Journal of Water Resources and Geosciences

Vol. 5, No. 2, 2026

Journal of Water Resources and Geosciences

Climate classification prediction based on compound drought and hot events and AI classification algorithms: Iraq as a case study

Anas Mahmood Al-Juboori*

Dams and water resources research center, University of Mosul, Mosul, Iraq.

*Corresponding author's email: anasmmr@uomosul.edu.iq

Abstract

Climate change is one of the most important challenges facing the world in this century due to its significant role in affecting the way of life on Earth in addition to its significant impact on the world's economies. The increase in extreme hydrological phenomena such as drought and floods has required more studies to analyse these phenomena and develop plans to reduce their impact on society and the economy. In the present research, the climate of Iraq was studied and analysed using the Standardized Precipitation Index (SPI), the Standard Temperature Index (STI), and the analysis of compound hot and dry events. Temperature and precipitation data from four meteorological stations representative of the climate of Iraq were used. Two classification algorithms were used to predict the annual climate classification of Iraq. The results showed that Iraq was affected by climate change, as the percentage of extreme hot and dry events reached 90% during the period 2012-2022, with the annual rainfall rate decreasing by more than 40%, and temperatures were observed to rise at a rate of 0.68 degrees Celsius per decade. The novelty of this research is the application of classification algorithms in machine learning using hot and dry categorical classification as input variables and the study demonstrated the efficiency of the Bayesian network algorithm in predicting the annual and monthly climate classification.

Keywords: STI, SPI, compound hot/dry, climate change, artificial inelegance algorithm.

Journal of Water Resources and Geosciences

Vol. 5, No. 2, 2026

التنبؤ بتصنيف المناخ بناءً على الاحداث الجافة-الحارة المركبة وخوارزميات التصنيف في الذكاء الاصطناعي: العراق كدراسة حالة

انس محمود الجبوري*

مركز بحوث السدود والموارد المائية، جامعة الموصل، موصل، العراق

*البريد الالكتروني للمؤلف المراسل: anasmmr@uomosul.edu.iq

الخلاصة

يعد تغير المناخ أحد أهم التحديات التي تواجه العالم في هذا القرن، نظرًا لدوره المحوري في التأثير على نمط الحياة على كوكب الأرض، فضلًا عن تأثيره الكبير على اقتصادات العالم. وقد استدعى ازدياد الظواهر الهيدرولوجية المتطرفة، كالجفاف والفيضانات، إجراء المزيد من الدراسات لتحليل هذه الظواهر ووضع خطط للحدّ من آثارها على المجتمع والاقتصاد. في هذا البحث، درس مناخ العراق وحلل باستخدام مؤشر هطول الأمطار القياسي (SPI) ومؤشر درجة الحرارة القياسي (STI)، بالإضافة إلى تحليل الظواهر المناخية المركبة من الحرارة والجفاف. استخدمت بيانات درجة الحرارة وهطول الأمطار من أربع محطات أرصاد جوية تمثل مناخ العراق. استخدمت خوارزميتان للتصنيف للتنبؤ بالتصنيف المناخي السنوي للعراق. أظهرت النتائج أن العراق تأثر بتغير المناخ، حيث بلغت نسبة الظواهر المناخية المتطرفة من الحرارة والجفاف 90% خلال الفترة 2012-2022، مع انخفاض معدل هطول الأمطار السنوي بأكثر من 40%، ولوحظ ارتفاع درجات الحرارة بمعدل 0.68 درجة مئوية لكل عقد. تكمن حداثة هذا البحث في تطبيق خوارزميات التصنيف في التعلم الآلي باستخدام التصنيف الفئوي الحار والجاف كمتغيرات إدخال، وقد أظهرت الدراسة كفاءة خوارزمية الشبكة البايزية في التنبؤ بالتصنيف المناخي السنوي والشهري.

الكلمات المفتاحية: STI، SPI، الاحداث الحارة-الجافة المركبة، تغير المناخ، خوارزميات الذكاء الاصطناعي.

1. Introduction

Iraq is one of the countries affected by climate change, as reports from international organizations indicate that Iraq is the fifth country in the world in terms of its impact on climate change. In recent years, Iraq has been exposed to noticeable changes in its climate, such as a significant rise in temperatures, a decrease in the average annual rainfall, and a decrease in the levels of its rivers, especially the Tigris and Euphrates rivers and their tributaries, which has increased the problem of agricultural land degradation and an increase in the salinity rate in water and soil, which has greatly affected the country's economy. Temperature and precipitation are among the most important factors influencing extreme hydrological events, especially droughts and floods, in addition to their significant impact on the hydrological cycle, water resources and quality, agricultural production and the economy (Gao et al. 2022). Scientists have proposed many methods to identify drought events, but the Standardized Precipitation Index (SPI) proposed by Mackie et al. (1993) is one of the most widely used methods in the literature. There are several studies on the use of the Standardized Precipitation Index to analyse dry and wet periods in different regions of the world. The accumulated drought monitoring data for 200 years in Seoul, the capital of South Korea, were analysed, and the results showed that the effective drought index is more efficient than the standard precipitation index in assessing drought, especially for short-term periods (Kim et al. 2009). An algorithm for detecting the change point in rainfall data in the class-to-class transition matrix was developed using data from four meteorological stations in Kenya to validate the proposed algorithm (Duggins et al. 2010).

The SPI was used to analyse drought in Greece using different time scale and the results showed that the SPI is an effective method for assessing drought in Greece (Karavitis et al. 2011). Due to the different time scales used in calculating the SPI, which may cause confusion for decision makers and researchers in choosing the most appropriate time scale to define drought, a multivariate approach has been developed that is characterized by its ability to aggregate different time series into a new time series called the Multivariate Standardized Precipitation Index (MSPI) (Bazrafshan et al. 2014). Two machine learning algorithms, namely artificial neural networks and support vector regression, were tested, as well as using wavelet transforms to process the inputs of artificial neural networks and SVR models to form WA-ANN and WA-SVR models. The results showed that WA-ANN model is superior to other models in forecasting meteorological drought (Belayneh et al. 2014). Projected changes in meteorological and hydrological drought in Spain were studied through SPI simulations using scientific climate change scenario data (Lorenzo et al., 2024).

The increasing impact of extreme hot and dry climate phenomena on the social, environmental and economic aspects of life requires more hydrological studies to analyze and predict these phenomena. The compound dry and hot events in China during the period 1961-2014 were studied on a daily time scale and it was found that the intensity of these events on China's climate is increasing (Feng et al. 2023). The association between temperature and drought was studied over four years in China and the results showed that drought coincided with a general increase in temperature in most of the study areas (Chen et al. 2019). The intensity of compound hot and dry events was studied using the Standardized Dry and Hot Index (SDHI) and showed a temporal increase in the average intensity of the world's hottest months (Hao et al. 2018). Since the world has been severely affected by compound dry and hot events since the 1990s, especially in Russia, the United States, China, India and other countries in the world, a group of scientists has developed a new Dry and Hot Event Index (BDHI). The new index has been examined over large areas of the world during the period 1950-2019 and the efficiency of the new index has been shown compared to traditional indices used for the same purpose (Wu et al. 2021). Hot and dry events in China were studied using two indices, the Standardized Compound Event Indicator (SCEI) and Standardized Dry Hot Index (SDHI) using monthly precipitation and temperature data from 1961 to 2012. The results showed a significant increase in extreme events in most parts of China (Wu et al. 2020). The Meta-Gaussian forecasting model was used to predict the intensity of extreme hot and dry weather events in South Africa based on the El Niño phenomenon. The results showed that the El Niño phenomenon significantly affects extreme weather events in South Africa (Hao et al. 2019). The causal links between dry and hot events were studied in terms of mechanism, time scale and impact range using the concept of system dynamics. The results helped researchers gain a deeper understanding of the mechanism and predictability of the causal links between extreme events (Mukherjee et al. 2023). The hot-dry events of 1921-2020 in the Mississippi River Basin were analyzed by developing a new mathematical method using time series of rainfall and temperature to determine the rates of change of extreme events using sensitivity analysis based on partial derivatives (Zhao et al. 2024). The uncertainty in extreme event indices was assessed based on several variables including runoff, soil moisture and precipitation using a multivariate index. The results showed the uncertainty in the prediction of the combined hot and dry event forecasts (Hosseinzadehtalaei et al. 2024). The ability of global climate models to simulate complex hot and dry events in South America during the period 1979–2014 was evaluated, the results were compared with observational data and the accuracy of the models was examined (Collazo et al. 2023). The impact of combined dry and warm events on several river basins in China was studied to determine the influence ratio of these events on the hydrological characteristics of these rivers by formulating the total differential equation with the Meta-gaussian model for the return period of combined dry and warm events (2024).

The relationship between vegetation cover and drought frequency in the Hamrin Basin of Iraq was studied under the influence of climate change, through the analysis of monthly weather data and a drought frequency index. Vegetation cover was analysed using satellite imagery. The results showed that the basin is sensitive to climate change and that vegetation cover is closely correlated with drought frequency (Al-Khafaji et al. 2023). Data on the discharges of the Tigris River in Baghdad for the period 1930-2022 were analysed to predict the impact of future drought on the river's discharges as a result of climate change and the construction of dams on its tributaries (AL-Shahrabale 2025).

The objectives of this study are: (1) Estimation and analysis of standardized precipitation index and standardized temperature index for four selected metrological stations covering most of Iraq, (2) Estimation and analysis of compound hot/dry events for the studied climate stations, (3) Predicting the annual and monthly climate classification in Iraq using machine learning algorithms. The novelty of this research is the application of classification algorithms in machine learning using hot and dry categorical classification as input variables.

2. Materials and methods

2.1. Study area and data

The climate of Iraq is divided into three main regions: the Mediterranean climate, which clearly affects the northern region of Iraq is characterized by cold winters and relatively mild summers with rainfall rates exceeding 400 mm annually; the steppe climate, which is a transitional climate between the Mediterranean climate and the desert climate; and the desert climate, which includes the alluvial plain and the western plateau of Iraq and covers most of the territory of Iraq and is characterized by high maximum daily temperatures in the summer ranging between 40 and 50 degrees Celsius. To study the annual classification of Iraq's climate, four climate stations were selected to cover most of Iraq's regions: Mosul station, which represents the climate of northern Iraq; Baghdad station, which represents the climate of central Iraq; Basra station, which represents the climate of southern Iraq; and Rutba station, which represents the climate of western Iraq. The spatial representation of Iraq's climate based on only four stations is insufficient to capture the strong climatic heterogeneity across the country. Figure 1 shows a map of Iraq with the locations of the metrological stations in this study. Data on mean annual temperature, annual rainfall, mean monthly temperature, monthly rainfall were collected for the four proposed meteorological stations. The statistical analysis of the data used in this study is summarized in Table 1.

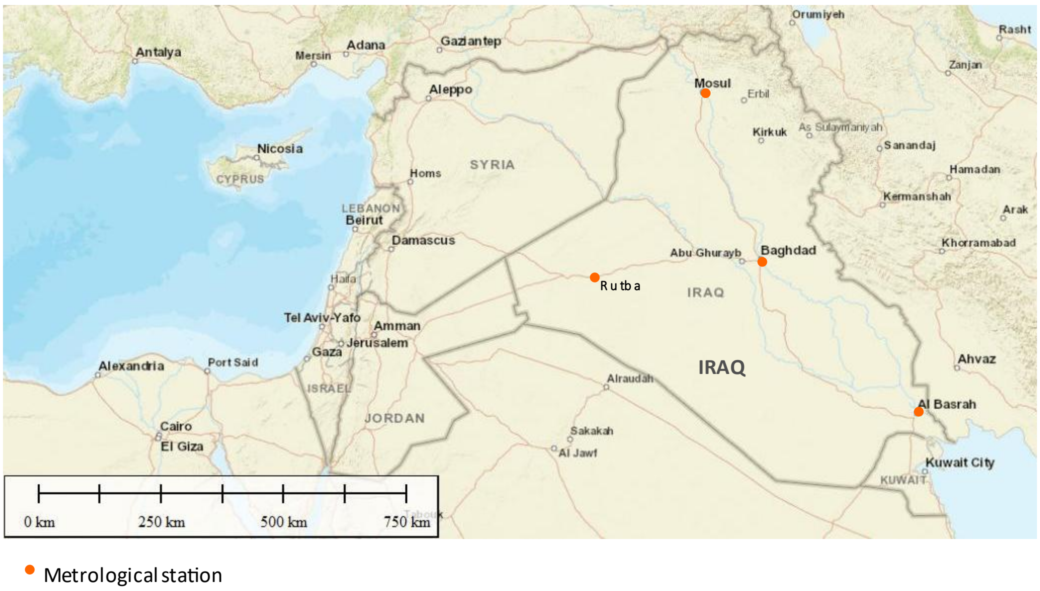


Figure 1: Study area.

Table 1: Statistical analysis of the data used in this study.

Station	Record length	Annual rainfall (mm)			Annual mean temperature (Degree Celsius)		
		Minimum	Maximum	Mean	Minimum	Maximum	Mean
Mosul	1988-2022	70.20	633.00	327.27	25.48	30.98	28.52
Baghdad	1988-2022	51.80	195.50	111.04	28.78	33.03	31.32
Rutba	1988-2022	31.90	247.10	115.23	29.90	34.23	32.74
Basrah	1988-2022	21.80	263.80	96.80	31.38	35.53	34.07

2.2. Definition of compound drought and hot events

The Standardized Precipitation Index (SPI) and the Standardized Temperature Index (STI) are used to classify the hot/dry events. The Standardized Precipitation Index developed by (McKee et al. 1993) is used to measure the dry condition and the Standardized Temperature Index is used to measure the hot condition. Both indices are calculated using the same procedure developed by (McKee et al. 1993). The sequence of SPI can be classified into dry and wet conditions and the sequence of STI can be classified into hot and cold conditions. The combined SPI and STI conditions are divided into four categories based on the threshold 0 to classify the dry/hot conditions. The SPI, STI and compound dry/hot classification conditions according to specified threshold are summarized in Table (2).

Table 2: Thresholds used for STI, SPI and compound dry/hot classification conditions.

Standardized Precipitation Index (SPI)		standardized Temperature Index (STI)	
Condition	**Threshold**	**Condition**	**Threshold**
Extreme wet	$SPI \geq 2.5$	Extreme hot	$STI \geq 2.5$
Severe wet	$1.5 \leq SPI < 2.5$	Severe hot	$1.5 \leq STI < 2.5$
Mild wet	$0.5 \leq SPI < 1.5$	Mild hot	$0.5 \leq STI < 1.5$
Normal wet	$0.0 \leq SPI < 0.5$	Normal hot	$0.0 \leq STI < 0.5$
Normal drought	$-0.5 \leq SPI < 0.0$	Normal cold	$-0.5 \leq STI < 0.0$
Mild drought	$-1.5 \leq SPI < -0.5$	Mild cold	$-1.5 \leq STI < -0.5$
Severe drought	$-2.5 \leq SPI < -1.5$	Severe cold	$-2.5 \leq STI < -1.5$
Extreme drought	$SPI < -2.5$	Extreme cold	$STI < -2.5$
Thresholds used for compound dry/hot event conditions			
Threshold	Condition	Threshold	Condition
$SPI \leq 0$, $STI > 0$	Dry hot	$SPI \leq 0$, $STI \leq 0$	Dry cold
$SPI > 0$, $STI \leq 0$	Wet cold	$SPI > 0$, $STI > 0$	Wet hot

2.3. Mann-Kendall test

The Mann-Kendall method (Mann 1945) is one of the most important statistical methods in trend analysis of hydrological time series by determining whether the variable has an upward or downward trend. The continuous increase in the variable with time shows the upward trend of the variable's time series, while the continuous decrease in the variable's time series shows the decreasing trend of the variable. Therefore, the alternative hypothesis (H_a) for this test means the presence of a slope in the variable's time series, unlike the null hypothesis (H_0), which indicates the absence of a slope in the data series. The following equations are used in Mann-Kendall test calculations:

$$S = \sum_{i=1}^{n-1} \sum_{j=j+1}^{n} sgn\left(x_j - x_i\right) \quad (1)$$

$$sgn(\varphi) = \begin{cases} 1; & if\left(x_j - x_i\right) > 0 \\ 0; & if\left(x_j - x_i\right) = 0 \\ -1; & if\left(x_j - x_i\right) < 0 \end{cases} \quad (2)$$

in which *xi* and *xj* are data quantities at time *i* and *j*, respectively, and n is the data lengt h.

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{p} t_i(t_i - 1)(2t_i + 5)}{18} \quad (3)$$

Where x_i and x_j are data quantities at time i and j respectively, as a result if S>0 this indicates that the time series of data is increasing with time while S<0 indicates that the time series is decreasing with time. The standard Z value can be determined as follows:

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}}; & S > 0 \\ 0; & S = 0 \\ \dfrac{S+1}{\sqrt{Var(S)}}; & S < 0 \end{cases} \quad (4)$$

The value of Z resulting from Equation 4 is tested using the following hypotheses.

1- If $|Z| \leq Z_{(1-\alpha/2)}$ This means that the time series data has no trend.

2- If $Z < -Z_{(1-\alpha/2)}$ This means that there is a downward trend in the time series data.

3- If $Z > -Z_{(1-\alpha/2)}$ This means that there is an upward trend in the time series data.

Where α represents the significance coefficient in calculating the confidence level, this means that α with a value of 0.05 indicates a 95% confidence level and this value is approximately equal to 1.96.

2.4. Theil-Sen Approach (TS)

The Thiel-Sen (TS) approach is one of the methods used to calculate the slope value of any time series. The following equation is used to estimate the (TS).

$$\beta = median\left(\frac{X_j - X_i}{j - i}\right) \tag{5}$$

where, X_i and X_j show the ordinal data quantities in the *ith* and *jth* year and β shows the trend slope size.

2.5. Bayesian network algorithm

A Bayesian network (BN), also called a belief network or decision network, is a probabilistic graphical model that represents a set of variables and their conditional dependencies through a directed periodic graph. The network is ideal for predicting an event that has occurred with the probability that one of the possible causes is the primary factor in that event. Therefore, the Bayesian network can represent the probabilistic relationship between the cause and the causes. A Bayesian lattice probability model in hydrology can be built by combining observed and recorded evidence with real-world knowledge to determine the probability of events using seemingly independent variables. The network mainly uses two mathematical models for classification: Tree Augmented Naive Bayes Networks (TAN) and Markov Blanket. A Bayesian network displays variables in a graphical model as nodes in a data set and the probabilistic or conditional independence between them and represents the causal relationships between nodes as a network. There are several reasons for using a Bayesian network in data mining, including that it helps identify causal relationships and thus enables understanding the philosophy of a problem and predicting solutions. The network has a highly effective approach to avoiding overfitting data. Since the Bayesian algorithm is based on graphics, it is easy to get a clear visualization of the relationships involved in the problem.

2.6. C5 algorithm

The C5.0 algorithm is developed based on the decision tree algorithm. The model in the C5.0 algorithm splits the sample based on the field that provides the maximum information gain. Each subsample identified by the first split is split again based on a different field that is likely to be different, and this process continues until the splitting process reaches a stage where the subsamples cannot be split further. The samples, especially those at the lower level, are examined, and the samples that do not contribute significantly to the efficiency of the model are modified or removed. The C5.0 algorithm is only predictive of the categorical target and is used for nominal or ordinal field data. There are two types of predictive models that can be developed in the C5.0 algorithm. The first model is the decision tree, which is a description of the partitions that the algorithm has reached, with each leaf node describing a subset of the training data sample, such that each instance in the training data belongs to a single leaf node in the tree. The second model is the rule set, which is a model based on a set of rules that attempt to make predictions for individual records. First, decision trees are analysed and then rule sets are extracted. This method represents a simplified analysis of the information in the decision tree in order to retain the important information from the full decision tree, but in a less complex way. Because the methodology of rule sets differs from the methodology of decision trees, they do not have the same characteristics as decision trees. The most important difference between the rule set methodology and the decision tree methodology is that the rule set can apply more than one rule to a given record or it can apply no rule at all. In case of applying multiple rules to a data record, each rule gets a vote based on the weight associated with that rule and the final prediction is then determined by adding the weighted votes of all the rules that apply to the given record.

3. Results and Discussion

3.1. Standardized temperature index

The annual average temperature data for the study area were analysed using the standard temperature index and the results are showed in Figure 2. The results in Figure 2 are shown a noticeable change in temperatures in the study area after 2005 in addition to the recent years are hotter than previous years. The results of STI analysis show the coldest period is started from 1988 to 1997. We note that the period from 1998 to 2012 is a period of temperature fluctuations between cold and hot, and in general the temperature classification during this period ranged from moderately hot to moderately cold, so this period is classified as a period of moderate temperatures in the Mosul, Baghdad, and Basra stations. Of the 35 years studied, the first decade from 1988 to 1997 was quite cold, and the next 15 years (1998-2012) were a transition period from cold to hot and the last decade (2013-2022) is classified as a hot period because all STI values are more than zero. This behaviour is evident for temperatures in Mosul, Baghdad and Basra stations. By analysing the STI results for Rutba station, it is clear that the first decade (1988-1997) is the coldest period because all STI values are less than zero for this period. The following years from 1998 to 2022 are classified as a hot period because all STI values are greater than zero. Table (3) summarizes the frequency of STI classification from 1988 to 2022 for the stations (Mosul, Baghdad, Rutba, and Basra). It is also noted that the percentage of hot years is 53%, 59%, 72%, and 63%, while the percentage of cold years is 47%, 41%, 28%, and 38% for the Mosul, Baghdad, Rutba, and Basra metrological stations, respectively. Most of the STI classifications during the studied period ranged between mild hot and mild cold, with only one year observed as severe hot and one year as severe cold, while the extreme cold category appeared in only one year during the studied period. The most STI classification categories observed during the study period for the four stations were mild hot and normal cold. Figure 3 shows the frequency of the STI classification classes for the studied metrological stations. The trend analysis of the STI for the four stations were analysed using Mann-Kendall test. The results showed that the Z value for Mosul, Baghdad, Rutba and Basra metrological stations are 4.66, 4.76, 3.24 and 4.56 respectively. The Z-value results for the four meteorological stations are greater than the $Z_{1-\alpha/2}$ value at the 95% confidence level of 1.96, which meets the first condition of the analysis method, and these results indicate a significant upward trend in temperatures in Iraq. The results showed that the mean annual temperatures increased by 0.69, 0.69, 0.59, and 0.74 °C/decade for Mosul, Baghdad, Rutba, and Basra stations, respectively.

3.2. Standardized precipitation index

The annual rainfall data for the four meteorological stations were analysed using the standard precipitation index and the results are showed in Figure 2. The results in Figure 2 show a significant trend in precipitation for the four studied meteorological stations. From the analysis of the SPI results of the Mosul Meteorological Station, it is clear that the period from 1988 to 1997 can be classified as a wet period, as the annual rainfall during this period were above the mean annual rainfall in that region, while the period from 1998 to 2022 is classified as a dry period, as most of the years are dry with some wet years that do not exceed six years during this period. The SPI results for Baghdad station showed stable characteristics of the annual rainfall time series with a slight decrease in rainfall amounts in recent years. The behaviour of the SPI results for the Rutba and Basra stations is similar to the characteristics of the SPI results for the Mosul station in terms of the frequencies of dry and wet periods. Table (3) summarizes the frequency of SPI classification from 1988 to 2022 for the stations (Mosul, Baghdad, Rutba, and Basra). It is also noted that the percentage of drought years is 57%, 60%, 63%, and 51%, while the percentage of wet years is 43%, 40%, 37%, and 49% for the Mosul, Baghdad, Rutba, and Basra metrological stations, respectively. Most of the SPI classifications during the studied period ranged between mild drought and mild wet. Figure 3 shows the frequency of the SPI classification classes for the studied metrological stations. The trend analysis of the SPI for the four stations was analysed using Mann-Kendall test. The results showed that the Z value for Mosul, Baghdad, Rutba and Basra metrological stations are -2.44, -0.47, -2.07 and -2.27 respectively. The Z-value results for the Mosul, Rutba, and Basra meteorological stations are less than the $-Z_{1-\alpha/2}$ value at the 95% confidence level of -1.96, which meets the second hypothesis of the analysis method, and these results indicate a significant downward trend. The Z value of Baghdad meteorological station is consistent with the first hypothesis of the analysis method, and this result indicates that there is no trend in annual rainfall at Baghdad station. The results of Theil-Sen showed decreasing trend 7.4, 0.64, 4.25, and 0.64 mm/year for Mosul, Baghdad, Rutba, and Basra stations, respectively as a result of climate change and the increasing number of dry years. As previously mentioned, the period from 1988 to 1997 is a wet period, while the period from 1998 to 2022 is a dry period. The average annual rainfall for both periods was compared for the four meteorological stations under study Figure 4. From the analysis of Figure 4, it is clear that the average annual rainfall decreased from 440 mm to 282 mm for Mosul station, from 125 mm to 106 mm for Baghdad station, from 153 mm to 73 mm for Rutba station, and from 160 mm to 97 mm for Basra station. These results indicate a significant decrease in the amount of annual rainfall by up to 36 % in northern Iraq, 15 % in central Iraq, 52 % in western Iraq, and 47 % in southern Iraq.

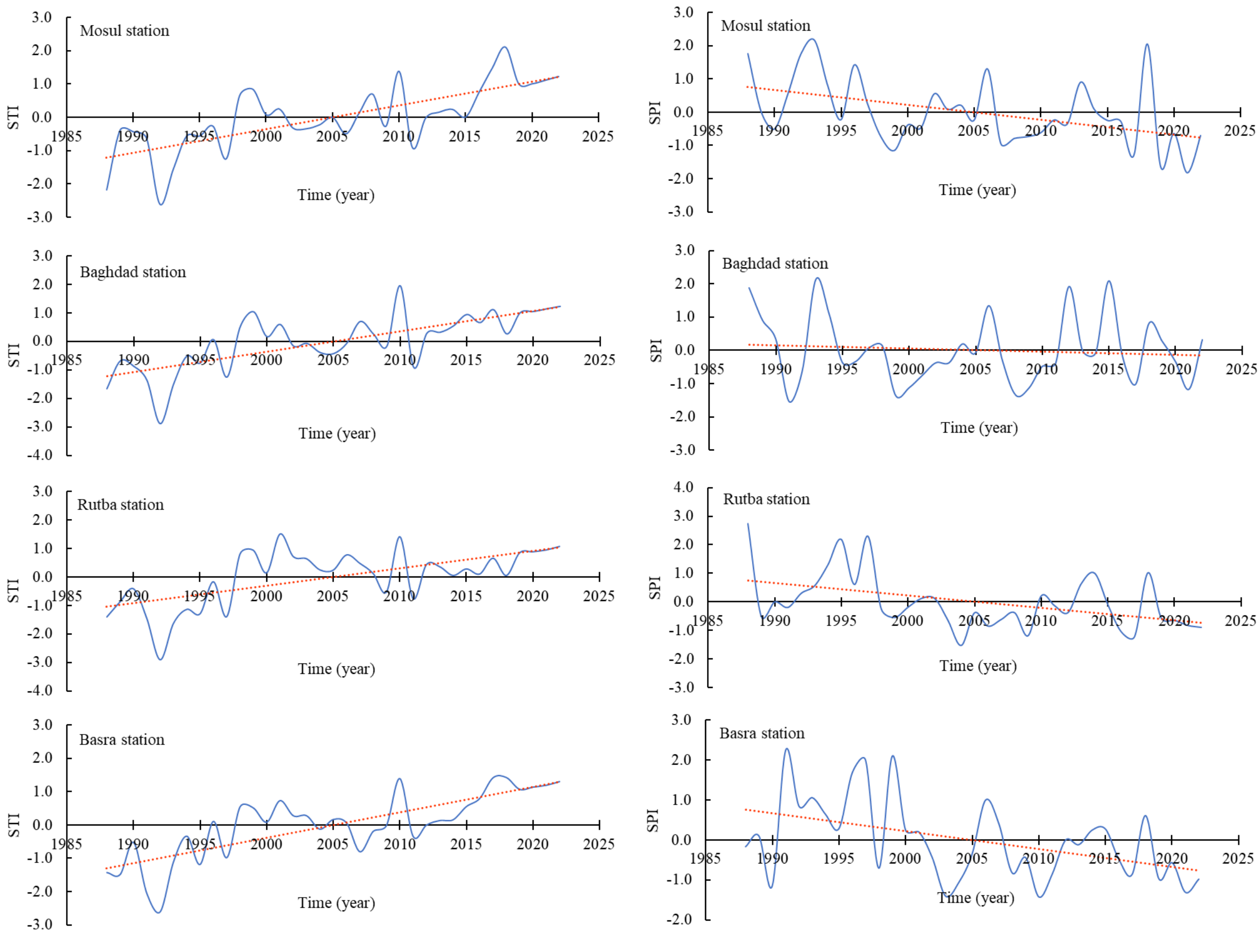
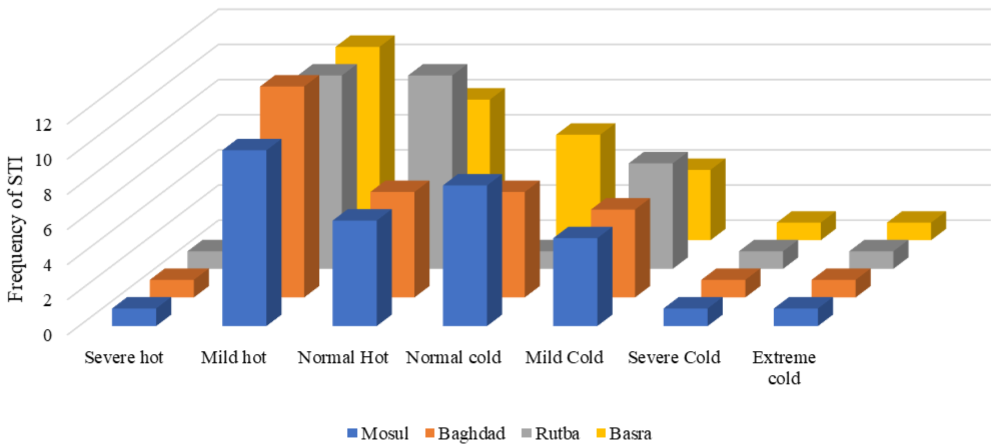


Figure 2: STI and SPI results for the studied meteorological stations.

Table 3: Frequency of STI and SPI classification for the studied meteorological stations.

STI classification	Mosul	Baghdad	Rutba	Basra	SPI classification	Mosul	Baghdad	Rutba	Basra
Extreme hot	0	0	0	0	Extreme drought	0	0	0	0
Severe hot	1	1	1	1	Severe drought	2	1	1	1
Mild hot	10	12	11	11	Mild drought	9	8	12	11
Normal Hot	6	6	11	8	Normal drought	9	12	9	6
Normal cold	8	6	1	6	Normal wet	5	6	4	8
Mild Cold	5	5	6	4	Mild wet	6	4	6	5
Severe Cold	1	1	1	1	Severe wet	4	4	3	4
Extreme cold	1	1	1	1	Extreme wet	0	0	0	0
Hot (total)	17	19	23	20	Drought (total)	20	21	22	18
Cold (total)	15	13	9	12	Wet (total)	15	14	13	17



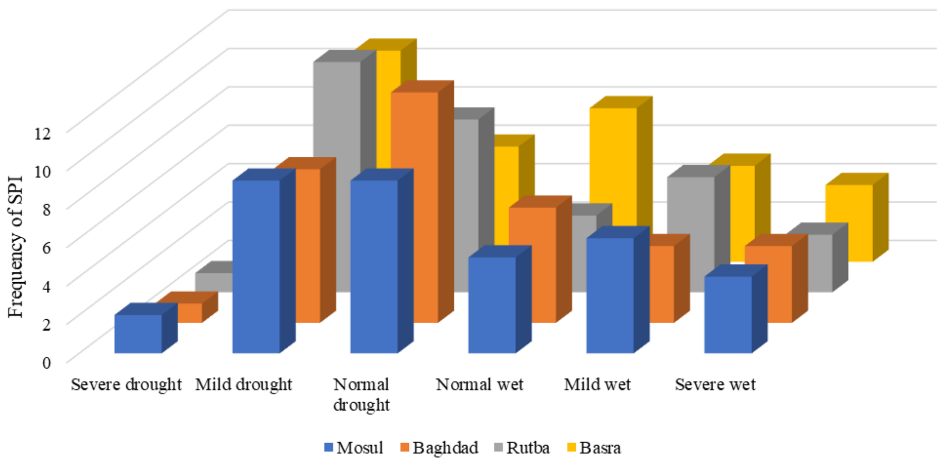


Figure 3: Frequency of STI and SPI classification classes for the studied meteorological stations.

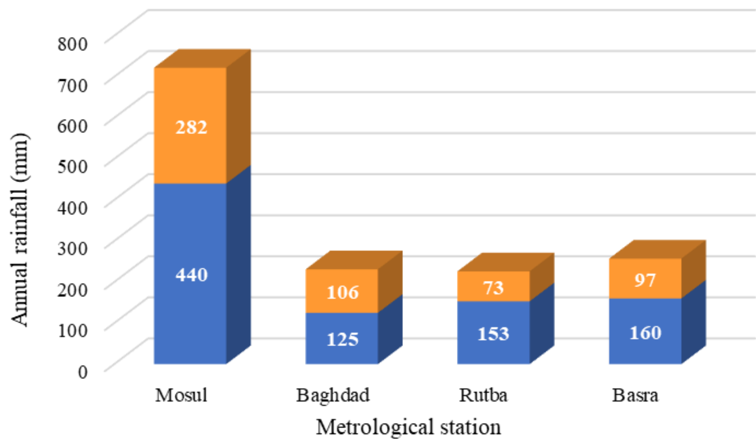


Figure 4: Average annual rainfall.

3.3. Hot and dry compound events

The interannual variability of hot and dry events was determined using combined of SPI and STI as described in the materials and methods paragraph. The results in Figure 5 show that the annual classification of Iraq's climate. The percentage of years classified as hot and dry was 31%, 41%, 53%, and 38% for Mosul, Baghdad, Rutba, and Basra metrological stations, respectively. The percentage of years classified as cold and wet was 31%, 16%, 19%, and 25% for Mosul, Baghdad, Rutba, and Basra metrological stations, respectively. The results also showed that the Rutba station has the highest hot and dry climate classification, meaning that the western region of Iraq is more affected by climate change. Most of the years with the hot and dry years recorded have occurred within the past decade, indicating the increasing impact of climate change on Iraq.

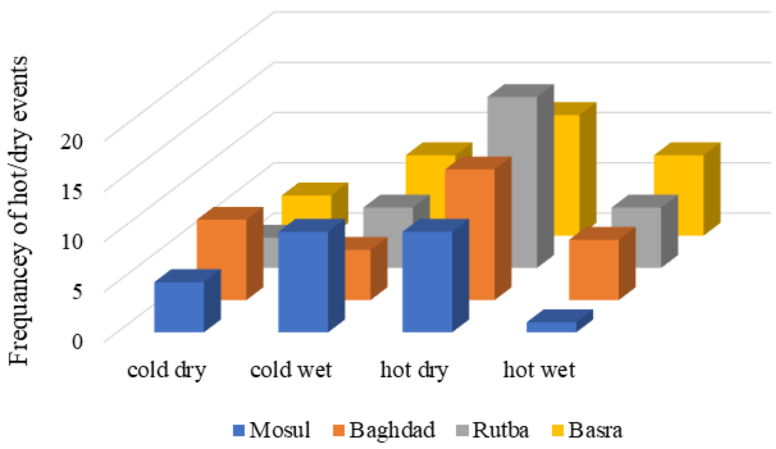


Figure 5: Frequency of annual climate classification categories for the studied meteorological stations.

The monthly variability in Iraq's climate was analysed by analysing the combined hot and dry events at the four stations using monthly temperature and rainfall data. From figure 6 we note that the percentage of months that fall within cold and wet classification is 35%, 29%, 29%, and 50%, while the percentage of months that fall within the hot and dry classification is 47%, 48%, 48%, and 50% for the Mosul, Baghdad, Rutba, and Basra stations, respectively.

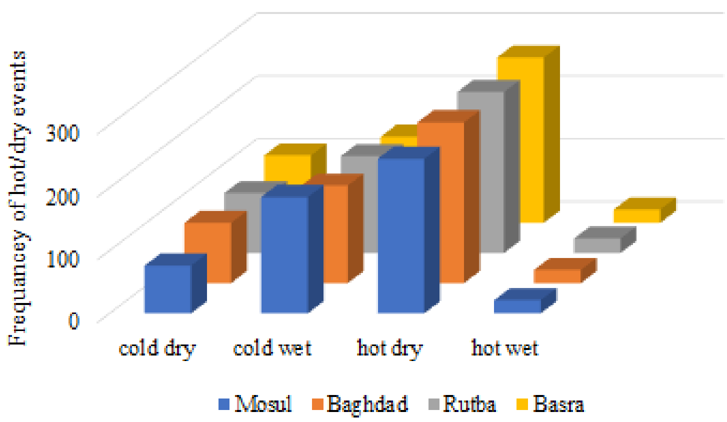


Figure 6: Frequency of monthly climate classification categories for the studied meteorological stations.

3.4. Climate state classification prediction

Bayesian network and C5 algorithm were used to predict the annual climate classification of the study area. The models input includes the categorical data represents the sequence of compound hot/dry events. To develop the models, the entire data at each metrological stations was separated into two subsets of data, the first 70% for training phase while the other 30% to validate the models. Model inputs include text data representing compound event such as hot dry, hot wet, cold dry, cold wet. More than one set of inputs was examined and it was found that the set with a lag time of three previous years gave the best results. The results are summarized in Table (4). From the analysis of the results in Table 4, it is clear that the Bayesian algorithm was more accurate than the C5 algorithm in predicting the annual climate classification. The results demonstrated the Bayesian algorithm's effectiveness in predicting annual climate classification at the studied stations. The accuracy of correct prediction of annual climate classification ranges from 81% to 95% in the training phase, while it ranges from 80% to 90% in the model validation phase. The results demonstrated the weakness of the algorithm in predicting the annual climate classification of the studied stations. The accuracy of correctly predicting the annual climate classification ranged between 40% and 50% in the training phase, while it ranged between 40% and 60% in the validation phase.

The Bayesian network and C5 algorithms were also tested to predict the monthly climate classification for the study area. The data was separated into two subsets, the first 70% for the training phase while the other 30% for the model validation. Several sets of input data were examined and it was found that inputs consisting of monthly classification data for the previous three months gave the best results. By analysing Table 5, it is clear that both models (Bayesian and C5) have similar performance in predicting the monthly climate classification. The accuracy of correct prediction of monthly climate classification ranges from 65% to 70% in the training phase, while it ranges from 65% to 72% in the model validation phase.

Table 4: Results of Bayesian network and C5 algorithms.

	Training		Validation	
	Correct	**Wrong**	**Correct**	**Wrong**
	Mosul metrological station			
Bayesian network	81.82%	18.18%	80%	20%
C5	50%	50%	60%	40%
	Baghdad metrological station			
Bayesian network	81.82%	18.18%	90%	10%
C5	45.45%	54.55%	50%	50%
	Rutba metrological station			
Bayesian network	86.36%	13.64%	90%	10%
C5	54.55%	45.45%	50%	50%
	Basra metrological station			
Bayesian network	95.45%	4.55%	90%	10%
C5	40.91%	59.09%	40%	60%

Table 5: Results of Bayesian network and C5 algorithms.

	Training		Validation	
	Correct	**Wrong**	**Correct**	**Wrong**
	Mosul metrological station			
Bayesian network	70.88%	29.12%	72.64%	27.36%
C5	70.17%	29.83%	71.7%	28.3%
	Baghdad metrological station			
Bayesian network	71.6%	28.4%	69.81%	30.19%
C5	69.21%	30.79%	66.04%	33.96%
	Rutba metrological station			
Bayesian network	68.26%	31.74%	70.75%	29.25%
C5	66.59%	33.41%	66.98%	33.02%
	Basra metrological station			
Bayesian network	67.3%	32.7%	66.98%	33.02%
C5	65.39%	34.61%	65.09%	34.91%

4. Conclusion

It is clear that Iraq is greatly affected by climate changes that have generally affected its water revenues by decreasing the annual rainfall and increasing temperatures. As a result of these changes, agricultural areas and green areas have shrunk, which has increased the problem of desertification, causing an increase in dust storms that affect Iraq. In recent years, an increase in extreme hydrological events has been observed, which requires in-depth studies to analyse extreme events, especially drought. From the analysis of the results, it is clear that Iraq is exposed to the impact of climate change, which greatly affects the characteristics of temperature and rainfall. From the analysis of the annual mean temperatures and annual rainfall of four meteorological stations distributed over the area of Iraq, the following conclusions were concluded:

1- A significant increase in temperatures at a rate of 0.68°C/decade, and the percentage of hot years ranges between 53% and 72% of the study period.
2- A significant decrease in the annual rainfall, with the percentage of drought years ranging between 51% and 63% of the study period.
3- Annual rainfall decreased by 36% in northern Iraq, 15% in central Iraq, 52% in western Iraq, and 47% in southern Iraq.
4- It is clear that the impact of climate change has become more severe in the period from 2012 to 2022, as the percentage of years classified as hot/dry has reached to 90%.
5- The results showed the efficiency of Bayesian network algorithm in predicting annual and monthly climate classification. The advance prediction of annual and monthly climate classification helps decision makers to make advance plans to overcome the impact of climate change on water resources management.

References

Al-Khafaji, M. S., Al-Ameri R. A., & Saeed F. H. (2023). Prediction of Land Cover Compliance to the Drought Frequency under Climate Change Conditions in Hemrin Watershed, *Journal of Water Resources and Geosciences,* 2(1), 68-93.

AL-Shahrabale, Q., AL-Shahrabalee S. (2025). Impacts of Climate Change and Dam Construction on Tigris River Streamflow: A Case Study at Baghdad Sarai Station, *Journal of Water Resources and Geosciences* 4(2), 140-158.

Bazrafshan, J., Hejabi, R. J. (2014) Drought Monitoring Using the Multivariate Standardized Precipitation Index (MSPI). *Water Resour Manage,* 28:1045–1060. https://doi.org/10.1007/s11269-014-0533-2

Belayneh, A., Adamowski, J., Khalil, B., & Ozga-Zielinski, B. (2014). Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models. *J Hydrol,* 508:418–429. http://dx.doi.org/10.1016/j.jhydrol.2013.10.052

Chen, L., Chen, X., Cheng, L., & Lin, Z. (2019). Compound hot droughts over China: Identification, risk patterns and variations. *Atmos Res*., 227:210-219. https://doi.org/10.1016/j.atmosres.2019.05.009

Collazo, S., Barrucand, M., & Rusticucci, M. (2023). Hot and dry compound events in South America: present climate and future projections, and their association with the Pacific Ocean. *Nat Hazards*, 119:299–323. https://doi.org/10.1007/s11069-023-06119-2

Duggins, J., Williams, M., Kim, D., & Smith, E. (2010). Changepoint detection in SPI transition probabilities. J Hydrol, 388:456-463. https://doi.org/10.1016/j.jhydrol.2010.05.030

Feng, Y., Wang, H., Sun, F., & Liu, W. (2023). Dependence of compound hot and dry extremes on individual ones across China during 1961–2014. *Atmos Res.,* 283:106553. https://doi.org/10.1016/j.atmosres.2022.106553

Gao, F., Chen, X., Yang, W. et al. (2022). Statistical characteristics, trends, and variability of rainfall in Shanxi Province, China, during the period 1957–2019. *Theor Appl Climatol,* 148:955–966. https://doi.org/10.1007/s00704-022-03924-w

Hao, Z., Hao, F., Singh, V. P., & Zhang, X. (2018). Changes in the severity of compound drought and hot extremes over global land areas. *Environ. Res. Lett,* 13:124022. https://doi.org/10.1088/1748-9326/aaee96

Hao, Z., Hao, F., Singh, V. P., & Zhang, X, (2019), Statistical prediction of the severity of compound dry-hot events based on El Niño-Southern Oscillation, *J Hydrol*, https://doi.org/10.1016/j.jhydrol.2019.03.001

Hosseinzadehtalaei, P., Termonia, P., & Tabari, H. (2024). Projected changes in compound hot-dry events depend on the dry indicator considered. *Commun Earth Environ,* 5:220. https://doi.org/10.1038/s43247-024-01352-4

Karavitis, C. A., Alexandris, S., Tsesmelis, D. E., & Athanasopoulos, G. (2011), Application of the Standardized Precipitation Index (SPI) in Greece. *Water,* 3:787-805. https://doi.org/10.3390/w3030787

Kim, D., Byun, H., & Choi, K. (2009). Evaluation, modification, and application of the Effective Drought Index to 200-Year drought climatology of Seoul, Korea. *J Hydrol,* 378:1–12. https://doi.org/10.1016/j.jhydrol.2009.08.021

Lorenzo, M. N., Pereira, H., Alvarez, I., & Dias, J. M. (2024). Standardized Precipitation Index (SPI) evolution over the Iberian Peninsula during the 21st century. *Atmospheric Research,* 297:1-19. https://doi.org/10.1016/j.atmosres.2023.107132

McKee, T. B., Doesken, N. J., & Kliest, J. (1993). The relationship of drought frequency and duration to time scales. Proc. *Eighth Conf. of Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184

Mukherjee, S., Mishra, A, K., Zscheischler, J. et al. (2023). Interaction between dry and hot extremes at a global scale using a cascade modeling framework. *Nat Commun,* 14:277. https://doi.org/10.1038/s41467-022-35748-7

Theil H. (1950). A rank-invariant method of linear and polynomial regression analysis. I, II, III. Nederl. Akad. Wetensch., Proc., 53: 386–392.

Wu, H., Su, X., & Singh, V. P. (2021). Blended Dry and Hot Events Index for monitoring dry-hot events over global land areas. *Geophysical Research Letters,* 48: e2021GL096181. https://doi.org/10.1029/2021GL096181

Wu, X., Hao, Z., Zhang, X., Li, C., & Hao, F. (2020). Evaluation of severity changes of compound dry and hot events in China based on a multivariate multi-index approach, *J Hydrol,* 583: 124580. https://doi.org/10.1016/j.jhydrol.2020.124580

Xiong, S., Zhao, T., Guo, C., et al. (2024). Evaluation and attribution of trends in compound dry-hot events for major river basins in China. Sci. *China Earth Sci,* 67:79–91. https://doi.org/10.1007/s11430-022-1174-7

Zhao, T., Xiong, S., Tian, Y., Wu, Y., Li, B., & Chen, X. (2024). Compound dry and hot events over major river basins of the world from 1921 to 2020, *Weather and Climate Extremes,* 44: 100679. https://doi.org/10.1016/j.wace.2024.100679