

# Applying the Hierarchical Gray Relational Clustering Method to Municipal Water Use in Turkey

Mehmet Şamil GÜNEŞ<sup>1</sup>  
Coşkun PARİM<sup>2</sup>  
Doğan YILDIZ<sup>3</sup>  
Ali Hakan BÜYÜKLÜ<sup>4</sup>

## ABSTRACT

This study aims to compare the municipal water distribution performance and classify the municipal water distribution systems in the provinces of Turkey using the Gray Relational Hierarchical clustering method for both 2006 and 2016. A correlation analysis was used to determine the variables affecting the distributed water. The Hierarchical gray clusters and the changes during the decade are presented with spatial distribution maps using statistically significant variables. Consequently, the findings reveal that the municipalities of Istanbul, Izmir, Bursa, Gaziantep, and Kocaeli managed water distribution effectively, whereas the municipalities of Konya, Şanlıurfa, Diyarbakır, Samsun, Trabzon, and Sakarya managed it poorly.

**Keywords:** Climate change, municipal water use, water management, sustainability.

## 1. INTRODUCTION

The importance of water distribution systems has increased gradually in recent years due to direct or indirect factors such as increasing population, water consumption, industrialization, seasonal conditions, and climate change [1, 2]. Currently, a critical concern is to ensure urban water sustainability and to improve the decision-making processes for sustainable development planning [3]. Water distribution modelling is used to predict what the future water consumption and requirements depend on, based on various socio-economic and

---

Note:

- This paper was received on July 19, 2020 and accepted for publication by the Editorial Board on November 9, 2020.
- Discussions on this paper will be accepted by May 31, 2022.
- <https://doi.org/10.18400/tekderg.771613>

1 Department of Statistics, Yıldız Technical University, Istanbul, Turkey - [mşgunes@yildiz.edu.tr](mailto:mşgunes@yildiz.edu.tr) - <https://orcid.org/0000-0001-5842-5181>

2 Department of Statistics, Yıldız Technical University, Istanbul, Turkey - [cparim@yildiz.edu.tr](mailto:cparim@yildiz.edu.tr) - <https://orcid.org/0000-0002-6412-1325>

3 Department of Statistics, Yıldız Technical University, Istanbul, Turkey - [dyildiz@yildiz.edu.tr](mailto:dyildiz@yildiz.edu.tr) - <https://orcid.org/0000-0001-7430-2368>

4 Department of Statistics, Yıldız Technical University, Istanbul, Turkey - [hbuyuklu@yildiz.edu.tr](mailto:hbuyuklu@yildiz.edu.tr) - <https://orcid.org/0000-0002-4174-4538>

climatic factors that affect water use. Domestic water consumption is the most significant component of municipal water use, and precise estimation techniques are required to accurately predict future water needs with appropriate methodologies [4].

Gray Relational Analysis (GRA) is a method that has been used recently in cases where the data set is small and contains insufficient information. Missing and insufficient information is very common in hydrological data in water research [5]. GRA was initially developed to be used for interval analysis in cases of insufficient information [6, 7]. Deng [8-10], on the other hand, produced the gray system theory, which can formally deal with small samples using fuzziness. GRA and its simple concept can be observed in many analyses that exhibit excellent performance when there is a small sample and insufficient information. In this study, Hierarchical Gray Relation Clustering (HGRC), a combination of GRA and Hierarchical Cluster Analysis, is applied to the distribution water model.

This study aims to compare the water distribution performance of the municipalities for 2006 and 2016 and to classify the provinces using HGRC. One of the main goals of the study is to use the HGRC method to determine the variables and regions that affect municipal water distribution.

This study is structured as follows. First, the scientific literature on water distribution systems and the basis of applying the method to municipal water distribution models is reviewed. Then, the study area and data are presented, followed by a discussion of the calculation procedure for hierarchical relation gray clustering analysis. Next, HGRC is applied to the variables that explain the distributed water, and the results are visualized with three spatial distribution maps. The cluster results in this study are then discussed, and the changes in the variables affecting the water distribution during the decade of study are determined. Additionally, the contribution of this analysis to water distribution management is discussed. Finally, suggestions are made for future steps by determining the levels of water management for the provinces and regions.

## **2. STUDY AREA AND DATA**

Turkey is a country with a total geographical area of 780,000 km<sup>2</sup>, comprising 81 provinces. Its municipalities are units with a population of over 2000. According to UN estimates, Turkey has a population of about 84 million, and 75.7% of its population lives in urban centers. Municipal water statistics are published by Turkstat [11] every two years and cover Turkey and its provinces. Therefore, instead of using a time series method, data from this source are analyzed separately for both 2006 and 2016 to illustrate changes, and the variables used are presented in Table 1. Municipal water statistics have been published in Turkey since 2003 (biennially after 2004) and the two years with complete data are 2006 and 2016. The study includes the annual data for 81 provinces.

Previous studies have shown that streamflow has a significant effect on the main variables (river, dam, lake, well, etc.) affecting water distribution [12, 13]. The streamflow is not included in this study due to the lack of data and discontinuities.

The variables used in the study and their correlations are presented in Table 1. The analysis in this study was concluded on the basis of 81 provinces, reflecting the general characteristics of the data, and the total values for Turkey are displayed in Table 1.

The amount of distributed water is the most convenient variable to represent the municipal water distribution system in Turkey. Therefore, the aspects that explain the amount of distributed water are also described in the water distribution system. Existing literature has revealed that there are many variables that directly or indirectly affect the amount of water distributed [14-16]. Correlation analysis was thus used to determine whether all the data received from Turkstat [11] are related to the amount of water distributed. These variables, their definitions, and the correlation table are presented in Table 1. As can be observed in this table, the variables that have a significant correlation with the amount of water distributed are TNM, RPSD, NS, WIN, TAWA, LAKE, RIV, DAM, WAPC, NTP, TC, and AWT, while the variables whose correlations with the AWD are insignificant are TNM, SPR, and WELL. The variables that were not significant in the correlation analysis were removed, and only the significant variables were used in the HGRC analysis.

Table 1 - Variables used in the study (with descriptives for 2006 and 2016) and their correlations.

Municipal Water Use Indicators, 2006-2016				
Variable Codes	Definition of Variables	2006	2016	
AWD	Amount of water distributed (m <sup>3</sup> )	2375043	3732875	
TNM	Total number of municipalities	3225	1397	
TMP	Total municipal population	58581515	74911343	
RPSD	Rate of the population served by drinking water treatment plants in total municipal population (%)	41	55	
NS	Number of subscribers	19358951	27486201	
WIN	Water income (Turkish Lira)	3096377755	14217798093	
TAWA	Total amount of water abstracted (Thousand m <sup>3</sup> )	5163500	5838561	
SPR	Spring (Thousand m <sup>3</sup> )	1380057	1000205	
LAKE	Lake (Thousand m <sup>3</sup> )	232621	104354	
RIV	River (Thousand m <sup>3</sup> )	305271	552624	
DAM	Dam (Thousand m <sup>3</sup> )	1843736	2618225	
WELL	Well (Thousand m <sup>3</sup> )	1401815	1563154	
WAPC	Water abstraction per capita in municipalities (liters/capita-day)	245	217	
NTP	Total number of treatment plants	139	519	
TC	Total capacity	3994060	6592863	
AWT	The total amount of water treated	2426639	3350389	

Source: Turkish Statistical Institute [21]

	TN M	TM P	RPS D	NS	WI N	TAW A	SPR	LAK E	RIV	DA M	WELL	WAP C	NT P	TC	AW T	
<b>AW D</b>	Correlation Coefficient	0.20	0.9 9	0.37	0.9 9	0.9 1	0.98	0.01	0.44	0.8 4	0.92	0.25	-0.11	0.4 1	0.9 0	0.95
	<i>p</i> - Value	<i>p</i> ≤ 0.05			<i>p</i> > 0.05		<i>p</i> ≤ 0.05		<i>p</i> > 0.05		<i>p</i> ≤ 0.05		<i>p</i> > 0.05			

HGRC analysis was then applied to examine the municipal water distribution system and yearly changing state of water management for all provinces in Turkey. Excel for HGRC was used to perform the analysis. Variables found to be significant through the correlation analysis in explaining the amount of water distributed were used in HGRC, and their spatial distribution maps are presented with the help of maps.

### 3. METHODS

#### 3.1. Hierarchical Gray Relational Clustering

Deng [17] first proposed the gray system theory in 1982. Gray means that a system provides partially known and partially unknown information; thus, it basically is an attempt to model uncertainty. Gray systems theory is used to examine systems that analyze relationships between systems, make predictions and decisions, and create models [18]. In 1987, Deng [19] first proposed gray cluster analysis (GCA). A new method has been developed known as gray relational clustering (GRC), which combines gray relational analysis and clustering to capture the complex factors and mixed structure in gray cluster analysis. The GRC method cannot use a tree diagram for classification without recalculation [20]. Hence, Wu et al. [20] combined the GRC method with hierarchical cluster analysis.

The calculation procedure for hierarchical gray relation clustering analysis proceeds as follows for the data set used in our study.

Let  $x_{jk}$  specify the  $k$ th coordinate axis of the  $j$ th province, and let  $x_j$  represent the indices for the  $j$ th province. It may be written as follows:

$$x_j = (x_{j1}, x_{j2}) \tag{1}$$

**Step 1.** Compute the difference series:

$$\Delta_{ij}(k) = |x_i(k) - x_j(k)| \tag{2}$$

**Step 2.** Compute the minimum and maximum of the difference series:

$$\begin{aligned} \Delta_{\max} &= \max_{\forall j \in I} \max_{\forall k} |x_i(k) - x_j(k)| \\ \Delta_{\min} &= \min_{\forall j \in I} \min_{\forall k} |x_i(k) - x_j(k)| \end{aligned} \tag{3}$$

**Step 3.** Compute the gray relation coefficient:

$$\gamma(x_i(k), x_j(k)) = \frac{\Delta_{\min} + \zeta \Delta_{\max}}{\Delta_j(k) + \zeta \Delta_{\max}} \quad \zeta = 0.1 \quad i = 1, 2, \dots, m \tag{4}$$

The value of  $\zeta$  is used to expand or compress the range of the gray relation coefficient; it is called the distinguishing coefficient and represents the significance of  $\Delta_{\max}$  [21, 22]. If the

difference between the analyzed data is large, the  $\zeta$  value should be determined to be close to 0 [23]. Therefore, the value of  $\zeta$  is taken as 0.1.

**Step 4.** Compute the gray relation grade to develop the matrix  $R(R=(\Gamma_{ij}), i, j=1,2,\dots,m)$ :

$$\Gamma_{ij} = \frac{1}{k} \sum_{k=1}^k \gamma(x_i(k), x_j(k)) \quad i=1,2,\dots,m, j \in i, k=1,2 \quad (5)$$

**Step 5.** Develop the matrix  $G(g_{ij} = (\Gamma_{ij} + \Gamma_{ij}) / 2)$  (known as the gray similar matrix, this is the crucial output for gray relational clustering):

$$G = [g_{ij}], \quad i, j = 1, 2, \dots, m \quad (6)$$

**Step 6.** Determine the two points (province) of the most near.

**Step 7.** Repeat steps 1–6 until all data are in one cluster [20].

## 4. RESULTS AND DISCUSSION

### 4.1. Hierarchical Gray Relational Clustering for 2006 and 2016 and Changes during the Ten-Year Period

In this study, the distances of the cluster centers to each other are calculated. Provinces are then divided into five classes to demonstrate the data's general structure and facilitate interpretation. As a result of the evaluations made based on the variables, clusters are named according to their characteristics from very poor to very high for both 2006 and 2016.

When the results of the HGRC analysis were evaluated for 2006 (Figure 2), Istanbul, which is Turkey's most populous city, emerged as the city with the best water management. Furthermore, it is notable that the provinces neighboring Istanbul were at poor and very poor levels. While the capital Ankara had medium-level management, the water management level of Izmir, which is another high-population province, was very poor. It was observed that Konya, Eskişehir, and Antalya followed a well-distributed water policy for 2006 in the Central Anatolia region. The provinces in the eastern and southeastern regions apart from Şanlıurfa and Diyarbakır showed a poor level, and only the province of Balıkesir performed well in the Aegean region.

Turning to 2016 (Figure 3), while Istanbul maintains a well-distributed water management system, it may be observed that Ankara, the capital city, remained at the same level. Contrary to its poor level in 2006, Izmir improved to a medium level in 2016, while its surrounding provinces dropped to a very poor level. It can be clearly seen from Figure 3 that there were severe problems in distributed water management system in the Mediterranean and Black Sea coastal regions. While Gaziantep continued to have a high standard of water management approach, the management approach of the other municipalities in the Southeastern Anatolia region in 2016 had deteriorated. It was determined that the distributed water management of Turkey had deteriorated in 2016 compared to 2006.

Figure 4 shows how the management of the distributed water changed over 10 years. This figure illustrates the general structure of the long-term change result of the HGRC spatial distribution maps depicted earlier in Figures 2 and 3. It demonstrates that Izmir, Bursa, Kocaeli, and Gaziantep Provinces displayed positive developments in distributed water management over the decade. The Aegean region, on the other hand, is the region where water management generally worsened. It was also determined that in the Central Anatolia region, Eskişehir and Konya managed to distribute water poorly. Figure 4 shows that in the Eastern Anatolian region, water distribution in Erzurum, Şanlıurfa, and Diyarbakır has deteriorated. On the other hand, the provinces colored white showed no change in distributed water management level between 2006 and 2016.



Figure 2 - Municipal water distribution model results for 2006 according to Hierarchical Gray Relational Clustering



Figure 3 - Municipal water distribution model results for 2016 according to Hierarchical Gray Relational Clustering

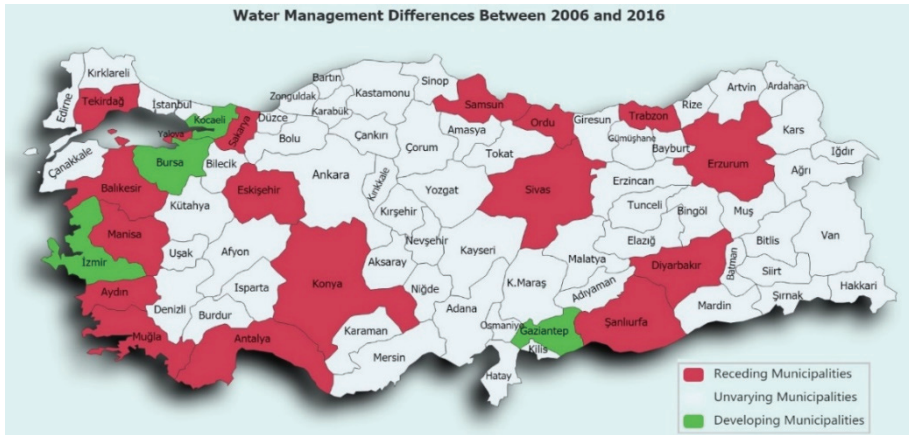


Figure 4 - Municipal water management results according to Hierarchical Gray Relational Clustering for a ten-year period

## 5. CONCLUSIONS

This study presented a comparative analysis of the relationships between such variables as water withdrawal, dams, lakes, wells, and treated water affecting the distributed water systems in Turkey over a decade (2006 and 2016). Population, water income, dams, streams, wells, and treated water can be considered as studied factors affecting water distribution, since their respective correlation coefficients are statistically significant.

According to the results of Hierarchical Gray Relational Clustering, we conclude that the municipalities of İzmir, Bursa, Gaziantep, and Kocaeli displayed an improved water management system over ten years. In contrast, the municipalities of Konya, Şanlıurfa, Diyarbakır, Samsun, Trabzon, and Sakarya experienced challenges in managing the water distribution. For example, the main reason for the shift in water management in Sakarya is the water withdrawal project for Istanbul. Istanbul, which is a megacity, is able to directly affect the environment in the distributed water management system and is part of the source of the distributed water from the surrounding municipalities. Large withdrawal water projects like the Melen River Project [24] can provide a valuable water source for Istanbul from the provinces of Düzce and Sakarya. The project is outlined for water withdrawal from the Melen River, which is located 180 kilometers east of Istanbul in Düzce Province in the Black Sea region [25]. For instance, in the province of Sakarya, water distribution management was in a poor condition in 2006, and this situation continued in 2016. It has been observed that the large amount of water withdrawal for the megacity of Istanbul affected Sakarya and the provinces surrounding Sakarya (this can be observed in Figures 2, 3, and 4). The amount of distributed water in the province of Düzce has doubled in 10 years (5,393,040 m<sup>3</sup> in 2006 and 11,293,906 in 2016 according to TUIK, Municipal Water Statistics [11]). When we consider the province of Düzce, whose current water management is poor, it can be expected that the withdrawal of large amounts of water will have a major effect on water management. One of the external factors affecting water distribution during this period was migration. Gaziantep, Şanlıurfa, Istanbul, and İzmir were some of the major cities witnessing immigration.

Migration was one of the factors in Şanlıurfa's deterioration in 2016, which had a good water management system in 2006. A reverse trend can be observed in Gaziantep [26]. Additionally, the province of Konya was one of the municipalities that witnessed the greatest groundwater withdrawal, which was responsible for the formation of a large number of sinkholes and the disappearance of groundwater resources [27]. For the surrounding provinces and the Central Anatolian region, this problem raises the problem of water scarcity, and this problem will worsen in the future if the distributed water is not managed well.

Most importantly, this study revealed that the distributed water management system was in a poor condition and that the number of provinces showing deterioration have been increasing.

From the study results, we note that the efficient management of the municipal water distribution is highly dependent on the environmental policy of the distributed water management, and this dependence reveals the necessity for an integrated water management system. Provinces sharing the same water basin and its source should thus follow a common distributed water policy. The results of this study will have a significant impact on research and analysis of distributed water system management in Turkey, which is an important contribution, due to the use of the Gray Relational Hierarchical clustering method.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Reference

- [1] Mayer, P. W., DeOreo, W. B., Opitz, E. M., Kiefer, J. C., Davis, W. Y. Dziegielewski, B., & Nelson, J. O., *Residential end uses of water*, 1999.
- [2] Corbella, H. M., & i Pujol, D. S. What lies behind domestic water use?: a review essay on the drivers of domestic water consumption. *BAGE: Boletín de la Asociación de Geógrafos Españoles*, (50), 297-314, 2009.
- [3] Childers, D. L., Pickett, S. T., Grove, J. M., Ogden, L., & Whitmer, A. Advancing urban sustainability theory and action: Challenges and opportunities. *Landscape and urban planning*, 125, 320-328, 2014.
- [4] Firat, M., Yurdusev, M. A., & Turan, M. E. Evaluation of artificial neural network techniques for municipal water consumption modeling. *Water resources management*, 23(4), 617-632, 2009.
- [5] Ip, W. C., Hu, B. Q., Wong, H., & Xia, J. Applications of grey relational method to river environment quality evaluation in China. *Journal of Hydrology*, 379(3-4), 284-290, 2009.
- [6] Moore, R.E., *Methods and applications of interval analysis*. SIAM, 1979.
- [7] Huang, Y. P., & Huang, C. C. The integration and application of fuzzy and grey modeling methods. *Fuzzy sets and Systems*, 78(1), 107-119, 1996.
- [8] Julong, D. Introduction to grey system theory. *The Journal of grey system*, 1(1), 1-24, 1989.
- [9] Deng, J. L. Grey Prediction and Decision, *Huazhong University of Science and Technology*, 1986.



- [10] Deng, J., Introduction to grey system theory. *The Journal of Grey System* vol, 1989.
- [11] TÜİK. Turkish Statistical Institute. Available from: <http://www.turkstat.gov.tr/>, 2020.
- [12] Piccolroaz, S., Calamita, E., Majone, B., Gallice, A., Siviglia, A., & Toffolon, M. Prediction of river water temperature: a comparison between a new family of hybrid models and statistical approaches. *Hydrological Processes*, 30(21), 3901-3917, 2016.
- [13] Zhu, S., Heddam, S., Nyarko, E. K., Hadzima-Nyarko, M., Piccolroaz, S., & Wu, S. Modeling daily water temperature for rivers: comparison between adaptive neuro-fuzzy inference systems and artificial neural networks models. *Environmental Science and Pollution Research*, 26(1), 402-420, 2019.
- [14] Villarín, M. C. Methodology based on fine spatial scale and preliminary clustering to improve multivariate linear regression analysis of domestic water consumption. *Applied Geography*, 103, 22-39, 2019.
- [15] House-Peters, L., Pratt, B., & Chang, H. Effects of urban spatial structure, sociodemographics, and climate on residential water consumption in hillsboro, oregon 1. *JAWRA Journal of the American Water Resources Association*, 46(3), 461-472, 2010.
- [16] Griffin, R. C., & Mjelde, J. W. Distributing water's bounty. *Ecological Economics*, 72, 116-128, 2011.
- [17] Ju-Long, D. Control problems of grey systems. *Systems & control letters*, 1(5), 288-294, 1982.
- [18] Tsai, M. T., Hsiao, S. W., & Liang, W. K. Using grey theory to develop a model for forecasting the demand for telecommunications. *Journal of Information and Optimization Sciences*, 26(3), 535-547, 2005.
- [19] Deng, J. Wuhan, China, *Basic method of grey system*. 1987.
- [20] Wu, W. H., Lin, C. T., Peng, K. H., & Huang, C. C. Applying hierarchical grey relation clustering analysis to geographical information systems—A case study of the hospitals in Taipei City. *Expert Systems with Applications*, 39(8), 7247-7254, 2012.
- [21] Hinduja, A., & Pandey, M. (2017). Multicriteria recommender system for life insurance plans based on utility theory. *Indian Journal of Science and Technology*, 10(14), 1-8.
- [22] Yildirim, B. F., Hepsen, A., & Onder, E. (2015). Grey Relational Analysis Based Ranking of Latin American and Caribbean Economies. *Journal of Economics Finance and Accounting*, 2(3).
- [23] Yıldırım, B. F. (2015). Gri ilişkisel analiz. Çok Kriterli Karar Verme Yöntemleri, 229-236.
- [24] Islar, M., & Boda, C. Political ecology of inter-basin water transfers in Turkish water governance. *Ecology and Society*, 19(4), 2014.
- [25] DSI, *The General Directorate of State Hydraulic Works*, Ankara, Turkey, 2020.
- [26] Doganay, M., & Demiraslan, H. (2016). Refugees of the Syrian civil war: impact on reemerging infections, health services, and biosecurity in Turkey. *Health security*, 14(4), 220-225.
- [27] Günay, G., Çörekçioğlu, İ., & Övül, G. Geologic and hydrogeologic factors affecting sinkhole (obruk) development in Central Turkey. *Carbonates and evaporites*, 26(1), 3-9, 2011.

