scale increases, the frequency of dry periods decreases. Therefore, while drought frequency decreases at longer time scales, increases in its severity and duration are observed. Depending on the type of water source available, response times to drought conditions vary significantly. In order to define the system response period, (McKee et al. 1993) proposed the notion of the drought timescale. For example, short time scales (1 to 3 months) are mainly related to soil water content and river discharge in headwater areas, medium time scales (3 to 12-months) are related to reservoir storage and discharge in the medium course of the rivers, and longtime scales (12 to 24-months) are related to variations in groundwater storage. As a result, different time scales can be used to observe drought conditions in different hydrological subsystems (Vicente-Serrano et al. 2010).

In recent years, ML algorithms have spurred significant advancements in drought prediction (Felsche and Ludwig 2021; Piri et al. 2023; Danandeh Mehr et al. 2023). A variety of ML algorithms have been developed and employed for this purpose, encompassing random forest, k-nearest neighbors, artificial neural networks, support vector machines, Gaussian Process Regression, adaptive neuro-fuzzy inference systems, decision trees, multivariate adaptive regression splines, M5 Trees, including long short-term memory, extreme learning machine, and extreme learning machines (Deo et al. 2017a; Kisi et al. 2019; Yaseen et al. 2021; Kikon and Deka 2022; Docheshmeh et al. 2022; Lotfifard et al. 2022; Moghaddasi et al. 2024; Lalika et al. 2024).

Katip (2018) applied a standardized precipitation index (SPI) to the Marmara region. Neural Network (NN) models were utilized. Five models, among six were successful (Katip 2018). Özger et al. (2020) applied the self-calibrated Palmer Drought Severity Index (sc-PDSI) to the Mediterranean region. Models were created via Empirical mode decomposition (EMD) and Wavelet decomposition (WD). WD was more successful (Özger et al. 2020). (Mehr et al. 2020) applied the SPI to the Ankara Province, Turkey. This paper presents a new hybrid model called ENN-SA for spatiotemporal drought estimation. In ENN-SA, an Elman neural network (ENN) is combined with simulated annealing (SA) optimization and support vector machine (SVM) classification algorithms for standardized precipitation index (SPI) modeling across multiple stations. In the study, researchers have shown that ENN-SA is promising and effective for multi-station SPI estimation. Altunkaynak and Jalilzadnezamabad (2021) applied the Palmer drought severity index (PDSI) to the Marmara region. Models were created via Discrete Wavelet Transform (DWT), fuzzy, k-Nearest Neighbour (kNN), and Support Vector Machine (SVM). The hybrid models outperformed stand-alone ones (Altunkaynak and Jalilzadnezamabad 2021). Evkaya and Kurnaz (2021) applied univariate drought index (UDI) and SPI to the Marmara region. External Input (NARX) type NN models were created. It was found that the drought index forecasting capacity could be increased using (NARX-NN) (Evkaya and Kurnaz 2021). Başakın et al. 2021a, b applied the self-calibrated Palmer Drought Severity Index (sc-PDSI) to the Mediterranean region. They proposed a new hybrid model via an adaptive neuro-fuzzy inference system (ANFIS) and EMD, namely EMD-ANFIS and stand-alone ANFIS. The hybrid model outperformed (Başakın et al. 2021a). Citakoglu and Coşkun (2022) applied SPI to Marmara region. Both stand-alone and hybrid models were created via ANFIS, Gaussian process regression (GPR), k-nearest neighbors (KNN), NN, and support vector machine regression (SVM). Hybrid GPR and NN models outperformed (Citakoglu and Coşkun 2022). Gholizadeh et al. (2022) applied SPEI to the Central Anatolia region. They proposed a new hybrid model via the Bat optimization algorithm and extreme learning machine (ELM), namely Bat-ELM. The proposed model improved the forecasting accuracy (Gholizadeh et al. 2022). Kilinc and Yurtsever (2022) proposed a grey wolf algorithm (GWO) based gated recurrent unit (GRU) hybrid approach for the Mediterranean region. GWO-GRU model was successful (Kilinc and Yurtsever 2022). Gul et al. (2023) applied SPI to the Aegean region. They proposed Extreme Gradient Boosting (XgBoost), Adaptive Boosting, and Gradient Boosting. XgBoost outperformed (Gul et al. 2023). Reihanifar et al. (2023) presented a new model named multi-objective multi-gene genetic programming (MOMGGP), compared with genetic programming and multi-gene genetic

programming. The same forecasting accuracy was obtained from MOMGGP (Reihanifar et al. 2023). Soylyu Pekpostalci et al. (2023) evaluated 71 drought monitoring and forecasting studies from 2010 to 2022 in Türkiye. The application of ML for short-term hydrological and meteorological drought forecasting was trending upward (Soylyu Pekpostalci et al. 2023). Danandeh Mehr et al. (2023) applied SPEI to the Central Anatolia. They proposed a new hybrid model via convolutional neural network (CNN) and long short-term memory (LSTM), namely convolutional long short-term memory (CNN-LSTM). The proposed model improved the forecasting accuracy (Danandeh Mehr et al. 2023). In addition, researchers have recently tried to suggest various hybrid models that can be used on in hydrological forecasting (Yuan et al. 2018; Adnan et al. 2021, 2022, 2023; Ikram et al. 2023; Mostafa et al. 2023; Mohammadi 2023).

There are various drought studies regarding the Euphrates basin in the literature. For example, (Katipoglu et al. 2020) compared SPI, SPEI, Statistical Z-Score Index (ZSI), Precipitation Anomaly Index (RAI), and Reconnaissance Drought Index (RDI) on a 3-month and 12-month time scale. (Katipoglu et al. 2021) mapped the SPI, ZSI, RAI, SPEI, and RDI using Kriging, Radial Basis Function (RBF), and Inverse Distance Weighting (IDW) methods at three and 12-month time scales. As seen in these studies from Türkiye, no existing studies predict the drought of the Euphrates basin. (Katipoğlu and Acar 2022). Calculated trends were calculated using Mann Kendall (MK) and Modified Mann Kendall (MMK) tests of the Standardized flow (SRI) index at three and 12-month time scales. Their studies also mapped drought trends using Kriging, RBF, IDW, Local Polynomial Interpolation (LPI) and Global Polynomial Interpolation (GPI) methods. (Katipoğlu et al. 2022) calculated and compared the trends with MK and MMK tests of SPI, SPEI, ZSI, RAI and RDI at three and 12-month time scales. (Esit et al. 2023) calculated and compared the trends of SPI, SPEI and the standardized streamflow index (SDI) with MK, Spearman Rho, and innovative trend analysis tests at a 12-month time scale. As seen in these studies from Türkiye, no existing studies have predicted the drought of the Euphrates basin using ML methods. In addition, although drought prediction studies have been carried out in many geographical regions of Türkiye, no current study has been found in the Euphrates basin, the coldest and largest geographical region. Therefore, this study aims to close this gap by estimating SPI and SPEI values obtained from existing stations in the Euphrates basin in Türkiye using ML methods.

This study selected the application units of two meteorological measurement stations in the Euphrates Basin, specifically for the Karasu sub-basin in the Eastern Anatolia region. Within the scope of the study, 1965–2022 monthly precipitation and temperature data of existing stations were used. Monthly SPI values (SPI-1, SPI-3, SPI-6, SPI-9, and

SPI-12) and SPEI values (SPEI-1, SPEI-3, SPEI-6, SPEI-9, and SPEI-12) were generated for 1, 3, 6, 9 and 12-time scales. SPI and SPEI outputs were utilized separately to monitor and describe the development of temporal variations of meteorological drought events over various time scales. A forecasting process was conducted using three different approaches: (i) Simple ML methods (ANN, GPR, SVM), (ii) Hybrid ML methods (TQWT-ANN, TQWT-GPR, TQWT-SVM) and (iii) Tribrid ML methods (TQWT-FSRFtest-ANN, TQWT-FSRFtest-GPR, TQWT-FSRFtest-SVM). Another critical purpose of this study, along with these three different approaches, was to develop high-performance drought prediction models and to reveal the advantages and disadvantages of the models by comparing the performances of simple, hybrid, and tribrid approaches. This study will contribute to the literature by offering a third alternative to the single-hybrid comparison.

Material & method

Study area and data

This study was conducted in the River Karasu, a sub-basin of the Euphrates Basin, which is situated between latitudes 40°20'–38°23' N and longitudes 37°17'–41°34' E. The Euphrates River, one of the most significant rivers in the Middle East, is in the most mountainous part of eastern Türkiye. The Karasu, on the other hand, is one of the primary tributaries contributing to the formation of the Euphrates River. The Karasu River Basin has an approximate surface area of 37,339 km², representing about 4.77% of Türkiye's total surface area and 30.70% of the Euphrates Basin. The Karasu River Basin is surrounded by mountainous areas consisting of volcanic masses. It is the most mountainous part of the Euphrates Basin, with a maximum altitude of 3,537 m in the eastern and interior areas and a minimum altitude of 800 m in the south. While 57.4% of the Basin has steep and very steep topography, 10.9% has relatively flat terrain. The basin's land structure consists of 0.7% urban areas, 14.2% agricultural land, 42.6% pastureland, 2.2% water masses, and the remaining portion is bare land (SYGM 2021). The most important agricultural lands of the basin are the Tercan, Erzurum, and Sarısu plains. The prevailing climate in the study sub-basin Karasu is continental and stands out for its harshness according to the central basin that the area belongs to, the Euphrates Basin. The basin experiences the highest rainfall during winter, while the least rainfall occurs in summer. Because of climate change, there is an anticipated rise in the frequency, severity, and spatial distribution of natural disasters, particularly those sensitive to changes in the water cycle, such as drought, across the country. In specific river

basins, a decrease in rainfall and a significant increase in temperature are observed, and consequently, a tendency for reduced water flow is observed. Research indicates that the Euphrates Basin is one of the basins witnessing a notable water deficit (Alivi et al. 2021). In this study, Tercan and Tunceli meteorological stations existing in the Karasu sub-basin were selected. Figure 1 presents an overview of the basin and the locations of the stations.

Monthly mean temperature and monthly total precipitation data measured at Tercan and Tunceli meteorological stations for 58 years from 1965 to 2022 were acquired from the General Directorate of Meteorology of Türkiye. The method used to estimate missing data is the Inverse Distance Method (IMD). This method assigns weights to neighboring stations inversely proportional to their distance from the target station. This means that distant stations are given lower weight, while closer stations are given more weight. The basic principle of this approach is the assumption that closer stations correlate better with the target station than distant stations (Mohamed Salleh et al. 2021). This study used the IMD to complete the stations' missing monthly precipitation and temperature data. While using the method, the four closest stations in the vicinity with high correlation were selected, and the missing data were tried to be completed using this station's data. Figure 2 illustrates temporal changes in the annual mean temperature and precipitation for the two selected meteorological stations in the basin.

Precipitation data analysis reveals that the long-term annual mean precipitation is approximately 430.7 mm at the Tercan station, below the country average, and 845.6 mm at the Tunceli station, above the country average. The recorded maximum and minimum annual mean precipitation was 684.2 mm in 1979 and 185.3 mm in 2022 in Tercan, 521.2 mm in 1989, and 1992.8 mm in 1967 in Tunceli. At the Tercan station, 65% of the annual precipitation occurred during autumn and spring, while approximately 22% and 13% occurred during winter and summer, respectively. At the Tunceli station, 77% of the annual precipitation occurred during winter and spring, while approximately 20% and 3% occurred during autumn and summer, respectively. Temperature data analysis indicated that the annual mean temperature is about 8.5 °C at the Tercan station and 12.9 °C at the Tunceli station. The highest and lowest recorded annual temperatures were observed at 11.4 °C in 2010 and 5.4 °C in 1992 in Tercan, 15.0 °C in 2018, and 10.1 °C in 1992 in Tunceli. Furthermore, the monthly mean temperature varies from 21.4 °C to 27.2 °C in August and from -1.7 °C to -10.6 °C in January at Tercan and Tunceli stations, respectively. In conclusion, it has been observed that the different geographical locations and hydrological conditions, such as altitude and high topographic conditions, of the study areas, significantly influence the amount of precipitation and temperature in the region. Table 1 presents geographical and

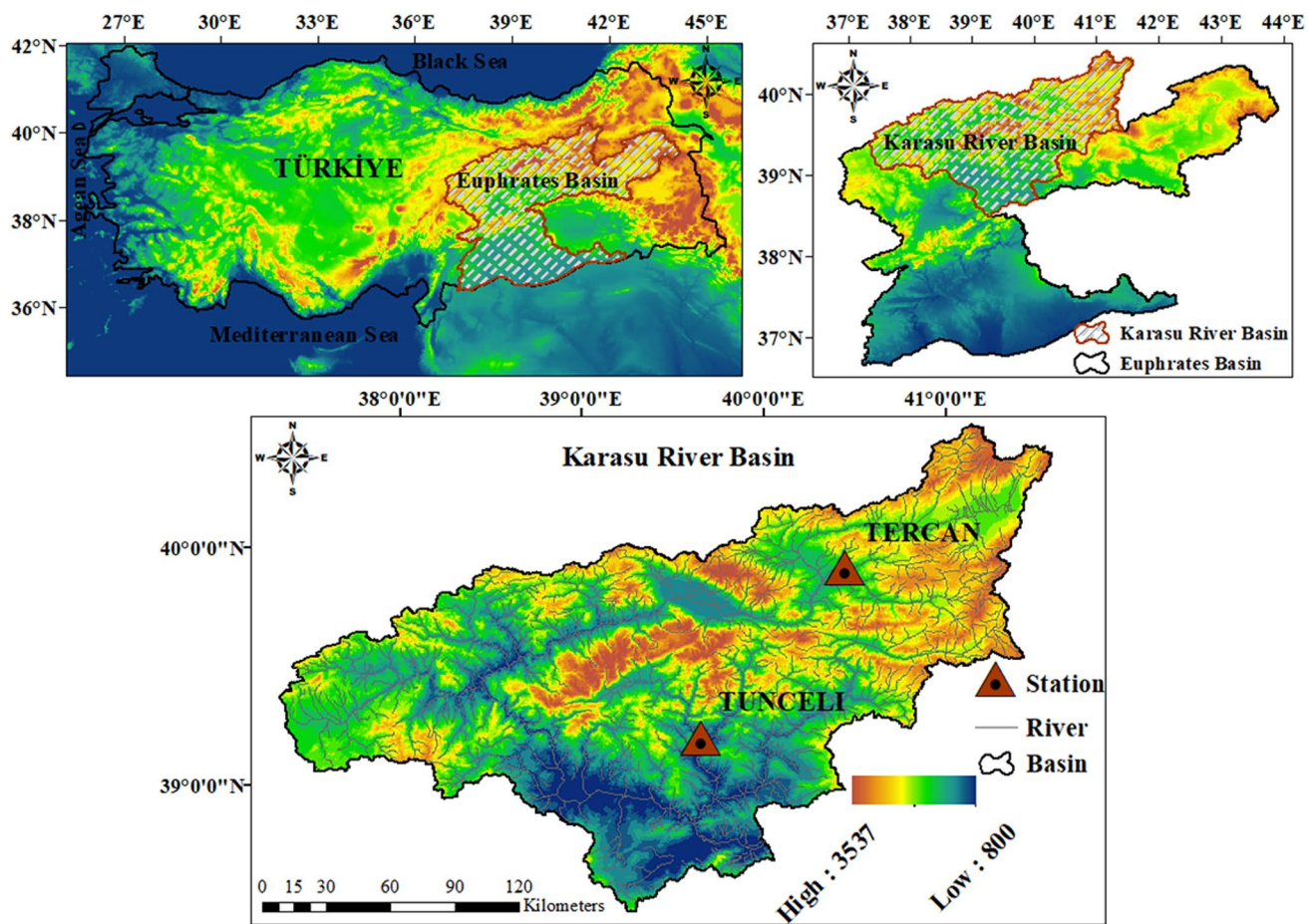


Fig. 1 General view of the study area

hydrological statistics characteristics related to monthly temperature and precipitation data between 1965 and 2022.

The selected stations have various climatic characteristics such as elevations ranging from 981 to 1,429 m., monthly mean precipitation from 35.9 to 70.5 mm., monthly mean temperature from 8.5 to 12.9 °C, standard deviation of monthly mean precipitation from 28.8 to 69.5 mm., and standard deviation of monthly mean temperature from 10.1 to 10.2 °C. On the other hand, the precipitation's C_v ranges from 101.3% to 124.7% for the selected stations, while temperature varies between 84.0% and 126.5%. The C_v coefficient of the temperature and precipitation data is considerably higher than zero. Therefore, it can be stated that available temperature and precipitation data are not homogeneously distributed around the arithmetic mean. In other words, these parameters indicate high variability in the probability density function. Only the temperature data of the Tercan station, which has the lowest C_v , shows a homogeneous distribution around the arithmetic mean. The precipitation data's C_k indicates a distribution with positive kurtosis, signifying a flatter distribution. On the other

hand, the C_k of the temperature data is negative, indicating a distribution with sharper peaks. For both stations, the precipitation data exhibit a right-skewed distribution (positive skewness), as evidenced by C_s greater than zero. Conversely, the temperature data demonstrate a left-skewed (negative skewness) distribution.

Standardized Precipitation Index (SPI)

SPI is a commonly utilized meteorological drought index. A dimensionless index standardizes precipitation data (Akturk et al. 2022; Aktürk et al. 2024). SPI reflects the precipitation deficit for a specific time frame and place. SPI can be calculated for various time scales, from short-term (e.g. one-month) to long-term (e.g. 24-months) (Svoboda et al. 2012). The index offers a comprehensive understanding of precipitation patterns by being applied to many time scales. Below-average precipitation is indicated by negative SPI readings, which may signal drought conditions. Positive SPI readings imply wetter circumstances since they show above-average precipitation

Fig. 2 Annual mean temperature and precipitation time series for Tercan and Tunceli stations

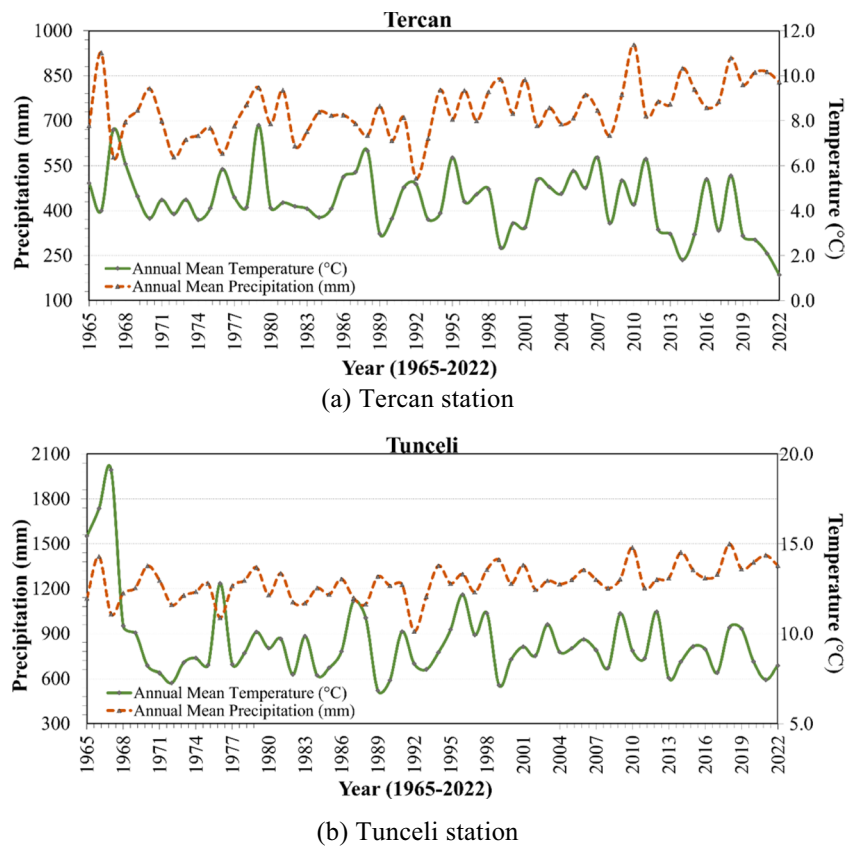


Table 1 Descriptive statistics of the selected stations

Station Name		Tercan		Tunceli	
Geographic Characteristics	Longitude (E)	40°23'26"		39°32'27"	
	Latitude (N)	39°46'36"		39°06'21"	
	Elevation (m)	1,429		981	
Hydrological Statistics (1965–2022)	Data	Precipitation (mm)	Temperature (°C)	Precipitation (mm)	Temperature (°C)
	Min	0	-16.40	0	-10.20
	Max	140.70	24.90	540.20	29.70
	Mean	35.89	8.47	70.46	12.89
	Standard Deviation (S_d)	28.77	10.09	69.55	10.19
	Kurtosis (C_k)	0.46	-1.06	3.91	-1.22
	Skewness (C_s)	0.89	-0.26	1.45	-0.08
	Coefficient of Variation (C_v), (%)	124.74	83.97	101.32	126.46

(Smakhtin and Hughes 2004). Precipitation data are assumed to have a gamma distribution for SPI purposes (McKee et al. 1993; Naumann et al. 2018). Precipitation's statistical characteristics can be described thanks to the gamma distribution (Guttman 1998; Zeybekoğlu and Aktürk 2021). SPI levels are frequently understood in terms of the length and severity of wet or dry conditions (Dabar et al. 2022). The measure helps track and contrast the intensity of droughts in various climatically

diverse areas. SPI is especially helpful in managing water resources, evaluating the effects of climate variability, and in agriculture (Tsesmelis et al. 2023). Early warning systems utilizing SPI can assist communities and policymakers in being ready for the possible effects of a drought. SPI is versatile enough to assess drought conditions for short-term agricultural impacts and long-term water resource management (Wilhite 2000; Loukas and Vasiliades 2004; Alemaw et al. 2013).

Standardized Precipitation Evapotranspiration Index (SPEI)

SPEI is a hydro climatic drought index. SPEI incorporates precipitation and possible evapotranspiration to evaluate water availability (Stagge et al. 2014). SPEI considers the atmospheric demand for moisture, in contrast to SPI, which considers precipitation (Reyniers et al. 2023). SPEI is, therefore, versatile, considering both arid and humid environments, making it suitable for various climates. SPEI estimates the difference between precipitation and potential evapotranspiration, standardizing it for various locations and periods. SPEI is particularly useful in regions where evapotranspiration significantly influences water availability. SPEI considers the combined probability distribution of possible evapotranspiration and precipitation. Log-logistic distribution is frequently employed to simulate potential evapotranspiration in SPEI computations (Kumanlioglu 2020). SPEI can capture the effects of climate variability on water availability and is sensitive to temperature fluctuations. The susceptibility of ecosystems to water availability variations can be evaluated using SPEI. The SPEI can be used to evaluate how regional water balances are affected by climate change. SPEI calculation is based on the water balance concept and incorporates the difference between precipitation (P) and weekly or monthly potential evapotranspiration (PET) as critical input parameters to determine the degree of humidity or aridity in a given area (Vicente-Serrano et al. 2010; Tirivarombo et al. 2018). This difference, denoted as D, describes the excess or shortage of water for a certain time interval, (i). The calculation of this difference can be identified mathematically as follows:

$$D_i = P_i - PET_i \tag{1}$$

PET values can be calculated using the Thornthwaite Equation (Thornthwaite 1948). The specified meteorological station must provide this equation's monthly average temperature and latitude data.

Pre-processing and Feature Selection (FS)

ML-based studies require classification, decomposition into components, and extraction of meaningless data, namely noise removal, from the observed input data. Applying inputs obtained through sub-band decomposition methods that reveal different data features can enhance the forecasting performance of ML models. The TQWT method is widely used for its advantages in separating oscillation components, tuning frequencies, representing data, and sensitivity (Selesnick 2011). The TQWT transform allows for the efficient examination of oscillating signals by adjusting the Q factor. There are three input variables in the TQWT method. These are the Q factor symbolized as Q, the redundancy factor represented by r, and the number of decomposition levels (Selesnick 2011). The TQWT subband decomposition method uses non-rational transfer functions, making it simple to utilize in the frequency domain. Q determines how often the wavelets oscillate, while r represents the frequency overlap. Multilevel decomposition can be performed by periodically applying two-band low-pass and high-pass filter banks to the lower-band and upper-band signals in TQWT (Selesnick 2011; Latifoğlu and Özger 2023).

The signal with sampling frequency f_s is decomposed into a high pass subband (HPS) signal with sampling frequency βf_s and a low pass subband (LPS) signal with sampling frequency αf_s at each level. These high and low-frequency subband signals are combined to form the final signal. The α scaling parameter for the low pass filter ($F_0(\omega)$) and the β scaling parameter for the high pass filter ($F_1(\omega)$) are used. A single-level TQWT filter bank and how the signals are combined at the output of the single-level TQWT filter bank is shown in Fig. 3:

- 1- The HPS and LPS signals are computed separately at each level.
- 2- Subsequently, the HPS and LPS signals are appropriately scaled and added together. This scaling process uses scale factors determined by parameters such as α and β .

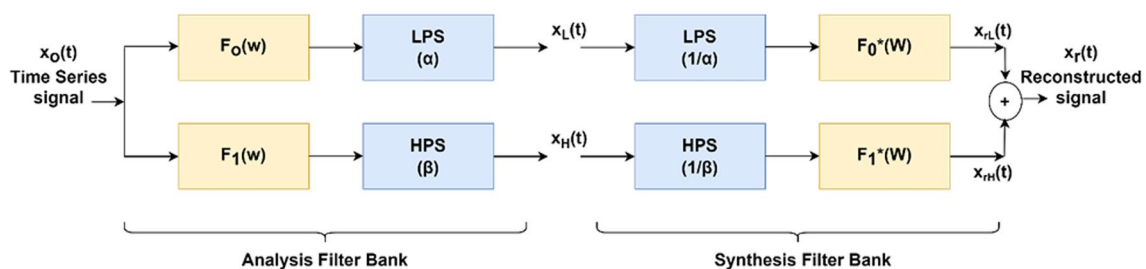


Fig. 3 Single-level TQWT filter bank

- 3- As a result, the combined signal from the HPS and LPS subbands is obtained. This combined signal constitutes the output of the TQWT and represents the processed data.

The Q and r-values decide the passband width of each filter. In addition, the wavelet is well localized in the time domain due to the oversampled filter bank feature. J + 1 subband signals are obtained for the J-Level TQWT-based decomposition approach.

The process of determining the most effective components among input features is FS, which aims to reduce the complexity of models and enhance their performance. FSRFtest is widely used among FS methods for its advantages, such as statistical reliability, ability to measure differences, and consideration of between-group discrimination (Ferrari and Yang 2015). FSRFtest evaluates the significance of each predictor individually by using an F-test. Each F-test analyzes the hypothesis of whether the response values, grouped by predictor variable values, come from the same population mean and proposes the alternative hypothesis that the population means are not all the same. A small p-value of the test statistic indicates that the corresponding predictor is significant. The output scores are $-\log(p)$; thus, a higher score value indicates the significance of the corresponding predictor.

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) are computer models designed for machine learning tasks, inspired by the structure and function of the human brain (Taye 2023). ANNs consist of interconnected nodes arranged in three layers: input, hidden, and output. During training, the network adjusts the weights assigned to each connection to learn from the data. The feedforward process transforms and propagates information through the network's layers, driven by mathematical operations and activation functions.

In this process, each hidden layer node computes its activation based on the weighted sum of inputs plus a bias term, as shown in Eq. (2):

$$h_j = \sigma \left(\sum_{i=1}^n w_{ij}x_i + b_j \right) \quad (2)$$

where, h_j is the activation of node j , w_{ij} is the weight between input node i and hidden node x_i is the input from node, b_j is the bias term, and σ is the activation function. This process continues through each layer until the output layer generates the final predictions. ANNs can model complex patterns in data using activation functions that introduce non-linearity. A key training technique is backpropagation, which updates weights based on the error gradient (Citakoglu et al. 2014). Optimization involves hyperparameter tuning, including

learning rates and batch sizes, weight regularization to prevent overfitting, and normalization to stabilize training (Haykin 1998).

Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) is a non-parametric Bayesian approach used for regression tasks, such as time series forecasting and spatial modeling (Xu and Zhang 2023). GPR models the relationship between input and output variables using a distribution of functions, providing uncertainty estimates by offering a distribution over potential functions rather than a single prediction. The GPR model is formulated as:

$$f(x) \sim GP(m(x), k(k, x')) \quad (3)$$

where $f(x)$ is the latent function, $m(x)$ is the mean function representing the expected value, and $k(k, x')$ is the kernel function capturing similarities between input points. The kernel function's hyperparameters, such as amplitude and length scales, are optimized during training. The choice of kernel significantly impacts the model's performance, with Bayesian model comparison aiding in selecting the most suitable kernel. For detailed insights into GPR, refer to (Rasmussen and Williams 2005).

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a versatile supervised learning algorithm used for both regression and classification tasks across various domains, including bioinformatics and text classification (Uncuoglu et al. 2022). SVM aims to find the hyperplane that best separates data points in a high-dimensional space.

The core objective is to maximize the margin between classes, represented as:

$$w^T x + b = 0 \quad (4)$$

where, w represents the weight vector, x denotes the input vector, and b is the bias term. The sign of the dot product between w and x determines the classification decision. SVM employs various kernel functions, like sigmoid, polynomial, linear, and radial basis functions, to handle different data distributions. Kernel trick allows implicit data transformation to higher dimensions, enhancing versatility and efficiency.

For regression, SVM uses a loss function, such as epsilon-insensitive loss, to predict continuous outputs. Preprocessing techniques like feature scaling ensure stable performance. SVM's convex optimization ensures a globally optimal solution, with support vectors setting the decision boundary. For detailed insights into SVM, refer to (Hearst et al. 1998).

The hyper parameters of the ANN, SVM, and GPR models of this study were tuned via the Bayesian Optimization Algorithm.

Performance criteria

Relative Root Mean Square Error (RRMSE) evaluates model accuracy by comparing predictions with true values, expressed as a percentage, facilitating cross-dataset comparison (Uvidia-Cabadiana et al. 2023). Lower RRMSE values indicate better performance, with 0% representing a perfect match (Bayram and Çıtakoğlu 2023). The coefficient of determination (R2) measures how much of the variance in the dependent variable is explained by independent variables, ranging from 0 (no explained variability) to 1 (perfect match), and assesses a regression model's goodness of fit (Figueiredo et al. 2011). Kling-Gupta Efficiency (KGE) offers a comprehensive performance evaluation considering correlation, bias, and variability, ranging from $-\infty$ to 1, with higher values indicating superior performance (Knoben et al. 2019; Bayram and Çıtakoğlu 2023). The Overall Index (OI) of Model Performance combines multiple criteria to evaluate a model's efficacy, where higher values indicate a more robust model (Citakoglu 2015; Bayram et al. 2016). Mean Absolute Error (MAE) measures the average error size between predicted and observed values, treating all errors equally and indicating better performance with lower values (Jierula et al. 2021; Schneider and Xhafa 2022). RRMSE, MAE, R2, KGE, and OI are calculated as follows:

$$RRMSE = \frac{\frac{1}{n} \sum_{i=1}^n \sqrt{D_{m,i} - D_{c,i}}}{\frac{1}{n} \sum_{i=1}^n D_{c,i}} \tag{5}$$

$$R^2 = \frac{[\sum_{i=1}^n (D_{c,i} - D_{cm}) \cdot (D_{m,i} - D_{mm})]^2}{\sum_{i=1}^n (D_{c,i} - D_{cm})^2 \cdot \sum_{i=1}^n (D_{m,i} - D_{mm})^2} \tag{6}$$

$$KGE = 1 - \sqrt{(R - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{7}$$

$$\beta = \frac{D_{mm}}{D_{cm}} \tag{8}$$

$$\gamma = \frac{\frac{\sigma_{mm}}{D_{mm}}}{\frac{\sigma_{cm}}{D_{cm}}} \tag{9}$$

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (D_{m,i} - D_{c,i})^2} \tag{10}$$

$$OI = \frac{1}{2} \left(2 - \frac{RMSE}{D_{c,max} - D_{c,min}} - \frac{\sum_{i=1}^n (D_{m,i} - D_{c,i})^2}{\sum_{i=1}^n (D_{c,i} - D_{cm})^2} \right) \tag{11}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |D_{m,i} - D_{c,i}| \tag{12}$$

While selecting the performance criteria mentioned above, criteria with different variables and, therefore, different units were considered as much as possible. RRMSE and MAE measure prediction errors, while R-squared assesses the explanatory power of the model. KGE provides a comprehensive performance evaluation by considering correlation, bias, and variability. OI combines multiple criteria to evaluate the overall effectiveness of the model, offering a more holistic performance analysis. This approach is thought to be vital in evaluating the results obtained from the models.

Application of the models

Model 1 (stand-alone approach, ML)

For Model 1, the SPI and SPEI data were split into training and test data, with 70% allocated for training and 30% for testing. In the study, the number of training and test data instances is given in Table 2. The training data consisted of lagged inputs ranging from 1 to 4, combined with the corresponding output (Monfort and Peña 2008; Irandoust 2019; Jiménez-Gómez and Flores-Márquez 2023). After experimenting with different lagged inputs, it was observed that utilizing up to 4 lagged inputs provided optimal results, achieving an R² value exceeding 0.99. Therefore, the analysis was focused on these 4-lagged data points to balance computational complexity and model performance.

The training data was then used to train ANN, SVM, and GPR models. Subsequently, these trained models

Table 2 The number of training and testing data

	TERCAN, TUNCELİ		
	SPI, SPEI Data		
	Train Data Number (70%)	Test Data Number (30%)	Total Data Number
SPI1	487	209	696
SPI3	485	209	694
SPI6	483	208	691
SPI9	481	207	688
SPI12	479	206	685

were applied to the test data for evaluation. The input-output configurations for Model 1 are summarized in Table 3, and the workflow is depicted in Fig. 4. Figure 4 illustrates the process of developing and applying machine learning models (Model1) using SPI and SPEI data for drought prediction. The data is divided into 70% for training and 30% for testing. In the training phase, input data (historical SPI and SPEI values) and output data (target variables) are used to train machine learning models, including ANN, SVM, and GPR. The input and output variables are defined as follows: for SPI, input data includes lagged values up to four-time steps (e.g., SPI(t-1) to SPI(t-4)), and the output data is SPI(t); for SPEI, the same lagging approach is applied. These models learn the patterns in the data to make accurate predictions. In the testing phase, the trained models are applied to test data to evaluate their performance by comparing predicted values with actual observations. This systematic approach ensures unbiased model evaluation and leverages advanced machine learning techniques to capture complex patterns in the data.

Model 2 (tribrid approach, TQWT- FSRFtest-ML)

In Model 2, similar to Model 1, the SPI and SPEI data were divided into training and test data, with 70% allocated for training and 30% for testing. However, a novel approach involving the TQWT method was employed for feature extraction (FE). The data were decomposed into six sub-bands using the TQWT method, and the obtained data was used as features for the prediction considering lagged inputs ranging from 1 to 4 (Latifoğlu 2022). To enhance the model's efficiency and effectiveness, FS was performed using the FSRFtest method to retain features contributing to 95% or more of the model's performance. A single ML algorithm was trained using the extracted features to estimate the data comprehensively. The input-output configurations for Model 2 are presented in Table 4, with an illustrative example for SPI data. A similar approach was adopted for predicting other SPIs and SPEIs, albeit not included in the table due to space constraints.

The selection of the TQWT method and FSRFtest for pre-processing and feature selection was driven by their specific

Table 3 Input and output variables of Model 1

SPI Input Data	SPI Output Data
1 Lagged Model	SPI(t)
2 Lagged Model	SPI(t)
3 Lagged Model	SPI(t)
4 Lagged Model	SPI(t)
SPEI Input Data	SPEI Output Data
1 Lagged Model	SPEI(t)
2 Lagged Model	SPEI(t)
3 Lagged Model	SPEI(t)
4 Lagged Model	SPEI(t)

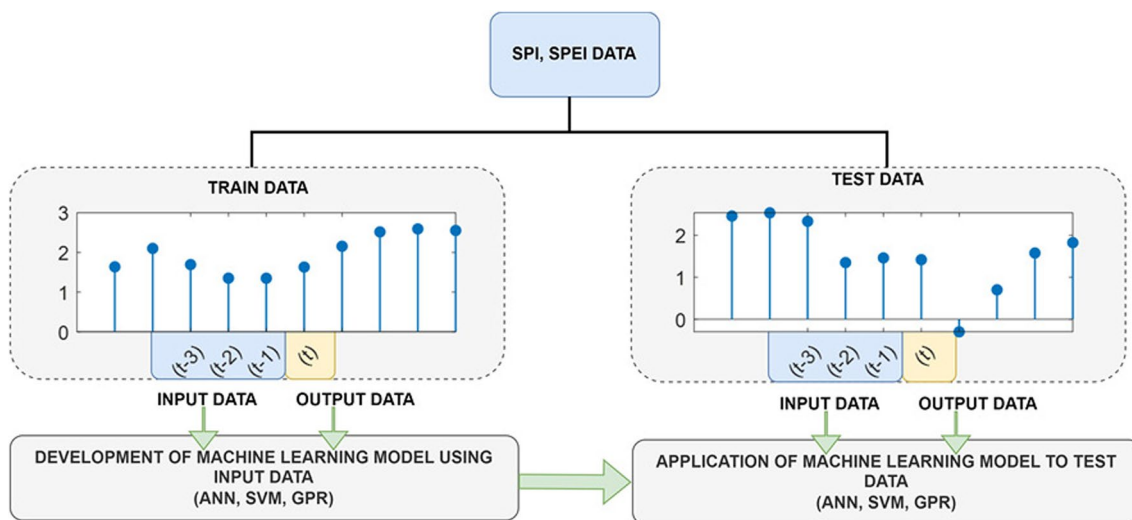


Fig. 4 Flowchart for Model 1

Table 4 Input and output variables of Model 2 for Tunceli Station, SPI1 data

SPI Input Data		SPI Output Data	
Model	Subband Data	Selected Features	
1 Lagged Model	TQWT1(t-1), TQWT2(t-1), TQWT3(t-1), TQWT4(t-1), TQWT5(t-1), TQWT6(t-1)	TQWT6(t-1), TQWT1(t-1), TQWT5(t-1), TQWT4(t-1)	SPI(t)
2 Lagged Model	TQWT1(t-1), TQWT1(t-2), TQWT2(t-1), TQWT2(t-2), TQWT3(t-1), TQWT3(t-2), TQWT4(t-1), TQWT4(t-2), TQWT5(t-1), TQWT5(t-2), TQWT6(t-1), TQWT6(t-2)	TQWT6(t-2), TQWT1(t-2), TQWT2(t-1), TQWT6(t-1), TQWT3(t-1), TQWT5(t-2), TQWT4(t-2)	SPI(t)
3 Lagged Model	TQWT1(t-1), TQWT1(t-2), TQWT1(t-3), TQWT2(t-1), TQWT2(t-2), TQWT2(t-3), TQWT3(t-1), TQWT3(t-2), TQWT3(t-3), TQWT4(t-1), TQWT4(t-2), TQWT4(t-3), TQWT5(t-1), TQWT5(t-2), TQWT5(t-3), TQWT6(t-1), TQWT6(t-2), TQWT6(t-3)	TQWT6(t-3), TQWT1(t-3), TQWT2(t-2), TQWT6(t-2), TQWT4(t-1), TQWT3(t-1), TQWT3(t-2), TQWT5(t-3), TQWT6(t-1), TQWT4(t-3), TQWT5(t-1)	SPI(t)
4 Lagged Model	TQWT1(t-1), TQWT1(t-2), TQWT1(t-3), TQWT1(t-4), TQWT2(t-1), TQWT2(t-2), TQWT2(t-3), TQWT2(t-4), TQWT3(t-1), TQWT3(t-2), TQWT3(t-3), TQWT3(t-4), TQWT4(t-1), TQWT4(t-2), TQWT4(t-3), TQWT4(t-4), TQWT5(t-1), TQWT5(t-2), TQWT5(t-3), TQWT5(t-4), TQWT6(t-1), TQWT6(t-2), TQWT6(t-3), TQWT6(t-4)	TQWT6(t-4), TQWT2(t-3), TQWT1(t-4), TQWT6(t-3), TQWT2(t-1), TQWT4(t-2), TQWT3(t-2), TQWT3(t-3), TQWT5(t-1), TQWT5(t-4), TQWT6(t-2), TQWT4(t-1), TQWT4(t-4), TQWT5(t-2), TQWT6(t-1)	SPI(t)

advantages in handling the characteristics of SPI/SPEI data and enhancing model performance in the proposed study.

The TQWT was chosen for its ability to effectively decompose time-series data into subbands with different frequency components. This method is particularly advantageous for analyzing SPI/SPEI data, which often exhibit non-stationary and multi-scale characteristics. TQWT allows for the extraction of meaningful features from various frequency bands, capturing both short-term fluctuations and long-term trends in the data. This decomposition aids in better understanding and modeling the underlying patterns, leading to improved prediction accuracy.

The Univariate Feature Ranking Using F-Tests (FSRFtest) was selected for feature selection due to its simplicity and effectiveness in identifying the most relevant features. FSRFtest evaluates the significance of each feature individually by comparing the variance between different groups, thus ranking features based on their importance. This method is computationally efficient and helps in reducing the dimensionality of the input data by selecting only the most informative features. By focusing on these key features, the models can be trained more effectively, resulting in enhanced performance and reduced computational complexity.

The combination of TQWT for data decomposition and FSRFtest for feature selection leverages the strengths of both

methods to address the specific challenges posed by SPI/SPEI data, ultimately leading to more accurate and efficient machine learning models.

If we were to explain an example from the provided data in Table 4, for instance, in the 2 lagged model, 12 features were extracted from the 6 sub-bands and 2 lagged inputs, resulting in 12 features ($6 \times 2 = 12$). Of these 12 features, 7 were selected for SPI1 data using F-Test. These selected features are: TQWT6(t-2), TQWT1(t-2), TQWT2(t-1), TQWT6(t-1), TQWT3(t-1), TQWT5(t-2), and TQWT4(t-2) in order of its efficiency. The lengths of the input and output signals in Model 2 are provided in Table 5, and the workflow is depicted in Fig. 15 in the Appendix.

Model 3 (hybrid approach, TQWT-ML)

In Model 3, the SPI and SPEI data were again divided into training and test data, with 70% allocated for training and 30% for testing. Utilizing the TQWT method, the data were decomposed into six sub-bands. Unlike Model 2, where each sub-band was individually predicted and aggregated, Model 3 involved predicting each sub-band separately using ML methods. The forecasted sub-band signals were then summed to obtain the final prediction.

The input-output configurations for Model 3 are outlined in Table 6 and the workflow is depicted in Fig. 16 in the

Table 5 The lengths of the input signal and the output signal in Model 2

	Before FE	After FE	Length of Output Signal	Before FE	After FE	Length of Output Signal
TRAINING DATA						
	One time lag			Two time lag		
SPI1	6×486	4×486	1×486	12×485	8×485	1×485
SPI3	6×484	4×484	1×484	12×483	7×483	1×483
SPI6	6×482	2×482	1×482	12×481	4×481	1×481
SPI9	6×480	2×480	1×480	12×479	3×479	1×479
SPI12	6×478	1×478	1×478	12×477	2×477	1×477
	Three-time lag			Four-time lag		
SPI1	18×484	12×484	1×484	24×483	15×483	1×483
SPI3	18×482	11×482	1×482	24×481	13×481	1×481
SPI6	18×480	6×480	1×480	24×479	10×479	1×479
SPI9	18×478	6×478	1×478	24×477	9×477	1×477
SPI12	18×476	3×476	1×476	24×475	4×475	1×475
TESTING DATA						
	One time lag			Two time lag		
SPI1	6×208	4×208	1×208	12×207	8×207	1×207
SPI3	6×208	4×208	1×208	12×207	7×207	1×207
SPI6	6×207	2×207	1×207	12×206	4×206	1×206
SPI9	6×206	2×206	1×206	12×205	3×205	1×205
SPI12	6×205	1×205	1×205	12×204	2×204	1×204
	Three time lag			Four time lag		
SPI1	18×206	12×206	1×206	24×205	15×205	1×205
SPI3	18×206	11×206	1×206	24×205	13×205	1×205
SPI6	18×205	6×205	1×205	24×204	10×204	1×204
SPI9	18×204	6×204	1×204	24×203	9×203	1×203
SPI12	18×203	3×203	1×203	24×202	4×202	1×202

Appendix, emphasizing SPI and SPEI data. Like the previous models, lagged inputs ranging from 1 to 4 were considered for prediction. The research flowchart for this study is depicted in Fig. 5. Pseudo code for the proposed methods is given in Tables 12 and 13 in the Appendix.

Both models utilize the TQWT (Tunable Q-factor Wavelet Transform) method to decompose SPI/SPEI data into subbands. The tribrid model focuses on feature selection from these subbands, using the selected features to train a single machine learning model (ANN, SVM, GPR) for direct prediction. Conversely, the hybrid model involves separately modeling each subband with machine learning models and then summing the individual subband predictions to form the final forecasted data. The key difference lies in the approach: tribrid models use feature selection to develop one comprehensive model, simplifying the process, while hybrid models entail training multiple models for each subband and aggregating their outputs, capturing detailed subband-specific patterns.

The specific advantages of tribrid models in reducing computational load and shortening calculation time compared to hybrid models are as follows:

Firstly, Tribrid model in this study employs feature selection from the decomposed subbands to identify the most relevant features for model training. This reduces the dimensionality of the input data, leading to a more efficient training process and faster computation times.

Secondly, only one machine learning model (ANN, SVM, GPR) is trained using the selected features. This contrasts with hybrid models, which require separate models to be trained for each subband. Training a single model significantly reduces the overall computational load and time required for model development.

Thirdly, during the testing phase, tribrid model uses the single trained model to make predictions directly, while hybrid model in this study must aggregate predictions from multiple subband-specific models. This aggregation step in hybrid models adds additional computational complexity and time.

Finally, proposed tribrid model handle fewer data points during training and testing due to the feature selection process. This reduction in data handling contributes to faster calculations and lower computational requirements.

Table 6 Input and output variables of Model 3

SPEI Input Data		SPEI Output Data
1 Lagged Model	SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
2 Lagged Model	SPEI_TQWT1(t-2), SPEI_TQWT2(t-2), SPEI_TQWT3(t-2), SPEI_TQWT4(t-2), SPEI_TQWT5(t-2), SPEI_TQWT6(t-2), SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
3 Lagged Model	SPEI_TQWT1(t-3), SPEI_TQWT2(t-3), SPEI_TQWT3(t-3), SPEI_TQWT4(t-3), SPEI_TQWT5(t-3), SPEI_TQWT6(t-3), SPEI_TQWT1(t-2), SPEI_TQWT2(t-2), SPEI_TQWT3(t-2), SPEI_TQWT4(t-2), SPEI_TQWT5(t-2), SPEI_TQWT6(t-2), SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
4 Lagged Model	SPEI_TQWT1(t-4), SPEI_TQWT2(t-4), SPEI_TQWT3(t-4), SPEI_TQWT4(t-4), SPEI_TQWT5(t-4), SPEI_TQWT6(t-4), SPEI_TQWT1(t-3), SPEI_TQWT2(t-3), SPEI_TQWT3(t-3), SPEI_TQWT4(t-3), SPEI_TQWT5(t-3), SPEI_TQWT6(t-3), SPEI_TQWT1(t-2), SPEI_TQWT2(t-2), SPEI_TQWT3(t-2), SPEI_TQWT4(t-2), SPEI_TQWT5(t-2), SPEI_TQWT6(t-2), SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
SPEI Input Data		SPEI Output Data
1 Lagged Model	SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
2 Lagged Model	SPEI_TQWT1(t-2), SPEI_TQWT2(t-2), SPEI_TQWT3(t-2), SPEI_TQWT4(t-2), SPEI_TQWT5(t-2), SPEI_TQWT6(t-2), SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
3 Lagged Model	SPEI_TQWT1(t-3), SPEI_TQWT2(t-3), SPEI_TQWT3(t-3), SPEI_TQWT4(t-3), SPEI_TQWT5(t-3), SPEI_TQWT6(t-3), SPEI_TQWT1(t-2), SPEI_TQWT2(t-2), SPEI_TQWT3(t-2), SPEI_TQWT4(t-2), SPEI_TQWT5(t-2), SPEI_TQWT6(t-2), SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)
4 Lagged Model	SPEI_TQWT1(t-4), SPEI_TQWT2(t-4), SPEI_TQWT3(t-4), SPEI_TQWT4(t-4), SPEI_TQWT5(t-4), SPEI_TQWT6(t-4), SPEI_TQWT1(t-3), SPEI_TQWT2(t-3), SPEI_TQWT3(t-3), SPEI_TQWT4(t-3), SPEI_TQWT5(t-3), SPEI_TQWT6(t-3), SPEI_TQWT1(t-2), SPEI_TQWT2(t-2), SPEI_TQWT3(t-2), SPEI_TQWT4(t-2), SPEI_TQWT5(t-2), SPEI_TQWT6(t-2), SPEI_TQWT1(t-1), SPEI_TQWT2(t-1), SPEI_TQWT3(t-1), SPEI_TQWT4(t-1), SPEI_TQWT5(t-1), SPEI_TQWT6(t-1)	SPEI(t)

Findings

Evaluation of drought analyses

During this phase of the study, meteorological drought events that occurred in the Karasu River basin between 1965 and 2022 were examined using SPI and SPEI values obtained from the data of the Tercan and Tunceli stations. This study focuses on moderate, severe, and extreme droughts, and the classification of droughts for SPI and SPEI is considered as extreme drought (≤ -2), severe drought ($-2 \sim -1.5$), and moderate drought ($-1.5 \sim -1$). This section of the study aims to monitor and assess the development of these meteorological drought events at one-, three-, six-, nine- and twelve-month time scales.

Temporal variation of meteorological drought events

SPI and SPEI outputs were utilized separately to monitor and describe the development of temporal variations of meteorological drought events over various time scales at selected stations within the basin. The temporal evolutions

of dry and wet periods computed at one-, three-, six-, nine- and twelve-month time scales for Tunceli and Tercan stations are presented in Figs. 6 and 7. According to the SPI and SPEI outputs obtained from two selected stations, with the increase in the time scale, the drought trend has generally become more evident, and also, a decrease in the frequency of dry periods and an increase in drought durations and magnitudes have been observed relatively.

As seen in Figs. 6 and 7, for both stations, the wet and dry periods, which were relatively unclear at one- and three-month time scales, started to become clearer starting from the six-month time scale. In addition, it was noteworthy that since the 2000s, there has been a significant increase in drought durations and magnitudes, especially in the Tercan station rather than the Tunceli station, and that this increase has progressively become more noticeable at six-month and longer time scales. When SPI and SPEI values were compared at the selected time scales, although the temporal development of the two indices is similar, there were slight differences in the characteristic features of drought periods (frequency, magnitude, and intensity). The differences

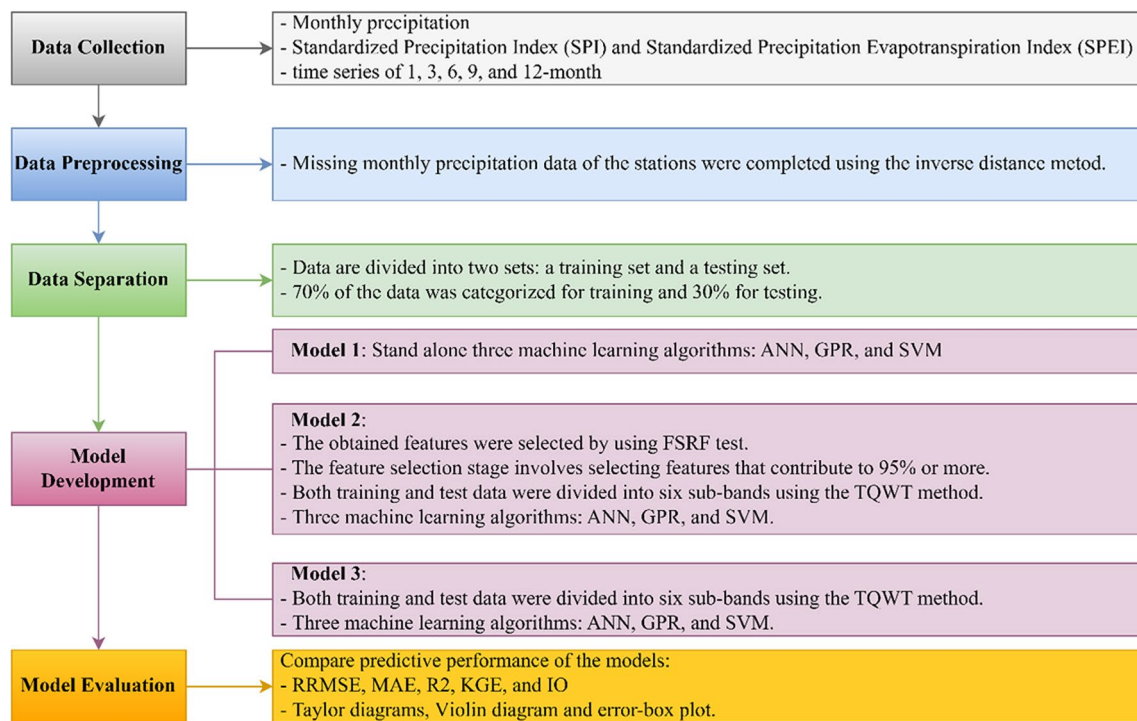


Fig. 5 Research flowchart of the study

in the fluctuation value and continuity properties of the mentioned drought indices, representing the drought features, decreased as the time scale increased.

Recent research on drought in Türkiye reveals that significant drought periods occurred in the early 1970s and the 1980s, 1990s and 2000s (Türkeş and Tatlı 2009; Akbaş 2014; Kurnaz 2014; Kumanlioglu 2020; Patel 2021), and at the same time, projection values propose that, under the assumption of climate change scenarios, the severity and frequency of droughts will increase across the country (Sen et al. 2012; Afshar et al. 2020; Danandeh Mehr et al. 2020; Soylu Pekpostalci et al. 2023). Despite their differences, SPI and SPEI were generally successful in accurately identifying the same primary drought years.

As a result of the evaluation of SPI and SPEI analysis performed on the twelve-month time scale, it was observed that significant drought events occurred at Tercan station between 1970–1971, 1983, 1985, 1989–1990, 1993–1994, 1999–2001, 2008, 2012–2018, 2020–2022 and at Tunceli station between 1970–1973, 1974–1975, 1982–1983, 1984–1986, 1989–1991, 1994, 1999–2001, 2007–2008, 2013–2014, 2016–2018, 2020–2022. According to the results obtained from both stations, it was concluded that the region was seriously affected by the droughts occurring throughout Türkiye.

Frequency analysis of drought indices

Frequency analysis results of SPI and SPEI values obtained at the selected time scales for Tercan and Tunceli stations were computed, and the relative frequency results are given in Table 7 and Fig. 17 in the Appendix. It was only classified based on the dry category. Based on all frequencies, drought in Tercan ranged from 13.9% to 15.9% for SPI and from 15.7% to 19.7% for SPEI. Similarly, in Tunceli, drought varied from 11.4% to 12.8% for SPI and from 13.4% to 17.2% for SPEI. Upon detailed examination, as the results of SPI analysis at Tercan station, the maximum relative frequency in the moderate drought class was obtained as 9.5% at the 12-month time scale, while the minimum relative frequency in the extreme drought class was 2.6% at the 1-month time scale. As the results of SPEI analysis at Tercan station, the maximum relative frequency in the moderate drought class was 13.4% at the twelve-month time scale, while the minimum relative frequency was 0.9% at the 1-month time scale.

At the same time, at Tunceli station, the results of SPI analysis demonstrated that the maximum relative frequency was obtained as 8.7% at the 3-month time scale in the moderate drought class, while the minimum relative frequency was 0.1% at the 12-month time scale in the extreme drought class. Furthermore, at the Tunceli station,

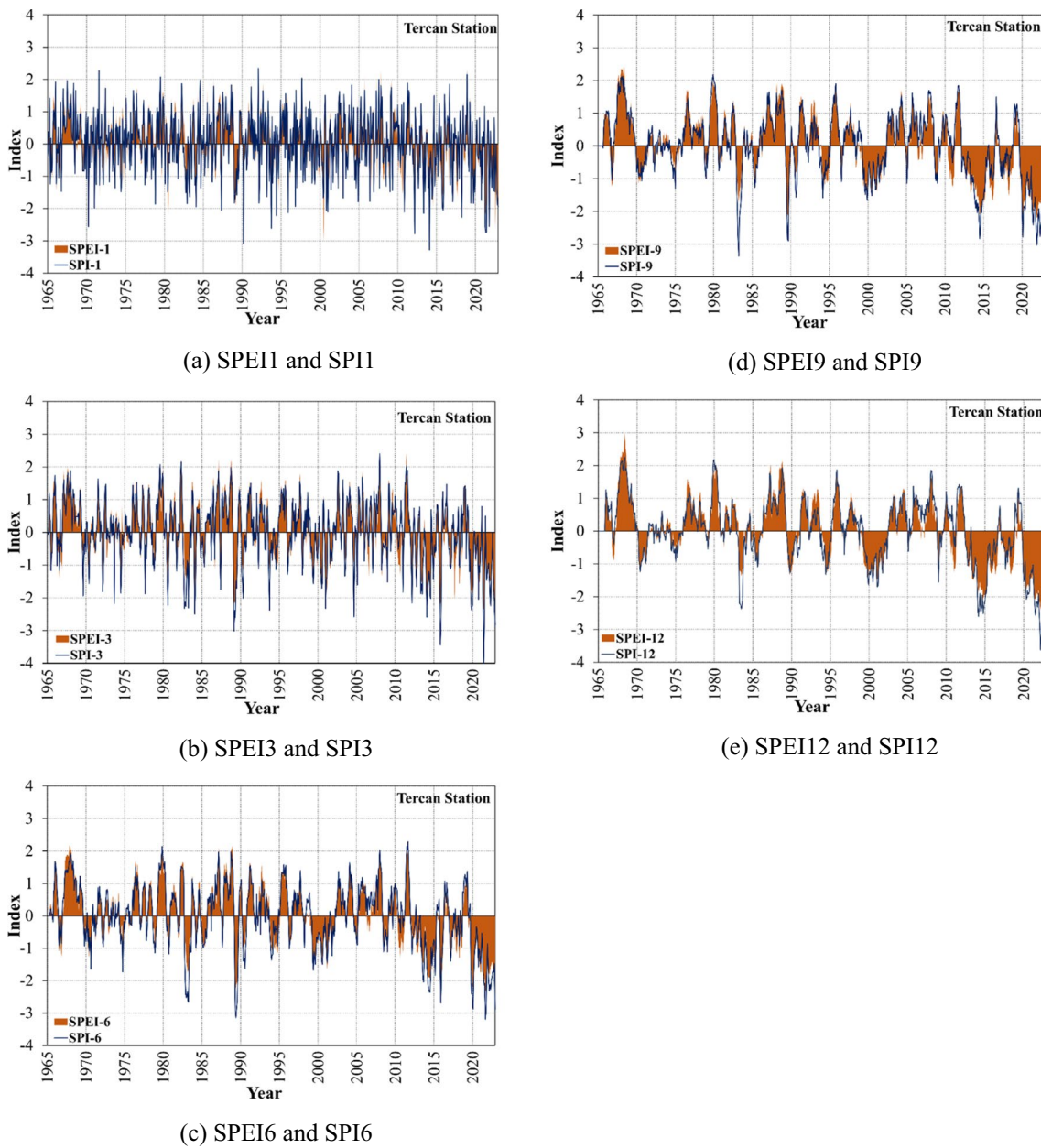


Fig. 6 The temporal evolution of the SPEIs and SPIs at 1-, 3-, 6-, 9- and 12-month timescales in the Tercan stations between 1965–2022

the results of the SPEI analysis revealed that the maximum relative frequency was obtained as 11.5% at the 12-month time scale in the moderate drought class, while the minimum relative frequency was 1.0% at the 1-month time scale in the extreme drought class. When the relative frequency percentages of all drought classes were compared, the relative frequency in the moderate drought class was more dominant than the relative frequencies in the severe and extreme drought classes. Findings obtained from the selected station showed that moderate drought events were mainly experienced in the basin.

Correlation of drought indices

The correlation coefficients between SPI and SPEI values of meteorological stations at 1-, 3-, 6-, 9- and 12-month time scales were computed, and the relationships between drought indices at different time scales were examined. The distribution of correlation coefficients calculated for this purpose was visualized in Fig. 8. The drought indices for both stations exhibited an extremely high correlation at the same time scale, indicating that the selected indices have a strong relationship. For example, the correlation

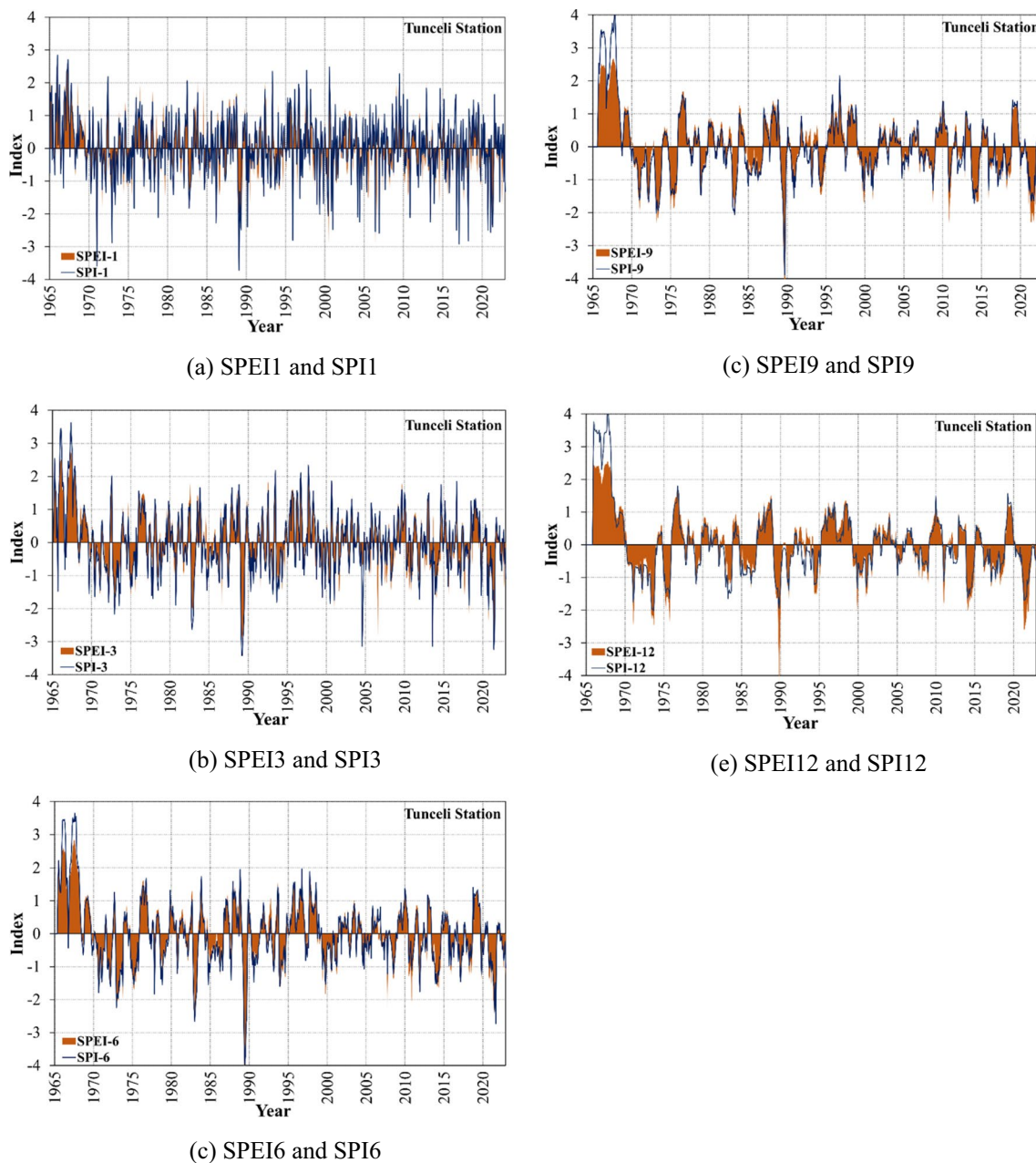


Fig. 7 The temporal evolution of the SPIs and SPEIs at 1-, 3-, 6-, 9- and 12-month timescales in the Tunceli stations between 1965–2022

value of SPI-1 with SPEI-1 was obtained as $r=0.92$ at Tercan station, and the correlation value of SPI-12 with SPEI-12 was obtained as $r=0.94$ at Tunceli station. When comparing the same time scales, the highest correlations at Tercan and Tunceli stations ($r=0.95$ and $r=0.96$, respectively) were observed at a 6-month time scale, and the lowest correlations ($r=0.92$ and $r=0.87$, respectively) were observed at 1-month time scale. All this means that these indices consistently assess drought conditions over specific periods.

Performance of machine learning models

This study forecasts drought data from Tercan and Tunceli meteorological stations using stand-alone, hybrid, and tribrid different model (version) structures. To evaluate the performance results of these models, (i) classical performance criteria, (ii) scatter and trajectory plots, (iii) Taylor diagrams, and (iv) violin and error-box plots were utilized. In the final stage, the Kruskal–Wallis test investigated the relationship between the calculated and forecasted drought values' average.

Table 7 Relative frequency of meteorological drought events in the Tercan and Tunceli stations

			Time Scales				
		Index	1	3	6	9	12
TERCAN	Moderate Drought	SPI	9.5%	7.5%	6.1%	6.1%	6.1%
		SPEI	13.4%	9.4%	9.1%	11.0%	13.3%
	Severe Drought	SPI	3.6%	4.3%	3.6%	3.8%	2.9%
		SPEI	4.3%	4.9%	5.2%	6.1%	5.1%
	Extreme Drought	SPI	2.6%	4.0%	4.2%	4.5%	5.1%
		SPEI	0.9%	1.4%	1.4%	1.3%	1.3%
TUNCELI	Moderate Drought	SPI	6.5%	7.3%	8.0%	8.7%	8.5%
		SPEI	11.5%	10.5%	9.7%	5.7%	6.6%
	Severe Drought	SPI	1.6%	3.3%	3.0%	2.9%	2.8%
		SPEI	4.7%	2.7%	3.2%	5.5%	4.5%
	Extreme Drought	SPI	3.6%	2.2%	1.6%	0.9%	0.1%
		SPEI	1.0%	2.2%	2.5%	2.2%	3.1%

Fig. 8 Correlations coefficients between drought indices at different time scales for Tercan and Tunceli stations



The MAE and R^2 criteria for the forecast results obtained for each lag in these three different models are provided in Figs. 9 and 10. The ML method that outperformed each version, depending on lag numbers, was also presented for SPI and SPEI indices of the Tercan and

Tunceli stations in these figures. As seen in Fig. 9, for the SPEI1 value of the Tercan station, version 1 with 4 lags via GPR method (V1-4GPR), version 2 with 3 lags via GPR method (V2-3GPR), and version 3 with 4 lags via GPR method (V3-4GPR) were selected as the most

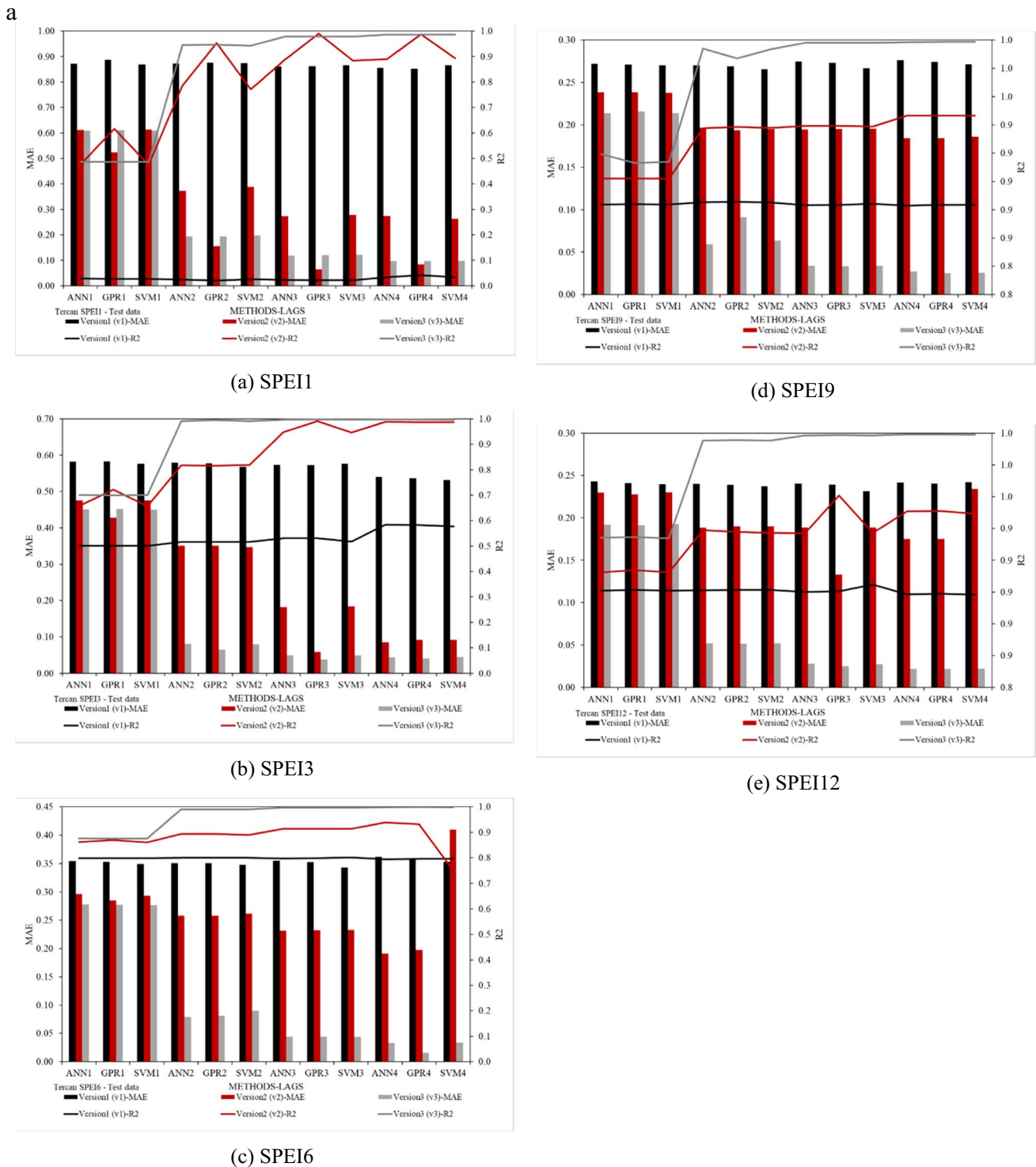


Fig. 9 **a** All combinations of three different versions of Tercan Station's SPEI indexes for test data. **b** All combinations of three different versions of Tercan Station's SPI indexes for test data

successful combinations. As seen in Fig. 10, on the other hand, for the SPEI1 value of the Tunceli station, versions 1, 2, and 3 with 4 lags via the ANN method (V1-4ANN, V2-4ANN, V3-4ANN) were selected as the

most successful combinations. The most successful combinations for the remaining time criteria of the SPEI index, besides all-time scales of the SPI index, are indicated in Tables 8 and 9. The optimum hyperparameters of each

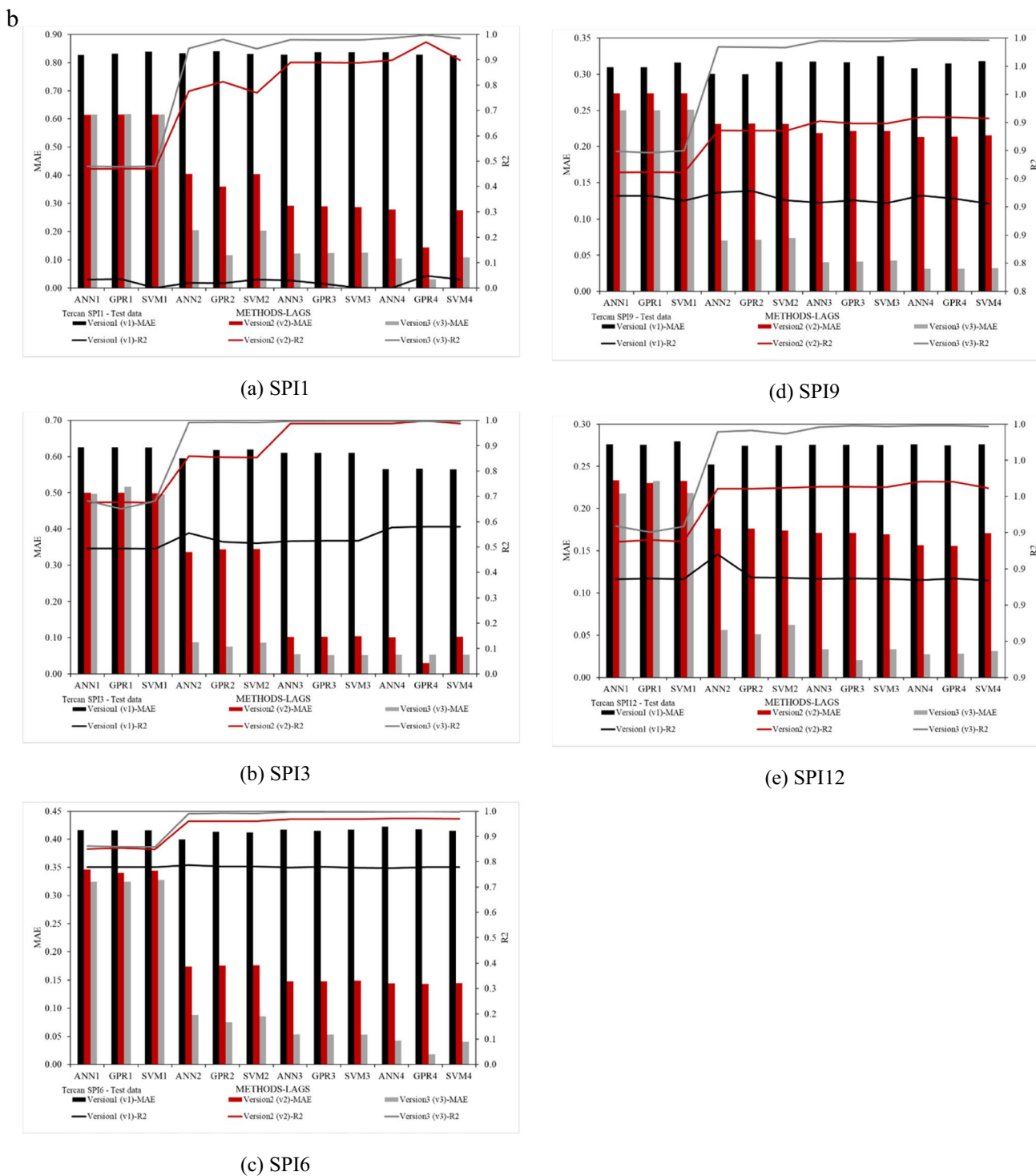


Fig. 9 (continued)

best model for the SPI and SPEI indices of Tercan and Tunceli stations are given in Tables 14, 15, 16 and 17 in the Appendix. Separate performance criteria are given for training and testing data. Performance criteria

for training data are given in Tables 18 and 19 in the Appendix.

Classical performance criteria were utilized in the first stage of evaluating the results. For Tercan and Tunceli

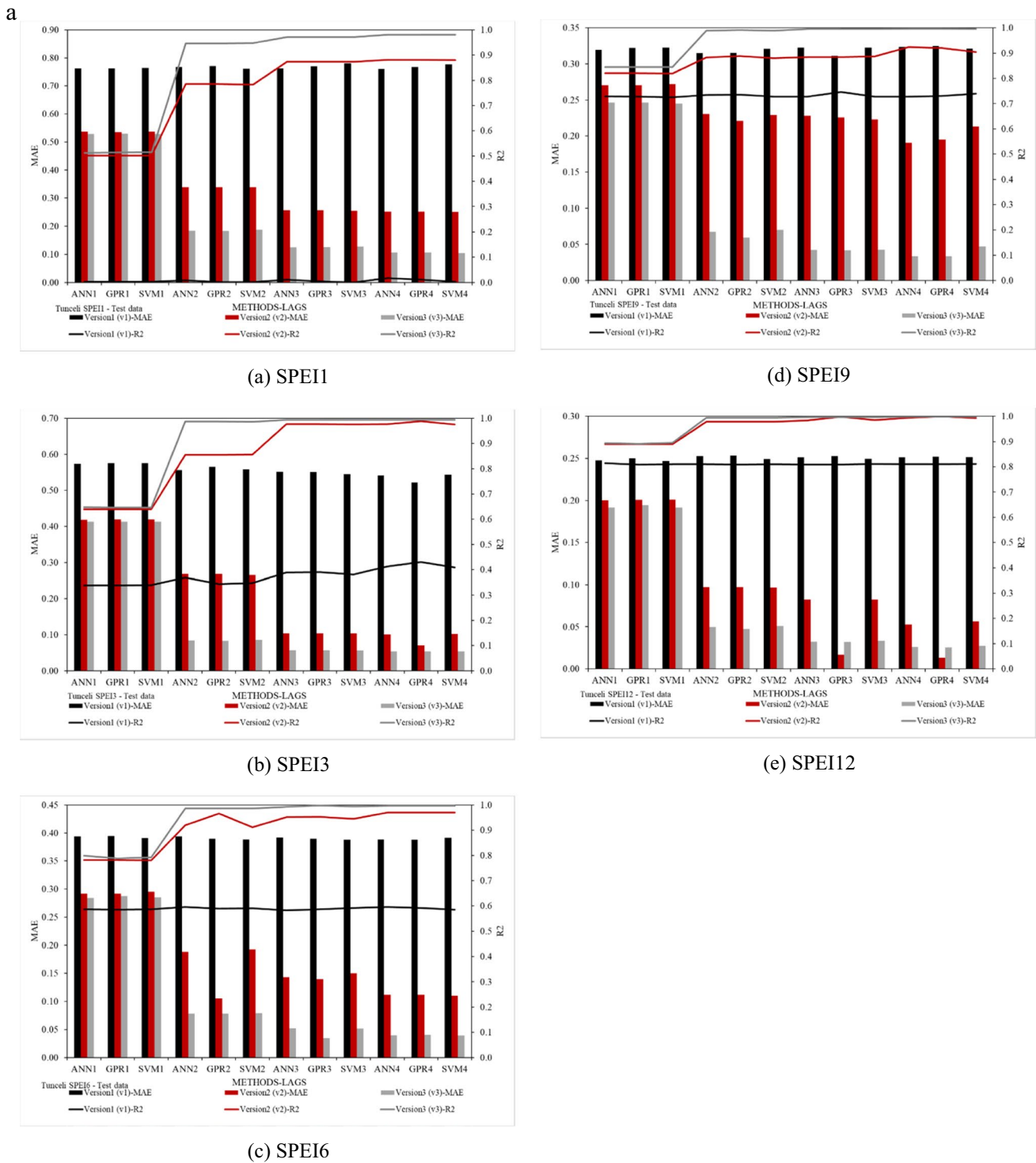


Fig. 10 **a** All combinations of three different versions of Tunceli Station's SPEI indexes for test data. **b** All combinations of three different versions of Tunceli Station's SPI indexes for test data

stations, RRMSE, MAE, R^2 , KGE, and OI criteria for forecasting SPI and SPEI indices at all time scales were provided in Tables 8 and 9. As seen in Table 8, for Tercan Station, the combination V2-3GPR for SPEI1 index, V3-3GPR for

SPEI3 index, V3-4GPR for SPEI6 index, V3-4GPR for SPEI9 index, and V3-4ANN for SPEI12 index provided the lowest RRMSE and MAE; and the highest R^2 , KGE, and OI, making them as the most successful combinations. The

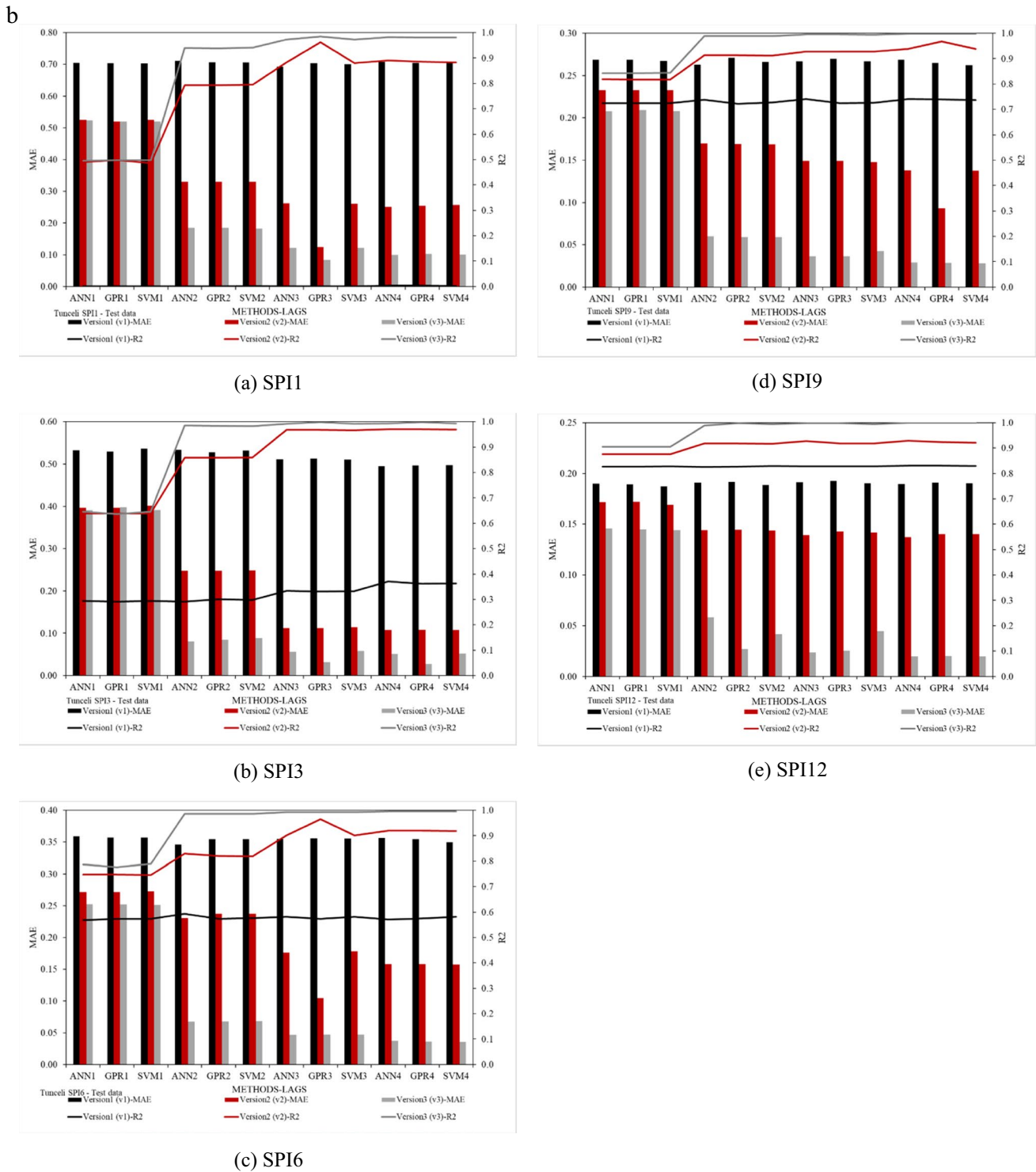


Fig. 10 (continued)

combination V3-4GPR for the SPI1 index, V2-4GPR for the SPI3 index, V3-4GPR for the SPI6 index, V3-4ANN for the SPI9 index, and V3-3GPR for the SPI12 index were identified as the most successful models. The RRMSE values for SPEI1, SPEI3, SPEI6, SPEI9, and SPEI12 indices were

calculated as 36.40, 11.34, 5.31, 6.07, and 5.01, respectively. The RRMSE values for SPI1, SPI3, SPI6, SPI9, and SPI12 indices were also calculated as 24.90, 17.08, 9.93, 9.94, and 7.39, respectively. According to (Bayram and Çıtakoğlu 2023); RRMSE values greater than 30% are considered poor,

Table 8 The most successful methods of three different versions of Tercan Station and the performance criteria of these methods for test data

	V1	V2	V3		V1	V2	V3	
SPEI1	4GPR	3GPR	4GPR	SPEI1	4GPR	4GPR	4GPR	RRMSE
	368.02	36.40	44.22		658.56	116.2	24.90	MAE
	0.853	0.065	0.096		0.83	0.14	0.03	R ²
	0.042	0.99	0.99		0.048	0.97	0.999	KGE
	-0.191	0.98	0.967		-9.575	0.950	0.986	OI
SPEI3	4SVM	3GPR	3GPR	SPEI3	4SVM	4GPR	3GPR	RRMSE
	161.67	23.57	11.34		264.46	17.08	24.02	MAE
	0.53	0.058	0.04		0.56	0.03	0.05	R ²
	0.578	0.991	0.998		0.580	0.998	0.997	KGE
	0.705	0.975	0.992		0.501	0.995	0.992	OI
SPEI6	3SVM	4ANN	4GPR	SPEI6	2ANN	4GPR	4GPR	RRMSE
	89.79	49.42	5.31		142.52	52.30	9.93	MAE
	0.34	0.19	0.02		0.40	0.14	0.02	R ²
	0.801	0.939	0.999		0.784	0.971	0.999	KGE
	0.853	0.952	0.993		0.807	0.981	0.996	OI
SPEI9	2SVM	4GPR	4GPR	SPEI9	2GPR	4ANN	4ANN	RRMSE
	61.34	41.54	6.08		91.67	65.65	9.94	MAE
	0.27	0.18	0.03		0.30	0.21	0.03	R ²
	0.884	0.946	0.999		0.891	0.944	0.999	KGE
	0.923	0.966	0.995		0.910	0.961	0.996	OI
SPEI12	3SVM	3GPR	4ANN	SPEI12	2ANN	4ANN	4GPR	RRMSE
	51.720	33.068	5.010		78.17	47.30	7.39	MAE
	0.232	0.133	0.022		0.25	0.16	0.02	R ²
	0.904	0.960	0.999		0.914	0.968	0.999	KGE
	0.926	0.975	0.997		0.909	0.978	0.995	OI
	0.942	0.967	0.996	0.940	0.970	0.997		

between 20–30% are considered fair, between 10–20% are considered good, and less than 10% are considered excellent.

Therefore, for the RRMSE values of Tercan station, the SPEI1 index is considered poor, the SPI1 index is fair, the SPEI3 and SPI3 indices are good, and the SPEI6, SPEI9, SPEI12, SPI6, SPI9, and SPI12 indices are considered excellent forecasts. Moreover, a perfect forecast is made when the R² criterion is greater than 0.8 and the KGE criterion is greater than 0.7 (Bayram and Çıtakoğlu 2023). Thus, it can be emphasized that the forecast results for the SPI and SPEI indices at all time scales for the Tercan station are perfect according to the R² and KGE criteria. While the RRMSE was obtained fair for the SPEI1, SPEI3, SPI1, and SPI3 indices of the Tercan station, the R² and KGE are interpreted as excellent. On the other hand, as the time scale increases, the RRMSE and R² and KGE support each other.

As seen in Table 9 for Tunceli Station, based on the classical performance criteria, the combination V3-4ANN for the SPEI1 index, V3-4SVM for the SPEI3 index, V3-3GPR

for the SPEI6 index, V3-4ANN for the SPEI9 index, and V2-4GPR for the SPEI12 index were identified as the most successful models. Based on the classical performance criteria, the combination V3-3GPR for the SPI1 index, V3-4GPR for the SPI3 index, and V3-4SVM for the SPI6, SPI9, and SPI12 indices were determined as the most successful models. The RRMSE values for SPEI1, SPEI3, SPEI6, SPEI9, and SPEI12 indices were calculated as 77.43, 30.95, 20.93, 15.53, and 7.65, respectively. According to (Bayram and Çıtakoğlu 2023), in terms of the RRMSE values of Tunceli station, the SPEI1, SPEI3, SPI1, and SPI3 indices are considered poor, the SPEI6 and SPI6 indices are fair, the SPEI9 and SPI9 indices are good, and the SPEI12 and SPI12 indices are considered excellent forecasts. It was also determined that the forecast results for the SPI and SPEI indices at all-time scales for the Tunceli station are excellent according to the R² and KGE criteria. Except for the SPEI12, SPEI9, SPI9, and SPI12 indices, while the RRMSE was obtained as fair in other drought indices of the Tunceli station, the R²

Table 9 The most successful methods of three different versions of Tunceli Station and the performance criteria of these methods for test data

	V1	V2	V3		V1	V2	V3	
SPEI1	4ANN	4ANN	4ANN	SPEI1	4ANN	3GPR	3GPR	
	556.34	192.44	77.43		610.32	140.13	85.58	RRMSE
	0.76	0.25	0.11		0.71	0.12	0.08	MAE
	0.017	0.881	0.981		0.003	0.964	0.985	R ²
	-0.207	0.912	0.969		-89.409	-2.710	-0.563	KGE
SPEI3	0.463	0.915	0.978	0.405	0.964	0.982	OI	
	4GPR	4GPR	4SVM	SPEI3	4GPR	4GPR	4GPR	
	303.65	42.11	30.95		676.97	147.10	35.73	RRMSE
	0.52	0.07	0.05		0.50	0.11	0.03	MAE
	0.429	0.989	0.994		0.372	0.970	0.998	R ²
0.548	0.980	0.983	0.544		0.965	0.995	KGE	
SPEI6	0.698	0.986	0.990	0.633	0.972	0.996	OI	
	4ANN	4GPR	3GPR	SPEI6	2ANN	3GPR	4SVM	
	213.12	56.77	20.84		294.32	85.95	29.38	RRMSE
	0.39	0.11	0.04		0.35	0.10	0.04	MAE
	0.595	0.970	0.996		0.591	0.965	0.996	R ²
0.732	0.973	0.985	0.735		0.966	0.993	KGE	
SPEI9	0.782	0.972	0.992	0.770	0.969	0.993	OI	
	3GPR	4ANN	4ANN	SPEI9	4SVM	4GPR	4SVM	
	158.38	86.36	15.54		206.65	70.14	19.72	RRMSE
	0.31	0.19	0.03		0.26	0.09	0.03	MAE
	0.749	0.925	0.998		0.740	0.970	0.998	R ²
0.844	0.920	0.991	0.821		0.959	0.993	KGE	
SPEI12	0.853	0.942	0.994	0.837	0.968	0.993	OI	
	1SVM	4GPR	4GPR	SPEI12	4GPR	4ANN	4SVM	
	135.42	7.65	11.77		153.35	99.16	13.27	RRMSE
	0.25	0.01	0.03		0.19	0.14	0.02	MAE
	0.812	0.999	0.999		0.831	0.929	0.999	R ²
0.866	0.995	0.993	0.890		0.934	0.980	KGE	
	0.888	0.997	0.996	0.892	0.945	0.996	OI	

and KGE are interpreted as excellent. On the other hand, as the time scale increases, the RRMSE, R², and KGE support each other.

Visual criteria such as scatter and trajectory plots were utilized in the second stage of evaluating the results. Scatter and trajectory plots for all time scales of SPEI and SPI indices for Tercan and Tunceli stations are presented in Figs. 11 and 12. Figure 11 indicates for SPEI1 and SPI1 indices that version 1 was mediocre, while version 2 and version 3 were similar. For SPEI3, SPI3, and SPI6 indices, improvement was observed in version 1 forecasts, while version 2 and version 3 were similar. For SPEI6, SPEI9, SPEI12, SPI9, and SPI12 indices, version 1 and version 2 were similar, while version 3 was the most successful method. The trajectory plots of Figs. 11 and 12 indicate that the calculated SPEI and SPI values and the forecasted SPEI and SPI values for version 2 and version 3 are remarkably close.

In the third stage of evaluating the results, the Taylor diagrams, widely used recently to visualize, were employed.

Taylor diagrams for all time scales of SPEI and SPI indices for the Tercan and Tunceli stations are provided in Figs. 13 and 14. Figure 13 presents for the SPEI1 index that the SPEI1 value calculated with the V2-3GPR combination was selected as the most successful model since it is the closest to the calculated SPEI1 value. According to the Taylor diagrams for the Tercan station, the most successful combinations were selected as follows: V3-3GPR for SPEI3 and SPI12 indices, V3-4GPR for SPEI6, SPEI9, SPI1, and SPI6 indices, V3-4ANN for SPEI12 and SPI9 indices, and V2-4GPR for SPI3 index. Moreover, Fig. 14 indicates for Tunceli station that the most successful combinations were V3-4ANN for SPEI1 and SPEI9 indices, V3-4SVM for SPEI3, SPI6, SPI9, and SPI12 indices, V3-3GPR for SPEI6 and SPI1 indices, and V3-4GPR for SPEI12 and SPI3 indices. The results obtained from the Taylor diagrams support the classical performance criteria.

In the fourth stage of evaluating the results, visual criteria of violin and error-box plots were utilized. Violin and

Fig. 11 **a** Scatter and trajectory plots of three different versions of Tercan Station's SPEI indexes for test data. **b** Scatter and trajectory plots of three different versions of Tercan Station's SPI indexes for test data

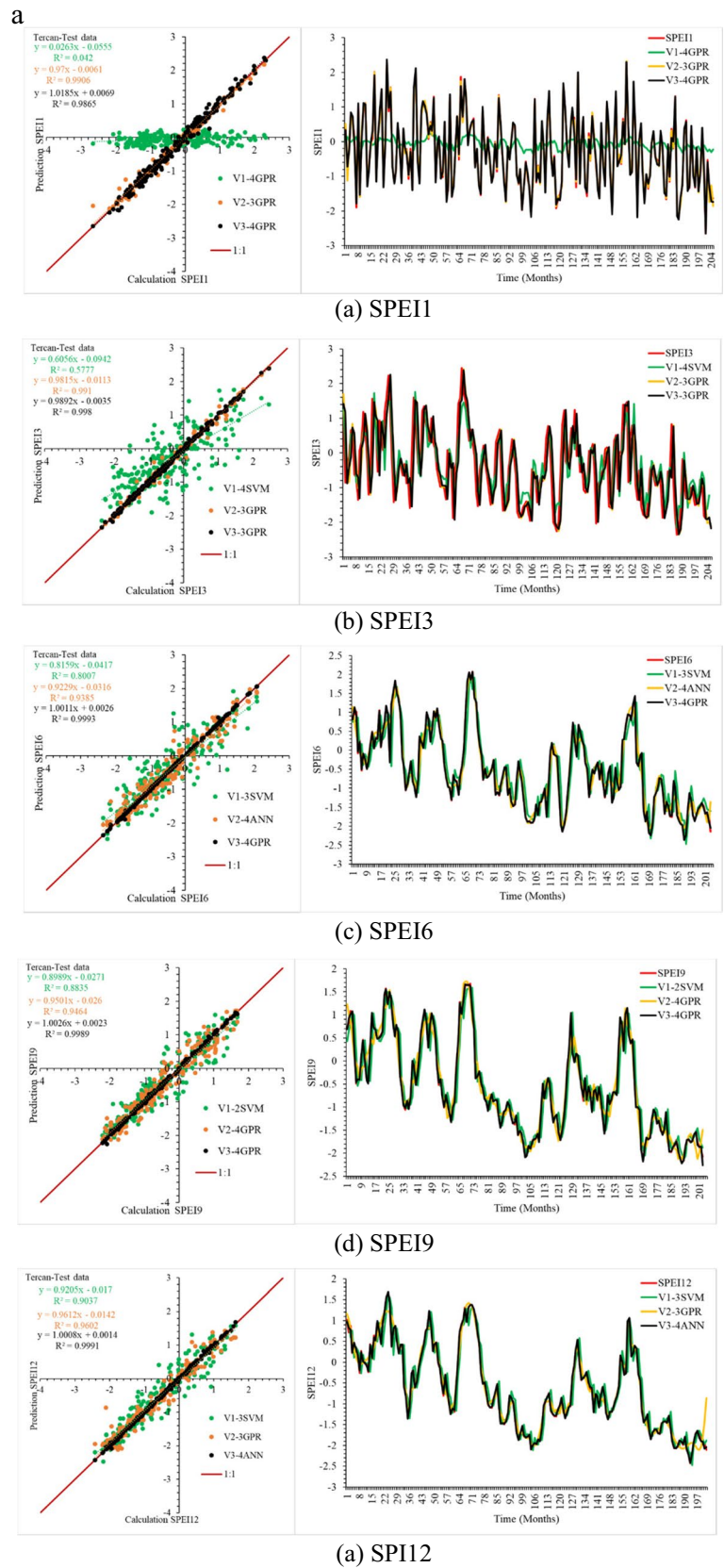


Fig. 11 (continued)

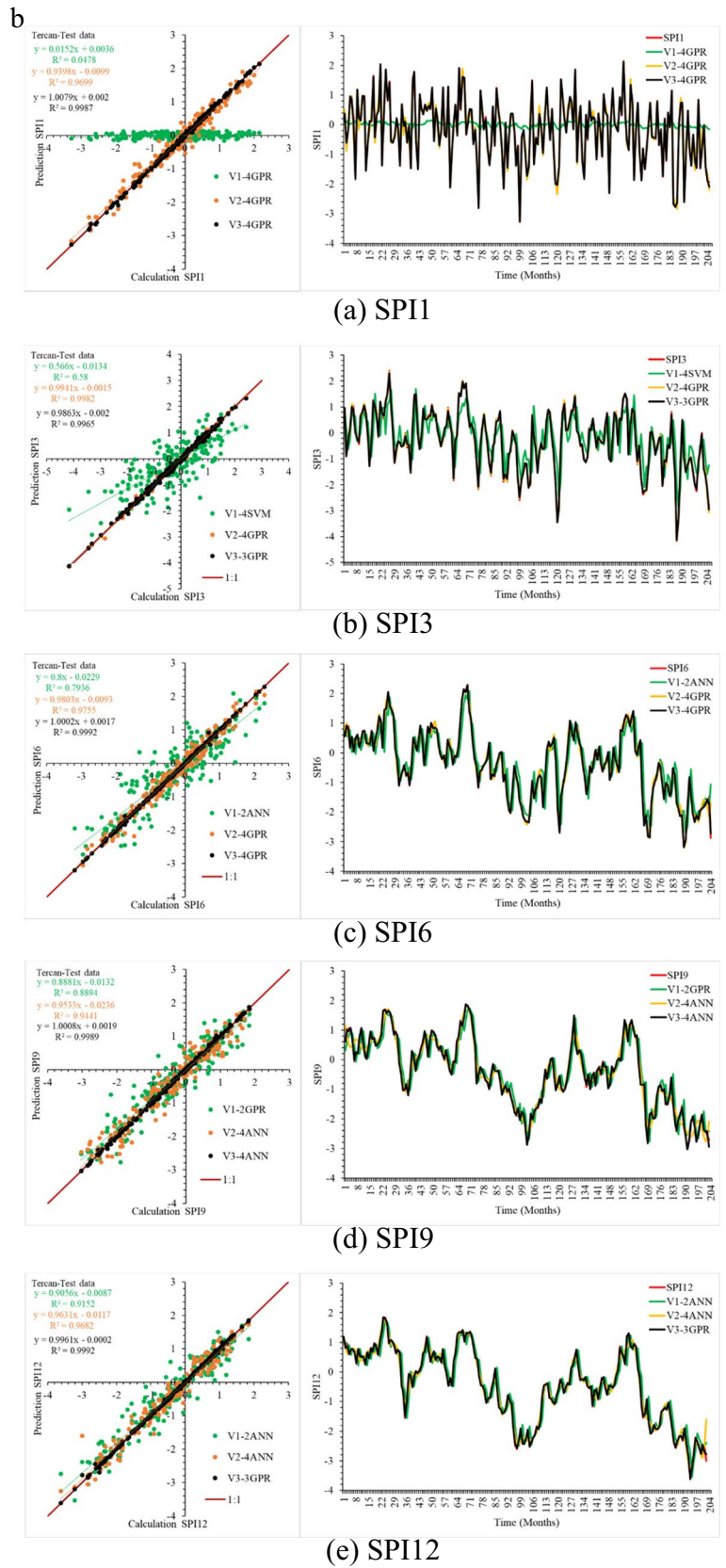


Fig. 12 **a** Scatter and trajectory plots of three different versions of Tunceli Station's SPEI indexes for test data. **b** Scatter and trajectory plots of three different versions of Tunceli Station's SPI indexes for test data

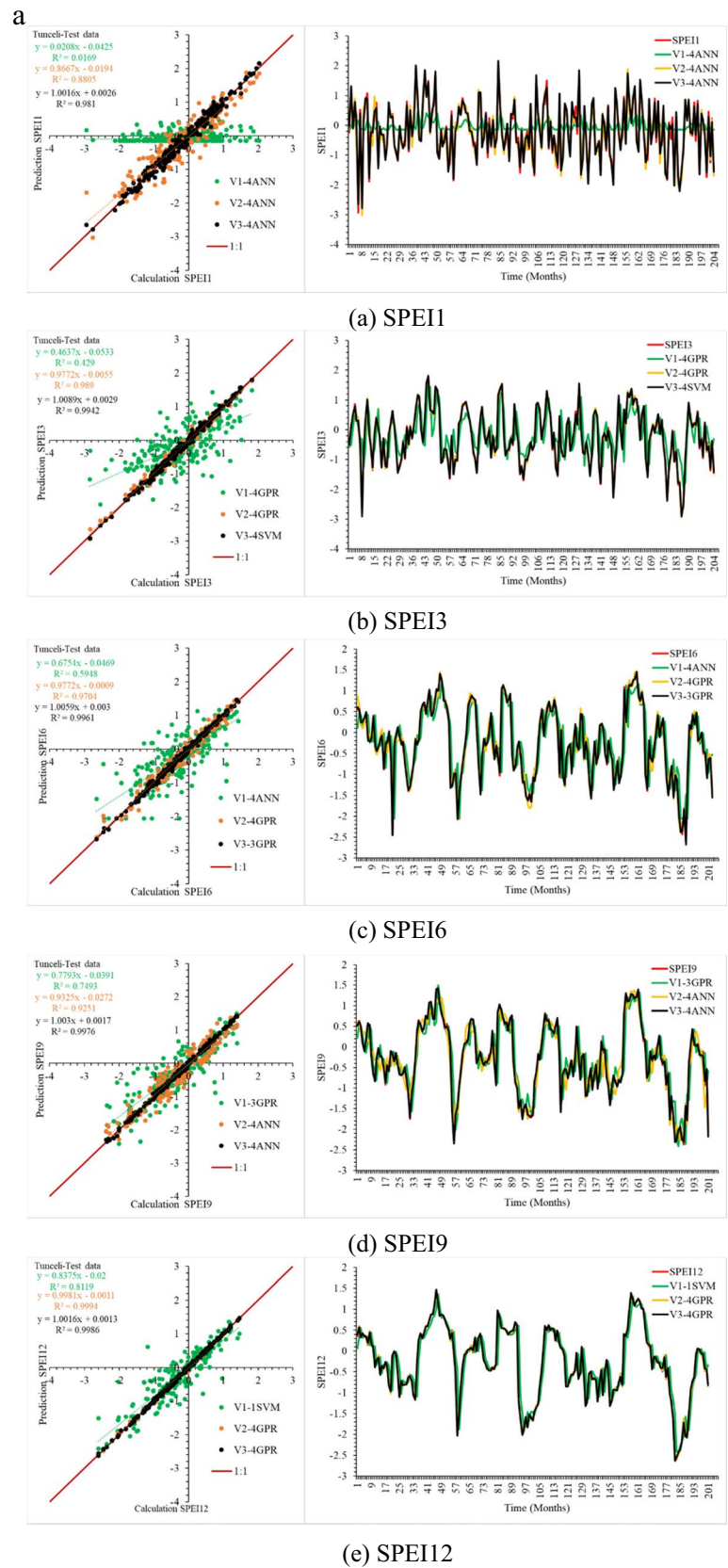


Fig. 12 (continued)

b

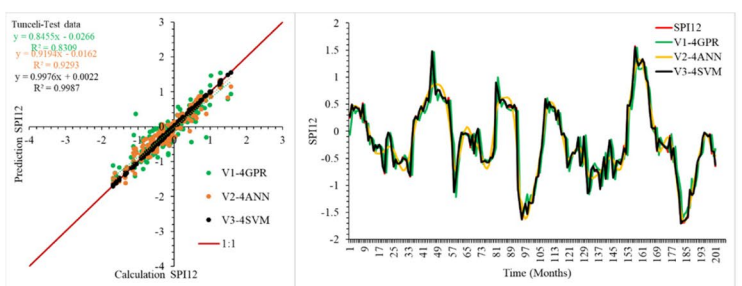
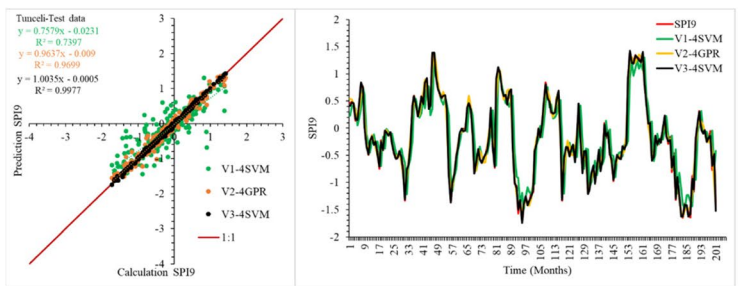
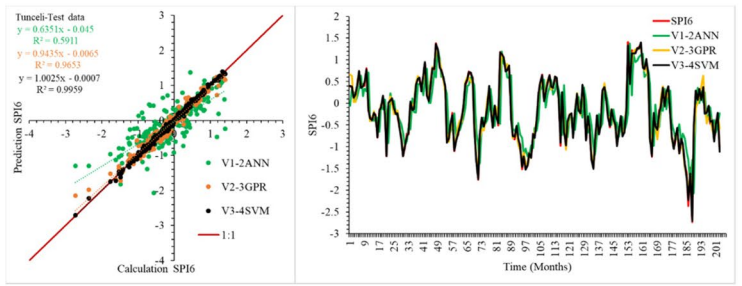
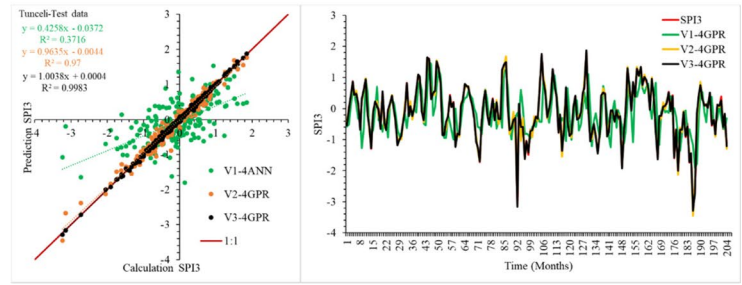
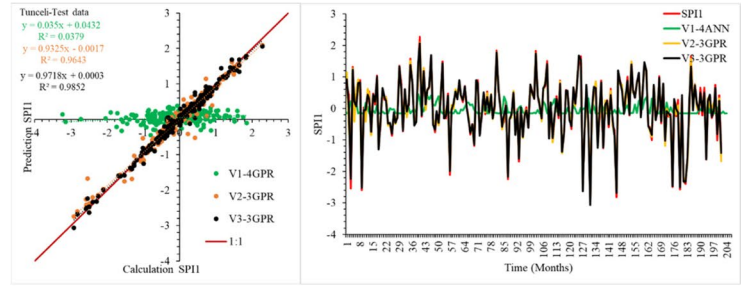
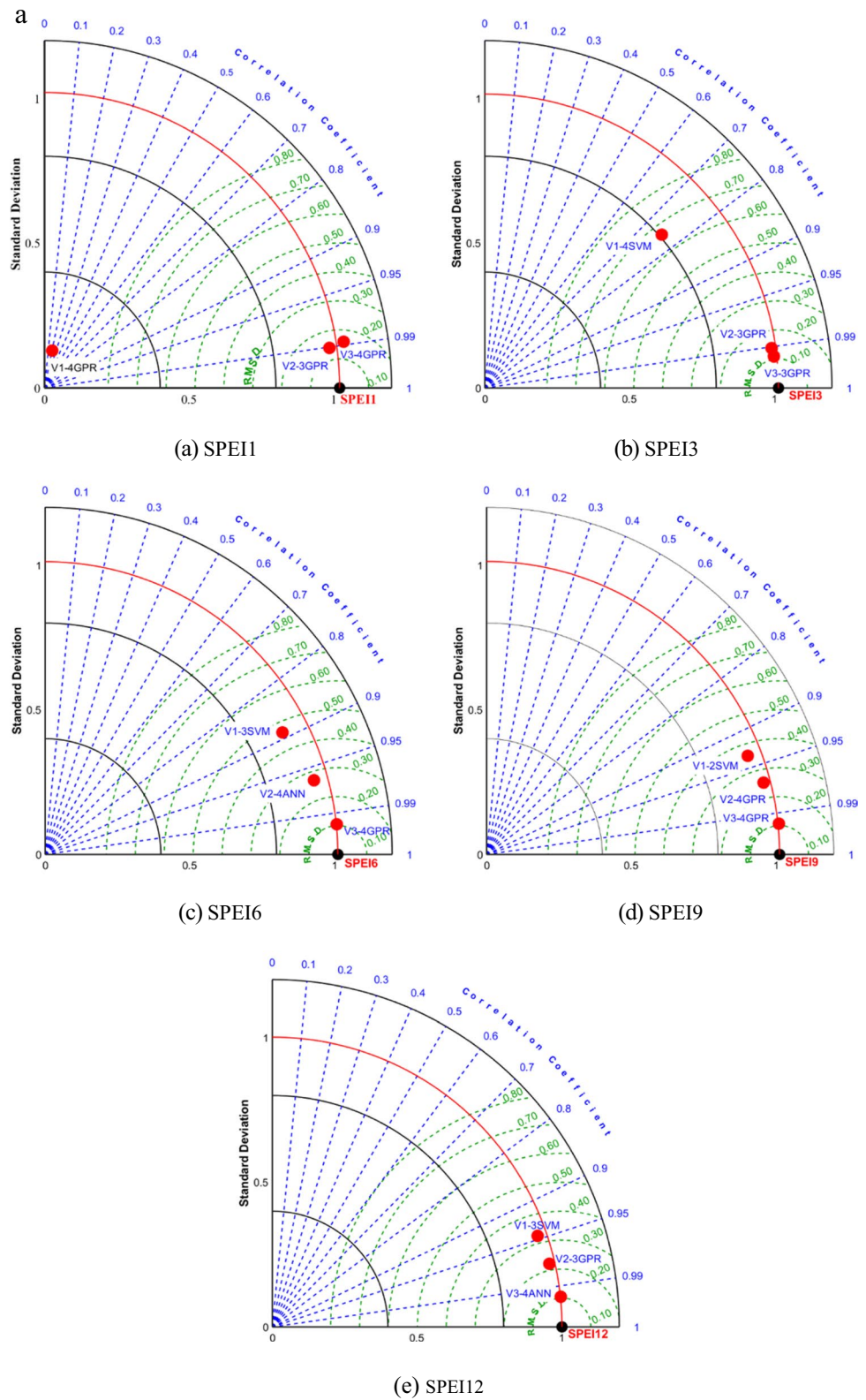


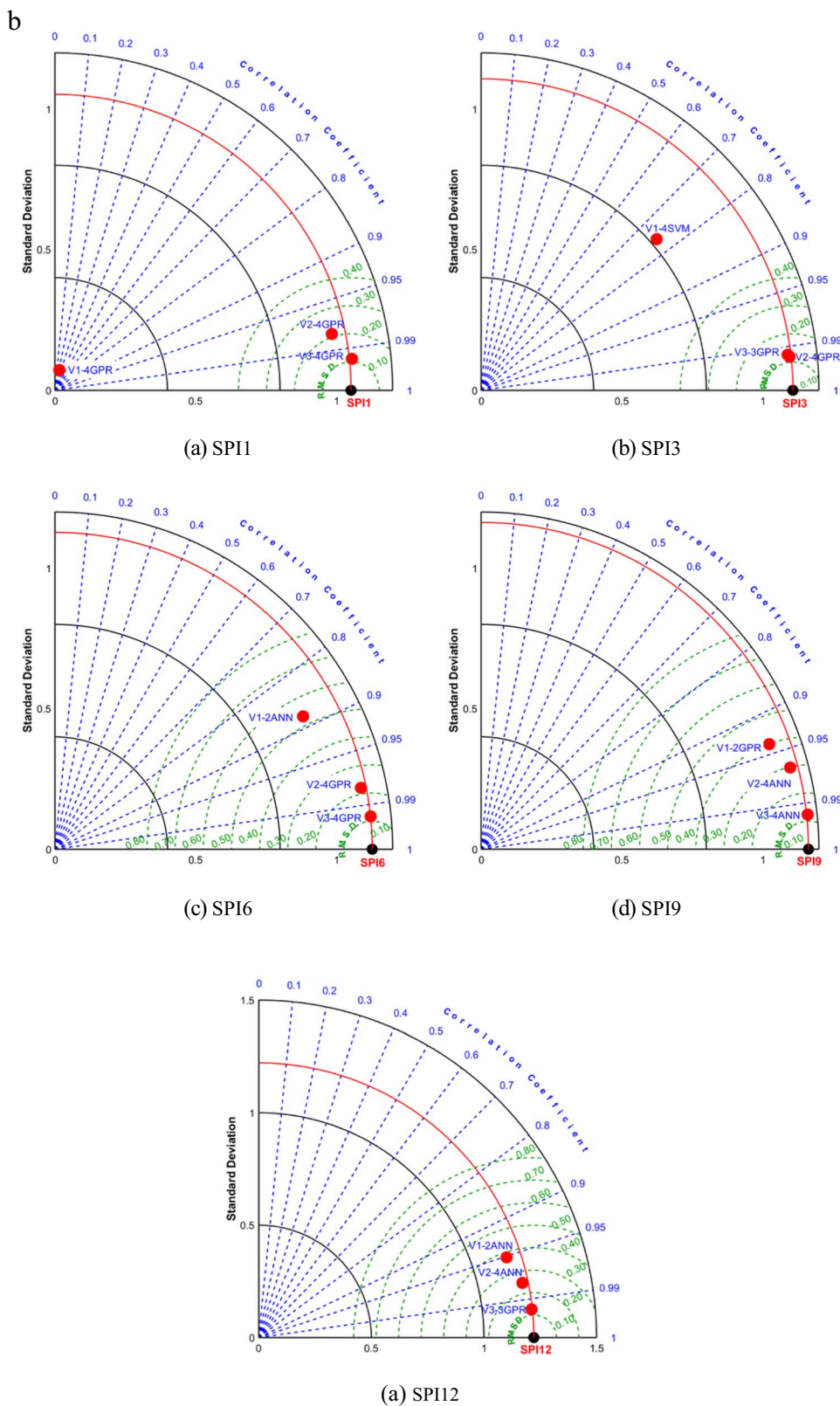
Fig. 13 **a** Taylor diagrams for Tercan station's SPEI indexes for test data. **b** Taylor diagrams for Tercan Station's SPI indexes for test data



error-box plot graphs for all time scales of SPEI and SPI indices for Tercan and Tunceli stations are presented in Figs. 18, 19, 20, 21, 22, 23 and 24. According to the violin

diagrams for Tercan station, version 2 model forecasts lower severity drought values for SPEI1, SPEI6, and SPEI9 indices. These results, however, were not visible in the scatter

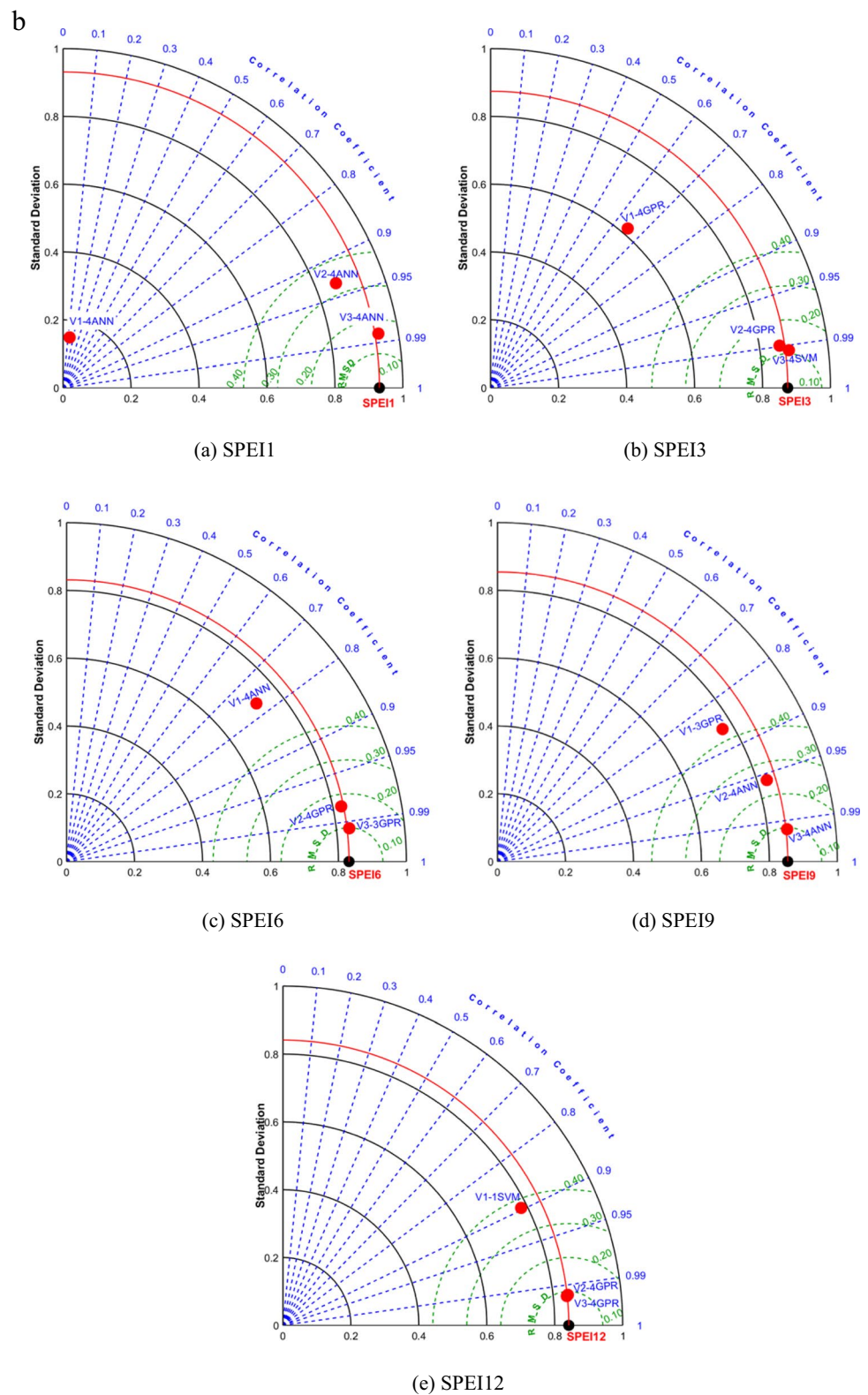
Fig. 13 (continued)



and trajectory plots for SPEI1, SPEI6, and SPEI9 indices. The violin diagrams for all time scales of SPEI and SPI indices for the Tercan and Tunceli stations indicated that version

2 and version 3 were similar. In the error-box plot graphs for all time scales except SPEI1, SPEI3, and SPI3 for the Tercan station, the error range of version 3 was lower than the other

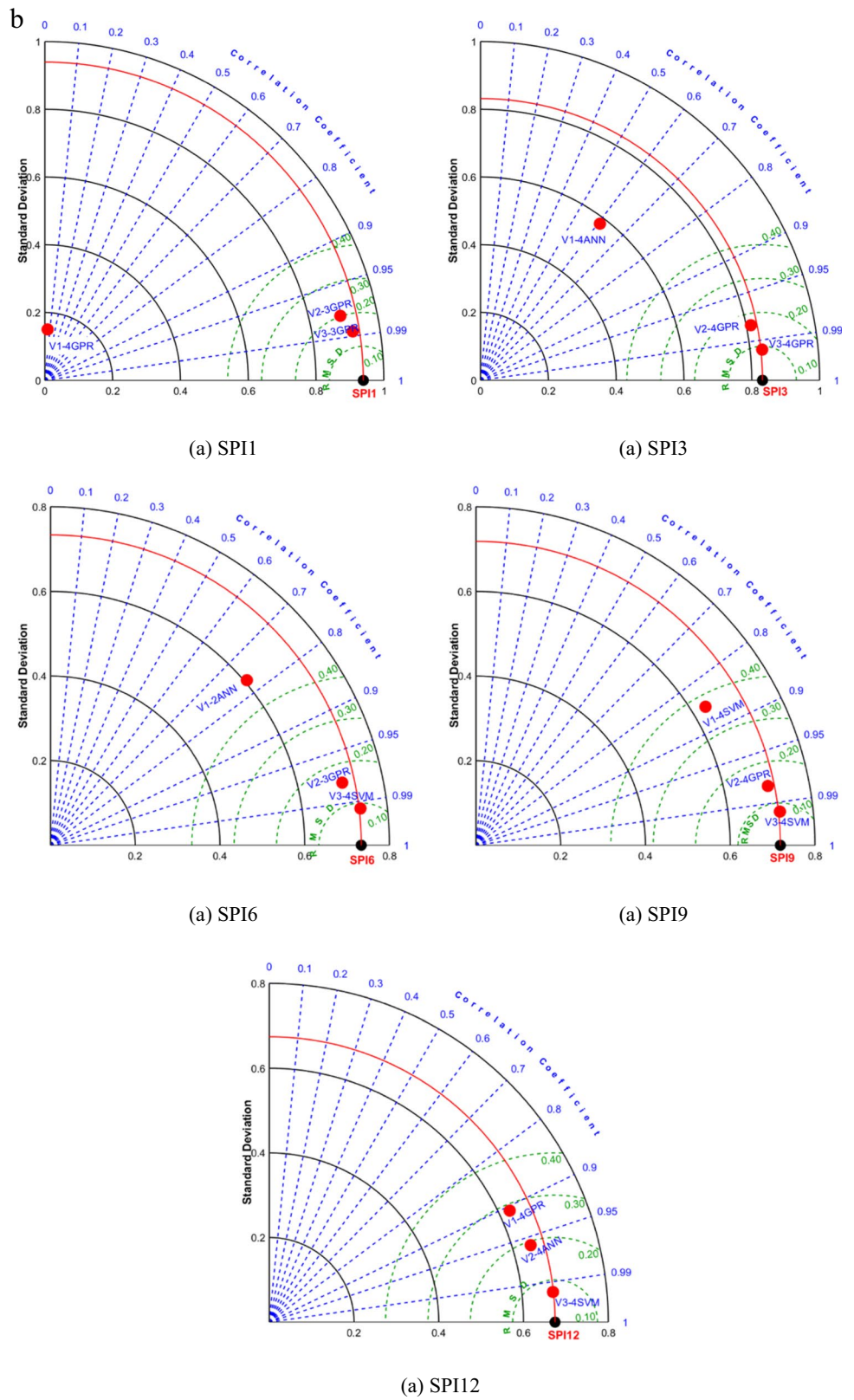
Fig. 14 **a** Taylor diagrams for Tunceli Station's SPEI indexes for test data. **b** Taylor diagrams for Tunceli Station's SPI indexes for test data



versions. However, for the SPEI1 and SPI3 indices for the Tercan station and SPEI3 and SPEI6 indices for the Tunceli station, the error ranges of versions 2 and 3 were quite close. The violin diagrams and scatter and trajectory plots for the

Tercan and Tunceli stations provided supportive results. In the error-box plot for Tercan and Tunceli stations, findings supporting the classical performance criteria and Taylor diagram were observed.

Fig. 14 (continued)



In the final stage, a Kruskal-Wallis test was conducted to determine whether there is a correlation between the forecasted SPEI and SPI indices and the calculated average of the

SPEI and SPI indices. Tables 10 and 11 present the results for all time scales of SPEI and SPI indices for the Tercan and Tunceli stations. Table 10 shows that for the Tercan station,

Table 10 Kruskal–Wallis test for Tercan Station's SPEI and SPI indexes for test data

	V1		V2		V3	
	<i>p</i> -value	H ₀	<i>p</i> -value	H ₀	<i>p</i> -value	H ₀
SPEI1	0.0037	Accept	0.9671	Reject	0.9950	Reject
SPEI3	0.3247	Reject	0.9977	Reject	0.9804	Reject
SPEI6	0.4339	Reject	0.9344	Reject	0.9879	Reject
SPEI9	0.6782	Reject	0.9747	Reject	0.9821	Reject
SPEI12	0.6725	Reject	0.9017	Reject	0.9980	Reject
SPI1	0.5415	Reject	0.9824	Reject	0.9757	Reject
SPI3	0.4857	Reject	0.9957	Reject	0.9890	Reject
SPI6	0.6509	Reject	0.9993	Reject	0.9826	Reject
SPI9	0.7585	Reject	0.9255	Reject	0.9828	Reject
SPI12	0.8340	Reject	0.9782	Reject	0.9694	Reject

Table 11 Kruskal–Wallis test for Tunceli Station's SPEI and SPI indexes for test data

	V1		V2		V3	
	<i>p</i> -value	H ₀	<i>p</i> -value	H ₀	<i>p</i> -value	H ₀
SPEI1	0.0905	Reject	0.9759	Reject	0.9913	Reject
SPEI3	0.5812	Reject	0.9625	Reject	0.9786	Reject
SPEI6	0.6683	Reject	0.9592	Reject	0.9882	Reject
SPEI9	0.9654	Reject	0.8532	Reject	0.9894	Reject
SPEI12	0.8224	Reject	0.9908	Reject	0.9983	Reject
SPI1	0.6038	Reject	0.9094	Reject	0.9709	Reject
SPI3	0.8911	Reject	0.9344	Reject	0.9772	Reject
SPI6	0.9632	Reject	0.9956	Reject	0.999	Reject
SPI9	0.7622	Reject	0.9846	Reject	0.9716	Reject
SPI12	0.9866	Reject	0.9818	Reject	0.9470	Reject

the *p*-value is greater than 5% for all indices except SPEI1, which leads to the rejection of the null hypothesis (H₀). This implies that the averages of the forecasted drought indices are the same as those of the calculated drought indices. However, for SPEI1 at Tercan station, the analysis results did not pass the Kruskal–Wallis test for version 1. Table 11 shows that for the Tunceli station, the *p*-value is greater than 5% for all indices, leading to the rejection of the H₀. Thus, it was concluded that the averages of the forecasted and calculated values were the same. These test results confirm that the usability of ML has been demonstrated.

The utilized numerical and visual criteria point out that the GPR method provides superior results for the SPEI and SPI indices of the Tercan station at ‘all-time scales’ compared to the ANN and SVM. Additionally, for the Tercan station, it was observed that the hybrid model (version 3) was more successful than the stand-alone and tribrid models (versions 1 and 2) for the SPEI and SPI indices at all-time scales. For Tunceli station, the GPR and ANN methods were determined as the most successful method in the SPEI index twice each, and the SVM method was determined as the most successful method once. As can be seen from the results, different machine-learning

methods were successful in five different time scales of the SPEI index at Tunceli station. This is due to the descriptive statistics of the Tunceli station given in Table 1., because the rainfall data of the Tunceli station is flatter, and the temperature data is more variable. For Tunceli station, the GPR method was the most successful twice in the SPI index, while the SVM method was the most successful three times. While the GPR method was the most successful at the Tercan station in the SPI index, the SVM method was the most successful at the Tunceli station. While Tunceli station is rainier, Tercan station is colder, and the climate characteristics of these two stations are different. Lag numbers from 1 to 4 have been considered input data, and for both Tercan and Tunceli stations, 4 lags at most have provided the best forecast results. These results are significant according to autocorrelation graphs.

Discussion

The concepts underlying the SPI and the SPEI share similarities, yet there are distinct variations in their calculation parameters. The SPI solely accounts for precipitation,

offering simplicity in computation and robust adaptability across temporal and spatial scales (Zhou et al. 2013; Tao et al. 2014; Pei et al. 2020). Conversely, the SPEI incorporates the cumulative disparity between precipitation and potential evapotranspiration, comprehensively depicting alterations in surface water equilibrium (Zhang et al. 2015). Nevertheless, the escalation of evaporation attributed to warming presents a non-negligible factor in accurately evaluating drought conditions amidst global warming concerns. While the SPEI stands out as notably superior to the SPI in drought monitoring due to its consideration of evapotranspiration (Shaowei et al. 2013; Shi et al. 2019), its applicability in arid regions may face constraints. Furthermore, despite the advantages of the SPEI, the SPI continues to enjoy widespread usage globally (Zhou et al. 2013; Yuan et al. 2016; Mohammed et al. 2018; Sobral et al. 2019; Li et al. 2021).

There is an increasing potential related to the crucial impact of signal processing techniques on achieving enhanced accuracy in drought modeling. For instance, various forms of wavelet transformations, such as CWT (Özger et al. 2020), DWT (Altunkaynak and Jalilzadnezamabad 2021; Achite et al. 2023), EMD (Özger et al. 2020; Başakın et al. 2021a), and VMD (Citakoglu and Coşkun 2022; Ekmekcioğlu 2023) algorithms have been frequently employed. However, the literature review indicates drought studies do not commonly utilize the TQWT technique. (Latifoğlu and Özger 2023) considered the TQWT technique, emphasizing its success over other signal processing methods due to its lower computational load and shorter processing time. Given the abundance of surface and groundwater and the presence of irrigable flat land, this study integrated the TQWT sub-band decomposition technique with ANN, GPR, and SVM for the Tunceli and Tercan Stations. The results of this study confirm similar findings of (Latifoğlu and Özger 2023) that the TQWT sub-band decomposition technique yields successful outcomes.

An evaluation in terms of the index used indicates that (Altunkaynak and Jalilzadnezamabad 2021) utilized the PDSI for the Marmara region; (Özger et al. 2020) and (Başakın et al. 2021a) employed the sc-PDSI for the Mediterranean region; (Katipoğlu 2023) introduced the SDI for the Yeşilırmak basin; (Evkaya and Kurnaz 2021) utilized the SPI for the Marmara region; (Citakoglu and Coşkun 2022; Coşkun and Citakoglu 2023) applied SPI for the Marmara region; (Gholizadeh et al. 2022; Danandeh Mehr et al. 2023) used the SPEI for the Central Anatolia region; (Gul et al. 2023) employed SPI for the Aegean region; (Reihanifar et al. 2023) conducted SPI for the Mediterranean region, using various ML methods. Previous studies have not addressed the drought indices forecast for the Eastern Anatolia region in Türkiye. Despite being the coldest geographical region in Türkiye, the Eastern Anatolia region remains under snow for most months of the year, which limits the agricultural

window. Given the significant temperature variations, drought forecasting for the Eastern Anatolia region is crucial. This study is the first to forecast two different drought indices in Turkish studies since previous research focused on a single one. This study evaluated the performance of various stand-alone, hybrid, and tribrid machine-learning methods via different drought indices. The current study is considered original in this respect. While previous studies have generally been confined to a single time scale, the current research provides a comprehensive perspective by considering five different scales in addressing the problem.

Turkish-origin previous drought forecast studies have not investigated the accuracy of model results. However, in some studies, the forecasting results have been meticulously examined and evaluated regarding statistical significance. For instance, they utilized the Kolmogorov-Smirnov test to assess whether the obtained forecasts were statistically acceptable, comparing the observed and forecasted time series to determine their accuracy (Özger et al. 2020; Başakın et al. 2021a, b; Coşkun and Citakoglu 2023).

The performance superiority between any ML techniques is something that researchers have not yet determined in previous studies. The studies (Liu et al. 2017; Mokhtarzad et al. 2017; Moghaddasi et al. 2024) suggest that SVM outperforms ANN in certain contexts. Depending on the specific application of these techniques, it remains unclear which one is definitively superior to the other. However, what is evident from this study is that the GPR model surpasses both ANN and SVM, although they remain competitive during model training.

Previous studies have underscored the potential of ML models in elucidating intricate relationships between meteorological and hydrological variables and drought occurrence (Mishra and Desai 2005; Morid et al. 2006; Bacanlı et al. 2009; Marj and Meijerink 2011). Among the array of models employed in research endeavors, the GPR model has showcased commendable efficacy across numerous studies (Sihag et al. 2017; Mishra and Kushwaha 2019; Shabani et al. 2020; Ghasemi et al. 2021), consistently delivering high performance, as in this study.

In this study, as the time scale increases, the performances of the RRMSE, KGE and R2 criteria in the most successful ML models support each other. In other words, as the time scale increases, the predictive power of the ML models is included in the perfect prediction classification. Except for some time scales, there is a decrease in the RRMSE and MAE criteria as the time scale increases. The results obtained from this study support the findings of the study conducted by (Anshuka et al. 2019). (Anshuka et al. 2019) also found that the performance criteria decreased with the increase in the time scale.

This study validated the results of the ANN, GPR, and SVM approaches, which were used within stand-alone,

hybrid, and tribrid models, using Kruskal-Wallis tests. Although the highest forecast performance was achieved in the recommended hybrid models (version 3), it was observed that the forecast performance of the recommended tribrid models (version 2) was close to that of hybrid models and performed well. Besides, the hegemony of hybrid models over stand-alone models (version 1 in this study) has been proven again, as expressed in the literature review. Moreover, using a FS approach in the tribrid models could reduce the computational load and shorten the processing time by selecting the most compelling features instead of forecasting each sub-band separately. Thus, the superiority of tribrid models was demonstrated quantitatively in the literature and qualitatively in terms of reduced computational load and processing time in this study.

Conclusion

This study focuses on drought forecasting in the Eastern Anatolia region of Türkiye, referred to as Upper Mesopotamia, which is a geographically significant area not investigated in previous research. The data utilized in the study comes specifically from two different meteorological measurement stations in the Karasu sub-basin of the Euphrates Basin. Various ML models, including stand-alone, hybrid, and tribrid approaches, are evaluated to develop high-performance drought forecasting models, and their performances are compared across different time scales.

Analyses were conducted using two widely accepted drought indices, namely the SPI and the SPEI, at different time scales (1-, 3-, 6-, 9-, and 12-months). Three different approaches were employed for forecasting: stand-alone ML methods (ANN, GPR, SVM), hybrid ML methods (TQWT-ANN, TQWT-GPR, TQWT-SVM), and tribrid ML methods (TQWT-FSRFtest-ANN, TQWT-FSRFtest-GPR, TQWT-FSRFtest-SVM). The temporal variation of meteorological drought events was assessed through SPI and SPEI outputs, revealing an increase in drought trends, durations, and magnitudes, especially since the 2000s. Frequency analysis of drought indices indicated that the East Anatolia Region mainly experienced moderate drought events.

ML models were applied to forecast the drought indices, including stand-alone, hybrid, and tribrid versions. The evaluation of these models involved classical performance criteria (such as R^2 , RRMSE, MAE, etc.), visual criteria (scatter and trajectory plots), Taylor diagrams, violin, and a statistical test (Kruskal-Wallis's test). The results showed that the hybrid and tribrid models, particularly those utilizing the TQWT sub-band decomposition technique, performed high forecasting accuracy.

The analysis concluded that the GPR approach outperformed the ANN and SVM methods for SPEI and SPI

outputs at the Tercan station. Conversely, for the Tunceli station, the GPR and ANN methods were identified as the most successful approaches for SPEI output twice each, while the SVM method achieved this designation once. Thus, no clear superiority was observed between the GPR and ANN methods for the SPEI output at the Tunceli station. Regarding the SPI output at the Tunceli station, the GPR method was identified as the most successful approach twice, whereas the SVM method achieved this designation three times. Consequently, the SVM method emerged as the most successful approach for the SPI output at the Tunceli station.

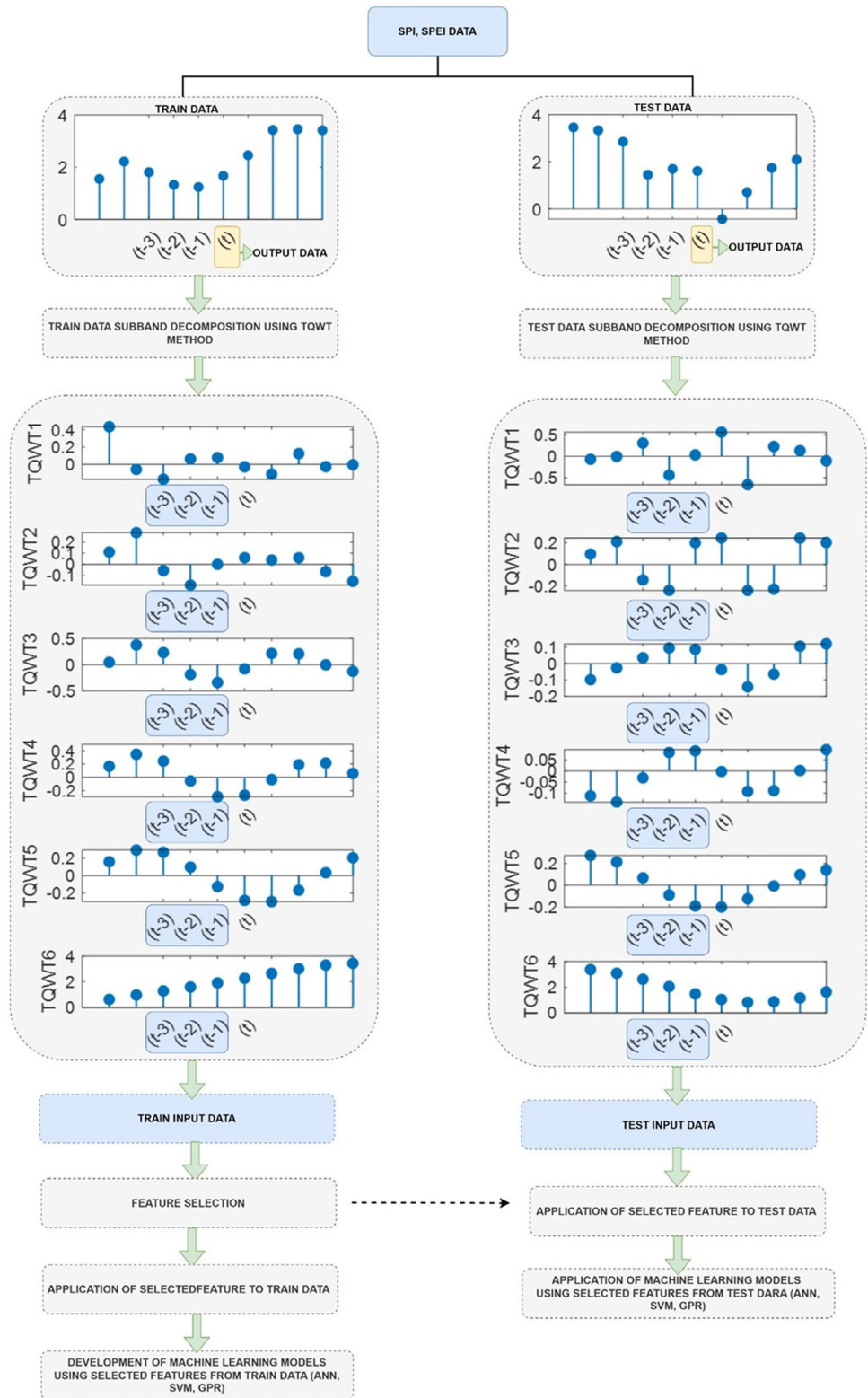
This study contributes to the existing literature by focusing on high-performance drought forecasting, proposing hybrid and tribrid modeling and traditional approaches, and emphasizing the efficiency of the TQWT sub-band decomposition technique. The findings underscore the significance of considering multiple time scales and utilizing advanced ML methods for accurate drought forecasting, providing valuable insights for tackling climate change, efficient water resource management, and environmental planning in the Eastern Anatolia region and similar semi-arid climatic zones.

This study can be characterized by the following seven fundamental limitations: (i) Utilization of two meteorological stations specific to the Eastern Anatolia Region, where agriculture is limited. (ii) Utilization of 58 years of data. (iii) Variation in precipitation and temperature conditions for the Tercan station ranging from 0 to 140.7 mm and -16.4 to 24.9 °C, respectively, and for the Tunceli station ranging from 0 to 540.2 mm and -10.2 to 29.7 °C, respectively. (iv) Calculation of two different drought indices, namely SPI and SPEI. (v) Calculation of temporal scales including 1-, 3-, 6-, 9-, and 12-months. (vi) Utilization of three ML methods, namely ANN, GPR, and SVM, as well as the TQWT sub-band decomposition method. (vii) Employment of the FSRFtest for selecting effective input parameters for drought prediction.

In future studies, drought prediction in the Eastern Anatolia Region can be enhanced by employing other sub-band decomposition methods and different ML models available in the literature. Drought prediction models can be increased with ML methods by calculating different drought indices of other stations in the Eastern Anatolia Region and making comprehensive analyses. Different climate and hydrological data can increase the accuracy of drought models for local and regional predictions with the help of different indices used. This study demonstrates the capability of the TQWT sub-band decomposition method to address such problems. Utilizing the TQWT method for similar analyses in different regions and predicting different drought indices can provide valuable insights. Moreover, using the FSRFtest can contribute to effectively selecting input variables in various problems.

Appendix

Fig. 15 Flowchart for Model 2



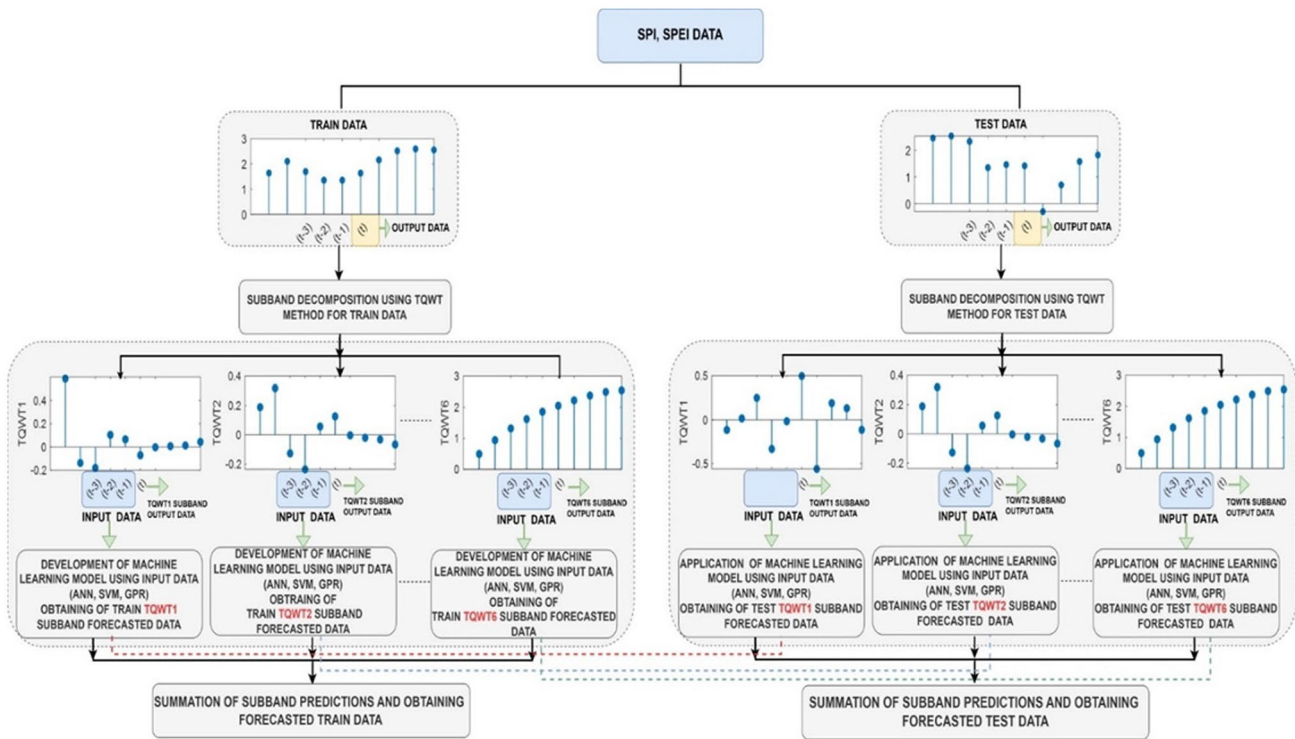


Fig. 16 Flowchart for Model 3

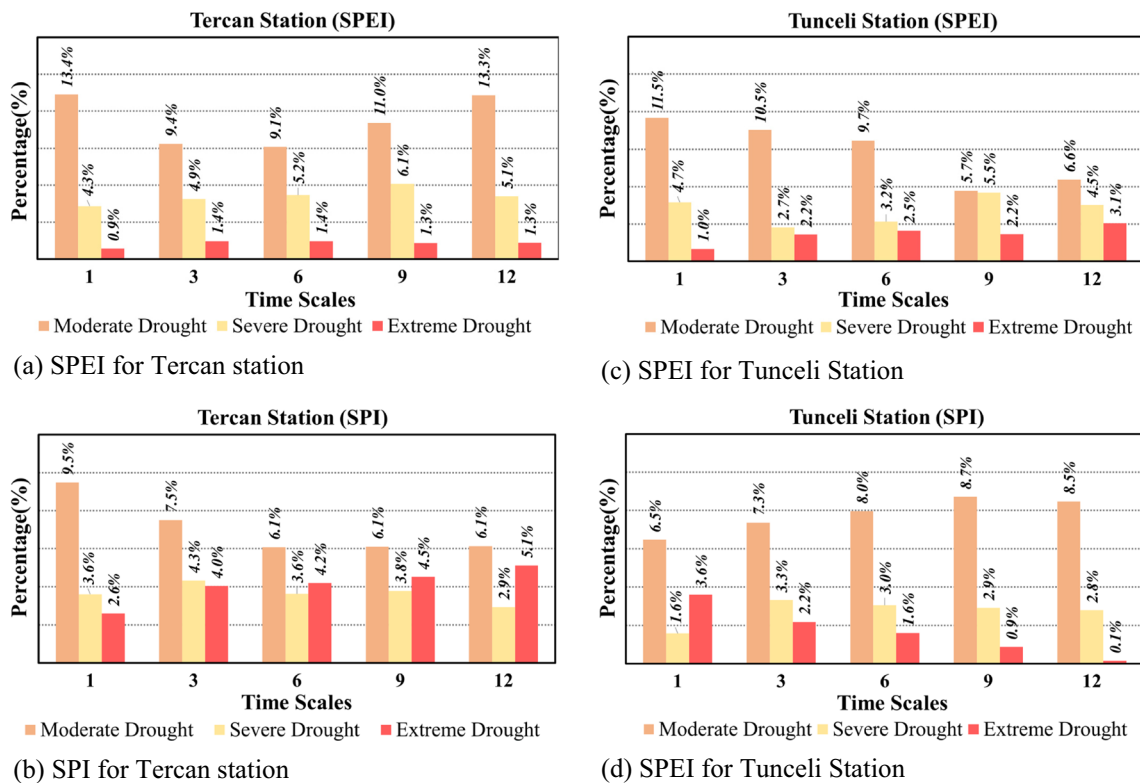


Fig. 17 Relative frequency of meteorological drought events in the Tercan and Tunceli stations

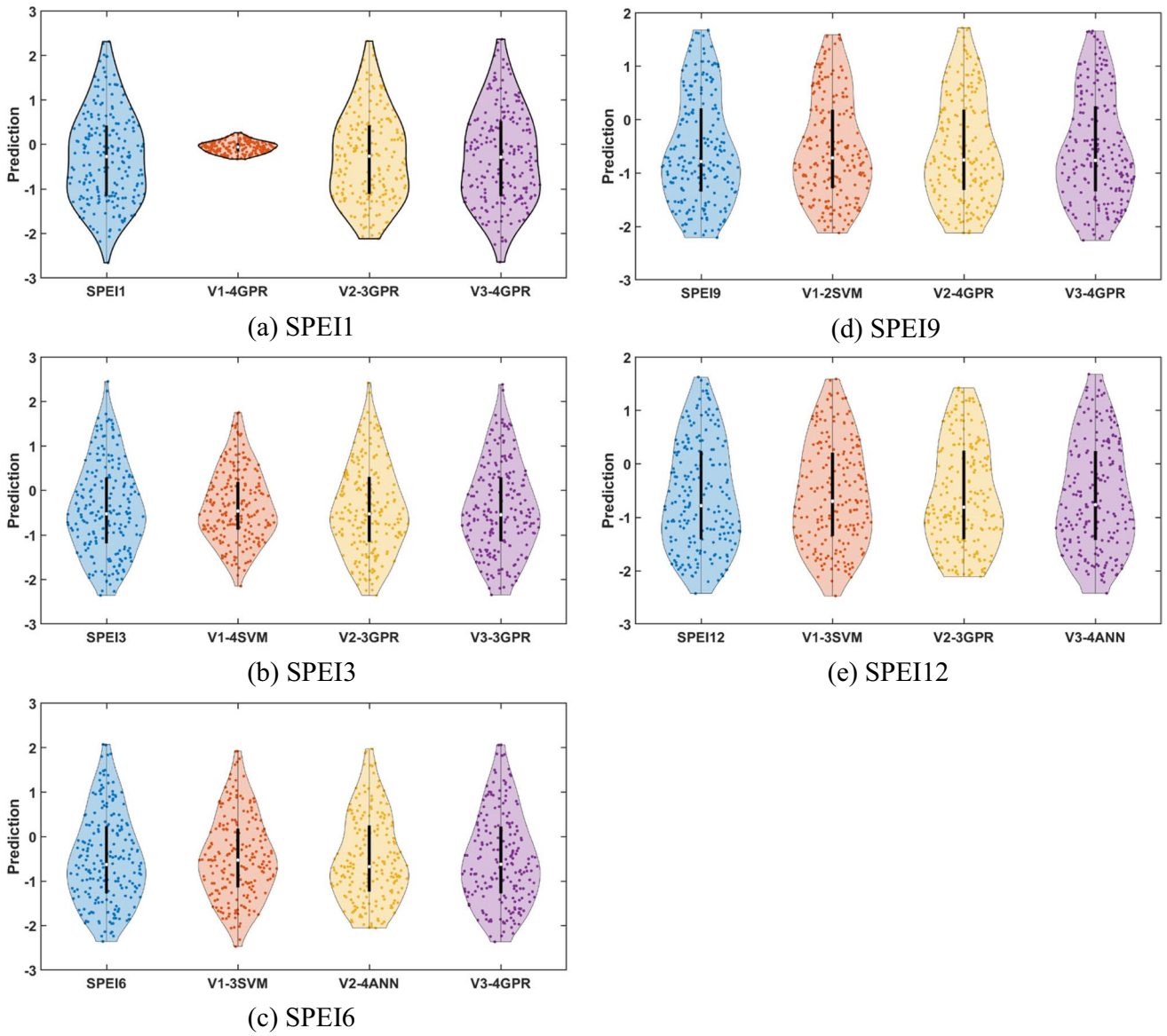


Fig. 18 Violin diagram for Tercan Station's SPEI indexes for test data

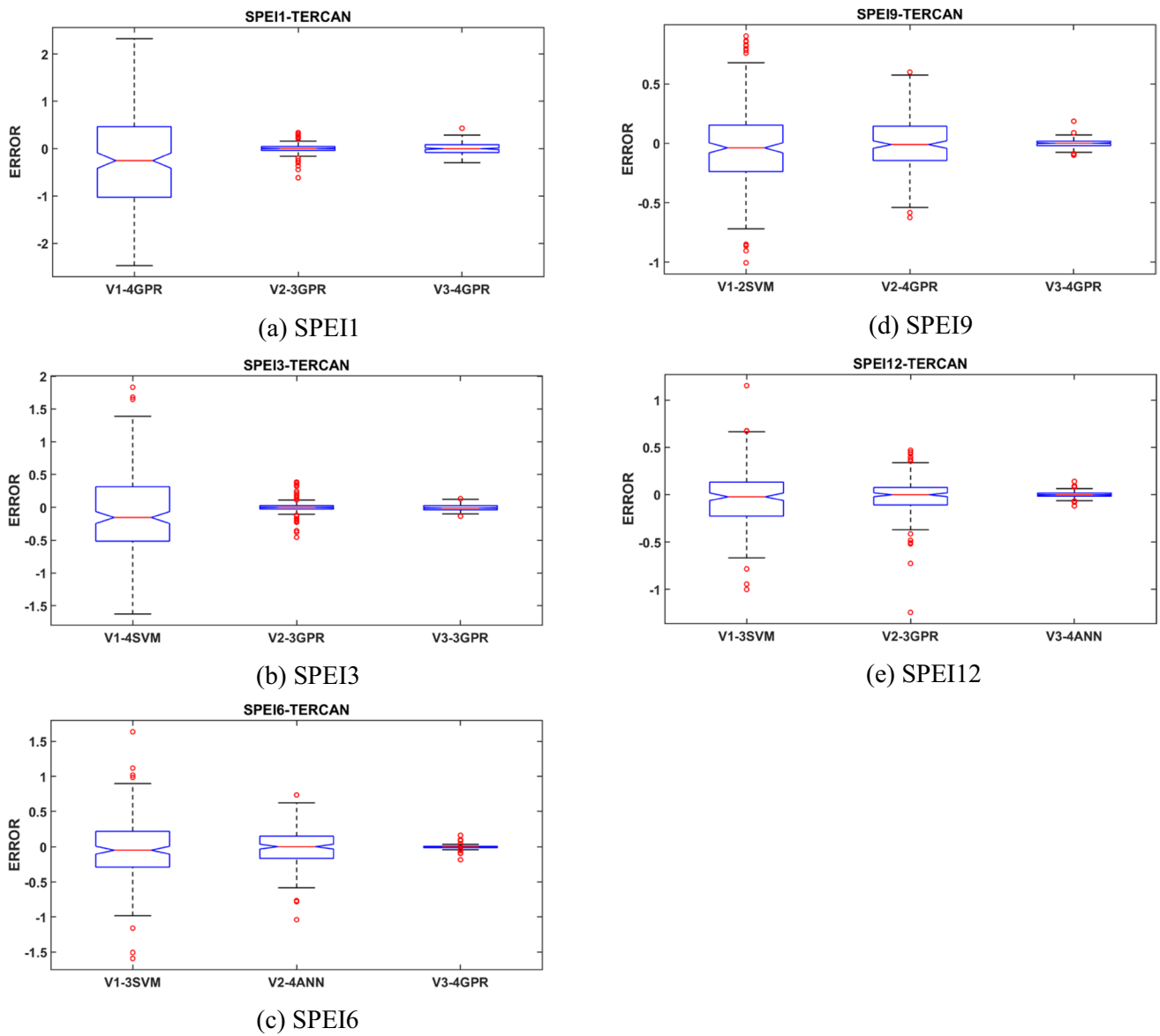
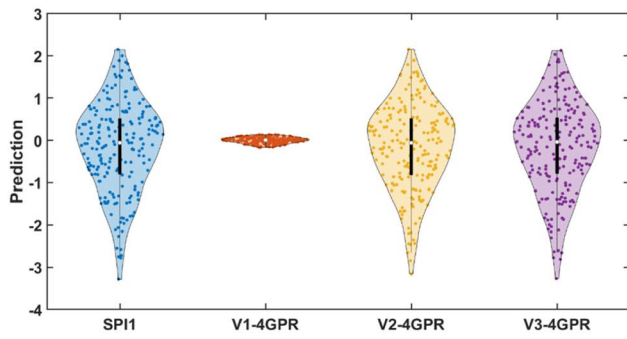
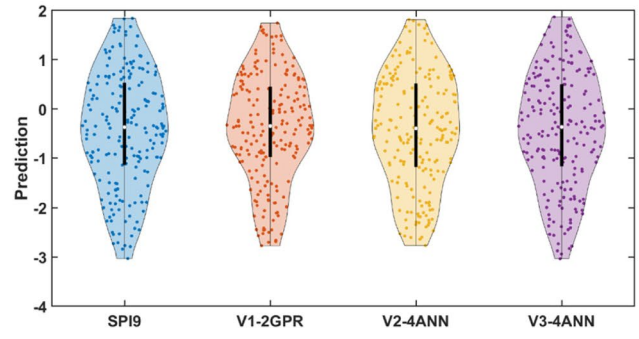


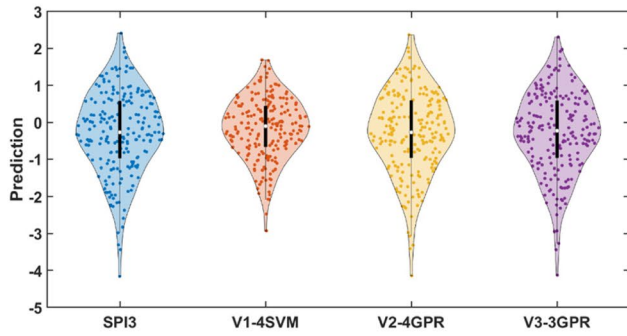
Fig. 19 Error-box plot for Tercan Station's SPEI indexes for test data



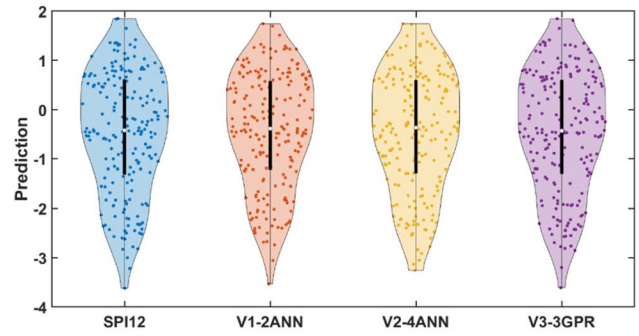
(a) SPI1



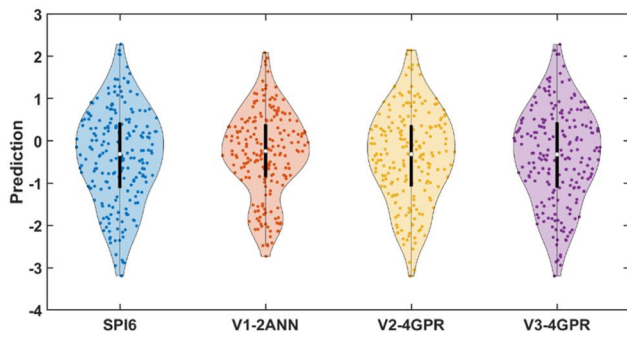
(d) SPI19



(b) SPI3



(e) SPI12



(c) SPI6

Fig. 20 Violin diagram for Tercan Station's SPI indexes for test data

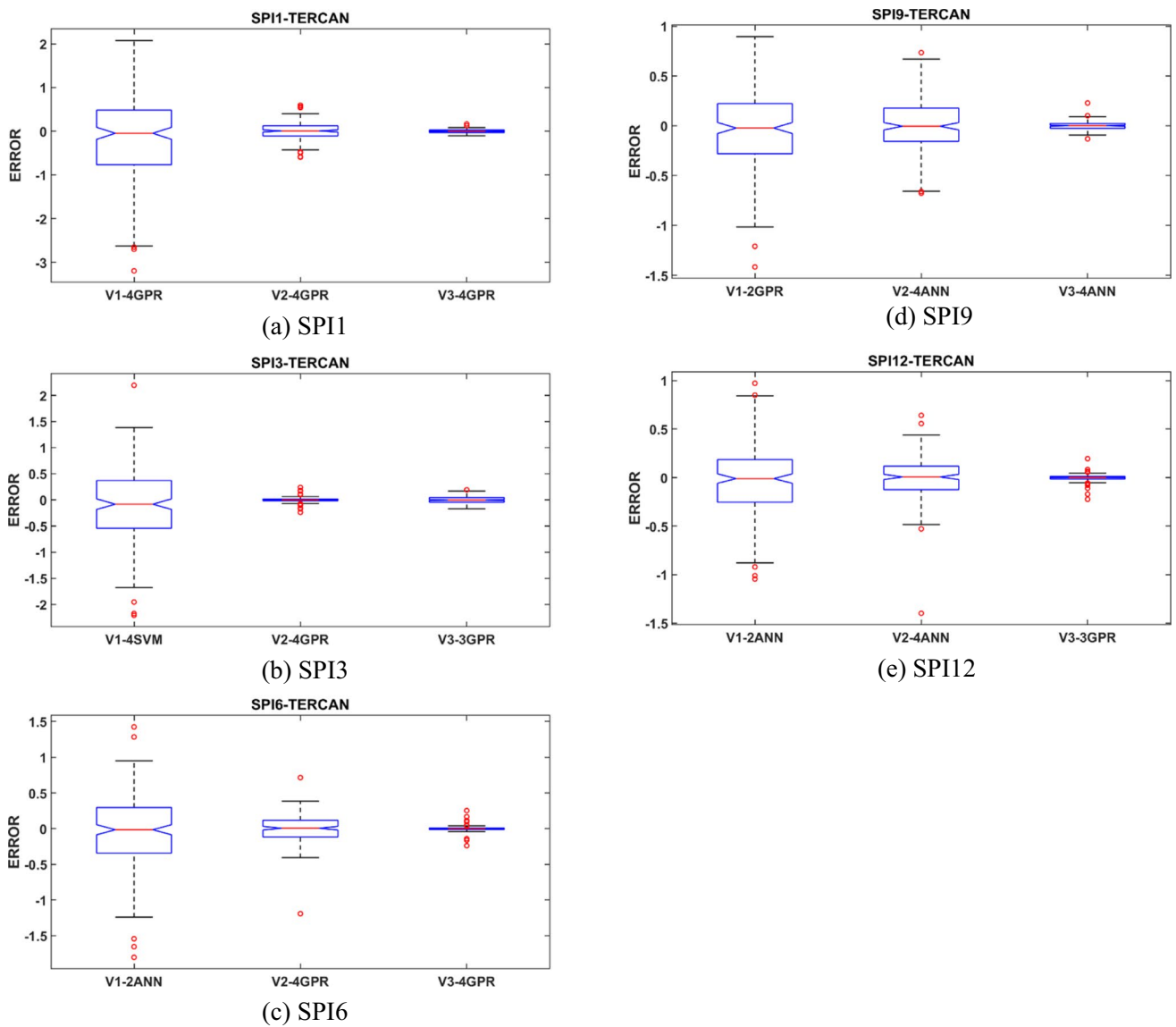


Fig. 21 Error-box plot for Tercan Station's SPI indexes for test data

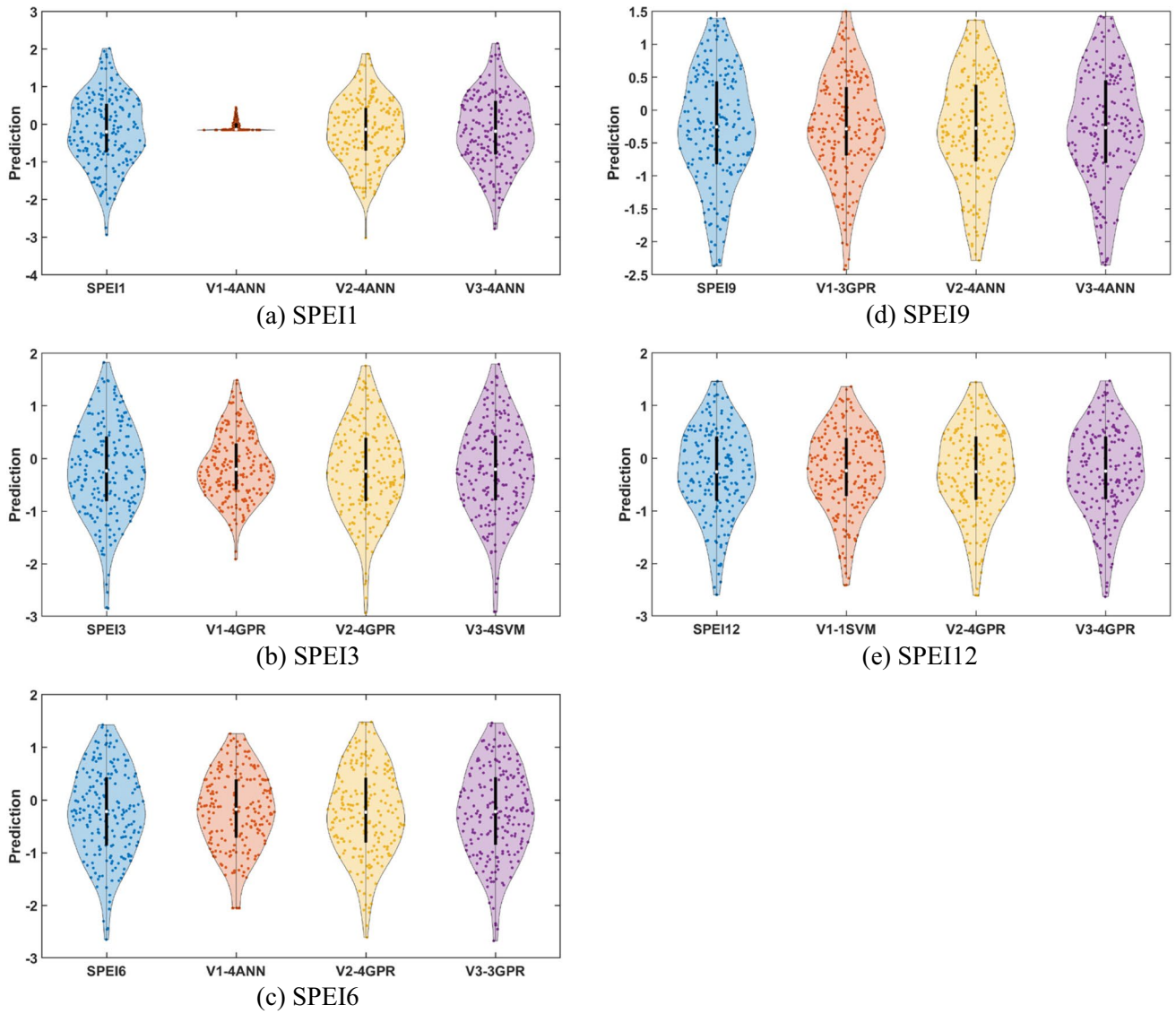


Fig. 22 Violin diagram for Tunceli Station's SPEI indexes for test data

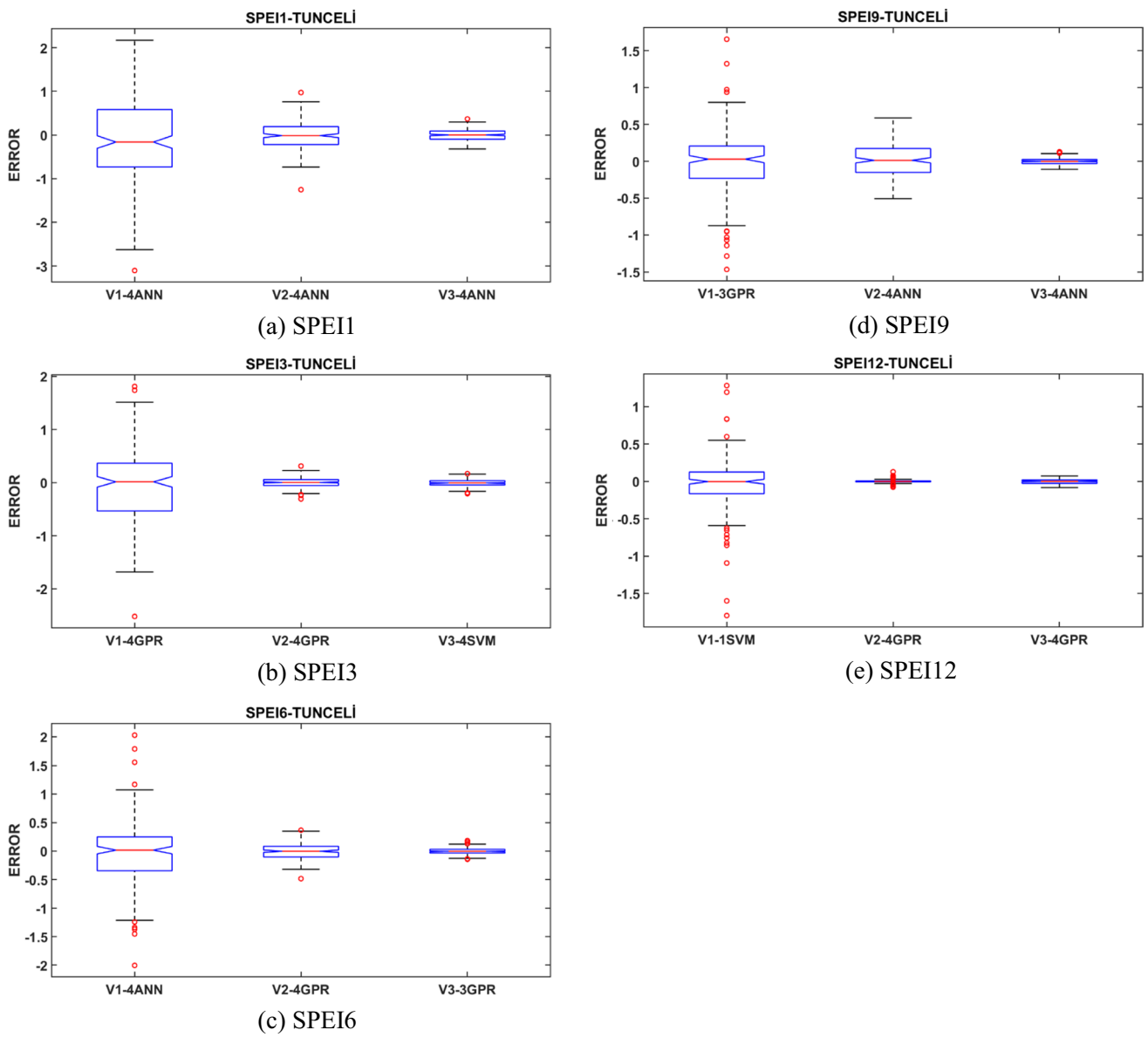


Fig. 23 Error-box plot for Tunceli Station's SPEI indexes for test data

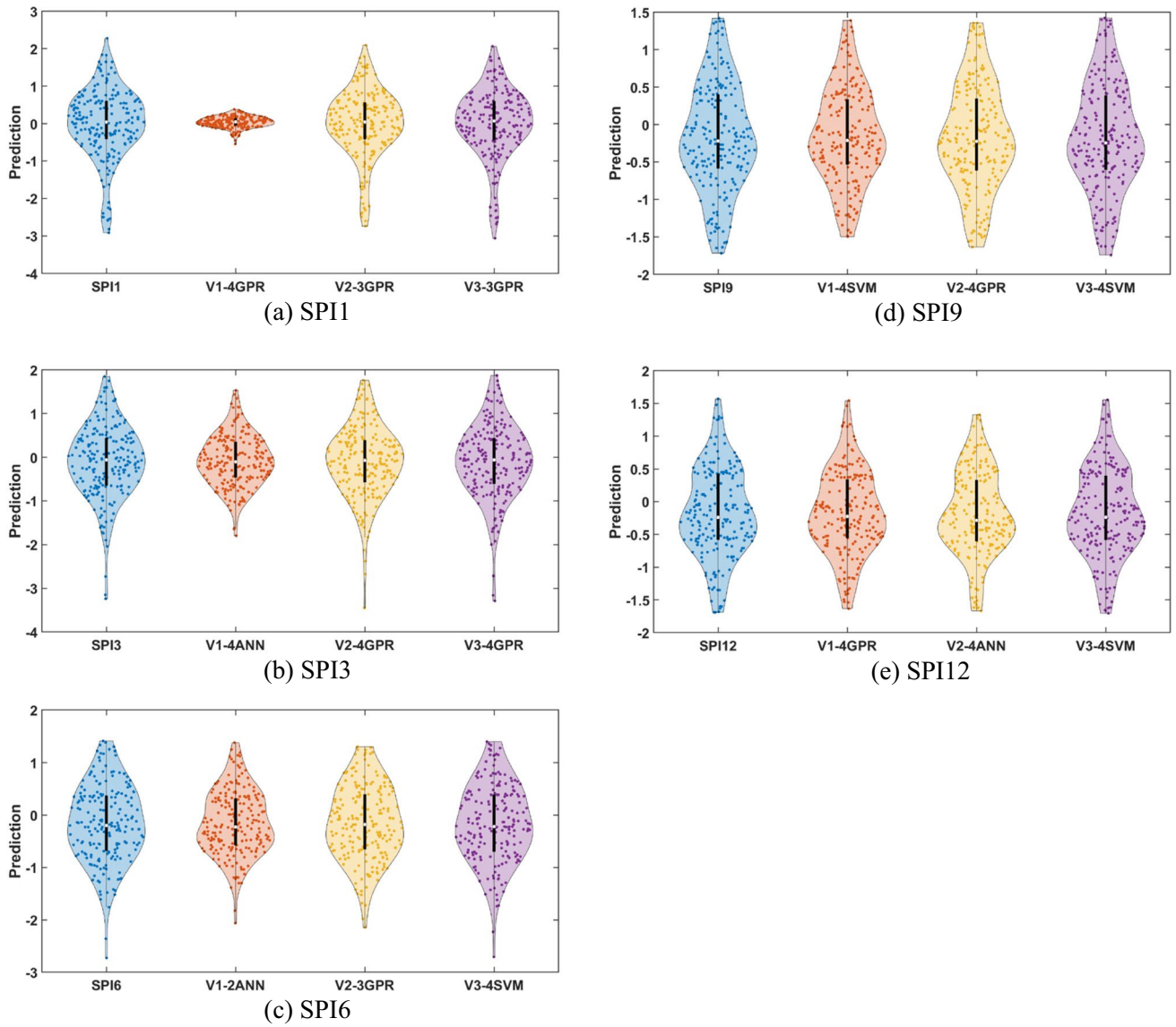


Fig. 24 Violin diagram for Tunceli Station's SPI indexes for test data

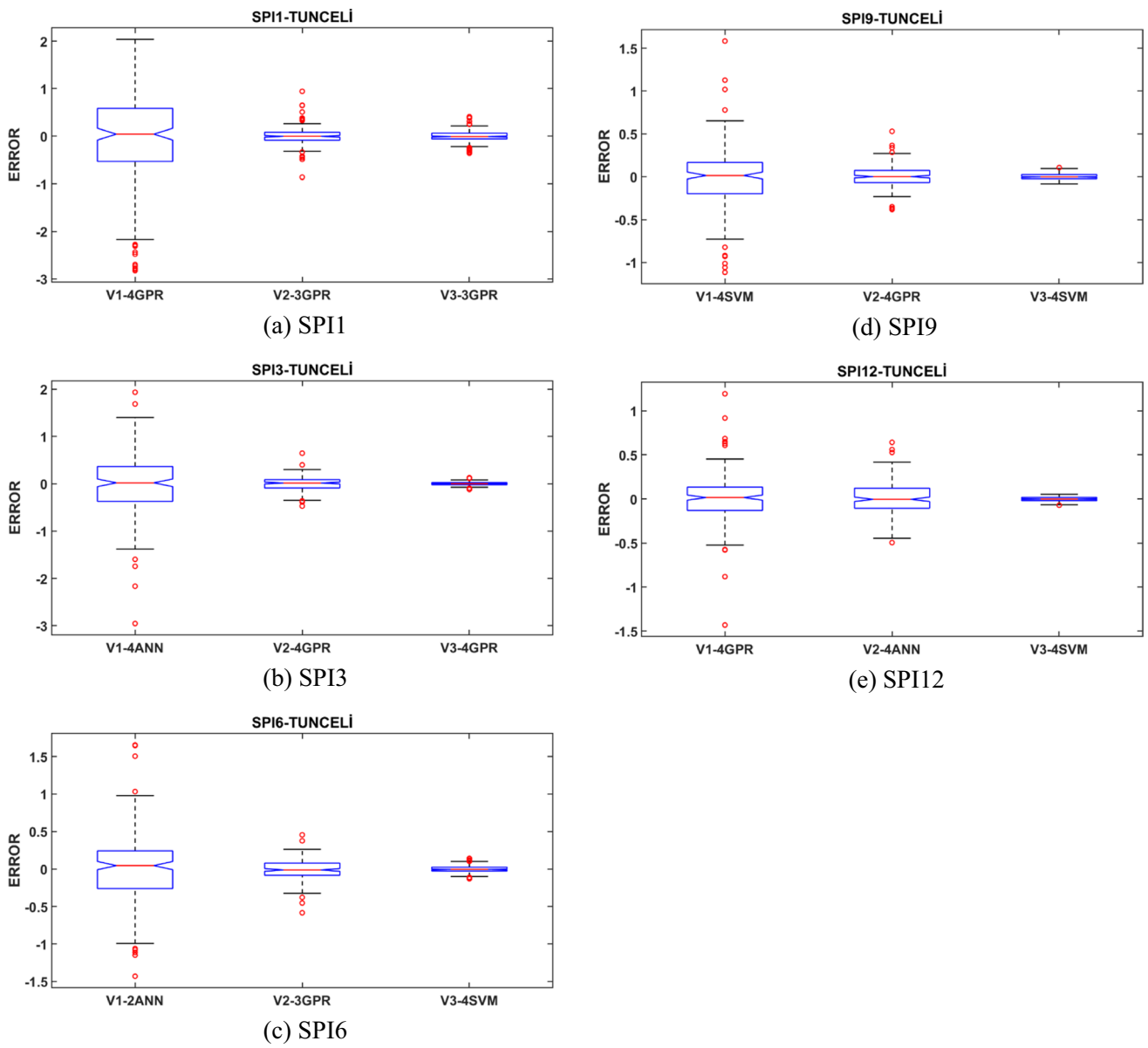


Fig. 25 Error-box plot for Tunceli Station's SPI indexes for test data

Table 12 Pseudo Code for Hybrid Model

```
Initialize variables and parameters
For each machine learning model (ANN, SVR, GPR):
  Set the model name based on the loop iteration
  For each SPI data type (SPI1, SPI3, SPI6, SPI9, SPI12):
    For each lag value (1 to 4):
      Read data from the file
      Set SPI data based on the selected type
      Split the data into training (70%) and testing (30%) sets
      Normalize the training and testing data:
        Detect and handle outliers in the training data:
          Use a method (e.g., z-score) to identify outliers
          For each outlier, replace it with the average of its previous and next values
        Apply TQWT to decompose the training and testing data into subbands
      For each subband:
        Prepare the input and output data for training and testing based on the specified lag
        Train the model using the training data and optimize hyperparameters
        Save the best hyperparameters for the model
        Predict the output for the training and testing data using the trained model
        Denormalize the predictions:
          Calculate performance metrics (MSE, MAE, correlation coefficient, R-squared) for the model predictions
          Plot and save the observed and forecasted data for each subband
      Aggregate the predictions from all subbands to obtain the final forecasted data
      Calculate performance metrics for the aggregated predictions
      Plot and save the observed and forecasted data for the aggregated predictions
      Save the performance metrics and results
End
```

Table 13 Pseudo Code for Tribrid Model

```
Initialize variables and parameters
For each machine learning model (ANN, SVR, GPR):
  Set the model name based on the loop iteration
  For each SPI data type (SPI1, SPI3, SPI6, SPI9, SPI12):
    For each lag value (1 to 4):
      Read data from the file
      Set SPI data based on the selected type
      Split the data into training (70%) and testing (30%) sets
      Normalize the training and testing data:
        Detect and handle outliers in the training data:
          Use a method (e.g., z-score) to identify outliers
          For each outlier, replace it with the average of its previous and next values
      Apply TQWT to decompose the training and testing data into subbands
      Prepare the input and output data for training and testing based on the specified lag for each subband
      Combine the prepared data from all subbands
      Select features using FSRFtest:
        Rank features based on their statistical significance
        Select top features that contribute to 95% of the total importance
      Train the model using the selected features and optimize hyperparameters
      Save the best hyperparameters for the model
      Predict the output for the training and testing data using the trained model
      Denormalize the predictions:
        Multiply the predicted values by the standard deviation and add the mean of the original training data to obtain the actual values
      Calculate performance metrics (MSE, MAE, correlation coefficient, R-squared) for the model predictions
      Plot and save the observed and forecasted data for the aggregated predictions
      Save the performance metrics and results
End
```

Table 14 Optimized Hyperparameters for Machine Learning Models for Tunceli station SPI data

SPI	Model	Hyperparameters
SPI1	ModelV3-3GPR	<p>Subband1: Sigma: 0.00010036, BasisFunction: linear, KernelFunction: matern32, KernelScale: 0.023595, Standardize: true</p> <p>Subband2: Sigma: 0.093995, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: false</p> <p>Subband3: Sigma: 0.030197, BasisFunction: none, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true</p> <p>Subband4: Sigma: 0.037382, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: false</p> <p>Subband5: Sigma: 0.006064, BasisFunction: linear, KernelFunction: rationalquadratic, KernelScale: 0.0042276, Standardize: true</p> <p>Subband6: Sigma: 0.054773, BasisFunction: linear, KernelFunction: exponential, KernelScale: 1.6811, Standardize: true</p>
SPI3	ModelV3-4GPR	<p>Subband1: Sigma: 2.4181, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true</p> <p>Subband2: Sigma: 0.00010878, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: true</p> <p>Subband3: Sigma: 0.23979, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.023405, Standardize: false</p> <p>Subband4: Sigma: 0.14109, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true</p> <p>Subband5: Sigma: 0.00021258, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: true</p> <p>Subband6: Sigma: 0.00022727, BasisFunction: none, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true</p>
SPI6	ModelV3-4SVM	<p>Subband1: BoxConstraint: 62.899, KernelScale: -, Epsilon: 0.012722, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband2: BoxConstraint: 240.34, KernelScale: -, Epsilon: 0.0022281, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband3: BoxConstraint: 4.0754, KernelScale: -, Epsilon: 0.00012636, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband4: BoxConstraint: 79.502, KernelScale: -, Epsilon: 9.6547e-05, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband5: BoxConstraint: 61.096, KernelScale: -, Epsilon: 0.00017696, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband6: BoxConstraint: 72.131, KernelScale: -, Epsilon: 0.014428, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p>
SPI9	ModelV3-4SVM	<p>Subband1: BoxConstraint: 997.77, KernelScale: -, Epsilon: 0.0096886, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband2: BoxConstraint: 0.1828, KernelScale: -, Epsilon: 0.0014426, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband3: BoxConstraint: 130.31, KernelScale: -, Epsilon: 0.0040077, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband4: BoxConstraint: 3.7081, KernelScale: -, Epsilon: 0.0027411, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband5: BoxConstraint: 270.17, KernelScale: -, Epsilon: 0.00075659, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband6: BoxConstraint: 37.077, KernelScale: -, Epsilon: 0.0013042, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p>

Table 14 (continued)

SPI	Model	Hyperparameters
SPI12	ModelV3-4SVM	<p>Subband1: BoxConstraint: 973.6, KernelScale: -, Epsilon: 9.0973e-05, KernelFunction: polynomial, PolynomialOrder: 3, Standardize: false</p> <p>Subband2: BoxConstraint: 1.3232, KernelScale: -, Epsilon: 0.001128, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband3: BoxConstraint: 913.18, KernelScale: -, Epsilon: 0.00012501, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband4: BoxConstraint: 205.42, KernelScale: -, Epsilon: 0.0020522, KernelFunction: polynomial, PolynomialOrder: 3, Standardize: false</p> <p>Subband5: BoxConstraint: 0.5319, KernelScale: -, Epsilon: 0.00024295, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband6: BoxConstraint: 205.71, KernelScale: -, Epsilon: 0.0085808, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p>

Table 15 Optimized Hyperparameters for Machine Learning Models for Tunceli station SPEI data

SPI	Model	Hyperparameters
SPEI1	ModelV3-4ANN	<p>Subband1: NumLayers: 1, Activations: relu, Standardize: true, Lambda: 2.6214e-05, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 1, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband2: NumLayers: 2, Activations: none, Standardize: true, Lambda: 9.5443e-05, LayerWeightsInitializer: he, LayerBiasesInitializer: ones, Layer_1_Size: 42, Layer_2_Size: 1, Layer_3_Size: -</p> <p>Subband3: NumLayers: 1, Activations: none, Standardize: true, Lambda: 8.8991e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: ones, Layer_1_Size: 6, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband4: NumLayers: 2, Activations: tanh, Standardize: false, Lambda: 4.6464e-07, LayerWeightsInitializer: he, LayerBiasesInitializer: zeros, Layer_1_Size: 1, Layer_2_Size: 41, Layer_3_Size: -</p> <p>Subband5: NumLayers: 2, Activations: none, Standardize: false, Lambda: 1.6448e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 12, Layer_2_Size: 271, Layer_3_Size: -</p> <p>Subband6: NumLayers: 1, Activations: relu, Standardize: false, Lambda: 5.8551e-06, LayerWeightsInitializer: he, LayerBiasesInitializer: zeros, Layer_1_Size: 15, Layer_2_Size: -, Layer_3_Size: -</p>
SPEI3	ModelV3-4SVM	<p>Subband1: BoxConstraint: 42.452, KernelScale: -, Epsilon: 0.00019191, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband2: BoxConstraint: 259.76, KernelScale: -, Epsilon: 0.0053228, KernelFunction: linear, PolynomialOrder: -, Standardize: false</p> <p>Subband3: BoxConstraint: 0.65564, KernelScale: -, Epsilon: 0.00097464, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband4: BoxConstraint: 182, KernelScale: -, Epsilon: 0.010144, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p> <p>Subband5: BoxConstraint: 9.7561, KernelScale: -, Epsilon: 0.0011977, KernelFunction: polynomial, PolynomialOrder: 2, Standardize: true</p> <p>Subband6: BoxConstraint: 21.729, KernelScale: -, Epsilon: 0.0050129, KernelFunction: linear, PolynomialOrder: -, Standardize: true</p>
SPEI6	ModelV3-3GPR	<p>Subband1: Sigma: 0.081961, BasisFunction: linear, KernelFunction: ardmatern52, KernelScale: -, Standardize: false</p> <p>Subband2: Sigma: 0.5997, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: false</p> <p>Subband3: Sigma: 0.0001292, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0039853, Standardize: true</p> <p>Subband4: Sigma: 0.30387, BasisFunction: linear, KernelFunction: ardmatern32, KernelScale: -, Standardize: true</p> <p>Subband5: Sigma: 0.0025498, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0011036, Standardize: true</p> <p>Subband6: Sigma: 0.025046, BasisFunction: none, KernelFunction: ardmatern32, KernelScale: -, Standardize: false</p>
SPEI9	ModelV3-4ANN	<p>Subband1: NumLayers: 2, Activations: sigmoid, Standardize: true, Lambda: 3.3861e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 1, Layer_2_Size: 43, Layer_3_Size: -</p> <p>Subband2: NumLayers: 2, Activations: none, Standardize: true, Lambda: 4.4021e-08, LayerWeightsInitializer: glorot, LayerBiasesInitializer: ones, Layer_1_Size: 4, Layer_2_Size: 55, Layer_3_Size: -</p> <p>Subband3: NumLayers: 1, Activations: tanh, Standardize: false, Lambda: 2.2429e-08, LayerWeightsInitializer: he, LayerBiasesInitializer: zeros, Layer_1_Size: 2, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband4: NumLayers: 2, Activations: none, Standardize: true, Lambda: 2.4968e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 47, Layer_2_Size: 115, Layer_3_Size: -</p> <p>Subband5: NumLayers: 1, Activations: relu, Standardize: false, Lambda: 3.202e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: ones, Layer_1_Size: 121, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband6: NumLayers: 2, Activations: none, Standardize: false, Lambda: 2.1172e-08, LayerWeightsInitializer: he, LayerBiasesInitializer: ones, Layer_1_Size: 4, Layer_2_Size: 144, Layer_3_Size: -</p>

Table 15 (continued)

SPI	Model	Hyperparameters
SPEI12	ModelV2-4GPR	Sigma: 0.001036, BasisFunction: pureQuadratic, KernelFunction: exponential, KernelScale: 2.9023, Standardize: true
	ModelV3-4GPR	<p>Subband1: Sigma: 0.2897, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: false</p> <p>Subband2: Sigma: 0.056197, BasisFunction: linear, KernelFunction: ardmatern32, KernelScale: -, Standardize: false</p> <p>Subband3: Sigma: 0.0039347, BasisFunction: linear, KernelFunction: rationalquadratic, KernelScale: 0.26083, Standardize: true</p> <p>Subband4: Sigma: 0.0063754, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.00076651, Standardize: true</p> <p>Subband5: Sigma: 0.0005224, BasisFunction: linear, KernelFunction: matern32, KernelScale: 0.091704, Standardize: false</p> <p>Subband6: Sigma: 1.2321, BasisFunction: pureQuadratic, KernelFunction: matern52, KernelScale: 0.13398, Standardize: false</p>

Table 16 Optimized Hyperparameters for Machine Learning Models for Tercan station SPI data

SPI	Model	Hyperparameters
SPI1	ModelV3-4GPR	<p>Subband1: Sigma: 0.081961, BasisFunction: linear, KernelFunction: ardmatern52, KernelScale: -, Standardize: false</p> <p>Subband2: Sigma: 0.5997, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: false</p> <p>Subband3: Sigma: 0.0001292, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0039853, Standardize: true</p> <p>Subband4: Sigma: 0.30387, BasisFunction: linear, KernelFunction: ardmatern32, KernelScale: -, Standardize: true</p> <p>Subband5: Sigma: 0.0025498, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0011036, Standardize: true</p> <p>Subband6: Sigma: 0.025046, BasisFunction: none, KernelFunction: ardmatern32, KernelScale: -, Standardize: false</p>
SPI3	ModelV2-4GPR	Sigma: 0.0022595, BasisFunction: linear, KernelFunction: exponential, KernelScale: 1.0345, Standardize: true
SPI6	ModelV3-4GPR	<p>Subband1: Sigma: 0.026299, BasisFunction: none, KernelFunction: ardmatern52, KernelScale: -, Standardize: true</p> <p>Subband2: Sigma: 0.01351, BasisFunction: linear, KernelFunction: ardexponential, KernelScale: -, Standardize: false</p> <p>Subband3: Sigma: 0.0083267, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.083388, Standardize: true</p> <p>Subband4: Sigma: 0.00040844, BasisFunction: linear, KernelFunction: ardexponential, KernelScale: -, Standardize: false</p> <p>Subband5: Sigma: 0.15225, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true</p> <p>Subband6: Sigma: 0.0012887, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.89336, Standardize: true</p>
SPI9	ModelV3-4ANN	<p>Subband1: NumLayers: 1, Activations: none, Standardize: false, Lambda: 2.3007e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 3, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband2: NumLayers: 1, Activations: tanh, Standardize: true, Lambda: 8.1343e-06, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 28, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband3: NumLayers: 1, Activations: tanh, Standardize: true, Lambda: 5.4029e-05, LayerWeightsInitializer: he, LayerBiasesInitializer: zeros, Layer_1_Size: 3, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband4: NumLayers: 2, Activations: none, Standardize: false, Lambda: 4.6831e-08, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 4, Layer_2_Size: 103, Layer_3_Size: -</p> <p>Subband5: NumLayers: 1, Activations: sigmoid, Standardize: true, Lambda: 1.6234e-06, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 48, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband6: NumLayers: 2, Activations: sigmoid, Standardize: false, Lambda: 2.104e-08, LayerWeightsInitializer: glorot, LayerBiasesInitializer: ones, Layer_1_Size: 8, Layer_2_Size: 33, Layer_3_Size: -</p>
SPI12	ModelV3-4GPR	<p>Subband1: Sigma: 0.00010762, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.037874, Standardize: true</p> <p>Subband2: Sigma: 0.075587, BasisFunction: linear, KernelFunction: squaredexponential, KernelScale: 0.032323, Standardize: false</p> <p>Subband3: Sigma: 0.0041516, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: false</p> <p>Subband4: Sigma: 0.0082025, BasisFunction: none, KernelFunction: ardmatern32, KernelScale: -, Standardize: true</p> <p>Subband5: Sigma: 0.0082025, BasisFunction: none, KernelFunction: ardmatern32, KernelScale: -, Standardize: true</p> <p>Subband6: Sigma: 0.00010035, BasisFunction: pureQuadratic, KernelFunction: matern52, KernelScale: 0.39109, Standardize: false</p>

Table 17 Optimized Hyperparameters for Machine Learning Models for Tercan station SPEI data

SPEI	Model	Hyperparameters
SPEI1	ModelV2-3GPR	Sigma: 0.0026904, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0034877, Standardize: true
	ModelV3-4GPR	<p>Subband1: Sigma: 5.2772, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0035171, Standardize: false</p> <p>Subband2: Sigma: 0.086128, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0018849, Standardize: true</p> <p>Subband3: Sigma: 0.001868, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.0028089, Standardize: false</p> <p>Subband4: Sigma: 1.4959, BasisFunction: linear, KernelFunction: rationalquadratic, KernelScale: 0.0030061, Standardize: false</p> <p>Subband5: Sigma: 0.00021076, BasisFunction: constant, KernelFunction: ardmatern52, KernelScale: -, Standardize: false</p> <p>Subband 6: Sigma: 0.0052899, BasisFunction: linear, KernelFunction: squaredexponential, KernelScale: 0.15722, Standardize: true</p>
SPEI3	ModelV3-4GPR	Subband 1: Sigma: 0.0077782, BasisFunction: linear, KernelFunction: ardexponential, KernelScale: -, Standardize: false
		Subband 2: Sigma: 0.18169, BasisFunction: linear, KernelFunction: ardmatern32, KernelScale: -, Standardize: true
		Subband 3: Sigma: 0.010885, BasisFunction: linear, KernelFunction: matern32, KernelScale: 0.082801, Standardize: false
		Subband 4: Sigma: 0.056423, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.046173, Standardize: false
		Subband 5: Sigma: 0.0016388, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.015278, Standardize: false
		Subband 6: Sigma: 0.00033675, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.16067, Standardize: false
SPEI6	ModelV3-4GPR	Subband1: Sigma: 0.00010886, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.0050868, Standardize: false
		Subband2: Sigma: 0.045138, BasisFunction: linear, KernelFunction: ardmatern32, KernelScale: -, Standardize: true
		Subband3: Sigma: 0.001327, BasisFunction: linear, KernelFunction: exponential, KernelScale: 0.059611, Standardize: true
		Subband4: Sigma: 0.00047403, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true
		Subband5: Sigma: 0.00015389, BasisFunction: linear, KernelFunction: matern52, KernelScale: 0.015701, Standardize: false
		Subband6: Sigma: 0.000121, BasisFunction: none, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: false
SPEI9	ModelV3-4GPR	Subband1: Sigma: 0.063203, BasisFunction: none, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: true
		Subband2: Sigma: 0.00049638, BasisFunction: linear, KernelFunction: ardsquaredexponential, KernelScale: -, Standardize: true
		Subband3: Sigma: 0.020183, BasisFunction: linear, KernelFunction: matern32, KernelScale: 0.017979, Standardize: false
		Subband4: Sigma: 0.0035356, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: true
		Subband5: Sigma: 0.0011987, BasisFunction: linear, KernelFunction: ardrationalquadratic, KernelScale: -, Standardize: false
		Subband6: Sigma: 0.00094898, BasisFunction: pureQuadratic, KernelFunction: matern52, KernelScale: 0.38533, Standardize: true

Table 17 (continued)

SPEI	Model	Hyperparameters
SPEI12	ModelV3-4ANN	<p>Subband1: NumLayers: 1, Activations: tanh, Standardize: false, Lambda: 7.412e-06, LayerWeightsInitializer: he, LayerBiasesInitializer: ones, Layer_1_Size: 194, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband2: NumLayers: 1, Activations: tanh, Standardize: true, Lambda: 2.6982e-05, LayerWeightsInitializer: he, LayerBiasesInitializer: ones, Layer_1_Size: 19, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband3: NumLayers: 2, Activations: none, Standardize: true, Lambda: 3.7272e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: ones, Layer_1_Size: 57, Layer_2_Size: 157, Layer_3_Size: -</p> <p>Subband4: NumLayers: 1, Activations: relu, Standardize: false, Lambda: 6.2578e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 264, Layer_2_Size: -, Layer_3_Size: -</p> <p>Subband5: NumLayers: 3, Activations: tanh, Standardize: false, Lambda: 1.8427e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 64, Layer_2_Size: 139, Layer_3_Size: 26</p> <p>Subband6: NumLayers: 2, Activations: none, Standardize: false, Lambda: 8.4422e-07, LayerWeightsInitializer: glorot, LayerBiasesInitializer: zeros, Layer_1_Size: 92, Layer_2_Size: 15, Layer_3_Size: -</p>

Table 18 The most successful methods of three different versions of Tercan Station and the performance criteria of these methods for training data

	V1	V2	V3		V1	V2	V3	
SPEI1	4GPR	3GPR	4GPR	SPI1	4GPR	4GPR	4GPR	RRMSE
	25.61	3.43	9.21		4.36	0.04	1.30	
	0.815	0.104	0.282		0.773	0.007	0.000	MAE
	0.030	0.983	0.875		0.032	0.999	0.999	R ²
	-0.201	0.984	0.901		-0.238	0.999	0.999	KGE
	0.425	0.979	0.905		0.425	0.999	1.000	OI
SPEI3	4SVM	3GPR	3GPR	SPI3	4SVM	4GPR	3GPR	RRMSE
	37.60	1.25	5.87		78.52	7.08	13.87	
	0.502	0.017	0.079		0.526	0.048	0.095	MAE
	0.599	0.999	0.990		0.552	0.996	0.986	R ²
	-22.685	0.998	0.621		-23.481	0.993	0.746	KGE
	0.734	0.998	0.985		0.725	0.994	0.984	OI
SPEI6	3SVM	4ANN	4GPR	SPI6	2ANN	4GPR	4GPR	RRMSE
	50.81	1.51	30.41		63.99	1.72	21.40	
	0.350	0.011	0.214		0.377	0.010	0.130	MAE
	0.791	0.999	0.925		0.767	0.999	0.974	R ²
	0.844	0.800	0.256		0.777	0.996	0.981	KGE
	0.846	0.998	0.933		0.839	0.999	0.972	OI
SPEI9	2SVM	4GPR	4GPR	SPI9	2GPR	4ANN	4ANN	RRMSE
	6.56	0.63	3.94		18.55	1.77	11.98	
	0.290	0.028	0.172		0.303	0.029	0.194	MAE
	0.864	0.999	0.951		0.856	0.999	0.940	R ²
	0.904	0.984	0.965		0.771	0.999	0.955	KGE
	0.893	0.996	0.952		0.894	0.996	0.948	OI
SPEI12	3SVM	3GPR	4ANN	SPI12	2ANN	4ANN	4GPR	RRMSE
	3.41	0.29	2.16		6.20	0.57	3.56	
	0.241	0.021	0.158		0.238	0.023	0.142	MAE
	0.896	0.999	0.958		0.900	0.999	0.967	R ²
	0.411	0.994	0.970		0.708	0.969	0.977	KGE
	0.918	0.997	0.961		0.923	0.997	0.968	OI

Table 19 The most successful methods of three different versions of Tunceli Station and the performance criteria of these methods for training data

	V1	V2	V3		V1	V2	V3	
SPEI1	4GPR	3GPR	4GPR	SPEI1	4GPR	4GPR	4GPR	
	116.92	15.70	37.80		1.79	0.82	0.57	RRMSE
	0.783	0.103	0.244		0.712	0.034	0.236	MAE
	0.063	0.984	0.902		0.034	0.998	0.991	R ²
	-3.007	0.985	0.929		-0.163	0.995	0.992	KGE
0.445	0.980	0.923	0.446	0.996	1.000	OI		
SPEI3	4SVM	3GPR	3GPR	SPEI3	4SVM	4GPR	3GPR	
	36.91	3.50	7.31		7.78	3.55	1.60	RRMSE
	0.499	0.047	0.095		0.531	0.025	0.110	MAE
	0.589	0.996	0.984		0.505	0.999	0.979	R ²
	0.662	0.993	-2.062		0.581	0.997	0.985	KGE
0.740	0.993	0.981	0.703	0.997	0.979	OI		
SPEI6	3SVM	4ANN	4GPR	SPEI6	2ANN	4GPR	4GPR	
	38.47	3.50	7.35		4.32	5.99	1.87	RRMSE
	0.523	0.047	0.098		0.372	0.052	0.165	MAE
	0.553	0.996	0.984		0.751	0.995	0.953	R ²
	0.636	0.993	0.825		0.799	0.991	0.190	KGE
0.719	0.993	0.981	0.843	0.993	0.963	OI		
SPEI9	2SVM	4GPR	4GPR	SPEI9	2GPR	4ANN	4ANN	
	4.11	4.73	2.11		2.31	2.70	1.11	RRMSE
	0.298	0.334	0.168		0.258	0.298	0.128	MAE
	0.834	0.790	0.956		0.870	0.828	0.970	R ²
	0.877	0.186	0.946		-0.369	0.197	0.437	KGE
0.890	0.858	0.964	0.913	0.886	0.974	OI		
SPEI12	3SVM	3GPR	4ANN	SPEI12	2ANN	4ANN	4GPR	
	3.08	2.51	0.23		1.26	1.12	0.96	RRMSE
	0.246	0.022	0.002		0.199	0.019	0.163	MAE
	0.868	0.999	1.000		0.918	0.999	0.953	R ²
	-0.057	0.998	1.000		0.940	0.999	0.621	KGE
0.908	0.997	1.000	0.937	0.998	0.959	OI		

Acknowledgements The authors thank the Turkish State Meteorological Service (MGM) for the statistics provided.

We would like to thank the Proofreading & Editing Office of the Dean for Research at Erciyes University for copyediting and proofreading service for this manuscript.

Authors' contributions Conceptualization: SB; Methodology: LL; Data collection: HC; Analysis: GA; Writing—original draft preparation: LL, GA; Writing—review and editing: SB; Supervision: HC.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Code availability The codes were developed from the MATLAB website.

Declarations

Ethical approval The authors paid attention to ethical rules in the study. There is no violation of ethics.

Consent to participate Not applicable.

Consent to publish Not applicable.

Competing interests The authors declare no competing interests.

References

- Achite M, Elshaboury N, Jehanzaib M et al (2023) Performance of machine learning techniques for meteorological drought forecasting in the Wadi Mina Basin, Algeria. *Water* 15:765. <https://doi.org/10.3390/w15040765>
- Adnan RM, Mostafa RR, Islam ARMT et al (2021) Estimating reference evapotranspiration using hybrid adaptive fuzzy inferencing coupled with heuristic algorithms. *Comput Electron Agric* 191:106541. <https://doi.org/10.1016/j.compag.2021.106541>
- Adnan RM, Dai H-L, Mostafa RR et al (2022) Modeling multistep ahead dissolved oxygen concentration using improved support vector machines by a hybrid metaheuristic algorithm. *Sustainability* 14:3470. <https://doi.org/10.3390/su14063470>
- Adnan RM, Mostafa RR, Dai H-L et al (2023) Pan evaporation estimation by relevance vector machine tuned with new metaheuristic

- algorithms using limited climatic data. *Eng Appl Comput Fluid Mech* 17. <https://doi.org/10.1080/19942060.2023.2192258>
- Afshar MH, Şorman AÜ, Tosunoğlu F et al (2020) Climate change impact assessment on mild and extreme drought events using copulas over Ankara, Turkey. *Theor Appl Climatol* 141:1045–1055. <https://doi.org/10.1007/s00704-020-03257-6>
- Aibaidula D, Ates N, Dadaser-Celik F (2022) Modelling climate change impacts at a drinking water reservoir in Turkey and implications for reservoir management in semi-arid regions. *Environ Sci Pollut Res* 30:13582–13604. <https://doi.org/10.1007/s11356-022-23141-2>
- Akbaş A (2014) Important drought years over Turkey. *Turkish J Geogr Sci* 12:101–118. https://doi.org/10.1501/Cogbil_0000000155
- Aktürk G, Çıtakoğlu H, Demir V, Beden N (2024) Meteorological drought analysis and regional frequency analysis in the Kızılırmak Basin: creating a framework for sustainable water resources management. *Water* 16:2124. <https://doi.org/10.3390/w16152124>
- Akturk G, Zeybekoglu U, Yildiz O (2022) Assessment of meteorological drought analysis in the Kızılırmak River Basin, Turkey. *Arab J Geosci* 15. <https://doi.org/10.1007/s12517-022-10119-0>
- Alemaw BF, Kileshye-Onema JM, Love D (2013) Regional drought severity assessment at a basin scale in the Limpopo drainage system. *J Water Resour Prot* 05:1110–1116. <https://doi.org/10.4236/jwarp.2013.511116>
- Alivi A, Yıldız O, Aktürk G (2021) Fırat-Dicle havzasında yıllık ortalama akımlar üzerinde iklim değişikliği etkilerinin iklim elastikiyeti metodu ile incelenmesi. *Gazi Üniversitesi Mühendislik Mimar Fakültesi Derg* 36:1449–1466. <https://doi.org/10.17341/gazimmfd.739556>
- Altunkaynak A, Jalilzadnezamabad A (2021) Extended lead time accurate forecasting of palmer drought severity index using hybrid wavelet-fuzzy and machine learning techniques. *J Hydrol* 601:126619. <https://doi.org/10.1016/j.jhydrol.2021.126619>
- Anshuka A, van Ogtrop FF, Willem Vervoort R (2019) Drought forecasting through statistical models using standardised precipitation index: a systematic review and meta-regression analysis. *Nat Hazards* 97:955–977. <https://doi.org/10.1007/s11069-019-03665-6>
- Bacanli UG, Firat M, Dikbas F (2009) Adaptive neuro-fuzzy inference system for drought forecasting. *Stoch Environ Res Risk Assess* 23:1143–1154. <https://doi.org/10.1007/s00477-008-0288-5>
- Başakın EE, Ekmekcioğlu Ö, Özger M (2021a) Drought prediction using hybrid soft-computing methods for semi-arid region. *Model Earth Syst Environ* 7:2363–2371. <https://doi.org/10.1007/s40808-020-01010-6>
- Başakın EE, Ekmekcioğlu Ö, Özger M, Altınbaş N, Şaylan L (2021b) Estimation of measured evapotranspiration using data-driven methods with limited meteorological variables. *Ital J Agrometeorol-Riv Ital Agrometeorologia* 1:63–80
- Bayram S, Ocal ME, Laptali Oral E, Atis CD (2016) Comparison of multi layer perceptron (MLP) and radial basis function (RBF) for construction cost estimation: the case of Turkey. *J Civ Eng Manag* 22:480–490. <https://doi.org/10.3846/13923730.2014.897988>
- Bayram S, Çıtakoğlu H (2023) Modeling monthly reference evapotranspiration process in Turkey: application of machine learning methods. *Environ Monit Assess* 195. <https://doi.org/10.1007/s10661-022-10662-z>
- Belayneh A, Adamowski J, Khalil B, Quilty J (2016) Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. *Atmos Res* 172–173:37–47. <https://doi.org/10.1016/j.atmosres.2015.12.017>
- Citakoglu H (2015) Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation. *Comput Electron Agric* 118:28–37. <https://doi.org/10.1016/j.compag.2015.08.020>
- Citakoglu H, Coşkun Ö (2022) Comparison of hybrid machine learning methods for the prediction of short-term meteorological droughts of Sakarya Meteorological Station in Turkey. *Environ Sci Pollut Res* 29:75487–75511. <https://doi.org/10.1007/s11356-022-21083-3>
- Citakoglu H, Cobaner M, Haktanir T, Kisi O (2014) Estimation of monthly mean reference evapotranspiration in Turkey. *Water Resour Manag* 28. <https://doi.org/10.1007/s11269-013-0474-1>
- Coşkun Ö, Citakoglu H (2023) Prediction of the standardized precipitation index based on the long short-term memory and empirical mode decomposition-extreme learning machine models: the case of Sakarya, Türkiye. *Phys Chem Earth* 131. <https://doi.org/10.1016/j.pce.2023.103418>
- Dabar OA, Adan A-BI, Ahmed MM et al (2022) Evolution and trends of meteorological drought and wet events over the Republic of Djibouti from 1961 to 2021. *Climate* 10:148. <https://doi.org/10.3390/cli10100148>
- Danandeh Mehr A, Sorman AU, Kahya E, Hesami Afshar M (2020) Climate change impacts on meteorological drought using SPI and SPEI: case study of Ankara, Turkey. *Hydrol Sci J* 65:254–268. <https://doi.org/10.1080/02626667.2019.1691218>
- Danandeh Mehr A, Rikhtehgar Ghiasi A, Yaseen ZM et al (2023) A novel intelligent deep learning predictive model for meteorological drought forecasting. *J Ambient Intell Humaniz Comput* 14:10441–10455. <https://doi.org/10.1007/s12652-022-03701-7>
- Deo RC, Kisi O, Singh VP (2017a) Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmos Res* 184:149–175. <https://doi.org/10.1016/j.atmosres.2016.10.004>
- Deo RC, Tiwari MK, Adamowski JF, Quilty JM (2017b) Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stoch Environ Res Risk Assess* 31:1211–1240. <https://doi.org/10.1007/s00477-016-1265-z>
- Docheshmeh GA, Alizamir M, Kisi O, Elshafie A (2022) Drought modelling by standard precipitation index (SPI) in a semi-arid climate using deep learning method: long short-term memory. *Neural Comput Appl* 34:2425–2442. <https://doi.org/10.1007/s00521-021-06505-6>
- Ekmekcioğlu Ö (2023) Drought forecasting using integrated variational mode decomposition and extreme gradient boosting. *Water* 15:3413. <https://doi.org/10.3390/w15193413>
- Esit M, Çelik R, Akbas E (2023) Long-term meteorological and hydrological drought characteristics on the lower Tigris-Euphrates basin, Türkiye: relation, impact and trend. *Environ Earth Sci* 82:1–19. <https://doi.org/10.1007/s12665-023-11182-w>
- Evkaya O, Kurnaz SF (2021) Forecasting drought using neural network approaches with transformed time series data. *J Appl Stat* 48:2591–2606. <https://doi.org/10.1080/02664763.2020.1867829>
- Felsche E, Ludwig R (2021) Applying machine learning for drought prediction in a perfect model framework using data from a large ensemble of climate simulations. *Nat Hazards Earth Syst Sci* 21:3679–3691. <https://doi.org/10.5194/nhess-21-3679-2021>
- Ferrari D, Yang Y (2015) Confidence sets for model selection by F-testing. *Stat Sin*. <https://doi.org/10.5705/ss.2014.110>
- Field CB (2012) Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the Intergovernmental Panel on Climate Change. Cambridge Univ Press Cambridge
- Figueiredo F, Britto D, Júnior S et al (2011) What is R2 all about? *Leviathan (São Paulo)* 60. <https://doi.org/10.11606/issn.2237-4485.lev.2011.132282>

- Ghasemi P, Karbasi M, Zamani Nouri A et al (2021) Application of Gaussian process regression to forecast multi-step ahead SPEI drought index. *Alex Eng J* 60:5375–5392. <https://doi.org/10.1016/j.aej.2021.04.022>
- Gholizadeh R, Yılmaz H, Danandeh Mehr A (2022) Multitemporal meteorological drought forecasting using Bat-ELM. *Acta Geophys*. <https://doi.org/10.1007/s11600-022-00739-1>
- Gul E, Staiou E, Safari MJS, Vaheddoost B (2023) Enhancing meteorological drought modeling accuracy using hybrid boost regression models: a case study from the Aegean Region, Türkiye. *Sustainability* 15:11568. <https://doi.org/10.3390/su151511568>
- Guttman NB (1998) Comparing the palmer drought index and the standardized precipitation index 1. *JAWRA J Am Water Resour Assoc* 34:113–121. <https://doi.org/10.1111/j.1752-1688.1998.tb05964.x>
- Haykin S (1998) *Neural networks: a comprehensive foundation*. Prentice Hall PTR
- Hearst MA, Dumais ST, Osuna E et al (1998) Support vector machines. *IEEE Intell Syst Their Appl* 13:18–28. <https://doi.org/10.1109/5254.708428>
- Ikram RMA, Mostafa RR, Chen Z et al (2023) Water temperature prediction using improved deep learning methods through reptile search algorithm and weighted mean of vectors optimizer. *J Mar Sci Eng* 11:259. <https://doi.org/10.3390/jmse11020259>
- Intergovernmental Panel on Climate Change (IPCC) (2014) *Climate change 2014 synthesis report summary for policymakers*.
- Irandoust M (2019) Saving and investment causality: implications for financial integration in transition countries of Eastern Europe. *Int Econ Econ Policy* 16:397–416. <https://doi.org/10.1007/s10368-017-0390-6>
- Jierula A, Wang S, Oh T-M, Wang P (2021) Study on accuracy metrics for evaluating the predictions of damage locations in deep piles using artificial neural networks with acoustic emission data. *Appl Sci* 11:2314. <https://doi.org/10.3390/app11052314>
- Jiménez-Gómez A, Flores-Márquez H (2023) Determinants of the remittances sent to Mexico 1980-2022: was there a structural change?. *EconoQuantum* 20(2):1–30
- Katip A (2018) Meteorological drought analysis using artificial neural networks for Bursa city, Turkey. *Appl Ecol Environ Res* 16:3315–3332. https://doi.org/10.15666/aeer/1603_33153332
- Katipoglu OM, Acar R, Şengül S (2020) Comparison of meteorological indices for drought monitoring and evaluating: a case study from Euphrates Basin, Turkey. *J Water Clim Chang* 11:29–43. <https://doi.org/10.2166/wcc.2020.171>
- Katipoglu OM, Acar R, Şenocak S (2021) Spatio-temporal analysis of meteorological and hydrological droughts in the Euphrates Basin, Turkey. *Water Sci Technol Water Supply* 21:1657–1673. <https://doi.org/10.2166/WS.2021.019>
- Katipoğlu OM (2023) Prediction of streamflow drought index for short-term hydrological drought in the semi-arid Yesilirmak Basin using Wavelet transform and artificial intelligence techniques. *Sustainability* 15:1109. <https://doi.org/10.3390/su15021109>
- Katipoğlu OM, Acar R (2022) Space-time variations of hydrological drought severities and trends in the semi-arid Euphrates Basin, Turkey. *Stoch Environ Res Risk Assess* 36:4017–4040. <https://doi.org/10.1007/s00477-022-02246-7>
- Katipoğlu OM, Acar R, Şenocak S, Şengül S (2022) Assessment of meteorological drought trends in the Euphrates Basin, Turkey. *Arab J Geosci* 15. <https://doi.org/10.1007/s12517-021-08482-5>
- Kikon A, Deka PC (2022) Forecasting of meteorological drought using machine learning algorithm. In: Rao CM, Patra KC, Jhajharia D, Kumari S (eds) *Advanced modelling and innovations in water resources engineering: select proceedings of AMIWARE 2021*. Springer Singapore, Singapore, pp 43–52
- Kilinc HC, Yurtsever A (2022) Short-term streamflow forecasting using hybrid deep learning model based on grey wolf algorithm for hydrological time series. *Sustainability* 14:3352. <https://doi.org/10.3390/su14063352>
- Kisi O, Docheshmeh Gorgij A, Zounemat-Kermani M et al (2019) Drought forecasting using novel heuristic methods in a semi-arid environment. *J Hydrol* 578:124053. <https://doi.org/10.1016/j.jhydrol.2019.124053>
- Knoben WJM, Freer JE, Woods RA (2019) Technical note: inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrol Earth Syst Sci* 23:4323–4331. <https://doi.org/10.5194/hess-23-4323-2019>
- Kumanlioglu AA (2020) Characterizing meteorological and hydrological droughts: a case study of the Gediz River Basin, Turkey. *Meteorol Appl* 27. <https://doi.org/10.1002/met.1857>
- Kurnaz L (2014) *Drought in Turkey*. Istanbul Policy Cent Sabancı Üniversitesi, İstanbul, Turkey
- Lalika C, Mujahid AUH, James M, Lalika MCS (2024) Machine learning algorithms for the prediction of drought conditions in the Wami River sub-catchment, Tanzania. *J Hydrol Reg Stud* 53:101794. <https://doi.org/10.1016/j.ejrh.2024.101794>
- Latifoğlu L (2022) Application of the novel circulant singular spectrum analysis ensemble model for forecasting of streamflow data. *Arab J Geosci* 15. <https://doi.org/10.1007/s12517-022-10230-2>
- Latifoğlu L, Özger M (2023) A novel approach for high-performance estimation of SPI data in drought prediction. *Sustainability* 15:14046. <https://doi.org/10.3390/su151914046>
- Li W, Duan L, Wang W et al (2021) Spatiotemporal characteristics of drought in a semi-arid grassland over the past 56 years based on the Standardized Precipitation Index. *Meteorol Atmos Phys* 133:41–54. <https://doi.org/10.1007/s00703-020-00727-4>
- Liu B, Wei Y, Zhang Y, Yang Q (2017) Deep neural networks for high dimension, low sample size data. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization, California, pp 2287–2293
- Lotfirad M, Esmaeili-Gisavandani H, Adib A (2022) Drought monitoring and prediction using SPI, SPEI, and random forest model in various climates of Iran. *J Water Clim Chang* 13:383–406. <https://doi.org/10.2166/wcc.2021.287>
- Loukas A, Vasilades L (2004) Probabilistic analysis of drought spatiotemporal characteristics in Thessaly region, Greece. *Nat Hazards Earth Syst Sci* 4:719–731. <https://doi.org/10.5194/nhess-4-719-2004>
- Marj AF, Meijerink AMJ (2011) Agricultural drought forecasting using satellite images, climate indices and artificial neural network. *Int J Remote Sens* 32:9707–9719. <https://doi.org/10.1080/01431161.2011.575896>
- McKee T, Doesken N, Kleist J (1993) The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*, Anaheim, California, 179–184. In: *Eighth Conference on Applied Climatology*. CA. American Meteorological Society, Boston, pp 17–22
- Mehr AD, Vaheddoost B, Mohammadi B (2020) ENN-SA: A novel neuro-annealing model for multi-station drought prediction. *Comput Geosci* 145:104622. <https://doi.org/10.1016/j.cageo.2020.104622>
- Mishra AK, Desai VR (2005) Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess* 19:326–339. <https://doi.org/10.1007/s00477-005-0238-4>
- Mishra N, Kushwaha A (2019) Rainfall prediction using gaussian process regression classifier. *Int J Adv Res Comput Eng Technol* 8:392–397
- Mishra AK, Singh VP (2010) A review of drought concepts. *J Hydrol* 391:202–216. <https://doi.org/10.1016/j.jhydrol.2010.07.012>
- Mishra AK, Singh VP (2011) Drought modeling – a review. *J Hydrol* 403:157–175. <https://doi.org/10.1016/j.jhydrol.2011.03.049>
- Mo KC, Schemm J-KE, Yoo S-H (2009) Influence of ENSO and the Atlantic multidecadal oscillation on drought over the United

- States. *J Clim* 22:5962–5982. <https://doi.org/10.1175/2009JCLI2966.1>
- Moghaddasi F, Moghaddasi M, Ghaleni MM, Yaseen ZM (2024) Fusion-based approach for hydrometeorological drought modeling: a regional investigation for Iran. *Environ Sci Pollut Res* 31:25637–25658. <https://doi.org/10.1007/s11356-024-32598-2>
- Mohamed Salleh MK, Ahmad Radib NF, Mohd Amin NA (2021) Spatial interpolation for missing rainfall data in Northern Region of Peninsular Malaysia. *J Phys Conf Ser* 1863:012049. <https://doi.org/10.1088/1742-6596/1863/1/012049>
- Mohammadi B (2023) Modeling various drought time scales via a merged artificial neural network with a firefly algorithm. *Hydrology* 10:58. <https://doi.org/10.3390/hydrology10030058>
- Mohammed Y, Yimer F, Tadesse M, Tesfaye K (2018) Meteorological drought assessment in north east highlands of Ethiopia. *Int J Clim Chang Strateg Manag* 10:142–160. <https://doi.org/10.1108/IJCCSM-12-2016-0179>
- Mokhtarzad M, Eskandari F, Jamshidi Vanjani N, Arabasadi A (2017) Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. *Environ Earth Sci* 76:729. <https://doi.org/10.1007/s12665-017-7064-0>
- Monfort B, Peña S (2008) Inflation determinants in Paraguay: cost push versus demand pull factors. *IMF Work Pap* 08:1. <https://doi.org/10.5089/9781451871289.001>
- Morid S, Smakhtin V, Moghaddasi M (2006) Comparison of seven meteorological indices for drought monitoring in Iran. *Int J Climatol* 26:971–985. <https://doi.org/10.1002/joc.1264>
- Mostafa RR, Kisi O, Adnan RM et al (2023) Modeling potential evapotranspiration by improved machine learning methods using limited climatic data. *Water* 15:486. <https://doi.org/10.3390/w15030486>
- Naumann G, Alfieri L, Wyser K et al (2018) Global changes in drought conditions under different levels of warming. *Geophys Res Lett* 45:3285–3296. <https://doi.org/10.1002/2017GL076521>
- Özger M, Başakın EE, Ekmekcioğlu Ö, Hacısüleyman V (2020) Comparison of wavelet and empirical mode decomposition hybrid models in drought prediction. *Comput Electron Agric* 179:105851. <https://doi.org/10.1016/j.compag.2020.105851>
- Patel K (2021) Turkey experiences intense drought. NASA Earth Observatory. Available online: <https://earthobservatory.nasa.gov/images/147811/turkey-experiences-intense-drought>. Accessed 29 Jan 2021
- Pei Z, Fang S, Wang L, Yang W (2020) Comparative analysis of drought indicated by the SPI and SPEI at various timescales in inner Mongolia, China. *Water (Switzerland)* 12:1925. <https://doi.org/10.3390/w12071925>
- Piri J, Abdollahipour M, Keshtegar B (2023) Advanced machine learning model for prediction of drought indices using hybrid SVR-RSM. *Water Resour Manag* 37:683–712. <https://doi.org/10.1007/s11269-022-03395-8>
- Rasmussen CE, Williams CKI (2005) Gaussian processes for machine learning. The MIT Press
- Reihanifar M, Danandeh Mehr A, Tur R et al (2023) A new multi-objective genetic programming model for meteorological drought forecasting. *Water* 15:3602. <https://doi.org/10.3390/w15203602>
- Reyniers N, Osborn TJ, Addor N, Darch G (2023) Projected changes in droughts and extreme droughts in Great Britain strongly influenced by the choice of drought index. *Hydrol Earth Syst Sci* 27:1151–1171. <https://doi.org/10.5194/hess-27-1151-2023>
- Schneider P, Xhafa F (2022) Anomaly detection and complex event processing over IoT data streams: with application to ehealth and patient data monitoring. Academic Press
- Selesnick IW (2011) Wavelet transform with tunable Q-factor. *IEEE Trans Signal Process* 59:3560–3575. <https://doi.org/10.1109/TSP.2011.2143711>
- Sen B, Topcu S, Türkeş M et al (2012) Projecting climate change, drought conditions and crop productivity in Turkey. *Clim Res* 52:175–191. <https://doi.org/10.3354/cr01074>
- Shabani S, Samadianfard S, Sattari MT et al (2020) Modeling pan evaporation using Gaussian process regression K-nearest neighbors random forest and support vector machines; comparative analysis. *Atmosphere (Basel)* 11:66. <https://doi.org/10.3390/atmos11010066>
- Shaowei Z, Hongchao Z, Pengcheng R et al (2013) Application of standardized precipitation evapotranspiration index in China. *Clim Environ Res* 18:617–625
- Shi P, Li G, Yuan Y et al (2019) Prediction of dissolved oxygen content in aquaculture using Clustering-based Softplus Extreme Learning Machine. *Comput Electron Agric* 157:329–338. <https://doi.org/10.1016/J.COMPAG.2019.01.004>
- Sihag P, Tiwari NK, Ranjan S (2017) Modelling of infiltration of sandy soil using Gaussian process regression. *Model Earth Syst Environ* 3:1091–1100. <https://doi.org/10.1007/s40808-017-0357-1>
- Smakhtin VU, Hughes DA (2004) Review, automated estimation and analyses of drought indices in South Asia. Working Paper 83. International Water Institute, Colombo, Sri Lanka
- Sobral BS, de Oliveira-Júnior JF, de Gois G et al (2019) Drought characterization for the state of Rio de Janeiro based on the annual SPI index: trends, statistical tests and its relation with ENSO. *Atmos Res* 220:141–154. <https://doi.org/10.1016/j.atmosres.2019.01.003>
- Soylu Pekpostalci D, Tur R, Danandeh Mehr A et al (2023) Drought monitoring and forecasting across Turkey: a contemporary review. *Sustainability* 15:6080. <https://doi.org/10.3390/su15076080>
- Stagge JH, Tallaksen LM, Xu CY, Van Lanen HAJ (2014) Standardized precipitation–evapotranspiration index (SPEI): Sensitivity to potential evapotranspiration model and parameters. *IAHS-AISH Proc Reports* 363:367–373
- Svoboda M, Hayes M, Wood D (2012) Standardized precipitation index: user guide, World Meteorological Organization, (WMO-No. 1090), Geneva
- SYGM (2021) Fırat-Dicle Havzası Taşkın Yönetim Planı. In: Republic Türkiye Ministry Of Agriculture and Forestry General Directorate of Water Management. <https://www.tarimorman.gov.tr/SYGM>
- Tao H, Borth H, Fraedrich K et al (2014) Drought and wetness variability in the Tarim River Basin and connection to large-scale atmospheric circulation. *Int J Climatol* 34:2678–2684. <https://doi.org/10.1002/joc.3867>
- Taye MM (2023) Understanding of machine learning with deep learning: architectures, workflow. *Appl Future Dir Comput* 12:91. <https://doi.org/10.3390/computers12050091>
- Thornthwaite CW (1948) An approach toward a rational classification of climate. *Geogr Rev* 38:55. <https://doi.org/10.2307/210739>
- Tirivarombo S, Osupile D, Eliasson P (2018) Drought monitoring and analysis: Standardised Precipitation Evapotranspiration Index (SPEI) and Standardised Precipitation Index (SPI). *Phys Chem Earth Parts A/B/C* 106:1–10. <https://doi.org/10.1016/j.pce.2018.07.001>
- Tsesmelis DE, Leveidioti I, Karavitis CA et al (2023) Spatiotemporal application of the standardized precipitation index (SPI) in the Eastern Mediterranean. *Climate* 11:95. <https://doi.org/10.3390/cli11050095>
- Türkeş M, Tatlı H (2009) Use of the standardized precipitation index (SPI) and a modified SPI for shaping the drought probabilities over Turkey. *Int J Climatol* 29:2270–2282. <https://doi.org/10.1002/joc.1862>
- Uncuoglu E, Citakoglu H, Latifoglu L et al (2022) Comparison of neural network, Gaussian regression, support vector machine, long short-term memory, multi-gene genetic programming, and M5 Trees methods for solving civil engineering problems. *Appl Soft Comput* 129:109623. <https://doi.org/10.1016/j.asoc.2022.109623>

- Uvidia-Cabadiana HA, Estrada-Jiménez PM, Herrera-Herrera R del C et al (2023) Adaptability of regression algorithms to protein Plant behavior. *Enfoque UTE*. <https://doi.org/10.29019/enfoqueute.861>
- Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J Clim* 23:1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>
- Wang L, Yu H, Yang M et al (2019) A drought index: the standardized precipitation evapotranspiration runoff index. *J Hydrol* 571:651–668. <https://doi.org/10.1016/j.jhydrol.2019.02.023>
- Wilhite DA (2000) *Drought: a global assessment, vol I*. Routledge Press, London New York
- Xu X, Zhang Y (2023) A Gaussian process regression machine learning model for forecasting retail property prices with Bayesian optimizations and cross-validation. *Decis Anal J* 8:100267. <https://doi.org/10.1016/j.dajour.2023.100267>
- Yaseen ZM, Ali M, Sharafati A et al (2021) Forecasting standardized precipitation index using data intelligence models: regional investigation of Bangladesh. *Sci Rep* 11:3435. <https://doi.org/10.1038/s41598-021-82977-9>
- Yuan X, Chen C, Lei X et al (2018) Monthly runoff forecasting based on LSTM–ALO model. *Stoch Environ Res Risk Assess* 32:2199–2212. <https://doi.org/10.1007/s00477-018-1560-y>
- Yuan X, Jian J, Jiang G (2016) Spatiotemporal variation of precipitation regime in China from 1961 to 2014 from the standardized precipitation index. *ISPRS Int J Geo-Inf* 5. <https://doi.org/10.3390/ijgi5110194>
- Zeybekoğlu U, Aktürk G (2021) A comparison of the China-Z Index (CZI) and the Standardized Precipitation Index (SPI) for drought assessment in the Hirfanli Dam basin in central Turkey. *Arab J Geosci* 14. <https://doi.org/10.1007/S12517-021-09095-8>
- Zhang Q, Qi T, Singh VP et al (2015) Regional frequency analysis of droughts in China: a multivariate perspective. *Water Resour Manag* 29:1767–1787. <https://doi.org/10.1007/s11269-014-0910-x>
- Zhang Y, Li H, Reggiani P (2019) Climate variability and climate change impacts on land surface. *Hydrol Process Water Manag Water* 11:1492. <https://doi.org/10.3390/w11071492>
- Zhou Y, Ning L, Zhong-hui J et al (2013) Temporal and spatial patterns of droughts based on standard precipitation index (SPI) in Inner Mongolia during 1981–2010. *J Nat Resour* 28:1694–1706

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.