

Groundwater Nitrate Estimation Using Geostatistical Methods and Artificial Intelligence Models

Asad Davarpanah¹✉

¹ Agricultural Bank of Iran, Information Technology Department, Zahedan, Iran

Corresponding Author, email: (✉ asaddavarpanah81@gmail.com)

Received: 10 April 2025

Accepted: 03 March 2025

Available online: 16 March 2025

Abstract: Increased water demands, limitations of surface water resources, and excessive groundwater extraction have caused irreparable damage to the environment and natural resources. One of the major water quality issues in agriculture is nitrate toxicity. The Mashhad plain aquifer is one of the most important aquifers for water supply in Razavi Khorasan Province, Iran. This study aimed to investigate nitrate variations with some groundwater quality characteristics such as sodium, chloride, and electrical conductivity using geostatistical methods and to determine groundwater pollution trends in the Mashhad plain. Additionally, artificial intelligence models were employed to estimate nitrate concentrations. The latest nitrate measurements from 246 wells during 2006-2008 formed the basis for spatial and temporal analysis. Cross-validation techniques with evaluation criteria including Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were used to compare methods. Results showed that nitrate concentration exhibited moderate spatial correlation in the plain. Cross-validation results indicated that among the methods used, log-normal kriging with minimum RMSE and MAE had higher accuracy, although the difference with ordinary kriging was negligible. Indicator kriging had the lowest accuracy in estimating groundwater nitrate levels. In simulating nitrate levels using artificial intelligence models with sodium, chloride, and electrical conductivity as inputs, neural networks demonstrated superiority over neuro-fuzzy systems.

Keywords: Mashhad plain, Kriging, Artificial Intelligence, Groundwater quality, Nitrate contamination

1. Introduction

The issue of declining water quality has compelled many countries worldwide, including Iran, to review and revise their perspectives on integrated water resource management. Groundwater constitutes a significant portion of the world's renewable water resources. Global water resource calculations indicate that groundwater accounts for approximately 0.6% of total water resources and 60% of accessible renewable water resources. Optimal utilization of existing water resources requires comprehensive and informed planning, which is only possible through accurate knowledge of water quantity and quality. In developing and third-world countries, most attention has been focused on finding suitable groundwater aquifers to meet drinking and agricultural water needs, while less attention has been paid to maintaining aquifer quality (Kangaroglu and Gunay, 1997).

Groundwater nitrate (NO_3^-) pollution has emerged as a critical environmental concern globally, driven primarily by intensive agricultural practices, livestock waste, and inadequate wastewater management (Islam et al., 2025). Recent studies have shown that nitrate contamination poses significant health risks even at levels far below regulatory limits, with concentrations as low as 2 mg/L $\text{NO}_3\text{-N}$ being associated with adverse health outcomes including cancer and birth defects (Hamlin et al., 2022). The Mashhad plain, one of the most critical plains in Razavi Khorasan Province, Iran, contains over 5,000 wells, making it the densest plain in

the province in terms of water extraction wells. From 1991 to 2011, the groundwater level in the Mashhad plain decreased by approximately 14 meters, and if current trends continue, it is projected that the groundwater level will drop by an additional 14.5 meters in the next 10 years. This means that the level of decline that occurred over 20 years could be compressed into just 10 years at the current rate. In addition to water level decline, water quality in the region has also deteriorated over these years. Currently, 90% of Mashhad city's water supply comes from deep wells, highlighting the importance of quality assessment of this water source.

Since classical statistics cannot account for the spatial distribution of groundwater quality and quantity factors, geostatistics is used as a technique for this purpose (Madani, 1994). Interpolation of groundwater nitrate with limited measured data requires the application of various interpolation methods. Kriging is a weighted average estimator in which weights are determined based on a model fitted to a function such as a semi-variogram that expresses spatial variations of the variable. Ordinary kriging, also known as the Best Linear Unbiased Estimator, is the most common type of kriging (Journel and Huijbergts, 1978). Recent advances in geostatistical methods have introduced sophisticated techniques such as extension variance combined with Voronoi tessellation regionalization, which provides more reliable probability estimates for threshold exceedance compared to traditional interpolation methods (Hydrogeology Journal, 2023). Furthermore,

the integration of geostatistics with network connectivity and spatial random forests has shown promising results in environmental contaminant mapping (Environmental Modelling & Software, 2025).

In Iran, numerous studies have used semi-variograms and ordinary kriging estimators to investigate spatial variations and interpolate groundwater quality characteristics in various fields including soil science (Delbari 2010, Delbari et al. 2012, Heshmati and Bigi Herchegani 2012, Delbari and Afrasiab 2014). However, using ordinary kriging for variables with highly skewed frequency distributions is not recommended. On the other hand, many characteristics in geosciences do not follow normal distribution due to a few very large or small values among them. These small or large values may affect the statistical summary of the variable such as mean, variance, or measures of spatial correlation (such as semi-variogram). Various methods exist to eliminate or adjust the adverse effects of these values (Goovaerts, 1997). Recent comparative studies have evaluated multiple geostatistical methods including Inverse Distance Weighting (IDW), spline, natural neighbor, and ordinary kriging for nitrate estimation, with findings indicating that the choice of method significantly impacts prediction accuracy (Water Journal, 2023).

The first approach is to remove outliers from the dataset, provided that their incorrectness has been demonstrated. The second approach is to study this category of numbers under a separate statistical population, provided that data density allows extraction of reliable statistics for each subset. Another method is to transform data using an appropriate transformation function to numbers with normal or near-normal distribution. Among the most common transformations is logarithmic transformation used for data with positively skewed histograms (Saito and Goovaerts, 2000). Nevertheless, the most general solution is using indicator kriging, which, unlike ordinary kriging, makes no assumptions about the normality of data distribution and is essentially considered the non-parametric counterpart of ordinary kriging (Journel, 1983). Thus, indicator kriging can overcome both the parametric nature and smoothing effect issues of ordinary kriging (Journel, 1983).

Kresic (1997) introduced geostatistical methods as the best and most powerful tools for data interpolation and preparation of groundwater level maps. Geostatistical methods have been used for zoning and modeling many regional variables including groundwater depth (Mariño, 1998; McKenna, 2002; Desbarats et al., 2002; Kelin et al., 2005; Koike and Moukanaa, 2008; Bhat, 2012) and soil chemical properties and groundwater quality (Yang et al., 1998; Lee et al., 2008; Hakan, 2012; Delbari, 2013; Delbari, 2014). Jang (2005) used geostatistical methods to estimate nitrate pollution and other water chemical

properties in the Choushui River in Taiwan. Sarukkalige (2012) investigated groundwater quality in western Australia using geostatistical techniques, fitting four kriging variogram models including spherical, circular, exponential, and Gaussian models to the study area data and performing parameter interpolation using these models. According to this study, the kriging method with exponential variogram model provides the best results for interpolating groundwater quality parameters. Delbari et al. (2012) investigated spatial variation and zoning of groundwater quality including SAR and EC using geostatistical methods, with the ultimate goal of predicting variability in water and soil infiltration rates in the study area of Kerman Province considering groundwater salinity and sodium conditions.

Regarding artificial intelligence in groundwater quality modeling, Kuo et al. (2005) simulated arsenic levels in groundwater in the Black Foot area of Taiwan using neural network models. Similarly, Nasr and Zahran (2014) investigated groundwater salinity based on TDS and pH parameters in the New Burg El-Arab plain in Egypt using artificial neural networks, with results indicating good modeling performance. The application of machine learning techniques for groundwater quality prediction has seen remarkable advancement in recent years. Studies have demonstrated that ensemble machine learning models, particularly Random Forest (RF), XGBoost, and LightGBM, combined with explainable AI techniques, can effectively predict groundwater nitrate contamination and identify key influencing factors (Geocarto International, 2025). Furthermore, hybrid models combining artificial neural networks (ANN) with adaptive neuro-fuzzy inference systems (ANFIS) have shown exceptional accuracy in water quality index prediction, with some studies reporting R^2 values exceeding 0.95 (Water Journal, 2022; Scientific Reports, 2024).

Recent systematic reviews have highlighted the growing preference for deep learning architectures such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) for groundwater level and quality prediction, although traditional feedforward neural networks and ANFIS models continue to demonstrate competitive performance (Discover Applied Sciences, 2025; Frontiers in Environmental Science, 2024). This research, considering various interpolation methods and the importance of their results in water sciences, as well as the fact that successful plant production requires appropriate water quality, has led to a comparison of interpolation methods for some important quality characteristics of Mashhad plain groundwater and zoning of water quality characteristics using the best obtained method. Finally, given the importance of nitrate levels in groundwater, simulation was performed using artificial intelligence (ANN and ANFIS).

2. Materials and Methods

The study area in this research is Mashhad Plain located in Razavi Khorasan Province, Iran. Mashhad Plain is an extensive valley with dimensions exceeding 100 km in length, an average width of approximately 25 km, and an area of about 2,500 km² (Ghahreman et al., 2003). This plain is situated in northeastern Iran with geographical coordinates of 58°20' to 60°08' East longitude and 35°40' to 36°03' North latitude, within the Kashafrud River basin in northern Razavi Khorasan Province (Figure 1).

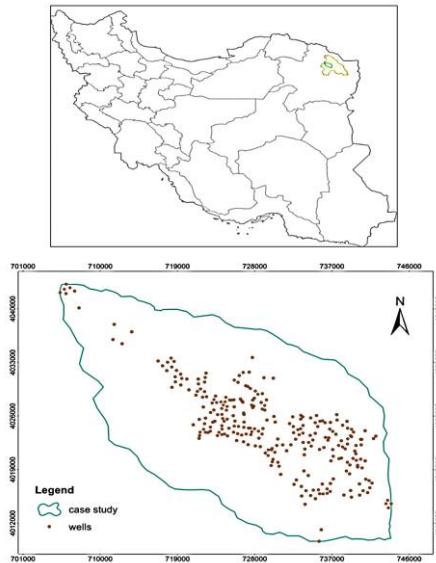


Figure 1. Geographical location map of Mashhad Plain study area

Soil stratification studies were compiled by the Water Engineering Consulting Company in 2008. In this research, groundwater quality data including nitrate (NO₃), chloride (Cl), sodium (Na), and electrical conductivity (EC), along with well locations in UTM coordinates, were collected on a monthly basis for 252 observation wells during 2006-2008 from the Mashhad Water and Wastewater Authority.

2.1. Statistical Analysis of Data

The shape of data frequency distribution is of paramount importance in estimation procedures. Therefore, in this research, statistical indices including mean, minimum, maximum, standard deviation, coefficient of variation, skewness coefficient, and frequency distribution of the parameters under study were initially determined. SPSS software was utilized for statistical analysis of the data.

To investigate the spatial variations of groundwater nitrate data in Mashhad Plain, the experimental semivariogram was calculated and the most appropriate theoretical semivariogram was fitted to it. The semivariogram formula is expressed as Equation (1):

$$\gamma(h) = (1/2N(h)) \sum [Z(x_i) - Z(x_i + h)]^2 \quad (1)$$

Where $\gamma(h)$ = experimental semivariogram, N = number of sample pairs at each lag distance h , $Z(x_i)$ = variable

value at point x_i , and $Z(x_i + h)$ = variable value at a point at distance h from x_i . The semivariogram is essentially a tool that demonstrates the distance-related variations or the variability structure of a particular variable (Mahdian, 2006).

2.2. Geostatistical Analysis

Variations in earth properties exhibit spatial correlation. These variations can be examined primarily from the perspective of variance. Fundamentally, the variance of a variable between points at distance h from each other can express the mutual correlation of values at two points at a specified distance. In the presence of spatial structure, it is natural that the dependence of variables at nearby points is greater than at distant points. Such variance can serve as a criterion for representing the influence or susceptibility of variable values on adjacent points in the environment. This property, which can be utilized in earth sciences and particularly soil science, is termed continuity (Kaidani, 2009).

The spatial correlation criterion is determined based on the ratio $C/(C_0+C)$, where C_0 is the nugget effect and (C_0+C) is the sill. If this ratio exceeds 75%, the data demonstrate strong spatial correlation. When this ratio falls between 25% and 75%, the data exhibit moderate spatial correlation, and if below 25%, the data show weak spatial correlation (Kaidani, 2009).

Among the interpolation methods employed, Ordinary Kriging and Cokriging were utilized for interpolating groundwater depth data in the region. Ordinary Kriging is one of the most important and commonly used geostatistical methods. This technique performs spatial interpolation using semivariograms and fitted models. The general formula for this method is defined as:

$$Z^*(x_i) = \sum \lambda_i Z(x_i) \quad (2)$$

Where $Z^*(x_i)$ is the estimated variable value, λ_i is the weight or importance of sample i , and $Z(x_i)$ is the observed variable value.

2.3. Artificial Intelligence Methods

MATLAB 2014 software was employed for artificial intelligence methods implementation.

2.3.1. Artificial Neural Network (ANN)

In this research, a feedforward neural network with three input layers (electrical conductivity, chloride, and sodium), one hidden layer, and one output layer (nitrate) was utilized. The transfer functions employed for training included Gaussian, Sigmoid, and Spherical functions, with the Levenberg-Marquardt algorithm as the training algorithm (Piri et al., 2009). The dataset comprising 746 monthly well quality measurements was divided into training (60%), testing (20%), and validation (20%) subsets.

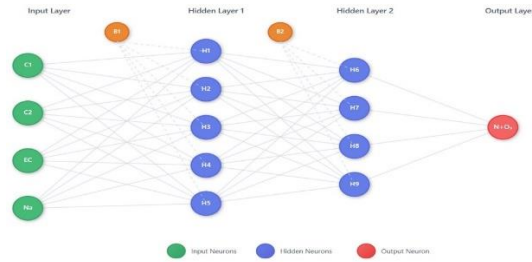


Figure 2. Architecture of the artificial neural network model used in this study

2.3.2. Adaptive Neuro-Fuzzy Inference System (ANFIS)

For designing the neuro-fuzzy network model in this research, a multi-layer feedforward neural network, Levenberg-Marquardt algorithm, Sugeno fuzzy inference system, and linear output function were employed. The weighted average function was used for defuzzification. This network contained 3 membership functions, with 60% of data allocated for training, 20% for testing, and 20% for validation (Piri and Kisi, 2015).

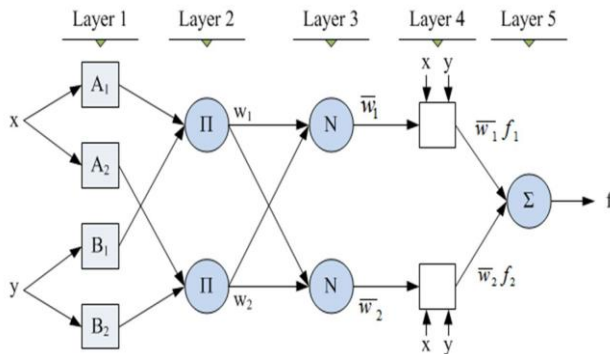


Figure 3. Structure of the ANFIS model used in this study

2.3.3. Evaluation Methods and Criteria

The evaluation method employed for geostatistical approaches was cross-validation. For comparing simulation and prediction results from artificial intelligence and geostatistical models, Root Mean Square Error (RMSE), Mean Bias Error (MBE), and coefficient of determination (R^2) were utilized:

$$RMSE = \sqrt{\frac{1}{n} \sum (Z^0(x_i) - Z(x_i))^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum |Z^0(x_i) - Z(x_i)| \quad (4)$$

Where n is the number of points, $Z^0(x_i)$ is the estimated value, and $Z(x_i)$ is the actual value of variable Z at the given point.

3. Results

Prior to applying various geostatistical models to the data, it is essential to ensure data normality or transform the data to approximate a normal distribution. The Kolmogorov-Smirnov (K-S) test, a statistical method for assessing data normality, was employed in this research. In cases where data did not follow a normal distribution, various methods such as log-normal transformation and Cox-Box transformation were applied to achieve normality.

Table 1. Statistical analysis of study data

Transfo rm	Io n	Mi n	Ma x	Mea n	Skewn ess	Kurto sis	Std. Dev.	Medi an
Raw	N O ₃	4.6	218 .8	41.6	1.78	3.31	38.7	27.7
Log		2.4 5	5.4 2	3.63	0.45	-0.66	0.69	1.44
Raw	EC	31 9	373 5	914. 97	2.14	6.10	504. 39	771
Log		5.7 9	8.2 3	6.72	0.73	0.23	0.45	2.88
Raw	Na	16. 7	495 .1	98.3 4	2.19	5.78	71.9	81.3
Log		3.1 7	6.2 2	4.29	0.43	0.2	0.57	1.91
Raw	Cl	13. 5	430 .9	70.7 9	2.01	5.86	58.3 5	51.8
Log		3.0 2	6.0 8	4.14	0.49	-0.6	0.64	1.72

As shown in Table 1, the NO_3 and Na variables in the study area approximately followed a normal distribution; however, logarithmic transformation was applied for greater accuracy and computational precision. The Cl and EC data did not follow a normal distribution and required normalization procedures.

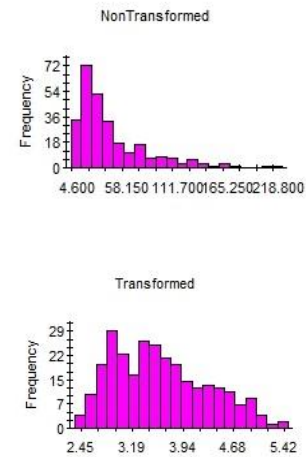


Figure 4. Histograms of NO_3 data: (a) raw data, (b) log-transformed data

3.1. Investigation of Isotropy and Anisotropy of Quality Data

To examine the anisotropy of groundwater quality data, the surface semivariogram and directional semivariograms for four directions (0° , 45° , 90° , and 135°) with an angular tolerance of 22.5° were plotted using GS+ software.

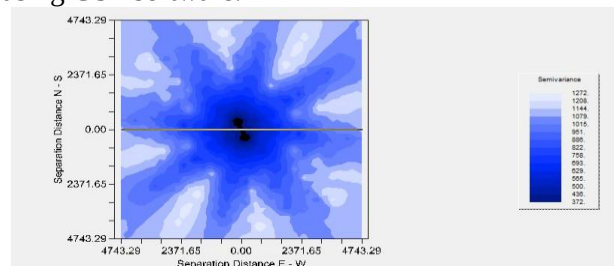


Figure 5. Surface semivariogram of nitrate data in the study area

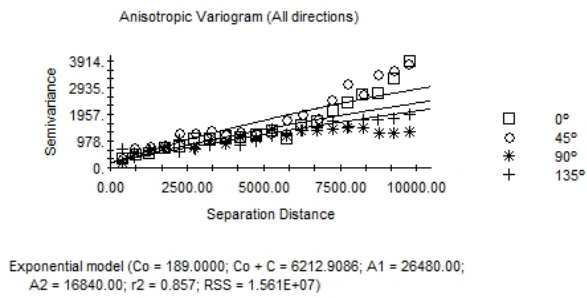


Figure 6. Directional semivariograms of nitrate data in the study area

As evidenced by the figures, the NO_3 data under investigation exhibit anisotropy. Given the critical role of lag size selection in accurately representing the spatial structure of a variable in an environment, and considering that this distance should clearly identify the range of influence while maintaining an optimal number of point pairs for estimation, the lag size selected for groundwater nitrate variations in the region was 500 m (minimum distance between points) with an active lag distance of 5,000 m (10 times the lag).

3.2. Selection of the Best Fitted Model for Experimental Semivariogram

The characteristics of experimental semivariograms for groundwater nitrate concentration in both raw and log-transformed states are presented in Table 2. Among the spherical, exponential, and Gaussian models fitted, the spherical model was selected as the best theoretical model for the estimation process due to having minimum RSS and maximum R^2 .

Table 2. Characteristics of fitted models to experimental semivariograms of nitrate (mg/L)

Model	Nugget (C_0)	Sill ($C_0 + C$)	Range (m)	$C/(C_0 + C)$	R^2	RSS
Spherical	322	1088	4270	0.704	0.987	7179
Linear	450	1201.19	4753	0.625	0.924	107600
Gaussian	425	1087	2060	0.609	0.983	8885

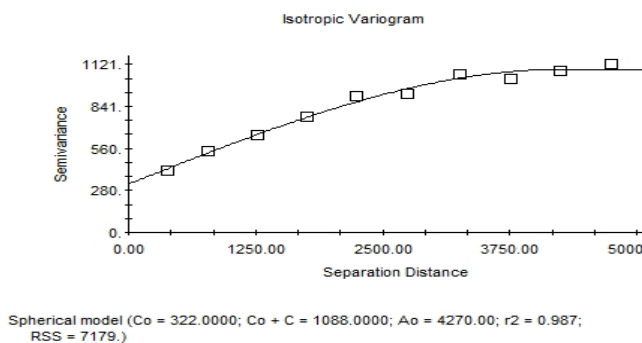


Figure 7. Experimental semivariogram and best fitted model for groundwater nitrate

As Figure 7 demonstrates, the spatial correlation criterion $C/(C_0 + C)$ for measured groundwater nitrate data, with semivariogram characteristics including a nugget effect of 322 and range of influence of 4,270 m, equals approximately 70%. These results indicate moderate spatial correlation for groundwater nitrate.

3.3. Selection of the Best Estimation Method for Quality Data

The results of cross-validation for estimating groundwater nitrate concentration demonstrate that the Cokriging method exhibits higher accuracy in nitrate estimation among the methods evaluated, as it possesses the lowest error values. The calculated RMSE, MAE, and R^2 values presented in Table 3 indicate the optimal interpolation method for spatial mapping of NO_3 quality parameters.

Table 3. RMSE, MAE, and R^2 values of geostatistical methods for NO_3

Model	RMSE	MAE	R^2
Ordinary Kriging	0.1319	4.6235	0.8994
Indicator Kriging	0.0865	3.8402	0.9012
Cokriging	0.0766	2.8566	0.9514

3.4. Spatial Distribution Maps

Risk assessment maps for nitrate data in the plain were generated based on the optimal method using ArcGIS version 10 software.

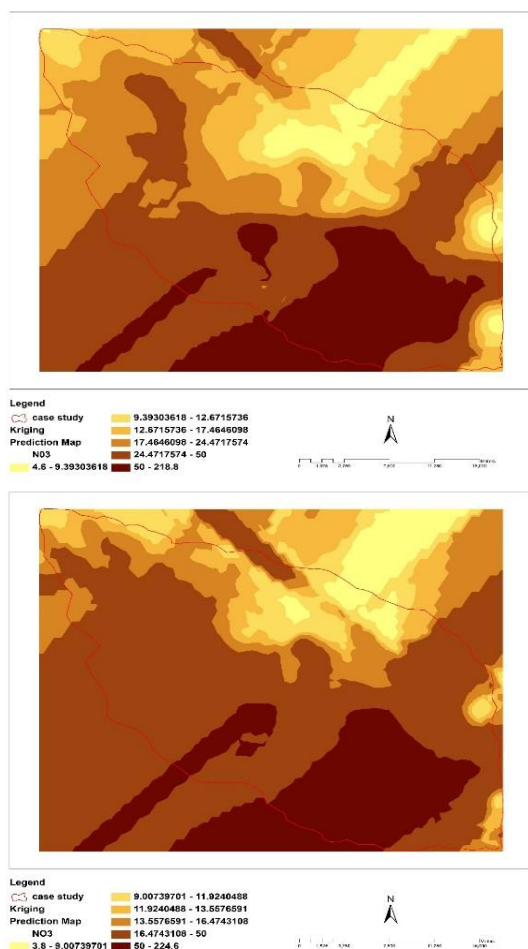


Figure 8. Spatial distribution of NO₃ using Cokriging method: (a) 2006, (b) 2008

According to Iranian Standard 1053, the nitrate threshold for drinking water is 50 mg/L. As shown in Figure 8, in certain areas of the plain, particularly in the southern regions spatially and from wet to dry seasons temporally, nitrate concentrations exhibited an increasing trend, likely attributable to infiltration of urban wastewater from Mashhad, waste disposal, and chemical fertilizer application (Latif et al., 2005; Shirvani et al., 2013).

3.5. Indicator Kriging

The indicator kriging equation system is similar to ordinary kriging, with the distinction that indicator kriging utilizes indicator semivariograms instead of conventional semivariograms. This method can be employed to determine the desired nitrate concentration along with its probability of occurrence. Furthermore, spatial distribution maps showing the probability of nitrate exceeding specified thresholds can be readily generated (HasaniPak, 2003).

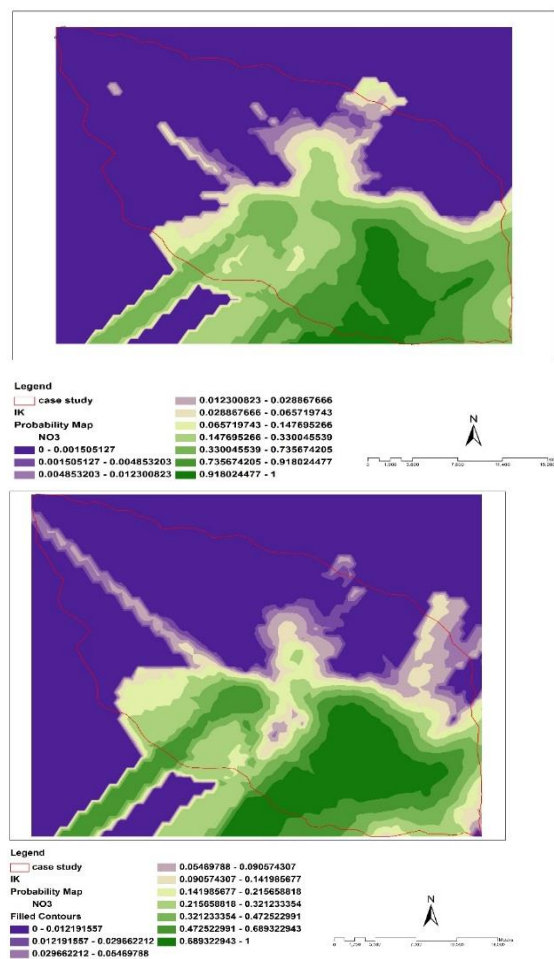


Figure 9. Spatial distribution of NO₃ using Indicator Kriging method: (a) 2006, (b) 2008

According to Figure 9, in certain areas of the plain, particularly in the southern regions spatially and from wet to dry seasons temporally, nitrate concentrations exhibited an increasing trend.

3.6. Artificial Intelligence Results

Nitrate concentrations were simulated using neural network and neuro-fuzzy models with chloride, sodium, and electrical conductivity as input parameters. The results indicate the superiority of neural networks over neuro-fuzzy models. The architecture was implemented in MATLAB 2014 with three inputs (sodium, chloride, and electrical conductivity) and 6 hidden neurons in the hidden layer.

Table 4. Statistical parameters for training, testing, and validation of ANN and ANFIS models for NO₃ estimation (mg/L)

Model	Input	Function	Structure	Train R MSE	Train MAE	Train R ²	Test R MSE	Test MAE	Test R ²
AN FIS	EC, C, I, Na	Gbel lmf	3-9-1	22.57	16.57	0.67	20.06	13.85	0.68
AN FIS	EC, C, I, Na	Gaus smf	3-9-1	24.81	18.36	0.61	20.77	14.81	0.66

Model	Input	Function	Structure	Train RMSE	Train MAE	Train R ²	Test RMSE	Test MAE	Test R ²
AN FIS	EC, Cl, Na	Psigmf	3-9-1	21.82	16.37	0.70	18.63	13.93	0.73
AN N	EC, Cl, Na	Spherical	3-6-1	20.26	14.75	0.74	19.20	13.69	0.71
AN N	EC, Cl, Na	Gaussian	3-6-1	20.05	14.17	0.74	17.96	12.50	0.74
AN N	EC, Cl, Na	Sigmoid	3-6-1	19.29	13.52	0.76	22.17	13.37	0.63

The artificial intelligence results presented in Table 4 demonstrate the superiority of neural networks. The simulation of nitrate parameters in agricultural wells of Mashhad Plain in Razavi Khorasan was investigated. The results indicate moderate performance in the validation phase for this data. A comparison of the two models based on R², RMSE, and MAE reveals the superiority of neural networks over neuro-fuzzy systems. Among the neural network models, the network with sigmoid transfer function achieved the best results in nitrate simulation. The ANN model with sigmoid function achieved validation results of RMSE = 23.96 mg/L, MAE = 14.85 mg/L, and R² = 0.67, representing the optimal configuration for nitrate estimation in this study.

4. Conclusions

Statistical analysis of groundwater nitrate data in the study area indicated positive skewness of data. Geostatistical analysis of groundwater nitrate levels in Mashhad plain showed that nitrate levels have moderate spatial correlation. Results showed that the nitrate semi-variogram structure follows a spherical model (with minimum RSS and maximum R²) with a range of influence of about 5,000 meters. Cross-validation results showed that among the interpolation methods used, ordinary kriging and co-kriging methods with sodium as auxiliary variable had similar accuracy (RMSE) in estimating groundwater nitrate levels.

The nitrate estimation map showed plain contamination in southern parts of the region, most likely due to sewage and use of chemical fertilizers. Results showed that the indicator kriging method with the capability to provide probability maps of exceeding a certain threshold for any data frequency distribution is applicable, and especially when the research objective is more to determine the credibility of estimation rather than the estimation itself, it should be preferred over ordinary kriging. However, this method is more complex than ordinary kriging and requires more precision and time. Therefore, when only obtaining a general estimate of variations in a desired parameter is needed, ordinary kriging can still be used as a quick and easy method.

Results from nitrate level simulation using artificial intelligence models showed moderate accuracy of these models in estimating nitrate levels, and among them, the neural network with structure of 3 inputs, 1 hidden layer with 6 hidden neurons, one output layer, and sigmoid function had better estimation in terms of RMSE, MAE, and R² equal to 23.96, 14.85, and 0.67, respectively. These findings contribute to the growing body of evidence supporting the integration of geostatistical and machine learning approaches for comprehensive groundwater quality assessment and management, particularly in regions experiencing intensive agricultural activities and urban expansion.

References

- Bhat, S., Motz, L., Pathka, C., & Kuebler, L. (2012). Designing Groundwater Level Monitoring Network Using Geostatistical A Case Study and Central Florida U.S.A. World Environmental and Water Resources Congress, 48-58.
- Delbari, M. (2007). Estimation and stochastic simulation of soil properties for case studies in Lower Austria and Sistan plain, southeast of Iran. Doctorate thesis. University of Natural Resources and Applied Life Sciences, Vienna, Austria.
- Desbarats, A.J., Logan, C.E., Hinton, M.J., & Sharpe, D.R. (2002). On the kriging of water table elevations using collateral information from a digital elevation model. *Journal of Hydrology*, 255, 25-38.
- Elmotawakkil, A., Enneya, N., Bhagat, S.K., Ouda, M.M., & Kumar, V. (2025). Advanced machine learning models for robust prediction of water quality index and classification. *Journal of Hydroinformatics*, 27(2), 299-319.
- Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Hamlin, K.E., et al. (2022). Examining Relationships Between Groundwater Nitrate Concentrations in Drinking Water and Landscape Characteristics to Understand Health Risks. *GeoHealth*, 6(5), e2021GH000524.
- HasaniPak, A.A. (2003). *Geostatistics*. Tehran University Press, 325p.
- Islam, A., Sarkar, B., Shit, P.K., & Gazi, H.A.R. (2025). Groundwater nitrate contamination in agroecosystems: A comprehensive review of detection methods, pollution mechanisms, and policy implications. *Progress in Physical Geography*, doi:10.1177/03091333251337298.
- Jia, H., & Qian, H. (2025). Groundwater nitrate response to hydrogeological conditions and socioeconomic load in an agriculture dominated area. *Scientific Reports*, 15, 1315.
- Journal, A.G. (1983). Nonparametric estimation of spatial distributions. *Mathematical Geology*, 15, 445-468.
- Journal, A.G., & Huijbregts, Ch.J. (1978). *Mining Geostatistics*. Academic Press, London.

- Kangaroglu, F., & Gunay, G. (1997). Ground water nitrate pollution in an alluvial aquifer, Eskir urban area and its vicinity, Turkey. *Environmental Geology*, 31, 178-184.
- Kelin, H., Huang, Y., Li, H., Li, B., Chen, D., & Robert, E.W. (2005). Spatial variability of shallow ground water level, electrical conductivity and nitrate concentration, and risk assessment of nitrate contamination in North China Plain. *Environment International*, 31, 893-903.
- Kresic, N. (1997). *Quantitative Solutions in Hydrogeology and Groundwater Modeling*. Lewis Publishers, Boca Raton, FL, 461p.
- Kuo, Y.M., Liu, C.W., & Lin, K.H. (2005). Evaluation of the ability of an artificial neural network model to assess the variation of groundwater quality in an area of blackfoot disease in Taiwan. *Water Research*, 38, 148-158.
- Mariño, M.A. (1998). California groundwater degradation: agricultural practices and management strategies. In: *Agricultural Threats to Groundwater Quality*, L. Candela and A. Aureli (eds.), Mediterranean Agronomic Institute of Zaragoza, Spain, pp. 61-70.
- McKenna, S.A. (2002). Simulating geological uncertainty with imprecise data for groundwater flow and advective transport modeling. Department of Geology and Geological Engineering, Colorado School of Mines Golden, Colorado, U.S.A., pp. 1-15.
- Nasr, M., & Zahran, H.F. (2014). Using of pH as a tool to predict salinity of groundwater for irrigation purpose using artificial neural network. *Egyptian Journal of Aquatic Research*, 40, 111-115.
- Piri, J., Amin, S., Moghaddamnia, A., Han, D., & Remesun, D. (2009). Daily pan evaporation modelling in hot and dry climate. *Journal of Hydrologic Engineering*, 14, 803-811.
- Piri, J., & Kisi, O. (2015). Modelling solar radiation reached to the Earth using ANFIS, NN-ARX and empirical models (Case studies: Zahedan and Bojnurd stations). *Journal of Atmospheric and Solar-Terrestrial Physics*, 123, 39-47.
- Saito, H., & Goovaerts, P. (2000). Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site. *Environmental Science & Technology*, 34, 4228-4235.
- Scholz, M., et al. (2023). Extension variance: an early geostatistical concept applied to assess nitrate pollution in groundwater. *Hydrogeology Journal*, 31, 1393-1411.
- Sentenac, P., et al. (2023). Addressing Nitrate Contamination in Groundwater: The Importance of Spatial and Temporal Understandings and Interpolation Methods. *Water*, 15(24), 4220.
- Singh, A., et al. (2024). Predicting groundwater level using traditional and deep machine learning algorithms. *Frontiers in Environmental Science*, 12, 1291327.
- Spatial prediction of groundwater nitrate by explainable machine learning models and remote sensing indicators under conditions of imbalanced datasets. (2025). *Geocarto International*, 40(1), Article 2557953.
- Systematic review of neural network applications for groundwater level prediction. (2025). *Discover Applied Sciences*, 7, Article 06817.
- Wang, J., et al. (2020). A Review of the Artificial Neural Network Models for Water Quality Prediction. *Applied Sciences*, 10(17), 5776.
- Yang, F.G., et al. (1998). Comparison of different methods for predicting soil salinity. *Geoderma*, 83, 73-85.
- Zhou, Y., et al. (2022). An Artificial Neural Network Model for Water Quality and Water Consumption Prediction. *Water*, 14(21), 3359.
- Zounemat-Kermani, M., et al. (2024). Machine learning predictive insight of water pollution and groundwater quality in the Eastern Province of Saudi Arabia. *Scientific Reports*, 14, 70610.
- Mapping specific groundwater nitrate concentrations from spatial data using machine learning: A case study of Chongqing, China. (2024). *PMC*, Article 10958364.
- Environmental Modelling & Software. (2025). An assessment of spatial random forests for environmental mapping: the case of groundwater nitrate concentration. Vol 193