Employing nonlinear dynamic concepts for catchment classification using runoff response of catchments

Hadi Delafrouz[1], Abbas Ghaheri[2] and Mohammad Ali Ghorbani[3]

1. PhD candidate, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran,

2. Professor, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran,

3. Associate professor, Department of Water Engineering, University of Tabriz, Tabriz, Iran.

Corresponding Author: H. Delafrouz

Email: delafrooz@gmail.com

Mobile number: +989125379066

Tel number: +982166906467

Fax number: +982166906468

Abstract:

Classification has been considered as a fundamental step towards improved science and management data. Introducing methods which describe the underlying dynamics of runoff could be a promising way for catchment classification. In this respect chaos theory and correlation dimension was applied to test its ability to construct a concept to introduce a catchment classification framework in this study. The correlation dimension, as an indicator, was calculated for the daily river flow of sixty grouping stations in different catchments in Iran, ranging in size from 8 to 36500 (km^2). The results confirmed that applying this indicator to catchments in varied ranges, from low to high complexity, can also be classified. The results showed that Iran's catchments can be classified into four groups based on the complexity degree of runoff time series. The group is as flows: low dimension (D_2 <= 4) as group 1, medium dimension (D_2 = 5) as group 2, high dimension (D_2 => 6) as group 3 and unidentifiable as group 4.The spatial pattern classification of Iran's catchments indicates that catchments with different climate characteristics which are located at a far distance from each other might yield similar responses along with the same level of complexity.

Keywords: Catchment Classification, Chaos theory, Correlation dimension, System Complexity, Streamflow variability

1. Introduction

Integrated catchment management is of great importance to manage the resulted runoff of rainfalls. Several attempts have been made to develop new approaches for hydrologic planning and watershed management purposes. Moreover, the invention of powerful computers and measurement devices leads to a more rapid and significant growth of these approaches. Despite all these attempts, the hydrological modelling state has not changed significantly and seems to need more investigations. Although the model complexity presented is increasing, there are no significant changes in its efficiency and accuracy. As stated by McDonnell and Woods [1] there is no clear guidance to verify which model or model structure is suitable for each catchment type or management questions. In other sciences, this problem was solved by classification concepts such as Carolus Linnaeus' organism classification system in biology, the periodic table in chemistry, or dimensionless numbers in fluid dynamics. Gani et al. studied 45 methods for classifying "big data"[2]. Basically, the classification can provide valuable information and is one of the measures of the science maturity [3]. A classification framework in which each catchment increases the understanding of the definition of catchment functions is necessary [4]. Therefore, an appropriate classification framework would be beneficial in catchment management to development and application of a suitable model in hydrology science.

Two general approaches of catchment classification are categorized as deductive and inductive [5]. The deductive approach is based on studying and analysing the effects of the environmental influences (such as catchments' physical, hydrological, and climatic similarities) on discharge. In the inductive approach, hydrologic time series are analysed directly to identify the *"complexity"* (see Sivakumar and Singh [6] for further information) and dynamic pattern properties of discharge (see Figure 1).

According to the complexity of the rainfall-runoff phenomenon, catchments show a varied range of response behaviours, even if they are similar or situated near each other. The effects of all factors in the catchment system are embedded in the discharge time series which reflects both climatic and morphologic characteristics. Therefore, in cases with time series data available, the inductive approach seems to be more suitable to find catchments similarities and classification.

There are two main procedures in the inductive approach. The first procedure is based on statistical methods, which uses data properties such as size, frequency, duration of seasons and variation of flow regime characteristics (i.e. seasonal pattern, frequency and duration of droughts, floods, annual runoff variability and variation rate) [5,7-10].

The second procedure is based on the application of nonlinear dynamics and complexity concepts to discharge time series. Zoppou et al. [11] applied wavelet spectral analysis to some Australian runoff stations and classified the basins into twelve categories by using the spectral properties of the outflow time series. Krasovskaia [12] classified the hydrologic time series in Scandinavia by using the entropy concept. Sen [13] used Lempel-Ziv Complexity (LZC) measures to analyse the complexity of monthly runoff data in England and proposed that the catchments can be classified using this measure. Chou [14] applied sample entropy (SampEn) to rainfall and runoff time series to study the evolution of complexity of different temporal scales.

To introduce the concepts of chaos theory, as well as the reconstruction of the phase-space in hydrology, together with the similarities between the hypothesis and hydrologic phenomena such as rainfall-runoff, a number of researchers proposed the use of this theory for classification. Sivakumar [15] proposed the use of nonlinear concepts to identify the number of dominant processes in hydrologic systems. Sivakumar et al. [16] showed the efficiency of phase-space reconstruction in identifying deterministic and random time series with two artificial time series, and proposed the use of this approach for catchment classification. They also analysed and tested this idea for daily and monthly streamflow time series under varying conditions. Sivakumar and Singh [6] classified a number of catchments in the western United States by reconstructing the phase-space and employing the correlation dimension (D_2) of monthly discharge time series. They classified the discharge of the catchments into four categories based on the following correlation dimension values: low dimension ($D_2 < 3$), medium dimension ($3 < D_2 \leq 6$), high dimension ($D_2 > 6$) and unidentifiable (D_2 not identifiable). They recommended testing this approach in other parts of the world with different time scales for future researches. Sivakumar et al. [17] classified a number of monthly rainfall time series in Western Australia into five groups based on correlation dimension values. Vingnesh et al. [18] employed false nearest neighbours to study spatial variability of streamflow over across the United States.

During recent years, there has been an increasing interest on developing a framework for catchment classification and thus a number of attempts have been made to reach a more efficient modelling [see 4, 6, 18-20]. Our investigation of the literature indicates that there are only a few studies investigating the spatial variation of discharges data at different locations with spatial climatic variability, through employing the concept of chaos theory (see [6] and [21] as examples). To the best of authors' knowledge, there is only one study in which Vignesh et al [18] used the chaos method to classify the United States catchments over the whole area of the USA. Therefore, we believe that the present study is suitable because of the number and distribution of data tested. In addition, since Iran is a developing country where there are obviously problems with recorded data, this study can appropriately identify this classification pattern.

The results of this research can be used to convert large volumes of information into smaller and homogeneous volumes for easy modelling, complete missing and incomplete data and expanding point information to regional information for ungauged sites, using raw data instead of calculating index or other criteria, grading basins using the basins functions instead of flow regimes, easily determining the complexity of stream flows using the nonlinear dynamic method in order to classify the basins and suggest appropriate models, and determining suitable basins for the models and appropriate models for homogeneous basins.

This approach is used to classify the catchments in the whole area of Iran with different climates across the entire country. The mentioned method is applied to daily discharges of sixty stations which are spatially distributed all over the country. As Iran is subject to a wide range of climatic variety, the country's catchments have been chosen to examine the ability of the chaos concept for different climatic catchment classifications.

This study aims to present a classification pattern based on chaos theory and dynamic phase-space reconstruction. For this purpose, the following are addressed: a) employ the phase-space reconstruction concept and calculate the correlation dimension (D_2), b) evaluate D_2 as a criterion for measuring the complexity and classification, c) present a pattern for the classification (identify the limit of group) .

In the next sections of the paper, the methodology including dynamic phase space, average mutual information and correlation dimension are described. Subsequent sections present the study area and data. The results obtained via the proposed modelling framework are then presented and discussed, followed by the conclusion.

2. Reconstruction of Dynamic Phase Space

The concept of phase space is a useful tool for conducting research into the dynamic systems and several researchers have used this concept to study various hydrological phenomena in the past two decades [22]. According to this concept, a dynamic system can be represented by a set of effective variables influencing the system's behaviour, and is modelled by a phase space diagram, in which each point represents the system behaviour at a specific time. The common method was proposed by Takens [23]. According to this method, given one single-variable time series present in a system, namely X_t, the state vector can be reconstructed as:

$$Y_t = (X_t, X_{t-\tau}, X_{t-2\tau}, \ldots, X_{t-(m-1)\tau}) \quad (1)$$

where m is the dimension of vector Y_t, called embedding dimension; τ is referred to as the delay time and is usually used as an appropriate multiple of the sampling interval for discrete time series. The dimension m can be considered as the minimum number of variables required to represent the system.

The Grassberger-Procaccia method [24], which is one of the common methods for embedding dimention estimation and chaos identification, is used in this study. Average Mutual Information (AMI) and Auto Correlation Function (ACF) methods are two typical methods which are used to estimate the delay time. Since the runoff generation of a catchment is a complex and non-linear process, it is better to use a non-linear method to estimate the delay time. Therefore, the AMI method which has a non-linear structure was used to calculate the delay time in this study. A brief description of the AMI and correlation dimension methods is presented below.

2.1 Average Mutual Information

This method measures the amount of information about the value of $X(t+\tau)$, if the value of $X(t)$ is known [25]. For a discrete time series, with $X(t)$ and $X(t+\tau)$, the average mutual information (AMI), $I(\tau)$, is defined as:

$$I(\tau) = \sum_{X(t), X(t-\tau)} P(X(t), X(t+\tau)) \times \ln\left[\frac{P(X(t), X(t+\tau))}{P(X(t)) \times P(X(t+\tau))}\right]$$

(2)

In which $P(X(t))$ and $P(X(t+\tau))$ are individual probabilities of $X(t)$ and $X(t+\tau)$, respectively, and $P(X(t),X(t+\tau))$ is the joint probability density. In this method, the time of the first minimum in average mutual information function is defined as the optimum delay time.

2.2 Correlation Dimension

The correlation dimension method is also named as correlation integral analysis. The correlation integral *C(r)* is calculated using Grassberger-Procaccia approach [24] as the common method. According to this method, for an m-dimensional phase space, the correlation integral *C(r)* is defined as:

$$C(r) = \lim_{N\to\infty} \frac{2}{N(N-1)} \sum_{\substack{i,j \\ (1\le i<j\le N)}} H\left(r - \| Y_i - Y_j \|\right) \tag{3}$$

where H is the Heaviside function, with $H(u)$=1 for u>0 and $H(u)$=0 for $u\le 0$, where $u=r-|Y_i - Y_j|$, r is the radius of the sphere centred on Y_i or Y_j and N is the number of points on the reconstructed attractor. If the time series is described as an attractor, for positive values of r, the correlation function *C(r)* and *r* are related to each other as:

$$C(r) = \alpha r^{D_2} \tag{4}$$

where α is a constant value and D_2 is a correlation exponent or the slope of ln *C(r)* versus ln *r* given by:

$$D_2 = \lim_{r\to 0} \frac{\ln C(r)}{\ln r} \tag{5}$$

This gradient is generally calculated by fitting a line in the main range of *r* (i.e., scaling range). The type of the system can be determined by studying the behaviour of D_2 versus dimension *m*. If D_2 varies linearly with increasing m without achieving a saturation value, the system can be considered as a stochastic system. In contrast, if D_2 is saturated at a definite value of *m*, the system is considered as a deterministic system [25].

3. Data Sets

The study area is the whole area of Iran consisting of different climate catchments. Iran's territory is divided into six main catchments called Sarakhs, Caspian Sea, Orumieh Lake, Persian Gulf, Gulf of Oman, and Central Basin. The climate in each catchment is completely different from the other. The average rainfall in these catchments varies from 337 mm/year in the Caspian Sea catchment to 102 mm/year in the Oman Sea catchment. In this study the catchment areas vary from 8

km^2 (station 19-148 at Ardebil province) to 36500 km^2 (station 24-029 at Bushehr province). The daily recorded runoffs data ranging from 18 to 60 years were recorded in sixty 60 synoptic stations were used in this study (see Figure 2). As shown in Table 1, the properties of streamflow, such as maximum and average, are of high variation values.

The spatial distribution of the selected stations and their climate properties are shown in Figure 2. Table 1 shows the elevation, area and the outflow properties of each station. As can be seen from Figure 2 and Table 1 a wide range of topography, land use, geology, soil types, and climate data are used in this study. The minimum and maximum elevations are -23 and 2068 meters, respectively. The basins area ranges from 8 to 36500 km^2. The maximum flow rate is between 0.95 and 2866 m^3/s and the average flow rate is between 0.07 and 48.33 m^3/s.

4. Results and Discussion

The correlation dimension of each time series in each station was calculated by using the Grassberger-Procaccia method (Table 1). The comments of some researchers in the calculations of the correlation dimensions have been considered in this study [26-29]. Three stations (21-163, 21-966 and 19-055) were selected as examples to illustrate the correlation dimension explanation.

Figure 3 shows how the scalar time series is reconstructed in a higher dimensional phase-space, according to Eq. (1), to represent the underlying dynamics. In fact, Figure 3 represents the phase-space diagram in two dimensions (m=2) with a delay time of τ=1, i.e., the projection attractor on the plane (X_i, X_{i+1}). As can be seen in Figure 3, while the projection produces a well-defined structure attractor in station 19-055 (Figure 3-c), the structures of the station 21-163 are scattered in the phase space (Figure 3-a). The state of station 21-966 is approximately between the other two stations (Figure 3-b). It should be noted that according to the presented projections in Figure 3, the complexity of the system in station 19-055 is less than that of station 21-163. Therefore, the station 19-055 would be an approximate deterministic system while the station 21-163 is a stochastic one. This result was verified by using the correlation dimension results which are discussed in the following step.

The delay time was estimated using the AMI method. Then the correlation function for each time series was calculated using the embedding dimensions (m) values from 1-40 and the calculated delay time. Figure 4 shows the relationship between the correlation function $C(r)$ and the radius r for the increasing values of m (log $C(r)$ versus log r). In Figure 4, there is a middle region where correlation function $C(r)$ starts to stray from linearity. The correlation function from these regions is calculated by least square error method as shown in Figure 5. Moreover, the relationship between the correlation exponent and the embedding dimension values is shown in Figure 5.

As can be seen in Figure 5 while the correlation exponent values increase with embedding dimension values without any bound in station 21-163, the correlation exponent is saturated in a finite value in station 19-055. The infinite increase of the correlation exponent in station 21-163 is an indicator of stochastic behaviour in the streamflow time series. In contrast, the correlation dimension in station 19-055 is saturated at D_2 =3, suggesting the possible presence of the chaotic behaviour in the streamflow time series. In the station 21-966 the correlation dimension is equal to 5 and as can be observed in Figure 3 the behaviour of this station is somehow between the other two stations. The results of the other investigated stations are presented in Table 1. The sign "∞" in the Table shows the indefinability of the correlation dimension parameter.

According to Table 1, the correlation dimension values have a wide range of variation between 3 (very small) and ∞. Considering the stations 29-011 and 12-033 with undefined values of D_2, it can be seen that these stations are located far from each other and also in two different climates (dry and wet regions, respectively) in Iran (see Figure 2). This condition occurs with a number of other couples of stations, such as 47-023 and 24-021. In contrast, two other stations (i.e. 18-091 and 18-093) which are located near each other and have similar climates, elevations and areas have different correlation dimension values. Some important points can be concluded from these results. Firstly, the parameter D_2 can be used to identify inherent characteristics of the time series that are different due to the prevailing conditions in every catchment and can be used to determine the similarity or dissimilarity of the two considered stations. Secondly, the similarity in the location or climate of the stations does not result in similarities in their systems' behaviours.

As already mentioned, the correlation dimension values contain important information about the behaviours of the fundamental system dynamics that could be distinguished between catchment hydrological behaviours. Therefore, the correlation dimension values can be used as base criteria for the foundation of a catchment classification pattern. In this manner, catchments would be categorized into groups and subgroups using their similar hydrologic behaviours which in turn are specified by correlation dimensions calculated for monthly outflow data.

With regards to dimensionality of the studied stations in Table 1, it can be seen that the catchments in the study area range widely from stochastic (undefined) to deterministic (low dimensions). Stations in the study area can be divided into four groups based on the correlation dimensions: low dimension (D_2 <= 4) as group 1, medium dimension (D_2 = 5) as group 2, high dimension (D_2 => 6) as group 3 and unidentifiable as group 4. The selection of the group's number and the dimension value ranges are nearly arbitrary. However, the limit of the dimension values for each group is reasonable, because too many groups with very small differences or only two groups presenting just high and low dimension values are not efficient for the purpose of classification.

Figure 2 presents the classification of stations in the study area into four mentioned groups which indicates the variability and patterns across the study area. The number of stations is not large enough and more data are needed to analyse the regional results more confidently, but the stations are scattered all over the country and homogeneity can be seen among them.

Having considered Figure 2, a clear relationship between the correlation dimension and the spatial pattern could not be established, but a general view of the spatial pattern of groups and similarities is detectable. In southeast Iran, the stations are classified as group 4. The stations located in the south part with a semi-dry climate belonging to group 4 decrease in group number (degree of complexity) as you move closer to the sea and dry climates. Moreover, none of the stations in the south part of Iran belong to group 1. Overall, the stations located in the west part of Iran mostly belong to groups 1 and 2. Although all of the groups are observed in the northwest, most stations with wet or very wet climates in this region belong to group 1. Most stations with wet (very wet, wet, and semi-wet) climates that are located in the north part of Iran are classified as group 1.

Evident from an analysis in this study is while the runoff dynamic of the south eastern stations exhibit very high complexity, more stations located in very wet climates (especially in the north part of Iran) are less complex than the other stations. The tributaries of the south eastern rivers, unlike the north western rivers, are mostly seasonal. Thus, changes in discharges of the south eastern rivers are very considerable which lead to increase their system's complexity. Furthermore, the climate changes resulting from the Oman sea currents affect the rivers flow significantly. In general, a decrease in rainfall (drought increasing) leads to an increase in the complexity value.

As mentioned, one of the advantages of this classification pattern is selecting the appropriate model before starting the modelling operation. In other words, by determine the classification group the user can predict the appropriate model such as stochastic, neural network, etc. At this stage, it is required that the behaviours of different time series be compared with other groups while considering the model, and researchers, in addition to presenting the accuracy criteria, should mention the classification criteria so that other researchers can specify which model is appropriate for which complexity level (or classification group). As an example, Tongal and Berndstoon [30] studied the impact of streamflow complexity on the forecasting result of three type models (chaos, stochastic and black-box). They showed that determining the degree of complexity can help with the pre-determination of model to be applied. Therefore, the proposed classification framework can be a proper candidate for selecting the type and quantity of the model. While more than six effective variables ($D_2 \geq 6$) would result in a high complexity requiring the use of more complex models, four or fewer variables ($D_2 \leq 4$) would lead to a lower complexity which can be analysed by simpler models. As an example, the obtained results in this research suggest that the stochastic models could be a first-order selection to model the south-eastern stations in Iran.

Figure 6 shows how to use the classification pattern. Accordingly, after calculating the delay time, D_2 is calculated and based on its value the time series in classification is specified. Having the times series specified, this classification can be used in cases such as selecting the appropriate model for the time series modelling (and other classification benefits).

5. Conclusion

An important hydrological challenge is to establish a catchment classification framework which serves as a basis for many purposes in hydrology (e.g., for generalization and developing models). In this regard, finding a sensitive and interpretable criterion that could be used for properly distinguishing various catchments is required. In this study, the capability of the chaos theory using the correlation dimension value (D_2) was tested for this purpose.

The results of this study showed that the application of the correlation dimension is suitable for assessing the differences in catchment behaviours (runoffs) resulting from various climatic and hydrological conditions. Based on the discussed results, this criterion can be used as an index or a measure to obtain a catchment classification framework. In the present study, the result of the dimensional analysis was used to categorize Iran's catchments into four groups based on the degree of time series complexity: low dimensional (low complexity), medium dimensional (medium complexity), high dimensional (high complexity) and unidentifiable (very high complexity or stochastic).

The spatial pattern classification of Iran's catchments indicates that the catchments with different climate characteristics which are located at a far distance from each other might yield similar responses along with the same degree of complexity. Furthermore, spatial plot results showed some degree of homogeneity and connectivity between catchments.

The presented classification method in this study can be used by engineers and researchers to study hydrological modelling methods more effectively through matching the degree of complexity between models and catchment responses (streamflow time series).

In future studies, finding a new approach that could suggest definitive guidelines in specifying the boundaries of groups and sub-groups in a classification framework is necessary. In addition, the efficiency of classification in an existing model could be studied. The time scale plays one of the most important roles in dimensionality. Thus, every time scale could be studied to identify the group and sub-group boundaries. The ability of these criteria must be tested with different climatic and catchment properties in other countries. For future study, it is suggested that the results of this study with the results of another classification method be compared and validated. Applying the method under climate change conditions and multivariate chaotic method in order to optimize the results obtained by using data of precipitation, runoff, sediment and etc. can be another issue. We emphasize that the finding of “similarity measures” or an agreed-upon “classification framework” should come first and be developed more.

References:

[1] McDonnell, J.J. and Woods, R.A. “On the need for catchment classification”, *Journal of Hydrology*, 299, 2-3 (2004).

[2] Gani, A., Siddiqa, A., Shamshirband, S. and Hanum, F. “A survey on indexing techniques for big data: taxonomy and performance evaluation”, *Knowledge and Information Systems*, 46, 241–284 (2016).

[3] Grigg, D.B. “The logic of regional systems”, *Annals of the Association of American Geographers*, 465–491 (1965).

[4] Wagener, T., Sivapalan, M., Troch, P.A. and Woods, R.A. “Catchment classification and hydrologic similarity”, *Geography Compass*, 1(4), 901–931 (2007).

[5] Olden, J.D., Kennard, M.J. and Pusey, B.J. “A framework for hydrologic classification with a review of methodologies and applications in ecohydrology”, *Ecohydrology*, 5, 503-518 (2011).

[6] Sivakumar, B. and Singh, V.P. “Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework”, *Hydrology and Earth System Sciences*, 16, 4119–4131 (2012).

[7] Zhang, Y., Xia, J., Bunn, S.E., Arthington, A.H., Mackay, S. and Kennard, M. “Classification of flow regimes for environmental flow assessment in regulated rivers: the Huai River Basin, China”, *River Research and Applications*, 28, 989-1005 (2011).

[8] Bejarano, M.D., Marchamalo, M., De Jalón, D.G. and De Tánago, M.G. “Flow regime patterns and their controlling factors in the Ebro basin (Spain) ”, *Journal of Hydrology*, 385, 323–335 (2010).

[9] Monk, W.A., Wood, P.J., Hannah, D.M., Wilson, D.A., Extence, C.A. and Chadd, R.P. “Flow variability and macro invertebrate community response within riverine systems”, *River Research and Applications*, 22, 595–615 (2006).

[10] Rao, A.R. and Srinivas, V.V. “Regionalization of watersheds by fuzzy cluster analysis”, *Journal of Hydrology*, 318 (1), 57–79 (2006).

[11] Zoppou, C., Nielsen, O.M. and Zhang, L. “Regionalization of daily streamflow in Australia using wavelets and K means analysis, *CMA Research Report MRR02-003*”, *Australian National University*, Canberra Available at: http://wwwmaths.anu.edu.au/research.reports/mrr/02/003/ (2002).

[12] Krasovskaia, I. “Entropy-based grouping of river flow regimes”, *Journal of Hydrology*, 202, 173–191 (1998).

[13] Sen, A.K. “Complexity analysis of river flow time series”, *Stochastic Environmental Research and Risk Assessment*, 23, 361-366 (2009).

[14] Chou, C., “Complexity analysis of rainfall and runoff time series based on sample entropy in different temporal scales”, *Stochastic Environmental Research and Risk Assessment*, 28, 1401-1408 (2014).

[15] Sivakumar, B. “Dominant processes concept in hydrology: moving forward”, *Hydrological Processes*, 18, 2349-2353 (2004).

[16] Sivakumar, B., Jayawardena, A.W. and Li, W.K. “Hydrologic complexity and classification: a simple data reconstruction approach”, *Hydrological Processes*, 21, 2713–2728 (2007).

[17] Sivakumar B., Woldemeskel F.M. and Puente C.E. “Nonlinear analysis of rainfall variability in Australia”, *Stochastic Environmental Research and Risk Assessment*, 28, 17-27 (2014).

[18] Vignesh R., Jothiprakash V. and Sivakumar B. “Streamflow variability and classification using false nearest neighbour method”, *Journal of Hydrology*, 531, 706-715 (2015).

[19] Hrachowitz M., Savenije H.H.G., Blöschl G., McDonnell J.J., Sivapalan M., Pomeroy J.W., Arheimer B., Blume T., Clark M.P., Ehret U., Fenicia F., Freer J.E., Gelfan A., Gupta H.V., Hughes D.A., Hut R.W., Montanari A., Pande S., Tetzlaff D., Troch P.A., Uhlenbrook S., Wagener T., Winsemius H.C., Woods R.A., Zehe E. and Cudennec C. "A decade of predictions in ungaged basins (PUB) – a review", *Hydrological Science Journal*, 58, 1198–1255 (2013).

[20] Sivakumar B., Singh V.P., Berndtsson R. and Khan S.K. "Catchment classification framework in hydrology: Challenges and directions", *Journal of Hydrologic Engineering*, doi: 10.1061/(ASCE)HE.1943-5584.0000837 (2015).

[21] Liu Q., Islam S., Rodriguez-Iturbe I. and Le Y. "Phase-space analysis of daily streamflow: characterisation and prediction", *Advances in Water Resources*, 21, 463–475 (1998).

[22] Fan, Q., Wang, Y. and Zhu, Li. "Complexity analysis of spatial–temporal precipitation system by PCA and SDLE", *Applied Mathematical Modelling*, 37, 4059-4066 (2013).

[23] Takens, F. "Detecting strange attractors in turbulence", *lecture notes in mathematics*, 898, 366-381 (1981).

[24] Grassberger, P. and Procaccia, I. "Measuring the strangeness of strange attractors", *Physica D: Nonlinear Phenomena*, 9 (2-1), 189–208 (1983).

[25] Ghorbani, M.A., Kisi, O. and Aalinezhad M. "A probe into the chaotic nature of daily streamflow time series by correlation dimension and largest lyapunov methods", *Applied Mathematical Modelling*, 34, 4050-4057 (2010).

[26] Schertzer, D., Tchiguirinskaia, I., Lovejoy, S., Hubert, P. and Bend-joudi, H. "Which chaos in the rainfall-runoff process? A discussion on "Evidence of chaos in the rainfall-runoff process" by Sivakumar et al., *Hydrological Sciences Journal*, 47, 139–147 (2002).

[27] Sivakumar, B. "Correlation dimension estimation of hydrologic series and data size requirement: myth and reality", *Hydrological Sciences Journal*, 50, 591–604 (2005).

[28] Sivakumar, B., Berndtsson, R., Olsson, J. and Jinno, K. "Reply to "Which chaos in the rainfall-runoff process?" by Schertzer et al. ", *Hydrological Sciences Journal*, 47, 149-158 (2002).

[29] Sivakumar, B., Persson, M., Berndtsson, R. and Uvo, C.B. "Is correlation dimension a reliable indicator of low-dimensional chaos in short hydrological time series? ", *Water Resources Research*, 38 (2), doi:10.1029/2001WR000333 (2002).

[30] Tongal, H. and Berndtsson, R. "Impact of complexity on daily and multi-step forecasting of streamflow with chaotic, stochastic, and black-box models", *Stochastic Environmental Research and Risk Assessment*, doi: 10.1007/s00477-016-1236-4 (2016).

Figures caption

Figure 1. General Categorize of catchment classification study

Figure 2. Scattering Map of Study Stations and classification map

Figure 3. Phase space diagram for three selected stations
a) 21-163, b) 21-966, c) 19-055
Figure 4. logC(r) versus log(r) plots for daily river flow data for three selected stations
a) 21-163, b) 21-966, c) 19-055
Figure 5. Relation between correlation dimension and embedding dimension for three selected stations
a) 21-163, b) 21-966, c) 19-055

Figure 6. Schematic representation of proposed classification pattern

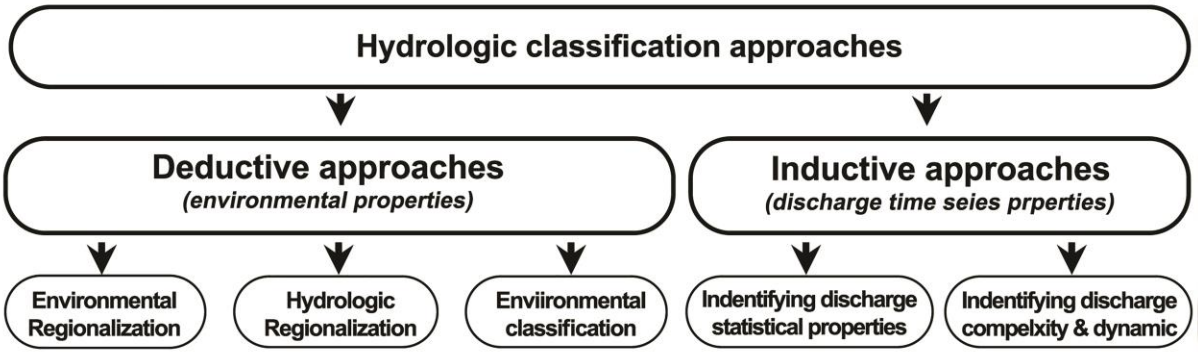
Hydrologic classification approaches
Deductive approaches
(environmental properties)
Inductive approaches
(discharge time seies prperties)
Environmental Regionalization
Hydrologic Regionalization
Enviironmental classification
Indentifying discharge statistical properties
Indentifying discharge compelxity & dynamic

Figure 1.

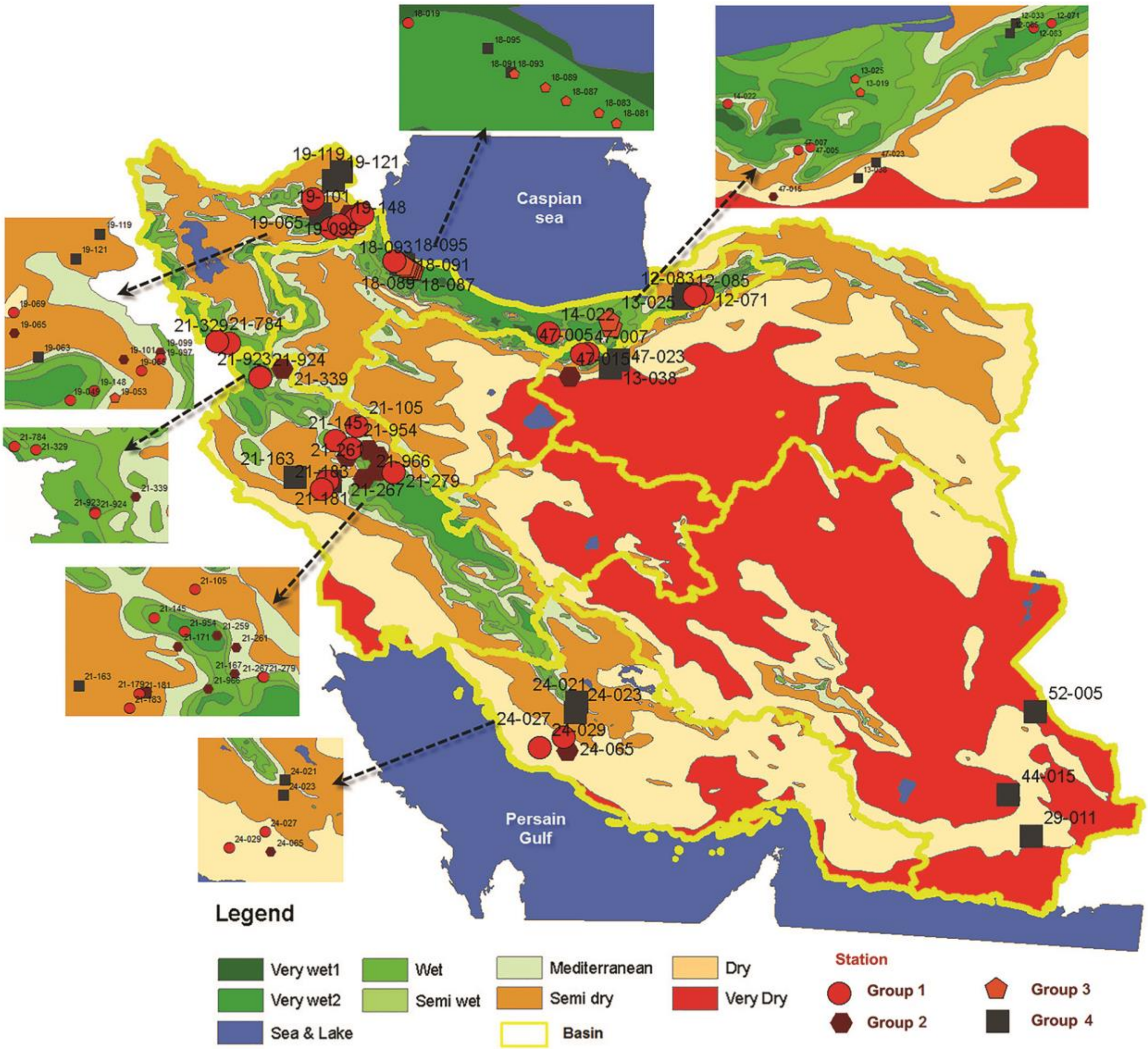
Caspian sea
Persain Gulf
Legend
Very wet1
Very wet2
Sea & Lake
Wet
Semi wet
Mediterranean
Semi dry
Basin
Dry
Very Dry
Station
Group 1
Group 2
Group 3
Group 4
19-119
19-121
19-148
19-065
18-095
18-091
18-087
12-085
12-071
14-022
47-007
47-023
13-038
21-784
21-924
21-339
21-105
21-954
21-163
21-966
21-279
24-021
24-023
24-027
24-029
24-065
52-005
44-015
29-011

Figure 2.

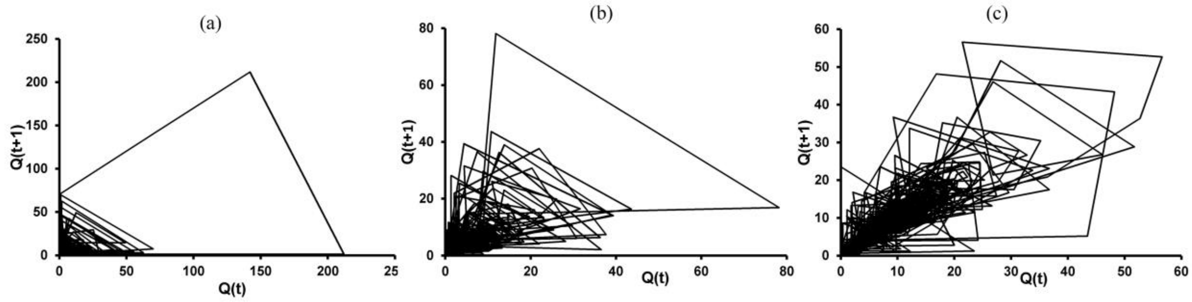

Figure 3.

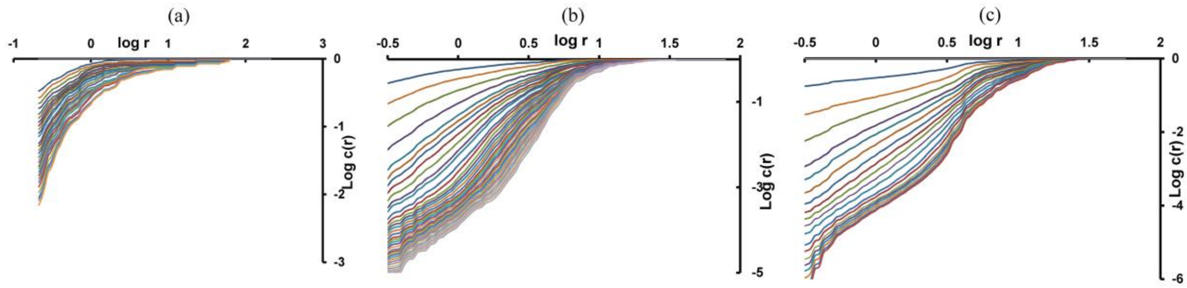

Figure 4.

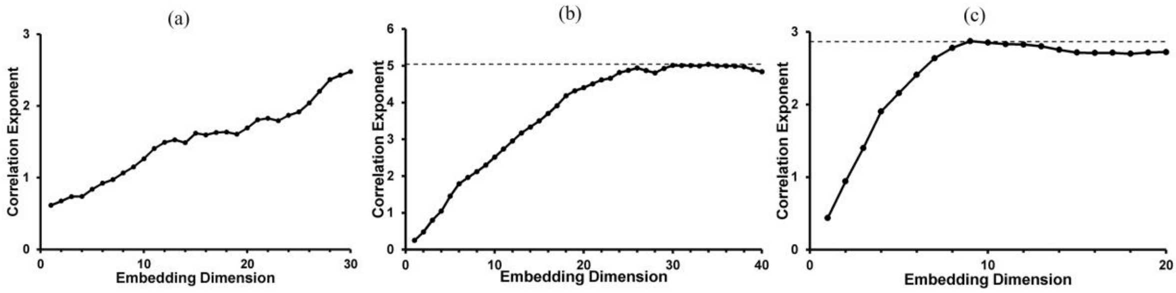

Figure 5.

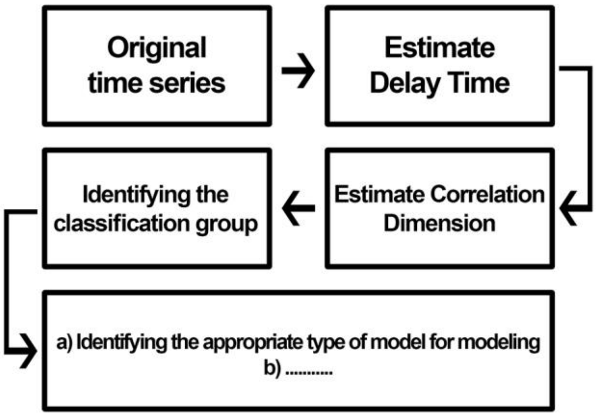
Original time series
Estimate Delay Time
Identifying the classification group
Estimate Correlation Dimension
a) Identifying the appropriate type of model for modeling
b) ....

Figure 6.

Table caption

Table 1. Details and results of stations analysis

Table 1.

Station	Catchment Properties		Outflow Properties		D_2	Station	Catchment Properties		Outflow Properties		D_2
	Elevation (m)	Area (km^2)	Average (m^3/s)	Max (m^3/s)			Elevation (m)	Area (km^2)	Average (m^3/s)	Max (m^3/s)	
12-033	100	114.5	0.39	67	∞	21-339	1575	280	1.53	39	5
12-071	300	335	2.24	78.1	6	21-784	955	162	1.08	29.8	3
12-083	280	387.5	1.20	52	6	21-923	1310	996	9.66	438.82	4
12-085	465	195.4	0.54	51.05	∞	21-924	1311	379.5	1.90	69.6	4
13-038	1800	248	0.62	40	∞	21-105	1800	901	4.21	37.97	3
47-015	1040	3209	7.75	320	5	47-005	1924	1910	1.13	72	6
47-023	1280	360	0.22	22	∞	47-007	1814	587	4.36	34.2	4
19-049	2068	36	0.11	5.27	6	18-019	135	126	2.07	75	6
19-053	1440	1070.6	3.61	142.48	7	18-081	-20	780.5	21.33	415.1	11
19-055	1332	1638	2.60	56.6	3	18-083	-20	441.8	10.68	266	7
19-063	1420	98.1	0.78	116	∞	18-087	-22	355	5.57	125	7
19-065	780	7311	8.09	185.2	5	18-089	-23	206.2	6.89	167	8
19-069	705	11267	15.19	232	4	18-091	-19	341.1	4.55	66	8
19-097	1352	44	0.13	5.11	6	18-093	-19	318.3	3.59	360	∞
19-099	1459	40	0.08	2.11	5	18-095	-15	100.3	1.93	103	∞
19-101	1290	4003.7	6.28	351.4	5	21-145	1780	615	3.37	99.5	6
19-119	334	710	0.16	66.3	∞	21-163	880	568	1.45	212	∞
19-121	820	156	0.16	18.6	∞	21-167	1770	270	2.71	104.61	5
19-148	1800	8	0.07	0.949	4	21-171	1520	773	7.70	125.37	5
13-019	400	1256	6.26	108.91	8	21-177	820	6700	41.86	923	5
13-025	270	2715	10.12	321	8	21-179	800	800	3.50	472.25	∞
14-022	570	524	2.31	46.7	4	21-181	790	1108	1.44	142	6
24-021	1585	415	1.82	264	∞	21-183	650	9140	48.33	945	6
24-023	1376	1410	3.75	665	∞	21-259	2000	60.4	2.27	37.12	5
24-027	384	4300	10.13	671	6	21-261	1490	1000	5.72	119	5
24-065	222	18525	19.64	1825.27	5	21-267	1450	3400	15.85	445	5
29-011	778	2420	2.57	1922	∞	21-279	1450	2655	9.36	427.5	4
44-015	651	3670	2.39	1019	∞	21-954	1720	166	2.68	34.6	4
52-005	1158	1350	0.24	60	∞	21-966	1420	234	1.35	78.16	5
21-329	1493	111	1.58	38.83	3	24-029	70	36500	38.94	2866	6

∞ = unidentifiable

Biographies

Hadi Delafrouz is PhD candidate at Iran University of science and technology, school of civil engineering. He obtained his BS degree from Shahrood University in 2000, and MS degrees degree from Tabriz University in 2007. His research area is study on hydrological, hydraulic modelling and catchment management. At present, he is actively involved in use complexity concept in catchment classification. He has published 14 national and international journals, conference proceedings.

Abbas Ghaheri is Professor at Iran University of science and technology, school of civil engineering, where he has been academic Experiences 1976 up to now. He obtained his BS degree from Tehran University, Iran, and MS degrees and PhD degree from Colorado State University, USA. His research interests include Design of Run off Collection Network, River Mechanics and Flood Plain Analysis and Management, Water Resources Development, Planning, and Management, Ground Water Analysis, Planning for Ground Water Supply. The results of his research have been presented in over 80 publications in international journals, book chapters, conference proceedings, and research reports. He has also advised and supervised numerous projects.

Mohammad Ali Ghorbani is Associate Professor of Hydrology at the University of Tabriz, where he firstly joined in 2006. His main research focuses on Hydrology, Water Resources and Artificial Intelligence, etc. He has more than 10 years research experience and has guided PhDs and MScs in the field of Hydrology. He received his B.Sc. in Irrigation from University of Tabriz in 1997, and M.Sc. in hydraulic structures from University of Tehran in 2000. He received his Ph.D. in Hydrology from University of Tehran in 2006. He has published 54 international and 60 national peer-reviewed papers, as well as one chapter in international book, two books in Persian.