# DESIGN OF THE GROUNDWATER LEVEL MONITORING NETWORK USING PRINCIPAL COMPONENT ANALYSIS (PCA) TECHNIQUE

**Atefeh sayadi shahraki[1], Saeed Broomandnasab[1], Abd Ali Naseri[1] and Amir Soltani Mohammadi[1]**
[1]Irrigation and Drainage Department, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran
Sayadi.Atefeh@gmail.com

(Received June 2019 – Accepted April 2021)

## *ABSTRACT*

**Shahraki, A. Broomandnasab, S. Naseri, A. & Mohammedi, A. (2021). Design of the groundwater level monitoring network using principal component analysis (PCA) technique. *Lebanese Science Journal,* 22(1), 22- 37.**

*Using an appropriate monitoring network is considered as an efficient option to manage the groundwater resources and reduce drilling of costly sampling wells. Principal component analysis (PCA) is one of the data reduction techniques used to extract essential components. The used techniques are based on the identification of those describing the variance of the system. In this paper, the PCA technique has been employed in order to identify the effective wells and remove the less important ones. For this purpose, 160 wells were constructed in the Salman Farsi Agro-Industry, located in Khuzestan province of Iran. The data are measured twice a month for 12 months. In this technique, variation factors called principal components are identified through considering the data structures. Using the PCA, the relative importance of each well has been calculated for the groundwater depth estimation. In the present study, the acceptable threshold has been taken to be 0.8 and therefore the number of wells in determining groundwater depth was reduced to 33 ones. Identifying the essential wells, the important points for sampling are identified and groundwater depth monitoring is performed only in these wells. This will save time and cost of groundwater level monitoring within the study area.*

**Keywords**: Groundwater, Principal Component Analysis, Monitoring.

http://dx.doi.org/10.22453/LSJ-022.1.022-037

lsj.cnrs.edu.lb/vol-22-no-1-2021/

## INTRODUCTION

The importance of water in all parts of human life, including the drinking, industry and agriculture, has led to a great deal of focus on the critical areas such as water demand management and water quality conservation (Elçi & Ayvaz, 2014). Due to the complexities of the underground water environment and considerable costs of conventional monitoring methods, the invention of new technologies and use of advanced methods will help improve the underground water systems. The design approaches of groundwater monitoring network are divided into two main categories of land-based and statistical analysis (Helena et al., 2000). The geological methodology is based on the quantity as well as quality of the geological and groundwater information and advanced statistical methods are not used in (Lucas and Jauzein, 2008). The installation and maintenance costs associated with the groundwater level readings in the observation wells are proportional to their number in the network (Sheikhy Narany et al., 2013). Designing the groundwater monitoring network is the same as determining the number of wells and their distribution as well as density in the area. The groundwater monitoring goals include the ambient resource condition monitoring, compliance monitoring, risk detection monitoring, research monitoring, or a combination of these (Gangopadhyay et al., 2001). In general, the monitoring network optimization problem is a nonlinear hybrid one and optimization algorithms are appropriate for solving it. The purpose of designing the network model is to find an optimal sampling map from many cases which are solved using iterative processes and thus require complicated computational efforts.

PCA is an optimal mathematical method for reducing the volume of data and transforming primary variables into several components (Richards & Jia, 1999; Jolliffe, 2002). In this method, according to the data structure, the factors causing changes which are called the principal or implicit components are identified (Pearson, 1901). In this way, after identifying the components that make the most variance changes, the variables with the highest correlation coefficient with the main components can be extracted. Khan et al. (2008), developed a process for employing PCA in order to optimize a groundwater monitoring network in an irrigation area of Australia. Their results indicated that the overall difference of groundwater level between the original piezometer and optimized networks after the PCA process is less than 20%. At the same time, the total number of piezometers in the optimized network is reduced by 63%, which will save the time and cost of the groundwater levels monitoring within the irrigation area. The real strength of PCA lies in its dimensionality reduction capability for multivariate data sets (Richards and Jia, 1999). By reducing the dimensionality, a smaller principal component (PC) set may calculate the contributions of individual variables rather than the original multivariate data set. It is a common practice to aggregate the variance of first few PCs in order to obtain more than 80% of the variance

http://dx.doi.org/10.22453/LSJ-022.1.022-037

of the original multivariate data. Adil et al. (2011) conducted a study on optimizing the groundwater quality monitoring network in the Indian Maheshwaram basin using PCA and Kriging analysis. The aim of this study was to evaluate the relationship between variables and to minimize the collection of information for monitoring the groundwater quality. The Kriging method was also used to estimate the standard error. Results indicated that 13 of total 61 wells can be removed from future sampling. Also, Selle et al. (2013) examined the spatial and temporal patterns of PC scores in order to improve the understanding of processes governing the groundwater quality in the Ammer catchment located in southwest Germany. The results indicated the influences of land use and geology on the groundwater quality. Sheikhy Narany et al. (2015) used the multi-objective based approach for the groundwater quality monitoring network optimization in Amol-Babol plain. In this study, the natural contaminant factors (NCF) and anthropogenic contaminant factors (ACF) were selected for the optimal design of monitoring wells based on the PCA. The results illustrated that the contamination mass detection capacity of around 86% can be estimated by sampling 114 wells instead of 154 ones in the initial existing monitoring wells within the study area, for both the natural and anthropogenic contaminations. Sánchez-Martos et al. (2001), Debels et al. (2005), Iscen et al. (2008), Wan (2009), Stathis and Myronidis (2009), Gvozdić et al. (2012), Nguyan et al. (2013), Hu et al. (2013), Vonberg et al. (2014) and Wang et al. (2019) also used the main components technique in their research. The present study aims to identify the effective wells in determining the groundwater depth of Salman Farsi Sugarcane Agro-Industry using PCA method.

## MATERIALS AND METHODS

**Case study:** Salman Farsi Sugarcane Agro-Industry is located 40 km south of Ahvaz city, Khuzestan province, Iran. Its agricultural area is about 12,000 hectares, 10,000 hectares of which are annually harvested and the remaining 2,000 hectares are grown and re-cultivated. Salman Farsi Agro-Industry is limited to the Debal Khazaee sugarcane agro-industry from the north, Ahvaz-Abadan road from the east and Karun River from the west. The research area has a dry climate with scorching summers and mild winters. The coldest and warmest months of the year are January (with the lowest temperature of 7.5 °C) and July (with the highest temperature of 47 ° C) respectively. Moreover, the average precipitation of the study area is 266 mm and the annual evaporation is reported as 2788 mm. The only irrigation source of the farms is the large Karun River. Physiologically, the study area is flat and relatively of low altitude. The general slope of the land is from the northern highlands to the shores of the Persian Gulf in the south, so that it is shallow in the south of Ahvaz to Abadan (about 1/10 m per kilometer). In addition, the altitude of this agro-industry is about 2-4 m above the sea level. The soil texture is dense in the area, mostly composed of clay and silt. The position of the Salman Farsi Sugarcane Agro-Industry is shown in Figure 1.

http://dx.doi.org/10.22453/LSJ-022.1.022-037

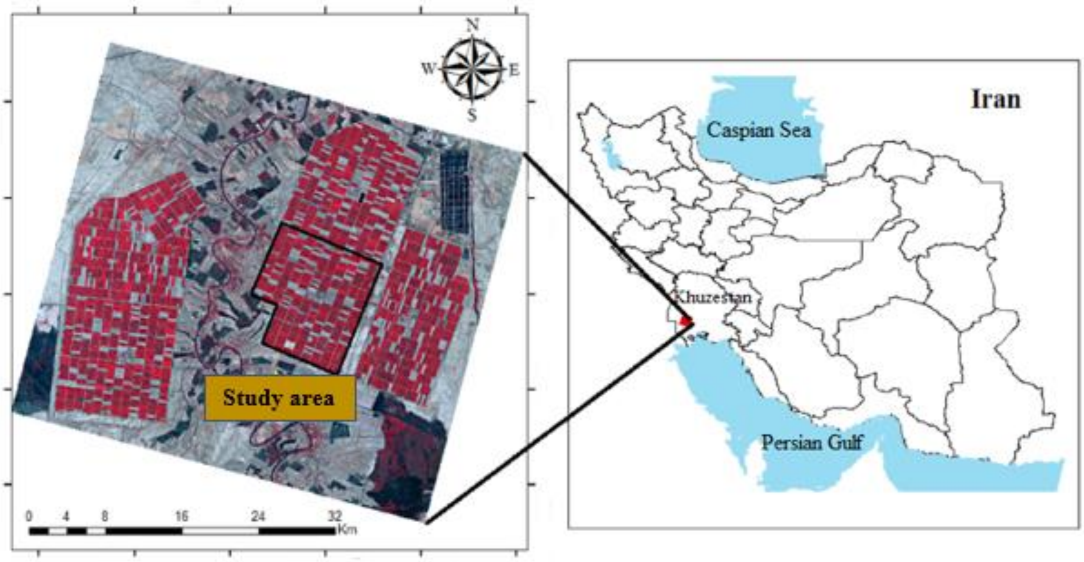


**Figure 1. Location of Salman Farsi Agro Industry unit in Southwest Iran.**

**Principal Components Analysis (PCA):** PCA is a multivariate statistical method used for data reduction as well as deciphering patterns within the large data sets (Hull, 1984; Joliffe, 1986; Wold et al., 1987; Stetzenbach et al., 1999). The purpose of this analysis is to transform the data table into a new orthogonal and uncorrelated set of factors (PC) and to extract important information from them. This new representation compresses the data by keeping only the most important information, simplifying the explanation and exploring the structure as well as composition of the data (Abdi & Williams, 2010). It is assumed that $X$ is an $n * p$ matrix, where n is the number of observations for $p$. In the proposed scheme, $n$ is the number of statistical periods in which the groundwater level is measured and $p$ denotes the number of wells or stations. Using the correlation of water level in the $p$-well adjacent, the relative importance of each well will be determined in the display of groundwater depth changes using PCA. In this analysis, the PCs are defined as the following linear functions (Sauquet, 2000):

$$\begin{aligned} Z_1 &= Xa_1 = a_{1,1}X_1 + a_{2,1}X_2 + \cdots + a_{p,1}X_p \\ Z_2 &= Xa_2 = a_{1,2}X_1 + a_{2,2}X_2 + \cdots + a_{p,2}X_p \\ &\vdots \\ Z_p &= Xa_p = a_{1,p}X_1 + a_{2,p}X_2 + \cdots + a_{p,p}X_p \end{aligned} \qquad (1)$$

where $a_{i,j}$ is the $i$th element of the $j$th principal component and $a_j$ stands for the conversion coefficient of the principal variable ($X$) to the $j$th main component ($Z_j$).

Using properties of the matrices, it can be proved that the coefficients of the principal components ($a_j$) are special vectors related to the covariance matrix S. The value and special vector of matrix $S$ are calculated from the following equations:

$$[S - \partial I] = 0 \quad (2)$$

In the above relation, if *p* is the number of wells, then *I* is an identity matrix of *P* * *I* and *S* is the covariance of order *P* which can be calculated from the following equation:

$$S = \frac{X^T X}{n-1} \quad (3)$$

In the above equation, *T* stands for the transpose of the matrix. The following constraints are given while solving equation 2:

The eigenvectors must be orthogonal$(a_j^T a_i = a_i^T a_j = 0, i \neq j)$
The eigenvectors must be orthonormal or unit $(a_j^T a_j = 1)$

The above constraints make the answers to equation 2 unique and therefore the main components of $Z_j$ are independent. If $a_1, a_2, \ldots, a_p$ are the *eigenvectors* corresponding to the eigenvalues $\partial_1, \partial_2, \ldots, \partial_p$ (as for $i < j$ ، $\partial_i > \partial_j$) respectively, then equation 1 is rewritten as follows:

$$Z = XA \quad (4)$$

where $Z = (z_1, z_2, \ldots, z_p)$ and $A = (a_1, a_2, \ldots, a_p)$.

The variances of the main components, $Z_j$, are the eigenvalues corresponding to them. Therefore, the variance of the first component ($Z_1$) is equal to $\partial_1$. So, the first main component has the highest variance, indicating the high capability of the first main component in identifying the groundwater level changes. The first major component is a line whose extension corresponds to the most visible scattering in the original data. Also, the second main component ($Z_j$) has a variance of $\partial_2$, which is in the second place in terms of the variance, and its extension is in the direction that the visible scatter of data is in the second place. The rest of the main components are described in the same way. The main components pass through the center of the main data and are perpendicular to each other.

http://dx.doi.org/10.22453/LSJ-022.1.022-037

In order to calculate the relative importance of each well, the correlation coefficient between the main components and observed data is used. The correlation coefficient between well $I$ ($X_i$) and $j$th component ($Z_j$) can be calculated from the following equation:

$$Cor\left(Z_j, X_i\right) = \partial_{i,j}^{1/2} a_{i,j} \qquad (5)$$

The higher the correlation coefficient, the greater the relative importance of the well (Sauquet, 2000).

**Research method:** In summary, the steps for determining the relative importance of observation wells in Salman Farsi Sugarcane Agro-Industry include the following steps:

1. Selection of wells for the PCA: In this research, 160 observation wells were constructed in the study area and groundwater depth information was collected twice a month during the first 12 months of 2018 in order to identify the significant wells in determining the groundwater depth. Wells have been constructed in 100-m apart strips on both sides of the land border (inside and outside). The wells were 3 m deep with a radius of 4 in. Figure 2 shows the position of observation wells in the area. In the PCA, the number of variables (wells) must be smaller or equal to the observed data (number of sampling periods) (Patterson, 2001). Since the number of statistical periods for measuring the groundwater level is 24, the corresponding information of 24 adjacent wells has been used for monitoring each of them in the current study. For example, for monitoring well No. 105, the data of 24 adjacent wells, i.e. 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 106, 107, 108, 109, 110, 111, 112, 113, 114 and 115, 116, 117 have been used. Therefore, the data matrix X, will be a 24*24 one for analyzing the main components of well No. 105. At this point, a matrix X will be constructed for each well. Also, the data of a well is not used while constructing the corresponding matrix X but only those of the 24 adjacent ones are used.
2. Calculation of the correlation coefficient: The PCA is performed for each of the matrices defined in step 1 and the correlation coefficient is calculated using equation (3).
3. Selection of the effective wells in each PCA: In this research, if the correlation coefficient of a PCA is higher than 0.9 that well is selected as the main or effective well in the monitoring process.
4. Calculation of the relative importance of each well: In steps 2 and 3, PCA is performed for each available well and efficiency is determined in each analysis. The high frequency of well-known wells indicates their

high relative importance. While monitoring the groundwater level, the relative importance of each well is defined by the ratio of the number of times that the well has been identified as an effective one in the PCA to that it is involved in this analysis. This ratio shows the importance of each well over other ones. Therefore, in order to save time and cost, it would be possible to eliminate low-level wells while monitoring the groundwater levels.

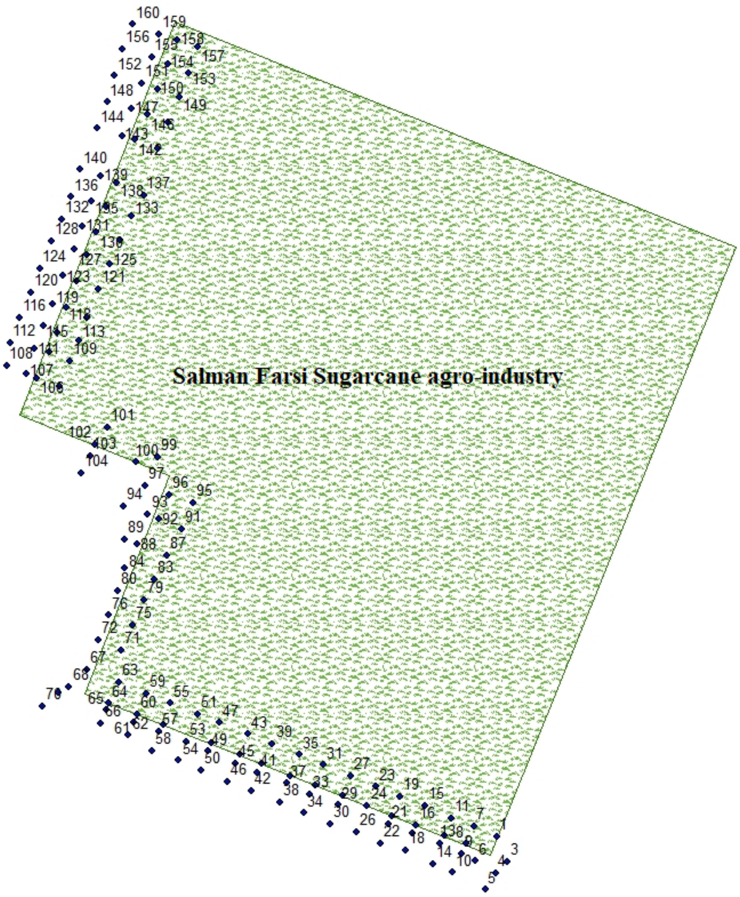


**Figure 2. Location of observation wells.**

5. Analysis of the method used: Removing any well may remove part of the groundwater table information. In order to investigate the uncertainty in selecting the number of wells, the average coefficient of variation of the groundwater level is used. For this purpose, due to the accuracy (or relative importance) of the wells identified and using their corresponding data, the coefficient of groundwater depth changes for each of the years is calculated and it's average is achieved as well. Therefore, as long as the average coefficient of variation does not significantly change, the inefficient wells can

be discarded. Using this method alone will be acceptable if it is assumed that the coefficient of variation increases upon removing the inefficient wells. In this way, the threshold on which the least difference in the coefficient of variation occurs is selected. However, it is important to note that the removal of inefficient wells does not always increase the coefficient of variation and this value increases or decreases depending on the nature of the information removed from the wells. If a well with greater water depth is removed on a certain threshold, the obtained coefficient of variation will be smaller than that associated with all wells. Hence, in order to choose an acceptable threshold level in the present study, in addition to calculating the coefficient of variation, the average amount of uncertainty or monitoring error was calculated using the average water level of the wells (equation 5). Using equations (6) and (7), the coefficient of variation and monitoring error in the case of the removal of inert wells on a certain threshold were obtained from the through comparing the average wells of that threshold with the mean of all wells as:

$$CV = \frac{\sigma}{\mu} \quad (6)$$

$$Error = \frac{(m_n - m_o)}{m_o} * 100 \quad (7)$$

where σ and μ stand for the standard deviation and mean respectively. Also, $m_n$ is the average water level after the removal of ineffective wells on a specified threshold and $m_o$ stands for the average water level of all wells.

## RESULTS AND DISCUSSION

**Calculate the correlation coefficient and Selection of the effective wells in each PCA:** Optimal design of a monitoring network is a controversial issue due to difficulties in the selection of spatial and temporal variations, variables to be monitored and sampling objectives (Mogheir et al., 2005; Masoumi & Kerachian, 2010). Although several techniques have been developed, such as the geostatistical technique, entropy theory, optimization algorithms, statistical methods and hydrogeological evaluation, but no particular one is applicable to all situations. Therefore, the process of monitoring well selection should be representative of the entire groundwater system and also cost-effective (Chadalavada et al., 2011) with a minimum number of sampling wells. In this research, the PCA method was used for the groundwater monitoring. For this purpose, having 24 statistical periods of groundwater level information, 24 adjacent wells were analyzed for each well to determine the effective ones. For example, during the monitoring process of the groundwater level of well No. 105, 24 adjacent ones have been analyzed. In this process, 24 main components were extracted, the value of each is equal to its variance (which is the same as in equation 2). Only a few of these main components, which have the highest variances, describe the system changes and can be used to determine the effective wells and the rest with less variances are ignored. In other words, since the goal is to calculate the correlation coefficient between the wells and main components for determining the effective ones, only the first two components, for which the correlation coefficient is significant, are used. Table (1) presents the analysis results of monitoring well No. 105.

**Table 1. Monitoring correlation matrix of well No. 105.**

| | | Principal components | | | | Principal components | |
|---|---|---|---|---|---|---|---|
| **Row** | **Well** | **$z_1$** | **$z_2$** | **Row** | **Well** | **$z_1$** | **$z_2$** |
| 1 | 93 | 0.35 | 0.4 | 13 | 106 | 0.58 | 0.32 |
| 2 | 94 | 0.58 | 0.39 | 14 | 107 | 0.67 | 0.39 |
| 3 | 95 | 0.45 | 0.22 | 15 | 108 | **0.95** | 0.27 |
| 4 | 96 | 0.27 | 0.35 | 16 | 109 | 0.62 | 0.24 |
| 5 | 97 | **0.89** | 0.09 | 17 | 110 | 0.41 | 0.45 |
| 6 | 98 | 0.31 | 0.13 | 18 | 111 | 0.74 | 0.52 |
| 7 | 99 | 0.57 | 0.27 | 19 | 112 | 0.3 | 0.16 |
| 8 | 100 | 0.22 | 0.26 | 20 | 113 | 0.37 | 0.29 |
| 9 | 101 | 0.36 | 0.5 | 21 | 114 | 0.25 | 0.61 |
| 10 | 102 | **0.9** | 0.11 | 22 | 115 | 0.21 | 0.24 |
| 11 | 103 | 0.77 | 0.3 | 23 | 116 | 0.43 | 0.41 |
| 12 | 104 | 0.71 | 0.09 | 24 | 117 | 0.51 | 0.29 |

According to table 1, the highest correlation coefficient is related to wells No. 108, 102 and 97, which can be considered as the most effective ones in monitoring the groundwater level of well No. 105. The same analysis was carried out for all wells.

**Calculation of the relative importance of each well:** Finally, the relative importance of each well was then determined from the ratio of the number of times being effective to that of being involved in the analysis and the obtained results are presented in Figure 3.

http://dx.doi.org/10.22453/LSJ-022.1.022-037

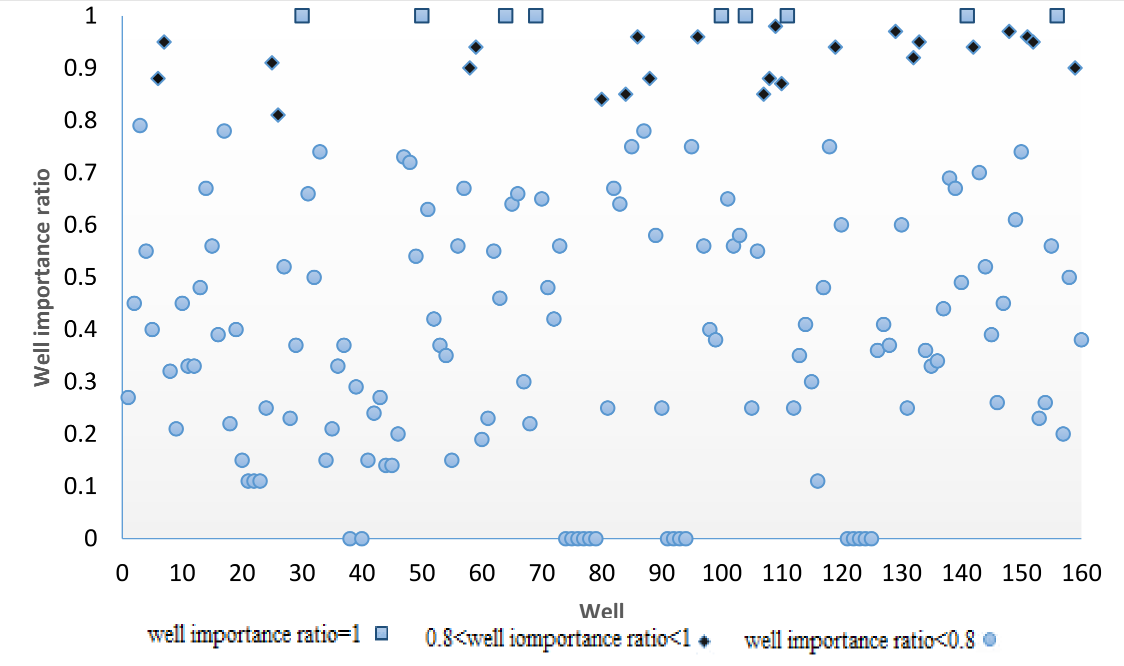


**Figure 3. The relative importance of wells based on the PCA.**

According to figure 3, there are nine wells as the top-ranked ones (30, 50, 64, 69, 100, 104, 111, 141, and 156). These wells have been identified as effective ones each time they have been included in the analysis. Therefore, they are more critical in the groundwater monitoring rather than other wells. This indicates that, they have a higher groundwater level compared to the other wells. If the acceptable threshold is considered to be 0.9, the number of wells will be reduced to 25. For rating the wells, the thresholds are considered to be 0, 0.2, 0.4, 0.6, 0.8 and 1. On the threshold of 1, only wells that are ranked 1 (wells that are known as effective in all analyses) remain while the threshold of 0 contains all wells (effective and ineffective).

**Analysis of the method used:** Figures (4) and (5), depict the coefficient of variation corresponding to the specified thresholds and the monitoring error.

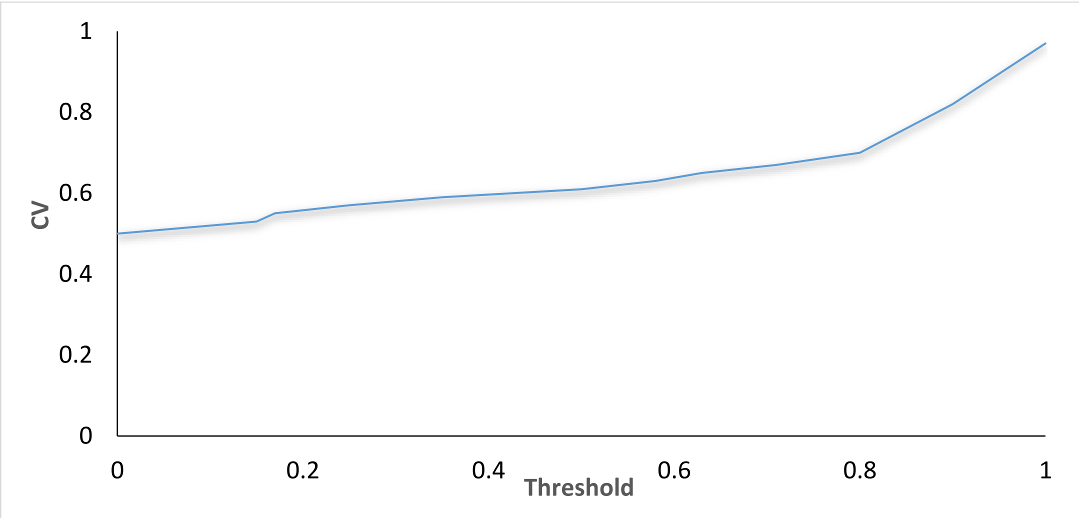


**Figure 4. The coefficient of variation in the specified thresholds.**

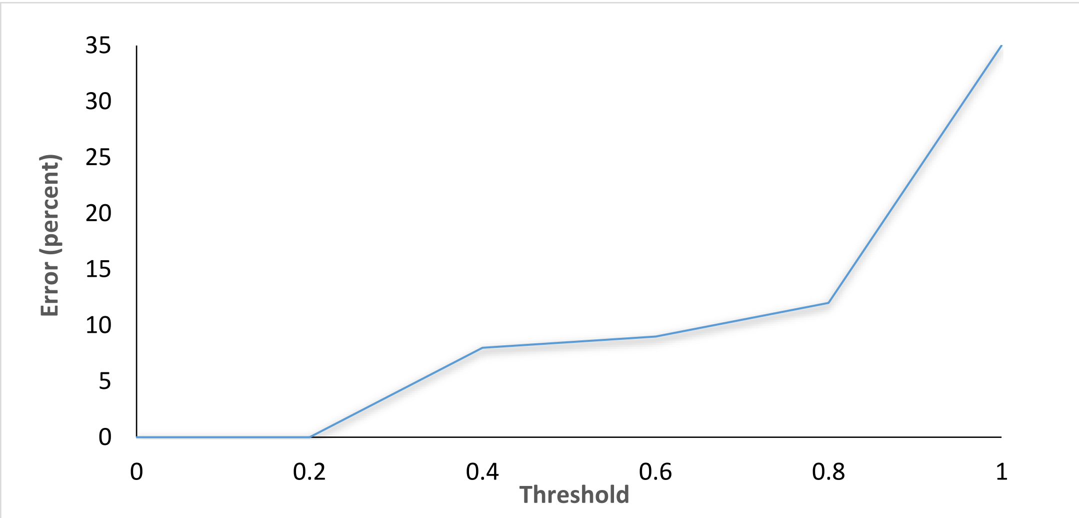


**Figure 5. The amount of monitoring error in the specified thresholds.**

According to figure 4, the average coefficient of variation does not significantly change from the threshold of 0 to 0.8 (where the number of effective wells decreases from 160 to 33 loops). However, if the threshold exceeds 0.8, the groundwater level variation coefficient will increase. Also, as would be observed from the figure, the coefficient of variation increases by removing the ineffective wells on each threshold. This is mainly due to the fact that the eliminated wells had less water depth than the average value. To confirm the results, the monitoring error value was calculated for each threshold. According to figure 5, the error rate gradually increases when the thresholds fall within the range of 0-0.8. However, on the thresholds of 0.9 and 1, the

http://dx.doi.org/10.22453/LSJ-022.1.022-037

error magnitude was found to increase with a higher slope. The error value on the threshold of 0.8 is equal to 12% (therefore, by eliminating the 127 less important wells, the groundwater estimation error will increase by 12% compared to the state in which all wells are included). However, these values are estimated as 25 and 34% on the thresholds of 0.9 and 1, respectively. In the present study, the threshold value was considered to be 0.8 and the number of effective wells in the groundwater level monitoring was reduced from 160 to 33 (the effective wells are shown in figure 3). Also, no significant error is observed in the surface water table display by reducing the number of wells identified. In this way, 127 wells can be removed from the study area in the groundwater modeling to save time and cost. Then, performing measurements and readings of the critical wells may lead to the increased modeling accuracy (Kumar Singh et al. 2009; Subba Rao, 2014; Sheikhy Narany et al. 2015; Zakhem, 2017; Jung et al. 2019).

In order to evaluate the results, the design of the groundwater level monitoring network was performed using the method employed in the study of Khan et al. (2008). The PCA functions, distributed with the MATLAB package, were used to determine the relative contributions of individual wells in capturing the spatiotemporal variation of the groundwater levels. MATLAB has a built-in PCA multifunctional library, which is quite suitable to accomplish PCA of the data set. This MATLAB PCA multifunctional library can provide flexible options for data inputs and visualizing outputs in association with other softwares. Kriging interpolation algorithm was used to render the data surface presentations and determine the spatial differences in wells surfaces using different numbers of data sets. In other words, the kriging method estimates the groundwater level at existing points and elsewhere. Then, using it, the results of the principal component analysis are evaluated and excess monitoring points are determined. Due to the fact that in the present study the number of monitoring points is sufficient (160 points in the 4.1 hectares), to compare the results, the kriging method was used only to evaluate the results of principal component analysis. According to the results, the number of wells removed in the present method was equal to 29 and the error rate was estimated as 19%. This means that the ratio of the overall difference between the groundwater levels corresponding to the original wells and optimized networks after the PCA process to the groundwater level of the original wells network is less than 20%. It should be noted that out of 29 removed wells, 26 ones are shared between this method and the proposed method. By comparing the obtained results, the groundwater level monitoring method used in this study, in addition to its ease of doing, has a lower error rate than the method of Khan et al. (2008).

http://dx.doi.org/10.22453/LSJ-022.1.022-037

## CONCLUSION

The groundwater monitoring network is an essential step in describing the underground water system. The inappropriate distribution of monitoring sites or inadequate numbers of information collected on one side and the waste of cost and time used to measure additional data on the other side, emphasize the necessity of conducting monitoring operations. In this research, effective wells have been identified in determining the groundwater depth of Salman Farsi Sugarcane Agro-Industry using the principal component analysis method. For this purpose, 160 wells were constructed in different parts of the area, and groundwater depth information was measured twice a month for a one-year period. Using the principal components analysis, the relative importance of each well was calculated for the groundwater depth estimation. According to the relative importance of them, one can consider some of them as effective in terms of the acceptable thresholds. In the present study, the acceptable threshold was taken to be 0.8. Based on this value, the number of effective wells in determining the groundwater level depth was reduced to 33. It was found that the average coefficient of variation does not significantly change for the thresholds within the range of 0-0.8 and the associated error rate increases gradually. However, on the thresholds of 0.9 and 1, the error magnitude increases with a higher slope.

## ACKNOWLEDGMENTS

We are grateful to the Research Council of the Shahid Chamran University of Ahvaz and Center of Excellence of the Network Improvement and Maintenance of the Faculty of Water and Environmental Engineering for financial support (GN: SCU.WI99.399). We also thank two anonymous reviewers and editor of this journal for their helpful comments and suggestions. Also, great thanks of Khuzestan Water and Power Authority (KWPA) for financial support.

## REFERENCES

Adil, N., Adrian, H. and Shakeel, A. (2011). Optimization of a Groundwater Monitoring Network for a Sustainable Development of the Maheshwaram Catchment, India. Sustainability, 3(2), 396-409.

Abdi, H. and Williams, L. J. (2010). Principal Component Analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433-459.

Chadalavada, S., Datta, B. and Naidu, R. (2011). Uncertainty based optimal monitoring network design for a chlorinated hydrocarbon contaminated site. Environmental Monitoring and Assessment, 173(1–4): 929–940.

Debels, P., Figueroa, R., Urrutia, R., Barra, R. & Niell, X. (2005). Evaluation of water quality in the Chilia River using physicochemical parameters and a modified water quality index. Environmental Monitoring and Assess, 110, 301–322.

Elçi, A. and Ayvaz, M. T. (2014). Differential-evolution algorithm based optimization for the site selection of groundwater production wells with the consideration of the vulnerability concept. Journal of Hydrology 511, 736–749.

Gangopadhyay, S., Das Gupta, A. & Nachabe, M. H. (2001). Evaluation of ground water monitoring network by principal component analysis. Ground Water, 39 (2), 181–191.

Gvozdić, V., Brana, J., Malatesti, N. & Roland, D. (2012). Principal component analysis of surface water quality data of the River Drava in eastern Croatia (24 year survey). Journal of Hydroinformatics, 14 (4), 1051-1060.

Helena, B., Pardop, R., Vega, M., Barrado, E., Manuel, J. & Fernandez, L. (2000). Temporal evolution of groundwater composition in an alluvial aquifer by principal component analysis. Water Resource, 34 (3), 807-816.

Hu, S., Luo, T. & Jing, C. (2013). Principal component analysis of fluoride geochemistry of groundwater in Shanxi and Inner Mongolia, China. Journal of Geochemical Exploration, 135, 124–129.

Hull, L.C. 1984. Geochemistry of ground-water in the Sacramento Valley, CA. USGS Professional Paper 1401-B. Reston, Virginia: USGS.

Iscen, C., Emiroğlu, O., Ilhan, S., Arslan, N., Yilmaz, V. & Ahiska, S. (2008). Application of multivariate statisticaltechniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environmental Monitoring and Assess, 144 (1-3), 269–276.

Jolliffe, i.t. 2002. Principal Component Analysis. Springer series in statics, ISBN 978-0-387-95442-4.

Joliffe, I.T. 1986. Principal Component Analysis. Berlin, Germany: Springer.

Jung, H., Koh, D. C., Kim, Y., Ha, K. and Lee, J. (2019). Interpretation of Groundwater Level Variations in Jeju Island by Principal Component Analysis. Geophysical Research Abstracts, 21. 2019EGUGA.21.6332J.

Khan, S., Chen, H F. & Rana, T. (2008). Optimizing Ground Water Observation Networks in Irrigation Areas Using Principal Component Analysis. Ground Water Monitoring & Remediation, 28 (3), 93-100.

Kumar Singh, S., Kumar Singh, C., Kumar, K., Gupta, R. and Mukherjee, S. (2009). Spatial-Temporal Monitoring of Groundwater Using Multivariate Statistical Techniques in Bareilly District of Uttar Pradesh, India. Journal of Hydrology and Hydromechanics, 57(1), 45-54.

Lucas, L. & Jauzein, M. (2008). Use of principal component analysis to profile temporal and spatial variations of chlorinated solvent concentration in groundwater. Environmental Pollution, 151 (1), 205-212.

Masoumi, F. and Kerachian, R. (2010). Optimal redesign of groundwater quality monitoring networks: a case study. Environmental Monitoring and Assessment, 161(1–4), 247–257.

Mogheir, Y., De Lima, J. L. M. P. and Singh, V. P. (2005). Assessment of in formativeness of groundwater monitoring in developing regions (Gaza strip case study). Water Resources Management, 19(6), 737–757.

Nguyan, T. T., Nakagawa, A. K., Amaaguchi, H. & Gilbuena, R. (2013). Temporal chenges in the hydrochemical facies of groundwater quality in tow main aquifers in Hanoi, Vietnam. *Environmental Earth Sciences*, 73(8), 4287–4303.

Pearson, K. (1901). On lines and plans of closest fit to systems of points in Space. *Philosophical Magazine*, 2(6), 559-572.

Petersen, W. (2001). Process identification by principal component analysis of river water-quality data. *Ecological Modelling,* 138(1-3), 193-213.

Richards, J.A. & Jia, X. (1999). Remote Sensing Digital Image Analysis, An Introduction, 3rd ed. Berlin, Germany, Springer.

Sánchez-Martos, F., Jiménez-Espinosa, R. & Pulido-Bosch, A. (2001). Mapping groundwater quality variables using PCA and geostatistics: a case study of Bajo Andarax, southeastern Spain. Hydrological sciences journal, 46(2), 227-242.

Sauquet, E. (2000). Mapping mean monthly runoff pattern using EOF analysis. *Hydrology and Earth System Sciences*, 4(1), 79-93.

Selle, B., Schwientek, M., Lischeid, G. (2013). Understanding processes governing water quality in catchments using principal component scores. *Journal of Hydrology,* 486, 31–38.

Sheikhy Narany, T., Ramli, M. F., Aris, A. Z., Sulaiman, W. N. A. and Fakharian, K. (2013). Spatial assessment of groundwater quality monitoring wells using indicator kriging and risk mapping, amol-babol plain Iran.Water, 6(1), 68–85.

Sheikhy Narany, T., Ramli, M. F., Fakharian, K., Aris, A. Z., Sulaiman, W. N. A. (2015). Multi-Objective Based Approach for Groundwater Quality Monitoring Network Optimization. *Water Resources Management*, 29(14): 5141–5156.

Stathis, D. & Myronidis, D. (2009). Principal component analysis of Precipitation in Thessaly Region (Central Greece). *Global NEST Journal,* 11(4), 467-476.

Stetzenbach, K. J., I. M. Farnham, V. F. Hodge., and Johannesson, K. H. (1999). Using multivariate statistical analysis of groundwater major cation and trace element concentrations to evaluate groundwater flow in a regional aquifer. Hydrological Process, 13(17), 2655–2697.

Subba Rao, N. (2014). Spatial control of groundwater contamination, using principal component analysis**.** Journal of Earth System Science, 123(4), 715-728.

Vonberg, D., Vanderborght, J., Cremer, N., Pütz, T., Herbst, M. & Vereecken, H. (2014). 20 years of long-term atrazine monitoring in a shallow aquifer in western Germany. Water Research, 50, 294–306.

Wan, K. L. (2009). A new variable for climate change study and implications for the built environment. *Renewable Energy,* 34(3), 916-919.

Wang, X., Chen, X., Cui, Q. & Yang, Z. (2019). An improved two-step parameter adjustment method for the optimization of a reservoir operation function model based on repeated principal component analysis and a genetic algorithm. Journal of Hydroinformatics, 21(1), 1-12.

Wold, S., K. Esbensen, and Geladi, P. (1987). Principal component analysis. Chemometrics and Intelligent Laboratory System, 2(1-3), 37–52.

Zakhem, B. A. (2017). Using principal component analysis in the investigation of groundwater hydrochemistry of Upper Jezireh Basin, Syria. Hydrological Sciences Journal, 62(14), 1-30.

http://dx.doi.org/10.22453/LSJ-022.1.022-037