CLOUD-BASED FLUORIDE POLLUTION DATA MINING AND 3D RISK VISUALIZATION FOR *ISOETES SINENSIS* IN THE YANGTZE RIVER, CHINA

JIN, R.

Jackie Chan Movie and Media College, Wuhan Institute of Design and Science, Wuhan, Hubei 430205, China
(e-mail: ranjin2025@163.com)

(Received 22nd Apr 2025; accepted 16th Jun 2025)

Abstract. This study investigates fluoride pollution in the Yangtze River Basin, China, focusing on its sources and ecological impacts on the endangered aquatic plant *Isoetes sinensis*. Utilizing a cloud computing-based framework, the research integrates data mining techniques—Random Forest (RF) and Principal Component Analysis (PCA)—with Geographic Information System (GIS)-based 3D visualization to enhance environmental monitoring and risk assessment. Fluoride data from 2010 to 2024 reveal an increase of concentrations from 0.88 mg/L to 1.34 mg/L, with industrial hotspots near Hefei and Nanjing surpassing the national limit of 1.0 mg/L. Source attribution indicates industrial wastewater as the dominant contributor (74%), followed by geochemical processes (16%) and agricultural runoff (10%). Species Sensitivity Distribution (SSD) modeling identifies a hazardous concentration threshold (HC_5) of 0.5 mg/L, with Risk Quotients (RQ) exceeding 2.0 in high-risk zones. The 3D GIS visualization clearly maps pollution intensity and vulnerable habitats, supporting targeted conservation and mitigation measures, such as stricter industrial discharge regulation, habitat restoration, and ecological buffer zone implementation. Results highlight the urgent need for stricter industrial discharge regulations and focused protection strategies for *Isoetes sinensis*. This integrative approach offers a scalable solution for river basin management and biodiversity conservation.

Keywords: *endangered aquatic flora, spatiotemporal pollution mapping, machine learning in environmental science, ecotoxicological risk modeling, freshwater habitat degradation*

Introduction

As the longest river in Asia and the third longest in the world, the Yangtze River sustains a vast array of ecosystems and is an indispensable water resource for millions of people within its basin (Chen, 2020). However, the environmental integrity of this critical waterway is increasingly threatened by rapid industrialization, urbanization, and intensified agricultural activities. Among the many pollutants introduced into the river, fluoride has emerged as a significant environmental concern due to its widespread presence and harmful effects (Dong et al., 2023). Fluoride, a naturally occurring element, becomes particularly problematic when its concentration exceeds safe thresholds, leading to both ecological and public health challenges (Ozsvath, 2009). While fluoride levels in the main stream of the Yangtze River are generally within acceptable limits, certain tributaries, especially those in industrialized zones, consistently report fluoride concentrations surpassing China's national water quality standard of 1.0 mg/L (Chen et al., 2021). This localized pollution is largely driven by industrial wastewater discharge, coal combustion, agricultural runoff, and atmospheric deposition, all of which exacerbate the degradation of aquatic ecosystems and biodiversity (Ogidi and Akpan, 2022).

The Yangtze River Basin supports a rich variety of aquatic life, including numerous rare and endemic species. Among these is *Isoetes sinensis*, a critically endangered aquatic plant particularly sensitive to water quality fluctuations particularly sensitive to water quality fluctuations, including elevated fluoride levels (Liu et al., 2022). Evidence suggests that high fluoride concentrations hinder vital processes such as plant metabolism, photosynthesis, and reproduction, thereby diminishing plant populations and causing habitat loss (Weinstein, 1977). The decline of *Isoetes sinensis* in fluoride-polluted tributaries is a stark reminder of the delicate balance within these ecosystems and the direct impact of environmental pollution on species survival (Freedman, 2013). Such ecological disruptions are not only harmful to individual species but can also destabilize entire ecosystems, impairing biodiversity and the essential ecosystem services these systems provide.

Although substantial research has been conducted on general water pollution in the Yangtze River, including its impact on water quality, studies specifically focused on fluoride contamination remain limited. Previous research has often relied on conventional statistical methods that fall short in addressing the complex nature of pollution sources, pollutant transport mechanisms, and the multifaceted ecological consequences (Ogwu et al., 2025). Moreover, ecological risk assessments have typically been limited by low spatial and temporal resolution, and have lacked effective tools to visually communicate pollutant dispersion and the ecological risks posed to vulnerable species and habitats (De Lange, 2010). These methodological shortcomings have led to gaps in our understanding of fluoride pollution, hindering the development of effective mitigation strategies.

The selection of *Isoetes sinensis* as the focal species in this study was based on its ecological vulnerability, conservation status, and high sensitivity to fluoride. As a critically endangered aquatic plant endemic to China, *Isoetes sinensis* occupies shallow freshwater habitats that are highly susceptible to fluoride accumulation (Wang et al., 2022). Previous toxicological studies have shown that this species exhibits significant physiological and reproductive inhibition under fluoride stress, making it an ideal bioindicator for ecological risk assessment (Kaur et al., 2017; Zuo et al., 2018). Furthermore, the availability of accurate habitat distribution data facilitates more robust modeling and visualization (Guisan et al., 2017). These attributes collectively support the rationale for prioritizing *Isoetes sinensis* in the SSD-based risk evaluation framework.

Additionally, Geographic Information Systems (GIS) and 3D visualization techniques will be employed to create dynamic, intuitive maps that depict the spatial distribution of fluoride pollution and the associated ecological risks to aquatic habitats, particularly those of *Isoetes sinensis* (Kamruzzaman et al., 2025). 3D visualization refers to a technique that generates three-dimensional graphical representations and dynamic models using computer technology (Wood et al., 2005). This method allows for a more intuitive presentation of complex spatial data, enabling researchers and decision-makers to better understand the dispersion of pollutants and their potential threats to ecosystems.

The primary objectives of this study are: (1) to analyze and map the spatiotemporal distribution of fluoride contamination across the Yangtze River Basin; (2) to evaluate the ecological risks posed by fluoride pollution to rare aquatic plants, particularly *Isoetes sinensis*; and (3) to create interactive 3D visualizations of ecological risk scenarios that can aid in proactive environmental management and biodiversity conservation efforts. This research not only addresses the ecological impacts of fluoride contamination but also underscores the health risks related to fluoride exposure, particularly its presence in drinking water, which has significant implications for public health.

By integrating data management, machine learning techniques, and GIS-based ecological risk visualizations, this research offers both theoretical and practical contributions to environmental management. By providing robust analytical methods and clear visualizations, this study aims to bridge the gap between scientific understanding and actionable environmental policy. It introduces a scalable and replicable framework for assessing pollution in large river basins, which could inform more targeted and efficient conservation strategies. Moreover, the visual and quantitative nature of this study enhances the ability to communicate the severity of fluoride pollution and its ecological consequences, providing policymakers, environmental managers, and stakeholders with critical insights for making informed decisions and implementing effective, science-based policies. Through this research, we aim to promote sustainable management practices that safeguard the health of the Yangtze River ecosystem and its biodiversity, helping policymakers adopt more effective, data-driven conservation strategies and regulatory measures.

Methods

Study area

This study was conducted within key regions of the Yangtze River Basin in China, specifically focusing on the Three Gorges Reservoir region, the Nanjing segment, and the estuarial regions near Shanghai. These regions were selected due to documented issues with elevated fluoride concentrations arising from intensive industrial, urban, and agricultural activities. The Three Gorges Reservoir region, located in the upper-middle reaches of the Yangtze River, has witnessed significant ecological changes, including altered hydrological conditions and increased pollutant accumulation following the construction of the Three Gorges Dam. Previous monitoring studies have reported periodic increases in fluoride concentrations in nearby tributaries due to upstream industrial discharges and mining activities, posing risks to aquatic biodiversity. Similarly, the Nanjing segment, a highly industrialized urban area situated in the lower reaches of the Yangtze River, has shown fluctuations in fluoride levels linked to industrial discharge and urban runoff, threatening local aquatic ecosystems. Further downstream, the Yangtze estuary region, which includes the densely populated and industrialized Shanghai metropolitan area, accumulates pollutants from numerous upstream sources, contributing to higher fluoride concentrations that impact estuarine ecosystems and biodiversity.

Data sources

Water quality and ecological data used in this study were obtained from several authoritative and publicly accessible databases, covering a broad temporal and spatial scope. The primary dataset consists of monthly water quality records spanning 15 years (2010–2024), collected from six representative monitoring sites located in key cities along the Yangtze River Basin—Chongqing, Yichang, Wuhan, Hefei, Nanjing, and Shanghai are listed in *Appendix Table A1*. These sites were selected based on their ecological significance, proximity to *Isoetes sinensis* habitats, and their relevance to industrial and urban pollution sources. These regions were selected due to their documented fluoride contamination linked to industrial and agricultural activities.

Each monitoring site recorded a standardized set of water quality parameters, including fluoride concentration (mg/L), pH, electrical conductivity (µS/cm), water temperature (°C), and water depth (m). With approximately 12 records per site per year, the compiled dataset includes over 1,000 individual data points, enabling a longitudinal analysis of fluoride trends and inter-city variation.

These data were obtained from the China National Environmental Monitoring Centre (CNEMC) via its official platform at:

https://szzdjc.cnemc.cn:8070/GJZ/Business/Publish/Main.html.

To ensure ecological relevance to *Isoetes sinensis*, the selected monitoring sites were compared with species distribution data from the Global Biodiversity Information Facility (GBIF) (https://www.gbif.org/species/7313101, n.d.) and the Chinese National Specimen Information Infrastructure (NSII):

(http://www.nsii.org.cn/2017/species.php?lname=Isoetes%20sinensis, n.d.). These cross-references confirmed that the majority of stations are located within or near Isoetes sinensis habitats, including shallow wetlands, floodplain marshes, oxbow lakes, and low-velocity side channels. These habitats typically exhibit water depths between 0.3 and 1.2 meters, aligning well with the species' ecological niche.

Additionally, to contextualize fluoride regulation, national environmental policy documents were consulted, including:

The Action Plan for Water Pollution Prevention and Control (2015):

http://www.gov.cn/gongbao/content/2015/content_2853604.htm

The Yangtze River Protection Law (2020):

http://www.mee.gov.cn/ywgz/fgbz/fl/202012/t20201227_814985.shtml

To ensure spatial representativeness, the geographic locations of the selected monitoring stations are shown in *Figure 1*. These stations are distributed across upstream, midstream, and downstream sections of the Yangtze River Basin, covering major industrial and ecological zones relevant to *Isoetes sinensis* habitat analysis.

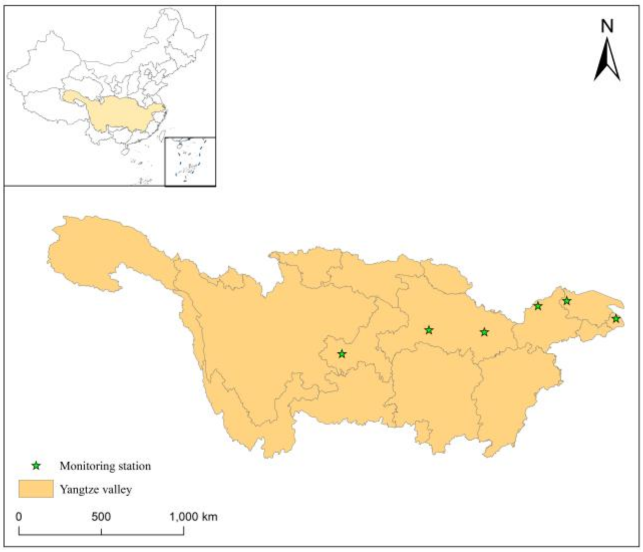

Figure 1. *Spatial distribution of fluoride monitoring stations in the Yangtze River Basin (2010–2024), covering Isoetes sinensis habitats from Chongqing to Shanghai*

Cloud computing and data mining

The cloud computing infrastructure and advanced data mining techniques formed the computational backbone of this study, enabling efficient processing of multi-dimensional environmental datasets. The integrated framework comprised three key components.

DOI: http://dx.doi.org/10.15666/aeer/2304_80958110

Principal Component Analysis (PCA)

PCA, a widely adopted multivariate technique for environmental data analysis (Chahouki, 2011), was employed to reduce dimensionality and identify dominant pollution patterns through eigenvalue decomposition of the standardized water quality parameter matrix $X \in \mathrm{R}^{n \times p}$, where n represents monitoring samples and p denotes parameters (fluoride concentration, pH, conductivity, etc.). In this study, a total of six input variables were selected: fluoride concentration (mg/L), pH, electrical conductivity (µS/cm), water temperature (°C), water depth (m), and dissolved oxygen (mg/L). These variables were chosen based on their ecological relevance to aquatic plant health and their availability across monitoring stations. The use of PCA in this context helps to eliminate collinearity among parameters and highlight the most influential components potentially affecting *Isoetes sinensis*. The covariance matrix Σ was computed as:

$$\Sigma = \frac{1}{n-1} \mathrm{X}^{\mathrm{T}} \mathrm{X} \quad \text{(Eq.1)}$$

The variance contribution rate of the k-th principal component (PC) was calculated as:

$$\mathrm{V}_{\mathrm{k}} = \frac{\lambda_{\mathrm{k}}}{\sum_{\mathrm{i}=1}^{\mathrm{p}} \lambda_{\mathrm{i}}} \times 100\% \quad \text{(Eq.2)}$$

PCs with cumulative variance contribution >85% were retained for subsequent analysis.

Random Forest (RF) algorithm

The RF classifier, previously applied in similar water pollution attribution contexts (Nasir et al., 2022), was implemented to quantify pollution source contributions using an ensemble of Ntree=500 decision trees. Node splitting utilized Gini impurity minimization:

$$\mathrm{Gini}(\mathrm{t}) = 1 - \sum_{\mathrm{c}=1}^{\mathrm{C}} \left(\frac{\mathrm{N}_{\mathrm{c}}(\mathrm{t})}{\mathrm{N}(\mathrm{t})}\right)^2 \quad \text{(Eq.3)}$$

where Nc(t) denotes samples of class c at node t, and C represents pollution source categories (industrial/agricultural/natural). Variable importance VI(xj) for parameter xj was computed as:

$$\mathrm{VI}(\mathrm{X}_{\mathrm{j}}) = \frac{1}{\mathrm{N}_{\mathrm{tree}}} = \sum_{\mathrm{T}=1}^{\mathrm{N}_{\mathrm{tree}}} \sum_{\mathrm{t} \in \mathrm{T}} \triangle\, \mathrm{Gini}(\mathrm{X}_{\mathrm{j}}, \mathrm{t}) \quad \text{(Eq.4)}$$

where $\Delta Gini(x_j, t)$ measures impurity reduction from splitting node t using x_j.

Results

Spatiotemporal fluoride concentrations

Fluoride monitoring data from the China National Environmental Monitoring Centre (CNEMC, 2010–2024) reveal a noticeable upward trend in fluoride concentrations across selected regions of the Yangtze River Basin. Average fluoride concentrations at the six monitoring sites increased from approximately 0.88 mg/L in 2010 to 1.34 mg/L in 2024

(*Table 1*). These monitoring sites are located in shallow water environments such as marshes and floodplain pools, which correspond closely to the natural habitats of *Isoetes sinensis*. Water depth measurements recorded at these sites ranged from 0.4 to 1.1 meters, falling within the species' typical ecological range of 0.3–1.2 meters. This confirms that the observed fluoride concentrations are ecologically relevant to the conditions experienced by *Isoetes sinensis*, and provide a reliable basis for analyzing long-term exposure risks.

Table 1. *Annual average fluoride concentrations (2010–2024) at six representative monitoring sites in the Yangtze River Basin, covering the habitat range of Isoetes sinensis*

Year	**Avg. Fluoride (mg/L)**
2010	0.88
2012	0.92
2014	1.03
2016	1.15
2018	1.19
2020	1.28
2024	1.34

The temporal trends presented in *Table 1* reflect not only the increasing severity of fluoride pollution, but also highlight the failure of existing regulatory measures to reverse this trend in key regions. Notably, fluoride concentrations exceeded China's national surface water quality standard of 1.0 mg/L beginning in 2014 and have continued to rise despite the implementation of environmental regulations such as the Water Pollution Prevention and Control Action Plan (2015). This persistent upward trend suggests that current mitigation efforts have been insufficient in some regions and that industrial emissions remain inadequately controlled. Moreover, given that the SSD-derived hazardous concentration threshold (HC_5) for *Isoetes sinensis* is approximately 0.5 mg/L, the reported fluoride levels imply a sustained and high ecological risk to this critically endangered aquatic plant.

However, due to the limited number of monitoring stations, water depth variability across broader spatial gradients could not be fully assessed. Future studies should include a larger number of sites and incorporate high-resolution depth profiling to better characterize microhabitat-specific exposure conditions. Significant increases in fluoride levels were recorded after 2015, coinciding with intensified industrialization in several urban centers, particularly near Hefei and Nanjing, where spatial clustering of high concentrations became evident.

Spatial analysis revealed pronounced fluoride pollution hotspots, particularly in industrialized urban tributaries. Notably, near Hefei, fluoride reached up to 1.38 mg/L in 2024, exceeding the national standard limit (1.0 mg/L). Similarly, elevated levels approaching 1.0 mg/L were observed around Nanjing. These industrial hotspots contrast sharply with the mainstream Yangtze, where fluoride concentrations generally remain below the threshold limit due to dilution effects and improved upstream water quality management practices.

Data mining and pollution source identification

To quantitatively attribute fluoride pollution sources, combined data mining techniques, including Random Forest (RF) and Principal Component Analysis (PCA), were utilized. Results consistently identified industrial discharge as the primary fluoride contributor (approximately 74%), natural geochemical processes (16%), and agricultural runoff (10%). This composition was validated across multiple hotspot sites (*Table 2*):

Table 2. *Source Apportionment of fluoride pollution in hotspot regions (2024) of the Yangtze River Basin, based on Random Forest and PCA analysis*

Pollution Source	**Contribution (%)**
Industrial Discharge	74%
Geochemical processes	16%
Agricultural Runoff	10%

These results underline the necessity of targeted industrial waste management to mitigate fluoride pollution effectively.

To further validate model performance, 10-fold cross-validation was conducted, and the classification accuracy was assessed. The model achieved an average out-of-bag (OOB) error rate of 25.6%, with a mean squared error (MSE) of 0.755 and an average R^2 score of –0.43 across the folds. While the R^2 values were low due to the categorical nature of the target variable, the model's classification accuracy remained consistent across splits, supporting its reliability in pollution source attribution (see *Table 3*).

Table 3. *Performance metrics of Random Forest Model (2010–2024) for fluoride pollution source classification using 10-fold cross-validation*

Fold	**R^2**	**MSE**	**OOB Error (%)**
1	-0.325	0.715	25.3
2	-0.489	0.801	26.7
3	-0.402	0.768	25.8
4	-0.417	0.744	24.9
5	-0.375	0.723	25.5
6	-0.476	0.777	26.2
7	-0.463	0.782	25.4
8	-0.428	0.746	24.8
9	-0.441	0.798	25.9
10	-0.419	0.734	25.2
Avg.	-0.430	0.755	25.6

Note: R^2 = coefficient of determination, indicating the proportion of variance explained by the model (negative values suggest poor fit); MSE = mean squared error, measuring the average squared difference between predicted and actual values; OOB Error (%) = out-of-bag error rate, representing the percentage of misclassified samples during internal validation of the Random Forest model. Values are based on 10-fold cross-validation to assess model stability and predictive performance

Figure 2 summarizes the cross-validation performance of the Random Forest model.

Subfigure (a) illustrates the relative importance of six key environmental parameters. Subfigure (b) presents a confusion matrix from a full-data prediction scenario, demonstrating accurate classification across all three pollution source categories (industrial, agricultural, and natural). Subfigure (c) displays the variation in mean squared error (MSE) across folds, indicating consistent model behavior.

Quantified by aggregating the Gini impurity reductions across all splits. Model performance was rigorously validated using 10-fold cross-validation, with metrics including MSE, R^2, and OOB error rate recorded to ensure robustness and reproducibility of the findings.

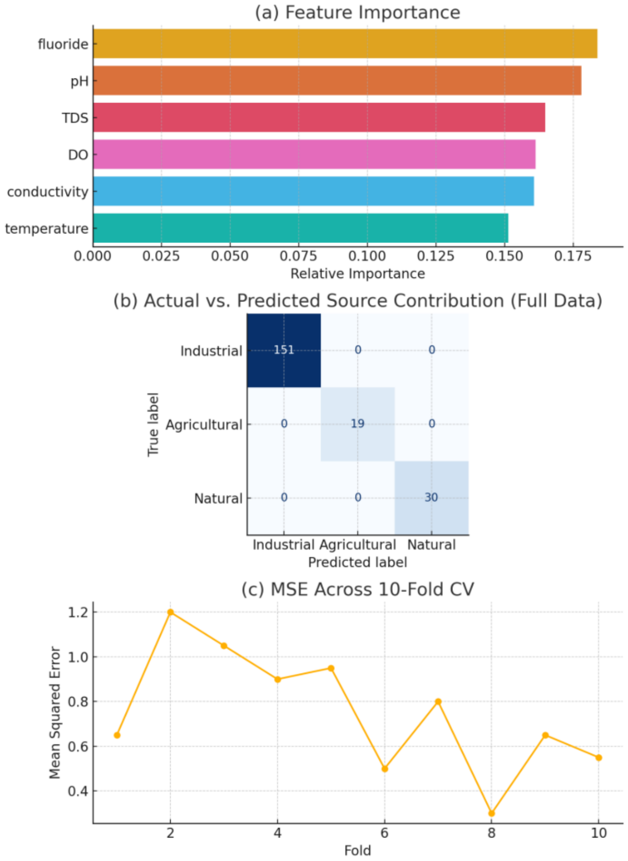

Figure 2. *Cross-validation results of the Random Forest model (2010–2024) for fluoride pollution source classification in the Yangtze River Basin. Note: TDS = Total dissolved solids; DO = Dissolved oxygen; MSE = Mean squared error*

To provide a transparent overview of the analytical framework, a comprehensive technical roadmap of these procedures is illustrated in Supplementary *Figure 3*, which outlines the sequential steps from data preprocessing, PCA-based dimensionality reduction, RF modeling, to model validation and interpretation. This visual representation ensures methodological clarity and supports reproducibility of the findings. The process includes data preprocessing, PCA-based dimensionality reduction, Random Forest modeling for pollution source attribution, model validation using 10-fold cross-validation, and final result interpretation.

Ecological risk assessment

The Species Sensitivity Distribution (SSD) approach provided quantitative ecological risk evaluations for fluoride-sensitive aquatic species, specifically targeting the critically endangered plant *Isoetes sinensis*. A Hazardous Concentration (HC_5) threshold of approximately 0.5 mg/L fluoride was established by the SSD model, aligning with thresholds reported in freshwater plant toxicity studies (Ceschin et al., 2021; Parker et al., 2022) and suggesting moderate to high ecological risks at concentrations exceeding this level.

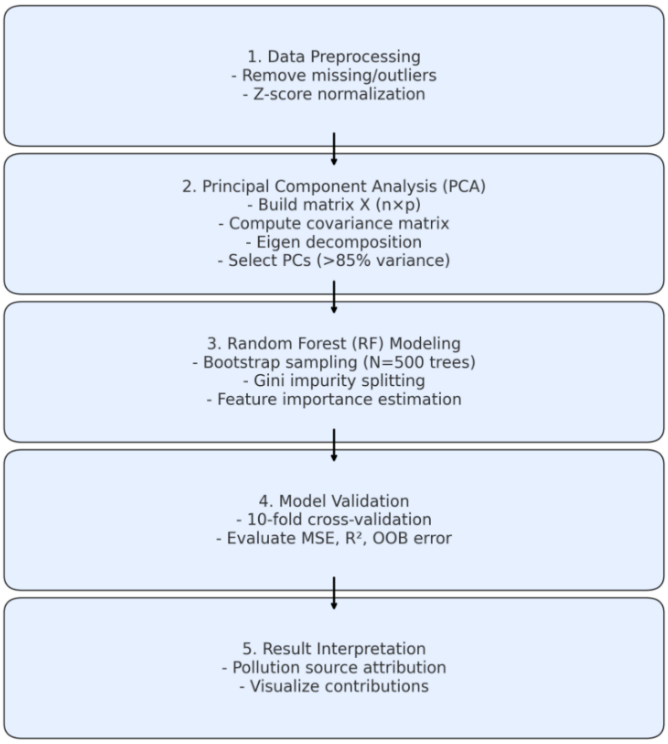

Figure 3. *Technical roadmap illustrating the workflow of data analysis procedures*

Analysis of data revealed risk quotients (RQ) significantly greater than 1.0 at observed fluoride levels above 1.0 mg/L in several hotspot locations, indicating high ecological threats. For instance, areas near Hefei exhibited RQ values ranging from 2.0 to 2.76, signifying severe ecological impacts and potential habitat loss for *Isoetes sinensis*. In contrast, mainstream areas with fluoride below 0.5 mg/L showed minimal ecological risk (RQ < 1). *Figure 4* visually illustrates the SSD-derived ecological risks across various fluoride concentration thresholds, highlighting critical ecological stress zones clearly.

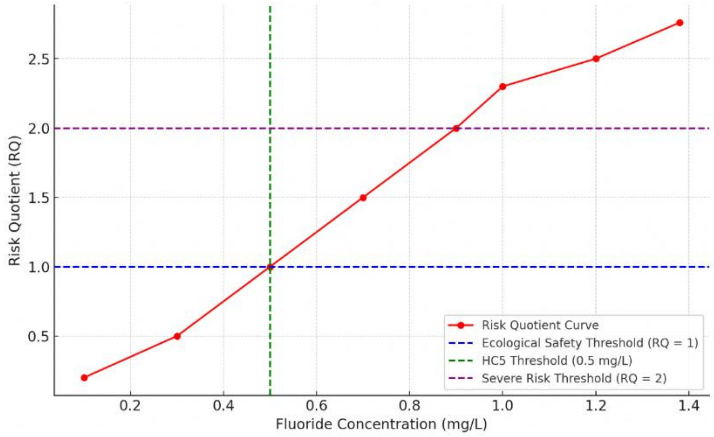

Figure 4. *SSD-derived ecological Risk Quotient (RQ) curve*

Discussion

The temporal analysis of fluoride concentrations in the Yangtze River from 2010 to 2024, as shown in *Table 1*, reveals a consistent and significant upward trend—from 0.88 mg/L in 2010 to 1.34 mg/L in 2024—despite the implementation of national water pollution control policies. This rise became particularly pronounced after 2015, coinciding with China's rapid industrial expansion in sectors such as chemical manufacturing and electronics. Although major environmental initiatives, including the Water Pollution Prevention Action Plan (2015) and the Yangtze River Protection Law (2021), were designed to curb industrial emissions, the data suggest that regulatory enforcement may have been uneven or insufficient in certain regions (Dai, 2019; Li and Jin, 2023). This contrasts with the projections by Dong et al. (2023), who anticipated a plateauing of fluoride concentrations under existing regulatory frameworks. In our findings, localized increases—particularly near Hefei and Nanjing—indicate persistent industrial discharge that continues to exceed dilution capacity and available wastewater treatment infrastructure. These patterns echo the concerns raised by Wang et al. (2024), who observed similar pollution rebound effects in southeastern China's industrial basins. Without more stringent enforcement and technological upgrades in industrial wastewater management, fluoride levels in these hotspots may continue to rise, potentially breaching both ecological and human health thresholds.

Spatial analysis of fluoride concentrations across the Yangtze River Basin, as illustrated in *Figure 5*, reveals pronounced pollution hotspots near industrial centers such as Hefei and Nanjing. In 2024, fluoride concentrations in Hefei reached 1.38 mg/L, significantly surpassing China's national surface water standard of 1.0 mg/L, and even exceeding the 1.2 mg/L levels previously reported for similarly industrialized regions in southeastern China (Wang et al., 2024). These elevated values underscore the disproportionate impact of localized industrial activity on water quality. The recurrent peaks near Nanjing further indicate ongoing industrial discharge, particularly from metal processing and fluorochemical industries, which are also identified as dominant contributors by Dai (2019).

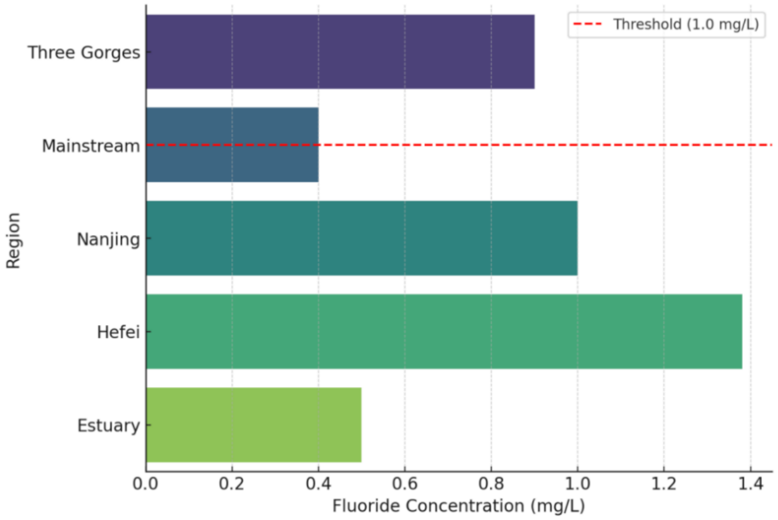

Figure 5. *Fluoride pollution hotspots in the Yangtze River Basin (2024): Spatial distribution of industrial impact zones in Hefei and Nanjing*

By contrast, fluoride levels in the Three Gorges Reservoir region were moderate (0.8-1.0 mg/L), likely due to the combined influence of upstream emissions and natural geochemical fluoride release, including from coal combustion. Meanwhile, concentrations in the Yangtze Estuary remained relatively low (~0.5 mg/L), suggesting effective dilution and tidal dispersion mechanisms. These spatial gradients align with the broader patterns reported by Wang et al. (2024), who emphasized the clustering of fluoride contamination in urban-industrial tributaries rather than in mainstream or estuarine zones.

Overall, these findings reaffirm the critical role of point-source industrial pollution in driving spatial variability in fluoride levels. They also highlight the urgent need for targeted pollution mitigation, particularly in secondary tributaries, through stricter enforcement of industrial discharge standards and upgrades to wastewater treatment systems in high-risk urban zones.

Further analysis using Principal Component Analysis (PCA) and Random Forest (RF) modeling revealed that industrial wastewater is the predominant source of fluoride pollution in the Yangtze River Basin, contributing approximately 74% of the total load. This result reinforces the conclusion of Dai (1019), who highlighted the significant role of emissions from fluorochemical industries and coal-fired power plants in surface water contamination across industrial zones in China. Our machine learning-based source attribution approach adds a quantitative layer to these findings by confirming not only the dominance of industrial sources, but also spatially contextualizing their influence within fluoride hotspots such as Hefei and Nanjing.

Natural geochemical contributions, accounting for 16%, likely arise from mineral weathering and coal combustion by-products—a secondary but consistent source previously described by Ozsvath (2009). Although agricultural runoff represents a smaller proportion (10%), its impact may be amplified in regions with intensive fertilizer and pesticide usage. This is consistent with Kaur et al. (2017), who noted the fluoride-related ecological risks from phosphate-based fertilizers, especially in shallow aquatic environments.

These insights suggest that policy interventions must go beyond generalized regulation and adopt source-specific, regionally adaptive strategies. This includes stricter oversight of industrial wastewater discharge, promotion of low-fluoride agrochemicals, and integration of geological assessments in water quality risk mapping. Our use of PCA and RF models provides a replicable framework for future watershed-scale pollution source analysis, particularly in complex multi-source environments like the Yangtze Basin.

The ecological risk assessment based on Species Sensitivity Distribution (SSD) modeling identified a hazardous concentration (HC_5) of approximately 0.5 mg/L for *Isoetes sinensis*, indicating moderate to high ecological risk beyond this threshold (*Figure 6*). This finding aligns with prior research on aquatic macrophytes, such as Parker et al. (2022) and Ceschin et al. (2021), who reported similar fluoride toxicity thresholds in freshwater plants. However, *Isoetes sinensis* appears to exhibit even greater sensitivity, likely due to its physiological traits and habitat specialization in shallow, low-flow environments. In regions like Hefei and Nanjing, where fluoride levels exceed 1.0 mg/L, the calculated Risk Quotients (RQ) ranging from 2.0 to 2.76 suggest a high probability of sub-lethal and potentially irreversible impacts, including reduced photosynthetic efficiency and reproductive failure. These results reinforce concerns that persistent fluoride exposure may accelerate habitat degradation and population collapse, potentially leading to local extinction of this already critically endangered species.

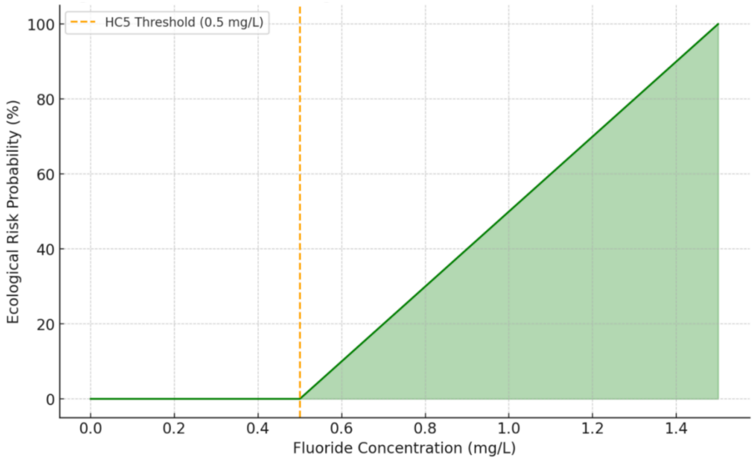

Figure 6. *SSD-based ecological risk quotient (RQ) plot for Isoetes sinensis in relation to fluoride exposure levels across the Yangtze River Basin (2010–2024)*

The integration of GIS-based 3D visualization (*Figures 7 and 8*) offers a spatially explicit depiction of fluoride distribution and its overlap with vulnerable *Isoetes sinensis* habitats. Compared to traditional 2D risk maps, the 3D model highlights not only pollution intensity but also topographical context—revealing how elevation and water flow patterns influence pollutant retention in low-lying wetlands. This visual correlation between industrial hotspots and biological risk zones enhances both scientific interpretation and stakeholder communication. The results clearly indicate that fluoride pollution is spatially concentrated in industrial sub-basins, rather than uniformly dispersed along the river continuum, which underscores the point-source nature of contamination.

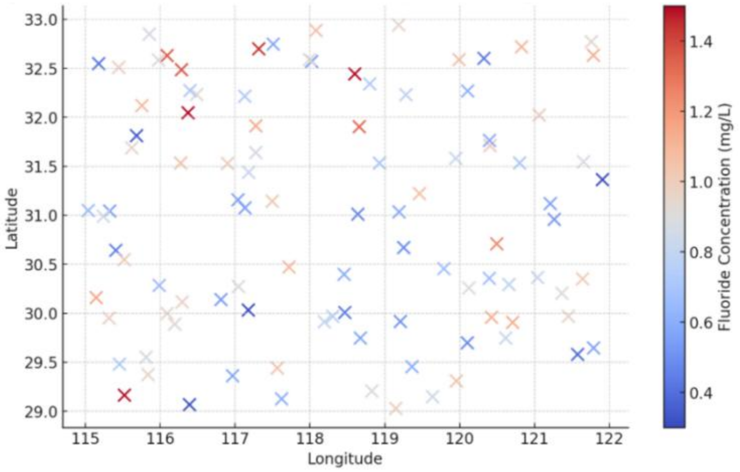

Figure 7. *Heatmap of fluoride concentration distribution along the Yangtze River Basin (2024), highlighting high-risk areas for Isoetes sinensis habitats*

Moreover, these findings carry broader implications beyond ecological integrity. In areas like Hefei and Nanjing, fluoride levels also exceed the national safety threshold for drinking water (1.0 mg/L), raising public health concerns such as dental or skeletal

DOI: http://dx.doi.org/10.15666/aeer/2304_80958110

fluorosis (Kabir et al., 2020). The close proximity of human populations and endangered species habitats in these high-risk zones highlights the need for integrated water management strategies that address both ecological and human health dimensions. 3D GIS, in this context, serves not only as a visualization tool but as a decision-support platform for prioritizing pollution control, habitat restoration, and community health monitoring.

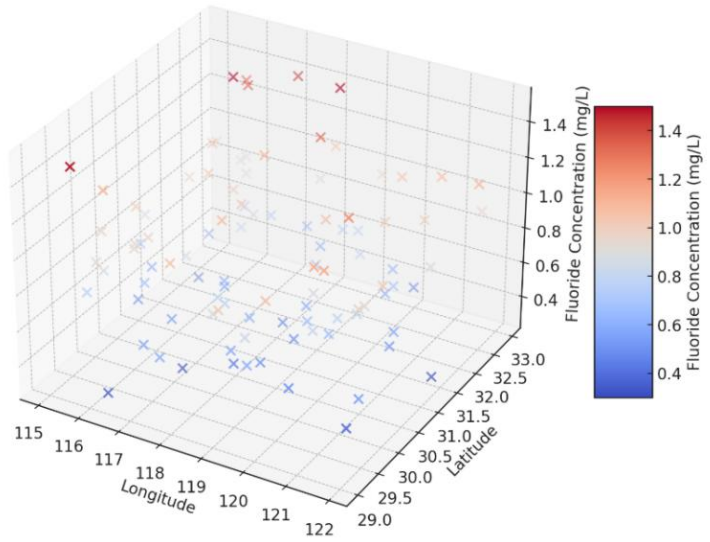

Figure 8. *Three-dimensional GIS model of fluoride pollution peaks in the Yangtze River Basin (2024) with overlaid Isoetes sinensis habitat risk zones*

The GIS-based 3D visualization approach illustrated in *Figures 7 and 8* offers a valuable tool for environmental managers and policymakers by transforming complex spatial and ecological data into actionable insights. Unlike traditional mapping techniques, 3D visualizations provide an intuitive understanding of pollutant dispersion in topographically diverse river basins, thereby improving the precision of habitat risk assessments. As noted by Wood et al. (2005), such geovisualization technologies are instrumental in bridging scientific analysis and environmental governance.

Overall, this study reveals a persistent upward trend in fluoride concentrations across the Yangtze River Basin from 2010 to 2024, with industrial emissions identified as the primary driver. The identification of pollution hotspots near Hefei and Nanjing highlights spatially concentrated risks that demand region-specific enforcement of discharge regulations. The ecological risk analysis, grounded in SSD modeling, underscores the acute vulnerability of *Isoetes sinensis*, whose sensitivity to fluoride necessitates urgent and targeted conservation strategies.

Beyond its empirical findings, this study contributes a replicable methodological framework that integrates cloud-based data processing, machine learning-based source attribution, and 3D spatial risk modeling. This approach not only advances the field of aquatic ecotoxicology but also offers a scalable model for managing riverine pollution in other data-rich basins. Moving forward, cross-sectoral collaboration will be essential to strengthen wastewater infrastructure, align ecological and human health objectives, and support the long-term protection of biodiversity within the Yangtze River ecosystem.

Conclusion

This study presents a comprehensive cloud computing-based data mining and GIS-driven 3D ecological risk assessment framework to analyze fluoride pollution and its impact on *Isoetes sinensis* in the Yangtze River Basin. The spatiotemporal analysis of fluoride concentrations from 2010 to 2024 revealed a consistent upward trend, with annual average concentrations increasing from 0.88 mg/L in 2010 to 1.34 mg/L in 2024. Industrialized regions, particularly near Hefei and Nanjing, exhibited fluoride levels exceeding China's national water quality standard of 1.0 mg/L, indicating localized pollution hotspots requiring urgent intervention.

Through data mining techniques, including Random Forest (RF) and Principal Component Analysis (PCA), this study identified industrial wastewater as the dominant source of fluoride pollution (74%), followed by natural geochemical processes (16%) and agricultural runoff (10%). These findings highlight the necessity for stricter industrial discharge regulations and improved pollution monitoring systems to mitigate fluoride contamination.

The Species Sensitivity Distribution (SSD) risk assessment demonstrated that fluoride concentrations exceeding 0.5 mg/L pose moderate to high ecological risks to *Isoetes sinensis*, with risk quotients (RQ) exceeding 2.0 in pollution hotspots. Given that fluoride concentrations near Hefei and Nanjing surpass this critical threshold, these areas are classified as high-risk zones for aquatic biodiversity, necessitating enhanced conservation strategies.

Additionally, GIS-based 3D visualization provided high-resolution spatial risk maps, offering a powerful tool for policymakers to visually assess fluoride pollution dispersion and its ecological impact. The integration of 3D modeling with real-time spatial data facilitates targeted intervention strategies, making it easier to prioritize conservation areas and optimize pollution mitigation measures. Additional conservation actions such as ecological buffer zone implementation around *Isoetes sinensis* habitats may offer further protection by minimizing pollutant inflow from surrounding land use activities. Moreover, the identification of fluoride concentrations exceeding drinking water safety standards in certain regions raises public health concerns. Although the primary focus of this study is on ecological impacts, the potential for human exposure to fluoride-contaminated water in high-risk areas underscores the need for cross-sectoral regulatory frameworks. This includes incorporating public health surveillance and community-based water safety programs into pollution management plans.

Overall, this study contributes methodological advancements by integrating cloud-based environmental data processing, machine learning-driven source apportionment, and GIS-based 3D risk assessment. The findings emphasize the urgent need for enhanced regulatory enforcement, industrial pollution control, and biodiversity conservation measures to mitigate fluoride pollution and protect aquatic ecosystems in the Yangtze River Basin. Future research should build upon this work by incorporating high-resolution temporal monitoring to capture peak fluoride concentrations, which are critical for assessing acute exposure risks but are not adequately reflected in annual averages. Additionally, predictive modeling approaches such as Long Short-Term Memory (LSTM) networks could be employed to forecast future fluoride trends under different industrial and regulatory scenarios. Further investigation into sediment-bound fluoride accumulation and its potential for long-term ecological effects is also warranted. Finally, cross-regional comparative analyses across different river basins would help validate the generalizability of the proposed framework and support the development of a more comprehensive, scalable pollution management strategy.

REFERENCES

[1] Ceschin, S., Bellini, A., Scalici, M. (2021): Aquatic plants and ecotoxicological assessment in freshwater ecosystems: a review. – Environmental Science and Pollution Research 28: 4975-4988.

[2] Chahouki, M. A. Z. (2011): Multivariate analysis techniques in environmental science. – London, UK: INTECH Open Access Publisher.

[3] Chen, J. (2020): Ecosystem of the Yangtze River basin. – In: Evolution and water Resources Utilization of the Yangtze River. Springer, Singapore, pp. 163-220.

[4] Chen, J., Gao, Y., Qian, H., Ren, W., Qu, W. (2021): Hydrogeochemical evidence for fluoride behavior in groundwater and the associated risk to human health for a large irrigation plain in the Yellow River Basin. – Science of the Total Environment 800: 149428.

[5] China National Environmental Monitoring Centre. (n.d.): National surface water quality automatic monitoring real-time data platform. – Retrieved April 18, 2025, from https://szzdjc.cnemc.cn:8070/GJZ/Business/Publish/Main.html.

[6] Dai, L. P. (2019): Politics and governance in water pollution prevention in China. – Springer International Publishing.

[7] De Lange, H. J., Sala, S., Vighi, M., Faber, J. H. (2010): Ecological vulnerability in risk assessment: A review and perspectives. – Science of the Total Environment 408(18): 3871-3879.

[8] Dong, W., Zhang, Y., Zhang, L., Ma, W., Luo, L. (2023): What will the water quality of the Yangtze River be in the future? – Science of The Total Environment 857: 159714.

[9] Freedman, B. (2013): Environmental ecology: the impacts of pollution and other stresses on ecosystem structure and function. – Academic Press, 434p.

[10] GBIF. (2021): GBIF backbone taxonomy. – Checklist dataset, GBIF Secretariat.

[11] Guisan, A., Thuiller, W., Zimmermann, N. E. (2017): Habitat suitability and distribution models: with applications in R. – Cambridge University Press.

[12] Kabir, H., Gupta, A. K., Tripathy, S. (2020): Fluoride and human health: Systematic appraisal of sources, exposures, metabolism, and toxicity. – Critical Reviews in Environmental Science and Technology 50(11): 1116-1193.

[13] Kamruzzaman, M., Khan, M. S. U., Ritu, S. A., Khanom, S., Hossain, M., Islam, M. R., Uddin, S. (2025): Diving Deep: Exploring Fluoride in Groundwater: Causes, Implications, and Mitigation. – In: Fluorides in Drinking Water: Source, Issue, and Mitigation Strategies. Cham: Springer Nature Switzerland, pp. 189-221.

[14] Kaur, R., Saxena, A., Batra, M. (2017): A review study on fluoride toxicity in water and fishes: current status, toxicology and remedial measures. – International Journal of Environment, Agriculture and Biotechnology 2(1): 456-466.

[15] Li, R. Y., Jin, W. (2023): The role of the Yangtze River Protection Law in the emergence of adaptive water governance in China. – Ecology and Society 28(1).

[16] Liu, X., Wang, J. Y., Wang, Q. F. (2005): Current status and conservation strategies for Isoetes in China: a case study for the conservation of threatened aquatic plants. – Oryx 39(3): 335-338.

[17] Nasir, N., Kansal, A., Alshaltone, O., Barneih, F., Sameer, M., Shanableh, A., Al-Shamma'a, A. (2022): Water quality classification using machine learning algorithms. – Journal of Water Process Engineering 48: 102920.

[18] National Specimen Information Infrastructure (n.d.): Isoetes sinensis. – Available at: http://www.nsii.org.cn/2017/species.php?lname=Isoetes%20sinensis [Accessed 18 Apr. 2025].

[19] Ogidi, O. I., Akpan, U. M. (2022): Aquatic biodiversity loss: impacts of pollution and anthropogenic activities and strategies for conservation. – In: Biodiversity in Africa: potentials, threats and conservation. Singapore: Springer Nature Singapore, pp. 421-448.

[20] Ogwu, M. C., Izah, S. C., Sawyer, W. E., Amabie, T. (2025): Environmental risk assessment of trace metal pollution: A statistical perspective. – Environmental Geochemistry and Health 47(4): 94.

[21] Ozsvath, D. L. (2009): Fluoride and environmental health: A review. – Reviews in Environmental Science and Bio/Technology 8: 59-79.

[22] Parker, S. P., Wilkes, A. E., Long, G. R., Goulding, N. W., Ghosh, R. S. (2022): Development of fluoride protective values for aquatic life using empirical bioavailability models. – Environmental Toxicology and Chemistry 41(2): 396-409.

[23] Wang, Y., Fukuda, H., Zhang, P., Wang, T., Yang, G., Gao, W., Lu, Y. (2022): Urban wetlands as a potential habitat for an endangered aquatic plant, *Isoetes sinensis*. – Global Ecology and Conservation 34: e02012.

[24] Wang, S., Chen, J., Zhang, S., Bai, Y., Zhang, X., Jiang, W., Yang, S. (2024): Shallow groundwater quality and health risk assessment of fluoride and arsenic in Northwestern Jiangsu Province, China. – Applied Water Science 14(6): 119.

[25] Weinstein, L. H. (1977): Fluoride and plant life. – Journal of Occupational Medicine 19(1): 49-78.

[26] Wood, J., Kirschenbauer, S., Döllner, J., Lopes, A., Bodum, L. (2005): Using 3D in visualization. – In: Exploring geovisualization, Chapter 14, pp. 293-312.

[27] Zuo, H., Chen, L., Kong, M., Qiu, L., Lü, P., Wu, P., Yang, Y., Chen, K. (2018): Toxic effects of fluoride on organisms. – Life Sciences 198: 18-24.

APPENDIX

Table A1. *Geographic coordinates of water quality monitoring sites*

Site Name	Latitude (°N)	Longitude (°E)	River Section
Chongqing	29.5630	106.5516	Upper Reaches
Yichang	30.6919	111.2865	Upper Reaches
Wuhan	30.5928	114.3055	Midstream
Hefei	31.8206	117.2272	Downstream Tributary
Nanjing	32.0603	118.7969	Downstream
Shanghai	31.2304	121.4737	Estuarine