Teknik Dergi, 2022 12067-12085, Paper 671

Entropy Based Regional Precipitation Prediction in the Case of Gediz River Basin

Özgür BOZOĞLU[1]
Türkay BARAN[2]
Filiz BARBAROS[3]

ABSTRACT

The Shannon's entropy concept is defined to measure the information content of hydrological processes in hydrology and water resources. Entropy concept has also provided an opportunity to solve several related topics in water resources engineering. The presented study aims to define the regional distribution of the expected long-term annual total precipitation, by using the entropy concept. For this purpose, the frequency analysis of the observed long-term total monthly precipitation data for each station is analyzed, and entropy values are calculated using the frequency histogram called " *Intensity Entropy – IE*". By using the IE values, it is possible to define the regional information of long-term expected precipitation even if the gauging stations have different observation periods without missing any information of available data. In addition, there is no need to complete the data of the missing observation years (months) to define the precipitation-elevation relations for producing an isometric map. In regional analysis, it is possible to create isoentropy map, by using the determined IE values for each gauging station. The IE method is performed and isoentropy map is created for Gediz Basin as a case study, and the a priori conditions of using IE method results for regional information are discussed in the paper.

Keywords: Gediz River Basin, water resources management, entropy, uncertainty, hydrological processes, intensity entropy.

1. INTRODUCTION

In water systems planning, using an objective criterion is essential to determine the information content of hydrological data. Hydrological processes must be measured in order to make optimal

Note:
- This paper has been received on April 28, 2020 and accepted for publication by the Editorial Board on January 18, 2021.
- Discussions on this paper will be accepted by September 30, 2022.

- https://doi.org/10.18400/tekderg.724164

1 Civil Engineering Department, Dokuz Eylul University, Izmir, Turkey - zgr_tmz@hotmail.com
https://orcid.org/0000-0002-8991-8445

2 Civil Engineering Department, Dokuz Eylul University, Izmir, Turkey - turkay.baran@deu.edu.tr
https://orcid.org/0000-0002-2325-2628

3 Civil Engineering Department, Dokuz Eylul University, Izmir, Turkey - filiz.barbaros@deu.edu.tr
https://orcid.org/0000-0002-2697-911X

decisions in the design and operation of water resources systems. However, in order to perform the measurement practice in an optimum way, the following questions must be answered: what, where, when and how often. In their studies, planners frequently use the concepts of expected information, available information, lack of information and increase in information and state that design parameters are based on the information brought by the measurements. However, there is no concrete definition of how information can be measured. Similarly, information cannot be defined directly like any statistical parameter; mostly variance, standard error or correlation coefficients are defined indirectly in terms of statistical magnitudes [1].

In order to demonstrate the measurability of the information content, the subject can be approached from another standpoint. Learning about any subject or event means learning about new situations that concern that subject or event. If previously known situations do not carry any innovation, they do not bring any new information. Therefore, if the degree of innovation of the situation in question can be measured, a criterion can be developed for the information it contains. Considering that, the simplest way to measure the degree of innovation in any situation is to determine the relative frequency of the situation therefore it can be concluded that information content can be defined by using the probability concept and probability rules [2].

Entropy is one of the methods to measure information content and has been used by researchers in wide range of disciplines, such as hydrology and water resources planning [3]. Entropy is a measure of the uncertainty or disorder of a variable. The probability mass function (p.m.f.) or probability density function (p.d.f.) of a random variable in a discrete or continuous form can be used to determine the entropy.

The entropy theory has been applied in hydrology and water resources for measuring the information content of random processes, evaluation of information transfer between hydrological processes, assessment of recharge systems for a river basin, evaluation of data acquisition systems, assessment of model performance, assessment of regional information on floods, and designing water quality monitoring network. Entropy concept is also an excellent tool for computing the information content of time series. The structure of the time series should be determined before the calculations are made. The comprehensive review of the application of entropy theory in hydrology and water resources is given by Singh [4, 5].

Determining the spatiotemporal variability of precipitation is essential in water resources research. For this reason, many researchers have contributed to this issue for years. One of them is Kawachi et al. [6], in which Shannon's informational theory is applied for determining the temporal variability of daily rainfall at 1107 stations in Japan to delineate water resources zones. Thus, the obtained entropy based climatic map presents the characteristics of nationwide rainfall of Japan.

Mishra et al. [7] analyzed monthly, seasonal and annual precipitation data sets by entropy method to assess the space-time variability of rainfall. The results indicate that the disorder in precipitation amount and the number of days with rain show a strong spatial gradient and could be related to significant historical drought periods.

Zhang et al. [8] used the optimal entropy model to analyze the spatiotemporal differences of monthly precipitation complexity in Heilongjiang, China. In this study, it has been demonstrated that agricultural production change is the main reason for the spatial difference of monthly precipitation complexity in the plain areas of the study.

Dey and Mujumdar [9] demonstrated that higher fraction of grid cells across the country shows an increasing trend in Relative Entropy (RE) which is measured as the degree of uniformity of rainfall. A significant increasing trend in RE affects the timing and amount of rainfall, which are crucial for sustainable planning and management of water in rainfed-agriculture practices.

Wang et al. [10] presented a multi-objective optimization technique based on information theory realized for a rainfall network design method in Wei basin, China. To design a multi-objective optimization technique, an integrated network design framework is developed for streamflow and precipitation networks. The results of the optimization method were further examined using an artificial neural network (ANN) model for streamflow simulation.

In information theory, uncertainty of a variable or the probability distribution of that variable (entropy) is the negative expected value of the logarithm of the probability density function of the variable. Entropy is also defined as the amount of *information gained ≡ reduction of uncertainty* through observations; in other words, the entropy criterion indirectly measures the information content. This concept is thought to be a measure of determining how much information can be gained from a measured parameter. Hence, Kawachi *et al*. [6] - *for daily precipitation* - and later Maruyama *et al* [11] - *for monthly precipitation* - used the apportionment entropy values to illustrate water resources zones and assessment of potential water resources availability.

The concept of entropy performs many roles in hydrology and water resources [12]:

i) determines the uncertainty of hydrological processes in water resources engineering,

ii) determines the probability distributions that will represent the process,

iii) transfers the information between hydrological processes,

iv) optimizes the measurement networks and defines optimal measurement intervals,

v) defines the knowledge content of mathematical models,

vi) performs the relations between processes and the operational performance of feeding systems and measurement networks, whereas providing the application of optimization and decision theory.

A review by Singh [13] discussed the frequency distributions, which have different origins based on different hypotheses and belong to different generating systems. This paper suggests various frequency distributions, which have been derived using different approaches, which include differential equations, distribution elasticity, genetic theory, generating functions, transformations, Bessel function, expansions, and entropy maximization. Each frequency distribution is used for several data analysis in hydrologic, hydraulic, environmental and water resources engineering.

These studies [6 - 11] have addressed the disorder (or uncertainty) in the intensity and occurrence of rainfall in time, which is the main constraint in water resources engineering. The most important information of developing a strategy for water use is to define the priorities, and make a decision for water allocation so as to define "water quantity". The entropy concept, which is applicable to any distribution even if the probability distribution is unknown, is a convenient tool for defining disorder of rainfall.

On the other hand, if researchers aim to define regional information without missing any information from the available data with entropy concept, one of the fundamental steps of determining entropy is to define validated regional class interval Δx for every single data set. Otherwise, the determined entropy values represent only the local information [12, 13] and the obtained regional information will be inaccurate. Another important problem for determining the regional information is the decay of data stationarity, which depends on climate change. To eliminate these problems, it is necessary to define validated regional class interval Δx and examine all of the data with trend analyses.

Hence, in the presented paper, an entropy method application on precipitation pattern is studied from a case area located in Turkey. Turkey is below the earth's average in terms of per capita water amount and among the water-stressed countries. A large number of reasons such as population growth, industrial wastes and adverse environmental conditions have affected water resources negatively. In Turkey, especially west and middle regions are in danger of drought. Among 25 river basins in Turkey, the Gediz River Basin, with its limited water resources, increasing population and industrial activities, was selected as a case study for water shortage conditions.

The objectives of this study are executing the degree of variability of the precipitation pattern using entropy theory in the Gediz River Basin and applying the intensity entropy (IE) method to monthly total precipitation series to discuss how and why entropy analysis can give reliable insights into multivariate hydrological processes. The study focused on increasingly obvious complexity evolution characteristics of the monthly precipitation series in Gediz River Basin, under the long-term influence of various natural factors and major agricultural developments. The approach in the presented study is to estimate the frequency analysis of precipitation records to define the region-wide water availability map.

The study will be significant for analyzing the variability of observed precipitation using the concept of entropy and forming a map by linking entropy and precipitation. In the following sections, applied method is presented in Section 2; available data and the study area are introduced in Section 3; the application and results are given in Section 4, and finally conclusions and future research suggestions are summarized in Section 5.

2. METHOD

Entropy is a value that can be determined by knowing or calculating the probability distribution function or probability density function of a random variable given in a discrete or continuous form using the information theory defined by Shannon [3]. Characteristically, increasing randomness or irregularity of the variable reduces the skewness of the probability distribution, thereby leading to an increase in entropy. Entropy takes the maximum value if the probability distribution is uniform without any deviation, whereas the occurrence of a specific value of the variable with a probability leads to the approach of entropy value to zero [13].

In this study, monthly rainfall intensities (monthly total rainfall) are considered as a random variable. The relationship between the probability of occurrence (frequency) of the variable at a selected time period and the monthly rainfall intensity is defined as entropy based expressions. Thus, the calculated entropy values are also called *Intensity Entropy* (IE).

An entropy value (H) for a set of probabilities $P_1, P_2, \dots, P_n$ is defined by Shannon (1948) as;

$$H = -\sum_{i=1}^{n} p_i \log p_i \qquad \text{(Eq. 1)}$$

The negative sign ensures that the result is always positive or zero. If the monthly precipitation intensity (or sum) is considered as a random variable and the probability of occurrence in a precipitation series takes a value such as $p_i$then the intensity entropy (*IE*) value at a station can be determined using equation (1). The p_i probabilities for a given station are defined for the discrete form by taking into account all available monthly precipitation values at this station and the probability of occurrence of these values. *IE* is calculated by following steps;

1. Complete years of monthly total precipitation series are collected for each station. Then, the monthly precipitation values are sorted in ascending order to form a dataset without any serial features (Hence, the total length of the series for the evaluated gauging stations will be reached, N = 12 months * Number of observation years).

2. The most important point in the frequency series is to determine what the class number will be or what the class range is. Various authors suggest various formulas about the class interval. Nevertheless, these formulas are only suggestive, they are not exact. The small number of classes leads to the loss of information provide in/by the series. It is recommended that due to the information loss, the number of classes should not be less than four. On the other hand, too many classes give difficulty in interpreting and operating the series, thus the number of classes should not be more than eight. One of the suggested rules is that the number of classes is chosen as the square root of number of observations [14]. In presented study, the observed monthly total rainfall values is divided into k equally spaced classes.

3. The class width is obtained by subtracting the smallest value from the largest value in the sequence and dividing this difference by the number of classes. However, in the proposed study, the smallest value in the sequence, that is, the smallest monthly total rainfall value is zero for each station. So that the class interval is obtained by dividing the largest monthly total precipitation value R_{max} in the ascending sorted series by the number of class expressed above (k).

   $$\Delta x = \frac{R_{max}}{k} \qquad \text{(Eq. 2)}$$

4. The frequency value for each class is calculated and a frequency distribution table is generated.

5. The relative frequency values in each range are calculated to obtain a discrete probability density function for all monthly precipitation values.

6. *IE* entropy value can be calculated using these relative frequencies and the formula is as follows.

   $$IE = \sum_{i=1}^{k} \frac{f_i}{N} ln \frac{f_i}{N} \qquad \text{(Eq. 3)}$$

   k in the Eq. 3 means the number of classes and the f_i is frequency value for each class. Since the logarithm is used at base e, the *IE* entropy value unit is “napier”.

The *IE* value, defined in the semi-infinite range $0 \leq IE < \infty$, is a measure used to express the irregularity of the monthly rainfall intensity. The less disordered intensity value is measured with a smaller *IE* value, thus indicating that the frequency distribution of the monthly precipitation value is more skewed. On the other hand, definition of monthly precipitation with a wider range leads to much more irregular density values and larger *IE* values. However, the probability density function of the precipitation intensity is defined as a positive x-axis (horizontal axis), which always includes the zero origin due to the non-negative precipitation. The increase in *IE* value creates an expectation for an increase in the expected monthly rainfall value. In other words, there is a positive correlation between the *IE* value and the expected monthly rainfall. For this reason, the *IE* value is indicative of the total amount of rainfall [11]. It is considered that observations are random and so there is no autocorrelation between them [15].

On the other hand, it is known that the selection of Δx is a crucial problem, such that each specified class interval size gives a different reference level of zero uncertainty with respect to the computed entropies. While researchers approximate the discrete probabilities p_i by using $f(x)\Delta x$, where $f(x)$ is the relative class frequency and Δx, the size of class intervals for continuous variables, various entropy measures become relative to the discretizing interval Δx and change in value as Δx changes [12, 16]. Hence, to define the regional information, same class intervals are used for all frequency analyses. If individual class ranges are selected for stations based on their own data and a specific figure is used for the whole basin stations, regional information can be obtained, as described in Fig. 1.

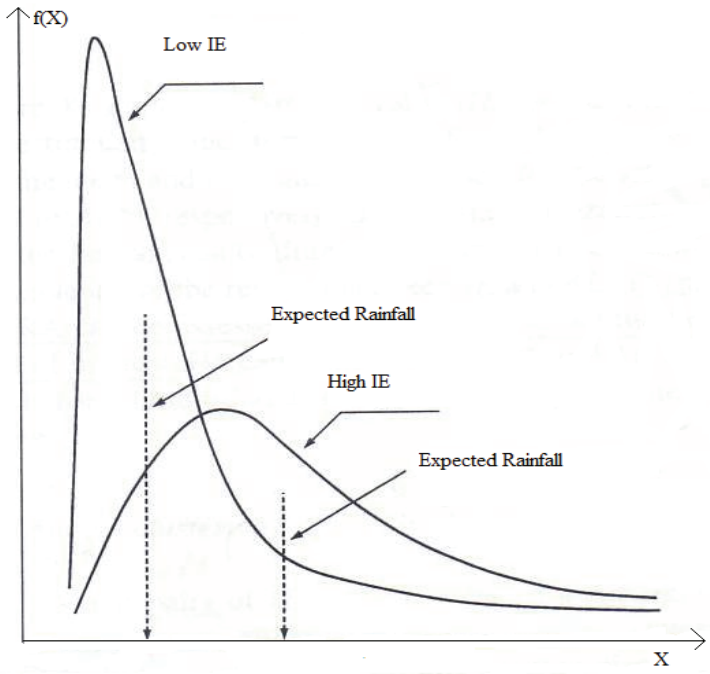

Figure 1 - Expected Precipitation with IE.

Expected rainfall values increase in the form given in Figure 1 with increasing *IE* results. This suggests that *IE* has a positive correlation with the expected precipitation and can therefore be an alternative to the prediction of total precipitation [11]. Thus *IE* values come from the frequency analyses of data sets, *IE* can be used to define the regional information even if stations have different observation periods. Other advantages of this method are;

i) there is no need to complete the missing observation data, since, no need to define mutual observation period for the basin,

ii) no need to define elevation–precipitation relations and orographic maps for homogenizing the determined expected precipitation values, since *IE* value is indicative of the total amount of rainfall at observed site.

3. DATA AND STUDY AREA

The presented study is carried out with the available precipitation data in the Gediz Basin, which is located in the western part of Turkey within the borders of the Aegean Region. The average annual precipitation in the Gediz Basin is approximately 635 mm. The Gediz basin is the 20th of Turkey's 25 major river basins, with a drainage area of 17,500 km^2 and the 21st with an annual surface water potential of 1.95 billion m^3. These values are 2.2 % of the country's surface area and 1.7 % of the country's water potential [17].

The Gediz River Basin is one of the most important water basins located within İzmir Metropolitan Municipality boundaries. The Gediz River meets both agricultural and domestic water demand of Izmir city. The Gediz delta and wetlands (İzmir Bird Paradise provide habitat for 205 bird and 308 plant species) are conservation areas protected by law. It is also of particular importance that it is an area included in the Ramsar Convention in 1997 [17].

In Gediz Basin, monthly total precipitation data of 44 gauging stations, which have more than 19 years of observation (number of monthly data are greater than 228), are evaluated. The State Hydraulic Works (SHW) operates 22 of stations and the others 22 operates by the Turkish State Meteorological Service (TSMS). The locations of the stations are presented in Figure 2.

As it is mentioned before, to define the decay of data stationarity depends on the effects of climate change, the first step all of the data sets must examined with trend analyses. As given in Tables 1 and 2, the earliest historical observation throughout the region begins in 1927 and the latest data evaluated belongs to 2006. In trend analyses, the available data from SHW and TSMS stations were investigated for a prori trends; both for monthly and annual total precipitation values. The trend analyses are performed by parametric and non-parametric methods. The results showed that most of the monthly data have negative/positive trends but the annual data series do not have statistically significant trends. To test the trend parameters, Student-t and Mann-Kendall test statistics were used [18].

The results of trend analyses showed that the annual data can be acceptable first order stationary [18]. Hence, the monthly data, which verified that there is no trend in annual period of the earliest historical observation throughout the region from 1927 to 2006 are used for defining the regional information. The unbiased statistical parameters related to the observation values of the examined stations are presented in Tables 1 and 2, for annual total precipitation values.

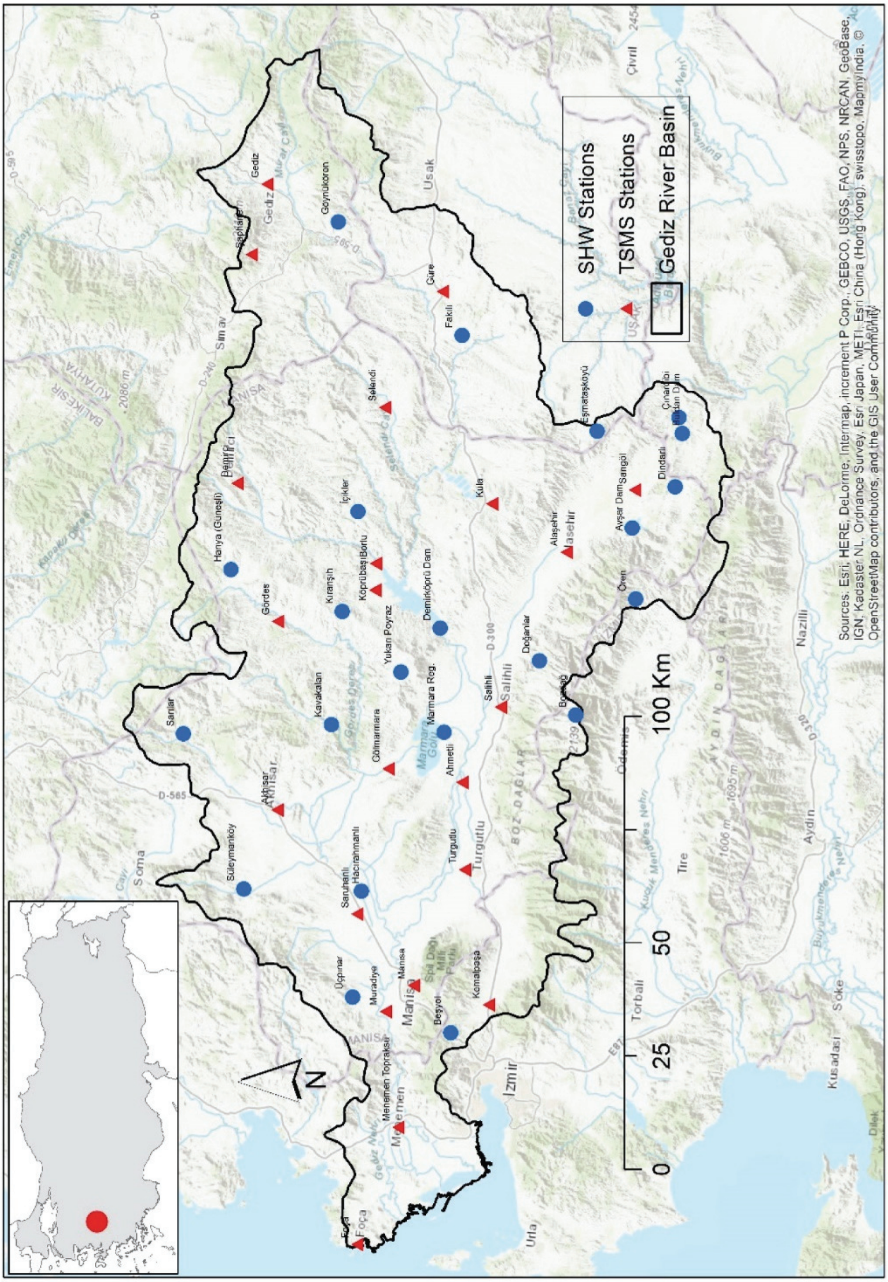

Figure 2 - Gediz Basin and the Station Locations of SHW and TSMS [18].

4. APPLICATION

Although the stations have different observation durations, Intensity Entropy (*IE*) can be used in the definition of regional information since it defines the total monthly rainfall intensity obtained as a result of frequency analysis by defining validated regional class interval Δx. Therefore, it is possible to use the available data (except very short-term stations) in the basin. However, the reliability of the data should be investigated by analyzing the stations one by one before performing frequency analysis. If the data can be acceptable as first order stationary, than *IE* values can be calculated without any extension/completion for missing years from frequency analyses. There is no need to define mutual observation period for the basin and also no need to determine precipitation/evaluation relationship for the region.

Table 1 - Observation period, Altitude and Main Statistical Parameters of Total Annual Precipitation Values of TSMS.

Station Name	Observation Period	Altitude	Mean (mm)	Std.Dev. (mm)	Skewness	Variability (%)	Kurtosis
Ahmetli	1938 – 1988	100	485.65	125.12	-1.204	25.76	2.600
Akhisar	1929 -	93	588.76	132.87	0.464	22.56	-0.198
Alaşehir	1931 -	189	482.87	111.77	0.387	23.14	-0.631
Borlu	1964 – 1982	250	541.12	100.46	0.242	18.56	-0.189
Demirci	1929 -	851	646.55	143.30	-0.014	22.16	-0.708
Foça	1957 – 1994	10	552.39	115.75	0.051	20.95	0.040
Gediz	1934 -	825	602.37	119.54	0.202	19.84	-0.531
Gölmarmara	1939 – 1992	150	546.22	129.23	0.602	23.65	-0.018
Gördes	1929 – 1997	550	638.46	151.61	0.473	23.74	0.860
Güre	1964 – 1995	650	452.26	78.89	0.105	17.44	-1.282
Kemalpaşa	1938 – 1997	200	1071.88	405.50	2.347	37.83	7.311
Köprübaşı	1967 -	250	447.80	94.63	0.415	21.13	0.418
Kula	1929 – 1991	675	590.44	142.36	0.578	24.11	0.033
Manisa	1929 -	71	727.03	162.75	0.514	22.38	-0.024
Menemen Topraksu	1929 – 1995	10	537.57	144.89	0.556	26.95	0.031
Muradiye	1964 – 1987	25	648.65	160.36	0.196	24.72	-0.388
Salihli	1939 -	111	489.44	98.32	0.485	20.09	0.109
Sarıgöl	1963 – 1986	225	486.71	99.40	0.250	20.42	-1.204
Saruhanlı	1959 – 1996	50	454.73	92.09	1.197	20.25	1.651
Selendi	1955 – 1991	575	514.41	110.21	0.199	21.42	-0.763
Şaphane	1964 – 1990	925	677.84	112.28	0.272	16.56	-0.874
Turgutlu	1929 -	120	584.02	140.34	0.578	24.03	-0.098
		MEAN	**580.33**	**135.08**	**0.404**	**22.62**	**0.279**

To define the regional information, the critical period of observations (number of years) is defined first; then raingauges which have less than the designated critical number of observation years are screened out and the selected stations which have enough data to be processed are determined significant statistical analysis were evaluated.

Table 2 - Observation period, Altitude and Main Statistical Parameters of Total Annual Precipitation Values of SHW.

Station Name	Observation Period	Altitude	Mean (mm)	Std. Dev. (mm)	Skewness	Variability (%)	Kurtosis
Avşar Dam	1980 -	275	430.53	75.55	-0.039	17.54	-1.314
Beşyol	1976 -	530	772.68	183.39	0.563	23.73	-0.061
Bozdağ	1961 -	1150	1258.62	338.44	1.014	26.88	1.600
Buldan Dam	1967 -	470	468.34	95.07	0.486	20.30	-0.625
Çınardibi	1961 – 2002	705	932.89	215.89	0.486	23.14	0.520
Demirköprü Dam	1962 – 1993	290	486.65	101.40	0.009	20.83	-0.576
Dindarlı	1962 -	685	455.36	96.10	0.298	21.10	0.045
Doğanlar	1970 -	650	630.46	152.52	0.323	24.19	0.497
Eşmataşköyü	1962 -	930	468.62	97.70	0.754	20.84	0.293
Fakılı	1962 -	715	447.13	77.95	0.436	17.43	-0.755
Göynükören	1966 – 2003	1020	467.92	133.60	-0.264	28.55	-0.290
Hacırahmanlı	1961 – 1997	45	483.33	108.81	0.778	22.51	1.084
Hanya (Güneşli)	1961 – 1995	640	634.88	157.95	0.231	24.88	-0.243
İçikler	1961 -	710	568.13	131.65	0.253	23.17	-0.742
Kavakalan	1962 – 1998	460	625.48	145.26	0.579	23.22	1.019
Kıranşıh	1962 -	670	589.24	124.79	0.320	21.17	-0.323
Marmara Lake Reg.	1961 – 2001	75	435.22	77.01	0.227	17.69	-0.239
Ören	1961 -	940	735.21	256.56	-0.212	34.89	0.731
Sarılar	1962 -	340	595.22	146.70	0.519	24.64	0.273
Süleymanköy	1962 – 1997	240	472.56	99.52	0.231	21.06	-0.708
Üçpınar	1961 -	100	547.92	165.78	0.570	30.25	0.045
Yukarı Poyraz	1962 - 2003	630	588.52	115.66	0.727	19.65	0.611
		MEAN	**595.22**	**140.79**	**0.377**	**23.08**	**0.038**

Since the presented study is an example application of the methodology proposed for developing a map of water availability, trend analysis of monthly total precipitation data (up to 2006) are used to determine the *IE* values. In the case study of Gediz Basin, 44 precipitation-gauging stations with long-term data are evaluated for determining the *IE* values. The results show that the expected (mean) monthly total precipitation and the mean annual total precipitation values have strong statistical relations between the *IE* values (r = 0.996 and r = 0.990 respectively) on the Gediz catchment basis. Consequently, *IE* values can be used to define regional information for the expected value of annual total precipitation, which have the same meaning of potential water resources availability of the Gediz basin.

4.1. Turkish State Meteorological Service (TSMS) Rainfall Observation Stations in the Basin

Throughout the twenty-three precipitation observation stations belonging to the Turkish State Meteorological Service (TSMS) in the Gediz Basin, Yarbasan was not suitable for use in entropy calculations due to its eight-year total monthly rainfall observation values. The same interval width (15 mm/month) was used at all stations for frequency analysis. The longest observed observations (76 years $*$ 12 months) belong to Akhisar and Manisa; the shortest observations (19 years $*$ 12 months) belong to Borlu station.

Relative frequency histogram and frequency density functions were obtained and plotted for all stations used in the study. In the present study, only graphs belonging to Ahmetli station are presented as an example.The absolute frequency histogram $\{f_i\}$ obtained for the total monthly rainfall and frequency density values calculated using relative frequency $\{f_i/N\}$ at Ahmetli station are presented in Figure 3.

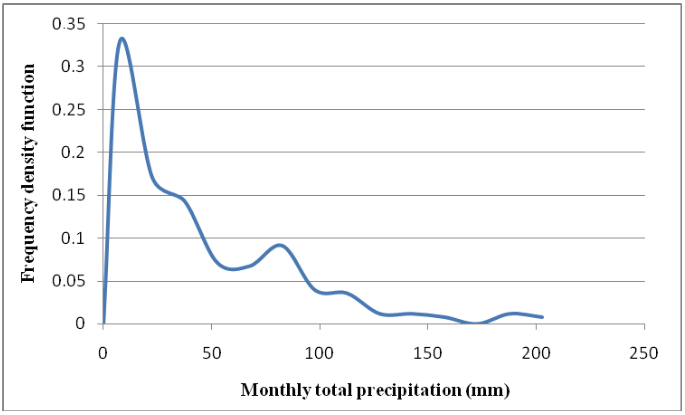

Figure 3 - Frequency Density Function of Ahmetli Station.

4.2. State Hydraulic Works (SHW) Precipitation Observation Stations in the Basin

The longest evaluated observations belong to Kıranşıh (44 years $*$ 12 months), Bozdağ (44 years $*$ 12 months) and Üçpınar (44 years $*$ 12 months) stations, and the shortest observation period (25 years $*$ 12 months) belong to Avşar Dam station. Although there are no shortcomings in the observations of the six stations used in the study, some years of the observations of the remaining sixteen stations are missing due to some observational errors. However, in the calculation of *IE*, it is considered that there is no autocorrelation between years and each data is random. Therefore, the years with observational errors or deficiencies on a monthly basis were ignored. If any year or years are missing within the observation range, the present observations were studied as if there were no missing data in the series.

4.3. Isometric Maps

The isometric map, which refers to the isometric lines of the mean annual total precipitation for observation period (Table 1 and 2) are drawn by using Surfer mapping tools by using linear interpolation approach and presented in Figure 4. Since the isohyetal maps need to define elevation–precipitation relations and use orographic maps for homogenizing of data, this map is diverged from standard isohyetal lines. Figure 4 was drawn by using determined long-term mean of the raw data for the whole observation periods.

In Figure 5, isoentropy maps are presented which are drawn by using same approach above. For isoentropy map, the determined *IE* values (Table 3) are used.

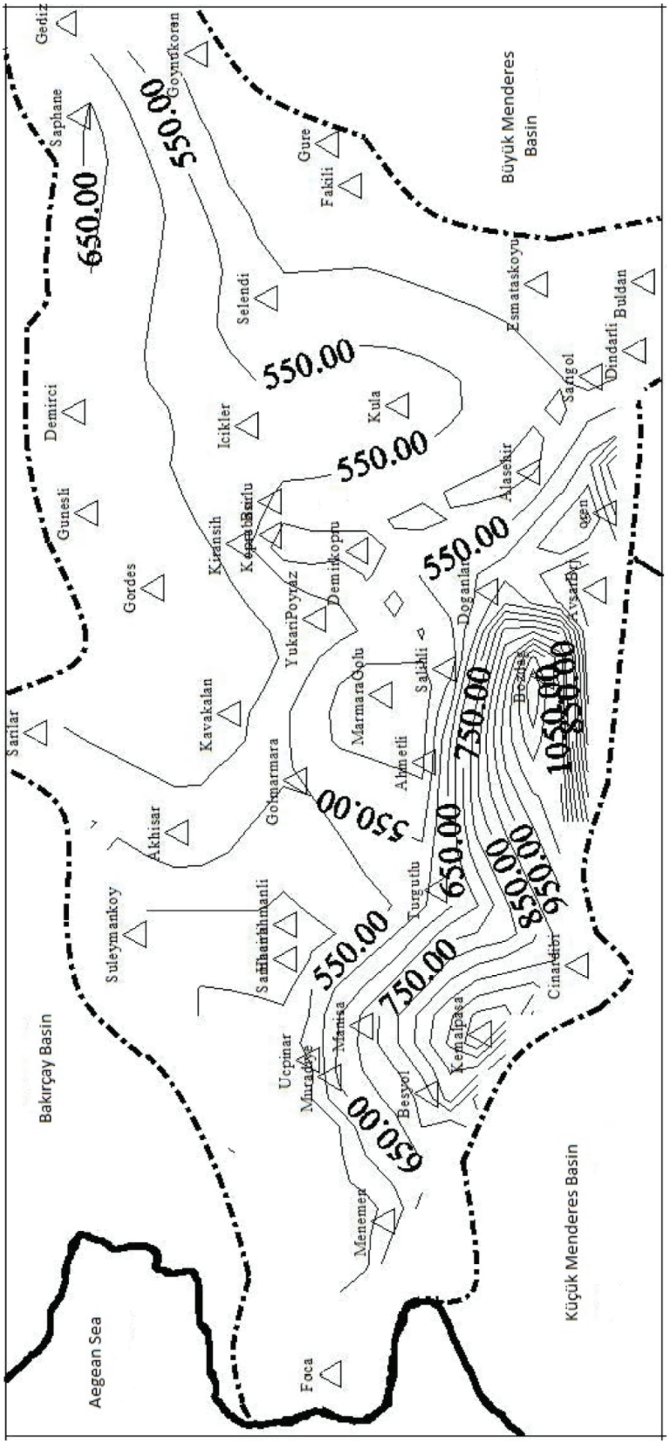

Figure 4 - Long-term Mean of Annual Total Precipitation in Gediz Basin.

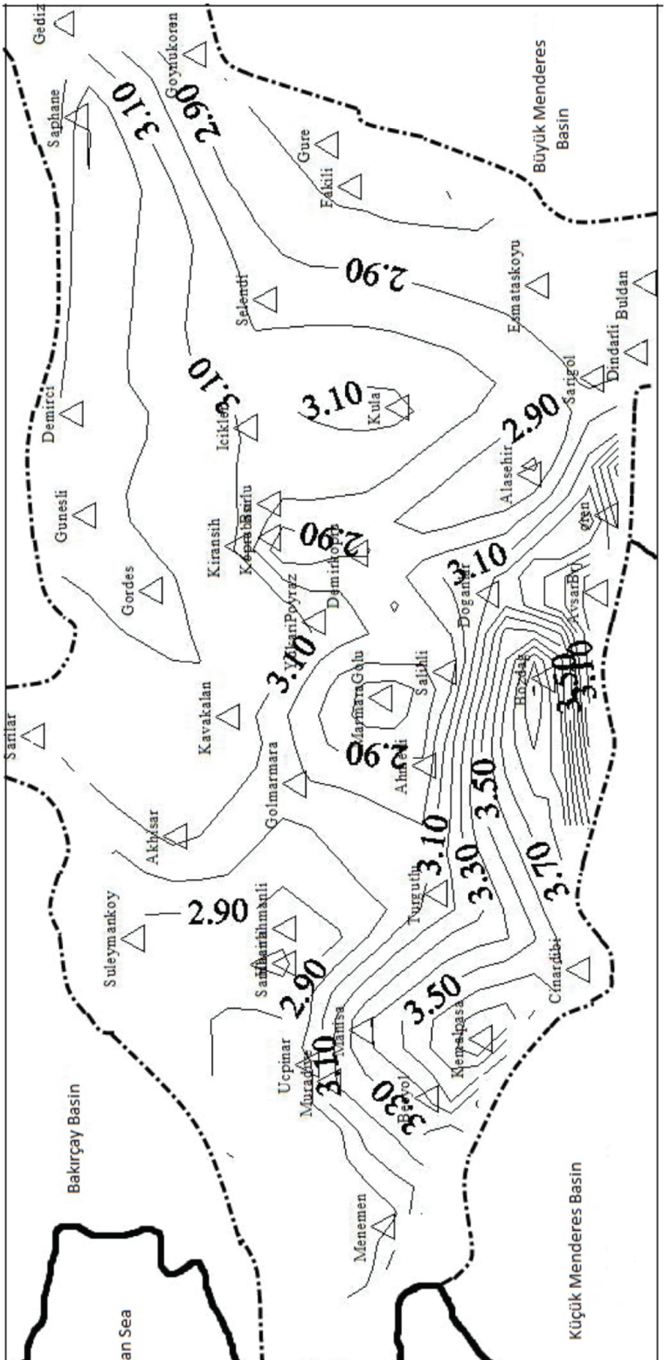

Figure 5 - Isoentropy Map for Intensity Entropy- IE in Gediz Basin.

Table 3 - Total Annual Precipitation and Entropy Values of Stations.

Station Name	Altitude (m)	Annual Total Average RainFall	IE values
Ahmetli	100	485.66	1.9667
Akhisar	93	585.11	2.1725
Alaşehir	189	467.22	1.9221
Avşar Dam	275	430.54	1.8622
Beşyol	530	772.68	2.3399
Borlu	250	541.12	2.0823
Bozdağ	1150	1258.62	2.8409
Buldan Dam	470	468.34	1.9572
Çınardibi	705	932.89	2.5249
Demirci	851	581.42	2.2826
Demirköprü Dam	290	466.32	2.0095
Dindarlı	685	455.37	1.9459
Doğanlar	650	630.46	2.2410
Eşmataşköyü	930	468.63	1.9653
Fakılı	715	447.13	1.9151
Foça	10	555.24	2.0808
Gediz	825	569.79	2.1322
Gölmarmara	150	546.23	2.1026
Gördes	550	638.46	2.2399
Göynükören	1020	467.93	1.9484
Güre	650	452.27	1.9066
Hacırahmanlı	45	483.33	1.9653
Hanya (Güneşli)	640	634.88	2.2409
İçikler	710	568.13	2.1477
Kavakalan	460	625.49	2.2097
Kemalpaşa	200	1071.88	2.6545
Kıranşıh	670	589.25	2.1659
Köprübaşı	250	447.80	1.9166
Kula	675	590.45	2.1728
Manisa	71	728.83	2.3722
Marmara Lake Reg.	75	435.23	1.8667
Menemen Topraksu	10	537.57	2.0571
Muradiye	25	648.74	2.2117
Ören	940	735.21	2.3539
Salihli	111	490.15	1.9979
Sarıgöl	225	486.72	2.0030
Sarılar	340	595.23	2.1563
Saruhanlı	50	454.74	1.9198
Selendi	575	514.42	2.0522
Süleymanköy	240	472.56	1.9604
Şaphane	925	677.85	2.3085
Turgutlu	120	584.02	2.1441
Üçpınar	100	547.93	2.0710
Yukarı Poyraz	630	588.53	2.1617

5. RESULTS

The entropy values of the precipitation gauging stations with long-term observations in the Gediz basin, define the expected values for the long-term average monthly precipitation. The mean annual total precipitation and the *IE* values, which were determined by using the monthly series (Table 3) are also performed for the mutual information. The result for the linear regression analysis between the *IE* values and mean monthly and annual total precipitation are presented in Figure 6 and 7 respectively.

At all stations evaluated both the mean monthly total precipitation (Figure 6) and mean annual total precipitation during the observation period (Figure 7) show that the results are quite consistent. The conversion correlations between intensity entropy and mean monthly/annual total precipitation show a high statistical dependence for logarithmic relationship ($R^2 = 0.991 \rightarrow r = 0.996$ for monthly values and $R^2 = 0.979 \rightarrow r = 0.990$).

The annual total precipitation values versus calculated IE values ($r = 0.99$) are evaluated by graphical analysis and the relationship as given in Eq. 4 was obtained from the graph given in Figure 7.

$$P = 52{,}58.\, e^{1{,}12IE} \qquad \text{(Eq. 4)}$$

The determined *IE* values and the regional *IE* distribution of the Gediz basin are presented in Figure 8 and 9. The regional distribution was achieved through basin-wide interpolation of the *IE* values by the Spline algorithm, and are presented in Figure 9. By using Fig. 9 and Eq. 4, researchers can estimate total annual precipitation, which also means gross water potential in the long-term period.

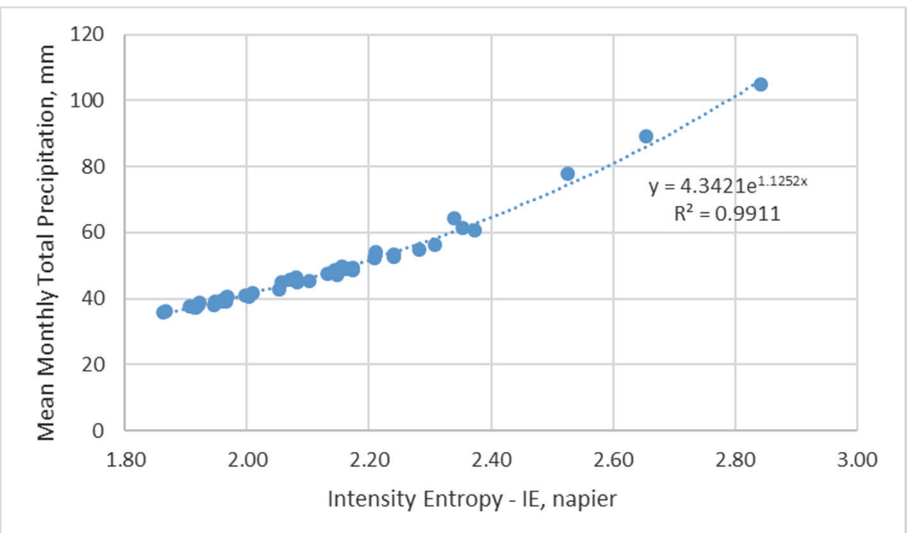

Figure 6 - The relationship between the Mean Monthly Total Precipitation and IE values in Gediz basin.

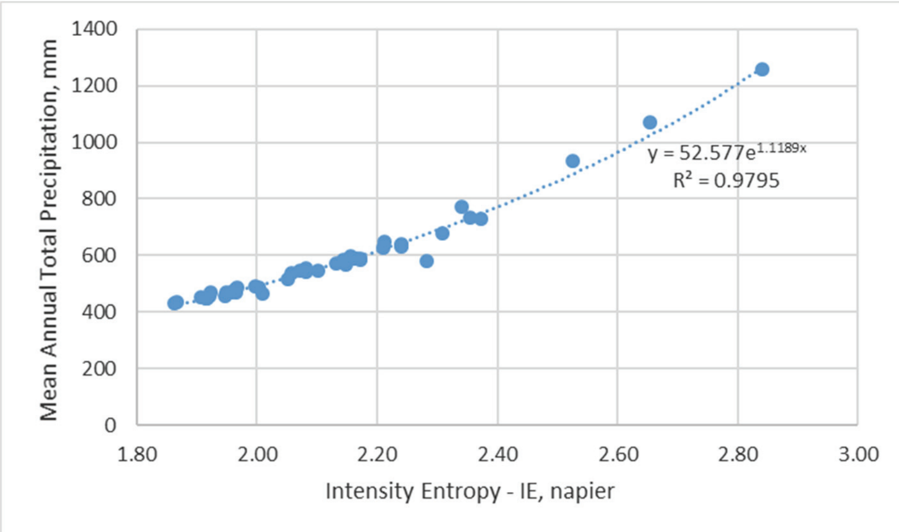

Figure 7 - The relationship between the Mean Annual Total Precipitation and IE values in Gediz basin.

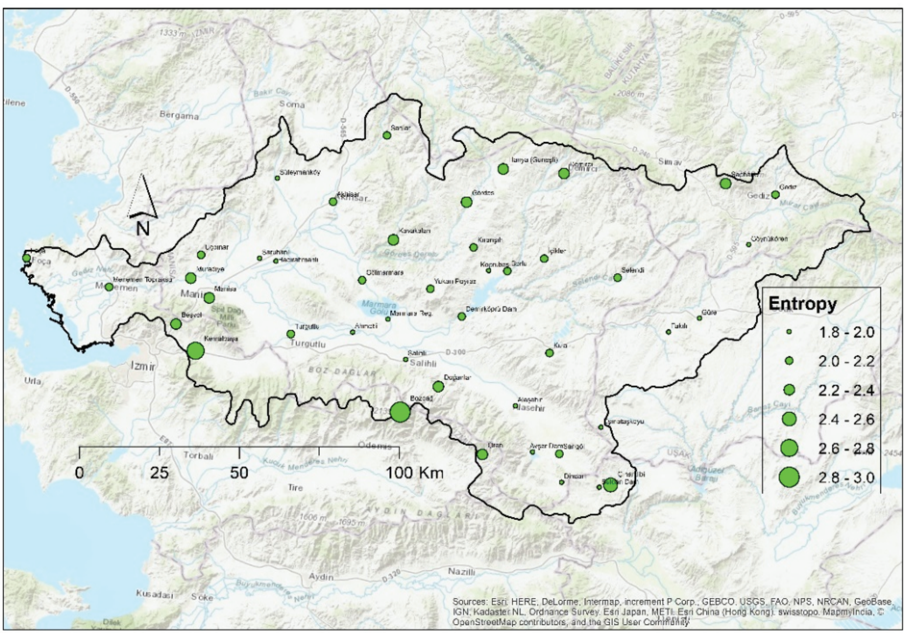

Figure 8 - The determined IE values in Gediz Basin.

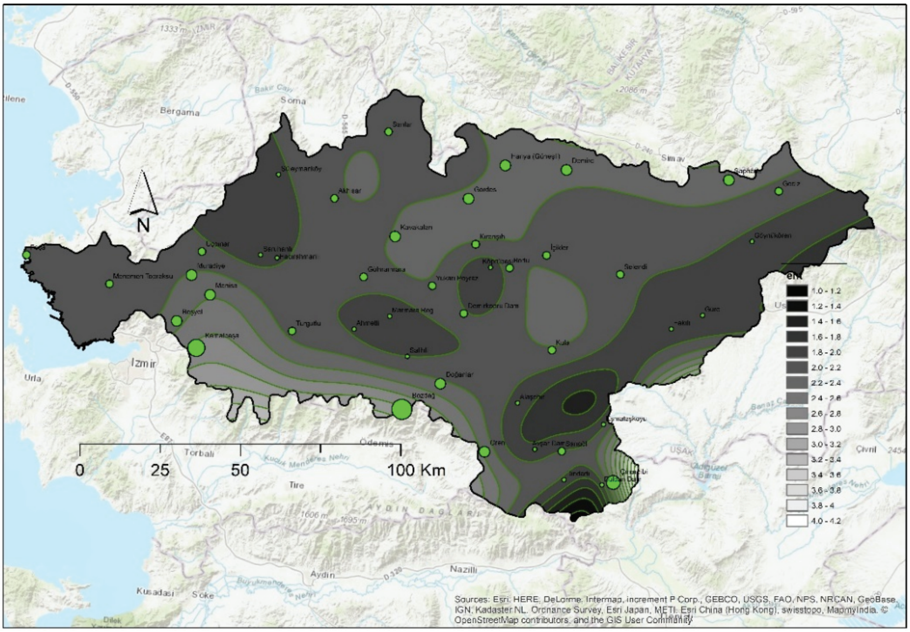

Figure 9 - Intensity Entropy -IE Distribution in Gediz Basin.

6. CONCLUSION

Employing intensity entropy method based on Shannon's informational entropy theory, a methodology is proposed for assessing potential water resources availability in terms of disorder of precipitation. Unlike traditional statistics such as mean, standard deviation, and variance, the entropy method focuses on the information content generated by variable monthly rainfall for the assessment of potential water resources availability.

Consequently, in the presented study, the expected value of long-term annual total precipitation is determined by entropy theory. The frequency analysis of the observed long-term total monthly precipitation is carried out and entropy values are determined by the histogram of frequency - "*Intensity Entropy - IE*".

Within the scope of this study, monthly rainfall intensities (monthly total rainfall precipitation) were considered as a random variable. The relationship between the probability of occurrence (frequency) of the variable at a selected time period and the monthly rainfall intensity is defined as entropy based expression. *IE* is used to define the regional information even if stations have different observation periods.

To avoid any effects of climate change, all available data were examined with trend analyses. To define regional information without missing any information from the available data and to obtain regional information accurately with entropy concept, a validated regional class interval Δx has been determined by using all available data.

The advantage of the proposed method is that it is possible for the researchers to use all the data sets from different periods in execution. In addition, a more simple-fast-reliable tool has been obtained as an alternative to similar methods.

The entropy-based methodology is applicable to any distribution whether the distribution is a priori known or unknown and although rain-gauges may not have the same number of observation years at every station.

The presented case study for the Gediz Basin offers a powerful tool to help detect water availability by relating the monthly total precipitation and intensity entropy. Local precipitation variations over the Gediz River Basin are reviewed by frequency analysis and consequently entropy concept.

Acknowledgement

Authors would like to acknowledge to Assoc. Prof. Dr. Ali GÜL for his great contribution to the mapping process of the manuscript.

References

[1] Harmancioglu N. B., Singh V. P., An Information-Based Approach to Monitoring and Evaluation of Water Quality Data. Athens, ECOWARM, European Conference on Advances in Water Resources Technology, A.A. Balkema Publishers, pp. 377-386, 1991.

[2] Cherry C., On Human Communication: A Review. A Survey and a criticism. Massachusetts, the Technology Press of Massachusetts Institute of Technology, 333 p, 1957.

[3] Shannon C.E., Mathematical Theory of Information. In The Mathematical Theory of Information, the University of Illinois Press: Urbana, IL, USA, 27: 170–180, 1948.

[4] Singh V. P., The Use of Entropy in Hydrology and Water Resources. Hydrological Processes, Hydrol. Process., 11: 587-626, DOI: 10.1002/(SICI)1099-1085(199705)11:6<587::AID-HYP479>3.0.CO;2-P, 1997.

[5] Singh V.P., The Entropy Theory as a Tool for Modelling and Decisionmaking in Environmental and Water Resources. Water S.A. 26(1):1-10., 2000.

[6] Kawachi T., Maruyama T., Singh V. P., Rainfall Entropy for Delineation of Water Resources Zones in Japan, Journal of Hydrology, 246: 36-44, DOI: 10.1016/S0022-1694(01)00355-9, 2001.

[7] Mishra A.K., Ozger M., Singh V.P., An Entropy Based Investigation into the Variability of Precipitation, Journal of Hydrology, 370:139–54, DOI: 10.1016/j.jhydrol.2009.03.006, 2009.

[8] Zhang L., Li H., Liu D., Fu Q., Li M., Faiz M., A., Khan M., I., Li T., Identification and application of the most suitable entropy model for precipitation complexity measurement. Atmospheric Research 221, 88-97, ISSN 0169-8095, DOI: 10.1016/j.atmosres.2019.02.002, 2019.

[9] Dey P., Mujumdar P.P., On the uniformity of rainfall distribution over India, Journal of Hydrology, 578, 124017, ISSN 0022-1694, DOI: 10.1016/j.jhydrol.2019.124017, 2019.

[10] Wang W., Wang D., Singh V.P., Wang Y., Wu J., Zhang J., Liu J., Zou Y., He R., Information theory-based multi-objective design of rainfall network for streamflow simulation. Advances in Water Resources. 135-103476, ISSN 0309-1708, DOI: 10.1016/j.advwatres.2019.103476, 2020.

[11] Maruyama T., Kawachi T., Singh V.P., Entropy-based Assessment and Clustering of Potential Water Resources Availability. Journal of Hydrology, 309, 104-113, DOI: 10.1016/j.jhydrol.2004.11.020, 2005.

[12] Baran T., Hidrolojik süreçlerin bilgi içeriğindeki değişim miktarı olarak entropi tanımı. Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İnşaat Mühendisliği Bölümü, Hidrolik-Hidroloji ve Su Kaynakları Doktora Tezi No: 10, 1993

[13] Singh V.P., Systems of Frequency Distributions for Water and Environmental Engineering. Physica A: Statistical Mechanics and its Applications, 506: 50-74, ISSN 0378-4371, DOI: 10.1016/j.physa.2018.03.038, 2018.

[14] Hines W.H., Montgomery D.C., Goldsman D.M., Borror C.M., Probability and Statistics in Engineering. John Wiley & Sons, 655 pp., ISBN 0-471-24087-7, 2003.

[15] Temiz Ö., Baran T., Determination of Expected Value for Monthly Total Precipitation by Entropy Based Method Case Study: Gediz Basin. 10th International Congress on Advances in Civil Engineering, Paper No: 769, 10 p, 2012.

[16] Baran T., Harmancioglu N.B., Cetinkaya C.P., Barbaros F., An Extension to the Revised Approach in the Assessment of Informational Entropy, Entropy, 19(12): 634, DOI: 10.3390/e19120634, 2017.

[17] SHW, Environmental Master Plan of Gediz Basin, II. Directorate of State Hydraulic Works, Izmir, 2005.

[18] Baran T., Temiz Ö., Gediz Havzası Yağışlarının Eğilim Analizi. IV. Ulusal Su Mühendisliği Sempozyumu, Bildiriler, s. 241 – 251, 2009.