Al-Rafidain Engineering Journal (AREJ) Vol.29, No.1, March 2024, pp. 132-143

The Disaggregation Model via Non-Parametric Approach

Aaisha Mohand Yahya
aaisha.almktar78@ gmail .com

Kamel A. AL-Mohseen
k.almohseen@uomosul.edu.iq

Shatha H.D.AL-Zakar
s.alzakar@uomosul.edu.iq

Department of Dams and Water Resources Engineering, College of Engineering, University of Mosul, Mosul, Iraq

Received: November 17th, 2023 Received in revised form: December 29th, 2023 Accepted: January 16th, 2024

ABSTRACT

Flow disaggregation models, which are one of the stochastic generation techniques, play a crucial role in the planning, design, and operation of water resource management systems and related projects. One distinguishing feature of these models is their ability to address the issue of missing observed data and compensate for it. They also enable the rescaling of data from a higher temporal level to a lower temporal scale. Data at lower temporal scales are typically required to address hydraulic and operational design problems in water resource projects. There are two main approaches to disaggregation flow data: the parametric approach and the non-parametric approach.One of the advantages of the disaggregation model is its ability to distribute flow data values from a key station to several sub-stations, both temporally and/or spatially, while preserving the basic statistical properties of the time series obtained from the model (mean, standard deviation, minimum, maximum, and correlation coefficient) for the observed data.In the current study, a non-parametric approach was used for the purpose of disaggregation approach. It is assumed that there is aggregated discharge data at a key station, and this data will be disaggregated into a corresponding series of discharges temporally and spatially at sub-stations that are statistically similar, using the SAMS 2010 platform program (Stochastic Analysis, Modeling, and Simulation). Annual and monthly discharge data for five stations measuring discharges on the Tigris River System in Iraq were used, including the Mosul Dam station on the Tigris River, the Asmawah station on the Khazir River, the Askiklik station on the Upper Zab, the Dibs Dam station on the Lower Zab, and the Baiji station on the Tigris River, covering a time span of twenty-three years. The statistical results of the disaggregation approach were compared with their observed counterparts and showed good agreement in most years and months and for all stations. Based on this, the method is recommended disaggregation of the data when decisions required water management strategies in these regions.

Keywords:

River discharge disaggregation, non-parametric approach, spatial methods, temporal methods, Tigris River.

https://rengj.mosuljournals.com
Email: alrafidain_engjournal2@uomosul.edu.iq

==

1. INTRODUCTION

The use of non-parametric models is essential to overcome limitations in the parametric methods employed in random stochastic models, such as reducing the required number of parameters and avoiding assumptions about data distribution [1]. These techniques have been applied to a variety of hydrological discharge data, including streamflow modeling [2, 3], flood estimation [4, 5], and river flow forecasting [6, 7]. Non-parametric methods provide an alternative to traditional linear parametric approaches, where a single linear model is assumed to fit the entire dataset [8].

The accuracy of non-parametric methods depends on the sample size, and estimating the probability density function (pdf) requires larger sample sizes to achieve similar accuracy for shorter time periods [3]. To mitigate the limitations of methods relying on pdf, a non-parametric approach called K-Nearest Neighbor (KNN) was developed to regenerate monthly discharges based on a weighted function, giving higher weights to closer values and lower weights to more distant values.

Researchers [9] combined KNN techniques with an idealized scheme for modeling spatial-temporal flow disaggregation, splitting monthly discharges from a main site into daily discharges at multiple basin locations. The results indicated that the continuity of flow across months at each site, in addition to the inter-site flow, yielded compelling results. A semi-distributed approach was employed to predict river flow at multiple locations on the Serawa River in Brazil, based on climatic factors. Seasonal and annual discharges were predicted and disaggregated into monthly discharges depending on the nearest value (k), which maintains spatial and temporal consistency across sites with high correlations (0.9) between the mean values of observed and generated annual discharges [10].

A model was developed by [11] to disaggregate annual discharge data into monthly data at multiple locations, analyzing temporal and spatial variations in statistics, including variances, standard deviations, droughts, and surplus-related statistics. Comparisons revealed that the SAMS program outperformed other methods in preserving both monthly and annual statistics.

Comparative analysis of time series generated using the KNN non-parametric approach and the Index Sequencing Method (ISM) for annual discharges in the Gunnison River showed that KNN produced a diverse set of time series, whereas ISM was limited in diversity [12].

Researchers [13] integrated two non-parametric approaches to provide seasonal flow predictions at multiple sites while maintaining aggregated data simultaneously. The first approach involved multi-model non-parametric ensemble prediction techniques [14], while the second was a non-parametric temporal disaggregation technique [15]. The results showed that the model maintained temporal and spatial correlations.

The researchers [16] reviewed the characteristics of non-parametric disaggregation methods and attempted to find a better disaggregation model to overcome the weaknesses of previous models. They proposed using the KNN model in conjunction with genetic algorithms and compared it with non-parametric disaggregation approaches on Colorado River data. The results showed that the proposed model could restore correlations between the last month of one year and the first month of the next year, producing discharges that were not observed in the data. Statistical capacity estimates for monthly data generated by different models were used for a temporal and spatial comparison.

Researchers [17] presented a study on the application of the KNN model for the nonparametric disaggregation method by dividing annual expenditures into monthly data for three measuring stations located along the Kizilirmak River Basin in Turkey, which was exposed to drought for a number of years. Both temporal and spatial approaches were used for the purpose of determining the potential of the model in conservation. On the statistical characteristics of the observed data for drought and storage analysis at these stations, the results showed that the spatial classification has the ability to reproduce the observed data better than the temporal approach for the selected sites and provides a variety of monthly sequential flows generated, which can then be used to analyze the performance of the water resource planning system.

The researchers [18] conducted a study that included the use of a random block length model (RB-NPD) combined with genetic algorithms for the purpose of overcoming a major weakness in the parametric approach, which is the repetition of similar values in disaggregated The model was compared with the parametric approach models for both VS and MR and the researchers' nonparametric models by [16].The results showed that the previous parametric and nonparametric models had some shortcomings in obtaining matching of all statistical measures. In contrast, the proposed nonparametric random model based on RB-NPD and mixed with the genetic algorithm can be comparable and is considered suitable for segmenting annual data in this area.

The researchers [19] used three different methods to estimate daily discharges in various rivers in Spain. The first method involved simulating daily discharges using the SWAT model, which was calibrated at the monthly scale. The second method involved obtaining daily discharges from monthly data using a disaggregation model, while the third method simulated daily data using the SWAT model, which was calibrated at the daily scale. Subsequently, model performance was assessed using several criteria, including Kling Gupta, KGE, Percent Bias Absolute Index (PBAIS), Nash-Sutcliffe Efficiency (NSE), Root Mean Square Error (RMSE), and Relative Standard Error (RSR). The results showed that the second model (disaggregation) outperformed the other models in all the statistics used. It was also noted that the disaggregation model exhibited better performance in terms of flow values close to the original values.

Within the framework of the Method of Fragments (MOF), Researchers [20] used a technique known as "Pattern Mapping" (PM) to disaggregate daily precipitation. Several statistics-based criteria are incorporated by PM into the fragment selection process to increase its accuracy. Three widely-known disaggregation approaches are utilized to assess the effectiveness of PM-MOF using reanalysis data from 14 global sites spanning different climatic zones: K-Nearest Neighbour-Based MOF (KNN-MOF), Bartlett-Lewis Rectangular Pulse Model, and Micro-Canonical Cascade (MCC). The PM-MOF strategy is deemed superior to the other models as it replicates the conventional statistics with a substantially lower percentage error (−50.2 to 50.3 %), according to their conclusion.

The researchers [21] employed the MuDRain model to investigate rainfall fragmentation at many stations in Hormozgan Province, as well as the impact of hourly data correlation between stations. By comparing the generated and observed hourly time series, they demonstrate that while the model correctly predicted daily precipitation levels, it typically simulated maximum precipitation values that were lower than the actual amounts. Furthermore, insufficiently intense precipitation fell. They demonstrate that a feasible way to generate a realistic hourly rainfall series at the target station is to make use of the hourly rainfall data that is currently available at nearby stations.

Based on a review of previous studies, it was found that various algorithms, techniques, and methodological approaches are available for generating flow data at different time scales. Different methods have been proposed for different time periods, with some modifications and improvements. Disaggregation models have been developed to disaggregate monthly discharge data into daily data, but each method has its pros and cons in terms of accuracy and limitations. Results from disaggregation techniques based on non-parametric methods have shown significant improvements over parametric models. However, some non-Parametric methods still have common shortcomings compared to parametric models, such as generating negative discharges in some cases and the challenge of capturing flow continuity between monthly and annual transitions.

The main goal of the current study is to disaggregate annual discharge data into monthly data at several stations along the Tigris River in Iraq using various non-parametric models. The study aims to analyze how successful these models are in preserving the statistical properties of the observed station data. This study is motivated by the absence of non-parametric approaches for disaggregating flow data in both temporal and spatial dimensions in Iraq, particularly using the SAMS 2010 program. The potential applications of such disaggregated data are crucial for predicting future discharges of the Tigris River, emphasizing its significance.

2. STUDY AREA

The Tigris River in Iraq was selected for this study, specifically the section between the Mosul Dam in the north and the Makhool Dam in the south. The Makhool Dam site represents the confluence point of discharges from other stations (subordinate stations). Four measurement stations were chosen along the Tigris River north of Baiji to measure the inflow into the dam reservoir. The study assumed that the discharge measurement station on the Tigris River in Baiji represents the sum of the inflows to the Makhool Dam reservoir, given the proximity between the two locations and the lack of data at the actual dam site [22]. The selected stations are: the Mosul Dam station located north of the city of Mosul; two stations on the Upper Zab tributary, Eski kalak station and Asmawah station located on the Khazir River; the Al-Dibs Dam station located on the Lower Zab tributary; and the main discharge collection station, the Baiji station located near the dam site. Figure 1 shows the selected measurement station locations for the study.

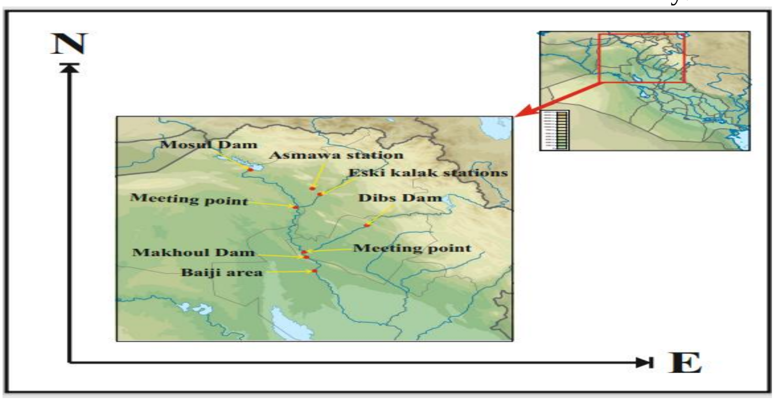

Fig. 1 Selected Measurement Station Locations for Discharge Measurement.

3. DATA USED IN THE STUDY:

The study relied on discharge records for the Tigris River basin at measurement stations (Mosul Dam, Al dibs, and Baiji), obtained from the National Center for Water Resources in Iraq, and two stations (Asmawah and Eski Kalak) from the Directorate of Water Resources and Irrigation in the Kurdistan region. Table (1) provides information about station locations in terms of longitude, latitude, and the historical observation period for each station.

Table 1: Longitude, Latitude, and Observation Duration of Stations

Station Name	Latitude	Longitude	Historical duration
Mosul Dam	42° 49' 23"	36° 37' 48"	2000-2022
Asmawa	43° 31' 49"	36° 31' 28"	2000-2022
Eski Kalak and	43° 34' 33"	36° 10' 43"	2000-2022
Dibs Dam	44° 06' 38"	35° 41' 21"	2000-2022
Baiji	43° 29' 35"	34° 55' 45"	2000-2022

4. RESEARCH METHODOLOGY:

Non-parametric methods are used in data disaggregation models for generating future river discharge data. These methods involve using data without making any assumptions about it (e.g., assuming linearity or following a probability density function (PDF)). There are two non-parametric approaches to data disaggregation: temporal and spatial disaggregation.

The non-parametric temporal disaggregation method combines the Non-Parametric Disaggregation (NPD) model, created by researchers [15], and the Accurate Adjusted Procedure (AAP) model proposed by [23]. The model starts by generating annual data for the aggregated variable X at the key station. Then, the K-Nearest Neighbor Resampling (KNNR) model is applied to generate seasonal data, ensuring that their sum is close to the total value of the aggregated variable X. The detailed steps for temporal disaggregation (from annual data to monthly data) are as follows:

1. Convert the observed monthly flow data for any given year into a proportion of the total annual flow to obtain Pyt. Repeat the same procedure for all months in the years to obtain the proportion matrix, Py, with dimensions (n × 12), where n represents the number of years in the series and the number 12 represents the number of months.

2. Apply a suitable model to historical data xi using one of the following models: Index Sequential Method (ISM), K-Nearest Neighbor Resampling (KNNR), KNN with Gamma Kernel Density Estimation (KDE) (KGK). Block Bootstrapping (BB). Then generate annual data ($X\upsilon$), υ = 1, …, NG, where NG represents the length of the generated series.

3. Impose the first generated value X1, and then calculate the distance Δi between X1 and the observed annual data xi, i = 1, …, N, N (representing the length of the observed time series), as follows:

$$\Delta i = \left| X1 - xi \right| \quad , \quad i=1,2,\ldots.N \qquad \ldots\ldots\ldots (1)$$

Then the distances are arranged from smallest to largest.

4. Finding the number of nearby values, k, where (k = N1/2) corresponding to weights w1, w2, …, wk.

$$W(i) = \frac{\left(1/_i\right)}{\left(\sum_{i=1}^{k} 1/_j\right)} \quad , \quad i = 1,2,\ldots., K \quad \ldots.. (2)$$

Then the smallest value of k for Δi is selected using the sum of the weights, . cw_l, l= 1, …, k,. Afterward, the observed values for a given year, corresponding to the disaggregated values, are considered candidates for the disaggregation process, resulting in (Y1 = Y1,1, Y1,2, …, Y1,d). The sum of these values should equal X1.

5. The next year, $X\upsilon$ (e.g., υ = 2), generated in the second step, is chosen for the purpose of generating monthly data for it.

6. Step five is repeated until obtaining the monthly data for the NG time series.

The steps for conducting spatial disaggregation are similar to temporal disaggregation, with a minor difference. The proportion matrix, Py, becomes (n × 12 × S), where S represents the number of sites to be disaggregated. Then, the matrix is formed as in temporal disaggregation, and the nearby annual discharges, matching the value to be disaggregated, are selected. The process is repeated to generate all discharges in the time series.

To apply the disaggregation methods, observed monthly discharge data for selected stations was used, which was obtained from the National Water Resources Center in Iraq/Baghdad and the General Water Resources Directorate in the Kurdistan Region/Erbil.

A software program called SAMS is used for modeling, simulation, and stochastic analysis of hydrologic time series. SAMS 2010 is the most recent version of SAMS. Three areas can be used to group SAMS capabilities: historical data analysis, fit modeling and parameter estimates, and generated synthetic time series. SAMS's data analysis functions include data transformation, data graphing, and data statistical properties. Model fitting is SAMS's second application. It covers model testing and parameter estimation for monthly and annual stochastic models. Data generation is SAMS's third primary use. The above-mentioned fitted models serve as the foundation for data creation [24].

5. RESULTS AND DISCUSSION:

The basic monthly statistical moments (mean, standard deviation, skewness, minimum, maximum, and correlation coefficient) were computed for the generated data resulting from the disaggregation process using the three models (KGK, BB, and ISM) for all stations and for data generated from 100 time series and a 23-year period (the same length as the observed data) using both temporal and spatial approaches. These statistics were compared with the statistics of the observed data at the same stations and for the same time period. The statistics were presented in the form of box plots representing the generated data, with the horizontal line inside the box representing the median of these generated data. The polygon represents the statistics for the historical data in temporal disaggregation, while it is represented as a triangle in spatial disaggregation. A good statistical performance is achieved when the statistical value of the observed data falls within the box range.

5.1 Temporal Disaggregation:

The temporal disaggregation process from annual data to monthly data using the three models (KGK, ISM, and BB) provided a lot of figures for the stations used in the study. Therefore, the focus was on presenting the shapes of a single station, the main station (Baiji), for all models and illustrating the differences between them. Figure (2) represents the statistical moments of the disaggregated data with their counterparts from the data observed using the KGK model. Meanwhile, Figure (3) represents the statistical results using the BB model, and Figure (4) represents the results using the ISM model.

Through these figures, a significant match is observed in the mean values in all models, as well as the ability of the models to generate data with a mean close to the mean of the observed data. This is evident through the tightening in the box, indicating a high match for the generated data in most months except April and May. Regarding the standard deviation, the results showed similarity between the generated values and the observed data, indicating the models' ability to accurately reproduce the observed data. Moreover, there was a significant match between the KGK and BB models, while the ISM model was superior, especially in April and May. The skewness values for the observed data were within the box in all months except April and May in the BB model, where they were higher than the generated values. The minimum values for the observed data did not match in some months, contrary to the maximum values, which matched in most months and for all models.

As for the correlation coefficient values for the observed data, they fell within the box range in all months and for all models. These results were similar to the results obtained for the other stations, as shown in Table (2), which represents the statistical values for some months (October, November, and December) for all stations. To assess the performance of the three models used in the study (KGK, ISM, and BB) and determine the best model, statistical criteria were used, including the coefficient of determination (R2) and the Nash-Sutcliffe Efficiency (NSE) [25, 26, 27]. Table (3) shows the values of these criteria.

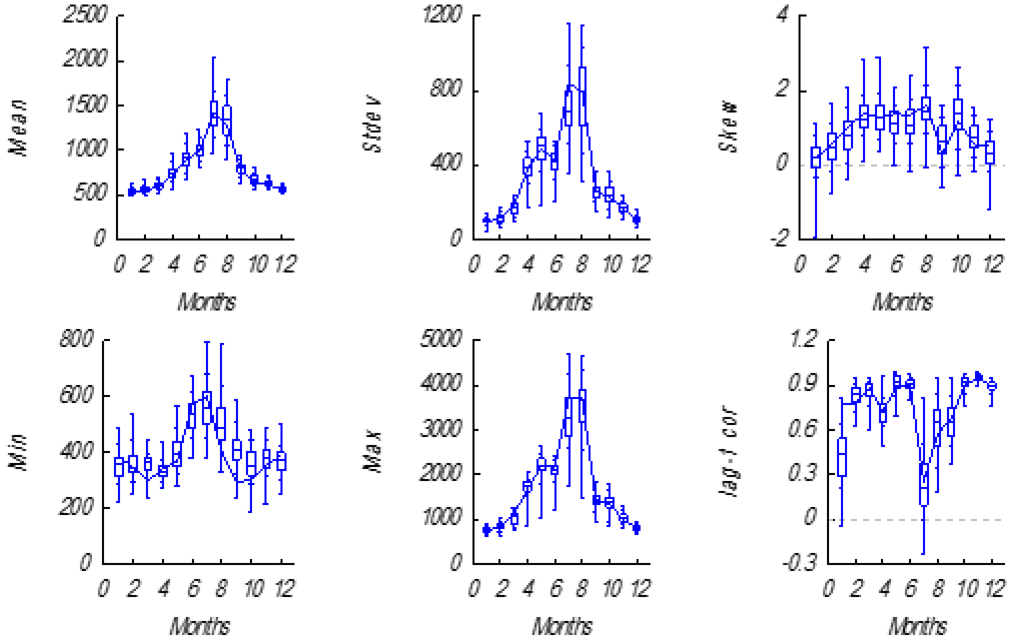

Fig. 2 Statistics of observed and simulated streamflow using (KGK) model (for Baiji Station). The polygon represents the observed values

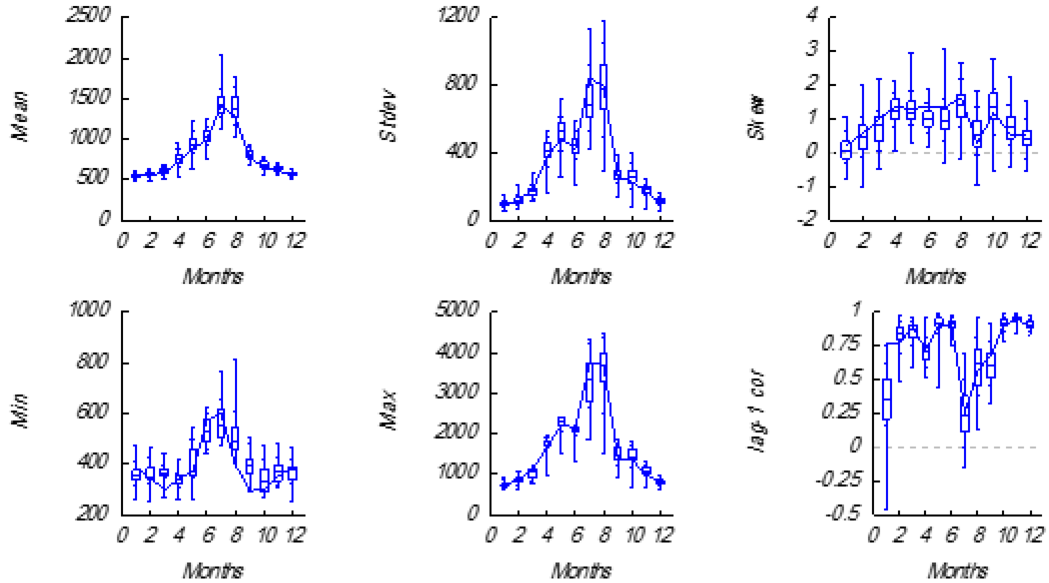

Fig. 3 Statistics of observed and simulated streamflow using (ISM) model (for Baiji Station). The polygon represents the observed values

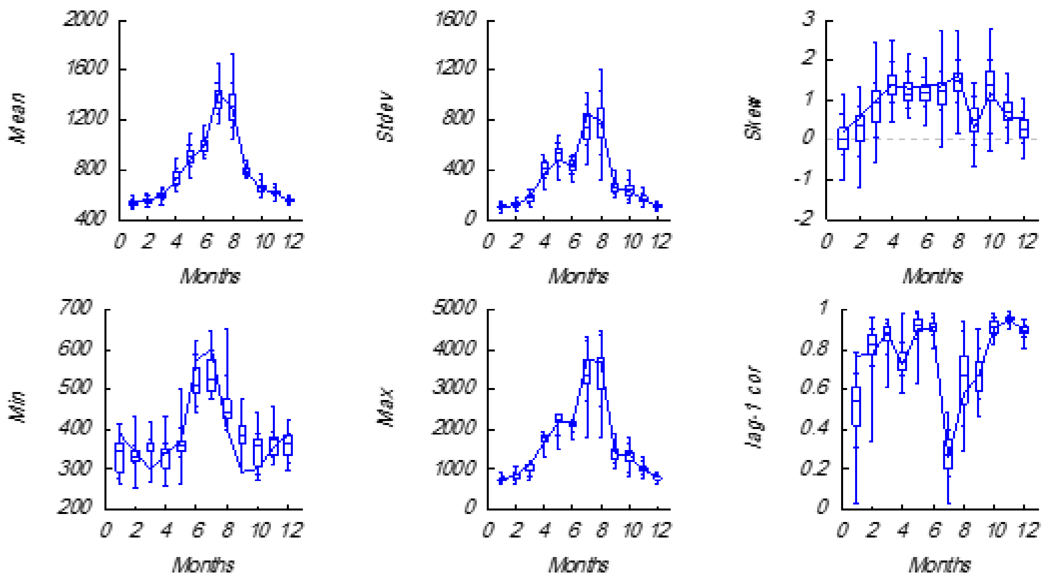

Fig. 4 Statistics of observed and simulated streamflow using (BB) model (for Baiji Station). The polygon represents the observed values

Table 2: Comparison of Observed and Model-Generated Values for Three Months in Substations.

Statis.	October				November				December			
	Obs. Data	Disaggregate Data			Obs. Data	Disaggregate Data			Obs. Data	Disaggregate Data		
		KGK	ISM	KGK		KGK	ISM	KGK		KGK	ISM	BB
Mosul Dam Station												
Mean	422.1	429.4	427.9	443	453.3	454.4	450.8	463.8	462.0	470	464.7	469.4
St.Dev	100.0	134.7	135	144.5	105.7	113.1	116.7	119.9	83.71	88.82	89.02	93.12
Skew	0.230	0.908	0.902	0.979	0.411	0.569	0.445	0.543	-0.04	0.037	0.000	0.046
Min	269.4	250.1	242	254.9	276.0	284.5	274.7	283.6	315.4	310	302	306.2
Max	604.9	783.9	785.3	829.1	649.8	697.5	688.1	721.7	598.8	635.7	625.6	639.1
Asmawa Station												
Mean	5.785	5.906	5.892	6.073	5.336	5.427	5.439	5.519	5.459	5.59	5.476	5.588
St.Dev	1.633	1.957	1.989	2.068	1.751	1.856	1.817	1.91	1.609	1.653	1.609	1.596
Skew	-0.04	0.685	0.189	0.842	0.189	0.350	0.256	0.330	0.535	0.427	0.448	0.334
Min	2.8	2.857	2.18	2.894	2.18	2.363	2.327	2.322	2.750	2.985	2.986	3.001
Max	8.37	10.79	8.6	11.49	8.6	9.274	9.117	9.485	9.00	8.976	8.804	8.776
Eski Kalak Station												
Mean	204.2	214.9	211.8	218.9	138.0	142	140.8	142.3	119.0	119	118.1	120.8
St.Dev	59.18	73.93	75.4	78.34	39.45	40.54	39.46	40.16	36.71	36.99	37.03	36.42
Skew	0.081	0.747	0.773	0.969	0.660	0.599	0.572	0.453	0.947	0.800	0.773	0.693
Min	89.20	97.29	90.15	99.33	69.61	73.16	74.02	72.9	63.07	64.06	61.05	63.26
Max	316.6	406.8	410.1	431.9	242.9	235.6	232.2	228.9	219.8	210.4	211.9	209.4
Dibs Dam Station												
Mean	105.3	114.9	111.6	114.9	117.5	125.7	123	126.1	78.38	80.31	76.25	79.77
St.Dev	71.32	69.73	71.82	70	69.61	70.82	70.46	70.78	55.36	53.27	53.22	54.4
Skew	0.949	0.903	0.880	0.724	0.186	0.2272	0.259	0.26	0.951	0.723	0.897	0.864
Min	11.60	20.42	18.36	19.31	18.0	22.27	21.32	22.4	18.97	18.89	19.05	19.3
Max	309.6	291.4	294.5	284.2	230.0	254.1	251.8	257.8	210.0	198.3	200.3	203.5

Table 3: Values of Statistical Criteria R2 and NS for the Three Models.

Station name	KGK model		ISM model		BB model	
	R2	NS	R2	NS	R2	NS
Mosul Dam	0.9938	0.9685	0.9945	0.9920	0.9839	0.90105
Asmawa	0.9983	0.9999	0.9974	0.9999	0.997	0.99999
Eski Kalak	0.9996	0.99635	0.9993	0.9991	0.9996	0.9909
Dibs Dam	0.9882	0.9998	0.9898	0.9999	0.9843	0.9997
Baiji	0.9969	0.9977	0.9955	0.9981	0.9953	0.9908

5.2 Spatial Disaggregation:

Spatial disaggregation was conducted from the annual data at the main station (Baiji Station) to the annual data at sub stations using the same models as in temporal disaggregation (KGK, ISM, and BB). Figure (5) represents the statistical moments of the data disaggregated with their counterparts from the observed data using the KGK model. It can be observed that the model can maintain the statistical characteristics of the observed data, where the values of the observed characteristics (indicated by the triangle symbol) fall within the box range, except for the correlation coefficient, where the generated data values were lower than the observed values. Figure (6) represents the statistical results using the BB model, which showed convergence in the results compared to the previous model, except for the correlation coefficient values, which were higher than the observed value. Figure (7) represents the results of the ISM model, which showed differences in the skewness values for the generated data being lower than the observed values and also generated lower minimum values than the observed minimum.

Table (4) illustrates the results of the statistical characteristics values for the rest of the stations, where the values showed a significant convergence between the generated and observed values in the three models, except for the correlation coefficient values, which exhibited some differences from historical values. Moreover, the KGK and BB models produced similar data.

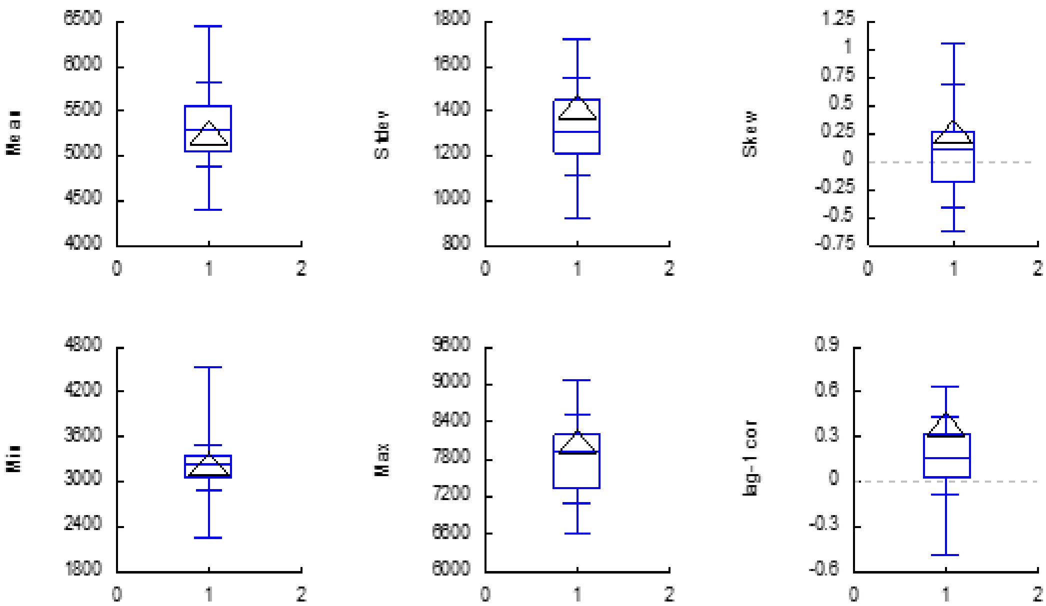

Fig. 5 Statistics of observed and simulated streamflow using (KGK) model (for Mosul Dam Station). The triangle represents the observed values.

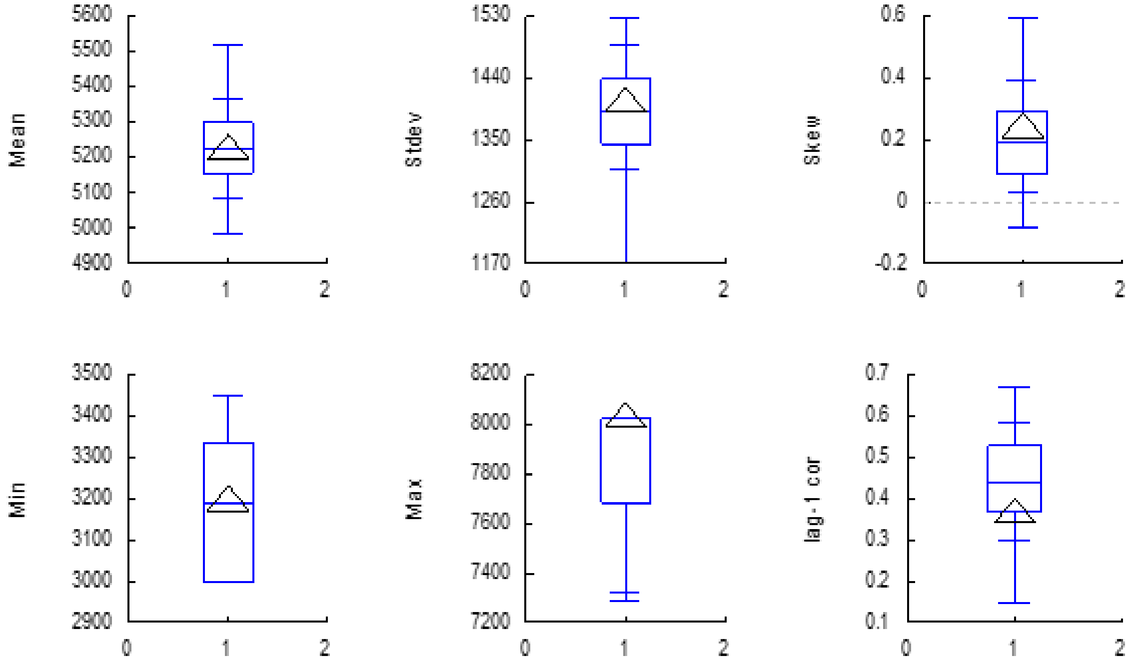

Fig. 6 Statistics of observed and simulated streamflow using (ISM) model (for Mosul Dam Station). The triangle represents the observed values.

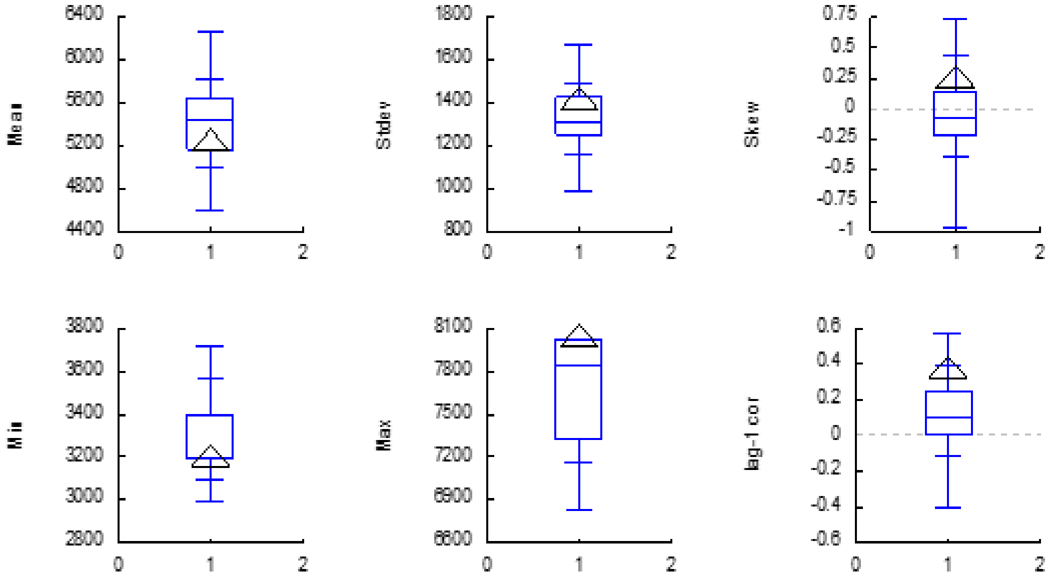

Fig. 7 Statistics of observed and simulated streamflow using (BB) model (for Mosul Dam Station). The triangle represents the observed values.

Table 4: Comparison of Observed and Model-Generated Values for Three Months in Substations

Statistics	Baigi Station				Asmawah Station			
	Observed	Disaggregation			Observed	Disaggregation		
		KGK model	ISM model	B.B model		KGK model	ISM model	B.B model
Mean	9595.0	9865	9595	10006	146.5	149.2	145.5	149.4
St.Dev	2832	2616	2832	2663	40.50	41.24	40.05	40.63
Skew	0.1001	0.0206	0.1001	-0.098	0.7601	0.6027	0.7426	0.6575
Min.	5110.0	5390	5110	5438	90.27	87.91	86.68	88.51
Max.	14390	14540	14390	14240	253.5	243.1	243.4	243.3
acf (1)	0.3973	0.2039	0.4597	0.1239	-0.1493	-0.0306	0.0071	-0.0746
Statistics	Eski Kalak Station				Dibs Dam Station			
	Observed	Disaggregation			Observed	Disaggregation		
		KGK model	ISM model	B.B model		KGK model	ISM model	B.B model
Mean	3360.0	3421	3315	3452.0	860.4	879.2	812.9	894.6
St.Dev	1089	1024	1056	1021	652.0	616	619.5	649
Skew	0.4846	0.3842	0.4837	0.3666	1.463	1.16	1.446	1.348
Min.	1493.0	1646	1540	1707	203.5	212.6	203.8	219.1
Max.	5901.0	5564	5603	5508	2826.0	2456	2543	2614
acf (1)	0.3638	0.1029	0.3135	0.0550	0.4778	0.0707	0.2467	0.0474

6. CONCLUSIONS:

In this study, the efficiency of using the non-parametric model for disaggregating annual discharge data to generate monthly time series was investigated. The model was applied to five stations located along the Tigris River in Iraq. The annual data at the key station (Baiji Station) was disaggregated into monthly data at the same station using a temporal approach. Additionally, spatial disaggregation was performed on substations using the same models (KGK, ISM, and BB). The model relied on three methods to regenerate monthly data. The results of various statistical properties of both observed and generated data demonstrated the performance of each approach in the disaggregation model.

The results show the model's capability to generate data with similar monthly statistical values both spatially and temporally, particularly regarding the first moments such as mean, standard deviation, and skewness across all three methods. Some differences were observed in the months of April and May due to increased discharges caused by rainfall and snowmelt. The main advantage of the model lies in its ability to obtain values at different time scales and its flexibility in acquiring values for several stations from a single station. However, one of the drawbacks of both spatial and temporal approaches is their inability to maintain continuity between the last month of one year and the first month of the following year. This limitation is common in many non-parametric disaggregation models but can be mitigated as long as the model can maintain other key statistical properties effectively.

The results indicated that the non-parametric disaggregation model is efficient and provides a good option for generating river flow time series. This knowledge can be instrumental in operating and managing the reservoir of the Makhool Dam in the future. It means that water resource experts will have a comprehensive understanding of how different tributaries contribute to the flow of the Tigris River into Makhool. Furthermore, if reservoir operation is integrated into a single system, it will be possible to control discharges from different components of the reservoir system to achieve the optimal operation policy for this reservoir.

In conclusion, the resulting recommendations include the following:

1. Apply the non-parametric models used in this study to other measuring stations located along the Tigris River in order to evaluate the effectiveness of these models in the field of management and operation.
2. It is suggested to use the SAMS-2010 program to study statistics related to storage, drought, and surplus for the selected stations in this study, to evaluate the performance of the disaggregation approach in capturing longer temporal properties and due to the importance of these characteristics in the management and operation of reservoirs.

REFERENCES:

[1] B. Rajagopalan, Salas, J. D.Salas, and U. Lall, "Stochastic methods for modeling precipitation and streamflow", Advances in data-based approaches for hydrologic modeling and forecasting , pp. 17–52, 2010.

[2] [2] J. R. Prairie, B. Rajagopalan, T. J. Fulp, E. A. Zagona, "Statistical nonparametric model for natural salt estimation", Journal of Environmental Engineering, Vol.131, Issue:1, pp.130–138, 2005. doi: 10.1061/(asce)0733-9372(2005)131(1)130-138

[3] U. Lall, and A. Sharma, " A nearest neighbor bootstrap for resampling hydrologic time series", Water Resources Research, Vol.32, Issue: 3, pp.679–693, 1996. doi: 10.1029/95WR02966.

[4] Y. I. Moon, and U. Lall, (1994) "A kernel quantite function estimator for flood frequency analysis", Water Resources Research, Vol.30, Issue :11, pp.3095-3103, 1994. doi: 10.1029/94WR01217.

[5] N. Şarlak, and Ş. Tiğrek, " Noktasal Taşkın Frekans Fanalizi: Göksu Nehri ve Kayraktepe Barajı Vaka Analizi", Journal of the Faculty of Engineering and Architecture of Gazi University, Vol.31, Issue: 4, pp.1095-1103, 2016. doi: 10.17341/gummfd.79436.

[6] K. Grantz, B. Rajagopalan, M. Clark, and E. Zagona " A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts" Water Resources Research ,Vol. 41,Issue:10, 2005. doi: 10.1029/WR003467.

[7] V.V. Srinivas, and K. Srinivasan, "Hybrid moving block bootstrap for stochastic simulation of multi-site multi-season streamflows" Journal of Hydrology, Vol.302,Issue: 1, pp.307–330,2005. doi: 10.1016/j.jhydrol.2006.01.023.

[8] U.Lall, "Recent advances in nonparametric function estimation: Hydrologic applications", Reviews of Geophysics, Vol.33, Issue:2, pp.1093–1102,1995. doi: 10.1029/95RG00343.

[9] D.N.Kumar, U. Lall, and M.R. Petersen,"Multisite disaggregation of monthly to daily streamflow", Water Resources Research,Vol.36 Issue:7, pp.1823–1833.2000. doi: 10.1029/2000WR900049.

[10] F.A.Filho, and U. Lall," Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a mutlivariate, semiparametric algorithm", Water Resource Research, Vol.39, Issue: 11, pp.1-11,2003. doi: 10.1029 /WR001373.

[11] T. Lee, J.D. Salas, J. Keedy, D. Frevert, and T. Fulp, "Stochastic Modeling and Simulation of the Colorado River Flows", World Environmental and Water Resources Congress, 10 p.2007. doi: 10.1061/40927(243)423.

[12] K.Nowak, J. Prairie, and B. Rajagopalan, "Development of stochastic flow sequences based on observed data", Washington, DC: Allen Institute.2008.

[13] C.Bracken, B. Rajagopalan, and J. Prairie, "A Multisite Seasonal Ensemble Streamflow Forecasting Technique", Journal of Water Resource Research, Vol.46, pp. 1-12,2009. doi: 10.1029/2009WR007965.

[14] S.K. Regonda, B. Rajagopalan, M. Clark, and E. Zagona, "A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin", Water Resources Research, Vol.42,Issue:9, 2006. doi: 10.1029/2005WR004653.

[15] J. Prairie, B. Rajagopalan, U. Lall, and T. Fulp, "A stochastic nonparametric technique for space-time disaggregation of streamflows",Water Resources Research, Vol.43, Issue:3,2007. doi: 10.1029/2005WR004721.

[16] T. Lee, J. D.Salas, J. Prairie, "An enhanced nonparametric streamflow disaggregation model with genetic algorithm", Water Resource Research, Vol.46, 14 p, 2010. doi: 10.1029/2009WR007761.

[17] S.H.AL-Zakar, N. Şarlak, O.M. Mahmood, "Disaggregation of Annual to Monthly Streamflow: A Case Study of Kızılırmak Basin (Turkey)", Advances in Meteorology, Vol.2017, 16 p, 2017. doi: 10.9790/4861-0901023443.

[18] T. Lee, and T. Ouarda, "Randomized block nonparametric temporal disaggregation of hydrological variables RB-NPD (version1.0) – model development" *Geoscientific Model Development Discussion,* 2023. doi.org/10.5194/gmd-2022-274,2023.

[19] G. Castellanos-Osorio , A. López-Ballesteros , J. Pérez- Sánchez , and J. Senent," Disaggregated monthly SWAT+ model versus daily SWAT+ model for estimating environmental flows in Peninsular Spain", Journal of Hydrology, Vol.623, August 2023. doi.org/ 10.1016 /j.jhydrol.2023.129837.

[20] M. Velpuri, G. Titas,and U. N.V. , " A Multi criteria Decision Making based nonparametric method of fragments to disaggregate daily precipitation", Journal of Hydrology, Vol. 617, Part A, 2023. doi: 10.24996/ijs.2023.64.6.22.

[21] H. Bolouki, and M. Fazeli, "Evaluation of Multivariate Rainfall Disaggregation Performance Using MuDRain Model (Case Study: North East of Hormozgan Province", A mirkabir Journal of Civil Engineering, Vol.54 (12), pp.4657–4676, 2023. doi: 10.1029/96WR00488.

[22] H. Khalid Hameed, K. Ahmed Abdullah ,and R.Hoobi Irzooki, "Seepage Simulation of the Proposed Makhool Dam in North Iraq", Iraqi Journal of Science, Vol. 64,Issue:6, pp.2934-2945,2023.

[23] D. Koutsoyiannis, and A. Manetas, "Simple disaggregation by accurate adjusting procedures", Water Resource Research, Vol.32, pp. 2105– 2117, 1996.

[24] O.G.B.Sveinsson , T.S. Lee, J. D. Salas, W. L. Lane, D. K. Frevert, and T.B.M.J. Ouarda," Stochastic Analysis, Modeling, and Simulation (SAMS) Version 2010 ".USER's MANUAL, Colorado State University, March 2011. doi: 10.1016/j.jhydrol.2011.08.027.

[25] R. Younus Ahmad Hassan, A. Mohammad Younes, "Prediction of Daily Flow to the Great Zab River Using Artificial Neural Network Models", Al-Rafidain Engineering Journal (AREJ) , Vol.28, Issue:.2, pp. 163-172,2023.

[26] S. Hazim Dawood, "Meteorological Estimations for selected stations in the North of Iraq", Al-Rafidain Engineering Journal (AREJ), Vol.17, Issue:1, 2009.

[27] A. Acharya, T.C. Piechota, H. Stephen, "Modeled streamflow response under cloud seedine in the North Platte River watershed", Journal of Hydrologic Engineering, Vol. 409, Issue:1-2, pp.305-314, 2011.

نموذج تجزئة للجريان باستخدام النهج اللامعلمي

عائشة مهند يحيى المختار
aaisha.almktar78@ gmail .com

كامل علي عبدالمحسن
k.almohseen@uomosul.edu.iq

شذى حازم داود الزكر
s.alzar@uomosul.edu.iq

قسم هندسة السدود والموارد المائية، كلية الهندسة، جامعة الموصل، الموصل، العراق

تاريخ الاستلام: 17 سبتمبر 2023 استلم بصيغته المنقحة: 29 نوفمبر 2023 تاريخ القبول: 16 يناير 2024

الملخص:

تعد نماذج تدفق الجريان المجزئة احدى تقنيات التوليد العشوائي، اذ ان لديها القدرة على معالجة مشكلة البيانات المرصودة المفقودة والتعويض عنها. كما انها تتيح إمكانية تقسيم بيانات تدفق الجريان من مستوى زمني اعلى الى مستوى زمني أقل والتي تكون مطلوبة لحل مشاكل التصاميم الهيدروليكية والمشاكل التشغيلية للمشاريع. يوجد نهجين رئيسيين لتجزئة بيانات الجريان وهما (Parametric and Non-parametric approachs)، أي النهج المعلمي واللامعلمي والذي يستخدم بشكل واسع لتجزئة تصاريف الأنهار المرصودة . من مميزات نموذج التجزئة قدرته على توزيع قيم بيانات الجريان من محطة واحدة الى عدة محطات زمانيا ومكانيا مع الحفاظ على الخصائص الاحصائية الاساسية للسلسلة الزمنية التي تم الحصول عليها من النموذج (المتوسط ، الانحراف القياسي، معامل الالتواء، الحد الادنى، الحد الاقصى، ومعامل الارتباط) للبيانات التاريخية المرصودة. في الدراسة الحالية، تم استخدام النهج اللامعلمي في تجزئة البيانات. حيث انه من المفترض ان هناك تصاريف مجمعة في محطة رئيسة وسيتم تجزئة هذه البيانات الى سلسلة مقابلة من التصاريف زمانياً ومكانياً في المحطات الفرعية والتي تشبه احصائياً تلك التي تم الحصول عليها من خلال استخدام منصة برنامج SAMS 2010 (Stochastic Analysis, Modelling and Simulation). تم استخدام بيانات التصاريف السنوية والشهرية لخمس محطات لقياس التصاريف على نهر دجلة في العراق وهي (محطة سد الموصل على نهر دجلة ، محطة اسماوا على نهر الخازر ، محطة اسكي كلك على الزاب الأعلى، محطة سد الدبس على آلزاب الأسفل، ومحطة بيجي على نهر دجلة) ولفترة زمنية مقدارها ثلاثة وعشرون عاما. تمت مقارنة النتائج الإحصائية لنهج التجزئة مع نظيراتها المرصودة وأظهرت توافق جيداً في معظم السنوات والأشهر ولجميع المحطات. وبناء على ذلك ، يوصى باستخدام نهج التجزئة عندما تتطلب القرارات استراتيجيات إدارة المياه في هذه المناطق.

الكلمات الدالة :
تجزئة تصاريف الانهر ، النهج اللامعلمي ، الطرائق المكانية ، الطرائق الزمانية، نهر دجلة .