remote sensing

MDPI

Article

A UAV-Assisted FP-xLSTM Framework for Spatio-Temporal Prediction of Multi-Depth Urban River Water Quality

Yishan Zhang [1,2,*] and Ruijun Chen [3]

[1] The Academy of Digital China, Fuzhou University, Fuzhou 350108, China
[2] Institute of Remote Sensing and Geographic Information System, Peking University, Beijing 100871, China
[3] Hongyou Software Co., Ltd., Beijing 100028, China; rjchy0728@163.com
* Correspondence: ys_zhang@pku.edu.cn

Highlights

What are the main findings?

- A feature propagation-based deep learning framework has been developed to estimate the vertical distribution of WQPs and achieve accurate multi-depth prediction with limited samples.
- Fine-grained spatio-temporal patterns of WQPs are quantified across three regions, achieving average MAPEs ranging from 9.33% to 19.06% at multiple depths.

What are the implications of the main findings?

- The proposed approach reduces sample dependence and enables efficient dynamic monitoring of water quality variations across different depths.
- Differentiated modeling strategies for optical active and non-optical active constituents improve prediction reliability and provide a scalable solution for comprehensive water quality management.

Abstract

Effective water quality monitoring is essential for protecting aquatic environments and supporting sustainable management of urban rivers. Accurate prediction and visualization of water quality parameters (WQPs) distributions across spatial, temporal, and vertical dimensions remained challenging for unmanned aerial vehicle (UAV)-based remote sensing approaches, which often neglected depth-wise variability and required extensive water sampling to maintain stable prediction performance for multiple WQPs. To address these limitations, this study proposed a feature propagation-based extended long short-term memory framework (FP-xLSTM) for unified prediction and visualization of multi-depth WQPs, including total phosphorus (TP), total nitrogen (TN), chlorophyll-a (Chl-a), chemical oxygen demand (COD), biochemical oxygen demand (BOD), and total suspended solids (TSS). By integrating spectral characteristics with spatio-temporal dependencies through feature propagation across consecutive temporal observations, FP-xLSTM effectively captured short-term fluctuations and learned vertical distribution patterns of heterogeneous WQPs from multi-source data while reducing sampling requirements. A field study with an eight-month interval between the first and last sampling campaigns was conducted to evaluate the proposed framework under cross-region experimental settings and against state-of-the-art baseline models. Results demonstrated that FP-xLSTM consistently outperformed competing methods, achieving average coefficients of determination (R^2) ranging from 0.66 to 0.95 and mean absolute percentage errors (MAPE) between 9.33% and 19.06%. Furthermore, the framework has potential for deployment in practical environmental monitoring applications to support local environmental authorities in tracking river water

Check for updates

Academic Editor: Dong Liu

Received: 29 August 2026
Revised: 2 October 2026
Accepted: 13 October 2026
Published: 16 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

quality dynamics and optimizing management strategies. Overall, the proposed method provided an effective and scalable intelligent framework for UAV-assisted urban water quality assessment and multi-depth environmental visualization.

Keywords: water quality parameters; hyperspectral images; long and short-term memory; spatio-temporal analysis; feature propagation

1. Introduction

Urban rivers constitute critical components of terrestrial aquatic systems and provide essential water resources for agricultural production, industrial activities, and human daily life [1]. However, rapid urbanization and industrialization have significantly degraded urban water environments, resulting in increasing nutrient enrichment and ecological deterioration [2]. Excessive accumulation of total phosphorus (TP) and total nitrogen (TN) from agricultural runoff, industrial discharge, and domestic wastewater frequently induced eutrophication and promoted rapid phytoplankton proliferation. Consequently, chlorophyll-a (Chl-a), a key indicator of phytoplankton biomass, often exhibited substantial increases in eutrophic aquatic systems [3,4]. Certain phytoplankton species might further trigger harmful algal blooms (HABs), releasing cyanotoxins that threatened both aquatic ecosystems and human health through ingestion, inhalation, or direct exposure. In addition, total suspended solids (TSS) strongly influenced underwater light penetration and habitat conditions, while also affecting algal growth dynamics through interactions with Chl-a concentrations [2]. Therefore, accurate monitoring and prediction of water quality parameters (WQPs) were essential for environmental protection and sustainable urban river management.

Fixed monitoring stations and manual water sampling approaches could provide reliable point-based measurements, but they were constrained by high labor and maintenance costs, limited spatial coverage, and insufficient temporal resolution [5]. They utilized statistical methods and analyzed WQPs results from some sampling points to estimate approximate distribution patterns of WQPs, which lacked theoretical foundation and usually deviated from pragmatic WQPs distributions. Satellite remote sensing technologies have consequently emerged as effective alternatives for large-scale and dynamic water quality monitoring [6,7]. Existing satellite remote sensing approaches estimated optically active constituents (OACs) WQPs such as Chl-a and TSS by featuring spectral band combinations with an average coefficient of determination (R^2) over 0.60 and mean absolute percentage error (MAPE) less than 30%. They usually suffered from inadequate physical and generalization foundations and needed abundant prior knowledge. A nonlinear multivariate correlation model was employed to predict Chl-a and TSS concentrations from Sentinel-2 with R^2 over 0.80 but suffered prediction instability in various regions [8]. Some methods utilized indirect inference of non-optically active constituents (non-OACs), including TP, TN, chemical oxygen demand (COD), and biochemical oxygen demand (BOD), using statistical relationships with spectral band combinations and correlated OACs, obtaining relatively low R^2 and high MAPE [9,10]. An extreme gradient boosting (BST) machine learning framework was proposed involving environmental factors analysis to retrieve TP concentration from different spectral band ratios of Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua level-1A images, obtaining R^2 of 0.31 and MAPE of 52.54% [11]. A two-stage multivariate regression model was established incorporating WQPs correlation and difference to retrieve TP, TN, and COD from satellite images but suffered R^2 around 0.60 and high MAPE due to low model robustness [12]. However, the

prediction of non-OACs remained challenging because their spectral responses were often indirect and highly nonlinear. Spatial distribution calculation accuracy of non-OAC WQPs remained unstable, causing low efficiency in tracking various non-OACs WQPs and water quality management.

Satellite-based remote sensing was restricted by long revisit cycles, limited temporal continuity, and insufficient spectral flexibility, hindering the capture of fine-grained temporal variations and subtle spectral changes associated with dynamic WQPs [13,14]. The relatively broad spectral bands of Landsat Operational Land Imager (OLI) lacked the fine spectral resolution and dedicated red-edge bands available in hyperspectral sensors. Unmanned aerial vehicle (UAV)-based remote sensing enabled high-frequency observations with flexible deployment and finer spatial resolution, making it increasingly suitable for urban water quality monitoring [15,16]. UAV-based remote sensing involving more spectral bands captured more spectral information and feature correlation, which enabled spectral-spatial relationships and temporal variations of WQPs to be modeled [17,18]. A hierarchical clustering analysis (HCA) framework was established with an excitation-emission matrix and parallel factor extraction to extract spectral features to retrieve Chl-a concentrations from UAV hyperspectral images, achieving R^2 of 0.82 and MAPE of 27.07% [19]. Deep learning-based UAV remote sensing enhanced correlation among WQPs through spatio-temporal distribution pattern learning and feature interaction, improving prediction accuracy and stability on non-OAC WQPs. Transformer and long short-term memory (LSTM) networks were integrated, introducing a cross-temporal attention mechanism and adaptive feature fusion module to enhance feature correlation, to predict TN, Chl-a, and TSS concentrations with R^2 over 0.9 and MAPE less than 26% [20]. Graph attention networks were proposed with LSTM (GAT-LSTM) to predict spatial distributions of TP and TN through inconsecutive sampling, which obtained average R^2 of 0.65 and MAPE of 15.91% [21]. These methods suffered from an inability to capture variation of WQPs at fine-grained time scales (e.g., day and month) and to predict WQPs in multiple depths.

Several critical challenges remained unresolved. First, most existing methods primarily focused on surface-level observations while neglecting vertical distributions of WQPs across multiple depths. Second, current models often exhibited limited generalization capability when simultaneously predicting both OACs and non-OACs within a unified framework. Third, deep learning-based UAV remote sensing was unable to capture tiny variations of WQPs in fine-grained time scales (e.g., day and month). To address these limitations, this study proposed a feature propagation-based extended long short-term memory framework (FP-xLSTM) for multi-depth (0.5 m, 2.0 m, 4.0 m) prediction and visualization of WQPs in spatio-temporal dimensions using UAV remote sensing data. The proposed framework enhanced temporal feature continuity through feature propagation across consecutive observations and captured subtle spatial variations in fine-grained temporal sequences. A multi-layer memory interaction mechanism was introduced to enable covariance-based information updating and collaborative feature learning among memory units across different layers. Furthermore, FP-xLSTM adopted differentiated learning strategies for OACs and non-OACs, improving unified prediction performance for heterogeneous WQPs. Spatial correlation networks and probability distribution transformation were additionally integrated to model depth-wise relationships and strengthen robustness in multi-depth prediction tasks. By jointly incorporating spectral characteristics, spatial dependencies, temporal dynamics, and feature interaction information, the proposed framework reduced dependence on large training sample sizes while improving prediction stability and generalization capability. The proposed method established an intelligent UAV-assisted water quality monitoring framework capable of visualizing multi-depth WQPs distributions on a monthly time scale. The framework provided effective

technical support for environmental management, pollution assessment, and early-warning systems in urban river environments.

2. Materials and Methods

2.1. Study Area

We conducted experiments on three watersheds in different regions as shown in Figure 1 including: (1) the main part of Yamen Watercourse in Jiangmen city, latitudinally ranging from 22°13′8″N to 22°21′39″N and longitudinally ranging from 113°3′43″E to 113°5′25″E, covering 19.79 km^2, which is bordered by the western estuary Port area of Xinhui in the north and Huangmao Sea in the south; (2) the main part of Huangyang River in Zhuhai city, latitudinally ranging from 22°14′10″N to 22°9′20″N and longitudinally ranging from 113°16′33″E to 113°19′13″E, covering 4.29 km^2, which starts from Minyusha and end at Jitimen; (3) the main part of Fen River in Taiyuan city, latitudinally ranging from 37°41′39″N to 37°48′55″N and longitudinally ranging from 112°30′46″E to 112°32′41″E, covering 4.04 km^2, which starts from Guancun Mountain and flows into Yellow River. Cross-region experiments on these three regions, used as study areas, were conducted to test the effectiveness and applicability of FP-xLSTM. These regions were representative (e.g., industrial, aquaculture, agricultural, and residential areas were densely distributed along the banks of these regions), where the water environment was contaminated in past years, causing rapid degradation and deterioration of water quality.

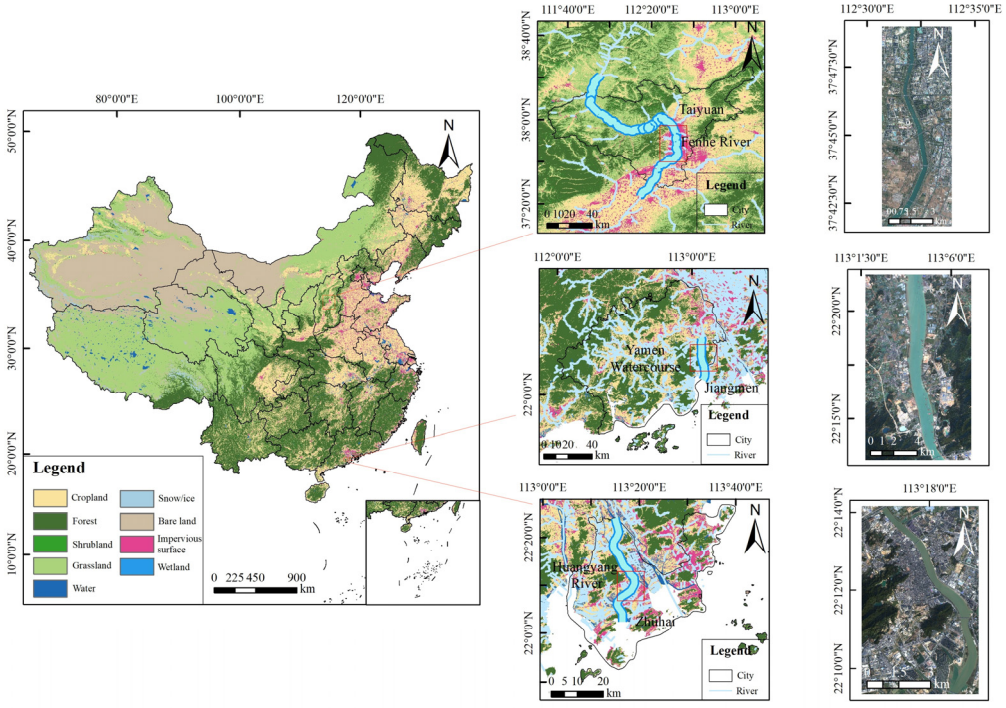


Figure 1. Study area. The red polygons cover cross-region watersheds.

2.2. Ground and UAV Data Sampling

In this study, ground and UAV data sampling were implemented simultaneously at each sampling site from March 2025 to November 2025, at two-month intervals in each region (Yamen Watercourse, Huangyang River, and Fen River). Water samples were collected in each region in each time period from March 2025 to November 2025 at three different depths. Groundwater sampling was implemented by an unmanned water sampling boat (UWSB). Three 500 mL water samples were collected from 0.5 m, 2.0 m, and 4.0 m below the water surface at each sampling site in each region, and laboratory chemical analysis

was performed on each water sample for six WQPs. Additionally, 125 water samples were collected at each depth in each period for each region, totaling 5625 samples (125 sampling points, 3 depths, 5 time periods in each of these three regions). Additionally, water samples were densely collected in relatively small regions and sparsely collected in relatively large regions so as to improve the self-adapting and transfer learning capability of the proposed method. By integrating these two sampling strategies with UAV hyperspectral remote sensing, intensive field measurements could be concentrated in representative areas while model-based predictions were extended to sparsely sampled regions, thereby reducing the dependence on large-scale and long-term field sampling. The corresponding chemical analysis methods were elaborated in Appendix A.2, and water quality classification standards were demonstrated in Table A1. Hyperspectral reflectance data collected by UAV-carried Gaia Sky-M imager covered 270 bands from 401 nm to 998 nm at each sampling site. The flying height of the DJI Matrice 600 Pro UAV was set to 120 m, and spatial resolution was resampled to 1.0 m. UAV hyperspectral data were calibrated by standard reference panels with reflectance of 0.1, 0.2, and 0.4.

2.3. Preprocessing of Hyperspectral Data

In this study, independent component analysis (ICA) was utilized to reduce medium interference such as dust and water vapor between the UAV and water surface to enhance the signal-to-noise ratio (SNR) [22]. The difference between UAV and ground reflectance was reduced. Thus, UAV-ground reflectance interpolation and transfer were unnecessary to denoise data and avoid quadratic error. The real source reflectance, representing the optical properties of the water body, was extracted as follows.

$$x_i = \sum_{j=1}^{k} \alpha_{i,j} s_{i,j} \tag{1}$$

$$\boldsymbol{x}_i = A_i \boldsymbol{s}_i \tag{2}$$

$$\hat{\mathbf{s}}_i = W_i^T \boldsymbol{x}_i \tag{3}$$

where x_i denotes the UAV-collected hyperspectral reflectance at wavelength i; $s_{i,j}$ denotes the source reflectance at wavelength i; $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{270})$; A is a mixing matrix which mixed components $\boldsymbol{s}_i$ including real source reflectance and medium interference; and $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_{270})$; $A = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{270})$; $\hat{\mathbf{s}}$ is estimated $\boldsymbol{s}$; $W = (W_1, W_2, \cdots, W_{270})$. The element with the greatest absolute value of coefficient was taken as the real source reflectance of $\hat{\mathbf{s}}_i$ due to smaller influence on collected reflectance by medium interference compared to optical properties. Then, the W iterative update process was formulated [23].

$$W^{(k+1)} = E\left\{ \boldsymbol{x} \frac{\partial G\left(W^{(k)^T} x \right)}{\partial W^{(k)}} \right\} - E\left\{ \boldsymbol{x} \frac{\partial^2 G\left(W^{(k)^T} x \right)}{\partial W^{(k)^2}} \right\} W^{(k)} \tag{4}$$

$$W^{(k)} = \frac{W^{(k+1)}}{\left|\left| W^{(k+1)} \right|\right|} \tag{5}$$

where $W^{(k)}$ indicates the result of kth step of W, and G denotes non-quadratic function. $W^{(k+1)}$ was obtained until $<W^{(k+1)}, W^{(k)}>= 1$. The details for obtaining weight matrix for $\hat{\mathbf{s}}_i$ are elucidated in Appendix A.3. Decorrelation operation was conducted on real source reflectance through principal component analysis (PCA) [24].

2.4. Model Framework

The proposed model mainly consisted of a model input construction part, predictive models for WQPs from 0.5 m below the water surface (FP-xLSTM$_s$), and predictive models for WQPs from more than 0.5 m below the water surface (FP-xLSTM$_d$) as shown in Figure 2. The input of FP-xLSTM$_s$ embraced hyperspectral reflectance, feature interaction, and Eigenvalue matrices. The feature interaction matrix was obtained from the upper-right model input construction workflow, and a small region scanner was employed to extract eigenvalues of each region in the reflectance matrix. The input to FP-xLSTM$_d$ was intermediate variables from the interaction and shift layer correlating the information from FP-xLSTM$_s$. The output was the predicted WQPs concentration. Eigenvalue, reflectance, and land cover embedding matrices worked as input to predict OAC WQPs; Eigenvalue, interaction, and land cover embedding matrices worked as input to predict non-OAC WQPs due to optical non-sensitivity. Layers consisted of an interaction layer, shift layer, and fully connected layer. Each time period went through mLSTM and sLSTM with hidden dimensions of 128, where the sequence length was four.

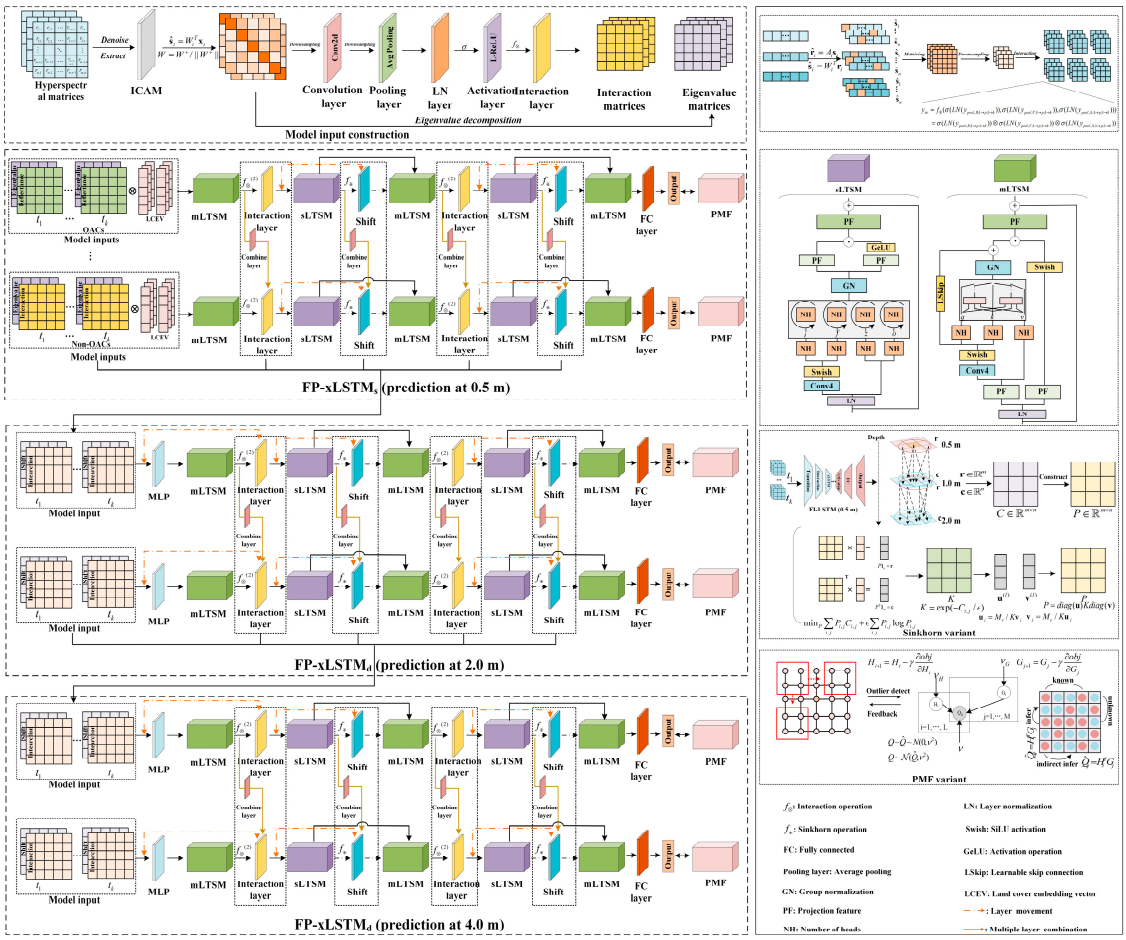


Figure 2. Model architecture. The left part is the workflow for predicting WQP concentrations, and the upper left is model input construction. The right part contains detailed structure of some modules including sLSTM, mLSTM, shift layer, and PMF. The upper right is a detailed demonstration of model input construction.

2.5. Model Formulation

2.5.1. Model Development of FP-xLSTM$_s$

In the model input construction part, raw reflectance of each pixel point of the UAV hyperspectral image was denoised through an ICA-variant module to obtain real source

reflectance. 2-dimensional convolution layer and an average pooling layer were employed to extract hyperspectral features from the real source reflectance.

$$y_{conv} = \text{Conv2d}\left(W_{conv}^{(1)} X + b_{conv}^{(1)}\right) \tag{6}$$

$$y_{pool} = Avgpooling\left(W_{pool}^{(1)} y_{conv} + b_{conv}^{(1)}\right) \tag{7}$$

where y_{conv} and y_{pool} indicate the results of the 2-dimensional convolution layer and average pooling layer, respectively; W indicates the weight matrix; and b indicates bias. Spectral information was correlated in inner space, where each entry was surrounded by its adjacent entries to form a region on the reflectance matrix. In each region, a scanner was utilized to extract eigenvalues to obtain the Eigenvalue matrix $\Lambda = \begin{bmatrix} \lambda_{1,1} & \cdots & \lambda_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda_{n,1} & \cdots & \lambda_{n,n} \end{bmatrix}$ from real source reflectance matrix, where each entry represents the largest Eigenvalue for each small area, with each entry value of the real source reflectance matrix as the center surrounded by adjacent entries in the scanned region (e.g., a scanned region of 3×3 where the center entry was surrounded by its adjacent 8 entries). The Eigenvalue matrices showed correlation of neighboring points within a certain area and differences among scanned regions. Interaction between the center point and its adjacent points elucidated their feature correlation and mutual effect. Real source reflectance matrices of spatially adjacent pixel points in a spatial neighborhood were correlated to one another through feature interaction. Furthermore, the feature interaction matrix was obtained through a model input construction workflow with a multi-dimensional cross product.

$$y_{it} = f_{\otimes}\left(y_{c,r},\, y_{a_1,r},\, \cdots,\, y_{a_j,r}\right) = y_{c,r} \otimes y_{a_1,r} \otimes, \cdots, \otimes y_{a_j,r} \tag{8}$$

where y_{it} denotes the results of feature interaction among the real source reflectance matrix, eigenvector matrix, and Eigenvalue matrix; $f_{\otimes}$ denotes the multi-dimensional cross product operation; and $y_{c,r}$ and $y_{a_j,r}$ denote the results of the center point and its adjacent points after the activation layer. The details to obtain feature interaction results can be found in Appendix A.4. Non-OACs, including TP, TN, COD, and BOD at 0.5 m below the water surface, were predicted by FP-xLSTM$_s$ using interaction and Eigenvalue matrices as model inputs since their variation did not affect spectral reflectance directly and OACs' variation indirectly affected non-OACs. OACs, including Chl-a and TSS, were predicted by FP-xLSTM$_s$ using reflectance and Eigenvalue matrices as model inputs. Thus, FP-xLSTM$_s$ were trained with different input variables for OACs and non-OACs, respectively. The Liang–Kleeman information flow framework quantified directional causal interactions between riparian land features and WQPs, revealing that anthropogenic land-cover types exerted stronger influences on WQP variability than natural vegetation because of their greater hydrological connectivity, nutrient export, and pollutant transport potential [1]. The derived information flow rates were subsequently integrated with spectral reflectance, Eigenvalue features, and interaction matrices to construct multi-source input representations, thereby enhancing the interpretability, robustness, and cross-regional generalization capability of the proposed model.

In the up-projection mLSTM module wrapped inside two multilayer perceptrons (MLPs) via a convolution and a learnable skip connection, the predicted results were obtained from the output gate at each time step. mLSTM employed matrix memory, allowing for parallel processing and improved storage capacity. mLSTM utilized matrices to store and retrieve information, which enhanced storage capacity and resolved the limitations of LSTM in handling spectral features that are hard to capture in vector form. Its

covariance update rules are suitable for handling large-scale parallel data, which analyzes subtle short-term variation.

$$C_t = \mathrm{f_t} C_{t-1} + \mathrm{i_t} \boldsymbol{v}_t \boldsymbol{k}_t^T \tag{9}$$

$$\boldsymbol{n}_t = \mathrm{f_t} \boldsymbol{n}_{t-1} + \mathrm{i_t} \boldsymbol{k}_t \tag{10}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \left(\mathrm{C}_t \boldsymbol{q}_t / \max\left(\left| \boldsymbol{n}_t^T \boldsymbol{q}_t \right|, 1 \right) \right) \tag{11}$$

$$\boldsymbol{x}_t = \mathrm{SiLU}\left(\mathrm{Conv}\left(\boldsymbol{w}_x^T f_{LN}(X_t) + \boldsymbol{b}_x \right) \right) \tag{12}$$

$$\boldsymbol{x'}_t = W_s \boldsymbol{x}_t + \boldsymbol{b}_s \tag{13}$$

$$\boldsymbol{y}_t = \left(\mathrm{GN}(\boldsymbol{h}_t) + \boldsymbol{x}'_t \right) \odot \mathrm{SiLU}\left(\boldsymbol{w}_r^T f_{LN}(X_t) + \boldsymbol{b}_r \right) \tag{14}$$

$$\boldsymbol{q}_t = \left(\boldsymbol{W'}_q \boldsymbol{x}_t + \boldsymbol{b}_q \right) \odot \left(\boldsymbol{w}_q^T E_t + \boldsymbol{b}_q \right) \tag{15}$$

$$\boldsymbol{k}_t = \frac{1}{\sqrt{d}} \left(\boldsymbol{W'}_k \boldsymbol{x}_t + \boldsymbol{b}_k \right) \odot \left(\boldsymbol{w}_k^T E_t + \boldsymbol{b}_k \right) \tag{16}$$

$$\boldsymbol{v}_t = \left(\boldsymbol{W'}_v \boldsymbol{x}_t + \boldsymbol{b}_v \right) \odot \left(\boldsymbol{w}_v^T E_t + \boldsymbol{b}_v \right) \tag{17}$$

$$\mathrm{i}_t = \exp\left(\boldsymbol{w}_i^T \boldsymbol{x}_t + b_i \right) \tag{18}$$

$$\mathrm{f}_t = \exp\left(\boldsymbol{w}_f^T \boldsymbol{x}_t + b_f \right) \tag{19}$$

$$\boldsymbol{o}_t = \mathrm{sigmoid}(W_o \boldsymbol{x}_t + \boldsymbol{b}_o) \tag{20}$$

where f, i, $\boldsymbol{o}$, C, $\boldsymbol{n}$, $\boldsymbol{h}$, $\boldsymbol{q}$, $\boldsymbol{k}$, v, X, E, W, $\boldsymbol{W'}$, and b denote forget gate, input gate, output gate, cell state, normalizer state, hidden state, query, key, value, interaction results of model inputs, Eigenvalue, weight, block diagonal linear projection, and bias, respectively. LN and GN indicate layer normalization and group normalization. $(\boldsymbol{W}X_t + \boldsymbol{b}) \odot (\boldsymbol{w}E^T + \boldsymbol{b})$ enhance relatedness between reflectance or interaction matrix and Eigenvalue matrix. The forget gate worked in consistence with decay rate; the input gate worked in consistency with the learning rate; the output gate scaled the predicted results. The normalizer state was the weighted sum of key vectors, with each key vector weighted by the input gate and all future forget gates, and it captured gate strength. The exponential activation function might obtain a larger value, thereby causing an overflow. Therefore, an additional state $\mathrm{m_t}$ was used to stabilize the gating signal.

$$m_t = \max(\log(\mathrm{f}_t) + m_{t-1}, \log(\mathrm{i}_t)) \tag{21}$$

$$\mathrm{i'}_t = \exp(\log(\mathrm{i}_t) - m_t) \tag{22}$$

$$\mathrm{f'}_t = \exp(\log(\mathrm{f}_t) - m_{t-1} - m_t) \tag{23}$$

where m_t is the stabilizer state; $\mathrm{i'}_t$ is the stabilized input gate; and $\mathrm{f'}_t$ is the stabilized forget gate. Thus, i_t and f_t could be superseded by $\mathrm{i'}_t$ and $\mathrm{f'}_t$ in the forward pass which changed neither the output nor the derivatives of the loss with respect to the parameters.

By extending the vector operations in the traditional LSTM to matrix operations, the memory ability of the model has been greatly enhanced. Thus, mLSTM improved prediction efficiency and the ability to capture tiny variations between features by incorporating information from two consecutive time steps. The calculation details were demonstrated in the mLSTM block diagram of Figure 2. Then, an interaction layer was employed to extract

the mutual effect of interaction and Eigenvalue matrices of non-OACs or reflectance and Eigenvalue matrices of OACs:

$$\boldsymbol{y}_{it}^{(2)} = f_{\otimes}^{(2)}\left(\boldsymbol{w}_{it}\boldsymbol{y}_{t}^{T},\ W_e E\right) \tag{24}$$

$$y_{out} = \sigma\left(\boldsymbol{w}^{T} y_{it}^{(2)} + b_{out}\right) \tag{25}$$

where $\boldsymbol{y}_{it}^{(2)}$ indicates interaction results and y_{out} indicates the predicted WQP concentration at the current time. $f_{\otimes}^{(2)}$ denotes row-wise cross product on multi-dimensional variables. Row-wise cross product enhanced the capability of capturing significant features and relatedness between previous and current time. The calculation process details were elaborated in Appendix A.5. With increasing depth, spectral reflectance scarcely enunciated the variation of OACs. A spatial neighborhood was determined to form a region (e.g., 3×3) that affected water below (e.g., 2.0 m and 4.0 m below the water surface), as shown in the Sinkhorn variant part of Figure 2. WQPs at 0.5 m below the water surface ($\text{WQPs}_{0.5}$) affected WQPs at 2.0 m below the water surface ($\text{WQPs}_{2.0}$), which affected WQPs at 4.0 m below the water surface ($\text{WQPs}_{4.0}$) due to water flowability and underwater substance exchange. Thereby, the spatial distribution of $\text{WQPs}_{0.5}$ and that of $\text{WQPs}_{2.0}$ were correlated. The effect imposed by $\text{WQPs}_{0.5}$ upon $\text{WQPs}_{2.0}$ would follow the easiest way to be achieved at the lowest cost. The results from the interaction layer were employed to establish a Sinkhorn network to transport the effect from $\text{WQPs}_{0.5}$ to $\text{WQPs}_{2.0}$ with a cost matrix that was formulated with respect to spatial distance among pixel points of shallow and deep levels in the shift layer. Thus, the objective of minimizing the total cost of transporting impact from $\text{WQPs}_{0.5}$ to $\text{WQPs}_{2.0}$ was formulated as:

$$obj = \min_{P} \sum P_{i,j} C_{i,j} + \epsilon \sum P_{i,j}\left(\log P_{i,j} - 1\right) \tag{26}$$

where $\boldsymbol{r} \in \mathbb{R}^{K}$ denotes $\text{WQPs}_{0.5}$ concentrations distribution; P denotes impact matrix; C denotes cost matrix; $\boldsymbol{c} \in \mathbb{R}^{N}$ denotes $\text{WQPs}_{2.0}$ concentrations distribution; constraint condition of $P\mathbf{1}_N = \boldsymbol{r}$ and $P\mathbf{1}_k = \boldsymbol{c}$ hold when effect by $\text{WQPs}_{0.5}$ us imposed upon $\text{WQPs}_{2.0}$. P and interaction matrices are retained at each time step and used as the input to FP-xLSTM_{d}. The calculation details of the impact matrix were demonstrated in Appendix A.6 Algorithm A1. Interaction matrices and the impact matrices were combined in the combine layer and retained to work as the input to FP-xLSTM_{d}. The sLSTM module added a scalar update mechanism based on the traditional LSTM. Fine-grained control of the internal memory unit was carried out, optimizing the gating mechanism to make it more suitable for processing sequence data with subtle time variation. Additionally, interaction matrices and the impact matrices worked as the input to sLSTM.

$$\boldsymbol{c}_t = \mathbf{f}_t \odot \boldsymbol{c}_{t-1} + \mathbf{i}_t \odot \boldsymbol{z}_t \tag{27}$$

$$\boldsymbol{n}_t = \mathbf{f}_t \odot \boldsymbol{n}_{t-1} + \mathbf{i}_t \tag{28}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \left(\boldsymbol{c}_t \odot \boldsymbol{n}_{t-1}^{-1}\right) \tag{29}$$

$$\boldsymbol{x}_t = \text{SiLU}\left(\text{Conv}\left(W_x f_{LN}\left(\boldsymbol{y}_{it}^{(2)}\right) + \boldsymbol{b}_x\right)\right) \tag{30}$$

$$\boldsymbol{z}_t = \tanh\left(\boldsymbol{W'}_z \boldsymbol{x}_t + \boldsymbol{R}_z \boldsymbol{h}_{t-1} + \boldsymbol{b}_z\right) \tag{31}$$

$$\mathbf{i}_t = \exp\left(\boldsymbol{W'}_i \boldsymbol{x}_t + \boldsymbol{R}_i \boldsymbol{h}_{t-1} + \boldsymbol{b}_i\right) \tag{32}$$

$$\mathbf{f}_t = \exp\left(\boldsymbol{W'}_f \boldsymbol{x}_t + \boldsymbol{R}_f \boldsymbol{h}_{t-1} + \boldsymbol{b}_f\right) \tag{33}$$

$$\boldsymbol{o}_t = \text{sigmoid}\left(\boldsymbol{W'}_o \boldsymbol{x}_t + \boldsymbol{R}_o \boldsymbol{h}_{t-1} + \boldsymbol{b}_o\right) \tag{34}$$

$$\boldsymbol{y}_t = W_y(W_l \text{GN}(\boldsymbol{h}_t) + \boldsymbol{b}_l) \odot GELU(W_r \text{GN}(\boldsymbol{h}_t) + \boldsymbol{b}_r) + X_t \tag{35}$$

The results kept going through the interaction layer and shift layer, then the combine layer, which were retained as input to FP-xLSTM$_d$. Finally, the fully connected (FC) layer took the results from the last mLSTM to calculate the WQPs concentrations. To improve calculation accuracy, the PMF variant feedback module was engaged to determine resulting WQPs concentration outliers and recalculate detected outliers, improving explainability and continuity of spatial distributions of WQPs concentrations. A scanning operation was implemented on a certain region covering spatially adjacent pixels with the obtained WQPs concentrations. PMF variant feedback module considered spatially continuous variation of WQPs concentrations in each scanned region, eliminating sudden changes in WQPs concentrations between adjacent points. The prediction accuracy was enhanced with detected outliers being recalculated by the PMF module. The detailed recalculation of WQPs concentration outliers detected by the PMF module was enunciated in Appendix A.7.

2.5.2. Model Development of FP-xLSTM$_d$

WQPs at more than 0.5 m below the water surface could not be retrieved from spectral reflectance, and thus, the retained interaction and shift information worked as input to FP-xLSTM$_d$ to predict variation of WQPs$_{2.0}$. Similarly, retained interaction and shift information from the prediction of WQPs$_{2.0}$ variation worked as input to WQPs$_{4.0}$. A multi-layer perceptron (MLP) was employed to extract features of interaction and shift information. Then, mLSTM, sLSTM, interaction layer, and shift layer were utilized to predict WQPs variation at different time stamps. Features from the interaction layer and shift layer were retained in the prediction of WQPs$_{2.0}$ and worked as input to the prediction of WQPs$_{4.0}$, the process of which was similar to that of WQPs$_{0.5}$ as shown in the left part of Figure 2.

The workflow of prediction of multi-depth spatio-temporal distribution of WQPs was summarized: (1) extraction Eigenvalue and interaction matrices from hyperspectral data; (2) utilization of reflectance and Eigenvalue matrices as input to prediction of OAC WQPs and interaction and Eigenvalue matrices as input to prediction of non-OAC WQPs; (3) retainment of intermediate information including featuring information from interaction and shift layer to work as input to prediction of WQPs at more than 0.5 m below water surface; (4) recalculation of detected WQPs concentration outliers through PMF module in each scanned region. The detailed algorithm for predicting multi-depth variation of WQPs over time was elucidated in Appendix A.8 Algorithm A2. FP-xLSTM employed different strategies to retrieve OAC and non-OAC WQPs and averted calculation instability due to the insensitivity of non-OAC WQPs to spectral reflection. Moreover, FP-xLSTM established feature sharing and delivery among WQPs at various depths, efficiently mitigating differences in calculation accuracy of WQPs at various depths due to optical signal attenuation with depth augmentation.

3. Results

3.1. Dataset Description

This study conducted multi-depth in-situ water sampling at depths of 0.5 m, 2.0 m, and 4.0 m, together with UAV image acquisition in March, May, July, September, and November 2025. Statistical distributions of water quality parameters (WQPs) based on water quality classification standards are shown in Figure 3.

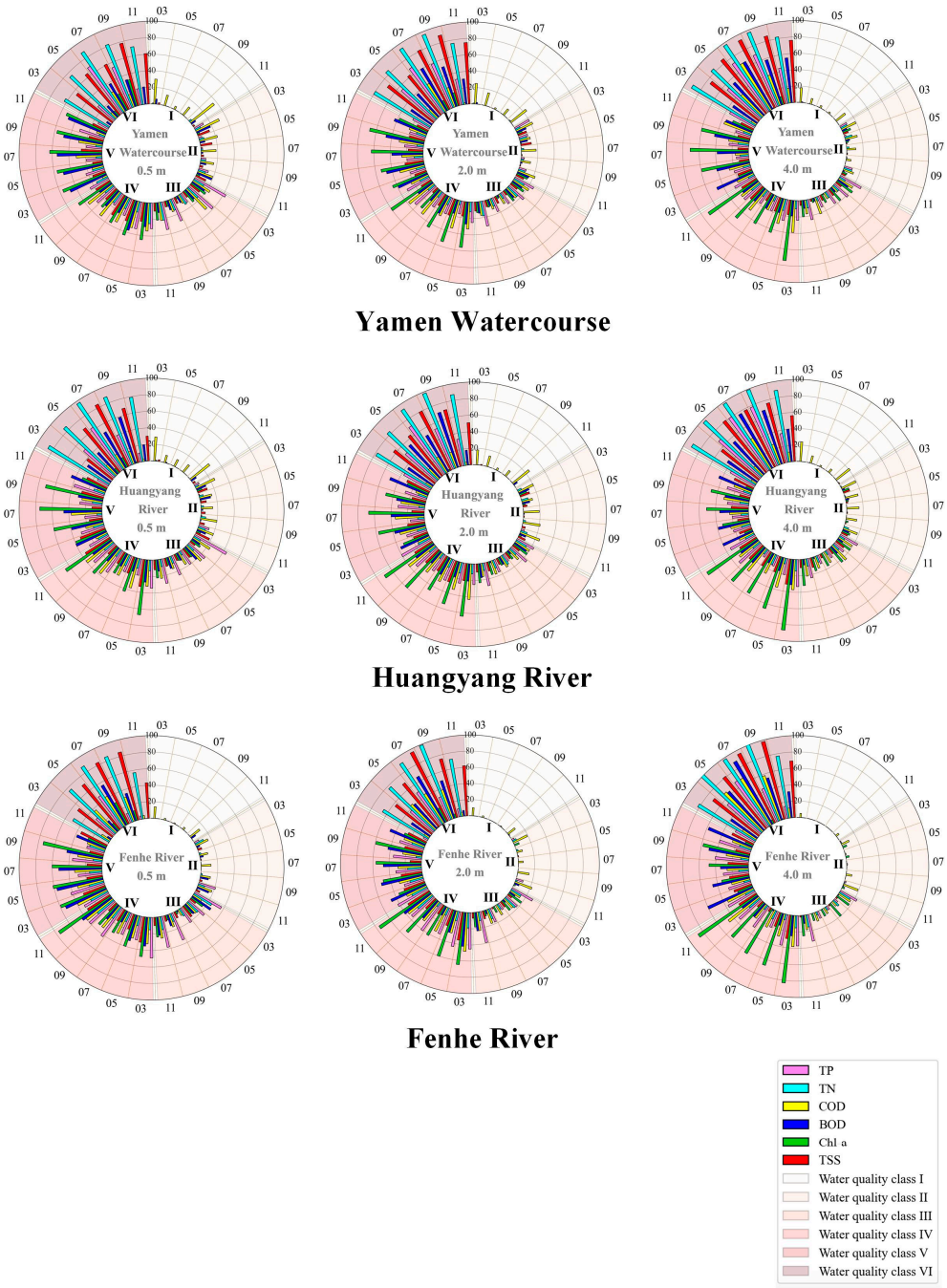


Figure 3. Statistical results of classification of WQPs concentrations in different regions from 03/2025 to 11/2025. Different sector colors indicate water quality classes with the darkest color indicative of the highest class. The darker the color turns, the worse the water quality is.

Temporal variability and consistency of WQPs were further evaluated using coefficients of variation (CVs) and Pearson correlation coefficients between consecutive monitoring periods, as shown in Figure 4. In Yamen Watercourse, COD and BOD exhibited relatively strong temporal correlations at all depths, suggesting stable pollution sources such as domestic sewage and industrial discharge. In contrast, TP, TN, Chl-a, and TSS showed weaker temporal correlations, indicating stronger seasonal variability. Similar patterns were observed in Huangyang River, where COD and BOD maintained relatively high temporal consistency, whereas TP and Chl-a displayed stronger fluctuations, particularly in surface layers. In Fenhe River, TP, TN, and COD demonstrated relatively strong correlations between May and September, while TP, TN, COD, and BOD at 4.0 m depth

showed greater temporal consistency than those in shallower layers, likely due to water stratification and sediment release processes.

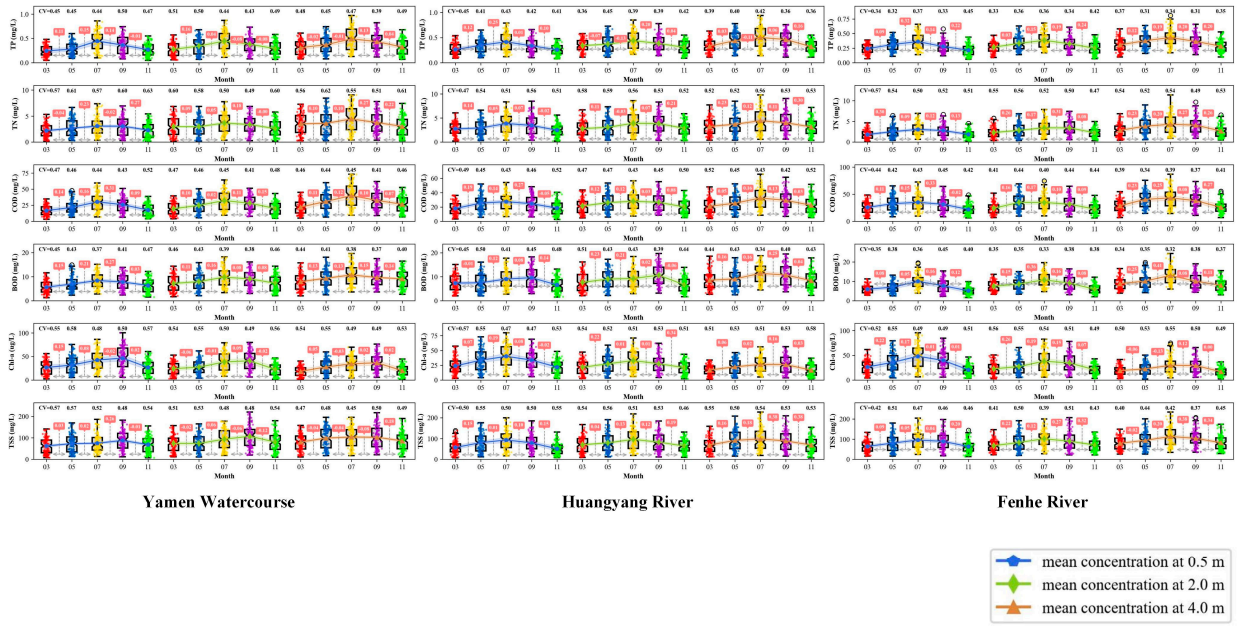


Figure 4. WQPs distribution from 03/2025 to 11/2025 in three different regions. The number in the red rectangle indicates the Pearson correlation coefficient value, where the larger the number, the higher the rectangle position. CV indicates the coefficient of variation, which is located over each bar. Every five months from left to right on each plot represents one depth; in total, three depths, including 0.5 m, 2.0 m, and 4.0 m below the water surface on each plot. All digits were rounded to two decimal places, with the original positive or negative sign kept.

CV analysis revealed significant spatial heterogeneity among WQPs. TN consistently exhibited the highest CVs across all regions and depths, indicating uneven nitrogen inputs and heterogeneous mixing conditions. In addition, concentrations of TP, TN, COD, BOD, and TSS generally increased with depth, whereas Chl-a concentrations decreased because of reduced light penetration and weaker phytoplankton activity in deeper layers.

Compared with Fenhe River, Yamen Watercourse and Huangyang River exhibited more uneven point-source pollution distributions according to their CV patterns. Furthermore, deeper water layers appeared to be more strongly influenced by internal processes, such as sediment release and redox reactions, than by seasonal external inputs. Consequently, WQPs below 0.5 m depth generally exhibited stronger temporal continuity between consecutive monitoring periods, reflecting the relatively slow response of deep-water environments to seasonal surface disturbances, including upstream pollutant transport and rainfall runoff.

Figure 5 showed the association of each WQP with spectral bands, where the Mann–Kendall (MK) Z-score was employed to test the changing trend of spectral difference of maximum and minimum concentration of each WQP from 401 nm to 998 nm. The spectral difference of TP scarcely showed resemblance across different regions, and large differences among MK Z scores of various months existed, indicating an inconspicuous changing trend [25]. There were spectral difference peaks of TP near 750 nm, 800 nm, and 950 nm, which might respond to variation in TP concentration. MK Z scores of TN differed significantly over different months and failed to demonstrate an upward or downward changing trend. Spectral difference peaks of TP near 600 nm, 900 nm, and 990 nm might correspond to TN variation to some extent, which might explain that TN varied by month. MK Z scores of COD were negative, indicating that a downward trend existed. Spectral

difference peaks of COD occurred near 580 nm and 630 nm, followed by a spectral difference drop near 750 nm, consistent with COD feature bands. Furthermore, the spectral difference of BOD near 950 nm possibly responded to the variation of BOD. Large differences in MK Z scores of BOD over various months showed that spectral reflectance differed from one another. Tiny differences in MK Z scores of Chl-a in three regions illustrated the consistency of spectral differences over various months, which proved stronger optical sensitivity of Chl-a compared to other WQPs. Spectral differences attained some peaks near 400 nm, 700 nm, 750 nm, 820 nm, and 950 nm, which might respond to Chl-a variation. MK Z scores of TSS differed little from one another, signifying a stable changing trend. Spectral difference peaks of TSS were presented near 750 nm, 830 nm, and 900 nm, which might correspond to TSS variation.

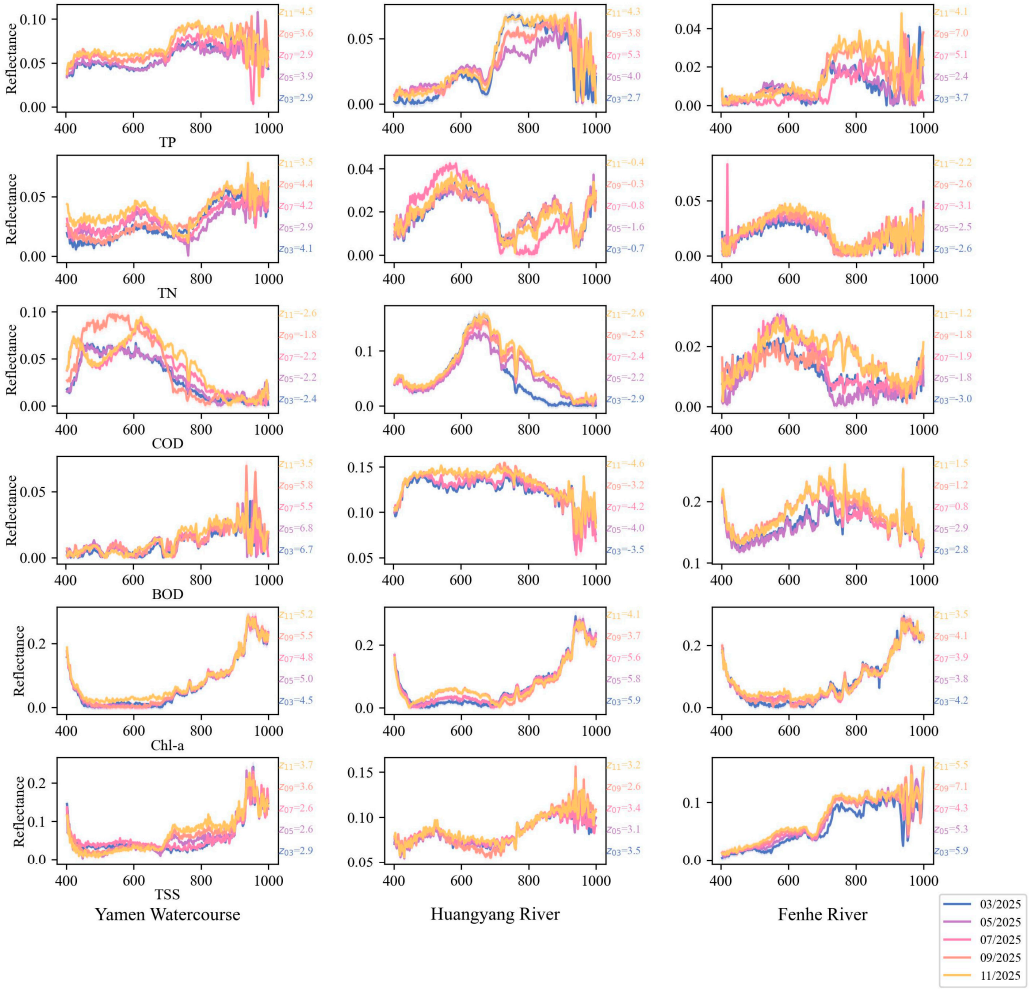


Figure 5. Spectral difference of maximum and minimum concentration of each WQP over wavelengths from 03/2025 to 11/2025. Z of different colors represents the Mann–Kendall Z-score with respect to the spectral reflectance difference of maximum and minimum concentration of each WQP at the same month.

3.2. Evaluation Metrics

In this study, we conducted performance experiments to test FP-xLSTM performance with respect to root mean squared error (RMSE), coefficient of determination (R^2), and mean absolute percentage error (MAPE). Welch's *t*-test, MK, and χ^2 test were employed to further verify the performance of FP-xLSTM by comparing the spatio-temporal distribution of obtained results by FP-xLSTM with that of ground truth. The null hypothesis of Welch's *t*-test was that the center of predicted values was equal to that of measured values, given

that measured and calculated values were independent with unequal variances [26]. Water samples were collected at different sampling sites and were collected in a certain order such that the MK test was applied to test if the dataset of each WQP in each period followed a certain changing trend [27]. χ^2 was utilized to compare the classification result distributions of measured and predicted values, which provided the overall consistency between measured and predicted WQPs distributions and discovered problems with the model in terms of category bias or imbalance. The null hypothesis assumed that the distribution of predicted classification results did not significantly deviate from that of the measured values [28]. The formulation of the aforementioned statistical tests was demonstrated in Appendix A.10.

The significance level of Welch's *t*-test, MK, and χ^2 was set to 0.05. FP-xLSTM was trained using the AdamW optimizer with an initial learning rate of 0.001 and a weight decay of 1×10^{-4}. The batch size was set to 16. The random seed was 100. The number of training epochs was 100, and early stopping with a patience of 15 epochs was employed based on validation loss. Activation functions included SiLU, GeLU, tanh, and Leaky-ReLU, as shown in Figure 2. Training, validation, and testing data accounted for 60%, 20%, and 20% of the total dataset. A total of 125 fixed sampling locations were monitored repeatedly in 03/2025, 05/2025, 07/2025, 09/2025, and 11/2025. These locations were divided into training, validation, and testing sets at 75, 25, and 25 unique sampling locations through a forward-validation approach, respectively, covering five time periods. All observations collected from the same sampling location across the five sampling campaigns were assigned exclusively to the same subset. Thus, no sampling location was shared among the training, validation, and testing sets, preventing spatial information leakage while preserving the temporal observations within each location.

3.3. Baseline Models

The state-of-the-art baseline models of FP-xLSTM employed information from previous time periods in time series, including spectral features and spatio-temporal features, as model inputs to predict WQPs concentration distributions of the next time period, and their descriptions are shown below.

- GCN [29]: A graph convolutional network was proposed for dynamic prediction of WQPs that leveraged spatial relatedness of each joint.
- ST-GCN [30]: A spatio-temporal graph convolutional network model was proposed to predict WQPs concentration distributions based on the spatio-temporal correlation between river stations.
- SpecFormer [31]: Spectral graph neural networks learned graph representations via spectral domain graph convolutions, which were utilized to couple WQPs and spectral features.
- CNN-Transformer [32]: The synergistic integration of an efficient compression CNN and transformer with a pooling strategy efficiently extracting spectral features was leveraged to predict WQPs concentrations.
- MSM-Transformer [33]: A multi-scale modeling transformer architecture with a decomposition attention mechanism in the time series integrated information from different scales to capture distribution patterns and dynamic changes of WQPs.
- LSTM [34]: LSTM extracted WQPs based on their respective pollution share rates and employed the Spearman method to identify the influential factors impacting each key parameter.
- Bi-LSTM [35]: A bidirectional LSTM method combined with a multi-layer perceptron was proposed to predict spatial-temporal variation of water quality.

- ConvLSTM [36]: Convolutional LSTM network extended traditional LSTM architecture by replacing fully connected layers with convolutional operations during input-to-state and state-to-state transitions to retrieve WQPs.
- CNN-LSTM-Attention [37]: The CNN-LSTM-Attention method with an attention mechanism extracted short-term abstract features of water quality data by CNN and selected an LSTM model to capture long-term dependencies between variables to detect water quality.
- xLSTM [38]: A method with xLSTM architecture integrated the scaled dot-product attention (SDPA) mechanism to improve the effectiveness of retrieving WQPs.

3.4. Model Evaluation

Traditional methods might be inaccurate in prediction due to insufficient samples in the target area. Transfer learning alleviated the problem of insufficient data in the target area by leveraging the feature representation of a large amount of data in other areas. Water bodies in different regions were affected differently by pollution sources, climate, rainfall, and river basin characteristics. Transfer learning enabled the model to quickly adapt to the environmental characteristics of the new region through the fine-tuning method. Model Training directly in a single area might overfit the environmental features of that area (e.g., water body types and seasonal variations). Transfer learning enabled the model to learn more general water quality patterns through cross-regional knowledge sharing.

A model evaluation experiment of FP-xLSTM was conducted to verify cross-region transfer learning capability through leave-one-basin-out cross-basin validation, which tested FP-xLSTM performance on one region dataset by training the model on two datasets from other two regions (e.g., the model was tested on the Yamen Watercourse dataset by being trained on other two datasets from Huangyang and Fenhe River). The model predicted WQPs concentration distributions in 11/2025 through the datasets of the first four periods in the time series. The randomly selected sampling points of each watershed, with sampling over five periods and three depths at each point, worked as the training dataset, and the rest of the sampling points of each watershed worked as the testing dataset due to the spatio-temporal variation continuity of water quality in the time series in each watershed. Prediction of WQPs concentration distributions for the next period in the time series was conducted based on information sharing and delivery of past periods. The proposed method design considered external and internal environmental variations in different regions, which enhanced generalization ability. MK Z-score of predicted values was approximated to that of measured values, indicating prediction effectiveness through analogous changing trends of predicted and measured values. The results of comparison between predicted values and measured values are shown in Figure 6. The prediction of Chl-a and TSS by FP-xLTSM outperformed that of other WQPs due to the direct and efficient response of OACs to spectral variation involved in FP-xLTSM at 0.5 m depth. The prediction of TP, TN, COD, and BOD was based on the interaction and Eigenvalue matrices of the spatial neighborhood. The average R^2 of OAC WQPs increased by 6.95% compared to that of non-OAC WQPs at 0.5 m depth in three regions, and the difference between OACs and non-OAC WQPs was insignificant to some extent, which proved the calculation stability and efficiency of FP-xLSTM.

Optical penetration decayed with increasing depth, which caused FP-xLSTM$_s$ to outperform FP-xLSTM$_d$ in predicting OAC WQPs. Moreover, the fitting degree of FP-xLSTM to measured values gradually decreased with increasing depth for non-OAC WQPs with respect to R^2 s, which illustrated that spatial relatedness and feature representation accuracy attenuated during information sharing and delivery. As the number of network layers increased, the input features went through multiple linear and nonlinear transformations,

and the useful original feature information might be diluted, which lowers prediction accuracy for WQPs. The average R^2 of OACs and non-OAC WQPs decreased by 8.72% and 17.10%, respectively, from 0.5 m to 4.0 m depth, indicating the difference in prediction accuracy decline between OACs and non-OAC WQPs, which further illustrated that spectral penetration decayed more slowly than the dilution of interaction information delivery and sharing. The signals of non-OACs were inherently weak, and attenuation during the gradient propagation process would make it more difficult for the previous layer to learn these weakly correlated features. The changing trends of predicted values for OAC WQPs were more closely aligned with the measured values than those of non-OAC WQPs, as indicated by smaller differences in MK Z scores between predicted and measured values [39].

Furthermore, the MK Z-score difference between predicted values and measured values increased with depth, increasing more for OAC WQPs than non-OAC WQPs. This was likely because OCAs were biologically and optically active, subject to more complex, nonlinear fluctuations influenced by light availability and biotic interactions. The confidence ellipse showed the population mean concentration of each WQP, and correlation existed among points near the long axis of the confidence ellipse owing to their similar spatial and spectral features. Predicted WQPs mainly aggregated inside the confidence ellipse of 1 standard deviation at the depth of 0.5 m and aggregated inside confidence ellipses of 2 and 3 standard deviations at depths greater than 0.5 m due to decreased R^2 and increased MAPE. Almost all points stayed inside these three circles, which further proved that the population mean of predicted WQPs approximated that of measured WQPs well, and prediction of each WQP was efficient at different confidence levels.

The predicted values of WQPs were classified into six classes according to the water quality classification standards shown in Table A1. Predicted WQPs concentrations were compared with measured values at various depths with respect to classification results as shown in Figure 7. The overall prediction performance of OAC WQPs was superior to that of non-OAC WQPs. The performance of FP-xLSTM$_s$ on OAC WQPs differed insignificantly from that of non-OAC WQPs in terms of classification results at 0.5 m depth. Additionally, the classification results of predicted concentrations exhibited closer correspondence to those of measured concentrations for TP and TN than for other WQPs due to a narrow range of predicted and measured values. The R^2 s of COD and BOD were less than those of other WQPs, signifying relatively low fitting degree to measured values and significant fluctuation around each measured value at a depth greater than 0.5 m. Moreover, the classification interval spans of TSS were wider than others, which gave rise to small classification differences between its measured and predicted values, improving classification accuracy and stability of predicted results to some extent. Marked differences in classification interval spans from class I to class VI led to relatively large discrepancies between measured and predicted Chl-a concentrations despite the resulting R^2 being greater than other WQPs. The differences between measured and predicted COD, BOD, and Chl-a in higher classes became more pronounced with depth increasing owing to the decreasing prediction accuracy and stability.

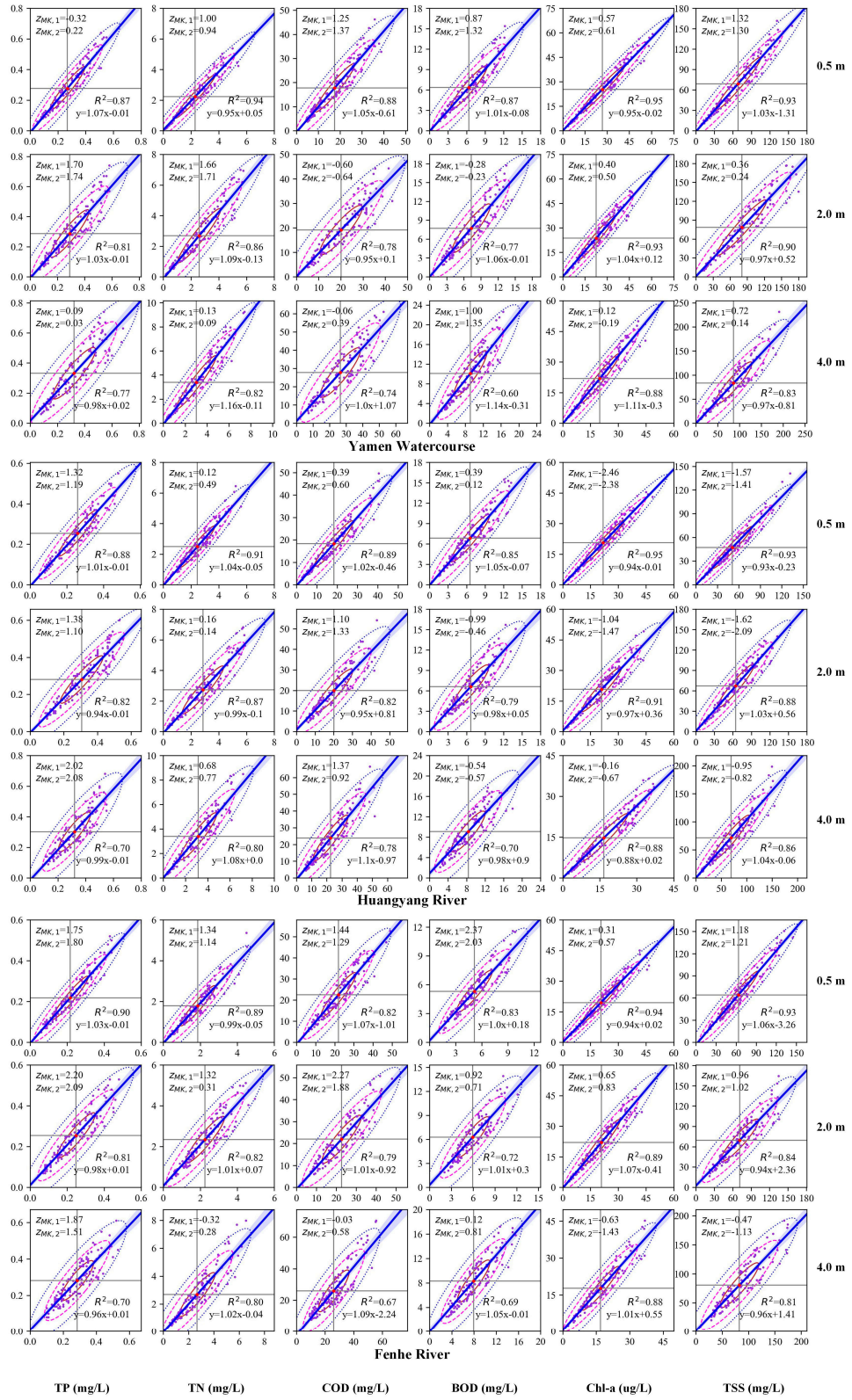


Figure 6. Model performance evaluation by comparison between predicted values and measured values at various depths in three regions. The blue line denotes the regression line. $Z_{MK,1}$ denotes the MK Z-score of the measured value. $Z_{MK,2}$ denotes the MK Z-score of the predicted value. Red, purple dashed, and blue dashed circles denote confidence ellipses of three types of standard deviation σ, 1 σ for 68.27% confidence in red dashed line, 2 σ for 95.45% confidence in purple dashed line, and 3 σ for 99.73% confidence in blue dashed line. The red dot denotes the center of the measured and predicted WQPs.

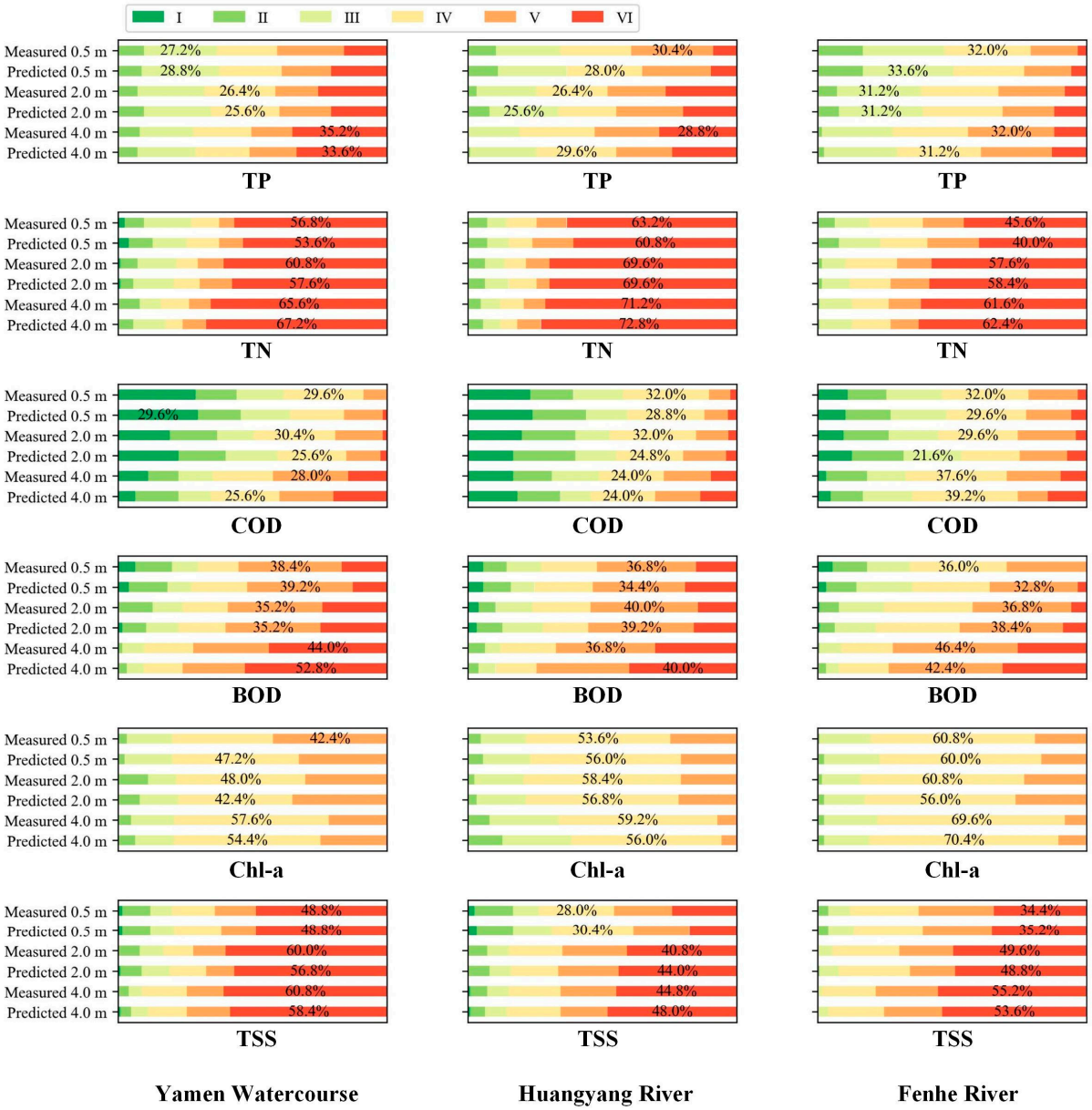


Figure 7. Comparison between predicted and measured WQP concentrations at various depths through water quality classification. The class accounting for the largest portion is digitally annotated.

4. Discussion

4.1. Model Comparison

Taylor diagram was employed to investigate the model performance difference among FP-xLSTM and its baseline models with respect to standard deviation, R^2, and centered root mean squared error (CRMSE), as shown in Figure 8. CRMSE quantified the deviation of predicted values from observed values and was related to the model's capability to capture patterns or variability of measured data. FP-xLSTM outperformed its baseline models with respect to the aforementioned evaluation metrics at various depths of the three regions. xLSTM showed the second-best performance due to smaller CRMSE and higher R^2 compared to other baseline models, signifying that it captured the variability of the measured values. Additionally, the ConvLSTM and CNN-LSTM-Attention resembled xLSTM in correlation but underperformed xLSTM in CRMSE, where they showed discrepancies in the variation trend of predicted values. ConvLSTM, CNN-LSTM-Attention, and xLSTM had similar design architecture but different embedded modules and connection modes,

such that they resembled each other in correlation with measured values. ConvLSTM and CNN-LSTM-Attention involved additional spatial convolutions and attention mechanisms, assuming data to have pronounced correlation among spatial features, whereas WQPs lacked direct spatial correlation and feature interaction. Thus, their convolutional layers might attenuate important signals or introduce additional noise. Under conditions of high data noise, short sequence lengths, or simple temporal patterns, ConvLSTM, CNN-LSTM-Attention, and xLSTM might fail to extract relevant features, potentially resulting in bias. FP-xLSTM was established based on the basic architecture of xLSTM and involved spatial correlation, feature interaction, substance transport mechanism, and feedback mechanism, making it more suitable for capturing dynamic variation of multi-depth WQPs in short time scales. The predicted values by ConvLSTM, CNN-LSTM-Attention, and xLSTM were evenly distributed, resembling those of FP-xLSTM with respect to standard deviation at 0.5 m depth, but their changing patterns were dissimilar to FP-xLSTM due to relatively high CRMSE. Bi-LSTM outperformed LSTM in terms of R^2 and CRMSE due to its bidirectional feature delivery, which enhanced calculation accuracy and stability. LSTM and Bi-LSTM underperformed ConvLSTM, CNN-LSTM-Attention, and xLSTM owing to a lack of spectral feature extraction and spatial correlation learning.

GCN and ST-GCN underperformed other baseline models with respect to R^2 and CRMSE since WQPs data were continuous time series or grid-like spatial data based on spatial positions and feature correlation without direct information flow and external connectedness. If the spatial correlation among water sampling points was weak, or the data mainly consisted of deep-water time series, their spatial convolution component contributed little and might even introduce noise. GCN-based models were designed to exploit graph structures, but the construction of a water quality graph introduced assumptions that might not reflect actual spatial interactions. The graph structure and node features needed to be fully enriched to capture more detailed correlation features during information sharing and delivery. Additionally, the higher complexity and parameter count of GCN and ST-GCN increased the risk of overfitting, especially with short and noisy time series. Thus, GCN and ST-GCN might not be suited to capturing temporal dynamics and local spatial patterns without direct information connectedness.

MSM-Transformer outperformed SpecFormer and CNN-Transformer as MSM-Transformer modeled both short-term and long-term changes simultaneously through multi-scale convolution and branching structures. Additionally, MSM-Transformer focused on the information at all positions in the sequence and iterated without relying on time steps. MSM-Transformer captured the potential dependencies and characteristic patterns within the sequence and multi-scale features, modeling extracted information through sequences of different time or space scales. Thereby, MSM-Transformer simultaneously captured short-term fluctuations and long-term trends, and was suitable for complicated multi-scale sequences. SpecFormer and CNN-Transformer showed similar performance with respect to R^2 and CRMSE due to the transformer encoding and decoding mechanism, which enhanced learning capability for spectral and temporal correlation. FP-xLSTM outperformed MSM-Transformer, SpecFormer, and CNN-Transformer, as effective training of these baselines required a large dataset, and excessive noise in the data hindered feature extraction, limiting the model performance. FP-xLSTM showed the best performance among baseline models with respect to R^2 and CRMSE for all WQPs in three regions, which further proved the reasonableness of the model's design architecture and applicability to various WQPs at multiple depths.

Multiple evaluation metrics were utilized to investigate and compare the performance of FP-xLSTM and baseline models shown in Table 1, with evaluation metrics of RMSE, MAPE, R^2, and the p-value of χ^2 test as p_{χ^2}, and the p-value of Welch's t-test as p_t. The

averaged statistical results of FP-xLSTM and its baseline models showed their performance in three different regions. The best result of each row was shown in boldface corresponding to each WQP at different depths. FP-xLSTM outperformed its baseline models with respect to RMSE, MAPE, and R^2, the performance of which explicated high prediction accuracy and stability for each WQP in different regions. MAPE and R^2 were employed to compare the FP-xLSTM performance variation across different WQPs due to significant differences in magnitude and range of different WQPs. Table 1 shows model performance for FP-xLSTM and its baseline models with respect to the average RMSE, R^2, MAPE, and statistical metrics over all regions and depths. The average R^2 for TP, TN, COD, BOD, Chl-a, and TSS by FP-xLSTM was 0.81, 0.86, 0.80, 0.76, 0.91, and 0.88 respectively across three different regions, and their individual differences of R^2 from 0.5 m to 4.0 m were 0.15, 0.11, 0.14, 0.19, 0.07, and 0.10. The corresponding descent of R^2 s for WQPs by FP-xLSTM was 17.05%, 11.96%, 16.09%, 22.35%, 7.37%, and 10.75%, respectively, the mean and standard deviation of which were 14.26% and 4.86%, illustrating stable prediction accuracy variation across different depths. MAPEs of FP-xLSTM increased by 5.11%, 5.60%, 5.47%, 5.13%, 4.88%, and 4.48% for TP, TN, COD, BOD, Chl-a, and TSS, respectively, from a depth of 0.5 m to 4.0 m, which averaged 5.11%, showing efficient prediction over multi-depth WQPs.

Moreover, MAPE of FP-xLSTM increased by 2.85% from 0.5 m to 2.0 m and 2.27% from 2.0 m to 4.0 m. R^2 of FP-xLSTM decreased by 0.07 from 0.5 m to 2.0 m and 0.06 from 2.0 m to 4.0 m. This elucidated that FP-xLSTM performance declined as depth increased. Ablation analysis was conducted by removing spatial correlation, transport impact, feature interaction, and PMF mechanisms from FP-xLSTM, which were denoted as Fx/SC, Fx/TI, Fx/FI, and Fx/PMF. The corresponding statistical results are shown in Table 1. MAPE of Fx/SC increased by 2.63%, 3.45%, and 5.47%, and R^2 of Fx/SC decreased by 0.08, 0.09, and 0.11 on average at the depths of 0.5 m, 2.0 m, and 4.0 m compared to FP-xLSTM. MAPE of Fx/TI increased by 2.34%, 5.78%, and 8.67%, and R^2 of Fx/SC decreased by 0.06, 0.13, and 0.16 on average at the depths of 0.5 m, 2.0 m, and 4.0 m compared to FP-xLSTM. MAPE of Fx/FI increased by 2.87%, 4.63%, and 7.06%, and R^2 of Fx/FI decreased by 0.10, 0.11, and 0.14 on average at the depths of 0.5 m, 2.0 m, and 4.0 m compared to FP-xLSTM. MAPE of Fx/PMF increased by 1.38%, 2.39%, and 3.18%, and R^2 of Fx/PMF decreased by 0.04, 0.05, and 0.08 on average at the depths of 0.5 m, 2.0 m, and 4.0 m compared to FP-xLSTM. The impact imposed upon model performance by removing each module according to MAPE and R^2 was ranked: Fx/TI > Fx/FI > Fx/SC > Fx/PMF. Fx/TI had a relatively weak influence on model performance at a depth of 0.5 m, whereas Fx/SC and Fx/FI exerted stronger influences. This might be attributed to the fact that feature interaction and spatial correlation facilitated more effective information sharing and propagation among neighboring features than feature impact transport at the horizontal level. Fx/TI exerted a stronger influence on model performance than Fx/SC and Fx/FI at depths of 2.0 m and 4.0 m, which might be attributed to the attenuation of optical signals with increasing depth, thereby reducing the effectiveness of information sharing and propagation across the vertical dimension. PMF variant module imposed the least impact on model performance since detected outlier concentrations accounted for a small fraction of all resulting concentrations. The effectiveness of all the aforementioned modules in improving model performance was demonstrated by the corresponding changes in MAPE and R^2.

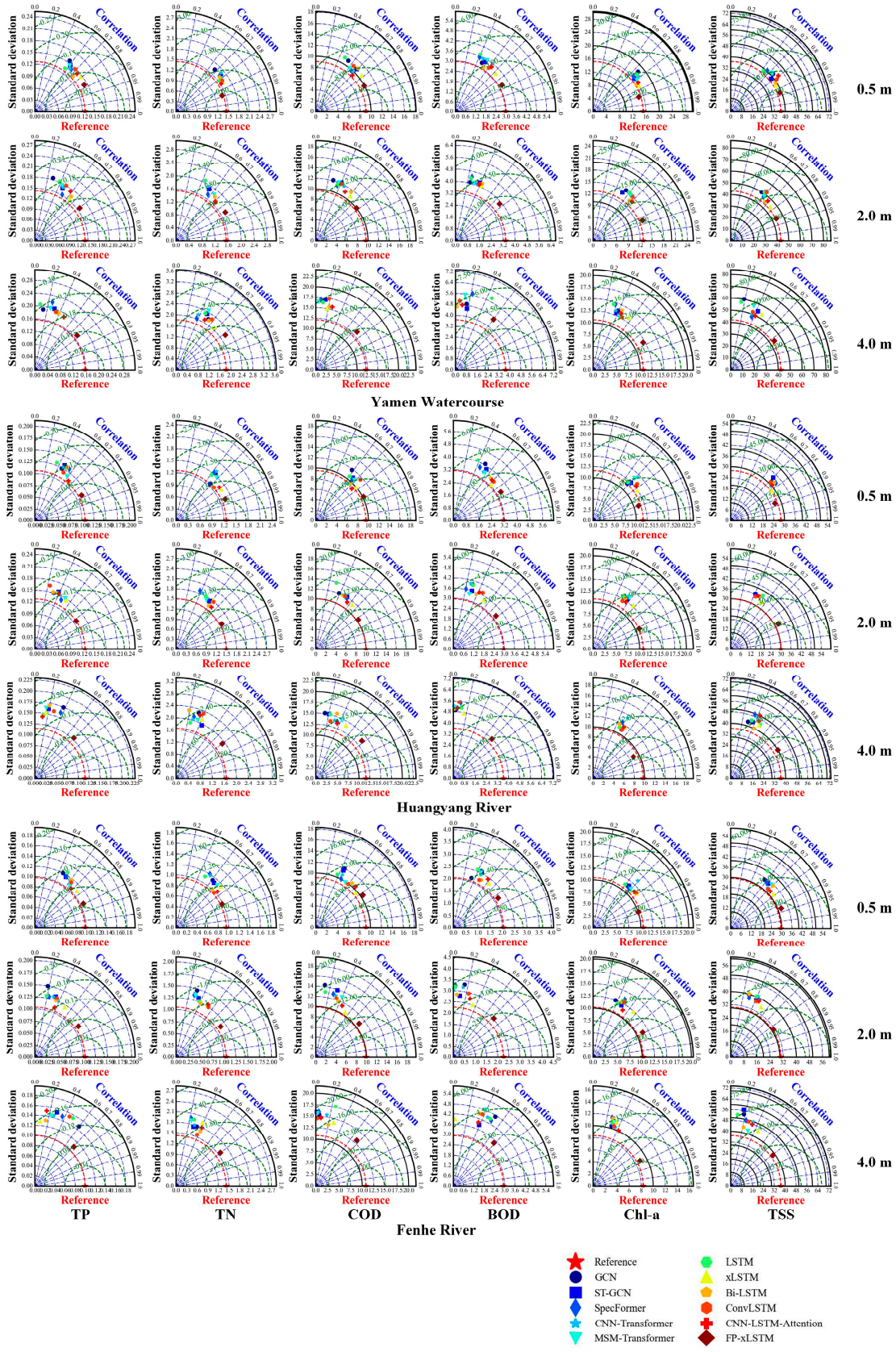


Figure 8. Taylor diagram of FP-xLSTM and baseline models. The blue radial dashed line denotes correlation between measured and predicted values, and the green dashed line denotes CRMSE between measured and predicted values. The reference denotes the standard deviation of measured values.

Table 1. Performance comparison between FP-xLSTM and baseline models.

WQPs	Depth	Evaluation Metrics	GCN	ST-GCN	SpecFormer	CNN-Transformer	MSM-Transformer	LSTM	xLSTM	Bi-LSTM	ConvLSTM	CNN-LSTM-Attention	Fx/SC	Fx/TI	Fx/FI	Fx/PMF	FP-xLSTM
TP	0.5 m	RMSE (mg/L)	0.08	0.07	0.07	0.07	0.06	0.08	0.06	0.07	0.07	0.06	0.05	0.05	0.05	0.04	**0.04**
		MAPE (%)	25.8	23.68	22.88	22.76	22.18	24.87	18.09	20.9	22.3	19.34	14.65	14.94	15.32	13.27	**11.90**
		R^2	0.49	0.53	0.57	0.56	0.61	0.48	0.72	0.62	0.61	0.69	0.79	0.81	0.76	0.85	**0.88**
		p_{χ^2}	0.21	0.24	0.14	0.07	0.51	0.41	0.47	0.33	0.08	0.38	0.55	0.48	0.52	0.79	0.82
		p_t	0.65	0.84	0.73	0.51	0.64	0.66	0.63	0.58	0.40	0.56	0.62	0.55	0.49	0.70	0.75
	2.0 m	RMSE (mg/L)	0.11	0.10	0.09	0.10	0.09	0.11	0.08	0.09	0.10	0.09	0.06	0.07	0.07	0.05	**0.05**
		MAPE (%)	30.12	28.00	26.55	26.34	25.72	28.27	22.86	26.15	28.51	25.35	18.66	21.75	20.98	17.59	**14.91**
		R^2	0.26	0.39	0.40	0.39	0.44	0.31	0.56	0.43	0.36	0.45	0.71	0.65	0.68	0.76	**0.81**
		p_{χ^2}	0.28	0.28	0.21	0.10	0.22	0.24	0.40	0.32	0.25	0.35	0.55	0.46	0.53	0.57	0.63
		p_t	0.64	0.69	0.27	0.42	0.47	0.61	0.66	0.45	0.76	0.39	0.62	0.42	0.56	0.51	0.56
	4.0 m	RMSE (mg/L)	0.14	0.12	0.13	0.12	0.11	0.13	0.11	0.12	0.13	0.11	0.08	0.10	0.09	0.07	**0.06**
		MAPE (%)	35.46	32.4	32.76	31.13	30.1	34.55	29.41	30.49	33.62	29.68	22.08	24.87	23.38	20.53	**17.01**
		R^2	0.23	0.25	0.3	0.26	0.31	0.24	0.36	0.32	0.24	0.34	0.64	0.59	0.61	0.66	**0.73**
		p_{χ^2}	0.01	0.07	0.11	0.25	0.09	0.14	0.16	0.20	0.09	0.26	0.68	0.63	0.66	0.75	0.82
		p_t	0.31	0.44	0.45	0.54	0.59	0.61	0.33	0.52	0.43	0.56	0.64	0.59	0.62	0.42	0.58
TN	0.5 m	RMSE (mg/L)	0.72	0.65	0.70	0.68	0.62	0.69	0.53	0.62	0.57	0.54	0.38	0.38	0.39	0.36	**0.34**
		MAPE (%)	23.36	22.54	25.06	24.84	21.88	24.48	18.23	21.44	18.97	18.87	14.53	14.27	15.73	13.25	**11.85**
		R^2	0.64	0.70	0.64	0.65	0.73	0.64	0.80	0.73	0.77	0.79	0.83	0.84	0.78	0.87	**0.92**
		p_{χ^2}	0.42	0.66	0.40	0.43	0.57	0.46	0.82	0.57	0.75	0.78	0.73	0.70	0.68	0.75	0.80
		p_t	0.67	0.84	0.92	0.69	0.79	0.76	0.75	0.59	0.56	0.93	0.69	0.65	0.63	0.66	0.70

Table 1. *Cont.*

WQPs	Depth	Evaluation Metrics	GCN	ST-GCN	SpecFormer	CNN-Transformer	MSM-Transformer	LSTM	xLSTM	Bi-LSTM	ConvLSTM	CNN-LSTM-Attention	Fx/SC	Fx/TI	Fx/FI	Fx/PMF	FP-xLSTM
TN	2.0 m	RMSE (mg/L)	0.95	0.95	1.02	1.03	0.94	1.07	0.84	0.88	0.81	0.82	0.57	0.60	0.59	0.54	**0.51**
		MAPE (%)	27.49	27.06	31.44	31.04	28.29	30.7	25.44	26.81	24.9	24.58	18.99	21.25	20.16	17.82	**15.37**
		R^2	0.48	0.49	0.41	0.42	0.49	0.44	0.60	0.56	0.63	0.66	0.77	0.72	0.74	0.80	**0.85**
		p_{χ^2}	0.62	0.46	0.28	0.55	0.46	0.32	0.76	0.64	0.25	0.65	0.75	0.69	0.73	0.79	0.87
		p_t	0.70	0.79	0.71	0.83	0.87	0.61	0.73	0.87	0.65	0.85	0.49	0.46	0.53	0.51	0.56
	4.0 m	RMSE (mg/L)	1.25	1.25	1.36	1.34	1.29	1.32	1.14	1.26	1.2	1.24	0.82	0.90	0.85	0.77	**0.72**
		MAPE (%)	32.75	31.46	36.28	35.96	33.51	36.07	30.15	33.02	29.41	30.57	22.19	25.23	24.05	20.89	**17.45**
		R^2	0.40	0.40	0.29	0.31	0.35	0.32	0.49	0.38	0.45	0.42	0.70	0.65	0.67	0.73	**0.81**
		p_{χ^2}	0.54	0.75	0.32	0.44	0.41	0.31	0.87	0.29	0.53	0.68	0.73	0.69	0.64	0.81	0.85
		p_t	0.48	0.69	0.77	0.78	0.77	0.58	0.66	0.74	0.44	0.71	0.42	0.39	0.37	0.44	0.46
COD	0.5 m	RMSE (mg/L)	6.38	5.95	5.87	5.76	5.58	6.58	4.44	5.29	5.39	4.91	4.08	3.93	4.02	3.72	**3.35**
		MAPE (%)	25.96	23.64	23.38	23.03	22.34	26.1	18.68	20.53	22.04	19.86	15.78	15.43	15.62	14.87	**13.59**
		R^2	0.52	0.58	0.59	0.60	0.62	0.48	0.77	0.67	0.65	0.72	0.79	0.82	0.81	0.84	**0.87**
		p_{χ^2}	0.15	0.29	0.58	0.58	0.76	0.16	0.67	0.61	0.20	0.56	0.63	0.66	0.64	0.68	0.72
		p_t	0.79	0.77	0.61	0.76	0.73	0.60	0.79	0.49	0.60	0.48	0.77	0.74	0.72	0.79	0.81
	2.0 m	RMSE (mg/L)	8.38	7.52	7.64	7.48	7.04	8.67	6.21	7.12	7.5	6.85	5.29	5.92	5.52	4.96	**4.43**
		MAPE (%)	30.08	28.37	28.24	27.99	26.97	31.01	23.29	26.41	28.25	25.33	19.33	21.96	20.19	18.89	**16.54**
		R^2	0.26	0.4	0.38	0.41	0.43	0.25	0.59	0.47	0.41	0.51	0.72	0.66	0.69	0.74	**0.79**
		p_{χ^2}	0.02	0.04	0.20	0.07	0.11	0.03	0.39	0.26	0.06	0.28	0.58	0.55	0.57	0.61	0.64
		p_t	0.63	0.81	0.56	0.87	0.77	0.53	0.57	0.4	0.45	0.47	0.44	0.46	0.41	0.55	0.62

Table 1. *Cont.*

WQPs	Depth	Evaluation Metrics	GCN	ST-GCN	SpecFormer	CNN-Transformer	MSM-Transformer	LSTM	xLSTM	Bi-LSTM	ConvLSTM	CNN-LSTM-Attention	Fx/SC	Fx/TI	Fx/FI	Fx/PMF	FP-xLSTM
COD	4.0 m	RMSE (mg/L)	11.09	10.77	10.46	10.59	10.16	11.6	9.43	9.85	10.6	9.98	7.83	8.87	8.24	7.22	**5.94**
		MAPE (%)	34.81	32.83	32.97	32.73	32.23	36.96	29.8	31.41	33.38	30.81	24.36	26.95	25.12	22.58	**19.06**
		R^2	0.09	0.12	0.16	0.23	0.17	0.19	0.32	0.24	0.15	0.23	0.61	0.57	0.60	0.66	**0.73**
		p_{χ^2}	0.13	0.18	0.11	0.18	0.17	0.09	0.34	0.21	0.14	0.24	0.38	0.33	0.36	0.41	0.45
		p_t	0.67	0.82	0.54	0.62	0.77	0.69	0.61	0.73	0.51	0.56	0.49	0.45	0.48	0.53	0.62
BOD	0.5 m	RMSE (mg/L)	2.01	1.9	2.04	1.94	1.88	2.12	1.48	1.76	1.67	1.63	1.23	1.21	1.26	1.13	**1.04**
		MAPE (%)	26.17	25	26.66	24.76	24.09	27.03	20.6	23.99	22.63	22.15	15.42	15.21	15.68	14.97	**13.63**
		R^2	0.44	0.50	0.43	0.48	0.51	0.39	0.69	0.56	0.60	0.63	0.79	0.81	0.77	0.82	**0.85**
		p_{χ^2}	0.20	0.34	0.10	0.31	0.33	0.17	0.49	0.27	0.51	0.30	0.56	0.58	0.53	0.59	0.62
		p_t	0.48	0.58	0.47	0.60	0.55	0.76	0.50	0.46	0.56	0.59	0.60	0.61	0.59	0.64	0.70
	2.0 m	RMSE (mg/L)	2.49	2.4	2.6	2.45	2.52	2.56	2.13	2.34	2.25	2.16	1.75	1.88	1.83	1.62	**1.39**
		MAPE (%)	30.1	29.61	31.4	29.82	30.02	31.25	25.86	29.42	27.76	27.22	19.83	21.52	20.37	18.72	**16.44**
		R^2	0.23	0.26	0.16	0.24	0.19	0.19	0.43	0.29	0.36	0.38	0.68	0.64	0.67	0.70	**0.76**
		p_{χ^2}	0.07	0.11	0.03	0.09	0.1	0.07	0.17	0.04	0.06	0.27	0.74	0.69	0.72	0.77	0.83
		p_t	0.72	0.5	0.38	0.71	0.38	0.75	0.38	0.29	0.62	0.69	0.45	0.39	0.43	0.48	0.50
	4.0 m	RMSE (mg/L)	3.8	3.66	3.96	3.55	3.27	3.79	3.32	3.57	3.46	3.39	2.59	2.93	2.68	2.39	**1.97**
		MAPE (%)	35.92	35.11	37.05	35.44	34.98	36.9	31.26	35.04	34.19	33.83	24.29	26.87	25.58	22.31	**18.76**
		R^2	0.24	0.22	0.27	0.2	0.28	0.35	0.39	0.36	0.33	0.37	0.55	0.51	0.53	0.58	**0.66**
		p_{χ^2}	0.02	0.01	0.02	0.03	0.02	0.03	0.10	0.01	0.07	0.10	0.41	0.36	0.39	0.44	0.49
		p_t	0.48	0.21	0.47	0.45	0.50	0.78	0.39	0.17	0.27	0.31	0.17	0.15	0.14	0.18	0.20

Table 1. *Cont.*

WQPs	Depth	Evaluation Metrics	GCN	ST-GCN	SpecFormer	CNN-Transformer	MSM-Transformer	LSTM	xLSTM	Bi-LSTM	ConvLSTM	CNN-LSTM-Attention	Fx/SC	Fx/TI	Fx/FI	Fx/PMF	FP-xLSTM
Chl-a	0.5 m	RMSE (ug/L)	6.88	6.57	6.63	6.34	6.05	6.94	5.04	5.49	5.37	5.28	3.31	3.19	3.24	3.02	**2.82**
		MAPE (%)	23.51	22.37	22.55	21.42	20.99	23.24	17.3	19.09	18.58	18.29	12.59	11.87	12.24	10.65	**9.33**
		R^2	0.68	0.71	0.71	0.73	0.75	0.67	0.83	0.78	0.77	0.79	0.86	0.89	0.87	0.91	**0.95**
		p_{χ^2}	0.48	0.66	0.77	0.24	0.56	0.31	0.83	0.79	0.44	0.60	0.59	0.62	0.60	0.64	0.67
		p_t	0.75	0.85	0.81	0.58	0.96	0.55	0.82	0.77	0.76	0.92	0.32	0.35	0.34	0.36	0.39
	2.0 m	RMSE (ug/L)	8.09	7.75	7.38	7.16	7.04	7.98	6.28	7.34	7.31	6.97	4.19	4.99	4.53	3.68	**3.33**
		MAPE (%)	28.71	28.05	27.86	26.72	26.2	28.48	22.15	25.61	25.16	24.55	15.27	17.89	16.37	13.95	**11.81**
		R^2	0.47	0.52	0.56	0.57	0.59	0.51	0.68	0.56	0.58	0.62	0.81	0.73	0.78	0.85	**0.91**
		p_{χ^2}	0.61	0.71	0.28	0.37	0.65	0.57	0.59	0.74	0.6	0.67	0.69	0.58	0.64	0.74	0.78
		p_t	0.83	0.72	0.77	0.53	0.86	0.70	0.83	0.76	0.51	0.52	0.48	0.42	0.49	0.55	0.59
	4.0 m	RMSE (ug/L)	7.35	7.27	7.62	7.04	6.95	7.68	6.78	7.41	7.07	6.99	4.68	5.65	5.13	3.99	**3.30**
		MAPE (%)	32.65	31.06	33.85	30.68	30.52	33.30	29.05	32.42	31.43	30.19	19.95	24.73	22.52	17.39	**14.21**
		R^2	0.40	0.39	0.35	0.42	0.44	0.35	0.48	0.39	0.44	0.46	0.77	0.69	0.71	0.81	**0.88**
		p_{χ^2}	0.35	0.27	0.10	0.29	0.23	0.28	0.21	0.22	0.31	0.22	0.65	0.62	0.67	0.72	0.78
		p_t	0.59	0.79	0.76	0.56	0.67	0.72	0.73	0.78	0.76	0.95	0.18	0.22	0.16	0.23	0.26
TSS	0.5 m	RMSE (mg/L)	17.52	16.81	17.6	16.56	16.2	18.72	13.2	14.87	15.71	15.18	9.78	9.55	9.67	8.98	**8.32**
		MAPE (%)	22.9	21.62	22.32	21.29	20.92	23.16	17.13	18.89	19.82	18.61	13.69	13.21	13.52	12.16	**10.59**
		R^2	0.68	0.70	0.68	0.72	0.73	0.64	0.82	0.76	0.74	0.77	0.85	0.87	0.84	0.89	**0.93**
		p_{χ^2}	0.35	0.79	0.62	0.51	0.49	0.51	0.87	0.85	0.78	0.84	0.81	0.83	0.84	0.87	0.91
		p_t	0.84	0.92	0.88	0.86	0.92	0.65	0.68	0.66	0.88	0.79	0.62	0.64	0.66	0.69	0.70

Table 1. *Cont.*

WQPs	Depth	Evaluation Metrics	GCN	ST-GCN	SpecFormer	CNN-Transformer	MSM-Transformer	LSTM	xLSTM	Bi-LSTM	ConvLSTM	CNN-LSTM-Attention	Fx/SC	Fx/TI	Fx/FI	Fx/PMF	FP-xLSTM
TSS	2.0 m	RMSE (mg/L)	26.03	24.7	23.89	23.97	23.47	24.27	20.76	24.09	23.43	22.92	15.43	16.95	16.33	14.39	**12.06**
		MAPE (%)	27.71	26.66	25.58	26	25.57	27.38	22.41	25.32	24.98	24.53	16.58	18.23	17.69	15.32	**12.89**
		R^2	0.41	0.47	0.50	0.49	0.52	0.49	0.61	0.47	0.51	0.55	0.78	0.73	0.76	0.82	**0.87**
		p_{χ^2}	0.40	0.60	0.47	0.37	0.39	0.28	0.64	0.45	0.63	0.51	0.82	0.71	0.78	0.87	0.89
		p_t	0.82	0.82	0.69	0.95	0.86	0.74	0.68	0.77	0.86	0.85	0.58	0.52	0.55	0.61	0.64
	4.0 m	RMSE (mg/L)	33.85	31.93	31.56	30.59	30.25	34.3	27.91	31.14	30.71	29.56	22.37	25.62	24.39	19.58	**15.65**
		MAPE (%)	33.99	32.93	31.97	31.09	30.6	33.77	28.46	30.92	30.25	29.39	21.87	24.92	23.25	18.95	**15.07**
		R^2	0.22	0.32	0.35	0.36	0.38	0.2	0.46	0.39	0.37	0.39	0.69	0.65	0.66	0.75	**0.83**
		p_{χ^2}	0.25	0.37	0.32	0.34	0.33	0.17	0.47	0.34	0.27	0.39	0.77	0.72	0.75	0.82	0.86
		p_t	0.60	0.76	0.77	0.84	0.81	0.86	0.85	0.60	0.78	0.64	0.55	0.51	0.47	0.57	0.61

The overall performance of FP-xLSTM on OAC WQPs was prioritized over that on non-OAC WQPs, indicating that spectral features were conducive to establishing information flow and correlating spatial neighborhoods, which further improved prediction accuracy and model robustness. The decline in model performance on OAC WQPs (Chl-a and TSS) with increasing water depth could be primarily attributed to the attenuation of optical signals within the water column. As depth increased, light underwent progressively stronger absorption and scattering, progressively attenuating the spectral information originating from deeper water layers before it reached the water surface and the UAV sensor. Consequently, the spectral information related to WQPs became weaker and more difficult to distinguish from background optical effects. Meanwhile, the reduced availability of depth-dependent spectral information limited the effectiveness of feature interaction and information propagation across the vertical dimension. These effects become more pronounced at greater depths, leading to increased prediction errors and a gradual decline in model performance. The depth-dependent decline in prediction performance for non-OAC WQPs (TP, TN, COD, and BOD) was mainly associated with their weak and indirect spectral responses. Because these parameters did not exhibit strong and distinctive optical signatures, their spectral information was largely represented through indirect relationships with optically active substances and other water quality characteristics. With increasing depth, absorption and scattering progressively attenuated the optical signals originating from deeper water layers, thereby weakening these indirect spectral relationships. The reduced reliability of depth-specific spectral information consequently limited the effectiveness of feature interaction and information propagation across the vertical dimension. This effect was particularly important for non-OAC WQPs, for which the available spectral information was already relatively weak, ultimately leading to greater prediction uncertainty and a decline in model performance at greater depths.

All p_{χ^2}s of FP-xLSTM were greater than 0.05, and the null hypothesis was not rejected, indicating that there was no significant difference between the distributions of classification results for measured and predicted values. Additionally, p_ts were greater than 0.05, and the corresponding null hypothesis was not rejected, demonstrating that the center of predicted values by FP-xLSTM was equal to that of measured values. p_{χ^2}s of baseline models were greater than 0.05 at the depths of 0.5 m and 2.0 m, but some were less than 0.05 at the depth of 4.0 m. This signified that there was no significant difference between distributions of classification results for predicted values by baseline models and measured values at the depth of 0.5 m, and predicted values of baseline models deviated significantly from measured values as depth increased. p_ts of baseline models were greater than 0.05, indicating the center of predicted values by baseline models was equal to that of measured values. These two hypothesis tests showed model performance on predicting WQPs in one region by using datasets from two other regions to train the model, which proved strong cross-region transfer learning capability of FP-xLSTM.

Moreover, xLSTM and CNN-LSTM-Attention models outperformed other baseline models with respect to RMSE, MAPE, and R^2 since they were extensions of the classic LSTM and were used to handle multivariate time series (multi-channel data). They were relatively flexible for multi-variable interaction and dynamically focused on important time steps or variables, which automatically assigned weights to key time steps or key variables. Thus, they were able to capture independent features of multiple variables and were suitable for applications with diverse parameters and pronounced differences in water quality attributes. Additionally, xLSTM and CNN-LSTM-Attention established spatio-temporal correlation and interaction of spectral features using information sharing and delivery. The MAPEs of xLSTM and CNN-LSTM-Attention were raised by 11.35% and 11.23%, respectively, on average for all WQPs from a depth of 0.5 m to 4.0 m, and

their performances were analogous. Their R^2 s decreased by 0.36, which showed relatively unstable prediction accuracy and low transfer learning capability compared with FP-xLSTM. Their best performance was observed on OAC WQPs, including Chl-a and TSS at the depth of 0.5 m, with performance declining sharply at greater depths, reflecting their dependence on optical information and its attenuation with depth. Their prediction accuracy on non-OAC WQPs was lower than that on OAC WQPs at the depth of 0.5 m but gradually approached that of OAC WQPs as depth increased. This was because optical characteristics did not directly reflect the variation of OAC WQPs, and prediction of all WQPs might be contingent on spectral feature correlation and interaction at depths greater than 0.5 m.

MSM-Transformer outperformed CNN-Transformer, SpecFormer, GCN, and ST-GCN with respect to RMSE, MAPE, and R^2 since it leveraged multi-channel data to establish an independent channel for each feature, and its self-attention mechanism captured complex nonlinear correlations among WQPs with multi-scale feature extraction. Thereby, MSM-Transformer was applicable to retrieve short-term variation of WQPs with high efficiency. CNN-Transformer and SpecFormer showed similar prediction performance due to their analogous architectures, which extracted spectral features and analyzed spatio-temporal correlations of WQPs via convolution and feature coupling. ST-GCN showed superiority over GCN since it considered spatio-temporal correlation of each WQP and spatial distribution resemblance of different WQPs. Above all, FP-xLSTM outperformed baseline models since the model structure design established internal and external multi-feature interaction, information flow in spatio-temporal distribution, multi-depth substance transport mechanism, and feedback mechanism. p_{χ^2}s of some methods for WQPs at depths greater than 0.5 m were smaller than 0.05, indicating that significant deviation existed between the distributions of classification results of measured and predicted values, and corresponding methods might not be suitable for predicting these WQPs [40,41].

4.2. Impact of Land Cover Features upon WQPs Retrieval

Based on the inverse distance-weighted land cover density (idLCD) index, the Liang–Kleeman information flow framework exhibited directional information transfer between dynamic variables and was employed to quantify causal information transfer between riparian land features and WQP concentrations. The idLCD assigned greater weights to land-cover objects closer to water bodies, reflecting stronger hydrological connectivity and pollutant transport potential. Different riparian land features exhibited varying information flow rates with WQPs due to differences in runoff generation, nutrient export, pollutant retention, and ecological buffering capacities. Consequently, anthropogenic land-cover types generally showed stronger causal influences on WQP variability, whereas natural vegetation features exhibited relatively weaker impacts. In this study, we conducted an extensive experiment to investigate the influence of riparian spatial features on multi-depth WQP retrieval through Liang–Kleeman information flow.

$$idLCD_i = \sum \frac{S_j}{Dis_j \cdot S_t} \tag{36}$$

$$T_{j \to i} = \frac{C_{x_i,\dot{x}_i} C_{x_j,x_i} - C_{x_j,\dot{x}_i} C_{x_i,x_i}}{C_{x_i,x_i} C_{x_j,x_j} - C^2_{x_i,x_j}} \tag{37}$$

where j refers to an individual patch belonging to land-use type i; Dis_j denotes the distance from the land feature object to the water body; S_j denotes the area of the corresponding land feature object; and S_t indicates the total land area. The idLCD values for different riparian land features could then be derived at various spatial scales. $T_{j \to i}$ is the information

flow rate from riparian land features x_j to WQPs concentrations x_i; C_{x_j,x_i} is the covariance between x_j and x_i; $\dot{x}_i$ is gradient in x_i; x is a series derived from idLCD; and a non-zero $T_{j\to i}$ is a causal relationship between x_j and x_i where a zero value indicates the relational absence. The magnitude of $T_{j\to i}$ indicated the strength of the causal relationship, with larger values indicating stronger causal influences. A 1 km riparian buffer zone was established according to the Technical Guidelines for Protection and Restoration of River and Lake Ecological Buffer Zones issued by the Ministry of Ecology and Environment of China. Based on this buffer setting, the relationships between riparian spatial features and WQP concentrations were systematically investigated, providing scientific support for urban land-use planning and ecological management optimization as shown in Figure A1. The details about the impact of land cover features on WQP retrieval at depths of 0.5, 2.0, and 4.0 m were demonstrated in Appendix A.11.

4.3. Visualization of Multi-Depth WQPs Distributions

The spatial distribution of multi-depth WQPs through FP-xLSTM was shown in Figure A2a. In Yamen Watercourse, water flowed from north to south during October to December. TP and TN were rapidly accumulated in the lower stream, where residential and agricultural areas were densely distributed, causing domestic waste, including phosphate-containing laundry detergent, farm waste, and phosphate and nitrogenous fertilizer, to enter the river. The accumulation of TP and TN gradually increased from 0.5 m to 4.0 m since strong sunlight, direct atmospheric contact, and wind-induced agitation promoted thorough mixing at the surface layer (0.5 m), enhancing nutrient redistribution in the water column. Additionally, abundant sunlight at the water surface promoted vigorous photosynthesis by phytoplankton, which strongly absorbed nitrogen and phosphorus, resulting in decreased TP and TN concentrations. At depths greater than 0.5 m, limited light and weak water disturbance usually occurred, while slow water flow allowed sediment accumulation and the release of organic matter and nutrients, resulting in nutrient-rich areas. Under anoxic conditions in deep water, sediments released TP and NH4+, leading to increased concentrations of TP and TN.

COD represented total organic matter, including non-biodegradable components, while BOD reflected the biodegradable fraction. Concentrations of COD and BOD were relatively high in the upper stream where industrial areas, agricultural areas, and fishing areas were densely distributed. Water volume was usually small, and the flow rate was unstable, causing weak dilution capacity and consequently leading to increased concentrations of COD and BOD. Meanwhile, COD and BOD were massively accumulated near riverbanks because urban surface runoff, slow flow rate, areas of backflow or stagnation, human disturbances, plant residues, and soil organic matter flow near riverbanks caused retention of organic matter and increased degradable organic matter. Furthermore, lower COD and BOD concentrations in the surface layer (0.5 m) could be attributed to higher DO levels, stronger hydrodynamic mixing, and enhanced microbial degradation. Additionally, low DO conditions in the deeper layer can inhibit microbial degradation, further contributing to higher organic matter concentrations. In contrast, the elevated concentrations in the deeper layer (e.g., 2.0 m and 4.0 m) were mainly associated with sediment release, weaker water exchange, and DO-limited conditions that inhibited organic matter decomposition, facilitating the accumulation of organic pollutants.

Chl-a of high concentration mainly aggregated near the lower stream, as large masses of TP and TN serving as essential nutrients for phytoplankton growth promoted quick accumulation of Chl-a. The proposed method presented high capability of feature capturing and learning, insomuch as predicted values of Chl-a concentration by FP-xLSTM covered a wider range than the measured values of the training dataset. This signified that the

proposed method was capable of predicting Chl-a of high and low concentration beyond the range of sampled data. The sunlight penetration in the water body decayed in the upper stream due to high turbidity, but sedimentation occurred in the lower stream, causing clearer water, which was beneficial to sunlight passage through the water body. Tidal action in the lower stream close to the estuary caused water bodies to flow back and forth, increasing retention time. Thus, Chl-a accumulated quickly near the lower stream, the spatial distribution of which resembled that of TP and TN to some extent. The concentration of Chl-a decreased with increasing depth because reduced sunlight penetration, low DO volume, and low biological activity inhibited the growth of phytoplankton in the deeper layer, leading to decreased Chl-a concentrations. As a result, Chl-a concentrations were significantly higher in the surface layer, while TN and TP were relatively lower due to biological consumption. This indicated a clear vertical decoupling, where the deeper layer acted as a nutrient regeneration zone, and the surface layer functioned as a primary production zone.

The concentration of TSS was higher near the upstream, corresponding to Chl-a of lower concentration near the upper stream since high TSS concentration caused high turbidity, lowering sunlight penetration in the water body, which reduced Chl-a accumulation. The higher TSS concentrations observed in the upstream region could be attributed to stronger soil erosion and enhanced resuspension, which increased the sediment load and runoff particles in suspension. The concentration of TSS increased with depth due to carried-away particles in surface water [42], gravitational settling of particles, and accumulation and resuspension in deeper water. The spatial distribution of TSS was partially analogous to that of COD and BOD, reflecting the association between organic matter and suspended particles. A significant portion of COD and BOD originated from particulate organic matter (POM), which was itself a component of TSS, highlighting the coupling mechanism between sediment release and organic pollutant distribution. However, there were some parts of the watercourse where TSS of higher concentration and COD of lower concentration concurred. This was mainly because dissolved organic matter (DOM) dominated, and BOD was rapidly depleted due to the highly degradable conditions caused by industrial effluents in the upstream region.

In Figure A2b, water of the Huangyang River flowed from north to south and ultimately discharged into the Pearl River Estuary. TP and TN showed similar concentration distributions. TP and TN of relatively high concentration occurred in the middle and lower stream due to densely distributed agricultural and residential areas discharging effluent from riverbanks into the river. Non-point-source pollution was mainly discharged into the river in the upper stream due to relatively high altitude. The higher TN and TP concentrations at the depth of 0.5 m were mainly driven by external inputs. The elevated concentrations observed at depths greater than 0.5 m were associated with sediment accumulation and the subsequent release of nutrients under weaker hydrodynamic conditions. Tidal action near the estuary at deeper layers in the middle and lower stream caused back-and-forth flow of water and extended retention time, which promoted nutrient salt deposition and released sediment, elevating TP and TN levels.

The concentration distribution difference between COD and BOD was mainly attributed to the concentration distribution difference of non-biodegradable organic matter. COD aggregated at relatively high concentrations near the upper stream and riverbanks in shallow water due to the aggregation of large masses of non-biodegradable organic matter near the upper stream. Quick accumulation of COD along riverbanks of the middle and lower stream at a depth of 0.5 m was attributed to plastic, chemical fertilizer, and dye plants mainly distributed near the middle and lower stream. Additionally, water flow rate near riverbanks was slower than in the river center, causing retention of discharged organic

matter and elevating COD levels. Profuse accumulation of BOD occurred in the middle and lower stream owing to enhanced microbial degradation and the transformation of organic matter into more biodegradable forms during downstream transport. The similar distribution patterns of COD and TN were attributed to their common sources, coupled transport processes, and shared association with suspended particles. Both were influenced by sedimentation–resuspension dynamics and hydrodynamic conditions, resulting in consistent spatial variations along the river.

Chl-a concentrations peaked at 0.5 m and gradually decreased with depth because favorable light conditions in surface waters promoted phytoplankton growth, whereas light limitation in deeper layers restricted algal production despite relatively higher nutrient availability. The vertical distribution of Chl-a was consistent with those of TP and TN, reflecting their roles as essential nutrients for phytoplankton growth. In contrast, TSS exhibited a different distribution pattern because it was primarily controlled by hydrodynamic conditions and sediment transport rather than light and nutrient availability. Elevated TSS reduced light penetration, thereby suppressing phytoplankton growth. Higher TSS concentrations in the middle and lower reaches were likely associated with industrial discharges from concrete, machinery, and textile plants. The co-occurrence of high TSS and Chl-a in deeper waters may result from sediment resuspension, flocculation between suspended particles and algae, and hydrodynamic retention in estuarine, reflux, and tidal zones.

In Figure A2c, high-level TP and TN in the Fenehe River occurred near the middle and upper stream, where untreated wastewater, agricultural waste, and fishery farm waste were mainly discharged into the river. TP and TN concentrations gradually increased overall with increasing depth, but TN level rose slowly from 0.5 m to 2.0 m, indicating unclear stratification of the water body in this range. Sediment release, weaker hydrodynamic mixing, low flow rate, and oxygen-limited conditions near the bottom were beneficial for TN and TP concentrations at deeper layers. Nitrogen- and phosphorus-containing particulate matter (organic debris, flocs) kept sinking, which caused TP and TN levels to increase with increasing depth. While the surface layer acted as a nutrient consumption zone due to active phytoplankton uptake, the deeper layer served as a nutrient regeneration zone, resulting in higher nutrient concentrations.

COD of high concentration mainly aggregated near the middle and upper stream and increased with increasing depth due to sediment resuspension, release of organic matter from bottom sediments, and reduced degradation under low DO conditions. In contrast, stronger mixing and higher microbial activity in the surface layer promoted the removal and dilution of organic matter. The concentration distribution difference between COD and BOD was mainly caused by varying composition and biodegradability of organic matter. While COD represented total organic matter including non-biodegradable components, BOD reflected the biodegradable fraction, which was strongly influenced by microbial activity, hydrodynamic conditions, and in-stream transformation processes. The higher BOD concentrations observed near the riverbanks could be attributed to enhanced external inputs, including domestic sewage discharge, agricultural runoff (e.g., fertilizer and organic matter), and runoff carrying organic debris. Furthermore, granular deposition and accumulation of organic debris usually occurred near riverbanks due to a lower water flow rate compared to the river center. Riverbanks were abundant in organic substrate and aquatic plants beneficial to microbial attachment and biofilm formation, enhancing the degradation process of organic matter and increasing BOD level.

High-level Chl-a mainly occurred near the upper stream, the spatial distribution of which resembled that of TP and TN, as they served as main nutrients for algal growth. Chl-a level gradually descended from a depth of 0.5 m to 4.0 m, mainly resulting from less sunlight availability constraining algal photosynthesis and growth. The co-occurrence of

high TSS and Chl-a concentrations in the upper stream region was attributed to strong external inputs (e.g., soil erosion, agricultural runoff including nutrients for algae, and sewage discharge including organic matter and microorganisms), sediment resuspension (e.g., TSS, deposited algae or organic flocs), and rapid hydrodynamic transport. Chl-a was not solely controlled by in-situ phytoplankton growth but also reflected the input and redistribution of algae-associated particles. Vertically, TP, TN, COD, BOD, and TSS concentrations were consistently higher in the deeper layer (2.0 m and 4.0 m), while Chl-a concentrations were higher in the surface layer (0.5 m) for all regions.

4.4. Limitations and Future Work

FP-xLSTM successfully predicted the spatial distributions of multi-depth WQPs at a fine-grained time scale with high accuracy and stability. The proposed framework effectively captured subtle short-term variations in WQPs and reduced temporal and financial costs by leveraging short-term observations from multiple regions, thereby enabling efficient cross-regional transfer learning. Experimental results demonstrated that FP-xLSTM consistently outperformed state-of-the-art baseline models across multiple evaluation metrics and achieved robust prediction performance for all investigated WQPs at different depths. Despite these advantages, several limitations remained. First, uncertainties associated with spectral unmixing, laboratory chemical analysis, and environmental inconsistencies during data acquisition might reduce prediction accuracy. Second, the framework relied on high-quality UAV remote sensing data, which might limit its applicability under adverse weather conditions or in regions with insufficient monitoring frequency and limited data availability. Third, the dependence on spectral features might constrain prediction performance for non-OAC WQPs under complex environmental conditions. Fourth, although the multi-layer mixing and spectral interaction mechanisms enhanced spatio-temporal feature learning, they also increased model complexity and computational cost, potentially restricting large-scale deployment in resource-constrained environments [43]. In addition, the spatial correlation network and feature propagation strategy might require further adaptation when applied to highly heterogeneous environments, such as heavily polluted rivers or coastal waters [44]. Finally, external factors including vegetation cover, topography, and solar radiation conditions might influence the spectral consistency and radiometric quality of UAV imagery.

The model was developed and evaluated using observations collected under the environmental conditions represented in the study period and did not explicitly include winter observations characterized by low solar elevation, reduced solar radiation, low water temperature, potential ice formation, and substantial changes in phytoplankton and suspended-matter conditions. These factors might alter the optical properties of water and reduce the consistency between the spectral features learned by the model and those observed under winter conditions. In regions where field observations were sparse, seasonal conditions differed substantially from the training domain, or UAV data acquisition was constrained by low solar elevation, cloud cover, snow, or ice, model performance might deteriorate, and additional local calibration or transfer learning might be required. Therefore, the practical application under winter and high-latitude data-scarce scenarios remained a constraint requiring further validation with seasonally and geographically expanded datasets.

Several improvements could therefore be considered in future studies. Advanced preprocessing approaches, including improved spectral unmixing, radiometric correction, and standardized laboratory protocols, could reduce measurement uncertainty and environmental bias. Multi-source data fusion strategies integrating UAV imagery, satellite observations, in-situ measurements, and numerical simulation outputs could further im-

prove data availability and model robustness under diverse environmental conditions. In addition, integrating physical mechanisms with data-driven approaches, such as physics-informed neural networks, may enhance prediction reliability for non-OAC WQPs by incorporating hydrodynamic, meteorological, and watershed-related variables. To improve computational efficiency and scalability, lightweight architectures and model compression techniques, including pruning and knowledge distillation, could be explored. Furthermore, transfer learning and domain adaptation strategies might improve model generalization across heterogeneous environments, while radiative transfer correction and bidirectional reflectance normalization could mitigate the influence of illumination and surface conditions on remote sensing observations.

Future research should further investigate the mechanisms through which environmental factors influence WQP distributions and extend data acquisition to estuaries, highly polluted rivers, and coastal environments to enhance model generalization. Long-term monitoring datasets could also be incorporated to analyze WQP dynamics under varying hydrological and climatic conditions. Moreover, integrating physical and biogeochemical constraints, uncertainty quantification, and explainable artificial intelligence techniques would improve model interpretability and predictive reliability. Finally, coupling the proposed framework with intelligent decision-support and early-warning systems could facilitate practical applications in environmental monitoring, water resource management, and policy-making. These efforts would further advance the integration of intelligent informatics and physically informed modeling for next-generation water quality monitoring systems.

5. Conclusions

In this study, an intelligent UAV remote sensing framework, FP-xLSTM, was developed for spatio-temporal prediction of multi-depth WQPs. The proposed framework effectively captured fine-grained temporal dynamics and spatial heterogeneity by strengthening feature propagation across consecutive time periods. Through the integration of multi-layer memory interaction and spectral feature fusion mechanisms, FP-xLSTM enabled efficient information sharing and adaptive feature learning in both temporal and spatial domains, thereby improving prediction robustness and stability for heterogeneous WQPs. In addition, differentiated learning strategies were designed for OACs and non-OAC WQPs, while spatial correlation networks and depth-wise probability distribution modeling further enhanced multi-depth prediction capability and representation of vertical WQP distributions. By jointly incorporating spectral characteristics, spatial dependencies, and interaction features at the pixel level, the proposed framework reduced dependence on large long-term training datasets and improved cross-regional transferability. Experimental results demonstrated that FP-xLSTM consistently outperformed state-of-the-art baseline models across multiple evaluation metrics, highlighting its effectiveness, robustness, and generalization capability in complex water environments. The proposed framework not only improved prediction accuracy and computational efficiency, but also reduced temporal and financial costs associated with large-scale water quality monitoring.

Compared with the relatively low-latitude Yamen Watercourse and Huangyang River, where WQPs showed relatively gradual vertical variations from 0.5 m to 4.0 m, the higher-latitude Fenhe River exhibited greater depth-dependent changes in WQPs, indicating greater vertical heterogeneity under different climatic and environmental conditions. The model's applicability under winter and high-latitude, data-scarce conditions might be constrained by seasonal changes (summer and winter) in water optical properties and limited UAV acquisition conditions, requiring further validation with seasonally and geographically expanded datasets.

Several limitations remained. Uncertainties in spectral unmixing, laboratory analysis, and environmental variability during data acquisition might reduce prediction accuracy. The framework also relied on high-quality UAV imagery, limiting its applicability under adverse weather or sparse monitoring conditions. Furthermore, dependence on spectral features might restrict the prediction of non-optically active constituents (non-OACs), while the multi-layer mixing mechanism increased computational cost despite improving spatio-temporal feature learning. In addition, further validation was needed in highly heterogeneous waters, where environmental factors such as vegetation cover, topography, and solar radiation might affect spectral consistency and model robustness. The experimental conditions evaluated in this study covered three representative urban river systems, namely the Yamen Watercourse, Huangyang River, and Fenhe River, and included repeated UAV hyperspectral observations and water quality measurements across multiple sampling periods and water depths. Six WQPs were measured: TP, TN, COD, BOD, Chl-a, and TSS. The experiments specifically examined spatio-temporal variations in WQPs across sampling locations, monitoring periods, and water depths, with particular attention to the differences in vertical WQP distributions among the three river systems. At a depth of 0.5 m, the overall WQP distributions among the three study areas showed relatively small differences, whereas more pronounced differences emerged with increasing depth. In general, the concentrations of TP, TN, COD, BOD, and TSS increased with depth, while Chl-a decreased, which was associated with reduced light availability for algal growth at greater depths. In addition, cross-regional experiments were conducted to assess the model's ability to transfer learned spectral and spatio-temporal representations across heterogeneous river environments. These experiments therefore focused on the conditions directly represented by the collected UAV hyperspectral and field-measured datasets. The results demonstrated the effectiveness of FP-xLSTM under the tested combinations of spatial heterogeneity, temporal variation, and multi-depth water quality conditions.

This study provided both methodological and practical contributions to intelligent environmental informatics and UAV-assisted monitoring systems. The proposed framework offered effective technical support for environmental assessment, water quality management, and early-warning applications in urban river systems. Furthermore, the framework established an intelligent and data-driven solution for visualization and prediction of multi-depth WQPs across heterogeneous spatio-temporal conditions. More broadly, the integration of spatio-temporal deep learning, multi-source feature interaction, and adaptive transfer learning advanced the development of scalable, interpretable, and intelligent monitoring frameworks for next-generation environmental management and decision-support systems.

Author Contributions: Conceptualization, Y.Z.; methodology, Y.Z.; software, Y.Z. and R.C.; validation, Y.Z.; formal analysis, Y.Z.; investigation, Y.Z. and R.C.; resources, Y.Z. and R.C.; data curation, Y.Z. and R.C.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z.; visualization, Y.Z.; supervision, Y.Z.; project administration, Y.Z. and R.C.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant number: 42301529) and the Fundamental Research Program of Shanxi Province (grant number: 202303021212039).

Data Availability Statement: Data will be made available on request.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (grant number: 42301529) and the Fundamental Research Program of Shanxi Province (grant number: 202303021212039). The authors would like to thank Hongyou Software Co., Ltd. for technical support.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Author Ruijun Chen was employed by the company Hongyou Software Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

Appendix A.1. Information About Classification Standards of Water Quality

Table A1. Classification standards of water quality.

Water Quality Classification	TP (mg/L)	TN (mg/L)	COD (mg/L)	BOD (mg/L)	Chl-a (μg/L)	TSS (mg/L)
I	(0, 0.02]	(0, 0.2]	(0, 10.0]	(0, 2.0]	(0, 1.0]	(0, 10.0]
II	(0.02, 0.1]	(0.2, 0.5]	(10.0, 15.0]	(2.0, 3.0]	(1.0, 4.0]	(10.0, 20.0]
III	(0.1, 0.2]	(0.5, 1.0]	(15.0, 20.0]	(3.0, 4.0]	(4.0, 10.0]	(20.0, 30.0]
IV	(0.2, 0.3]	(1.0, 1.5]	(20.0, 30.0]	(4.0, 6.0]	(10.0, 30.0]	(30.0, 50.0]
V	(0.3, 0.4]	(1.5, 2.0]	(30.0, 40.0]	(6.0, 10.0]	(30.0, 65.0]	(50.0, 70.0]
VI	(0.4, ∞)	(2.0, ∞)	(40.0, ∞)	(10.0, ∞)	(65.0, ∞)	(70.0, ∞)

Appendix A.2. Chemical Analysis Methods on WQPs

Laboratory chemical analysis methods included: TP was analyzed using ammonium molybdate spectrophotometry [45]; TN was measured through alkaline potassium persulfate ultraviolet spectrophotometry method [46]; COD was measured through dichromate oxidation method [47]; BOD was measured through microbial electrode method [48]; Chl-a was measured through high-performance liquid chromatography method [49]; TSS was measured through weighing method [50]. The measured concentration value of each WQP at each sampling site was determined by averaging the concentrations of each WQP. Researchers can request temporary database access through the local water supplies bureau and environmental protection department.

Appendix A.3. The Update Rule for Weight Matrix Involved in ICA Variant

Whitening transformation and singular value decomposition were utilized for $\boldsymbol{x}$ and A:

$$A = U\Sigma V^T \tag{A1}$$

$$\boldsymbol{x}\boldsymbol{x}^T = A\boldsymbol{s}\boldsymbol{s}^T A^T \tag{A2}$$

where U is the left singular matrix, Σ is the rectangular diagonal matrix, and V is the right singular matrix.

$$\boldsymbol{x}\boldsymbol{x}^T = A\boldsymbol{s}\boldsymbol{s}^T A^T = AA^T \tag{A3}$$

$$\boldsymbol{x}\boldsymbol{x}^T = U\Sigma V^T V\Sigma U^T = U\Sigma^2 U^T \tag{A4}$$

where $\boldsymbol{s}$ is an orthogonal matrix. Eigenvalue decomposition was then conducted on $E\{\boldsymbol{x}\boldsymbol{x}^T\}$:

$$\mathbf{x}\mathbf{x}^T = U\Sigma^2 U^T = FDF^T \tag{A5}$$

where F is the orthogonal matrix of eigenvectors of $E\{\boldsymbol{x}\boldsymbol{x}^T\}$, and D represents the diagonal matrix of its eigenvalues λs; $U = F$; $\Sigma = D^{1/2}$. The observed $\boldsymbol{x}$ was linearly transformed, and a new whitening matrix $\widetilde{\boldsymbol{x}}$ could be obtained. $E\left\{\widetilde{\boldsymbol{x}}\widetilde{\boldsymbol{x}}^T\right\} = I$ where I represents the identity matrix. Whitening could be formulated as:

$$\widetilde{A} = FD^{-1/2}F^T A \tag{A6}$$

$$\widetilde{\boldsymbol{x}} = FD^{-1/2}F^T\boldsymbol{x} = FD^{-1/2}F^TA\boldsymbol{s} = \widetilde{A}\boldsymbol{s} \tag{A7}$$

where $\Sigma = D^{1/2}$ is a diagonal matrix and all diagonal entries are $\lambda_1^{-1/2}$, $\lambda_2^{-1/2}, \cdots, \lambda_{270}^{-1/2}$. $\widetilde{A}$ is a new mixing matrix, and its orthogonality could be demonstrated as:

$$E\left\{\widetilde{\boldsymbol{x}}\widetilde{\boldsymbol{x}}^T\right\} = E\left\{\widetilde{A}\boldsymbol{s}\boldsymbol{s}^T\widetilde{A}^T\right\} = \widetilde{A}E\left\{\boldsymbol{s}\boldsymbol{s}^T\right\}\widetilde{A}^T = \widetilde{A}\widetilde{A}^T = I \tag{A8}$$

Moreover, maximization of negentropy based on the information-theoretic quantity of (differential) entropy determined $\boldsymbol{s}$, and the maxima of the approximation of negentropy of $W^T\boldsymbol{x}$ was determined at certain optima of $E\{G(W^T\boldsymbol{x})\}$. Thus, W could be formulated through maximization of the negentropy estimate:

$$\mathcal{J}\left(W^T\boldsymbol{x}\right) \approx \sum_{i=1}^{t}\gamma_i\left[E\left\{G\left(W^T\boldsymbol{x}\right)\right\} - E\{G(\delta)\}\right]^2 \tag{A9}$$

$$\mathcal{J}\left(W^T\boldsymbol{x}\right) \propto \left[E\left\{G\left(W^T\boldsymbol{x}\right)\right\} - E\{G(\delta)\}\right]^2 \tag{A10}$$

where γ_i is a constant; $G(x)$ is a non-quadratic function; and δ is a Gaussian random variable with zero mean and unit variance. The optimum of $E\{G(W^T\boldsymbol{x}\}$ and the maximum of $\mathcal{J}\left(W^T\boldsymbol{x}\right)$ were determined since $E\{G(\delta)\}$ was independent of W. As the Kuhn–Tucker conditions [51] were satisfied, the optimum of $E\{G(W^T\boldsymbol{x})\}$ with $E\left\{(W^T\boldsymbol{x})^2\right\} = ||W||^2 = 1$ was defined as:

$$\frac{\partial\left\{\mathcal{J}\left(W^T\boldsymbol{x}\right) - \beta\left(||W||^2 - 1\right)\right\}}{\partial W} = 2E\left\{\boldsymbol{x}\frac{\partial G\left(W^T\boldsymbol{x}\right)}{\partial W}\right\} - 2\beta W = 0 \tag{A11}$$

where β is a coefficient of W. Then, Newton's method could be applied:

$$\frac{\partial\left\{E\left[\boldsymbol{x}\frac{\partial G\left(W^T\boldsymbol{x}\right)}{\partial W}\right] - \beta W\right\}}{\partial W} = E\left\{\boldsymbol{x}\boldsymbol{x}^T\frac{\partial^2 G\left(W^T\boldsymbol{x}\right)}{\partial W^2}\right\} - \beta I \tag{A12}$$

$$E\left\{\boldsymbol{x}\boldsymbol{x}^T\frac{\partial^2 G\left(W^T\boldsymbol{x}\right)}{\partial W^2}\right\} \approx E\{\boldsymbol{x}\boldsymbol{x}^T\}E\left\{\frac{\partial^2 G\left(W^T\boldsymbol{x}\right)}{\partial W^2}\right\} \approx E\left\{\frac{\partial^2 G\left(W^T\boldsymbol{x}\right)}{\partial W^2}\right\}I \tag{A13}$$

$$W^{(n+1)} = W^{(n)} - \frac{E\left\{\boldsymbol{x}\frac{\partial G\left(W^{(n)T}\boldsymbol{x}\right)}{\partial W^{(n)}}\right\} - \beta W^{(n)}}{E\left\{\frac{\partial^2 G\left(W^{(n)T}\boldsymbol{x}\right)}{\partial\left(W^{(n)}\right)^2}\right\} - \beta I} \tag{A14}$$

$$W^{(n+1)} = E\left\{\boldsymbol{x}\frac{\partial G\left(W^{(n)T}\boldsymbol{x}\right)}{\partial W^{(n)}}\right\} - E\left\{\frac{\partial^2 G\left(W^{(n)T}\boldsymbol{x}\right)}{\partial\left(W^{(n)}\right)^2}\right\}W^{(n)} \tag{A15}$$

$$W^{(n)} = \frac{W^{(n+1)}}{||W^{(n+1)}||} \tag{A16}$$

where $W^{(n)}$ is W at nth iterative step; I denotes the identity matrix. $W^{(n+1)}$ was obtained until convergence (i.e., $<W^{(n)}, W^{(n+1)}>= 1$). $\hat{\boldsymbol{s}}_i = W_i^T\hat{\boldsymbol{x}}_i$ was demonstrated in Equation (3). Therefore, $W^{(n+1)}$ was updated through an iterative process with respect to $W^{(n)}$.

Appendix A.4. Multi-Dimensional Cross Product in Feature Interaction Layer

The details to calculate the multi-dimensional cross product among the resulting matrices could be obtained as follows:

$$\begin{aligned} y_{it} &= f_{\otimes}\left(y_{c,r}, y_{a_1,r}, \cdots, y_{a_j,r}\right) \\ &= \left(y_{c,r} \otimes y_{a_1,r} \otimes y_{a_2,r}, \cdots, \otimes y_{a_j,r}\right) \\ &= \begin{pmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n & & \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n \\ s_{1,1}^{(o)} & \cdots & s_{1,n}^{(o)} & & s_{n,1}^{(o)} & \cdots & s_{n,n}^{(o)} \\ \left| \begin{matrix} s_{1,1}^{(1)} \\ \vdots \\ s_{1,1}^{(j)} \end{matrix} \right. & \cdots & \left. \begin{matrix} s_{1,n}^{(1)} \\ \vdots \\ s_{1,n}^{(j)} \end{matrix} \right| & , \cdots , & \left| \begin{matrix} s_{n,1}^{(1)} \\ \vdots \\ s_{n,1}^{(j)} \end{matrix} \right. & \cdots & \left. \begin{matrix} s_{n,n}^{(1)} \\ \vdots \\ s_{n,n}^{(j)} \end{matrix} \right| \end{pmatrix}^T \\ &= \begin{pmatrix} \det(U_{1,1}) & \cdots & \det(U_{n,1}) \\ \det(U_{1,2}) & \cdots & \det(U_{n,2}) \\ \vdots & \cdots & \vdots \\ \det(U_{1,n}) & \cdots & \det(U_{n,n}) \end{pmatrix} \end{aligned} \tag{A17}$$

where $\det(U_{i,j})$ denotes the determinant of a matrix with jth column eliminated; $\boldsymbol{e}$ is the standard basis vector with a magnitude of 1.

Appendix A.5. Multi-Dimensional Row-Wise Cross Product

$$\begin{aligned} y_{it}^{(2)} &= f_{\otimes}^{(2)}\left(\boldsymbol{w}_{it}\boldsymbol{h}_t^T, \boldsymbol{w}_{it}\boldsymbol{h}_{t-1}^T\right) \\ &= \left(\widetilde{\boldsymbol{h}}_{t,1\to p} \otimes \widetilde{\boldsymbol{h}}_{t-1,1\to p}\right) \\ &= \left(\widetilde{\boldsymbol{h}}_{t,1} \otimes \widetilde{\boldsymbol{h}}_{t-1,1}, \cdots, \widetilde{\boldsymbol{h}}_{t,p} \otimes \widetilde{\boldsymbol{h}}_{t-1,p}\right) \\ &= \begin{pmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n & & \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n \\ \widetilde{\boldsymbol{h}}_{t,1,1} & \cdots & \widetilde{\boldsymbol{h}}_{t,1,n} & & \widetilde{\boldsymbol{h}}_{t,p,1} & \cdots & \widetilde{\boldsymbol{h}}_{t,p,n} \\ \left| \begin{matrix} \widetilde{\boldsymbol{h}}_{t-1,1,1} \\ \vdots \\ \widetilde{\boldsymbol{h}}_{t-1,p,1} \end{matrix} \right. & \cdots & \left. \begin{matrix} \widetilde{\boldsymbol{h}}_{t-1,1,n} \\ \vdots \\ \widetilde{\boldsymbol{h}}_{t-1,p,n} \end{matrix} \right| & , \cdots , & \left| \begin{matrix} \widetilde{\boldsymbol{h}}_{t-1,1,1} \\ \vdots \\ \widetilde{\boldsymbol{h}}_{t-1,p,1} \end{matrix} \right. & \cdots & \left. \begin{matrix} \widetilde{\boldsymbol{h}}_{t-1,1,n} \\ \vdots \\ \widetilde{\boldsymbol{h}}_{t-1,p,n} \end{matrix} \right| \end{pmatrix}^T \\ &= \begin{pmatrix} \det(U_{j,1}), & \cdots & , \det(U_{j,p}) \end{pmatrix}_{p\times 1}^T \end{aligned} \tag{A18}$$

Appendix A.6. Calculation Algorithm of Impact Matrix in Shift Layer

Algorithm A1: Impact matrix calculation algorithm

Input: Interaction vectors of each obtained $WQP_{0.5}$ as M_r, cost vectors of $WQPs_{1.0}$ as C, the impact distribution of $WQPs_{0.5}$ upon $WQPs_{1.0}$ as M_c, the number of iterations as n, entropy regularization coefficient as ϵ, the number of vectors in M_r as k, and convergence threshold as δ.
Output: Impact of each obtained $WQP_{0.5}$ upon $WQPs_{1.0}$ as matrices P
for i in k **do**
Initialize scaling factors $\boldsymbol{u}$ and $\boldsymbol{v}$ for M_r and M_c.
Construct kernel matrix as $K = \exp\left(\frac{C_{i,j}}{\epsilon}\right)$.
for j in n **do**
$\boldsymbol{u}_j \leftarrow M_r / K\boldsymbol{v}_j$;
$\boldsymbol{v}_j \leftarrow M_c / K^T\boldsymbol{u}_j$;
$\boldsymbol{u}_{prev} \leftarrow \boldsymbol{u}_j$;
if $\left|\mathbf{u}_j - \mathbf{u}_{prev}\right| < \delta$ **then**
break.
end
end
$f_*(\boldsymbol{u}, K, \boldsymbol{v}) \leftarrow diag(\boldsymbol{u}_j) K diag(\boldsymbol{v}_j)$.
$P_i \leftarrow f_*(\boldsymbol{u}, K, \boldsymbol{v})$.
end
return P.

Appendix A.7. Calculation Process of PMF Variant Module

A PMF variant utilized to improve calculation accuracy and stability was formulated as:

$$Q - \hat{Q} \sim \mathcal{N}\left(0, \nu^2\right) \tag{A19}$$

$$Q \sim \mathcal{N}\left(\hat{Q}, \nu^2\right) \tag{A20}$$

$$p\left(C \middle| H, G, \nu^2\right) = \prod_{i=1}^{L}\prod_{j=1}^{M}\left[\mathcal{N}\left(Q_{ij} \middle| H_i^T G_j, \nu^2\right)\right]^{I_{ij}} \tag{A21}$$

where Q, $\hat{Q}$, $\mathcal{N}$, I, H, and G indicate the ground truth of WQPs concentration, calculated WQPs concentration, Gaussian distribution function, indicator function that was equal to 1 if the calculated concentration was not 0, hidden matrix for scanned points, and hidden matrix for WQPs concentrations. $H \in \mathbb{R}^{M\times K}$ and $G \in \mathbb{R}^{M\times L}$ where $p\left(H|\nu_H^2\right) = \prod_{i=1}^{K}\mathcal{N}\left(H_i|0, \nu_H^2\boldsymbol{I}\right)$ and $p\left(G|\nu_G^2\right) = \prod_{j=1}^{L}\mathcal{N}\left(G_j|0, \nu_G^2\boldsymbol{I}\right)$ where ν denotes the standard deviation. A resulting WQP concentration outlier was detected if $\hat{Q}$ was outside the range of 3 standard deviations of the mean of WQP concentrations in the scanned region. The posterior distribution concerned with scanned points and WQP concentration features was formulated as:

$$\begin{aligned} p\left(H, G \middle| Q, \nu^2, \nu_H^2, \nu_G^2\right) &= \frac{p\left(Q \middle| H, G, \nu^2\right) p\left(H, G \middle| \nu_H^2, \nu_G^2\right)}{p\left(Q, \nu^2, \nu_H^2, \nu_G^2\right)} \\ &\propto p\left(Q \middle| H, G, \nu^2\right) p\left(H \middle| \nu_H^2\right) p\left(G \middle| \nu_G^2\right) \\ &= \prod_{i=1}^{K}\prod_{j=1}^{L}\left[N\left(Q_{ij} \middle| H_i^T G_j, \nu^2\right)\right]^{I_{ij}} \prod_{i=1}^{K} N\left(H_i|0, \nu_H^2\mathbf{I}\right) \prod_{j=1}^{L} N\left(G_j|0, \nu_G^2\mathbf{I}\right) \end{aligned} \tag{A22}$$

The log of the posterior distribution for the scanned points and WQPs concentration features was obtained:

$$\begin{aligned} &\ln p\left(H, G \middle| Q, \nu^2, \nu_H^2, \nu_G^2\right) \\ &= \frac{1}{-2\nu^2} \sum_{i=1}^{K} \sum_{j=1}^{L} I_{ij} \left(Q_{ij} - H_i^T G_j\right)^2 - \frac{1}{2\nu_H^2} \sum_{i=1}^{K} H_i^T H_i - \frac{1}{2\nu_G^2} \sum_{j=1}^{L} G_j^T G_j \\ &\quad -\frac{1}{2}\left(\left(\sum_{i=1}^{K} \sum_{j=1}^{L} I_{ij}\right) \ln \nu^2 + KM \ln \nu_H^2 + LM \ln \nu_G^2\right) + const \end{aligned} \tag{A23}$$

Furthermore, maximizing log-posterior over H and G was equivalent to minimizing the sum-of-squared-errors objective function:

$$obj = \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{L} I_{ij} \left(Q_{ij} - H_i^T G_j\right)^2 + \frac{\nu^2}{2\nu_H^2} \sum_{i=1}^{K} ||H_i||_{Fro}^2 + \frac{\nu^2}{2\nu_G^2} \sum_{j=1}^{L} ||G_j||_{Fro}^2 \tag{A24}$$

where *Fro* indicates the Frobenius norm. Minimization of *obj* was conducted through gradient descent:

$$\begin{aligned} H_i &= H_i - \gamma \frac{\partial obj}{\partial H_i} \\ &= H_i + \gamma \left(Q_{ij} - H_i^T G_j\right) G_j + \frac{\nu^2}{\nu_H^2} H_i \end{aligned} \tag{A25}$$

$$\begin{aligned} G_i &= G_j - \gamma \frac{\partial obj}{\partial G_j} \\ &= G_j + \gamma \left(Q_{ij} - H_i^T G_j\right) H_i + \frac{\nu^2}{\nu_G^2} G_j \end{aligned} \tag{A26}$$

Thereby, the detected WQP concentration outlier could be obtained as $\hat{Q}$:

$$\hat{Q}_{ij} = H_i^T G_j \tag{A27}$$

Appendix A.8. Prediction Algorithm of Depth-Wise Variation of WQPs over Time

Algorithm A2: Prediction of multi-depth variation of WQPs

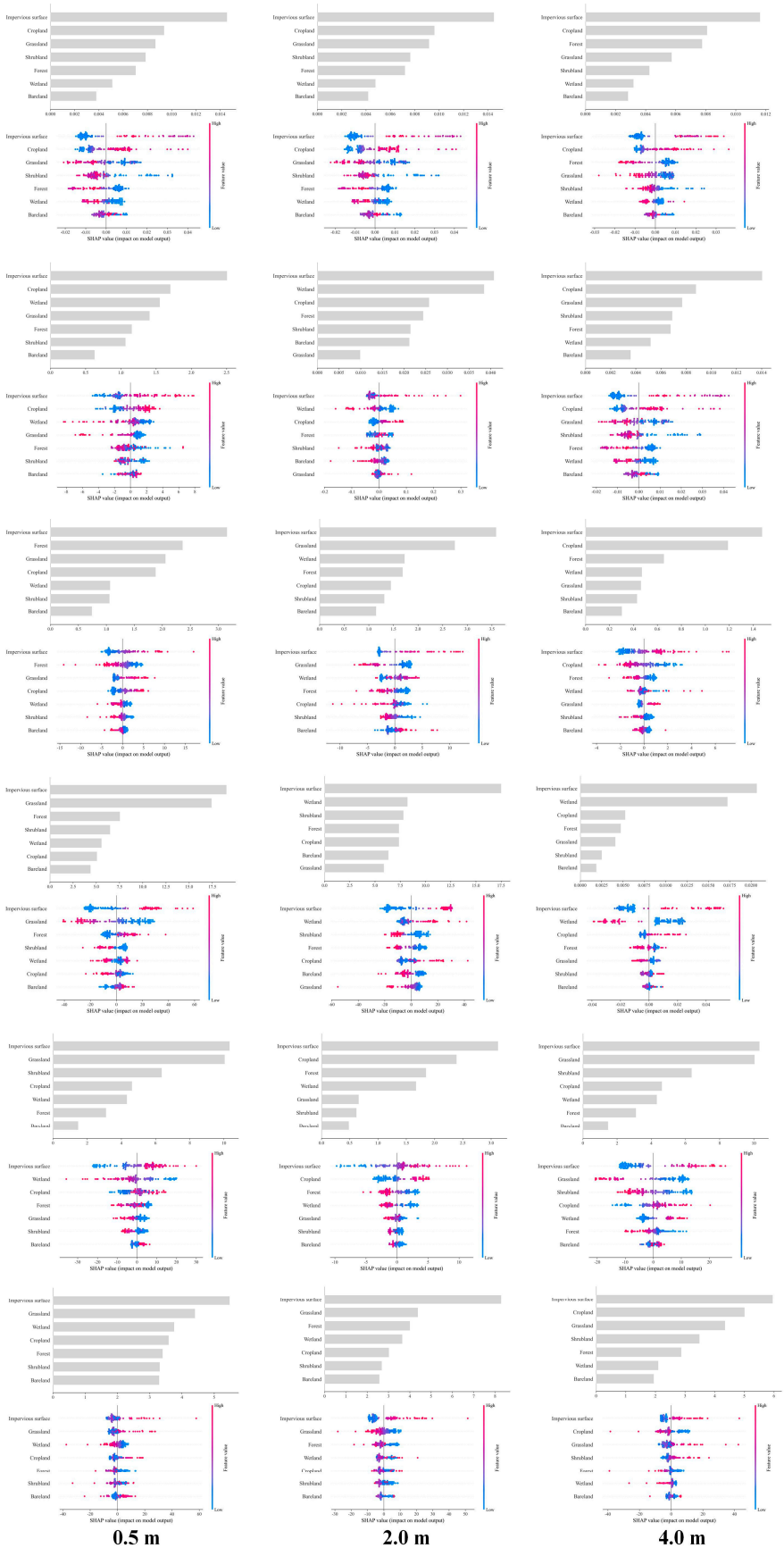
Input: Hyperspectral reflectance data R; multi-depth WQPs concentrations Φ; the number of iterations as n; the batch size as k.
Output: Predicted multi-depth WQPs concentrations.
Employ an ICA variant to extract real source reflectance R_{rs} from R.
Obtain Eigenvalue matrices E through Eigenvalue decomposition of R_{rs}.
Obtain interaction matrices T through the interaction operation $f_{\otimes}$ from R_{rs}.
Initialize all parameters Θ.
for i in n **do**
  for j in k **do**
    Train FP-xLSTM$_s$ with $R_{rs,j}$, E_j, and label Φ_j for OAC WQPs$_{0.5}$, and with T_j, E_j, and label Φ_j for non-OAC WQPs$_{0.5}$ using Equations (9)–(35);
    Combine and retain $f_{\otimes}^{(2)}(\cdot)$ and $f_{*}(\cdot)$ at each time stamp;
    Input $f_{\otimes}^{(2)}(\cdot)$ and $f_{*}(\cdot)$ to FP-xLSTM$_d$ to predict WQPs$_{1.0}$ and WQPs$_{2.0}$.
  end
end
Predict WQPs$_{0.5}$, WQPs$_{1.0}$, and WQPs$_{2.0}$ by their corresponding FP-xLSTM.
Employ the PMF variant to recalculate concentration outliers using Equations (A22)–(A27).
Return predicted multi-depth WQPs concentrations.

Appendix A.9. Impact of Land Cover Features upon WQPs Retrieval at Multiple Depths



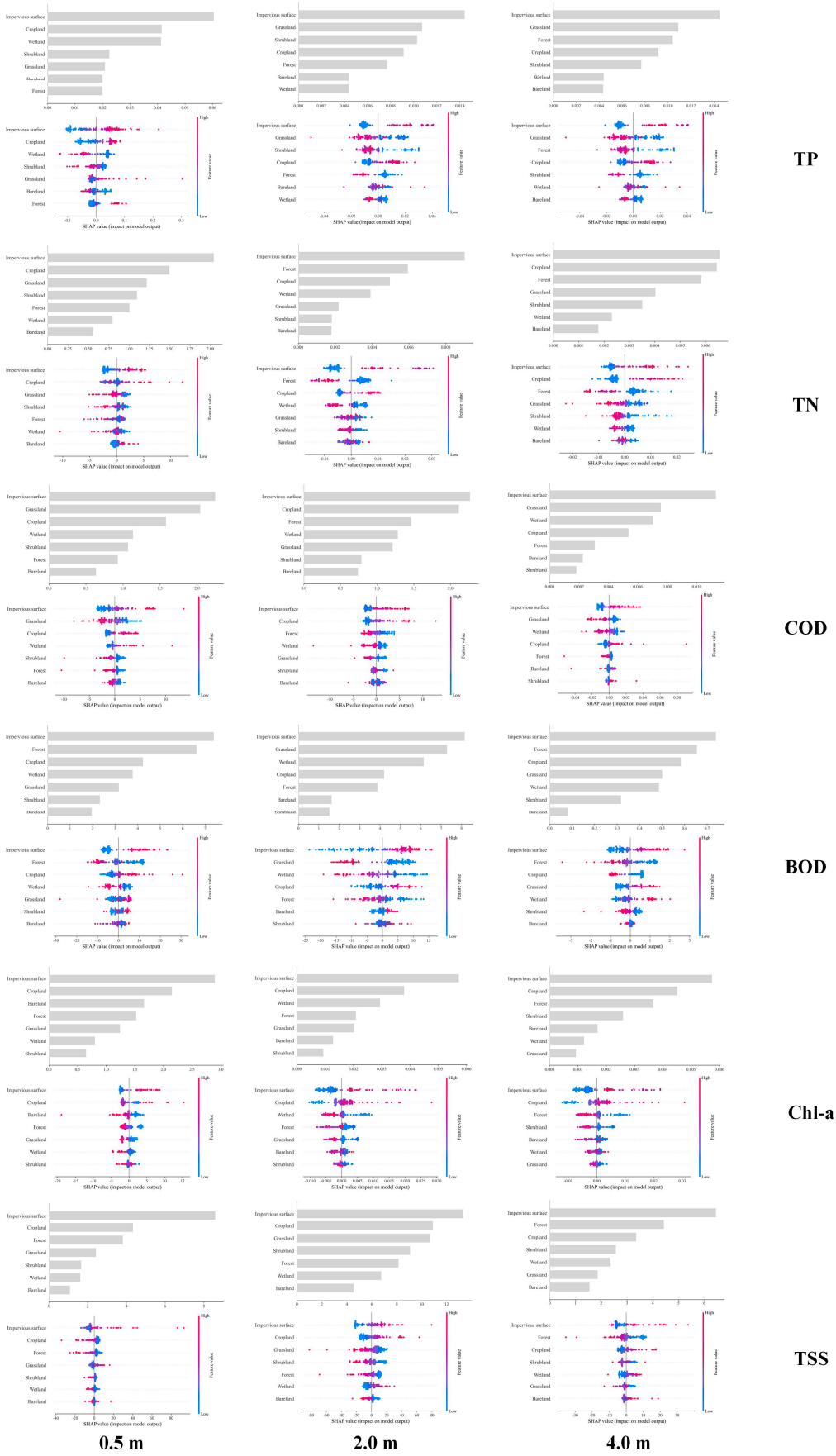
(**a**) Yamen Watercourse

Figure A1. *Cont.*



(**b**) Huangyang River

Figure A1. *Cont.*



(**c**) Fenhe River

Figure A1. Impact of land cover features upon WQPs retrieval at depths of 0.5, 2.0, and 4.0 m.

Appendix A.10. Formulation of Statistical Metrics

Welch's *t*-test was formulated as below:

$$t_{stat}(y,\hat{y}) = \frac{(\hat{y}_{mean} - y_{mean})}{\left(\frac{\sigma_{\hat{y}}^2}{n_{\hat{y}}} + \frac{\sigma_y^2}{n_y}\right)^{\frac{1}{2}}} \tag{A28}$$

$$df_{t-test}(y,\hat{y}) = \frac{\left(\frac{\sigma_{\hat{y}}^2}{n_{\hat{y}}} + \frac{\sigma_y^2}{n_y}\right)^2}{\frac{\sigma_{\hat{y}}^4}{n_{\hat{y}}^2(n_{\hat{y}}-1)} + \frac{\sigma_y^4}{n_{\hat{y}}^2(n_y-1)}} \tag{A29}$$

where y and $\hat{y}$ indicate measured values and calculated values; $df_{t-test}(y,\hat{y})$ indicate the degrees of freedom; σ was the standard deviation (SD); n indicates the number of samples. MK test was a non-parametric statistical method to detect upward or downward trends in data, where the null hypothesis was that the data did not show a monotonically changing trend. Water samples were collected at different sampling sites and in a certain order such that the MK test was applied to test whether the dataset of each WQP in each period followed a certain changing trend. The corresponding statistical information was obtained as below.

$$\mathrm{sgn}(x_j - x_i) = \begin{cases} +1, & x_j - x_i > 0 \\ 0, & x_j - x_i = 0, \\ -1, & x_j - x_i < 0 \end{cases} \quad i < j \tag{A30}$$

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}(x_j - x_i) \tag{A31}$$

$$\mathrm{Var}(S) = \frac{n(n-1)(2n+5)}{18} \tag{A32}$$

$$Z = \begin{cases} \frac{S-1}{\sqrt{\mathrm{Var}(S)}}, & S > 0 \\ 0, & S = 0 \\ \frac{S+1}{\sqrt{\mathrm{Var}(S)}}, & S < 0 \end{cases} \tag{A33}$$

where sgn denotes the sign function; S is the MK statistical value; $Var(S)$ is the variance; and Z is the normalized Z score. χ^2 was utilized to compare the classification result distributions of measured and predicted values, which provided the overall consistency between measured and predicted WQPs distributions and discovered problems with the model in terms of category bias or imbalance. The null hypothesis assumed that the distribution of predicted classification results did not significantly deviate from that of the measured values.

$$\chi^2 = \sum_{i}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{A34}$$

where O_i denotes the number of predicted values classified as class i; E_i denotes the number of measured values classified as class i; and k denotes the total number of classes.

Appendix A.11. Impact Analysis of Land Cover Features upon WQPs Retrieval

High SHapley Additive exPlanations (SHAP) values of impervious surfaces and croplands indicated their dominant contributions to TP variability. Impervious surfaces enhanced runoff-driven phosphorus transport from urban and industrial areas, whereas croplands supplied substantial phosphorus through fertilizer application, soil erosion, and agricultural drainage, resulting in stronger Liang–Kleeman information flow rates

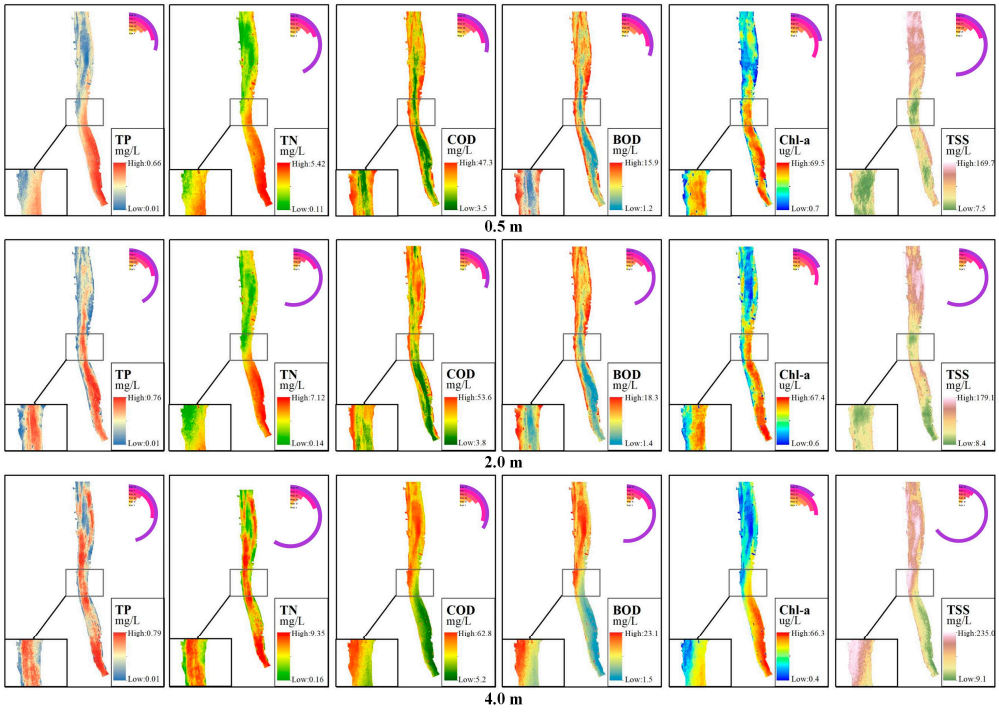
and higher model importance. The declining SHAP variation with increasing water depth suggests that land-cover effects weakened vertically due to hydrodynamic mixing, sediment interactions, and internal biogeochemical processes. Although wetlands exhibited high feature values, their SHAP values remained low because they primarily functioned as ecological buffers that retained and transformed phosphorus, thereby reducing direct impacts on TP fluctuations. Bareland and shrubland showed the lowest SHAP values owing to limited phosphorus loading and weak hydrological connectivity [52]. Similarly, grassland exhibited low SHAP values for TN despite relatively high feature values, as dense vegetation reduced runoff, enhanced nitrogen retention and infiltration, and promoted microbial assimilation, leading to weaker Liang–Kleeman information flow and limited contribution to model prediction.

High SHAP values of impervious surface and cropland in COD and BOD prediction indicate their strong influences on organic pollutant transport and accumulation. Impervious surfaces enhance runoff-driven delivery of urban and industrial organic contaminants, while croplands contribute organic matter through agricultural drainage, fertilizer residues, and soil erosion. These processes strengthen the Liang–Kleeman information flow rates with COD and BOD concentrations, resulting in higher feature importance. In contrast, wetlands and grasslands exhibit relatively low SHAP values because they mainly function as ecological buffers that promote pollutant retention, filtration, and biodegradation, thereby weakening their direct contributions to COD variability.

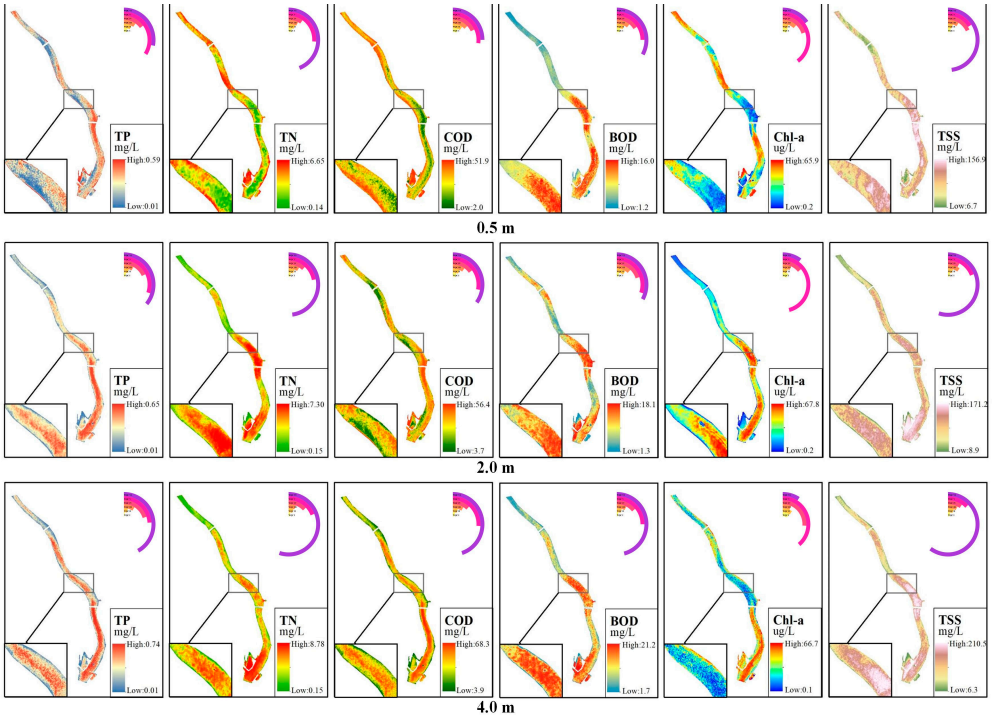
High SHAP values of impervious surfaces and croplands in Chl-a prediction indicated their dominant contributions to nutrient enrichment and algal growth. Impervious surfaces accelerated runoff-driven transport of nutrients and pollutants, whereas croplands supplied substantial nutrient inputs through fertilizer application and agricultural drainage, resulting in stronger Liang–Kleeman information flow rates with Chl-a concentrations. In contrast, bareland, forest, and shrubland exhibited lower SHAP values because of their limited nutrient export and weaker hydrological connectivity. Wetland and grassland showed similar positive effects on Chl-a at 0.5 m, reflecting stronger influences of ecological nutrient retention and surface runoff. However, these relationships became inconsistent at greater depths owing to hydrodynamic mixing, light attenuation, sediment interactions, and depth-dependent biogeochemical processes that altered causal information transfer. Likewise, positive Liang–Kleeman information flow rates between impervious surfaces, croplands, shrublands, barelands, and TSS indicated that these land-cover types enhanced sediment transport and suspended particle inputs. Impervious surfaces intensified runoff-driven erosion, whereas croplands, barelands, and shrublands increased sediment mobilization through soil disturbance and precipitation-induced erosion, thereby strengthening their contributions to TSS variability [52].

Above all, SHAP analysis revealed that impervious surface and cropland exerted the strongest influences on WQPs due to enhanced runoff generation, nutrient export, organic pollutant transport, and sediment delivery, resulting in higher Liang–Kleeman information flow rates. In contrast, wetlands and grasslands generally exhibited lower SHAP values because of their ecological buffering capacities, including pollutant retention, filtration, sediment stabilization, and nutrient assimilation. Bareland, forest, and shrubland showed relatively weak impacts owing to limited anthropogenic disturbance and weaker hydrological connectivity. Additionally, depth-dependent variations indicated that hydrodynamic mixing, sediment interactions, and biogeochemical processes progressively weakened surface-derived land-cover influences on subsurface WQP dynamics.

Appendix A.12. Spatial Distributions of WQPs at Multiple Depths

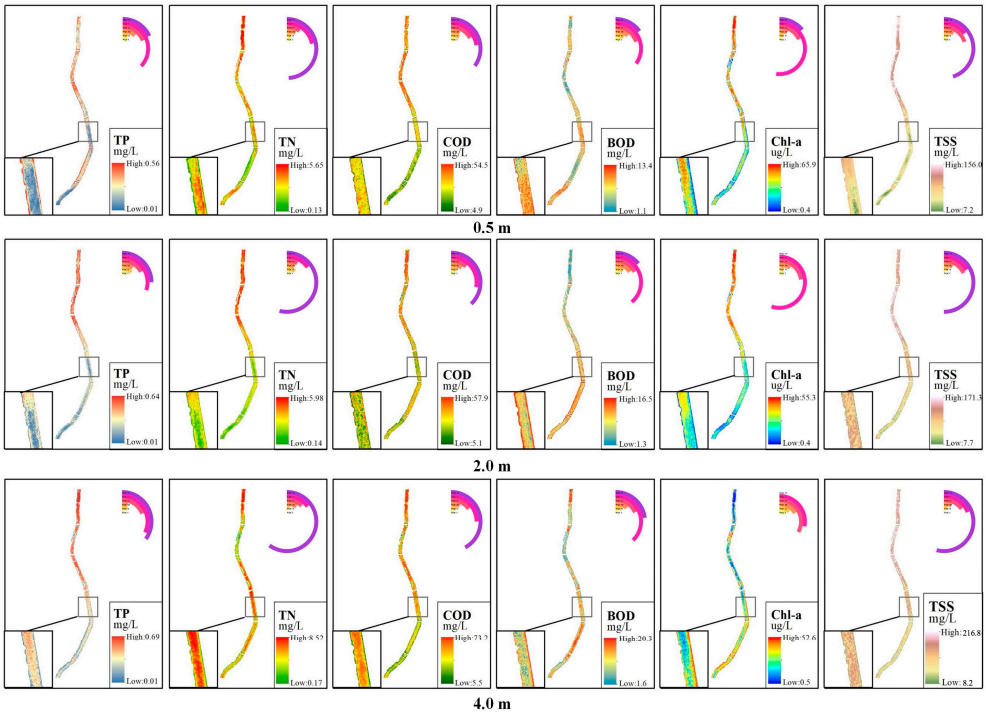


(**a**) Yamen Watercourse



(**b**) Huangyang River

Figure A2. *Cont.*



(**c**) Fenhe River

Figure A2. Concentration distributions of multi-depth WQPs in various regions through FP-xLSTM. Ring-shaped bars represent the proportions of each WQP of various classes and start from class I to VI from the inside out.

References

1. Zhang, H.; Zhang, X.; Xu, J.; Tao, S.; Li, C.; Yang, Y. Coupling stable isotopes and hydrochemical parameters to elucidate the linkages between hydrological connectivity and water quality in an urban river–lake system. *Environ. Pollut.* **2026**, *400*, 128308. [CrossRef] [PubMed]
2. Barasa, C.N.; Muse, G.S.; Khazenzi, J.A.; Gondwe, L.S.M. Influence of Nutrient Enrichment and Temperature on Aquatic Invertebrate Communities in a Tropical Riverine System in Kenya. *J. Watershed Ecol.* **2026**, *1*, 10010. [CrossRef]
3. Kim, H.G.; Hong, S.; Chon, T.-S.; Joo, G.-J. Spatial patterning of chlorophyll a and water-quality measurements for determining environmental thresholds for local eutrophication in the Nakdong River basin. *Environ. Pollut.* **2021**, *268*, 115701. [CrossRef] [PubMed]
4. Nguyen, T.H.; Fujiwara, T.; Yamashita, H.; Togawa, H.; Miyake, H.; Goto, M.; Nagare, H.; Nakamura, M.; Oritate, F.; Ihara, H. From sewage sludge to agriculture: Governmental initiatives, technologies, and sustainable practices in Japan. *J. Mater. Cycles Waste Manag.* **2025**, *27*, 2832–2849. [CrossRef]
5. Fascista, A. Toward integrated large-scale environmental monitoring using WSN/UAV/Crowdsensing: A review of applications, signal processing, and future perspectives. *Sensors* **2022**, *22*, 1824. [CrossRef] [PubMed]
6. Zhao, C.; Li, M.; Wang, X.; Liu, B.; Pan, X.; Fang, H. Improving the accuracy of nonpoint-source pollution estimates in inland waters with coupled satellite-UAV data. *Water Res.* **2022**, *225*, 119208. [CrossRef] [PubMed]
7. Wang, J.; Sun, D.; Wang, S.; Li, Z.; Zhang, Y.; Li, J.; Zhang, H. Satellite observations of suspended particulate matter concentration in Lake Gaoyou in the past four decades. *Water Res.* **2024**, *254*, 121442. [CrossRef] [PubMed]
8. Sivakumar, R.; Prasanth, B.S.V.; Ramaraj, M. An empirical approach for deriving specific inland water quality parameters from high spatio-spectral resolution image. *Wetl. Ecol. Manag.* **2022**, *30*, 405–422. [CrossRef]
9. Guo, H.; Tian, S.; Huang, J.J.; Zhu, X.; Wang, B.; Zhang, Z. Performance of deep learning in mapping water quality of Lake Simcoe with long-term Landsat archive. *ISPRS J. Photogramm. Remote Sens.* **2022**, *183*, 451–469. [CrossRef]
10. Zhuang, Y.; Ji, Y.; Kuang, Q.; Zhang, Z.; Li, P.; Song, J.; He, N. Oxidation treatment of shale gas produced water: Molecular changes in dissolved organic matter composition and toxicity evaluation. *J. Hazard. Mater.* **2023**, *452*, 131266. [CrossRef] [PubMed]
11. Xiong, J.; Lin, C.; Cao, Z.; Hu, M.; Xue, K.; Chen, X.; Ma, R. Development of remote sensing algorithm for total phosphorus concentration in eutrophic lakes: Conventional or machine learning? *Water Res.* **2022**, *215*, 118213. [CrossRef] [PubMed]

12. Cruz-Retana, A.; Becerril-Piña, R.; Fonseca, C.R.; Gómez-Albores, M.A.; Gaytán-Aguilar, S.; Hernández-Téllez, M.; Mastachi-Loza, C.A. Assessment of regression models for surface water quality modeling via remote sensing of a water body in the Mexican Highlands. *Water* **2023**, *15*, 3828. [CrossRef]
13. Li, Y.; Fu, Y.; Lang, Z.; Cai, F. A high-frequency and real-time ground remote sensing system for obtaining water quality based on a micro hyper-spectrometer. *Sensors* **2024**, *24*, 1833. [CrossRef] [PubMed]
14. Chen, L.; Liu, L.; Liu, S.; Shi, Z.; Shi, C. The application of remote sensing technology in inland water quality monitoring and water environment science: Recent progress and perspectives. *Remote Sens.* **2025**, *17*, 667. [CrossRef]
15. Rahul, T.; Brema, J.; Wessley, G.J.J. Evaluation of surface water quality of Ukkadam lake in Coimbatore using UAV and Sentinel-2 multispectral data. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 3205–3220. [CrossRef]
16. Moon, J.; Jung, S.; Suh, S.; Pyo, J. Development of deep learning quantization framework for remote sensing edge device to estimate inland water quality in South Korea. *Water Res.* **2025**, *283*, 123760. [CrossRef] [PubMed]
17. Varadharajan, C.; Appling, A.P.; Arora, B.; Christianson, D.S.; Hendrix, V.C.; Kumar, V.; Lima, A.R.; Müller, J.; Oliver, S.; Ombadi, M. Can machine learning accelerate process understanding and decision-relevant predictions of river water quality? *Hydrol. Process.* **2022**, *36*, e14565. [CrossRef]
18. Cheng, C.; Xie, Z.; Jin, X.; Peng, C.; Ren, C.; Mao, Y.; Zhou, K.; Li, Y. Urban Fine-Grained Water Quality Monitoring Based on Stacked Machine Learning Approach. *IEEE Access* **2024**, *12*, 77156–77170. [CrossRef]
19. Cai, X.; Wu, L.; Li, Y.; Lei, S.; Xu, J.; Lyu, H.; Li, J.; Wang, H.; Dong, X.; Zhu, Y. Remote sensing identification of urban water pollution source types using hyperspectral data. *J. Hazard. Mater.* **2023**, *459*, 132080. [CrossRef] [PubMed]
20. Li, H.; Wang, N.; Du, Z.; Huang, D.; Shi, M.; Zhong, Z.; Yuan, D. Multi-Parameter water quality inversion in heterogeneous inland waters using UAV-Based hyperspectral data and deep learning methods. *Remote Sens.* **2025**, *17*, 2191. [CrossRef]
21. Wang, S.; Xing, Y.; Zhu, J.; Li, Y.; Dong, F. Unveiling the river spatial topology in robust water quality prediction: A LSTM-based evaluation framework. *Ecol. Inform.* **2026**, *94*, 103659. [CrossRef]
22. Hashemi-Nasab, F.S.; Parastar, H. Vis-NIR hyperspectral imaging coupled with independent component analysis for saffron authentication. *Food Chem.* **2022**, *393*, 133450. [CrossRef] [PubMed]
23. Nordhausen, K.; Oja, H. Independent component analysis: A statistical perspective. *Wiley Interdiscip. Rev. Comput. Stat.* **2018**, *10*, e1440. [CrossRef]
24. Omuya, E.O.; Okeyo, G.O.; Kimwele, M.W. Feature selection for classification using principal component analysis and information gain. *Expert Syst. Appl.* **2021**, *174*, 114765. [CrossRef]
25. Modi, P.; Chintalacheruvu, M.R. Investigating river water quality assessment through non-parametric analysis: A case study of the Godavari River in India. *Environ. Qual. Manag.* **2024**, *33*, 239–264. [CrossRef]
26. Graham, K.E.; Anderson, C.E.; Boehm, A.B. Viral pathogens in urban stormwater runoff: Occurrence and removal via vegetated biochar-amended biofilters. *Water Res.* **2021**, *207*, 117829. [CrossRef] [PubMed]
27. Mohammed, S.; Arshad, S.; Bashir, B.; Vad, A.; Alsalman, A.; Harsányi, E. Machine learning driven forecasts of agricultural water quality from rainfall ionic characteristics in Central Europe. *Agric. Water Manag.* **2024**, *293*, 108690. [CrossRef]
28. V V R, M.R.; Silpa, N.; Reddy, S.S.; Kurada, R.R.; Hussain, S.M.; Sameera, E. A Robust XG-Boost machine learning model for water quality estimation system by leveraging with chi-square forward sequential feature selection technique. In Proceedings of the 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE), Ballari, India, 2–3 November 2023; pp. 1–7.
29. Salem, A.K.; Taha, A.F.; Abokifa, A.A. Graph neural networks-based dynamic water quality state estimation in water distribution networks. *Eng. Appl. Artif. Intell.* **2024**, *138*, 109426. [CrossRef]
30. Huan, J.; Liao, W.; Zheng, Y.; Xu, X.; Zhang, H.; Shi, B. A deep learning model with spatio-temporal graph convolutional networks for river water quality prediction. *Water Supply* **2023**, *23*, 2940–2957. [CrossRef]
31. Xiong, B.; Wen, S.; Lu, P.; Lin, K. A Spectral Domain Graph Transformer Model with Position-Encoded Information. *IAENG Int. J. Comput. Sci.* **2025**, *52*, 1683.
32. Xu, B.; Yan, J.; Tang, H.; Zhang, Y.; Wang, M.; Gao, J. An efficient compressive CNN and transformer hybrid framework for long-term dissolved oxygen prediction in aquaculture. *Inf. Process. Agric.* **2026**, *in press*. [CrossRef]
33. Li, D.; Ji, X.; Liu, L. An accurate forecasting model for key water quality factors based on Transformer with multi-scale attention mechanism. *Environ. Model. Softw.* **2025**, *191*, 106491. [CrossRef]
34. Gao, Z.; Chen, J.; Wang, G.; Ren, S.; Fang, L.; Yinglan, A.; Wang, Q. A novel multivariate time series prediction of crucial water quality parameters with Long Short-Term Memory (LSTM) networks. *J. Contam. Hydrol.* **2023**, *259*, 104262. [CrossRef] [PubMed]
35. Chavan, D.P.; Bhise, A.A.; Londhe, A.D. Exploring Hybrid Deep Learning Models for Water Quality Forecasting: A Comprehensive Review of Methods and Challenges. In Proceedings of the Congress on Intelligent Systems, Delhi, India, 6–7 September 2025; pp. 200–208.
36. Meng, H.; Zhang, J.; Chang, Y.; Zheng, Z. A new method for predicting chlorophyll-a concentration in a reservoir: Coupling EFDC hydrodynamic and water quality model with ConvLSTM-MLP network. *J. Hydrol.* **2025**, *660*, 133485. [CrossRef]

37. Mi, Z.; Li, Q.; Sha, Y.; Wu, Z. CNN-LSTM-attention water quality prediction hybrid model. In Proceedings of the Second International Conference on Sustainable Technology and Management (ICSTM 2023), Dongguan, China, 21–23 July 2023; pp. 19–24.
38. Zhu, L.; Lu, W. A quantum-inspired attention integrated scalar long short-term memory model for accurate and stable groundwater contaminant source inversion. *Environ. Monit. Assess.* **2026**, *198*, 124. [CrossRef] [PubMed]
39. Deng, M.; Ma, R.; Loiselle, S.A.; Hu, M.; Xue, K.; Cao, Z.; Wang, L.; Lin, C.; Gao, G. Monitoring salinity in inner Mongolian lakes based on Sentinel-2 images and machine learning. *Remote Sens.* **2024**, *16*, 3881. [CrossRef]
40. Garg, V.K.; Tanta, A.; Lal Srivastav, A.; Tiwari, M.; Sharma, A.; Kanwar, V.S. Water quality assessment using synchrotron-based TXRF. *Water Environ. Res.* **2022**, *94*, e10759. [CrossRef] [PubMed]
41. Kumar, A.; Singh, A. Pollution source characterization and evaluation of groundwater quality utilizing an integrated approach of Water Quality Index, GIS and multivariate statistical analysis. *Water Supply* **2024**, *24*, 3517–3539. [CrossRef]
42. Li, M.; Guo, Q.; Chen, L.; Fallgren, P.H. Effects of freshwater salinity on suspended sediment settling during different high-turbidity events. *Environ. Pollut.* **2025**, *386*, 127195. [CrossRef] [PubMed]
43. Tian, D.; Zhao, X.; Gao, L.; Liang, Z.; Yang, Z.; Zhang, P.; Wu, Q.; Ren, K.; Li, R.; Yang, C. Estimation of water quality variables based on machine learning model and cluster analysis-based empirical model using multi-source remote sensing data in inland reservoirs, South China. *Environ. Pollut.* **2024**, *342*, 123104. [CrossRef] [PubMed]
44. Zhang, S.; Wang, F.; Wang, R.; Cai, M. Spatial assessment of triazole organic compounds in surface water from the coastal estuaries to the East China sea. *Environ. Pollut.* **2023**, *320*, 121024. [CrossRef] [PubMed]
45. Zhang, M.; Zhu, J.; Fang, K.; Xu, T.; Chen, H.; Sun, X. Ozone-based photoactivated synergistic digestion method for the determination of total dissolved phosphorus in water. *J. Anal. Chem.* **2024**, *79*, 585–594. [CrossRef]
46. Zeng, S.; Qi, P.; Ai, S.; Sun, X.; Kang, H.; Bian, D. Precise determination of the total nitrogen content in activated sludge by ultrasonic pre-treatment assisted wet method. *Environ. Technol.* **2024**, *45*, 5233–5243. [CrossRef] [PubMed]
47. von Mühlen, L.; Prestes, O.D.; Ferrão, M.F.; Sirtori, C. Miniaturized method for chemical oxygen demand determination using the PhotoMetrix PRO application. *Molecules* **2022**, *27*, 4721. [CrossRef] [PubMed]
48. Zhou, Y.; Zheng, S.; Qin, W. Electrochemical biochemical oxygen demand biosensors and their applications in aquatic environmental monitoring. *Sens. Bio-Sens. Res.* **2024**, *44*, 100642. [CrossRef]
49. Merder, J.; Zhao, G.; Pahlevan, N.; Rigby, R.A.; Stasinopoulos, D.M.; Michalak, A.M. A novel algorithm for ocean chlorophyll-a concentration using MODIS Aqua data. *ISPRS J. Photogramm. Remote Sens.* **2024**, *210*, 198–211. [CrossRef]
50. Adjovu, G.E.; Stephen, H.; James, D.; Ahmad, S. Measurement of total dissolved solids and total suspended solids in water systems: A review of the issues, conventional, and remote sensing techniques. *Remote Sens.* **2023**, *15*, 3534. [CrossRef]
51. Zhou, R.; Han, J.; Li, T.; Guo, Z. Fast independent component analysis denoising for magnetotelluric data based on a correlation coefficient and fast iterative shrinkage threshold algorithm. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [CrossRef]
52. Wang, Y.; Zhao, Y.; Li, X.; Yan, L.; Deng, W.; Liu, Y.; He, F.; Zhai, J.; Xu, Z.; Chen, M. Integrated indices for sustainable water management in the Guanzhong Basin, China. *J. Hydrol. Reg. Stud.* **2025**, *59*, 102446. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.