J. Sus. Agri. and Env. Bio., 2026; 1(1): 9-15

Journal of Sustainable Agriculture and
Environmental Biology (JSAEB)

https://doi.org/10.57647/jsaeb.2026.0101.02

Research Article

Univariate Distribution Estimator for Predicting the Groundwater Level Fluctuations

Saviour Aletor

Department of Agricultural Economics, University of Uyo, Nigeria

***Corresponding authors:** savioural etro@gmail.com

Article History:

Received:
17 November 2025

Accepted:
12 December 2025

Published in Issue:
31 March 2026

Abstract

Today, with the growth of population and consequently the increasing human need for water resources, lack of rainfall and surface currents in arid and semi-arid regions of the world, the exploitation of groundwater resources has increased and caused many problems in these non-renewable resources. Therefore, it is necessary to study and predict the status of groundwater resources. In this regard, several approaches and methods such as modeling have been used. In this research, an attempt was made to investigate changes in water level using a hydrological model and numerical simulation of groundwater. Univariate frequency distribution functions and MODFLOW simulation model were used to create an operating system. Hydrological events including rainfall and drought were considered to evaluate the return periods and prediction of rate and time of the individual phenomena. The results showed that the 50-year return period is the best scenario for groundwater abstraction for agriculture. Groundwater balancing, artificial recharge and exploitation control are other ways of managing the water level. Rainfall was the main component of the decision system in the summer season and the depth of precipitation was evaluated as critical factor.

Keywords: Groundwater simulation; Sampling method; Univariate analysis; Water exploitation

Cite this article: Saviour A. 2026. Univariate Distribution Estimator for Predicting the Groundwater Level Fluctuations. Journal of Sustainable Agriculture and Environmental Biology 01(01): 9-15. http://doi.org/10.57647/jsaeb.2026.0101.02

1. Introduction

Since most of the world is located in arid and semi-arid regions, and due to low rainfall, mismatch between the time of rainfall and its location in the agricultural sector, which is the main water consumption, leads to a low crisis. Investigation of the trend of hydrological and hydrogeological factors is one of the basic requirements for the management of water area in aquifers and basins (Afshar et al. 2009; Dehghani et al. 2019).

To investigate the long-term effects and consequences of various management decisions on groundwater level, Singh et al. (2012) developed a two-dimensional finite difference simulation model and applied it together with an optimization model to study the problems of salinity and wetland of regional lands in India. El-Kadi et al. (2014) used the MODFLOW numerical model with the aim of achieving sustainable exploitation of groundwater resources in South Korea, under climatic scenarios, cultivation pattern and pumping volume. The results of this study showed that with the maximum amount of extraction with 38% yield and 26 years of drought conditions and reduction of groundwater extraction, the average amount of water level will increase. In the study of Toth et al. (2016) to investigate the interaction of groundwater and related ecosystems, a three-dimensional numerical model was used to simulate the level of groundwater, using the MODFLOW software package. The results of this modeling clarified the pattern of groundwater flow and hydraulic behavior as well as the hydraulic connections between the aquifer and its associated wetland.

Hydrological uncertainty is generally in the form of probabilistic models and leads to the production of quantities of variables with multi-year event estimates. Hence, numerous examples of the application of probabilities in phenomena related to water research topics can be traced in the last two decades (Montanari and Grossi 2008; Zhu et al. 2014; Kong et al. 2018; Lalehzari and Kerachian 2020). For example, probabilistic modeling to take into account hydraulic conduction uncertainty in groundwater resources (Singh and Minsker 2004), flow uncertainty analysis in estimating the average weight of flood damage (Karamouz et al. 2018) and investigation of rainfall uncertainty in the design of urban runoff control system (Sun et al. 2011) has been studied.

Kwon and Moon (2006) examined flood risk (hydrological risk) in the context of probability. In this research, important factors in risk analysis, uncertainty of random parameter values, type of uncertainty analysis method and type of probabilistic distribution of random parameters have been evaluated. Goodarzi et al. (2011) investigated the probability of flooding in southern Iran using frequency analysis of the two-variable logistic gamble distribution and Monte Carlo simulation. In this study, the stochastic variables of flood peak discharge, reservoir initial level and overflow coefficient are overflow. In flood management techniques, the use of probabilistic techniques has also been developed to evaluate structures in the face of destructive events of this phenomenon. The results showed that estimating the return period and developed risk functions can be effective in evaluating flood-related projects (Salas et al. 2019; Maurer et al. 2018; Gao et al. 2019; Rahman and Bowling 2019). In one study, a probabilistic method was used to study the behavior of the Dongnai River in Vietnam. The results showed that the assumption of time stagnation in flood characteristics leads to low flood risk estimation. For example, the estimated flood characteristics for a 50-year return period in the steady state are the same as the flood characteristics for a 10-year return period in a moving state (Shenouda et al. 2018; Dong et al. 2019). Also, flow routing in dam reservoirs, volume and flood discharge as well as optimal design of flood diversion system are among the issues in the field of water management that face hydrological, hydraulic and economic uncertainties (Roughani et al. 2007; Zhang et al. 2018; Zhang et al. 2019; Chen and Hobbs 2020).

Seasonal design floods that reflect seasonally variable information are critical to reservoir implementation and management. The seasonal design flood estimation method currently used in China is based on the frequency analysis of a variable and assumes that the seasonal and annual design frequencies are equal, which neither meets the flood prevention standards nor the interdependence between different seasonal floods (Xiao et al. 2009; Rizwan et al. 2019). A case study for a reservoir in China was created by dividing the entire research period into different subcategories, the dependent structure of summer and autumn floods (Yin et al. 2017). Ahmadisharaf and Kalyanapu (2019) evaluated the uncertainty in estimating flood loss using a combined hydraulic and hydrological probabilistic structure. They developed two separate models for two-dimensional simulation of unstable hydraulic conditions and the rainfall-runoff model. The depth of rainfall in this case is one of the main sources of uncertainty in the issue.

In this study a combined framework has been developed to estimate the groundwater fluctuations based on the different return periods of rainfall predicted by univariate frequency analysis.

2. Material and methods

2.1. Study area

Huai River Basin is located in the east of China, which mainly consists of Huai and Yishusi River system. The flood of Huai River system mainly comes from the upstream of Huai River, Huainan and Funiu mountains area. According to the relevant statistical information, there have been 28 times of floods and droughts disasters since 1974 in Huai River. Hongze Lake is the large lake in Huai River, which have storage function. Under the design conditions, the capacity of reservoir is about 12 billion cubic meters. Therefore, it is essential to control flood, resource utilization and ecological water. Transferring the flood water from Hongze Lake using a diversion tunnel is an advantage project for this area. Fig. 1 shows the main flood hydrograph generated by flood recorder from 1974 to 2019.

2.2. Determination of return periods

In the first step, flood events between 1974 and 2019 were evaluated and flood return periods were calculated using univariate distribution functions. To select an event from among the generated functions requires a selection method which will be explained below.

In this research, the sampling method used as a probabilistic model was as follows.

$$G = \int g(X) f_x(X) dX = E[g(X)] \tag{1}$$

where $X = (X_1, X_2, \ldots\ldots, X_K)^T$ = K-dimension vector of random variables and its density function. The developed sampling technique divides the feasible domain of each variable into an M interval with equal probability as shown in Fig. 2. Then, a random value is selected for

10.57647/jsaeb.2026.0101.02

each interval. As a result, random data will be generated for each variable M and the expected value is estimated as follows:

$$\bar{G} = \frac{1}{M}\sum_{K=1}^{M} g(X_{K1}, X_{K2}, \ldots\ldots, X_{Km}) \qquad (2)$$

where X_{Km} = the provided data for the random variable k in *m*th domain. The point intervals are obtained by:

$$P(X_{km} \leq X_k \leq X_{k,m+1} P\big(X_{km} \leq X_k \leq X_{k,m+1}\big) = \frac{1}{M} \quad m = 0,1,2,\ldots,M-1 \qquad (3)$$

$$X_K = X_{KO} < X_{K1} < X_{K2} < \cdots < X_{K,M-1} < X_{K,M} = \bar{X}_K \qquad (4)$$

where, F_k is a function of the cumulative density of the random variable X_k. In the next step, to obtain an M sample from the X_K variable, a sample is randomly selected from each of the intervals when the endpoints are determined for all intervals. For this purpose, variables with uncertainty must first be identified. For example, in flood deflection system, flood discharge is considered as uncertainty variables. Based on the defined sampling system, the estimation of maximum flood discharge in different return periods using probabilistic models compared to the experimental method. The estimated values of Inverse Gaussian were considered as the predicted amounts of the model.

The Mann-Kendall method is widely used in the analysis of hydrological and meteorological series trends. The Mann-Kendall nonparametric test was first developed by Mann in 1945 and then developed by Kendall in 1975 based on data rankings over a time series.

This method is widely and widely used in the analysis of hydrological and meteorological series trends. One of the strengths of this method is that it is suitable for time series that do not follow a specific statistical distribution. The low effectiveness of this method from the limit values that are observed in some time series is another advantage of using this method. The null hypothesis of this test indicates that the trend in the data series is random and there is no acceptance of hypothesis one (rejection of the null hypothesis) indicates the existence of a trend in the data series. This test was proposed by the World Meteorological Organization in 1988 and has been used to significantly study the trend of climatic series.

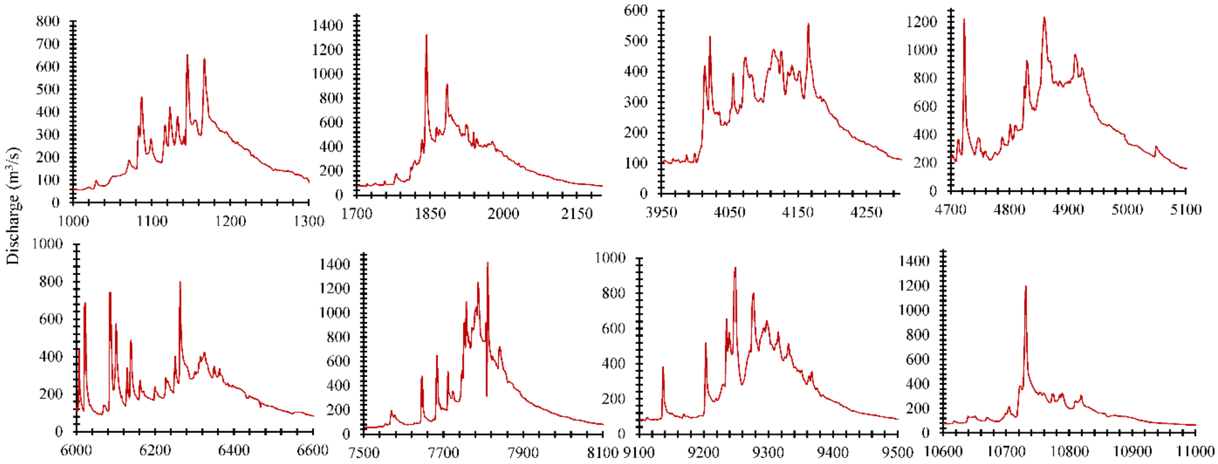

Figure 1. Main flood hydrographs in the study area

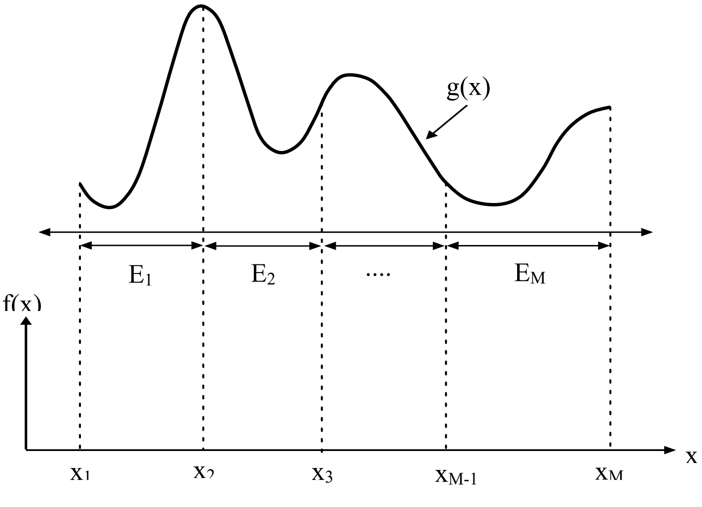

Figure 2. Schematic of the sampling method

Assumption zero in this test indicates the absence of trends in the data series and assumption one means the existence of trends in the data series. In this study, the computational method of Mann-Kendall test was used, which is determined from the following relation of S statistic.

$$S = \sum_{i=1}^{n-i} \sum_{j=i+1}^{n} sgn(x_j - x_i) \qquad (5)$$

where x_i and x_j are the ordered values and *n* is the number of observations in the measurement data of each factor. The sign function can be calculated as follows.

$$sgn(x_j - x_i) = \begin{cases} +1 & \text{if} \quad (x_j - x_i) > 0 \\ 0 & \text{if} \quad (x_j - x_i) = 0 \\ -1 & \text{if} \quad (x_j - x_i) < 0 \end{cases} \qquad (6)$$

Calculating the standardized values of the s-statistic requires calculating the variance, which can be calculated from the following relations of the values of variance and z:

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{n} i(i-1)(2i+5)}{18} \qquad n > 10$$

$$Var(S) = \frac{n(n-1)(2n+5)}{18} \qquad n \le 10 \qquad (8)$$

In the above equations, *i* indicates the frequency of assets. If the value of *Z* is greater than +1.96 or less than -1.96, the data has a trend and the null hypothesis is rejected. Another estimator method is similar to the Mann-Kendall method based on the analysis of differences between observations in a time series. The advantages of this method can be applied to this method in time series that are not subject to any particular distribution. In this method, the geometric slope between the data is calculated and then the median of the slopes is determined from the following equation.

$$\beta = m\left(\frac{X_j - X_i}{j-1}\right) \qquad \forall \qquad j > i \qquad (9)$$

In the mentioned relation, *B* is the estimation of the slope of the trend line and X_i, X_j are the consecutive observational values of *j* and *i*, where the negative values of *B* show the downward trend and the positive values show the upward trend in the series. Another method of calculating the trend in the data is the Petty test, which is to detect the point of failure change in the time series. The maximum values of s are selected and set to k and the value of *p* is determined from the following equation.

$$p = 2e^{-\frac{6k^2}{n^3+n^2}} \qquad (10)$$

Values of *p* are less than 0.05 and indicate a breaking point in the data series. If the data series does not follow the statistical distributions, they use the following test method. This method is based on the sum of the modified S_k^* series, which is calculated in the following equations.

$$s_k^* = 0 \qquad \forall \qquad k = 0 \qquad (11)$$

$$s_k^* = \sum_{i=1}^{k} (x_i - \mu) \quad \forall \quad k = 1, 2, \dots, t \qquad (12)$$

$$Q = \max|s_k^*/\partial| \qquad \forall \qquad 1 \le k < t \qquad (13)$$

In this method, xi is the values of time series, μ is the mean of time series, ∂ is the standard deviation and k is the point of change. From different routing methods, factors such as water table of piezometers, annual flow volume in the river and precipitation values have been evaluated. The water table hydrograph was investigated using MODFLOW software. This model is due to its high capabilities and having different subroutines that can analyze various factors and conditions of the aquifer. The equation governing the flow shows the saturated porous medium under unstable conditions and the heterogeneous medium.

$$K_{xx}\frac{\partial^2 H}{\partial x^2} + K_{yy}\frac{\partial^2 H}{\partial y^2} + K_{zz}\frac{\partial^2 H}{\partial z^2} - R = S_s\frac{\partial h}{\partial t} \qquad (14)$$

where K is the hydraulic conductivity in different directions, R is the input flow to the system, H is the hydraulic head, S_s is the specific yield and t is the time. In this study, the finite difference was used for simulation.

3. Results and discussion

3.1. Frequency analysis

The results section consists of three separate sections, which are: selecting the frequency distribution function fitted to the flood data, calibrating the groundwater model, and finally changing the water table based on the selected return period.

Fig. 3 shows the peak flow rates of the hydrograph in different return periods using different distribution functions. Univariate frequency distribution functions are a good criterion for proving the occurrence of hydrological events. Using this technique, predetermined values of events can be estimated that affect the design of hydraulic structures and their evaluation. Based on the analysis of evaluation criteria of the best distribution function in this study, generalized extreme values were obtained.

The 50-year return period with a flow rate of about 3800 cubic meters per second was selected as the base flow rate to calculate its impact on groundwater. The mechanism of effect of hydrograph on groundwater was calculated by simulating the interaction of surface and groundwater. Therefore, the hydrograph flow will be applied to the MODFLOW model based on surface feeding and will be transferred to groundwater. Therefore, in this section, first, the calibration results of the groundwater model are discussed.

3.2. Return periods

The groundwater model was simulated by calibrating hydraulic conductivity and special discharge. Table 1 shows the calibrated values of hydraulic conductivity and specific discharge for 10 observation wells. According to the results, the hydraulic conductivity varies between 1 and 10 meters per day. These values indicate that the porous medium is composed of silty loam soil. In addition, the specific discharge changes were similar to hydraulic conductivity. The range of its changes has increased or decreased by about 30%. The speed of surface water transfer to groundwater using these values reaches about 11 mm per hour. Therefore, in a surface feeding, it can be expected that a significant volume of flood to reach the peak point of the hydrograph to feed groundwater. Construction of flow collection and control structures in this area can increase the aquifer feeding potential.

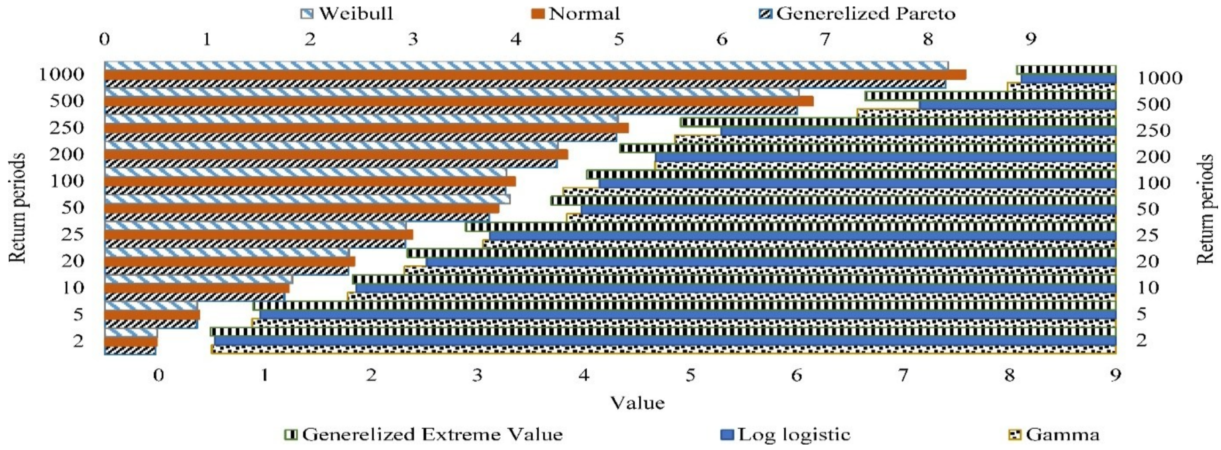

Figure 3. Peak flow (m^3/s) for the inlet hydrograph calculated using different distribution function

Table 1. Results of hydraulic conductivity and specific yield

Observational wells	Hydraulic conductivity					Specific yield (m/day)-				
	K	P%	Min	Max	Difference	S_s	P%	Min	Max	Difference
OW1	2.4	0.95	1.9	3.2	1.3	0.04	0.95	0.030	0.050	0.020
OW2	6.4	0.95	5.1	8.4	3.4	0.10	0.95	0.080	0.134	0.054
OW3	1.4	0.95	1.1	1.8	0.7	0.02	0.95	0.018	0.029	0.012
OW4	3.5	0.95	2.8	4.6	1.9	0.06	0.95	0.044	0.073	0.029
OW5	5.4	0.95	4.3	7.1	2.9	0.09	0.95	0.068	0.113	0.045
OW6	4.5	0.95	3.6	5.9	2.4	0.07	0.95	0.056	0.094	0.038
OW7	1.3	0.95	1.0	1.7	0.7	0.02	0.95	0.016	0.027	0.011
OW8	9.5	0.95	7.5	12.5	5.0	0.15	0.95	0.119	0.199	0.080
OW9	7.3	0.95	5.8	9.6	3.9	0.12	0.95	0.092	0.153	0.061

K = Hydraulic Conductivity; S_s = Specific Yield; *P%* = Probability Percentage

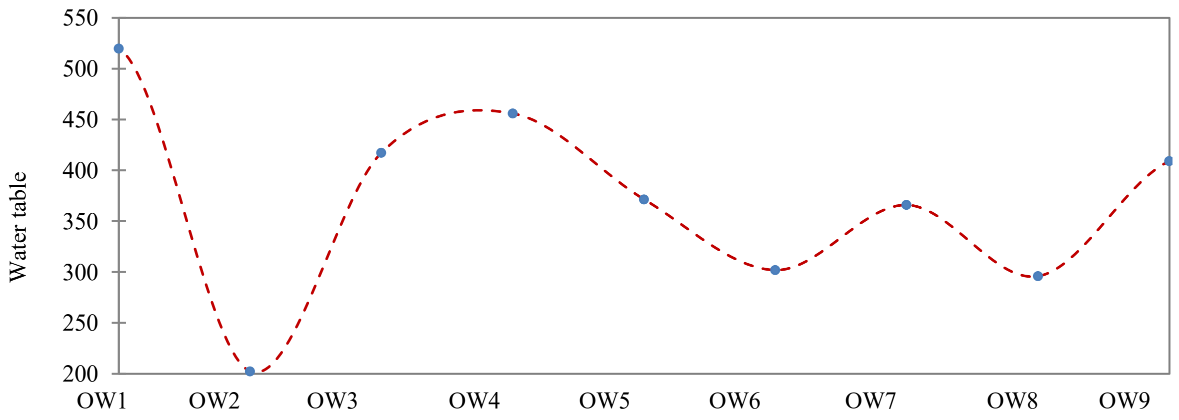

Figure 4. Groundwater level fluctuations in different observational wells

10.57647/jsaeb.2026.0101.02

3.3. Groundwater levels

Water table changes due to aquifer feeding with a return period of 50 years are shown in Fig. 4. Flood feeding to groundwater can increase the water table level in different observation wells by 80 to 230 cm. Well one had the highest increase in height and well number two had the lowest increase in height. The results showed that the 50-year return period is the best scale for aquifer feeding due to soil hydraulic conditions. In addition, the application of a higher return period causes water loss as surface runoff and the flow control structure is not economically viable. In the lower returns period, flood management will be easier and the aquifer will receive all the hydrograph water, but it can not be considered as a sustainable flood management option. Creating structures for such flows is not cost effective.

4. Conclusion

In this study, a three interconnected structure is planned to predict, control and infiltrate floodwater. It should be noted that the intended flood flow is taken into account after agricultural use and environmental needs. The frequency distribution function of generalized limit values was obtained by performing analysis as the best hydrograph prediction criterion. The predicted peak flow using this method for the 50-year return period was estimated at about 3800 cubic meters per second. This flow has increased the height of the water table to an average of about 120 cm. The use of this method is important for areas without statistics and with agricultural uses and can be recommended.

Authors Contribution
All the authors have participated sufficiently in the intellectual content, conception, and design of this work or the analysis and interpretation of the data (when applicable), as well as the writing of the manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable. request.

Conflict of interest
The author states that there is no conflict of interest.

References

Afshar, A., Rasekh, A., and Afshar, M.H., (2009) Risk-based optimization of large flood diversion systems, using genetic algorithms, Journal of Engineering Optimization, 41(3):259–273.

Ahmadisharaf E, Kalyanapu AJ (2019) A coupled probabilistic hydrologic and hydraulic modelling framework to investigate the uncertainty of flood loss estimates. J Flood Risk Management 12: 1-12.

Chen, L., & Hobbs, B. F. (2021). Robust Yellow River Delta flood management under uncertainty. Water, 13(16), 2226.

Dehghani, M. Saghafian, B. and Zargar, M (2019). Probabilistic hydrological drought index forecasting based on meteorological drought index using Archimedean copulas. Hydrology Research. 50: 1230-1251.

Dong, S., Wang, H., Guo, X., & Zhou, Z (2021). Characteristics of water hazards in China's coal mines: a review. Mine Water and the Environment, 40(2): 325-333.

El-Kadi AI, Tillery S, Whittier RB, Hagedorn B, Mair A, Ha K, & Koh G. W 2014. Assessing sustainability of groundwater resources on Jeju Island, South Korea, under climate change, drought, and increased usage. Hydrogeology Journal 22(3): 625-642.

Goodarzi, S., Settari, A., Zoback, M., & Keith, D (2011). A coupled geomechanical reservoir simulation analysis of carbon dioxide storage in a saline aquifer in the Ohio River Valley. Environmental Geosciences, 18(3): 189-207.

Gao, Y., Qian, H., Ren, W., Wang, H., Liu, F., & Yang, F. (2020). Hydrogeochemical characterization and quality assessment of groundwater based on integrated-weight water quality index in a concentrated urban area. Journal of cleaner production, 260, 121006.

Karamouz M, Doroudi S, Moridi A (2018) Developing a model for optimizing the geometric characteristics of water diversion systems. J Irrig Drain Eng 144(2): 04017062

Kong XM, Huang GH, Li YP, Fan YR, Zeng XT, Zhu Y (2018) Inexact copula-based stochastic programming method for water resources management under multiple uncertainties. J Water Resour Plann Manage 144(11): 04018069.

Kwon, H. H., & Moon, Y. I (2006). Improvement of overtopping risk evaluations using probabilistic concepts for existing dams. Stochastic Environmental Research and Risk Assessment, 20(4): 223-237.

Lalehzari, R. and Kerachian, R (2020) Developing a Framework for Daily Common Pool Groundwater Allocation to Demands in Agricultural Regions. Agricultural Water Management.

Maurer EP, Kayser G, Doyle L, Wood AW (2018) Adjusting flood peak frequency changes to account for climate change impacts in the western United States. J Water Resour Plann Manage 144(3): 05017025.

Montanari, A., and Grossi, G., (2008) Estimating the uncertainty of hydrological forecasts: A statistical approach, Journal of Water Resources Research, Volume 44, Issue 12, W00B08, http://doi.org/10.1029/2008WR006897

Rahman S, Bowling L (2019) Streamflow impacts of management and environmental change in the Upper Wabash River basin. J Hydrol Eng 24(3): 05018034.

Rizwan M, Guo S, Yin J, Feng X (2019) Deriving design flood hydrographs based on copula function: A case study in Pakistan. Water 11(1531): 1-18.

Roughani, M., Ghafouri, M., & Tabatabaei, M (2007). An innovative methodology for the prioritization of sub-catchments for flood control. International Journal of Applied Earth Observation and Geoinformation, 9(1): 79-87.

Salas JD, Obeysekera J (2019) Probability distribution and risk of the first occurrence of k extreme hydrologic events. J Hydrol Eng 24(10): 04019032.

Singh, A., Nath Panda, S., Flugel, W. A., & Krause, P. (2012). Waterlogging and farmland salinisation: causes and remedial measures in an irrigated semi-arid region of India. Irrigation and drainage, 61(3), 357-365.

Shenouda, I.M., Fleifle, A.E., Awad, H.M. and Younan, N.A. (2018). Simulation - Optimization Model for the Hydraulic and Structural Design of Barrages Regulatorsin Egypt. Journal of Irrigation and Drainage Engineering. 04018034.

Sun, G., Caldwell, P., Noormets, A., McNulty, S. G., Cohen, E., Moore Myers, J., ... & Chen, J (2011). Upscaling key ecosystem functions across the conterminous United States by a water-centric ecosystem model. Journal of Geophysical Research: Biogeosciences, 116(G3).

Tóth Á., Havril T, Simon S, Galsa A, Santos F. A. M, Müller I, & Mádl-Szőnyi J (2016). Groundwater flow pattern and related environmental phenomena in complex geologic setting based on integrated model construction. Journal of Hydrology 539:330-344.

Xiao Y, Guo SL, Liu P, Yan B, Chen L (2009) Design flood hydrograph based on multi characteristic synthesis index method. J Hydrol Eng 14(12): 1359–1364.

Yin J, Guo S, Liu Z, Chen K, Chang F, Xiong F (2017) Bivariate seasonal design flood estimation based on copulas. J Hydrol Eng 22(12): 05017028.

Zhang C, Ding W, Ming F, Fu G (2019) Cost-benefit framework design of water transfer systems. J Water Resour Plann Manage 145(5): 04019007.

Zhang X, Liu P, Xu CY, Ming B, Xie A, Feng M (2018) Conditional value-at-risk for nonstationary streamflow and its application for derivation of the adaptive reservoir flood limited water level. J Water Resour Plann Manage 144(3): 04018005.

Zhu X, Zhang C, Yin J, Zhou H, Jiang Y (2014) Optimization of water diversion based on reservoir operating rules: analysis of the Biliu River reservoir, China. J Hydrol Eng 19: 411-421.