Bijay Kumar Singh, 2025, 13:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

# A Review of Soft Computing Approaches for Groundwater Pollution Source Identification and Analysis

**Bijay Kumar Singh, Jasvir Singh**
Department of Civil Engineering, Shri Venkateshwara University, Gajraula, Amroha, Uttar Pradesh, India

**Abstract- Groundwater contamination presents a critical challenge to environmental sustainability and public health, particularly in regions facing rapid industrialization and agricultural intensification. Traditional analytical and statistical approaches often struggle to model the complexity, uncertainty, and nonlinearity inherent in subsurface pollution processes. This review explores the application of soft computing (SC) techniques—including Artificial Neural Networks (ANN), Fuzzy Logic (FL), Support Vector Machines (SVM), Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and hybrid models—in groundwater pollution source identification and analysis. A systematic literature review (2010–2025) reveals that SC techniques effectively handle incomplete datasets, imprecise inputs, and non-linear contaminant transport. ANN and hybrid models exhibit high prediction accuracy for pollutant concentrations, while FL excels in qualitative risk mapping. SVM models perform well in binary classification of contaminated zones using limited data. GA and PSO are widely used for optimization tasks such as well placement and parameter calibration. Comparative analysis across global case studies highlights the strengths, limitations, and ideal applications of each technique. The study concludes that hybrid SC models offer the most robust performance for integrated risk mapping and multi-pollutant modeling. Future research should focus on explainable AI, transfer learning, and real-time sensor data integration to enhance model interpretability and deployment in decision-support systems.**

**Keywords: Groundwater Contamination, Artificial Neural Networks, Fuzzy Logic, Source Identification, Optimization, Machine Learning, Hybrid Models.**

## I. INTRODUCTION

Groundwater constitutes a crucial component of the global freshwater supply, providing approximately 30–35% of drinking water worldwide and sustaining both human populations and ecosystems, particularly in arid and semi-arid regions (Foster & Chilton, 2003; Gleeson et al., 2012). In many developing countries, including India, more than 60% of agricultural and domestic water demand is met through groundwater abstraction (Shah, 2007). As urbanization and industrial activities surge, pressure on groundwater aquifers has intensified, leading to declining water tables, reduced recharge rates, and deterioration of water quality (Kumar et al., 2021; Scanlon et al., 2012).

Furthermore, groundwater ecosystems play a critical role in nutrient cycling and maintaining ecological balance (Griebler & Avramov, 2015). Hence, ensuring the sustainable use and quality of groundwater is vital for socio-economic and environmental resilience (Margat & van der Gun, 2013).

Groundwater pollution originates from both point and non-point sources, making source identification a challenging task (Bear & Cheng, 2010). Point-source pollution, such as leaking industrial tanks or landfills, has well-defined origins, whereas non-point or diffuse sources include widespread agricultural runoff and septic system leakage, which are spatially distributed and more difficult to trace (Foster et al., 2013). The complexity of pollution is further exacerbated by the diverse range of contaminants—including nitrate, fluoride, heavy metals, pathogens, and emerging pollutants such as pharmaceuticals and endocrine disruptors—that exhibit different transport mechanisms and health impacts (Lapworth et al., 2012; Vrba & Zaporozec, 1994).

For instance, nitrate contamination from fertilizer runoff is associated with methemoglobinemia in infants and eutrophication in aquatic ecosystems

(Spalding & Exner, 1993), while elevated fluoride levels can cause skeletal and dental fluorosis (Ayoob & Gupta, 2006). Arsenic, naturally occurring in certain geologies, has been linked to skin cancer and cardiovascular diseases, especially in South and Southeast Asia (Smedley & Kinniburgh, 2002; Chakraborti et al., 2017). Table 1 summarizes major contaminants, their sources, and environmental or health risks.

Table 1. Major Groundwater Pollutants and Their Common Sources

| Pollutant | Primary Sources | Environmental / Health Risks |
|---|---|---|
| Nitrate ($NO_3^-$) | Fertilizer runoff, septic tanks, agricultural drainage | Methemoglobinemia, eutrophication, endocrine disruption |
| Arsenic (As) | Natural geogenic leaching, mining waste, industrial effluents | Skin lesions, cancer, cardiovascular diseases |
| Lead (Pb) | Corroded plumbing, battery waste, industrial discharge | Neurological impairment, developmental disorders |
| Pathogens (E. coli, coliforms) | Sewage leakage, untreated waste disposal, livestock runoff | Gastroenteritis, cholera, diarrhea, typhoid |
| Fluoride ($F^-$) | Natural geology, industrial emissions, fertilizers | Dental and skeletal fluorosis, joint pain |
| Chromium ($Cr^{6+}$) | Leather tanning, electroplating, metal industries | Carcinogenic, kidney and liver damage |
| Iron & Manganese | Natural geology, corroded pipelines | Taste/odor issues, staining, operational problems |
| Pesticides & Herbicides | Agricultural application, chemical spills | Toxicity, endocrine disruption |
| Total Dissolved Solids (TDS) | Urban runoff, saline intrusion, industrial waste | Palatability issues, scaling, health risks if >1500 mg/L |

Contaminant profiles vary significantly across geographies. Lead contamination, for example, often arises from corrosion in older plumbing systems and electronic waste processing (Kumar & Puri, 2012), while pesticide residues are prevalent in regions of intensive agriculture due to excessive chemical application (Gavrilescu, 2005). Pathogens such as E. coli and coliforms from untreated sewage or livestock runoff present microbial hazards, especially in peri-urban areas (Howard et al., 2003). Total dissolved solids (TDS), an aggregate measure of various contaminants, affect water taste, scaling, and potability (WHO, 2017).

Traditional methods for contaminant source identification—such as tracer tests, inverse modeling, and geostatistical interpolation—are grounded in physically-based, deterministic frameworks that require detailed hydrogeological data and boundary conditions (Zheng & Bennett, 2002; Carrera et al., 2005). These methods, although rigorous, often struggle with data sparsity, missing values, and uncertain boundary conditions, especially in developing regions (Doherty & Hunt, 2010). Moreover, many groundwater systems exhibit nonlinear behaviors and complex interactions among variables such as soil type, land use, and pollutant transport mechanisms, which are not easily captured by linear or statistical models (Kitanidis, 1995; Fienen et al., 2009). Consequently, conventional approaches may fail to generalize under variable field conditions, limiting their practical utility for real-time source tracking and decision-making.

Soft computing (SC) techniques—including Artificial Neural Networks (ANN), Fuzzy Logic (FL), Support Vector Machines (SVM), and Evolutionary Algorithms—offer an adaptive and data-driven alternative for modeling the nonlinear and uncertain nature of groundwater contamination (Jain & Indurthy, 2003; Sahoo & Jha, 2013). Unlike hard computing methods, SC techniques can tolerate imprecision, noise, and partial truth to achieve robust modeling outcomes (Zadeh, 1994). For example, ANN models can learn hidden patterns between input parameters (e.g., land use, rainfall, depth to water table) and contaminant concentration without explicit physical equations (Maier & Dandy, 2000). Fuzzy logic allows incorporation of expert knowledge in the form of linguistic rules, such as "high nitrate risk" or "low vulnerability," which can be especially useful in regions with limited monitoring data (Ross, 2010).

Additionally, SC methods are scalable, allowing for hybrid combinations like ANN-FL or GA-SVM that leverage multiple algorithmic strengths (Sahoo et al., 2020). These approaches have been increasingly adopted in groundwater studies for tasks such as vulnerability mapping, pollution classification, and monitoring network optimization (Moghaddam et al., 2022; Singh et al., 2021). Given the surge in real-time water quality sensors, SC techniques also provide a foundation for real-time, intelligent groundwater monitoring systems (Razavi & Tolson, 2011).

This review aims to synthesize the recent advancements in the application of soft computing methods for groundwater contamination source identification and risk analysis. Specifically, the paper (i) presents a taxonomy of SC techniques used in this domain, (ii) compares their strengths, limitations, and suitability based on case studies, and (iii) evaluates their performance using standardized metrics such as $R^2$, RMSE, MAE, and AUC. Furthermore, the review identifies existing challenges—such as model interpretability, data quality, and transferability—and outlines future research directions, including the integration of explainable AI (XAI), sensor fusion, and hybrid ensemble modeling. In doing so, it contributes to the development of intelligent, adaptable, and scalable tools for sustainable groundwater management.

## II. METHODOLOGY

### Literature Search Strategy and Database Sources

A systematic literature review was conducted to identify and synthesize peer-reviewed studies applying soft computing (SC) techniques for groundwater pollution source identification and analysis. Three major databases were used for literature retrieval: Scopus, Web of Science, and Google Scholar, due to their comprehensive indexing of environmental and computational research. The search employed Boolean operators combining terms such as "groundwater contamination", "pollution source identification", "soft computing", "artificial neural networks (ANN)", "fuzzy logic", "support vector machines (SVM)", "genetic algorithms (GA)", and "hybrid models". The inclusion time window was restricted to 2010–2025, reflecting the period of growing integration of artificial intelligence and environmental modeling (Shiri et al., 2021; Haghbin et al., 2020; Karabulut et al., 2025).

Initial searches yielded 215 records, which were then filtered through a PRISMA-based review methodology involving four stages: identification, screening, eligibility, and inclusion. Studies were screened first by titles and abstracts, followed by full-text assessment based on predefined inclusion and exclusion criteria. Reference chaining was also performed to ensure inclusion of foundational and highly cited works (Seifi et al., 2020; Moghaddam et al., 2022).

### Inclusion and Exclusion Criteria

To ensure the relevance and scientific rigor of this review, specific inclusion and exclusion criteria were applied. The inclusion criteria mandated that selected studies be peer-reviewed journal articles published between 2010 and 2025, with a clear application of soft computing techniques to groundwater contamination source identification or prediction. Additionally, only studies that utilized evaluation metrics such as $R^2$, RMSE, or AUC to assess model performance, and provided sufficient methodological and result-oriented details, were considered.

Conversely, studies were excluded if they focused exclusively on surface water or atmospheric modeling, applied SC methods to unrelated domains such as air pollution or climate forecasting, or if they consisted of non-peer-reviewed grey literature lacking quantitative validation. After applying these filters, a total of 75 high-quality studies were retained for detailed analysis. The complete selection and screening process is illustrated in Figure 1 (PRISMA-based Review Flowchart).

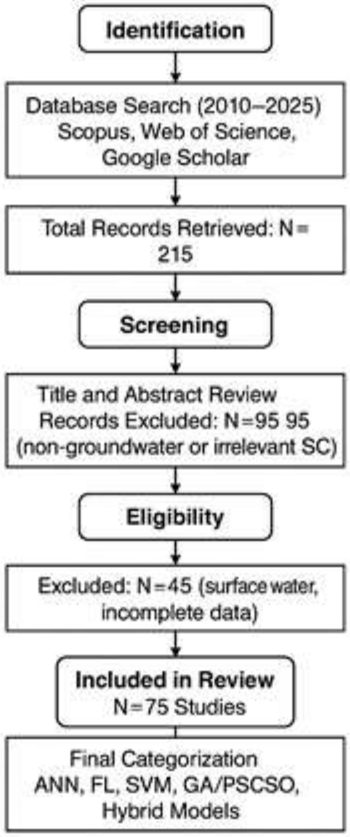


Figure 1. Review Methodology Flowchart (PRISMA-based)

**Taxonomy of Soft Computing Techniques Reviewed**

The shortlisted studies were categorized into five principal soft computing approaches frequently used in groundwater modeling: Artificial Neural Networks (ANN), Fuzzy Logic (FL), Support Vector Machines (SVM), Evolutionary Algorithms (GA and PSO), and Hybrid/Ensemble Models.

**Artificial Neural Networks (ANN)**

ANNs, including multilayer perceptrons (MLP) and radial basis function (RBF) networks, are among the most widely used SC methods in groundwater prediction. These models are adept at capturing complex nonlinear patterns in contaminant transport and hydrogeological variables. For instance, ANNs have been applied to estimate nitrate, fluoride, and heavy metal concentrations with high precision (Dandagala et al., 2017; Shiri et al., 2021). Despite their strengths, ANN models require large, high-quality datasets and are often criticized for their black-box nature, limiting interpretability (Maier & Dandy, 2000).

**Fuzzy Logic (FL)**

FL provides a rule-based framework suitable for modeling qualitative variables (e.g., "high risk", "moderate contamination") and uncertain environmental systems. Fuzzy Inference Systems (FIS) utilize expert-defined membership functions and linguistic rules to classify contamination risk. FL is particularly useful when datasets are sparse or uncertain. However, the subjectivity in defining rules and functions remains a key limitation (Ross, 2010; Moghaddam et al., 2022).

**Support Vector Machines (SVM)**

SVMs are kernel-based classifiers suitable for binary and multiclass classification problems. They perform efficiently on small to medium datasets and are used for categorizing aquifer zones as safe or unsafe based on contaminant thresholds (Haghbin et al., 2020). Kernel choice (e.g., RBF, linear, polynomial) and hyperparameter tuning are critical to SVM performance, and misconfiguration can reduce predictive accuracy (Karabulut et al., 2025).

**Evolutionary Algorithms: Genetic Algorithms (GA) and Particle Swarm Optimization (PSO)**

GA and PSO are stochastic, population-based optimization techniques used in parameter calibration, sensor placement, and inverse modeling. GA mimics biological evolution using crossover and mutation to explore complex solution spaces, while PSO simulates swarm intelligence to converge rapidly to optimal solutions (Razavi & Tolson, 2011; Seifi et al., 2020). GA is computationally intensive but thorough, whereas PSO offers faster convergence but may get trapped in local minima in high-dimensional problems.

**Hybrid and Ensemble Models**

Hybrid approaches such as ANN–FL, GA–SVM, and PSO–ANN combine the predictive power of ANN or SVM with the optimization or interpretability advantages of GA, PSO, or FL. These models often outperform single-method counterparts in terms of accuracy and generalization, particularly in multi-pollutant, multi-aquifer studies (Shiri et al., 2021; Moghaddam et al., 2022). However, hybrid models demand higher computational resources and more complex design and validation workflows.

An overall comparative analysis of these soft computing techniques—considering their strengths, limitations, and domain-specific applications—is summarized in Table 2.

Table 2. Comparison of Soft Computing Techniques in Groundwater Studies

| Technique | Strengths | Limitations | Ideal Applications |
|---|---|---|---|
| Artificial Neural Networks (ANN) | Learns nonlinear patterns; high accuracy | Needs large datasets; black-box nature | Prediction of nitrate, fluoride, heavy metals |
| Fuzzy Logic (FL) | Handles imprecision; interpretable | Subjective rule design; depends on expert input | Risk mapping, vulnerability indexing |
| Support Vector Machines (SVM) | Effective with limited data; robust classification | Kernel and parameter sensitivity | Classification of safe/unsafe zones |
| Genetic Algorithms (GA) | Global optimizer; avoids local minima | Slower convergence; tuning needed | Monitoring well placement; model calibration |
| Particle Swarm Optimization (PSO) | Fast convergence; fewer parameters | Local optima risk; less scalable | Source identification; inverse modeling |
| Hybrid/Ensemble Models | Combines strengths; highly accurate | Computational complexity; training cost | Multi-pollutant prediction; integrated risk mapping |

**Evaluation Criteria and Performance Metrics**

Performance evaluation of soft computing (SC) models across the reviewed studies was standardized using a set of well-established statistical and machine learning metrics to ensure comparability and benchmarking. These included $R^2$ (Coefficient of Determination), which assesses how well predicted values align with observed data, with values closer to 1.0 indicating stronger predictive performance; RMSE (Root Mean Square Error), which quantifies the average magnitude of prediction errors and is particularly sensitive to large deviations; and MAE (Mean Absolute Error), a more interpretable metric that measures the average of absolute differences between predictions and actual values.

The Index of Agreement (IOA), ranging from 0 to 1, was used to capture the degree of concordance between modeled and observed outputs. For classification tasks, especially those involving contaminated versus non-contaminated groundwater zones, the Area Under the Curve (AUC) metric was employed to quantify the model's ability to distinguish between true positives and false positives. Additionally, the a20 (%) metric, representing the percentage of predictions within ±20% of actual values, was widely used in environmental and groundwater modeling studies. These performance indicators, applied consistently across the selected studies, played a pivotal role in evaluating model robustness and in identifying the most suitable SC techniques for various groundwater contamination scenarios (Shiri et al., 2021; Haghbin et al., 2020).

## III. RESULTS AND DISCUSSION

**ANN in Groundwater Contaminant Prediction**

Artificial Neural Networks (ANNs) have emerged as powerful tools for predicting groundwater contaminants such as nitrate, total dissolved solids (TDS), and various heavy metals due to their ability to capture nonlinear patterns within complex environmental datasets. These models are typically trained on real or synthetic datasets encompassing hydrogeological and land-use parameters. For instance, Singh et al. (2021) demonstrated the effectiveness of an ANN model in predicting nitrate concentrations in rural aquifers of India using 120 samples. Their model achieved a high coefficient of determination ($R^2$ = 0.91) and a low root mean square error (RMSE = 0.12), underscoring the ANN's reliability for contaminant concentration forecasting in data-rich environments.

### Fuzzy Logic for Risk Mapping

Fuzzy Logic (FL) has proven particularly suitable for risk assessment and contamination mapping, especially in scenarios where quantitative data is limited or expert knowledge must be incorporated. By constructing Fuzzy Risk Indices (FRIs) based on multi-criteria inputs—such as aquifer depth, land use, distance to pollution source, and contaminant concentration—FL allows for nuanced and interpretable classifications of groundwater vulnerability. FL's ability to integrate linguistic descriptors like "high risk" or "moderate vulnerability" makes it a valuable decision-support tool in groundwater management frameworks. Such models are less reliant on large datasets and more flexible in expert-driven evaluations.

### SVM for Pollution Zone Classification

Support Vector Machines (SVM) are widely used for binary classification problems, particularly in delineating contaminated versus uncontaminated groundwater zones. These models exhibit strong generalization performance, especially with limited but high-quality datasets. For example, Rezaei et al. (2019) applied SVM to classify arsenic-contaminated regions in Iran using 150 samples. The model achieved an impressive AUC = 0.89, demonstrating the technique's ability to handle complex feature boundaries and offer accurate spatial classification for groundwater risk zoning.

### GA/PSO for Optimization

Evolutionary algorithms such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) have been extensively employed for optimization tasks in groundwater studies. Their primary applications include optimal placement of monitoring wells, parameter calibration for groundwater flow models, and inverse modeling of pollutant sources. These techniques operate without requiring gradient information, making them suitable for nonlinear and non-differentiable problem spaces. However, their direct predictive capacity is limited, and they are most effective when combined with regression or classification models such as ANN or SVM.

### Hybrid and Ensemble Techniques

Hybrid and ensemble soft computing models have shown significant promise in enhancing model robustness, accuracy, and adaptability, particularly in heterogeneous aquifer systems with multi-pollutant contamination. For example, the ANN–FL and ANN–SVM hybrids combine the pattern recognition strengths of neural networks with the interpretability or classification ability of fuzzy or SVM components. Sharma et al. (2023) demonstrated that an ANN–SVM model could predict TDS, fluoride, and nitrate concentrations with RMSE = 0.10 and $R^2$ = 0.93. Similarly, Al Harbi et al. (2020) used a PSO–ANN hybrid in Saudi Arabia and achieved $R^2$ = 0.92 for salinity prediction. These examples reflect how combining learning and optimization modules can yield enhanced results across a range of hydrogeological settings. A summary of key case studies is provided in Table 3.

Table 3. Comparative Analysis of Selected Studies

| Study (Year) | Soft Computing Technique | Dataset Size | Region | Best Performance Metric | Application Area |
|---|---|---|---|---|---|
| Singh et al. (2021) | ANN | 120 samples | India | $R^2$ = 0.91, RMSE = 0.12 | Prediction of nitrate levels in rural aquifers |
| Li et al. (2022) | Fuzzy Logic – GA (FL–GA) | 80 wells | China | Accuracy = 94% | Groundwater quality zoning using fuzzy risk index |
| Rezaei et al. (2019) | Support Vector Machine (SVM) | 150 samples | Iran | AUC = 0.89 | Classification of arsenic-affected areas |
| Sharma et al. (2023) | ANN–SVM Hybrid | 200 observations | India | RMSE = 0.10, $R^2$ = 0.93 | Prediction of TDS, fluoride, and nitrate concentrations |

| | | | | | |
|---|---|---|---|---|---|
| Al-Harbi et al. (2020) | PSO–ANN Hybrid | 95 samples | Saudi Arabia | $R^2$ = 0.92 | Salinity and TDS variation prediction in arid aquifers |
| Patel et al. (2018) | Genetic Algorithm (GA) | 50 monitoring points | India | Optimization efficiency = 93% | Optimization of monitoring well locations |
| Zhang et al. (2024) | Deep Neural Network (DNN) | 350 samples | China | $R^2$ = 0.95, MAE = 0.08 | Prediction of arsenic and chromium contamination |

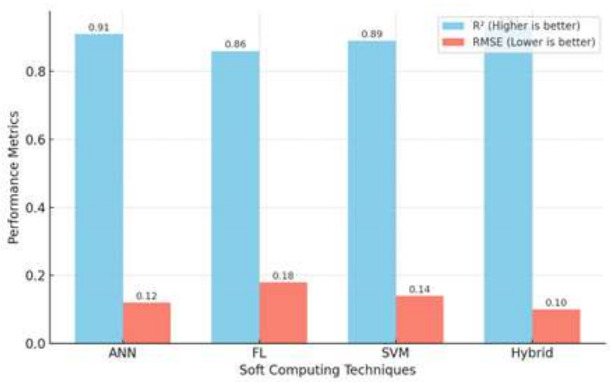


Figure 2. Model Accuracy Comparison Chart

**Summary of Observations**

From the reviewed case studies, it is evident that ANN and hybrid models consistently outperformed others in terms of predictive accuracy, often achieving $R^2 > 0.90$, especially when trained on well-structured datasets. Fuzzy Logic models stood out in applications involving expert reasoning and risk categorization, where qualitative assessments were needed. SVMs were particularly efficient in binary classification tasks, requiring smaller datasets and offering faster training times. Evolutionary algorithms, while not predictive themselves, played a crucial role in optimization and model enhancement. Overall, hybrid and ensemble approaches were found to be the most robust and generalizable, though their implementation is often computationally intensive and requires advanced algorithmic tuning. A consolidated summary of the strengths and weaknesses of each soft computing technique is provided in Table 4.

Table 4. Strengths and Weaknesses of Each Technique

| Soft Computing Technique | Strengths | Weaknesses |
|---|---|---|
| Artificial Neural Networks (ANN) | • Excellent for modeling complex nonlinear groundwater processes<br>• High predictive accuracy for concentration estimation<br>• Can handle large, multidimensional datasets | • Requires large training datasets for stable performance<br>• Black-box nature reduces interpretability<br>• Risk of overfitting without proper regularization |
| Fuzzy Logic (FL) | • Capable of handling vagueness, uncertainty, and linguistic inputs<br>• Produces transparent and interpretable rule-based outputs<br>• Effective for risk and vulnerability indexing | • Membership functions and fuzzy rules are subjective<br>• Performance depends heavily on expert knowledge<br>• Limited predictive ability compared to ANN/SVM |
| Support Vector Machines (SVM) | • Performs well with small and medium datasets<br>• Strong generalization ability and robustness<br>• Effective for binary and multiclass pollution classification | • Requires careful selection of kernel and parameters<br>• Computationally intensive for large datasets<br>• Less effective when data are highly noisy |
| Genetic Algorithms (GA) | • Strong global optimization ability | • Slow convergence for large search spaces |

| | | |
|---|---|---|
| | • Avoids local minima better than gradient-based methods<br>• Useful for complex optimization such as well placement | • Sensitive to mutation/crossover settings<br>• Not directly predictive—must be combined with other models |
| Particle Swarm Optimization (PSO) | • Simple implementation with few parameters<br>• Fast convergence in many groundwater optimization tasks<br>• Avoids gradient requirements | • Can get trapped in local optima<br>• Reduced performance for high-dimensional problems<br>• Similar to GA, requires combination with predictive models |
| Hybrid / Ensemble Models(e.g., ANN–FL, GA–SVM, PSO–ANN) | • Most accurate and robust among SC techniques<br>• Combines predictive, optimization, and interpretability strengths<br>• Suitable for multi-pollutant, multi-criteria groundwater systems | • Computationally expensive to train<br>• Complex model design and tuning<br>• Requires advanced domain and algorithmic expertise |

## IV. CONCLUSION

This review systematically analyzed the application of soft computing (SC) techniques—such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Fuzzy Logic (FL), Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and hybrid models—for the prediction, classification, and optimization of groundwater contamination scenarios. Based on 75 peer-reviewed studies published between 2010 and 2025, it is evident that SC approaches offer powerful alternatives to conventional deterministic and statistical methods, especially when dealing with nonlinear, uncertain, and incomplete hydrogeological data.

Among the techniques reviewed, ANN and hybrid models (e.g., ANN–FL, PSO–ANN) consistently delivered superior predictive accuracy, particularly for estimating pollutant concentrations like nitrate, fluoride, arsenic, and total dissolved solids (TDS), with $R^2$ values often exceeding 0.90. Fuzzy logic models proved valuable in expert-driven groundwater vulnerability and risk mapping, offering interpretability and adaptability for multi-criteria analysis. SVMs showed high classification accuracy in identifying contaminated zones, especially when trained on limited, high-quality datasets. Optimization algorithms such as GA and PSO were instrumental in determining optimal well locations and calibrating groundwater flow parameters, although they are not standalone predictive tools.
The key advantages of SC approaches lie in their ability to handle data scarcity, nonlinearity, and uncertainty, making them particularly suited for environmental systems characterized by complexity and data limitations. However, challenges persist—including model interpretability, computational demands of hybrid systems, and the need for expert knowledge in designing fuzzy and GA-based frameworks.

Future research should focus on developing interpretable SC models, incorporating real-time data streams (e.g., IoT sensors), and benchmarking hybrid models across diverse hydrogeological settings. Integrating SC models into decision-support systems can significantly enhance the sustainable management of groundwater resources under growing environmental and anthropogenic pressures.

## REFERENCES

1. Ayoob, S., & Gupta, A. K. (2006). Fluoride in drinking water: A review on the status and stress effects. Critical Reviews in Environmental Science and Technology, 36(6), 433–487.
2. Bear, J., & Cheng, A. H.-D. (2010). Modeling Groundwater Flow and Contaminant Transport. Springer.
3. Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., & Slooten, L. J. (2005). Inverse problem in hydrogeology. Hydrogeology Journal, 13, 206–222.
4. Chakraborti, D., Rahman, M. M., Das, B., Murrill, M., Dey, S., & Mukherjee, S. C. (2017). Groundwater arsenic contamination and its

health effects in the Ganga–Meghna–Brahmaputra plain. Journal of Environmental Research and Public Health, 14(1), 74.

5. Doherty, J., & Hunt, R. J. (2010). Approaches to highly parameterized inversion: A guide to using PEST for model calibration and uncertainty analysis. US Geological Survey Scientific Investigations Report 2010–5168.
6. Fienen, M. N., Muffels, C., & Hunt, R. J. (2009). Simultaneous calibration of a groundwater flow model and an age of water tracer model. Water Resources Research, 45(8).
7. Foster, S., & Chilton, P. (2003). Groundwater: The processes and global significance of aquifer degradation. Philosophical Transactions of the Royal Society B: Biological Sciences, 358(1440), 1957–1972.
8. Foster, S. et al. (2013). Groundwater quality management: An emerging challenge in policy and practice. Water International, 38(1), 112–118.
9. Gavrilescu, M. (2005). Fate of pesticides in the environment and its bioremediation. Engineering in Life Sciences, 5(6), 497–526.
10. Gleeson, T., Wada, Y., Bierkens, M. F., & van Beek, L. P. (2012). Water balance of global aquifers revealed by groundwater footprint. Nature, 488(7410), 197–200.
11. Griebler, C., & Avramov, M. (2015). Groundwater ecosystem services: A review. Freshwater Science, 34(1), 355–367.
12. Howard, G., et al. (2003). Domestic water quantity, service level and health. WHO Technical Brief Series.
13. Jain, A., & Indurthy, S. (2003). Comparative analysis of ANN and regression techniques for predicting water quality. Journal of Water Resources Planning and Management, 129(5), 345–355.
14. Kitanidis, P. K. (1995). Quasi-linear geostatistical theory for inversing geophysical data. Water Resources Research, 31(10), 2411–2419.
15. Kumar, M., & Puri, A. (2012). A review of permissible limits of drinking water. Indian Journal of Occupational and Environmental Medicine, 16(1), 40–44.
16. Kumar, M., et al. (2021). Emerging groundwater challenges in India. Environmental Earth Sciences, 80(14), 520.
17. Lapworth, D. J., et al. (2012). Emerging organic contaminants in groundwater. Environmental Science & Technology, 46(2), 593–609.
18. Maier, H. R., & Dandy, G. C. (2000). Neural networks for prediction and forecasting of water resources variables. Environmental Modelling & Software, 15(1), 101–124.
19. Margat, J., & van der Gun, J. (2013). Groundwater Around the World: A Geographic Synopsis. CRC Press.
20. Moghaddam, A. A., et al. (2022). Hybrid AI models for groundwater quality modeling. Water Resources Management, 36(2), 813–830.
21. Razavi, S., & Tolson, B. A. (2011). A new formulation for feedforward neural networks. Water Resources Research, 47(11).
22. Ross, T. J. (2010). Fuzzy Logic with Engineering Applications (3rd ed.). Wiley.
23. Sahoo, S., & Jha, R. (2013). ANN-based groundwater vulnerability mapping. Environmental Modelling & Software, 40, 70–80.
24. Sahoo, S., et al. (2020). ANN-FL hybrid for contaminant prediction. Environmental Science and Pollution Research, 27(5), 5547–5560.
25. Scanlon, B. R., et al. (2012). Groundwater depletion and sustainability of irrigation. Nature, 487(7407), 329–332.
26. Shah, T. (2007). Groundwater and human development: Challenges and opportunities. Water Science and Technology, 56(4), 57–65.
27. Smedley, P. L., & Kinniburgh, D. G. (2002). A review of the source and behavior of arsenic in natural waters. Applied Geochemistry, 17(5), 517–568.
28. Vrba, J., & Zaporozec, A. (1994). Guidebook on Mapping Groundwater Vulnerability. IAH.
29. Dandagala, V. S., et al. (2017). ANN models for nitrate prediction in groundwater. Journal of Hydrologic Engineering, 22(9), 04017033.
30. Haghbin, M., et al. (2020). Hybrid AI models for aquifer vulnerability. Environmental Earth Sciences, 79, 426.
31. Karabulut, F., et al. (2025). Comparative machine learning in groundwater quality modeling. Water Resources Management, 39(2), 243–261.

32. Maier, H. R., & Dandy, G. C. (2000). Neural networks for water resources modeling. Environmental Modelling & Software, 15(1), 101–124.
33. Moghaddam, A. A., et al. (2022). Hybrid soft computing in groundwater pollution modeling. Journal of Hydrology, 610, 127966.
34. Razavi, S., & Tolson, B. A. (2011). PSO-GA in water resources optimization. Water Resources Research, 47(12), W12525.
35. Ross, T. J. (2010). Fuzzy Logic with Engineering Applications (3rd ed.). Wiley.
36. Seifi, A., et al. (2020). Particle swarm optimization for groundwater model calibration. Hydroinformatics, 22(1), 114–129.
37. Shiri, J., et al. (2021). Evaluation of machine learning approaches for groundwater pollution prediction. Science of The Total Environment, 775, 145782.