ORIGINAL ARTICLE

Forecasting precipitation based on teleconnections using machine learning approaches across different precipitation regimes

Jalil Helali[1] · Milad Nouri[2] · Mehdi Mohammadi Ghaleni[3] · Seyed Asaad Hosseni[4] · Farshad Safarpour[5] · Azim Shirdeli[6] · Parisa Paymard[7] · Zahra Kalantari[8]

Abstract

Precipitation forecasts are of high significance for different disciplines. In this study, precipitation was forecasted using a wide range of teleconnection signals across different precipitation regimes. For this purpose, four sophisticated machine learning algorithms, i.e., the Generalized Regression Neural Network (GRNN), the Multi-Layer Perceptron (MLP), the Multi-Linear Regression (MLR), and the Least Squares Support Vector Machine (LSSVM), were applied to forecast seasonal and annual precipitation in 1- to 6-months lead times. To classify precipitation regimes, precipitation was clustered using percentiles. The indices quantifying El Niño-Southern Oscillation (ENSO) phasing showed the highest association with autumn, spring, and annual precipitation over the studied areas. The MLP and LSSVM algorithms provided satisfactory forecasts for almost all cases. However, our results indicated that the performance of LSSVM decreased in testing data, implying the tendency of this algorithm towards overfitting. The MLP showed a more balanced performance for the training and testing sets. Consequently, MLP seems best suited to be used for forecasting precipitation in our study area. The modeling algorithms provided less reliable forecasts for the regions corresponding to the 10–40th percentiles, mostly located in hyper-arid and arid environments. This underscores the inherent difficulty of precipitation forecasting in the hyper-arid and arid areas, wherein precipitation is very erratic and sparsely distributed. Our findings illustrate that clustering precipitation regimes to consider microclimate seems vital for reliable precipitation forecasting. Moreover, the results seem useful to design preventive drought/flood risk management strategies and to improve food-water security in Iran.

Keywords Drought · ENSO · Microclimate · Spatial inconsistency

✉ Milad Nouri
m.nouri@modares.ac.ir

Jalil Helali
jalilhelali@gmail.com

Mehdi Mohammadi Ghaleni
m-mohammadighaleni@araku.ac.ir

Seyed Asaad Hosseni
hosseini.asad8@gmail.com

Farshad Safarpour
farshadclima2008@gmail.com

Azim Shirdeli
shirdeli.azim@gmail.com

Parisa Paymard
parisa.paymard@tamu.edu

Zahra Kalantari
zahrak@kth.se

[1] Department of Irrigation and Reclamation Engineering, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

[2] Soil and Water Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran

[3] Department of Water Sciences and Engineering, Arak University, Arak, Iran

[4] Department of Climatology, Mohaghegh Ardabili University, Ardabil, Iran

[5] Department of Geography and Environmental Planning, Isfahan University, Isfahan, Iran

[6] Faculty of Agriculture, University of Zanjan, Zanjan, Iran

[7] Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843, USA

[8] Department of Sustainable Development, Environmental Science and Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

Introduction

Precipitation is the most important component of the hydrologic cycle on the globe, recharging renewable surface and ground water resources (Brutsaert 2005; Jin et al. 2021; Sospedra-Alfonso et al. 2015). As such, it is of paramount importance for domestic, industrial, and agricultural sectors, particularly in the context of changing climates (Medina et al. 2019; Myers et al. 2017; Qin et al. 2020). Precipitation anomalies, which can manifest as floods or droughts, may highly affect human life quality (Dai 2011; Petrucci 2022). Therefore, precipitation forecasts hold significant value in informing strategies aimed at adapting to drought and floods, ensuring food and water security, and promoting human health (Medina et al. 2019; Nguyen-Huy et al. 2017; Nouri et al. 2017a). Precipitation anomalies occur in association with local-scale microclimate anomalies and global teleconnections phasing. Climate change influences not only precipitation magnitude, but also its seasonality and type (Nouri and Homaee 2020, 2021b; Qin et al. 2020) through impacting the storm track activity and atmosphere–ocean teleconnection signals (Cai et al. 2015; Evans 2009; Trenberth 2020; Trenberth et al. 2013).

Teleconnection fluctuations and their correlation with precipitation provide a promising approach for forecasting precipitation amounts and anomalies as well as climate change impacts on precipitation. There are plenty of studies addressing the association between teleconnection signals and precipitation-related variables, e.g., dry/wet days, rainfall intensity, precipitation amount, precipitation extremity, precipitation duration, and precipitation type, on the globe (Casanueva et al. 2014; He and Guan 2013; Irannezhad et al. 2021; Nazemosadat and Cordery 2000; Nazemosadat and Ghasemi 2004; Nouri and Homaee 2021b; Skeeter et al. 2019; Helali et al. 2022; Xiao et al. 2017). Moreover, the precipitation behavior during different atmospheric teleconnection events (for instance the El Niño-Southern Oscillation, ENSO) has been leveraged to evaluate the variabilities of hydrological and biometeorological variables such as runoff and streamflow (Niu et al. 2014; Wang et al. 2022b), soil moisture (Nicolai-Shaw et al. 2016; Niu et al. 2014), groundwater level (Rust et al. 2018), heat waves (Choi et al. 2020; Jacques-Coper et al. 2021), evapotranspiration (Chai et al. 2018; Helali and Asadi Oskouei 2021; Zhao et al. 2020), land–atmosphere coupling strength (Holmes et al. 2017; Nouri and Homaee 2021a), and crop yield (Bannayan et al. 2011; Nouri et al. 2017a). The teleconnection signals have been also used to project the future changes in hydroclimatological variables (Rust et al. 2019; Yoon et al. 2015).

Spatial inconsistency in the relationship between precipitation and large-scale teleconnections refers to the phenomenon in which the relationship between precipitation and teleconnection is significant at certain points or regions but not at neighboring areas (Brown and Comrie 2004). This discrepancy can be a potential pitfall for understanding the complex interactions between precipitation and large-scale atmospheric phenomena. This can be explained by local impacts (particularly for point-scale studies) and/or the coupled effects exerted by other large-scale teleconnection patterns. Brown and Comrie (2004) explained spatial inconsistencies in the relationship of ENSO and winter precipitation by the Pacific Decadal Oscillation (PDO) phase shift over the western United States. Therefore, the phase of one teleconnection driver may modulate the impacts of another, affecting forecasting skill (Dannenberg et al. 2018; Theobald et al. 2018). This implicitly implies that forecasting precipitation based on one teleconnection driver seems to be highly uncertain, since it does not consider the coupled effects of teleconnections.

The investigation of the relationship between precipitation and teleconnections has been extended to regional scales, such as catchment-level studies, to overcome the spatial inconsistency observed in point-scale investigations (Räsänen and Kummu 2013; Sigaroodi et al. 2014; Xiao et al. 2015; Zhang et al. 2013, 2016a). However, conflicting results may still arise on regional scales. The relationships and forecasting models represent the average precipitation condition of a basin and might not be valid for each local site within that basin. This can be ascribed to the microclimate which can influence local precipitation regimes in a highly variable and non-uniform manner. Therefore, a sophisticated analytical approach is required to consider the microclimate impacts in different teleconnection events. Clustering techniques identify distinct groups of microclimate conditions associated with specific precipitation patterns (Roque-Malo and Kumar 2017). They can be particularly useful in reducing the uncertainty of precipitation forecasts, by providing a more detailed and nuanced understanding of the local factors that contribute to precipitation variability (Di et al. 2015; Gibson et al. 2021; Liu et al. 2020; Roque-Malo and Kumar 2017).

Various modeling approaches have been adopted to analyze the precipitation anomalies during different teleconnection events and to forecast precipitation. It is worth noting that modeling approach should physically explain the precipitation response to teleconnections. Linear approaches, such as linear regression, have been widely utilized in previous studies to associate precipitation-related variables with teleconnection indices (Hu et al. 2005; Kim et al. 2012; Wang et al. 2017, 2006a). This modeling approach is valid for the case that precipitation is linearly associated with teleconnection anomalies. However, it has been argued that the linear approaches seem unsuitable to be employed when

precipitation responds non-linearly to atmospheric teleconnection patterns (Chung and Power 2017; Kinouchi et al. 2018; Krishnaswamy et al. 2014). With recent advances in computational and data sciences, artificial intelligence has emerged as a promising tool for analyzing complex relationships with high inherent uncertainty in different fields (Boukabara et al. 2021; Shouval et al. 2021). Machine learning (ML) is a sub-domain of artificial intelligence that offers deep insights into complex non-linear data structures (Shouval et al. 2021). Use of machine learning is rapidly growing in weather forecasting and hydroclimatology literature (Arcomano et al. 2020; Ashley et al. 2019; Böhm et al. 2021; Chakraborty et al. 2021; Scher and Messori 2018). Recently, ML alternatives have been applied to forecast precipitation based on teleconnection signals (He et al. 2021). This study aimed to forecast the lagged responses of seasonal and annual precipitation to a wide range of teleconnection signals across different clustered precipitation regimes in Iran using sophisticated machine learning algorithms.

Methodology

Study area and data

Iran, situated in the Middle East, is a vast country characterized by a diverse range of climatic regimes, ranging from hyper-arid to humid. This remarkable climatic variety can largely be attributed to the presence of the Alborz Mountains in the north and the Zagros ranges in the west. The precipitating air masses blowing from the Caspian Sea are trapped by the Alborz, yielding sub-humid and humid conditions in northern Iran (Nouri and Homaee 2021b). The eastward Mediterranean humid air masses are also blocked by the Zagros Mountains, resulting in a semi-arid condition in western half of Iran. The arid and hyper-arid climates prevail in central, eastern, and southern Iran, where annual precipitation over 250 mm is infrequent (Nouri and Homaee 2021b). As a result, different seasonal and annual precipitation regimes can be observed in our study area. We clustered seasonal and annual precipitation based on the percentiles to portray different precipitation regimes and to consider the microclimate impacts across the studied sites (Fig. 1, and Table 1).

The Iran Ministry of Energy records precipitation data from more than 4360 locations (https://stu.wrm.ir/login.asp),

Fig. 1 Location of the studied sites (listed in Table 1 of supplementary material) classified in different precipitation percentiles
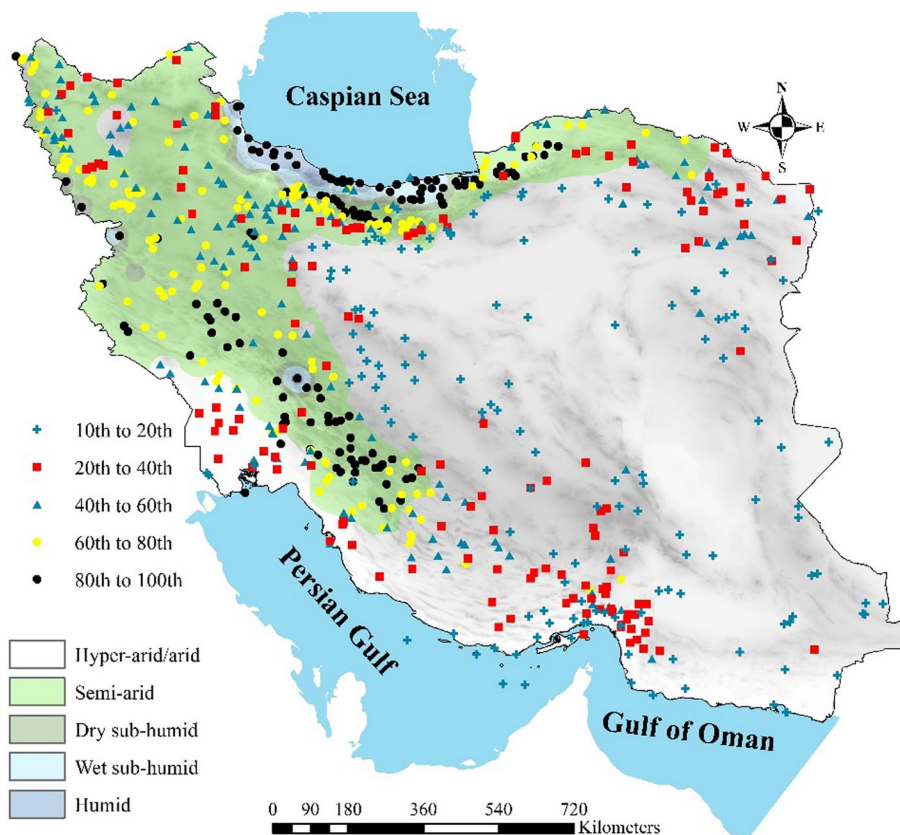
Table 1 Clustering criterion (Cr), the number of stations located in each cluster (SN), standard deviation (SD), and mean and coefficient of variation (CV) of seasonal and annual precipitation

No	Percentile	Autumn					Winter				
		Cr* (mm)	SN	SD	Mean	CV	Cr (mm)	SN	SD	Mean	CV
1	10th	28.4	73	14.0	19.7	71.0	57.9	72	14.4	45.2	31.8
2	20th	38.2	71	21.1	32.6	64.8	75.1	72	16.9	66.0	25.6
3	25th	43.0	36	22.7	40.7	55.8	82.2	36	26.3	78.8	33.4
4	30th	50.1	36	23.8	46.3	51.5	88.9	36	29.1	86.1	33.8
5	40th	66.0	71	22.4	57.8	38.7	102.4	72	31.5	96.2	32.8
6	50th	80.8	72	27.1	73.4	36.9	117.3	71	39.0	108.7	35.9
7	60th	94.4	72	36.9	87.7	42.1	134.6	72	45.3	126.1	35.9
8	70th	114.3	71	45.5	104.5	43.6	158.2	71	56.1	145.9	38.4
9	75th	124.3	36	52.1	118.6	44.0	173.6	37	49.6	166.7	29.7
10	80th	144.3	35	59.2	133.9	44.2	196.1	34	56.7	185.2	30.6
11	90th	194.1	73	69.9	165.9	42.2	255.8	72	59.1	223.3	26.4
12	95th	275.7	35	89.6	233.6	38.4	301.9	36	79.2	278.4	28.5
13	100th	813.0	36	96.2	453.0	21.2	653.6	36	138.8	409.6	33.9

No	Percentile	Spring					Annual				
		Cr (mm)	SN	SD	Mean	CV	Cr (mm)	SN	SD	Mean	CV
1	10th	20.2	73	16.0	15.4	104.3	135.3	72	34.4	102.1	33.7
2	20th	32.6	71	18.5	26.5	70.1	185.9	72	49.9	161.3	31.0
3	25th	38.1	38	19.8	35.3	56.0	202.2	37	75.3	194.8	38.7
4	30th	45.3	34	21.4	41.2	52.0	220.2	34	56.2	210.1	26.8
5	40th	63.1	72	27.0	53.8	50.1	253.9	72	54.1	237.5	22.8
6	50th	82.4	71	31.1	74.1	42.0	287.3	74	56.2	271.4	20.7
7	60th	96.2	71	33.3	89.7	37.1	328.3	69	65.7	307.1	21.4
8	70th	110.0	73	37.1	103.7	35.7	380.8	72	67.6	353.0	19.2
9	75th	117.4	36	37.6	113.0	33.2	416.2	36	80.6	398.8	20.2
10	80th	126.9	35	43.5	121.7	35.7	464.7	35	95.4	441.8	21.6
11	90th	152.7	71	44.8	139.1	32.2	653.2	72	109.0	549.9	19.8
12	95th	183.3	36	51.8	167.8	30.9	860.4	36	135.0	757.0	17.8
13	100th	305.4	36	62.6	215.3	29.1	1714.7	36	153.5	1178.0	13.0

*The criterion magnitude is the precipitation magnitude which the site grouped in a given percentile has the precipitation magnitude lower than that

and the Iran Meteorological Organization (IRIMO) collects precipitation data from more than 720 synoptic and climatology stations (https://data.irimo.ir/login/login.aspx) (Nouri and Homaee 2018, 2022; Saemian et al. 2021, 2022). We employed three main selection criteria: the dataset length, the presence of missing data, and the identification of outliers. The clustering facilitated the differentiation between outliers and anomalies. Specifically, precipitation spikes were observed in the data from some arid southern sites that initially appeared to be outliers but were, in fact, a typical behavior within specific clusters of these areas. Therefore, we classified such spikes as anomalies and retained the site. Nevertheless, the spikes that appeared unusual for a given cluster were deemed outliers and the corresponding sites were excluded from the analysis. Finally, the monthly precipitation data were gathered from 117 synoptic sites and 600 rain gauges with no missing data and outliers for the duration spanning from 1987 to 2015 (Table 1 in supplementary material).

The teleconnection indices are also listed in Table 2. These indices were retrieved from different sources given in Table 2 of supplementary material. For further details on the indices, one can refer to Hanley et al. (2003), Baldwin et al. (2001), Wang and Enfield (2001), Overland et al. (1999), Kutiel and Benaroch (2002), Barnston and Livezey (1987), and Wallace and Gutzler (1981). The datasets used to derive indices and their spatial resolution, as well as the latitude/longitude range of teleconnection patterns are listed in Table 3 of supplementary material. It is noteworthy that most indices were already computed by different

Table 2 List of teleconnection indices

No	Teleconnection indices	No	Teleconnection indices
1	AMO: Atlantic Multidecadal Oscillation	21	QBO: Quasi-Biennial Oscillation
2	AO: Arctic Oscillation	22	SCN: Scandinavia Index
3	CSST: Caspian Sea Surface Temperature	23	SFlux: Solar Flux
4	EA: East Atlantic pattern	24	SOI: Southern Oscillation Index
5	EAWR: Eastern Asia/Western Russia pattern	25	SSPOT: Sunspots
6	IOBSST: Indian Ocean Basin Sea Surface Temperature	26	SST1+2: Sea Surface Temperature in Niño1+2 region
7	MEI: Multivariate El Niño-Southern Oscillation Index	27	SST3: Sea Surface Temperature in Niño3 region
8	MSST: Mediterranean Sea Surface Temperature	28	SST3.4: Sea Surface Temperature in Niño3.4 region
9	NAO: North Atlantic Oscillation	29	SST4: Sea Surface Temperature in Niño4 region
10	NCP: North Sea-Caspian Sea Pattern	30	SSTas: Sea Surface Temperature anomalies in four Niño regions
11	Niño1+2: Extreme Eastern Tropical Pacific Sea Surface Temperature anomalies	31	SSTs: Sea Surface Temperature in four Niño regions
12	Niño3: Tropical Pacific Sea Surface Temperature anomalies	32	TNA: Tropical North Atlantic
13	Niño3.4: East Central Tropical Pacific Sea Surface Temperature anomalies	33	TNA-TSA: Tropical North and South Atlantic
14	Niño4: Central Tropical Pacific Sea Surface Temperature anomalies	34	TNI: Trans-Niño Index (standardized difference between Niño1+2 and Niño4)
15	NPI: North Pacific Index	35	TNIi: Difference between SST1+2 and SST4 indices
16	PDO: Pacific Decadal Oscillation	36	TPI: Tripole Index for the Interdecadal Pacific Oscillation
17	PNA: Pacific North American pattern	37	TSA: Tropical South Atlantic
18	POI: Pacific Ocean Index	38	WHWP: Western Hemisphere Warm Pool
19	POL: Polar/Eurasia pattern	39	WP: Western Pacific oscillation
20	PSST: Persian Gulf Sea Surface Temperature	40	WPSST: Western Pacific Sea Surface Temperature

[a]Niño1+2 region (0–10º S, 90º W–80º W)

[b]Niño3 region (5º N–5º S, 150º W–190º E)

[c]Niño3.4 region (5º N–5º S, 120º W–170º W)

[d]Niño4 region (5º N–5º S, 150º W–160º E)

sources, viz. National Center for Environmental Information (NCEI), Physical Sciences Laboratory (PSL), Climate Prediction Center (CPC) of National Oceanic and Atmospheric Administration (NOAA), National Center for Atmospheric Research (NCAR), Climatic Research Unit of University of East Anglia, Australian Bureau of Meteorology, and the Centre d'Estudis Ambientals del Mediterrani (CEAM). We, however, calculated Caspian Sea Surface Temperature (CSST), Indian Ocean Basin Sea Surface Temperature (IOBSST), Pacific Ocean Index (POI), Persian Gulf Sea Surface Temperature (PSST), and Western Pacific Sea Surface Temperature (WPSST) by applying gridded NOAA Optimum Interpolation SST v2 (https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html). These indices were equal to the average SST for the corresponding areas given in Table 3 of supplementary material.

In the current study, the aridity index map provided by Nouri and Homaee (2018) was utilized to classify the climates (Fig. 1). Accordingly, the sites corresponding to the 10–40th precipitation percentiles are mainly situated in central and eastern Iran and the strips of the Persian Gulf and the Gulf of Oman in the south of the country, with arid and hyper-arid climatic regimes (Fig. 1). The sites corresponding to the 40–60th percentiles are mostly located in northwestern semi-arid environments and southern and northeastern arid regions. The 60–80th percentiles dominantly comprise the northwestern semi-arid locations and the arid areas lying near the boundary between the semi-arid and arid climates in southern Iran. The 80–100th percentiles also encompass semi-arid, sub-humid, and humid areas in the central and northern Zagros and the sub-humid and humid areas located in the northern flanks of the Alborz. Accordingly, the precipitation percentiles are scattered across a broad range of climatic and topographic settings, illustrating that different precipitation regimes, and consequently microclimates, were included in our investigation (Fig. 1). Hence, the geographic and climatic indices appear not to be promising criteria to explain precipitation regimes. For instance, two sites, namely Arhan (No. 50 in Table 1 of supplementary material) and Saroo (No. 605 in Table 1 of supplementary material), both classified under the 70th percentile with an annual precipitation of approximately 360 mm, exhibited

significant differences in their climatic and geographic characteristics. Arhan is a cold semi-arid mountainous area situated at an elevation of 2028 m.a.s.l, while Saroo is an arid region located at an elevation of 1347 m.a.s.l. Despite these dissimilarities, the two sites were classified under the same precipitation regime, highlighting the limitations of relying solely on geographic and climatic indices for explaining precipitation patterns.

Summer precipitation was not considered in the current study, as it is less than 10% of annual precipitation in 80% of investigated sites. Given a high rate of evapotranspiration in summertime, water precipitated in summer appears not important to be forecasted in Iran.

Modeling framework

Pre-processing

The modeling approach consisted of three primary stages: the pre-processing, the forecasting, and statistical evaluation (Fig. 2). In the first phase, the correlations between seasonal and annual precipitation (as predictands) with 40 teleconnection indices (as candidate predictors in Table 2) were examined for 13 clusters. This was conducted in 1- to 6-months lagged times, resulting in the creation of matrices that encompassed a total of 240 associations. The predictands were autumn (October–November–December), winter (January–February–March), spring (April–May–June), and annual precipitation. It should be noted that the hydrologic year in Iran spans from 1st October to 30th September of the subsequent year (Nouri and Homaee 2020). The Pearson's correlation coefficient was used to evaluate the relationship between teleconnection indices and precipitation on seasonal and annual scales:

$$R = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 (y_i - \bar{y})^2}}, \tag{1}$$

where x_i is the independent variable (i.e., teleconnection indices), y_i is the dependent variable (precipitation), and $\bar{y}$ and $\bar{x}$ are the average of the dependent and independent variables, respectively.

The correlation matrix established in 1- to 6-month lead times is as follows:

$$R_{i,j} = \begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} & . & R_{1,13} \\ R_{2,1} & R_{2,2} & R_{2,3} & . & R_{2,13} \\ . & . & . & . & R_{3,13} \\ . & . & . & . & . \\ R_{40,1} & R_{40,2} & R_{40,3} & . & R_{40,13} \end{bmatrix}, \tag{2}$$

where i and j stand for the number of teleconnection indices (Table 2) and the number of percentiles (Table 1), respectively.

After constructing the correlation matrices, a three-step technique known as the forward selection was employed to identify the most significant predictors. This approach has been extensively utilized in the literature (Khan et al. 2007; Modaresi et al. 2018; Wang et al. 2006b). First, the predictor with the highest correlation coefficient was selected to model the predictand. Subsequently, other predictors were gradually introduced into the model in descending order of
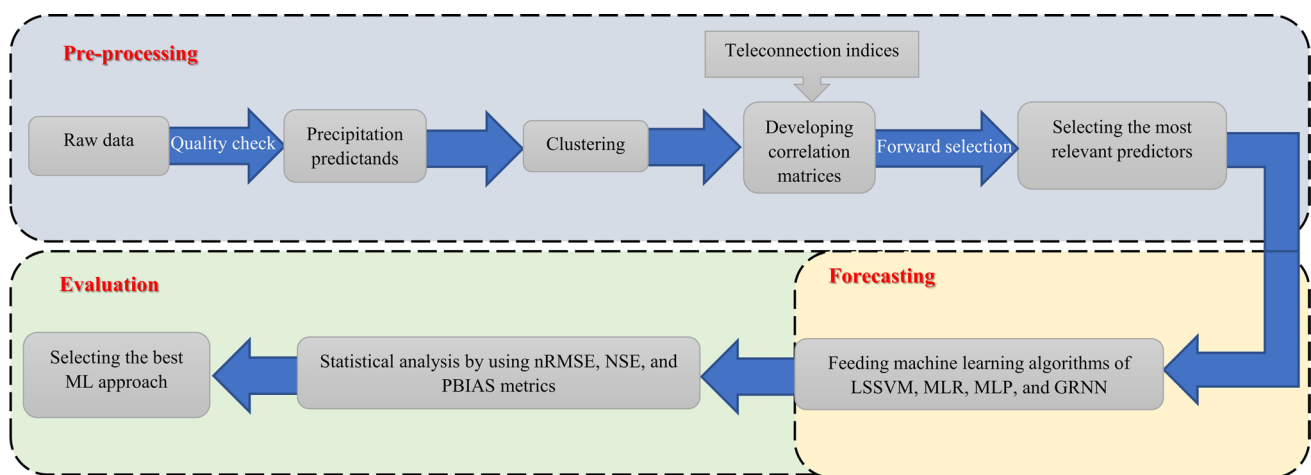

Fig. 2 The forecasting framework and stages based on the machine learning (ML) algorithms of the Generalized Regression Neural Network, GRNN, the Multi-Layer Perceptron, MLP, the Multi-Linear Regression, MLR, and the Least Squares Support Vector Machine, LSSVM. (The nRMSE, PBIAS and NSE denote normalized Root Mean Square Error, Percent Bias, and Nash–Sutcliffe Efficiency, respectively, explained in "Quantitative evaluation".)

Table 3 The most important teleconnection predictors applied to forecast seasonal and annual precipitation across different percentiles

Percentile	Autumn			Winter		
10th	$CSST_{Aug}$	$Niño4_{Apr}$	WP_{Jun}	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	WP_{Sep}
20th	$CSST_{Aug}$	NPI_{Aug}	$WHWP_{Apr}$	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	$MSST_{Dec}$
25th	$PSST_{Jul}$	$SSTas_{Sep}$	$WHWP_{Apr}$	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	$MSST_{Dec}$
30th	POL_{Jun}	$SST3.4_{Aug}$	$WHWP_{Apr}$	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	$MSST_{Dec}$
40th	AO_{Sep}	$PSST_{Jul}$	SOI_{Aug}	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	POL_{Dec}
50th	$MSST_{Sep}$	$Niño3.4_{Apr}$	SOI_{Jul}	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	$MSST_{Dec}$
60th	AMO_{Apr}	$MSST_{Sep}$	SOI_{Jul}	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	POL_{Dec}
70th	SOI_{Jul}	AO_{Sep}	PDO_{Jul}	PDO_{Aug}	$TNA\text{-}TSA_{Jul}$	TSA_{Oct}
75th	SOI_{Jul}	PDO_{Jul}	WP_{Aug}	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$	$MSST_{Dec}$
80th	SOI_{Jul}	AMO_{Apr}	$MSST_{Sep}$	PDO_{Aug}	TNA_{Jul}	TSA_{Oct}
90th	SOI_{Jul}	$CSST_{Jul}$	$WPSST_{Apr}$	$MSST_{Sep}$	POL_{Dec}	$TNA\text{-}TSA_{Jul}$
95th	PDO_{Jul}	SOI_{Aug}	TNA_{May}	POL_{Dec}	$MSST_{Sep}$	$TNA\text{-}TSA_{Jul}$
100th	PNA_{Apr}	SCN_{Sep}	SOI_{Jul}	POL_{Dec}	$TNA\text{-}TSA_{Jul}$	$NAO2_{Aug}$
Percentile	Spring			Annual		
10th	$CSST_{Dec}$	PNA_{Dec}	PNA_{Mar}	$CSST_{Aug}$	$NAO2_{Jul}$	NCP_{Apr}
20th	$CSST_{Dec}$	PNA_{Mar}	SOI_{Dec}	AO_{Jul}	$CSST_{Aug}$	$MSST_{May}$
25th	NPI_{Mar}	POL_{Jan}	$WPSST_{Mar}$	AO_{Jul}	$CSST_{Aug}$	$MSST_{May}$
30th	NPI_{Mar}	$NAO2_{Jan}$	$WPSST_{Mar}$	AO_{Jul}	$CSST_{Aug}$	$MSST_{May}$
40th	$Niño4_{Feb}$	NPI_{Mar}	$CSST_{Jan}$	$CSST_{Aug}$	PNA_{Aug}	TSA_{May}
50th	$NAO2_{Jan}$	$Niño4_{Feb}$	POL_{Oct}	$CSST_{May}$	$Niño4_{Jul}$	$PSST_{Jul}$
60th	$NAO2_{Jan}$	$Niño4_{Feb}$	$MSST_{Mar}$	PNA_{Aug}	$SST4_{Jul}$	$TNATSA_{Aug}$
70th	$NAO2_{Jan}$	POL_{Dec}	$SST4_{Nov}$	$CSST_{Jul}$	$Niño4_{Jul}$	$PSST_{Jul}$
75th	$NAO2_{Jan}$	TPI_{Mar}	$WHWP_{Mar}$	PNA_{Aug}	$SST4_{Jul}$	$TNA\text{-}TSA_{Aug}$
80th	$SST4_{Jan}$	POL_{Dec}	POL_{Jan}	PNA_{Aug}	SCN_{Jul}	$SST4_{Jul}$
90th	$MSST_{Mar}$	$NAO2_{Jan}$	$SST4_{Nov}$	EA_{Sep}	NCP_{Apr}	$SST4_{Jul}$
95th	$MSST_{Mar}$	$NAO2_{Jan}$	$SST4_{Nov}$	$EAWR_{Aug}$	$SST4_{Jul}$	WP_{Jul}
100th	$MSST_{Mar}$	$NAO2_{Jan}$	$SST4_{Nov}$	NCP_{Sep}	POI_{May}	TSA_{Aug}

their correlation coefficient. Finally, the three predictors that produced the minimum error were deemed the most appropriate combination (Table 3). Note that we experimentally realized that incorporating more than three predictors might elevate the risk of overfitting.
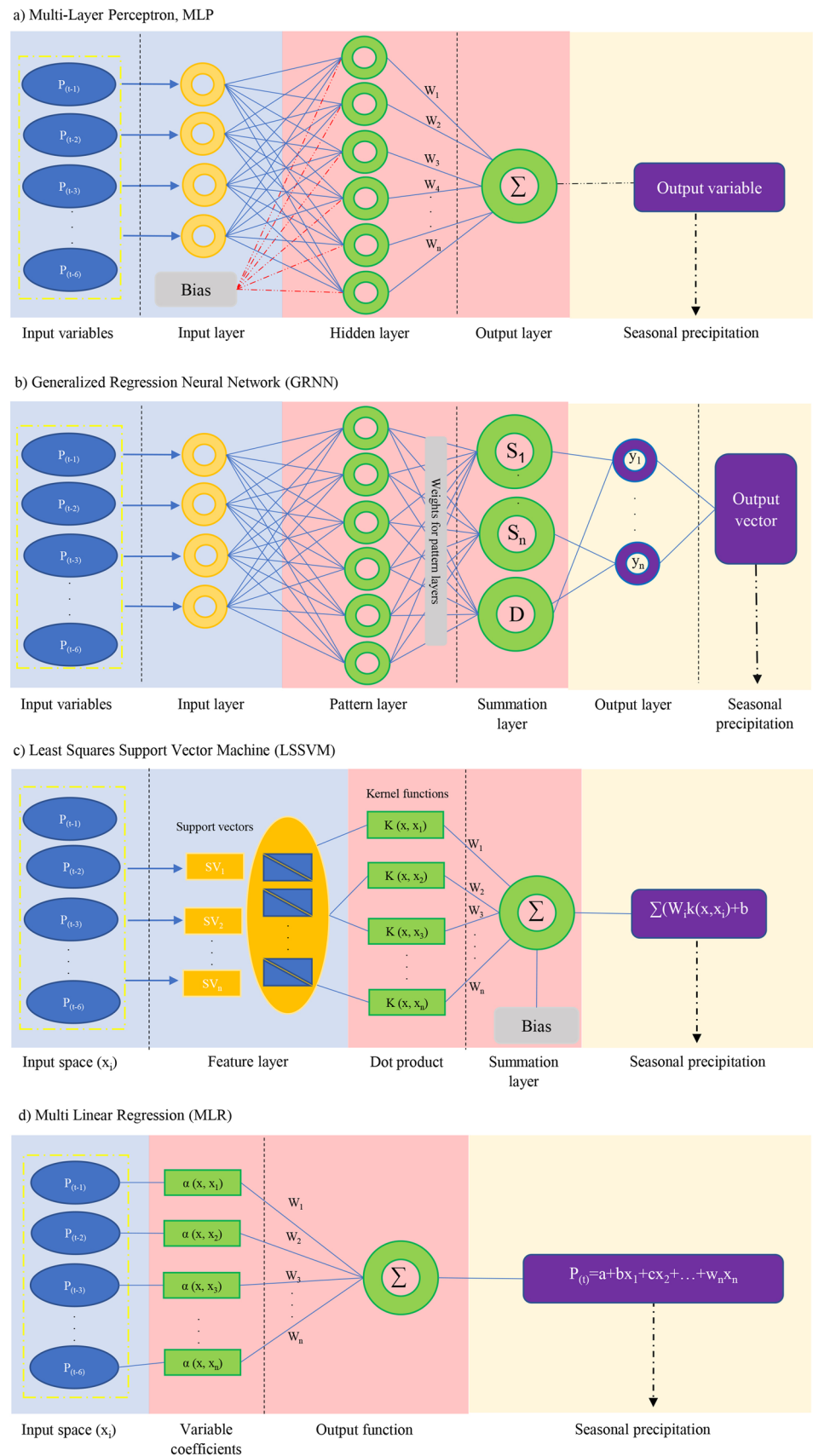
Forecasting

In the second step, the seasonal and annual precipitation was forecasted using four ML algorithms including the Generalized Regression Neural Network (GRNN), the Multi-Layer Perceptron (MLP), the Multi-Linear Regression (MLR), and the Least Squares Support Vector Machine (LSSVM). The MLP is a common Neural Network Model (ANN) with feed-forward network class including at least three layers of input, hidden, and output (Muni Kumar and Manjula 2019). It is composed of multiple interconnected nodes or neurons, arranged in a layered structure, where each neuron receives inputs from the previous layer and produces an output, which is transmitted to the next layer. The hidden layer of the MLP contains multiple neurons that are assigned weights and biases to optimize the model performance (Fig. 3a).

The GRNN is a kernel-based three-layer ANN in which the number of neurons in input (output) layers is equivalent to the input (output) vector dimensions (Lee and Resdi 2014). One of the key features of the GRNN is the use of a Radial Basis Function (RBF) in the pattern layer. The RBF is a type of activation function often employed in kernel-based methods. It is characterized by a center and a width, and it assigns a value to each input based on its distance from the center. The Gaussian kernel function is a common choice for the RBF due to its smoothness and symmetry. In the GRNN, the number of neurons in the input and output layers is fixed and equal to the dimension of the input and output vectors, respectively (Antanasijević et al. 2014; Antanasijevic et al. 2013). The number of neurons in the middle layer, however, is not fixed and is instead defined by observed data applied for calibration and validation (Fig. 3b).

The support vector machine (SVM) is a supervised ML applying the Structural Risk Minimization approach to

Fig. 3 The architecture of the machine learning algorithms ($\sum$ denotes the sum of weights for output layer in MLP, and the sum of weights for output functions in LSSVM, and the sum of weights for variable coefficients in MLR. S_1 to S_n are the summation units, and D stands for the division unit in GRNN. The weights (w) are given to layers in MPL and GRNN, to kernel functions in LSSVM, and to variable coefficients in MLR. SV denotes a support vector in LSSVM. $P_{(t-1)}$ and $P_{(t-6)}$ are the precipitation in 1- to 6-months time lags, respectively. $\alpha(x, x_i)$ is the variable coefficient of x_i in MLR, and $k(x, x_i)$ is the kernel function of x_i in LSSVM.)

minimize model error, whereas other methods such as artificial neural networks (ANN) use Empirical Risk Minimization principles (Cao et al. 2009; Kazem et al. 2013). The LSSVM is a variant of SVM that exploits linear equations in a forecasting algorithm. The LSSVM has been shown to achieve acceptable performance by applying an effective kernel function (Guo et al. 2012). The kernel function maps the input data into a higher-dimensional feature space, rendering LSSVM suitable for nonlinear problems (Fig. 3c).

The MLR is a supervised learning technique forecasting a continuous variable based on several independent variables. This algorithm uses the training dataset to estimate the coefficients of the linear equation that best fits the data (Jose et al. 2022; Najafi et al. 2011). The resulting model can then be applied to forecast the value of the dependent variable in unseen datasets. In the current work, the MLR algorithm using three independent variables was formulated to forecast seasonal and annual precipitation (Fig. 3d).

The data were split randomly into two subsets of training and testing. The training set comprised 70% of the data, while the testing set contained the remaining 30%. Two final outputs, one for each seasonal and annual timeframe, were generated for each model by averaging the results of the last 10 out of 30 iterations.

Quantitative evaluation

Three metrics of normalized Root Mean Square Error (nRMSE), Percent Bias (PBIAS), and Nash–Sutcliffe Efficiency (NSE) were employed to evaluate the forecasting skill of ML alternatives. The mathematical expressions of these statistics are

$$n\text{RMSE} = 100 \times \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(o_i - f_i\right)^2}}{\bar{o}}, \tag{3}$$

$$\text{PBAIS} = \frac{100}{\bar{o}} \times \frac{\sum_{i=1}^{n}\left(f_i - o_i\right)}{n}, \tag{4}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}\left(o_i - f_i\right)^2}{\sum_{i=1}^{n}\left(o_i - \bar{o}\right)^2}, \tag{5}$$

where f_i and o_i are, respectively, forecasts and observations, $\bar{o}$ denotes the average of the observed data, and n stands for the number of comparisons.

The nRMSE is oftentimes applied to quantify the absolute error of estimates. The performance of a model is unsatisfactory if nRMSE exceeds 30% (Dettori et al. 2011). The metric of PBIAS quantifies the model bias or systematic error. The positive (negative) PBIAS indicates the tendency of a modeling algorithm to overestimate (underestimate). The model forecast can be judged as satisfactory for the PBIAS values less than 25% (Moriasi et al. 2007). The NSE is used to evaluate the relative error, and a value lower than 0.5 suggests unreliable model performance (Moriasi et al. 2007). When forecasts match observations, the model performance is considered excellent and the values of nRMSE and PBIAS are equal to 0.0%, and the magnitude of NSE is 1.0. In this study, the forecasts with nRMSE of < 30%, NSE of > 0.5, and PBIAS of < 25% were deemed reliable.

Results and discussion

Predictors selection

Autumnal precipitation had a higher correlation with teleconnection indices in 1- to 3-months lagged periods (Fig. 4). Except for two indices, i.e., Tropical North Atlantic (TNA) and Atlantic Multidecadal Oscillation (AMO), the correlation coefficient was found to be insignificant for the 4- to 6-months lagged associations. Autumn precipitation showed a stronger association with the ENSO indices, such as the Southern Oscillation Index (SOI), Niño3.4, Sea Surface Temperature in four Niño regions (SST$_{as}$), Multivariate El Niño-Southern Oscillation Index (MEI), as well as Pacific Ocean Index (POI) in lead periods of 1 to 3 months (Fig. 4). Table 3 also displays that the indices quantifying ENSO activities, such as SOI, Niño3.4 and SST$_{as}$, were applied as three main teleconnection predictors to forecast autumn precipitation. The ENSO phasing has been shown to exert a strong impact on autumnal precipitation, particularly in western Iran (Helali et al. 2020; Nazemosadat and Cordery 2000; Nazemosadat and Ghasemi 2004; Nouri et al. 2017a). Winter precipitation was associated insignificantly ($p > 0.05$) with most teleconnection signals (Fig. 5). Nonetheless, winter precipitation showed a higher correlation with indices of Pacific Decadal Oscillation (PDO), and Mediterranean Sea Surface Temperature (MSST) in 4- to 6-months lead periods, as well as indices of Tropical South Atlantic (TSA), East Atlantic pattern (EA), SOI and Polar/Eurasia pattern (POL) in lagged times of 1–3 months. MSST and Tropical North and South Atlantic (TNA-TSA) were often considered as the main predictors of winter precipitation (Table 3). Unlike autumnal rainfall, sharp dry/wet epochs in different phases of ENSO are not anticipated in most regions in Iran, because ENSO impacts on winter precipitation are modulated by other teleconnections (Ghasemi and Khalili 2008; Nazemosadat and Ghasemi 2004).

There was a significant association between spring precipitation and Pacific Ocean Index (POI), WPSST, Tripole Index for the Interdecadal Pacific Oscillation (TPI), and ENSO-related indices, including SOI, Niño4, Sea Surface
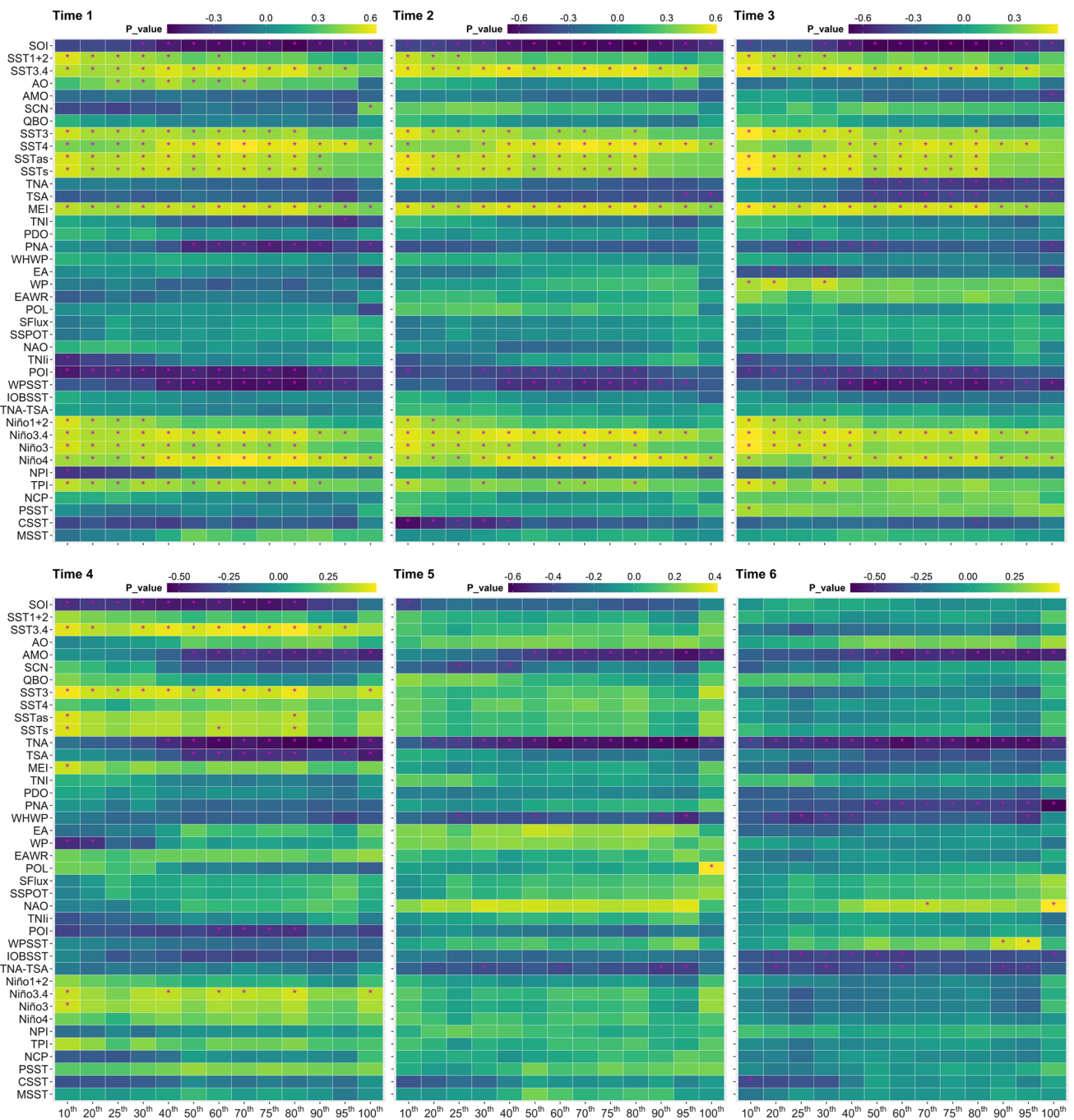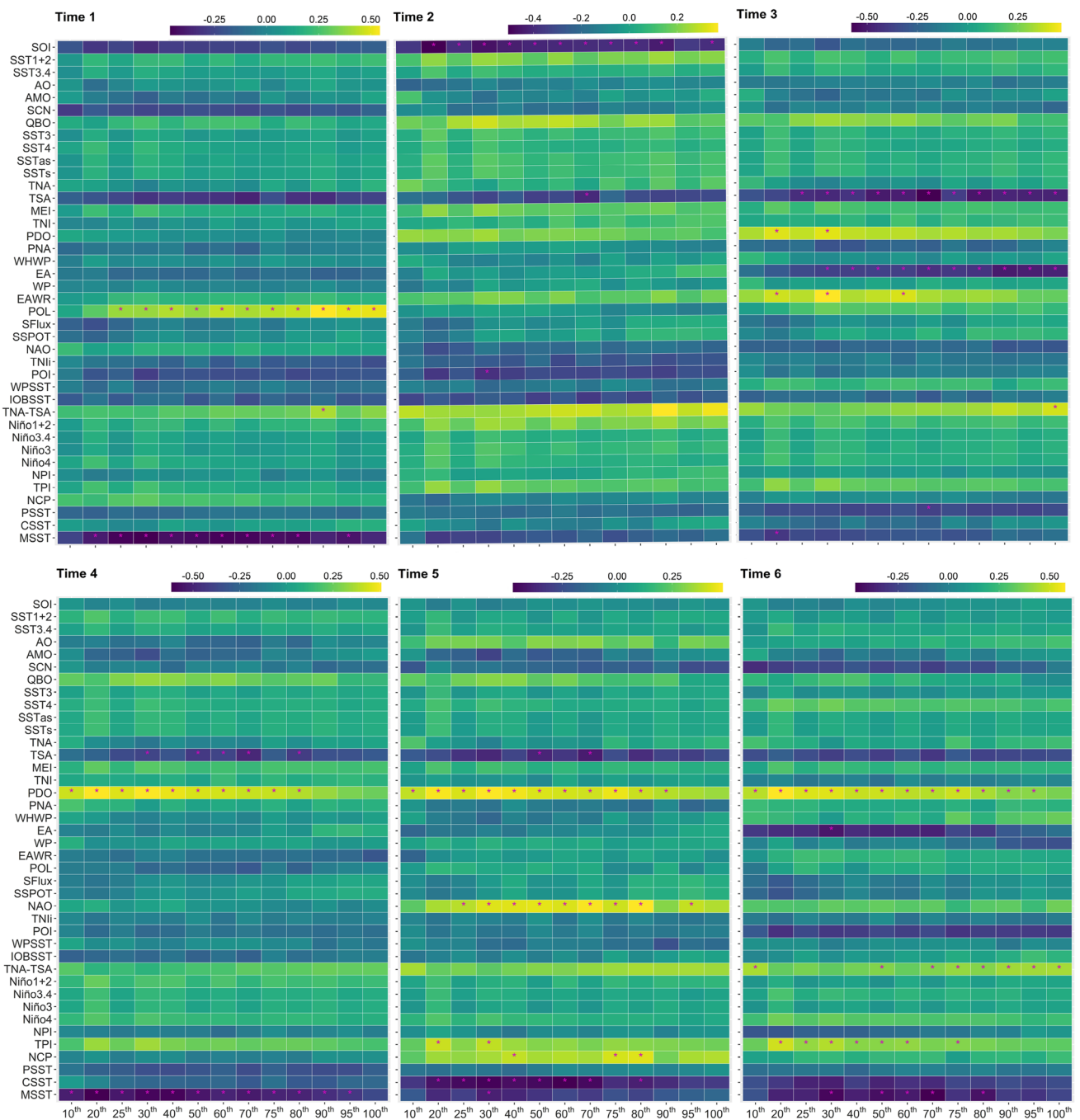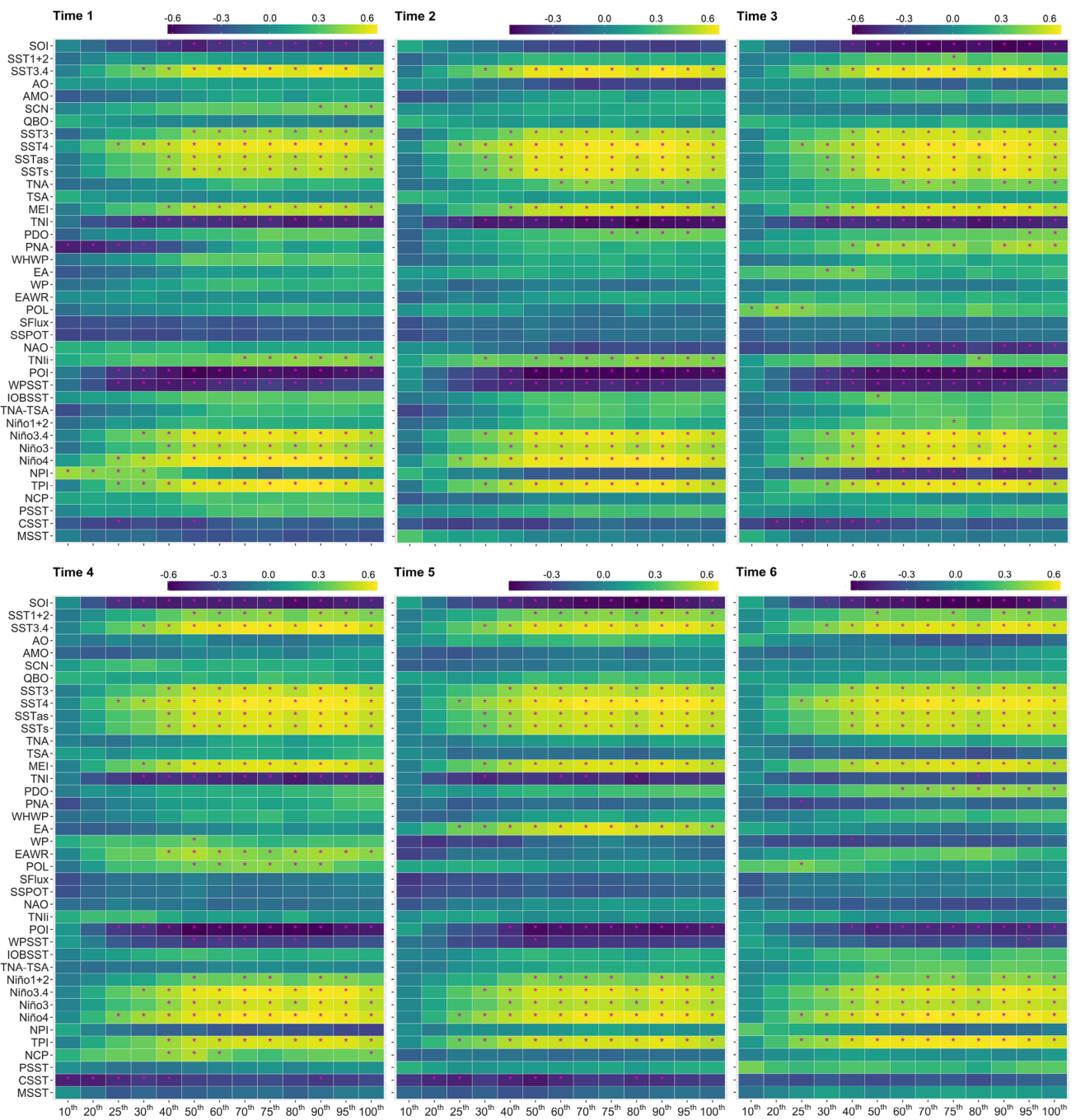
Fig. 4 The correlation coefficient of autumn precipitation-teleconnection indies for 1- to 6-months lead times (the asterisk indicates the significant correlation coefficient at the 95% confidence level)

Temperature in Niño4 region (SST4), and MEI (Fig. 6). Similar to autumnal precipitation, ENSO phasing seems to cause precipitation perturbations during spring, except for the 10–30th percentiles. The La Niña-triggered droughts are expected during springtime in western Iran (Ahmadi et al. 2019; Helali et al. 2021; Nouri and Homaee 2021a). However, ENSO anomalies do not explain precipitation variations occurring in hyper-arid/arid southern, eastern, and

central Iran, the regions mostly classified in the 10–40th percentiles (Fig. 1). Table 3 indicates that the ENSO-related indices, such as SOI, Niño4, and SST4, are among three main teleconnection indices applied to forecast spring precipitation. For most clusters, annual precipitation was associated significantly with ENSO-related indicators, POI, TPI, and WPSST (Table 3 and Fig. 7). Overall, ENSO seems to be one of the key teleconnections affecting precipitation
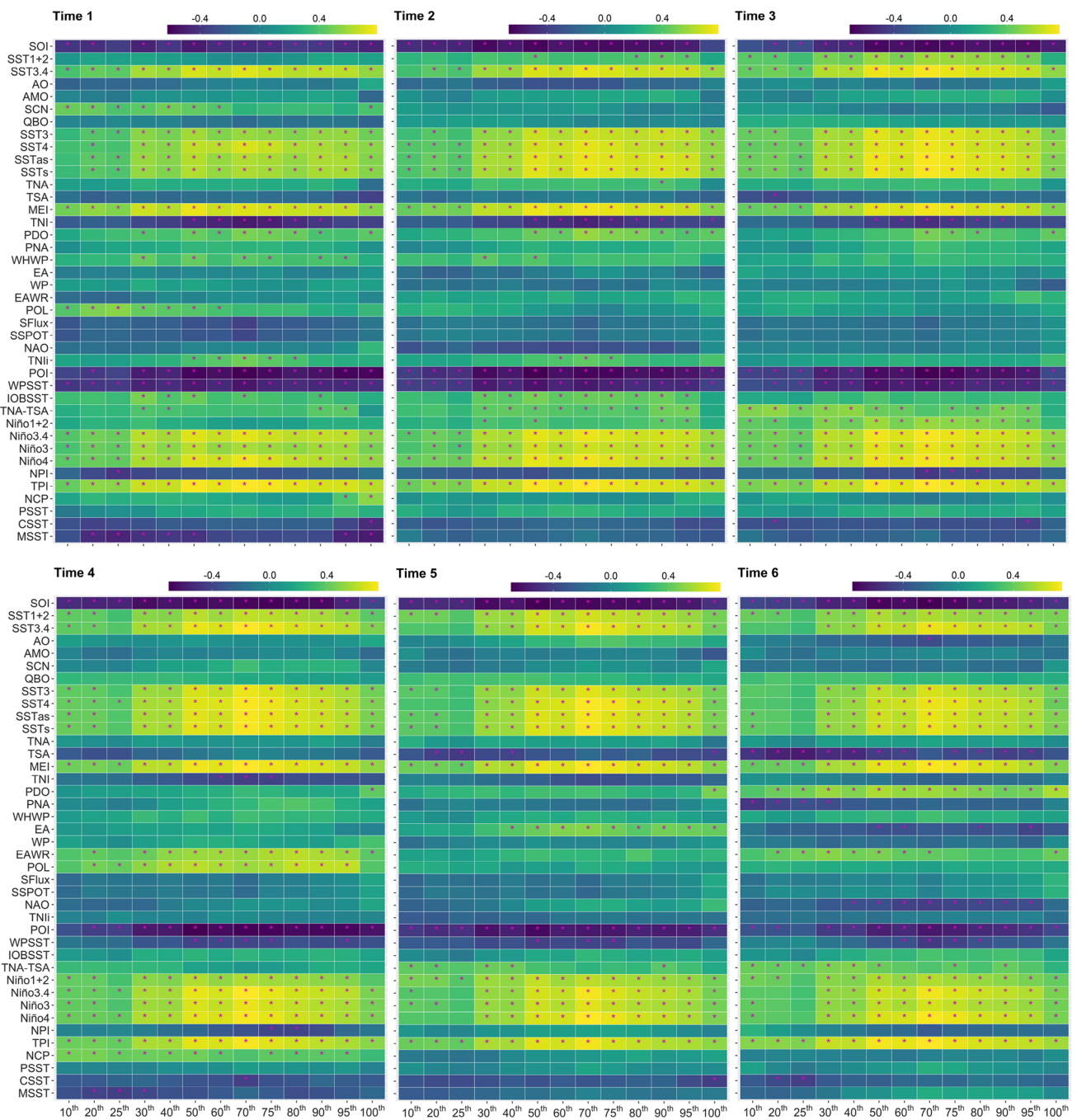
Fig. 5 The correlation coefficient of winter precipitation-teleconnection indies for 1- to 6-months lead times (the asterisk indicates the significant correlation coefficient at the 95% confidence level)

variabilities across the country (Table 3, and Figs. 4, 6, and 7).

Figure 8 shows the Pearson's correlation coefficient obtained for ML algorithms across different clusters in the training and testing steps. The GRNN, MLP, LSSVM, and MLR had an average Pearson's correlation coefficient of 0.85, 0.92, 0.99, and 0.82 in the training phase, and 0.65, 0.74, 0.48, and 0.74 in the testing step, respectively. The

results show that while LSSVM had superior performance in the training phase (Fig. 8a, c, e, and g), and its forecasting skill was severely impaired in the testing phase (Fig. 8b, d, f, and h). This discrepancy indicates a pronounced overfitting problem of LSSVM, despite a reasonable number of predictors included for training. In other words, LSSVM was too narrowly adjusted to the training dataset and could not effectively capture the underlying patterns in the unseen

Fig. 6 The correlation coefficient of spring precipitation-teleconnection indies for 1- to 6-month lead times (the asterisk indicates the significant correlation coefficient at the 95% confidence level)

dataset. This has been also argued in the ML literature (Peng and Wang 2009; Wei et al. 2008).

Machine learning modeling performance

The MLP and LSSVM had a reasonable forecasting skill for autumn precipitation in all clusters (Fig. 9b and c). However, the nRMSE of autumnal precipitation forecasted

by GRNN and MLR exceeded 30% for the 10th and 40th percentiles (Fig. 9c). Consequently, GRNN and MLR did not forecast autumnal precipitation reliably for the regions located in the 10th and 40th percentiles. The GRNN algorithm exhibited poor performance in forecasting wintertime precipitation, as evidenced by a NSE value of less than 0.5. However, MLP, LSSVM, and MLR algorithms demonstrated reliable performance across all percentiles for

Fig. 7 The correlation coefficient of annual precipitation-teleconnection indies for 1- to 6-months lead times (the asterisk indicates the significant correlation coefficient at the 95% confidence level)

forecasting winter precipitation, with nRMSE values below 30% and NSE values above 0.5 (Fig. 9e and f). The GRNN and MLR forecasted spring precipitation based on teleconnection indices unreliably (i.e., nRMSE > 30%) for the sites grouped in the 10th and 40th percentiles (Fig. 9i). However, MLP and LSSVM provided satisfactory forecasting results for springtime precipitation (i.e., nRMSE < 30% and NSE > 0.5) in all percentiles, except for the 10th percentile

(Fig. 9h and i). The average nRMSE of annual precipitation forecasted by GRNN, MLP, LSSVM, and MLR was 15.8%, 8.0%, 6.7%, and 12.6%, respectively. However, the NSE values for GRNN-forecasted annual precipitation were found to be less than 0.5 for some clusters, indicating poor model performance (Fig. 9k). Therefore, MLP, LSSVM, and MLR performed acceptably for all percentiles on annual scale (Fig. 9k and l).

Fig. 8 The Pearson's correlation coefficient of the Generalized Regression Neural Network, Multi-Layer Perceptron, Multi-Linear Regression, and Least Squares Support Vector Machine algorithms obtained for different precipitation percentiles in training and testing sets
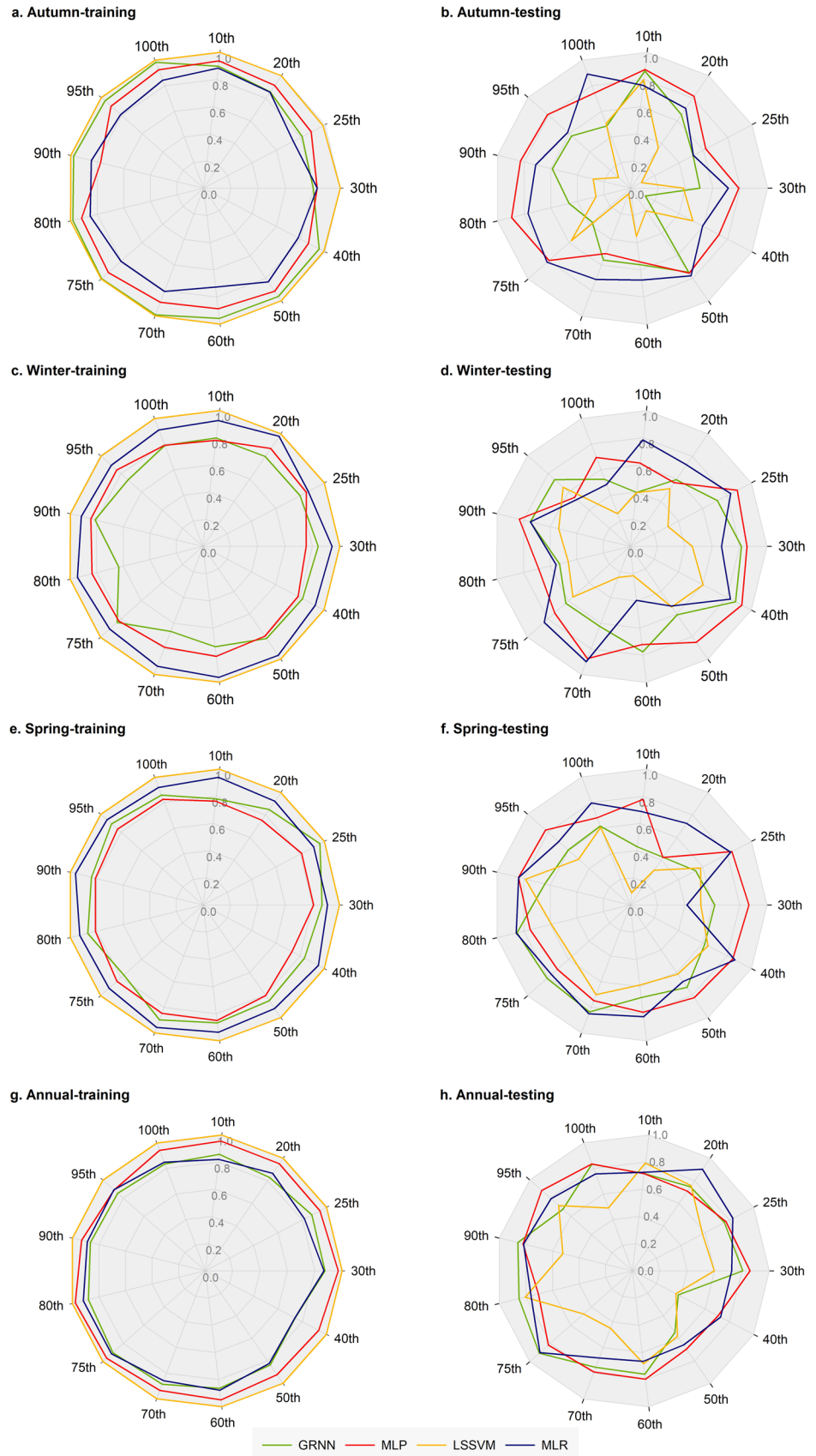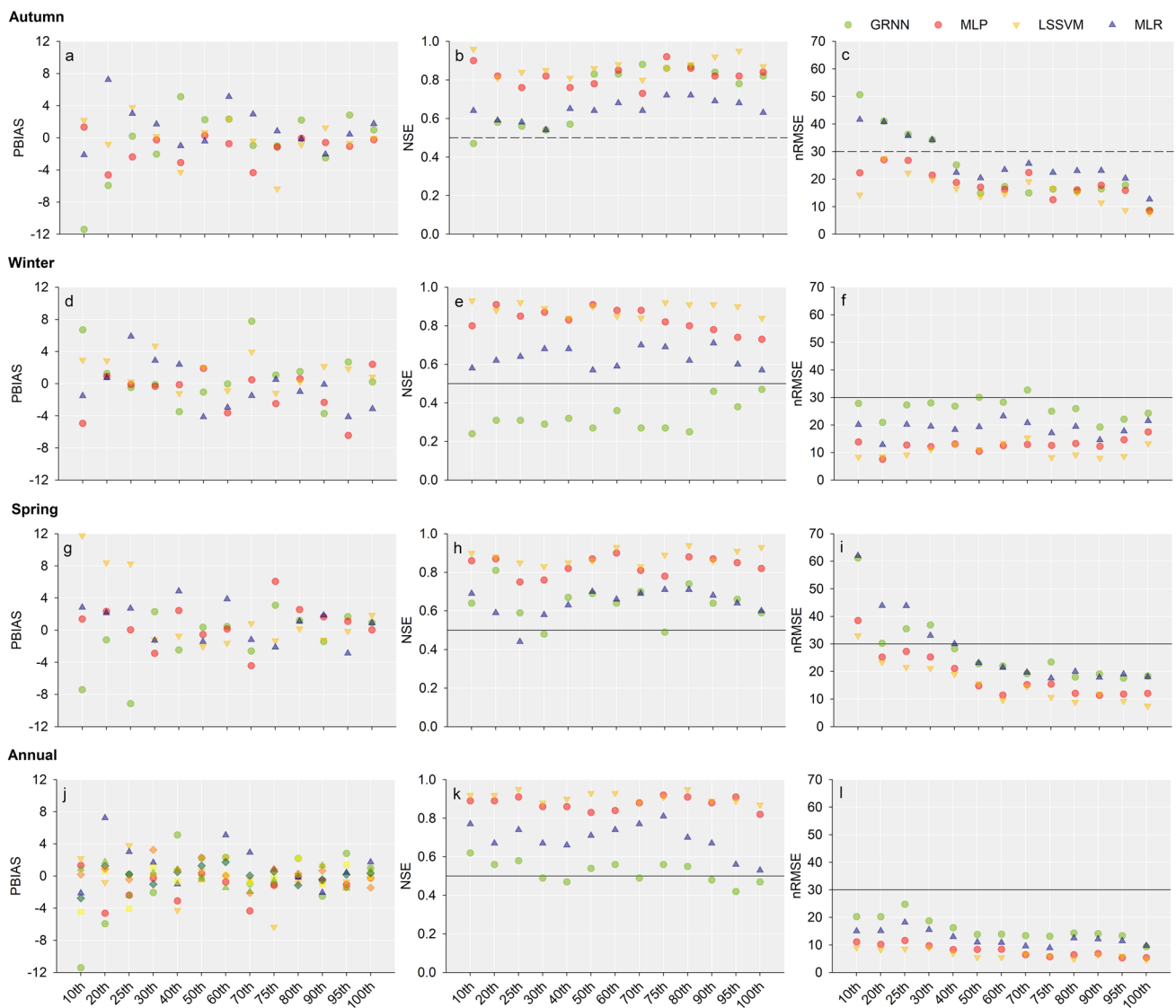
Fig. 9 The nRMSE, PBIAS, and NSE values of the Generalized Regression Neural Network (GRNN), the Multi-Layer Perceptron (MLP), the Multi-Linear Regression (MLR), and the Least Squares Support Vector Machine (LSSVM) for seasonal and annual precipitation across different clusters

The GRNN and MLR did not show a clear tendency to overestimate or underestimate seasonal and annual precipitation (Fig. 9a, d, g and j). The MLP algorithm, however, underestimated autumnal precipitation, and overestimated spring and annual precipitation in most percentiles (Fig. 9a, g and j). The LSSVM also tended to overestimate annual and wintertime precipitation for the majority of clusters (Fig. 9d and j). It is noteworthy that the absolute values of PBIAS did not exceed 25% for all cases (Fig. 9a, d, g and j), indicating acceptable bias errors of the studied ML alternatives in forecasting seasonal and annual precipitation.

Overall, LSSVM outperformed the other ML options in forecasting seasonal and annual precipitation. Except for spring precipitation of the areas in the 10th percentile, MLP and LSSVM gave reliable seasonal and annual precipitation forecasts (i.e., nRMSE < 30%, absolute PBIAS < 25%, and NSE > 0.5). Therefore, these algorithms are best suited to forecast precipitation using appropriate teleconnection signals listed in Table 3. The LSSVM has been identified as a robust algorithm for forecasting precipitation (Alizadeh and Farajzadeh 2018; Choubin et al. 2016; Tao et al. 2017). However, as discussed earlier, the LSSVM algorithm may experience overfitting, resulting in unsatisfactory precipitation modeling for regions not included in the training dataset. Therefore, caution should be exercised when applying LSSVM to forecast precipitation for unseen data. However, MLP provided a more balanced performance for training and testing steps (Fig. 8). This demonstrates that the MLP is less

susceptible to overfitting, which is a critical consideration when selecting an appropriate forecasting model. Therefore, in our study area, MLP is a preferable choice for precipitation forecasting compared to LSSVM.

The performance of ML models is generally inferior in the regions corresponding to the 10–40th percentiles, as compared to the sites grouped in the 40th to 100th percentiles (Fig. 9). For instance, the average nRMSE of autumn precipitation forecasts provided by GRNN, MLP and LSSVM and MLR was 37.5%, 23.4%, 20.0%, and 34.9% in the 10th to 40th percentiles, and 15.3%, 15.8%, 13.4%, and 21.4% in the 40–100th percentiles, respectively. As stated previously, the regions corresponding to the 10th and 40th percentiles are mostly situated in eastern and southern hyper-arid and arid areas (Fig. 1). These results indicate that forecasting skill of ML options is relatively lower in hyper-arid and arid areas. Precipitation modeling in hyper-arid/arid areas seems to be a challenging task. This can be ascribed to the nature of precipitation process in these environments, which is erratic, uneven and sparsely distributed (Al-Rawas and Valeo 2009; Altwegg and Anderson 2009; Attum et al. 2014). This renders these water-limited regions susceptible to flood and drought risks (Dai 2011). Moreover, precipitation may occur aloft, however, it is not detected by rain gauges. This can be attributed to sub-cloud evaporation which is the evaporation of raindrops prior to reaching to land surface owing to high atmospheric evaporative power in hyper-arid/arid areas (Dinku et al. 2011; Salamalikis et al. 2016; Wang et al. 2022a). This imposes high uncertainties on remotely-sensed precipitation products, and also teleconnection-based forecasts (Chen et al. 2020; Dinku et al. 2011). Given a relatively high forecasting skill of LSSVM and MLP in the 10–40th percentiles, these modeling approaches seem to be of much use for drought/flood risk management in the hyper-arid/arid areas studied.

The strength of precipitation and teleconnections correlations may change over time (Douville et al. 2017; Kamil et al. 2019; Nouri and Homaee 2020). For instance, Nouri and Homaee (2020) and Kamil et al. (2019) reported a sudden increase in the correlation coefficient between precipitation and ENSO as of the 1980s in central southwest Asia. As a result, ENSO-triggered droughts occurred more frequently in the twenty-first century with respect to the mid-twentieth century. It is worth noticing that three devastating dry episodes in Iran occurred in La Niña years of 1999–2001, 2007–2009, and 2010–2012 (Nouri and Homaee 2020; Trigo et al. 2010). Therefore, a change in time study might affect the performance of the ML algorithms by altering the strength of the association between precipitation and teleconnections.

Precipitation is a complex phenomenon influenced by several environmental factors, such as atmospheric conditions, geography, and topography (Kumari et al. 2016).

Precipitation is a conditional variable, meaning that it hinges on certain circumstances being met before it occurs (Zhang et al. 2016b). This adds an extra layer of complexity to precipitation forecasting, particularly in complex terrains. The ML algorithms can learn these relationships from historical data to forecast future precipitation. However, the above-mentioned complexities can impede the skill of ML algorithms to capture all the relevant information. Clustering identifies the underlying patterns and relationships that may be obscured in unstructured data (Ghorayeb et al. 2022; Kömüşcü et al. 2022; Mateo et al. 2013; Yang et al. 2022). In the current study, this technique facilitated the grouping of similar precipitation data into different subsets, enabling the identification of regions with similar precipitation patterns as recommended in the literature (Awan et al. 2015; Dehghan et al. 2018; Kömüşcü et al. 2022; Kumari et al. 2016). In addition, clustering can be applied to identify outliers or anomalies (Krleža et al. 2020; Mateo et al. 2013). In particular, we observed that some sites located in the southern and southeastern regions exhibited sudden spikes in springtime precipitation. However, since such patterns are typical for the corresponding clusters, we considered these data as anomalies rather than outliers. On the other hand, we also identified some stations with precipitation spikes that were not typical for a given cluster. Consequently, we considered these spikes as outliers and removed such stations from our study area. Overall, the application of clustering can significantly enhance the forecasting skill of ML models by reducing data complexity, and detecting patterns and outliers/anomalies.

Precipitation forecasting plays a crucial role in various fields, including agriculture. The availability of rainfall during the autumn and spring seasons is of utmost importance for dry farming in Iran. According to Nouri et al. (2017a), precipitation shortage during the months of October to December (OND) and March to May (MAM) can cause crops to fail in dry farming in Iran. Specifically, autumnal dry spells jeopardize crop establishment and finally result in crop failure under rainfed condition (Nouri et al. 2017a, b). Precipitation forecasts can also help the government to take well-informed decisions on food security such as importing cereal crops, croplands extension, and agricultural insurance. Overall, the forecasting methods described in this study can be beneficial for decision-makers to adopt proactive agricultural risk management in Iran. Furthermore, as climate change exerts its effects on precipitation patterns via impacting global teleconnection patterns, the forecasting frameworks can be used to design adaptation strategies in our study area.

As for water resources management, the results can provide valuable insights for policymakers seeking to ensure equitable distribution of water resources, particularly in densely populated water-scarce areas in Iran. By utilizing

precipitation forecasts, water managers can make informed decisions on how to manage water resources, including increasing water releases from reservoirs to create additional storage capacity based on expected inflow, or implementing water conservation measures to preserve water supplies during drought (Pattanaik and Das 2015; Ziervogel et al. 2010). The absence of reliable precipitation forecasts can have severe implications, as demonstrated by the disastrous flood in southwestern Iran during the spring of 2019 (Dezfuli 2020; Khosravi et al. 2020). In this event, the water held behind large dams constructed on the Dez and Karkheh rivers was not released in a timely manner, leading to severe flooding that inflicted substantial damage to the environment, infrastructure, and agriculture. Thus, precipitation forecasts can facilitate the development of early warning systems for floods and droughts and inform water management strategies to mitigate weather-related disasters. It is noteworthy that robust flood early warning systems require forecasts of low-frequency precipitation events on hourly and daily scales, as well as rare extreme events such as atmospheric rivers (Dezfuli 2020). These forecasts can aid decision-makers in identifying and mapping flood-prone areas, and can ultimately help to mitigate the impacts of floods. Therefore, we recommend that future research efforts be directed towards improving the low-frequency precipitation forecasts across Iran. This may involve developing new ML algorithms, as well as enhancing our understanding of the large-scale atmospheric anomalies.

In the present study, the point data were utilized for precipitation forecasting. However, in areas with limited data availability, gridded precipitation products offer several advantages, such as spatial continuity, long-term coverage, and access to a broader range of precipitation characteristics (Nouri 2023; Baatz et al. 2021; Valmassoi et al. 2022). Therefore, it is recommended to analyze the relationship between precipitation and teleconnections using gridded products. We also suggest further investigations on forecasting other precipitation characteristics such as precipitation type (e.g., snowfall), seasonality, and extremes using teleconnection signals in our study area.

Conclusions

We evaluated the precipitation forecasting skills of four machine learning (ML) approaches, i.e., the Generalized Regression Neural Network (GRNN), the Multi-Layer Perceptron (MLP), the Multi-Linear Regression (MLR), and the Least Squares Support Vector Machine (LSSVM), using a large number of teleconnection indices across different precipitation regimes in Iran. The precipitation percentiles were defined to cluster precipitation regimes. The El Niño-Southern Oscillation (ENSO) indices were applied as the predictors for most cases, denoting the ENSO phasing impacts on precipitation pattern, particularly during autumn and spring, in Iran. The LSSVM and MLP provided more reliable seasonal and annual precipitation forecasts relative to MLR and GRNN. However, as LSSVM showed high sensitivity to overfitting, MLP seems more suited to be applied for forecasting precipitation in the surveyed areas. Nonetheless, all ML algorithms showed weaker performance in the 10th and 40th percentiles, encompassing the southern and eastern hyper-arid and arid sites. Our results indicate that clustering precipitation regimes is a necessary step to overcome the spatial inconsistency frequently observed when investigating the association between precipitation and teleconnection anomalies. The findings can be of much use for developing proactive drought/flood risk management plans, highly required for maintaining food and water security in Iran.

Data availability All data are available as presented in supplementary material.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval All the authors provided ethical approval to submit the manuscript.

Consent to participate All the authors consent to participate in the present study.

Consent for publication All the authors consent to the publication of the article in Environmental Earth Sciences.

References

Ahmadi M, Salimi S, Hosseini SA, Poorantiyosh H, Bayat A (2019) Iran's precipitation analysis using synoptic modeling of major

teleconnection forces (MTF). Dyn Atmos Oceans 85:41–56. https://doi.org/10.1016/j.dynatmoce.2018.12.001

Alizadeh F, Farajzadeh J (2018) A hybrid linear–nonlinear approach to predict the monthly rainfall over the Urmia Lake watershed using wavelet SARIMAX-LSSVM conjugated model. J Hydroinf 20:246–262. https://doi.org/10.2166/hydro.2017.013

Al-Rawas GA, Valeo C (2009) Characteristics of rainstorm temporal distributions in arid mountainous and coastal regions. J Hydrol 376:318–326. https://doi.org/10.1016/j.jhydrol.2009.07.044

Altwegg R, Anderson MD (2009) Rainfall in arid zones: possible effects of climate change on the population ecology of blue cranes. Funct Ecol 23:1014–1021. https://doi.org/10.1111/j.1365-2435.2009.01563.x

Antanasijevic D, Pocajt V, Povrenovic D, Peric-Grujic A, Ristic M (2013) Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. Environ Sci Pollut Res Int 20:9006–9013. https://doi.org/10.1007/s11356-013-1876-6

Antanasijević D, Pocajt V, Perić-Grujić A, Ristić M (2014) Modelling of dissolved oxygen in the Danube River using artificial neural networks and Monte Carlo Simulation uncertainty analysis. J Hydrol 519:1895–1907. https://doi.org/10.1016/j.jhydrol.2014.10.009

Arcomano T, Szunyogh I, Pathak J, Wikner A, Hunt BR, Ott E (2020) A machine learning-based global atmospheric forecast model. Geophys Res Lett. https://doi.org/10.1029/2020gl087776

Ashley WS, Haberlie AM, Strohm J (2019) A climatology of quasi-linear convective systems and their hazards in the United States. Weather Forecast 34:1605–1631. https://doi.org/10.1175/waf-d-19-0014.1

Attum O, Ghazali U, El Noby SK, Hassan IN (2014) The effects of precipitation history on the kilometric index of Dorcas gazelles. J Arid Environ 102:113–116. https://doi.org/10.1016/j.jaridenv.2013.11.009

Awan JA, Bae D-H, Kim K-J (2015) Identification and trend analysis of homogeneous rainfall zones over the East Asia monsoon region. Int J Climatol 35:1422–1433. https://doi.org/10.1002/joc.4066

Baatz R, Hendricks Franssen HJ, Euskirchen E, Sihi D, Dietze M, Ciavatta S, Fennel K, Beck H, De Lannoy G, Pauwels VRN, Raiho A, Montzka C, Williams M, Mishra U, Poppe C, Zacharias S, Lausch A, Samaniego L, Van Looy K, Bogena H, Adamescu M, Mirtl M, Fox A, Goergen K, Naz BS, Zeng Y, Vereecken H (2021) Reanalysis in earth system science: toward terrestrial ecosystem reanalysis. Rev Geophys. https://doi.org/10.1029/2020rg000715

Baldwin MP, Gray LJ, Dunkerton TJ, Hamilton K, Haynes PH, Randel WJ, Holton JR, Alexander MJ, Hirota I, Horinouchi T, Jones DBA, Kinnersley JS, Marquardt C, Sato K, Takahashi M (2001) The quasi-biennial oscillation. Rev Geophys 39:179–229. https://doi.org/10.1029/1999RG000073

Bannayan M, Sadeghi Lotfabadi S, Sanjani S, Mohamadian A, Aghaalikhani M (2011) Effects of precipitation and temperature on crop production variability in northeast Iran. Int J Biometeorol 55:387–401. https://doi.org/10.1007/s00484-010-0348-7

Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon Wea Rev 115:1083–1126. https://doi.org/10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2

Böhm C, Schween JH, Reyers M, Maier B, Löhnert U, Crewell S (2021) Towards a climatology of fog frequency in the Atacama Desert via multi-spectral satellite data and machine learning techniques. J Appl Meteorol Climatol. https://doi.org/10.1175/jamc-d-20-0208.1

Boukabara S-A, Krasnopolsky V, Penny SG, Stewart JQ, McGovern A, Hall D, Ten Hoeve JE, Hickey J, Allen Huang H-L, Williams JK, Ide K, Tissot P, Haupt SE, Casey KS, Oza N, Geer AJ, Maddy ES, Hoffman RN (2021) Outlook for exploiting artificial intelligence in the earth and environmental sciences. Bull Am Meteorol Soc 102:E1016–E1032. https://doi.org/10.1175/bams-d-20-0031.1

Brown DP, Comrie AC (2004) A winter precipitation 'dipole' in the western United States associated with multidecadal ENSO variability. Geophys Res Lett. https://doi.org/10.1029/2003gl018726

Brutsaert W (2005) Hydrology: an introduction. Cambridge University Press

Cai W, Santoso A, Wang G, Yeh S-W, An S-I, Cobb KM, Collins M, Guilyardi E, Jin F-F, Kug J-S, Lengaigne M, McPhaden MJ, Takahashi K, Timmermann A, Vecchi G, Watanabe M, Wu L (2015) ENSO and greenhouse warming. Nat Clim Change 5:849. https://doi.org/10.1038/nclimate2743

Cao B, Zhan D, Wu X (2009) Application of SVM in financial research. 2009 International joint conference on computational sciences and optimization. pp 507–511

Casanueva A, Rodríguez-Puebla C, Frías MD, González-Reviriego N (2014) Variability of extreme precipitation over Europe and its relationships with teleconnection patterns. Hydrol Earth Syst Sci 18:709–725. https://doi.org/10.5194/hess-18-709-2014

Chai R, Sun S, Chen H, Zhou S (2018) Changes in reference evapotranspiration over China during 1960–2012: attributions and relationships with atmospheric circulation. Hydrol Processes 32:3032–3048. https://doi.org/10.1002/hyp.13252

Chakraborty D, Başağaoğlu H, Winterle J (2021) Interpretable vs. non-interpretable machine learning models for data-driven hydro-climatological process modeling. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2020.114498

Chen H, Chen Y, Li D, Li W (2020) Effect of sub-cloud evaporation on precipitation in the Tianshan Mountains (Central Asia) under the influence of global warming. Hydrol Processes 34:5557–5566. https://doi.org/10.1002/hyp.13969

Choi N, Lee M-I, Cha D-H, Lim Y-K, Kim K-M (2020) Decadal changes in the interannual variability of heat waves in East Asia caused by atmospheric teleconnection changes. J Clim 33:1505–1522. https://doi.org/10.1175/jcli-d-19-0222.1

Choubin B, Khalighi-Sigaroodi S, Malekian A, Kişi Ö (2016) Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals. Hydrol Sci J 61:1001–1009. https://doi.org/10.1080/02626667.2014.966721

Chung C, Power S (2017) The non-linear impact of El Niño, La Niña and the Southern Oscillation on seasonal and regional Australian precipitation. J South Hemisphere Earth Syst Sci 67:25–45. https://doi.org/10.22499/3.6701.003

Dai A (2011) Drought under global warming: a review. Wiley Interdiscip Rev Clim Change 2:45–65. https://doi.org/10.1002/wcc.81

Dannenberg MP, Wise EK, Janko M, Hwang T, Smith WK (2018) Atmospheric teleconnection influence on North American land surface phenology. Environ Res Lett. https://doi.org/10.1088/1748-9326/aaa85a

Dehghan Z, Eslamian SS, Modarres R (2018) Spatial clustering of maximum 24-h rainfall over Urmia Lake Basin by new weighting approaches. Int J Climatol 38:2298–2313. https://doi.org/10.1002/joc.5335

Dettori M, Cesaraccio C, Motroni A, Spano D, Duce P (2011) Using CERES-Wheat to simulate durum wheat production and phenology in Southern Sardinia, Italy. Field Crop Res 120:179–188. https://doi.org/10.1016/j.fcr.2010.09.008

Dezfuli A (2020) Rare atmospheric river caused record floods across the middle east. Bull Am Meteorol Soc 101:E394–E400. https://doi.org/10.1175/bams-d-19-0247.1

Di Y, Ding W, Mu Y, Small DL, Islam S, Chang NB (2015) Developing machine learning tools for long-lead heavy precipitation

prediction with multi-sensor data. 2015 IEEE 12th International Conference on Networking, Sensing and Control. pp 63–68

Dinku T, Ceccato P, Connor SJ (2011) Challenges of satellite rainfall estimation over mountainous and arid parts of east Africa. Int J Remote Sens 32:5965–5979. https://doi.org/10.1080/01431161.2010.499381

Douville H, Peings Y, Saint-Martin D (2017) Snow-(N)AO relationship revisited over the whole twentieth century. Geophys Res Lett 44:569–577. https://doi.org/10.1002/2016gl071584

Evans JP (2009) 21st century climate change in the Middle East. Clim Change 92:417–432. https://doi.org/10.1007/s10584-008-9438-5

Ghasemi AR, Khalili D (2008) The association between regional and global atmospheric patterns and winter precipitation in Iran. Atmos Res 88:116–133. https://doi.org/10.1016/j.atmosres.2007.10.009

Ghorayeb K, Ahmed Mawlod A, Maarouf A, Sami Q, El Droubi N, Merrill R, El Jundi O, Mustapha H (2022) Chain-based machine learning for full PVT data prediction. J Petrol Sci Eng. https://doi.org/10.1016/j.petrol.2021.109658

Gibson PB, Chapman WE, Altinok A, Delle Monache L, DeFlorio MJ, Waliser DE (2021) Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. Commun Earth Environ. https://doi.org/10.1038/s43247-021-00225-4

Guo Y, Li X, Bai G, Ma J (2012) Time series prediction method based on LS-SVR with modified Gaussian RBF. In: Huang T, Zeng Z, Li C, Leung CS (eds) Neural information processing. Springer, Berlin Heidelberg, pp 9–17

Hanley DE, Bourassa MA, O'Brien JJ, Smith SR, Spade ER (2003) A quantitative evaluation of ENSO indices. J Clim 16:1249–1258

He X, Guan H (2013) Multiresolution analysis of precipitation teleconnections with large-scale climate signals: a case study in South Australia. Water Resour Res 49:6995–7008. https://doi.org/10.1002/wrcr.20560

He C, Wei J, Song Y, Luo J-J (2021) Seasonal prediction of summer precipitation in the middle and lower reaches of the Yangtze River Valley: comparison of machine learning and climate model predictions. Water. https://doi.org/10.3390/w13223294

Helali J, Asadi Oskouei E (2021) Correlation analysis of large-scale teleconnection indices with monthly reference evapotranspiration of Iran synoptic stations. Iran J Soil Water Res 52:1629–1644. https://doi.org/10.22059/ijswr.2021.322853.668951 (In Persian)

Helali J, Salimi S, Lotfi M, Hosseini SA, Bayat A, Ahmadi M, Naderizarneh S (2020) Investigation of the effect of large-scale atmospheric signals at different time lags on the autumn precipitation of Iran's watersheds. Arab J Geosci. https://doi.org/10.1007/s12517-020-05840-7

Helali J, Momenzadeh H, Salimi S, Hosseini SA, Lotfi M, Mohamadi SM, Moghim GM, Pazhoh F, Ahmadi M (2021) Synoptic-dynamic analysis of precipitation anomalies over Iran in different phases of ENSO. Arab J Geosci. https://doi.org/10.1007/s12517-021-08644-5

Helali J, Ghaleni MM, Hosseini SA, Siraei AL, Saeidi V, Safarpour F, Mirzaei M, Lotfi M (2022) Assessment of machine learning model performance for seasonal precipitation simulation based on teleconnection indices in Iran. Arab J Geosci 15:1343. https://doi.org/10.1007/s12517-022-10640-2

Holmes A, Rüdiger C, Mueller B, Hirschi M, Tapper N (2017) Variability of soil moisture proxies and hot days across the climate regimes of Australia. Geophys Res Lett 44:7265–7275. https://doi.org/10.1002/2017gl073793

Hu Z-Z, Wu R, Kinter JL, Yang S (2005) Connection of summer rainfall variations in South and East Asia: role of El Niño-southern oscillation. Int J Climatol 25:1279–1289. https://doi.org/10.1002/joc.1159

Irannezhad M, Liu J, Chen D (2021) Extreme precipitation variability across the Lancang-Mekong River Basin during 1952–2015 in relation to teleconnections and summer monsoons. Int J Climatol 42:2614–2638. https://doi.org/10.1002/joc.7370

Jacques-Coper M, Veloso-Aguila D, Segura C, Valencia A (2021) Intraseasonal teleconnections leading to heat waves in central Chile. Int J Climatol 41:4712–4731. https://doi.org/10.1002/joc.7096

Jin H, Chen X, Wu P, Song C, Xia W (2021) Evaluation of spatial-temporal distribution of precipitation in mainland China by statistic and clustering methods. Atmos Res. https://doi.org/10.1016/j.atmosres.2021.105772

Jose DM, Vincent AM, Dwarakish GS (2022) Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. Sci Rep 12:4678. https://doi.org/10.1038/s41598-022-08786-w

Kamil S, Almazroui M, Kang I-S, Hanif M, Kucharski F, Abid MA, Saeed F (2019) Long-term ENSO relationship to precipitation and storm frequency over western Himalaya–Karakoram–Hindukush region during the winter season. Clim Dyn 53:5265–5278. https://doi.org/10.1007/s00382-019-04859-1

Kazem A, Sharifi E, Hussain FK, Saberi M, Hussain OK (2013) Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Appl Soft Comput 13:947–958. https://doi.org/10.1016/j.asoc.2012.09.024

Khan JA, Van Aelst S, Zamar RH (2007) Building a robust linear model with forward selection and stepwise procedures. Comput Stat Data Anal 52:239–248. https://doi.org/10.1016/j.csda.2007.01.007

Khosravi K, Panahi M, Golkarian A, Keesstra SD, Saco PM, Bui DT, Lee S (2020) Convolutional neural network approach for spatial prediction of flood hazard at national scale of Iran. J Hydrol. https://doi.org/10.1016/j.jhydrol.2020.125552

Kim J-S, Jain S, Yoon S-K (2012) Warm season streamflow variability in the Korean Han River Basin: links with atmospheric teleconnections. Int J Climatol 32:635–640. https://doi.org/10.1002/joc.2290

Kinouchi T, Yamamoto G, Komsai A, Liengcharernsit W (2018) Quantification of seasonal precipitation over the upper Chao Phraya River Basin in the past fifty years based on monsoon and El Niño/Southern Oscillation related climate indices. Water. https://doi.org/10.3390/w10060800

Kömüşcü AÜ, Turgu E, DeLiberty T (2022) Dynamics of precipitation regions of Turkey: a clustering approach by K-means methodology in respect of climate variability. J Water Clim Change 13:3578–3606. https://doi.org/10.2166/wcc.2022.186

Krishnaswamy J, Vaidyanathan S, Rajagopalan B, Bonell M, Sankaran M, Bhalla RS, Badiger S (2014) Non-stationary and non-linear influence of ENSO and Indian Ocean Dipole on the variability of Indian monsoon rainfall and extreme rain events. Clim Dyn 45:175–184. https://doi.org/10.1007/s00382-014-2288-0

Krleža D, Vrdoljak B, Brčić M (2020) Statistical hierarchical clustering algorithm for outlier detection in evolving data streams. Mach Learn 110:139–184. https://doi.org/10.1007/s10994-020-05905-4

Kumari M, Singh CK, Basistha A (2016) Clustering data and incorporating topographical variables for improving spatial interpolation of rainfall in mountainous region. Water Resour Manage 31:425–442. https://doi.org/10.1007/s11269-016-1534-0

Kutiel H, Benaroch Y (2002) North Sea-Caspian Pattern (NCP)—an upper level atmospheric teleconnection affecting the Eastern Mediterranean: identification and definition. Theor Appl Climatol 71:17–28. https://doi.org/10.1007/s704-002-8205-x

Lee W-K, Resdi TABT (2014) Neural network approach to coastal high and low water level prediction. In: Hassan R, Yusoff M,

Ismail Z, Amin NM, Fadzil MA (eds) InCIEC 2013. Springer, Singapore, pp 275–286

Liu Y-Y, Li L, Liu Y-S, Chan PW, Zhang W-H (2020) Dynamic spatial-temporal precipitation distribution models for short-duration rainstorms in Shenzhen, China based on machine learning. Atmos Res. https://doi.org/10.1016/j.atmosres.2020.104861

Mateo F, Carrasco JJ, Sellami A, Millán-Giraldo M, Domínguez M, Soria-Olivas E (2013) Machine learning methods to forecast temperature in buildings. Expert Syst Appl 40:1061–1068. https://doi.org/10.1016/j.eswa.2012.08.030

Medina H, Tian D, Marin FR, Chirico GB (2019) Comparing GEFS, ECMWF, and postprocessing methods for ensemble precipitation forecasts over Brazil. J Hydrometeorol 20:773–790. https://doi.org/10.1175/jhm-d-18-0125.1

Modaresi F, Araghinejad S, Ebrahimi K (2018) Selected model fusion: an approach for improving the accuracy of monthly streamflow forecasting. J Hydroinf 20:917–933. https://doi.org/10.2166/hydro.2018.098

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50:885–900

Muni Kumar N, Manjula R (2019) Design of multi-layer perceptron for the diagnosis of diabetes Mellitus Using Keras in deep learning. In: Satapathy SC, Bhateja V, Das S (eds) Smart intelligent computing and applications. Springer, Singapore, pp 703–711

Myers SS, Smith MR, Guth S, Golden CD, Vaitla B, Mueller ND, Dangour AD, Huybers P (2017) Climate change and global food systems: potential impacts on food security and undernutrition. Annu Rev Public Health 38:259–277. https://doi.org/10.1146/annurev-publhealth-031816-044356

Najafi MR, Moradkhani H, Wherry SA (2011) Statistical downscaling of precipitation using machine learning with optimal predictor selection. J Hydrol Eng 16:650–664. https://doi.org/10.1061/(asce)he.1943-5584.0000355

Nazemosadat MJ, Cordery I (2000) On the relationships between ENSO and autumn rainfall in Iran. Int J Climatol 20:47–61. https://doi.org/10.1002/(SICI)1097-0088(200001)20:1%3c47::AID-JOC461%3e3.0.CO;2-P

Nazemosadat MJ, Ghasemi AR (2004) Quantifying the ENSO-related shifts in the intensity and probability of drought and wet periods in Iran. J Clim 17:4005–4018. https://doi.org/10.1175/1520-0442(2004)017%3c4005:qtesit%3e2.0.co;2

Nguyen-Huy T, Deo RC, An-Vo D-A, Mushtaq S, Khan S (2017) Copula-statistical precipitation forecasting model in Australia's agro-ecological zones. Agric Water Manage 191:153–172. https://doi.org/10.1016/j.agwat.2017.06.010

Nicolai-Shaw N, Gudmundsson L, Hirschi M, Seneviratne SI (2016) Long-term predictability of soil moisture dynamics at the global scale: persistence versus large-scale drivers. Geophys Res Lett 43:8554–8562. https://doi.org/10.1002/2016gl069847

Niu J, Chen J, Sivakumar B (2014) Teleconnection analysis of runoff and soil moisture over the Pearl River basin in southern China. Hydrol Earth Syst Sci 18:1475–1492. https://doi.org/10.5194/hess-18-1475-2014

Nouri M (2023) Drought assessment using gridded data sources in data-poor areas with different aridity conditions. Water Resour Manage 37:4327–4343. https://doi.org/10.1007/s11269-023-03555-4

Nouri M, Homaee M (2018) On modeling reference crop evapotranspiration under lack of reliable data over Iran. J Hydrol 566:705–718. https://doi.org/10.1016/j.jhydrol.2018.09.037

Nouri M, Homaee M (2020) Drought trend, frequency and extremity across a wide range of climates over Iran. Meteorol Appl 27:e1899. https://doi.org/10.1002/met.1899

Nouri M, Homaee M (2021a) Contribution of soil moisture variations to high temperatures over different climatic regimes. Soil Tillage Res 213:105115. https://doi.org/10.1016/j.still.2021.105115

Nouri M, Homaee M (2021b) Spatiotemporal changes of snow metrics in mountainous data-scarce areas using reanalyses. J Hydrol 603:126858. https://doi.org/10.1016/j.jhydrol.2021.126858

Nouri M, Homaee M (2022) Reference crop evapotranspiration for data-sparse regions using reanalysis products. Agric Water Manage 262:107319. https://doi.org/10.1016/j.agwat.2021.107319

Nouri M, Homaee M, Bannayan M (2017a) Climate variability impacts on rainfed cereal yields in west and northwest Iran. Int J Biometeorol 61:1571–1583. https://doi.org/10.1007/s00484-017-1336-y

Nouri M, Homaee M, Bannayan M, Hoogenboom G (2017b) Towards shifting planting date as an adaptation practice for rainfed wheat response to climate change. Agric Water Manage 186:108–119. https://doi.org/10.1016/j.agwat.2017.03.004

Overland JE, Adams JM, Bond NA (1999) Decadal variability of the aleutian low and its relation to high-latitude circulation. J Clim 12:1542–1548. https://doi.org/10.1175/1520-0442(1999)012%3c1542:DVOTAL%3e2.0.CO;2

Pattanaik DR, Das AK (2015) Prospect of application of extended range forecast in water resource management: a case study over the Mahanadi River basin. Nat Hazards 77:575–595. https://doi.org/10.1007/s11069-015-1610-4

Peng X, Wang Y (2009) A normal least squares support vector machine (NLS-SVM) and its learning algorithm. Neurocomputing 72:3734–3741. https://doi.org/10.1016/j.neucom.2009.06.005

Petrucci O (2022) Review article: factors leading to the occurrence of flood fatalities: a systematic review of research papers published between 2010 and 2020. Nat Hazards Earth Syst Sci 22:71–83. https://doi.org/10.5194/nhess-22-71-2022

Qin Y, Abatzoglou JT, Siebert S, Huning LS, AghaKouchak A, Mankin JS, Hong C, Tong D, Davis SJ, Mueller ND (2020) Agricultural risks from changing snowmelt. Nat Clim Change 10:459–465. https://doi.org/10.1038/s41558-020-0746-8

Räsänen TA, Kummu M (2013) Spatiotemporal influences of ENSO on precipitation and flood pulse in the Mekong River Basin. J Hydrol 476:154–168. https://doi.org/10.1016/j.jhydrol.2012.10.028

Roque-Malo S, Kumar P (2017) Patterns of change in high frequency precipitation variability over North America. Sci Rep 7:10853. https://doi.org/10.1038/s41598-017-10827-8

Rust W, Holman I, Corstanje R, Bloomfield J, Cuthbert M (2018) A conceptual model for climatic teleconnection signal control on groundwater variability in Europe. Earth Sci Rev 177:164–174. https://doi.org/10.1016/j.earscirev.2017.09.017

Rust W, Holman I, Bloomfield J, Cuthbert M, Corstanje R (2019) Understanding the potential of climate teleconnections to project future groundwater drought. Hydrol Earth Syst Sci 23:3233–3245. https://doi.org/10.5194/hess-23-3233-2019

Saemian P, Hosseini-Moghari S-M, Fatehi I, Shoarinezhad V, Modiri E, Tourian MJ, Tang Q, Nowak W, Bárdossy A, Sneeuw N (2021) Comprehensive evaluation of precipitation datasets over Iran. J Hydrol. https://doi.org/10.1016/j.jhydrol.2021.127054

Saemian P, Tourian MJ, AghaKouchak A, Madani K, Sneeuw N (2022) How much water did Iran lose over the last two decades? J Hydrol. https://doi.org/10.1016/j.ejrh.2022.101095

Salamalikis V, Argiriou AA, Dotsika E (2016) Isotopic modeling of the sub-cloud evaporation effect in precipitation. Sci Total Environ 544:1059–1072. https://doi.org/10.1016/j.scitotenv.2015.11.072

Scher S, Messori G (2018) Predicting weather forecast uncertainty with machine learning. Q J R Meteorol Soc 144:2830–2841. https://doi.org/10.1002/qj.3410

Shouval R, Fein JA, Savani B, Mohty M, Nagler A (2021) Machine learning and artificial intelligence in haematology. Br J Haematol 192:239–250. https://doi.org/10.1111/bjh.16915

Sigaroodi SK, Chen Q, Ebrahimi S, Nazari A, Choobin B (2014) Long-term precipitation forecast for drought relief using atmospheric circulation factors: a study on the Maharloo Basin in Iran. Hydrol Earth Syst Sci 18:1995–2006. https://doi.org/10.5194/hess-18-1995-2014

Skeeter WJ, Senkbeil JC, Keellings DJ (2019) Spatial and temporal changes in the frequency and magnitude of intense precipitation events in the southeastern United States. Int J Climatol 39:768–782. https://doi.org/10.1002/joc.5841

Sospedra-Alfonso R, Melton JR, Merryfield WJ (2015) Effects of temperature and precipitation on snowpack variability in the Central Rocky Mountains as a function of elevation. Geophys Res Lett 42:4429–4438. https://doi.org/10.1002/2015gl063898

Tao L, He X, Wang R (2017) A hybrid LSSVM model with empirical mode decomposition and differential evolution for forecasting monthly precipitation. J Hydrometeorol 18:159–176. https://doi.org/10.1175/jhm-d-16-0109.1

Theobald A, McGowan H, Speirs J (2018) Teleconnection influence of precipitation-bearing synoptic types over the Snowy Mountains region of south-east Australia. Int J Climatol 38:2743–2759. https://doi.org/10.1002/joc.5457

Trenberth KE, Dai A, van der Schrier G, Jones PD, Barichivich J, Briffa KR, Sheffield J (2013) Global warming and changes in drought. Nat Clim Change 4:17. https://doi.org/10.1038/nclimate2067

Trenberth KE (2020) ENSO in the Global Climate System. El Niño Southern Oscillation in a Changing Climate. pp 21–37

Trigo RM, Gouveia CM, Barriopedro D (2010) The intense 2007–2009 drought in the Fertile Crescent: impacts and associated atmospheric circulation. Agric for Meteorol 150:1245–1257. https://doi.org/10.1016/j.agrformet.2010.05.006

Valmassoi A, Keller JD, Kleist DT, English S, Ahrens B, Ďurán IB, Bauernschubert E, Bosilovich MG, Fujiwara M, Hersbach H, Lei L, Löhnert U, Mamnun N, Martin CR, Moore A, Niermann D, Ruiz JJ, Scheck L (2022) Current challenges and future directions in data assimilation and reanalysis. Bull Am Meteorol Soc. https://doi.org/10.1175/bams-d-21-0331.1

Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the northern hemisphere winter. Mon Wea Rev 109:784–812. https://doi.org/10.1175/1520-0493(1981)109%3c0784:TITGHF%3e2.0.CO;2

Wang C, Enfield DB (2001) The tropical western hemisphere warm pool. Geophys Res Lett 28:1635–1638. https://doi.org/10.1029/2000GL011763

Wang X, Li C, Zhou W (2006a) Interdecadal variation of the relationship between Indian rainfall and SSTA modes in the Indian Ocean. Int J Climatol 26:595–606. https://doi.org/10.1002/joc.1283

Wang XX, Chen S, Lowe D, Harris CJ (2006b) Sparse support vector regression based on orthogonal forward selection for the generalised kernel model. Neurocomputing 70:462–474. https://doi.org/10.1016/j.neucom.2005.12.129

Wang S, Anichowski A, Tippett MK, Sobel AH (2017) Seasonal noise versus subseasonal signal: forecasts of California precipitation during the unusual winters of 2015–2016 and 2016–2017. Geophys Res Lett 44:9513–9520. https://doi.org/10.1002/2017gl075052

Wang L, Wang S, Zhang M, Duan L, Xia Y (2022a) An hourly-scale assessment of sub-cloud evaporation effect on precipitation isotopes in a rainshadow oasis of northwest China. Atmos Res. https://doi.org/10.1016/j.atmosres.2022.106202

Wang W, Yang P, Xia J, Zhang S, Cai W (2022b) Coupling analysis of surface runoff variation with atmospheric teleconnection indices in the middle reaches of the Yangtze River. Theor Appl Climatol 148:1513–1527. https://doi.org/10.1007/s00704-022-04013-8

Wei L, Mao J, Ma Y (2008) A new modeling method for nonlinear rate-dependent hysteresis system based on LS-SVM. 2008 10th International Conference on Control, Automation, Robotics and Vision. pp 1442–1446

Xiao M, Zhang Q, Singh VP (2015) Influences of ENSO, NAO, IOD and PDO on seasonal precipitation regimes in the Yangtze River basin, China. Int J Climatol 35:3556–3567. https://doi.org/10.1002/joc.4228

Xiao M, Zhang Q, Singh VP (2017) Spatiotemporal variations of extreme precipitation regimes during 1961–2010 and possible teleconnections with climate indices across China. Int J Climatol 37:468–479. https://doi.org/10.1002/joc.4719

Yang M, Zhao M, Huang D, Su X (2022) A composite framework for photovoltaic day-ahead power prediction based on dual clustering of dynamic time warping distance and deep autoencoder. Renew Energy 194:659–673. https://doi.org/10.1016/j.renene.2022.05.141

Yoon JH, Wang SS, Gillies RR, Kravitz B, Hipps L, Rasch PJ (2015) Increasing water cycle extremes in California and in relation to ENSO cycle under global warming. Nat Commun 6:8657. https://doi.org/10.1038/ncomms9657

Zhang Q, Li J, Singh VP, Xu C-Y, Deng J (2013) Influence of ENSO on precipitation in the East River basin, south China. J Geophys Res Atmos 118:2207–2219. https://doi.org/10.1002/jgrd.50279

Zhang Q, Wang Y, Singh VP, Gu X, Kong D, Xiao M (2016a) Impacts of ENSO and ENSO Modoki+A regimes on seasonal precipitation variations and possible underlying causes in the Huai River basin, China. J Hydrol 533:308–319. https://doi.org/10.1016/j.jhydrol.2015.12.003

Zhang Y, You Q, Chen C, Ge J (2016b) Impacts of climate change on streamflows under RCP scenarios: a case study in Xin River Basin, China. Atmos Res 178–179:521–534. https://doi.org/10.1016/j.atmosres.2016.04.018

Zhao Z, Wang H, Wang C, Li W, Chen H, Deng C (2020) Changes in reference evapotranspiration over Northwest China from 1957 to 2018: Variation characteristics, cause analysis and relationships with atmospheric circulation. Agric Water Manage. https://doi.org/10.1016/j.agwat.2019.105958

Ziervogel G, Johnston P, Matthew M, Mukheibir P (2010) Using climate information for supporting climate change adaptation in water resource management in South Africa. Clim Change 103:537–554. https://doi.org/10.1007/s10584-009-9771-3