Seasonal Source Apportionment of Dhaka river water and sediment heavy metal using novel graph based ensemble architecture of deep learning

Md Rakib Hasan[1], Arafat Rahman[1], Shoumik Zubyer[1], Yeasmin Nahar Jolly[1]

[a]*Department of Soil, Water And Environment, University of Dhaka, Dhaka-1000, Bangladesh*
[b]*Atmospheric And Environment Lab, Atomic Energy Center, Dhaka-1000, Bangladesh*

Abstract

This is where the abstract goes. Write a concise summary of the research and its results here.

Keywords: Heavy Metal, Source Apportionment, Graph Neural Network, Deep learning

1. Introduction

2. Methodology

2.1. Study Area Sample Collection

The study was conducted on five major rivers around Dhaka, Bangladesh—the Buriganga, Shitalakshya, Turag, Dhaleshwari, and Balu—which are heavily influenced by industrial and domestic activities. Sampling was performed during two distinct seasons: the winter (November–February) and the rainy season (June–September), to capture seasonal variations in pollution. Sediment and water samples were collected from pre-determined stations near industrial outfalls and urban settlements. Sediment samples were collected using a Van Veen grab sampler from a depth of 5–10 cm. Concurrently, surface water samples were collected using pre-cleaned polyethylene bottles. All samples were immediately placed in ice-cooled containers and transported to the laboratory to prevent physicochemical changes and preserve metal concentrations.

Sediments:. Upon arrival, sediment samples were air-dried at room temperature, homogenized using an agate mortar and pestle, and sieved through a 2 mm mesh to remove coarse debris. The fine fraction was stored in polyethylene bags for subsequent analysis.

Water:. Water samples were filtered using Whatman No. 42 filter paper to remove suspended particulates. On-site measurements were taken for key physicochemical parameters, including pH, turbidity, total dissolved solids (TDS), electrical conductivity (EC), and dissolved oxygen (DO), using portable multi-parameter meters. For metal analysis, a 0.5 g portion of the homogenized sediment was digested using a mixture of concentrated nitric acid (HNO3) and hydrochloric acid (HCl) in a closed-vessel microwave digestion system. Filtered water samples (50 mL) were acidified with ultrapure HNO3 to a pH of <2. The concentrations of heavy metals, including chromium (Cr), nickel (Ni), copper (Cu), arsenic (As), cadmium (Cd), and lead (Pb), were determined using Energy Dispersive X-Ray Fluorescence (EDXRF) spectrometry. The instrument was calibrated with certified reference materials (CRMs) for both sediments (e.g., MESS-4) and water to ensure accuracy. All measurements were performed in triplicate to ensure precision, and blank samples were analyzed to correct for any potential contamination.

2.2. *CNN GNN MLP Algorithm*

Data Preprocessing. The dataset consists of feature variables ($\mathbf{X}_i$) such as water quality parameters and spatial coordinates. The columns such as `Stations`, `River`, `Lat`, `Long`, and `geometry` are dropped, and only numeric columns are retained for further processing. Missing values are handled by filling them with zeros:

$$\mathbf{X}_i = \text{FillNaN}(\mathbf{X}_i, 0) \tag{1}$$

The data is split into training and testing sets using a random seed, and the target variable y (RI) is extracted.

CNN Input Data. The raster data for CNN input is collected from various layers (Indices, LULC, IDW). The coordinate pairs are used to extract patches from each raster layer using the windowing technique in `rasterio`:

$$\mathbf{patch}_i = \text{ExtractPatch}(\mathbf{coords}, \mathbf{raster\_files}) \tag{2}$$

The patches are normalized to avoid division by zero:

$$\mathbf{patch}_i = \frac{\mathbf{patch}_i}{\max(\mathbf{patch}_i) + \epsilon} \tag{3}$$

where $\epsilon = 1e - 8$ is a small constant added for stability.

MLP Input Data. For the MLP branch, features are extracted from numeric columns and standardized using the `StandardScaler`:

$$\mathbf{X}_i^{\text{mlp}} = \text{Standardize}(\mathbf{X}_i^{\text{numeric}}) \tag{4}$$

GNN Input Data. The spatial relationship between data points is modeled as a graph using a distance-based kernel:

$$\mathbf{G}_{ij} = \exp\left(-\frac{d_{ij}}{\tau}\right) \tag{5}$$

where d_{ij} is the Euclidean distance between coordinates i and j, and τ is a kernel parameter. The graph is used as input to the GNN branch.

CNN Branch:. The CNN branch consists of two convolutional layers followed by max-pooling layers. The output is flattened and passed through a dense layer:

$$\mathbf{h}_i^{\text{cnn}} = \text{Flatten}(\text{MaxPooling}(\text{Conv2D}(\mathbf{patch}_i))) \tag{6}$$

MLP Branch:. The MLP branch processes the numerical input through fully connected layers:

$$\mathbf{h}_i^{\text{mlp}} = \text{Dense}(\text{ReLU}(\mathbf{X}_i^{\text{mlp}})) \tag{7}$$

GNN Branch:. The GNN branch processes the graph input using graph-based convolutions or attention mechanisms:

$$\mathbf{h}_i^{\text{gnn}} = \text{Dense}(\text{ReLU}(\mathbf{G}_i)) \tag{8}$$

Fusion Layer. The outputs from all branches are concatenated:

$$\mathbf{z}_i = \text{Concatenate}(\mathbf{h}_i^{\text{cnn}}, \mathbf{h}_i^{\text{mlp}}, \mathbf{h}_i^{\text{gnn}}) \tag{9}$$

The concatenated vector is passed through dense layers:

$$\mathbf{f}_i = \text{Dense}(\text{ReLU}(\mathbf{z}_i)) \tag{10}$$

$$\hat{y}_i = \text{Dense}(\mathbf{f}_i) \tag{11}$$

2.3. CNN, GNN, MLP with PMF GWR Algorithm

Data Input:. The data input structure for this model remains the same as the standard CNN, GNN, and MLP model. However, the key addition here is the use of Probabilistic Matrix Factorization (PMF) and Geographically Weighted Regression (GWR) for graph feature embedding. The inputs are: - Raster data for CNN input: $\tilde{\mathcal{P}}_i$, - Tabular data for MLP input: $\tilde{\mathbf{x}}_i$, - Graph data for GNN input: $\mathbf{g}_i$.

PMF-based Graph Embedding:. PMF is used to embed the graph kernel $\mathbf{G}_{ij}$ computed from the graph distance. The graph kernel is defined as:

$$\mathbf{G}_{ij} = \exp\left(-\frac{d_{ij}}{\tau}\right), \tag{12}$$

where d_{ij} is the Euclidean distance between sites i and j, and τ is the kernel parameter. The graph-based features are then processed through the GNN branch after PMF embedding.

2.4. CNN, GAT, MLP Algorithm

Data Input:. The input structure remains the same as the previous model, with raster data for CNN, tabular data for MLP, and graph data for GAT. The key difference lies in the processing of graph features via Graph Attention Networks (GAT).

Graph Attention Network (GAT) Input:. In GAT, we compute the attention weight for each node i's neighbors j based on the graph kernel:

$$\mathbf{A}_{ij} = \frac{\exp\left(\mathbf{e}_{ij}\right)}{\sum_{k=1}^{n} \exp\left(\mathbf{e}_{ik}\right)}, \tag{13}$$

where $\mathbf{e}_{ij} = \text{LeakyReLU}(\mathbf{W}_{\text{att}}[\mathbf{h}_i \parallel \mathbf{h}_j])$ is the attention coefficient, $\mathbf{W}_{\text{att}}$ is the learnable weight matrix, and $\parallel$ represents concatenation. The attention mechanism aggregates the neighbors' features weighted by $\mathbf{A}_{ij}$:

$$\mathbf{h}_i^{\text{att}} = \sum_{j=1}^{n} \mathbf{A}_{ij}\mathbf{h}_j. \tag{14}$$

2.5. Transformer CNN GNN MLP Algorithm

Data Input:. The data input structure is similar to the previous models, with raster patches $\tilde{\mathcal{P}}_i$ for CNN, tabular features $\tilde{\mathbf{x}}_i$ for MLP, and graph features $\mathbf{g}_i$ for GNN.

Transformer Fusion (Multi-Head Attention):. The outputs from CNN, MLP, and GNN branches are concatenated into a tensor $\mathbf{T}_i \in R^{3\times d_p}$. The multi-head attention is applied to this tensor as follows:

$$\mathbf{Q}_h = \mathbf{T}_i\mathbf{W}_h^Q, \quad \mathbf{K}_h = \mathbf{T}_i\mathbf{W}_h^K, \quad \mathbf{V}_h = \mathbf{T}_i\mathbf{W}_h^V, \tag{15}$$

$$\text{Attn}_h(\mathbf{T}_i) = \text{softmax}\left(\frac{\mathbf{Q}_h\mathbf{K}_h^\top}{\sqrt{d_k}}\right)\mathbf{V}_h, \tag{16}$$

$$\text{MHA}(\mathbf{T}_i) = \|_{h=1}^{H}\text{Attn}_h(\mathbf{T}_i)\mathbf{W}^O. \tag{17}$$

The multi-head attention outputs are combined, followed by residual connection and layer normalization:

$$\tilde{\mathbf{T}}_i = \text{LayerNorm}\left(\mathbf{T}_i + \text{Dropout}(\text{MHA}(\mathbf{T}_i))\right). \tag{18}$$

Finally, the flattened result is passed through dense layers for prediction:

$$\mathbf{z}_i = \text{Flatten}(\tilde{\mathbf{T}}_i) \in R^{3d_p}, \tag{19}$$

$$\hat{y}_i = \mathbf{w}_3^\top\text{ReLU}\left(\mathbf{W}_3\text{Dropout}\left(\text{ReLU}\left(\mathbf{W}_4\mathbf{z}_i + \mathbf{b}_4\right)\right)\right) + \mathbf{b}_3. \tag{20}$$

2.6. Stacked CNN GNN MLP Algorithm

Data Input:. In this model, raster, tabular, and graph data are processed separately by the CNN, MLP, and GNN branches. The CNN branch takes $\tilde{\mathcal{P}}_i$, the MLP branch processes $\tilde{\mathbf{x}}_i$, and the GNN branch uses $\mathbf{g}_i$.

Model Architecture:. The outputs of each branch are concatenated into a single vector $\mathbf{z}_i$:

$$\mathbf{z}_i = \text{Concatenate}(\mathbf{h}_i^{\text{cnn}}, \mathbf{h}_i^{\text{mlp}}, \mathbf{h}_i^{\text{gnn}}). \tag{21}$$

This concatenated vector is passed through fully connected layers for prediction:

$$\mathbf{f}_i = \text{Dense}(\text{ReLU}(\mathbf{z}_i)), \tag{22}$$

$$\hat{y}_i = \mathbf{W}_5\mathbf{f}_i + \mathbf{b}_5. \tag{23}$$

2.7. GNN MLP Autoencoder Algorithm

Data Input:. In this model, the graph data $\mathbf{g}_i$ is passed through the GNN branch, and the tabular data $\tilde{\mathbf{x}}_i$ is passed through the MLP branch. These two branches' outputs are concatenated.

Latent Representation:. The GNN and MLP outputs are concatenated to form the latent representation $\mathbf{z}_i$:

$$\mathbf{z}_i = \text{Concatenate}(\mathbf{h}_i^{\text{gnn}}, \mathbf{h}_i^{\text{mlp}}). \tag{24}$$

Decoder:. The latent representation $\mathbf{z}_i$ is decoded to reconstruct the target $\hat{y}_i$:

$$\mathbf{f}_i = \text{Dense}(\text{ReLU}(\mathbf{W}_3\mathbf{z}_i + \mathbf{b}_3)), \tag{25}$$

$$\hat{y}_i = \mathbf{W}_4\mathbf{f}_i + \mathbf{b}_4. \tag{26}$$

2.8. *Mixture of Experts Algorithm*

Data Input:. The model consists of three experts: CNN, MLP, and GNN. Each expert processes different types of input data, and their outputs are combined using a gating network.

CNN Expert:. The CNN expert processes raster data as follows:

$$\mathbf{h}_i^{\text{cnn}} = \text{Flatten}(\text{MaxPool}(\text{Conv2D}(\tilde{\mathcal{P}}_i))). \tag{27}$$

MLP Expert:. The MLP expert processes tabular data:

$$\mathbf{h}_i^{\text{mlp}} = \text{ReLU}(\mathbf{W}_1\tilde{\mathbf{x}}_i + \mathbf{b}_1). \tag{28}$$

GNN Expert:. The GNN expert processes graph data:

$$\mathbf{h}_i^{\text{gnn}} = \text{ReLU}(\mathbf{U}_1\mathbf{g}_i + \mathbf{c}_1). \tag{29}$$

Gating Network:. The gating network computes a weight for each expert's output. The weights are computed as follows:

$$\text{Gate Input} = [\mathbf{h}_i^{\text{cnn}}, \mathbf{h}_i^{\text{mlp}}, \mathbf{h}_i^{\text{gnn}}], \tag{30}$$

$$\text{Gate Weights} = \text{Softmax}(\text{Dense}(\text{ReLU}(\text{Gate Input}))), \tag{31}$$

$$\hat{y}_i = \sum_{j=1}^{3} w_j \cdot y_j. \tag{32}$$

2.9. *Dual Attention Algorithm*

Data Input:. The data input consists of raster data $\tilde{\mathcal{P}}_i$ for CNN, tabular data $\tilde{\mathbf{x}}_i$ for MLP, and graph data $\mathbf{g}_i$ for GNN.

CNN with Spatial Attention:. Spatial attention is applied to the CNN output:

$$\text{Spatial Attention Map} = \sigma(\text{Conv1x1}(\tilde{\mathcal{P}}_i)), \tag{33}$$

$$X_{\text{cnn\_attn}} = X_{\text{cnn}} \cdot \text{Spatial Attention Map}. \tag{34}$$

Feature Attention:. Feature attention is applied to the MLP and GNN outputs:

$$X_{\text{mlp\_attn}} = X_{\text{mlp}} \cdot \text{Feature Attention Map}, \quad X_{\text{gnn\_attn}} = X_{\text{gnn}} \cdot \text{Feature Attention Map}. \tag{35}$$

Gating Network:. The outputs of the attention-modulated branches are combined:

$$\text{Gate Input} = [\mathbf{h}_i^{\text{cnn}}, \mathbf{h}_i^{\text{mlp}}, \mathbf{h}_i^{\text{gnn}}], \tag{36}$$

$$\text{Gate Network} = \text{Dense}(\text{ReLU}(\text{Dense}(\text{Gate Input}))), \tag{37}$$

$$\hat{y}_i = \sum_{j=1}^{3} w_j \cdot y_j. \tag{38}$$

2.10. Evaluation Metrics

For predictions $\hat{y}_i$ and truths y_i:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \qquad \text{RMSE} = \sqrt{\frac{1}{N} \sum_i (y_i - \hat{y}_i)^2}, \tag{39}$$

$$\text{MAE} = \frac{1}{N} \sum_i |y_i - \hat{y}_i|, \qquad \text{SMAPE} = \frac{100}{N} \sum_{i=1}^{N} \frac{|\, \hat{y}_i - y_i \,|}{\frac{1}{2}\big(|y_i| + |\hat{y}_i|\big) + \varepsilon}, \tag{40}$$

with a small ε to avoid division by zero.

2.11. Permutation Feature Importance

Permutation Feature Importance (PFI) evaluates the impact of individual features on the performance of the model. The key concept is to permute the values of a feature and observe how it affects the model's performance. The following steps describe the algorithm:

Train the model. on the original dataset and compute the baseline performance, R^2_{baseline}.

Permute the feature values. for each feature f_k in the dataset:

$$\mathbf{X}^{(k)}_{\text{shuffled}} = \text{Shuffle}(\mathbf{X}^{(k)}), \tag{41}$$

where $\mathbf{X}^{(k)}$ represents the feature column k of the dataset.

Re-evaluate the model. on the permuted data:

$$R^2_{\text{permuted}}{}^{(k)} = \text{Model}(\mathbf{X}^{(k)}_{\text{shuffled}}). \tag{42}$$

The importance of each feature f_k is computed as the difference between the baseline and the permuted model performance:

$$f_k = R^2_{\text{baseline}} - R^2_{\text{permuted}}{}^{(k)}. \tag{43}$$

The greater the reduction in performance, the more important the feature. This method is applied to three types of inputs: the CNN input (raster data), the MLP input (tabular data), and the GNN input (graph-based data). For each, the feature importance is calculated as follows:

$$\text{Importance}_{\text{CNN}} = R^2_{\text{baseline}} - R^2_{\text{permuted\_CNN}} \tag{44}$$

$$\text{Importance}_{\text{MLP}} = R^2_{\text{baseline}} - R^2_{\text{permuted\_MLP}} \tag{45}$$

$$\text{Importance}_{\text{GNN}} = R^2_{\text{baseline}} - R^2_{\text{permuted\_GNN}}. \tag{46}$$

This allows the identification of the most influential features in the model.

2.12. LIME (Local Interpretable Model-agnostic Explanations)

LIME provides local explanations for individual predictions by approximating the black-box model with an interpretable surrogate model in the neighborhood of the instance being explained. The process is as follows:

Select an instance. i to explain. For instance, choose the 5th instance in the test set. Generate perturbed instances by sampling around the instance i and perturbing the feature values:

$$\mathbf{X}_{\text{perturbed}}^{(i)} = \mathbf{X}_{\text{base}} + \epsilon, \tag{47}$$

where ϵ is a random perturbation.

Predict the outcomes. for each perturbed instance using the black-box model:

$$\hat{y}_{\text{perturbed}} = f(\mathbf{X}_{\text{perturbed}}), \tag{48}$$

where f represents the black-box model.

Fit a local surrogate model. , typically a linear model or decision tree, on the perturbed data and predicted outcomes:

$$\hat{y} = \mathbf{w}^{\top}\mathbf{z} + b, \tag{49}$$

where $\mathbf{w}$ is the coefficient vector, $\mathbf{z}$ is the perturbed feature vector, and b is the bias term.

The

local model coefficients. $\mathbf{w}$ provide the explanation of feature importance for the instance being explained. Larger absolute values of w_k indicate more important features. The final prediction is approximated by the surrogate model:

$$\hat{y}_i = \mathbf{w}_1^{\top}\mathbf{z}_i + b. \tag{50}$$

where $\mathbf{z}_i$ is the perturbed feature vector for the instance i, and $\mathbf{w}_1$ is the learned weight vector of the local model. In this case, the LIME explainer uses the MLP and GNN features, while keeping the CNN input fixed. The perturbation is done on the features of the MLP and GNN only, and the output is the feature importance of each individual feature in the neighborhood of the instance. The explanation is visualized as a bar plot showing the importance of each feature, as determined by the surrogate model's coefficients $\mathbf{w}_1$.

2.13. Data Post Processing

Following model training, a rigorous post-processing pipeline was implemented to evaluate predictive performance, interpret model behavior, and quantify the contribution of each data modality to the prediction of sediment risk index (RI). Model outputs (predicted RI values) were first compared against observed values using multiple complementary statistical indicators to capture both accuracy and robustness. The coefficient of determination (R2) was employed to quantify the proportion of variance in RI explained by the models, while the root mean square error (RMSE) and mean absolute error (MAE) provided measures of absolute prediction deviation and robustness against outliers. In addition, the symmetric mean absolute percentage error (SMAPE) was computed to standardize prediction errors relative to the magnitude of observed values, thereby ensuring comparability across different concentration ranges. Beyond predictive accuracy, feature attribution analyses were carried out to disentangle the contribution of each input modality–CNN-derived raster patches, MLP-based tabular attributes, and GNN-based spatial adjacency features–to the overall ensemble model. This was achieved through permutation feature importance, whereby the input values of each modality were randomly shuffled while holding the others constant, and the resulting decline in R 2 was recorded as a measure of relative importance. This procedure enabled quantification of how strongly each data stream (spectral indices, interpolated heavy metal rasters, anthropogenic buffers, hydrological distances, or spatial autocorrelation matrices) contributed to predictive skill. CNN-based features, for example, captured fine-scale spatial variability in environmental indices such as NDVI, NDWI, and SAVI; MLP-based attributes reflected standardized geochemical and textural variables; and GNN-based inputs encoded spatial dependence across sites. Together, the permutation results were summarized as importance scores, ranked by predictive loss, and used to interpret which environmental and anthropogenic drivers were most influential during the rainy season. To complement these global insights, Local Interpretable Model-agnostic Explanations (LIME) were employed to interpret individual model predictions. LIME works by approximating the complex black-box model with a simpler, more interpretable surrogate model, similar to linear regression, in the local vicinity of a specific prediction. This was achieved by generating a new, perturbed dataset by randomly sampling instances around a given observation point, weighting these new instances by their proximity to the original observation, and training a simple, local model on this weighted dataset. By analyzing the coefficients of this local model, it is possible to identify which features most strongly influenced the predicted RI value for a specific sampling site. For instance, LIME could reveal that a high predicted RI at a particular location was predominantly driven by high concentrations of interpolated lead and its proximity to an industrial area, providing fine-grained, sample-level explanations critical for targeted remediation efforts. This post-processing framework, combining statistical evaluation, feature attribution, and spatial visualization, ensured that the predictive outputs of the ensemble models were not only quantitatively validated but also ecologically interpretable, facilitating a robust discussion on the sources and spatial dynamics of heavy metal contamination in riverine sediments.

booktabs xcolor multirow graphicx

3. Results And Discussion

3.1. Heavy Metal Distribution

3.1. Heavy Metal Distribution

The results from the sediments and water of urban rivers around Dhaka highlight substantial seasonal and spatial variability in the concentration of heavy metals. The data indicates that winter typically exhibits higher concentrations of pollutants than the rainy season, particularly due to lower water flow and higher industrial discharge during the drier months. In sediments, Chromium (Cr) concentrations averaged 106.58 mg/kg in winter and 92.69 mg/kg during the rainy season, with the highest concentration (118.1 mg/kg) recorded at Station S-12 in winter. These values are higher than those found in similar studies of urban rivers, which often see elevated levels of Cr due to industrial discharges (Chowdhury et al., 2018). Similarly, Nickel (Ni)levels ranged from 11.72 to 85.22 mg/kg, with mean concentrations of 51.44 mg/kg in winter and 42.13 mg/kg in the rainy season. Copper (Cu) displayed a significant seasonal difference, with an average of 117.55 mg/kg in winter, compared to 79.92 mg/kg during the rainy season. The highest concentration (261.1 mg/kg) was found at Station S-12 during winter, which can be attributed to industrial runoff, particularly from metal processing and battery manufacturing (Hossain et al., 2021). Arsenic (As) concentrations ranged from 7.49 to 28.37 mg/kg, with an average of 16.74 mg/kg in winter and 13.12 mg/kg in the rainy season. Cadmium (Cd) levels, though lower, showed distinct seasonal variation, averaging 1.34 mg/kg in winter and 0.83 mg/kg in the rainy season. These values are concerning because Cd is highly toxic even at low concentrations, leading to long-term environmental degradation (Rahman et al., 2020). Lead (Pb) concentrations were particularly high, with a mean of 78.74 mg/kg in winter, decreasing to 54.41 mg/kg during the rainy season. The highest concentration (297.4 mg/kg) was recorded at Station S-12 during winter, pointing to significant anthropogenic contamination, likely from automotive emissions and industrial effluents (Chowdhury et al., 2018). In river water, heavy metals also demonstrated significant seasonal variation. Chromium (Cr)concentrations averaged 0.22 mg/L in winter and 0.18 mg/L during the rainy season, with most values exceeding the WHO standard of 0.05 mg/L (World Health Organization, 2011). Nickel (Ni)concentrations were higher in winter (0.31 mg/L) than in the rainy season (0.25 mg/L), with peaks observed at Stations S-9 and S-10, indicative of localized pollution sources. Copper (Cu) levels averaged 1.98 mg/L in winter and 1.66 mg/L in the rainy season, with some stations exceeding the WHO permissible limit of 2.0 mg/L. This is particularly concerning for aquatic health, as Cu is known to be toxic to aquatic organisms (USEPA, 2018). Arsenic (As) concentrations ranged between 0.10 and 0.50 mg/L, with mean values of 0.33 mg/L in winter and 0.29 mg/L in the rainy season, both exceeding safety limits and posing significant health risks to populations relying on these water sources for drinking and agriculture (WHO, 2011). Cadmium (Cd) concentrations were relatively low, averaging 0.019 mg/L in winter and 0.014 mg/Lin rainy, but these levels still surpass permissible limits for drinking water, emphasizing the long-term dangers of Cd exposure (IARC, 2012). Lead (Pb) concentrations were 0.13 mg/L in winter and 0.10 mg/L in the rainy season, consistently exceeding WHO standards for safe drinking water. The highest Pb concentration in water (0.25 mg/L) was recorded at Station S-15 during winter, indicating significant pollution and potential risks to human health, including neurological damage and developmental delays in children (ATSDR, 2007). These findings clearly indicate that both sediment and water samples from the urban rivers of Dhaka frequently exceed permissible limits for several heavy metals, especially during the winter season. This has important implications for both ecological health and human well-being, as exposure to these metals can cause long-term environmental damage and increase the risk of chronic diseases. The high concentrations of Pb, Cu, and Cd in particular are of major concern, as they pose significant ecological risks, including toxicity to aquatic organisms and bioaccumulation through the food chain (Rahman et al., 2020).

3.2. Igeo

The criterion to evaluate the metal pollution in sediments is the Igeo that has been widely used since the late 1960s which is calculated by Eq (X) shown in Table X. In the present study, Igeo for the elements Cr, Ni, Cu, As, Cd and Pb is measured and presented in Table 3. The calculated Igeo values reveal that Cadmium (Cd) and Lead (Pb) exhibited the highest geo-accumulation indices in all five rivers, indicating significant contamination. In Shitalakshya and Buriganga, the Igeo values for Cd were notably high, indicating strong to extreme pollution in both seasons, with values of 4.74 (highest) in winter and 4.16 in rainy season. Similarly, the Igeo values for Cd in the Turag, Dhaleshwari, and Balu rivers ranged between 2.06 and 3.08 in winter, indicating moderate to strong pollution, with slightly lower values in the rainy season, ranging from 1.74 to 2.63, representing moderate pollution. The Igeo values for Pb in the Buriganga, Turag, and Shitalakshya rivers were found to be 3.21, 2.96, and 2.39 in winter, indicating strong pollution, and 2.47, 2.02, and 1.58 in the rainy season, suggesting moderate to strong pollution. In the Dhaleshwari and Balu rivers, Igeo values for Pbwere relatively lower, 1.34 and 1.05 in winter, indicating moderate pollution, and 0.45 and 0.17 in the rainy season, suggesting unpolluted to moderate pollution. Copper (Cu) exhibited moderate pollution in Buriganga, with Igeo values of 1.67 in winter and 1.10in the rainy season. Arsenic (As) and Copper (Cu) in Shitalakshya and Turag during the rainy season also showed relatively low values, indicating unpolluted to moderately polluted conditions. Metals like Chromium (Cr), Nickel (Ni), and Copper (Cu), across all five rivers in both seasons, showed negative Igeo values, indicating no significant pollution by these metals. This suggests that their concentrations in the sediments were within natural background levels. These findings are consistent with previous studies that used Igeo to assess sediment pollution. For example, Rakib et al. (2021a) found that the Igeo values for metals like Ti, Fe, Cu, Rb, Sr, Zr, Pb, and Zn in the sediments from marine coastal areas in Sitakundo, Bangladesh, were classified as class zero, indicating unpolluted sediments by these metals.

3.3. EF

The Enrichment Factor (EF) is a widely used index for assessing the anthropogenic impact on sediment contamination. It quantifies the contribution of heavy metals to the sediment enrichment relative to background levels. The EF values were calculated for Cr, Ni, Cu, As, Cd, and Pb in the sediments from the Balu, Buriganga, Dhaleshwari, Shitalakshya, and Turag rivers during both the winter and rainy seasons. The EF values for each metal in the sediments are summarized in Table 2. The river system with the highest EF values was Shitalakshya, followed by Buriganga, Turag, Dhaleshwari, and Balu, indicating the degree of pollution in these rivers. The Shitalakshya river showed the highest enrichment in both seasons, particularly for Cd. In winter, the EF values for Cdranged from 28.03 to 48.96, and in the rainy season, they ranged from 28.66 to 38.63, indicating extremely high enrichment (EF 40) in winter and very high enrichment (20 EF < 40) in the rainy season. Similarly, Pb concentrations in Shitalakshya also exhibited significant enrichment, with EF values ranging from 4.47 to 12.20 in winter and 3.84 to 5.87 in the rainy season. These values indicate significant anthropogenic pollution, with industrial discharges and untreated waste likely contributing to the contamination (Hossain et al., 2021b). Following Shitalakshya, Buriganga exhibited the second-highest enrichment, especially for Cd and Pb. The EF values of Cd in Buriganga ranged from 18.23 to 21.01 in winter and 13.42 to 17.44 in the rainy season, indicating significant enrichment (5 EF < 20) in both seasons. The EF values of Pb in Buriganga ranged from 9.98 to 10.30 in winter and 5.63 to 5.86 in the rainy season, indicating moderate to strong enrichment. Turag exhibited moderate enrichment for Pb, with EF values ranging from 7.10 to 8.78 in winter and 3.32 to 5.50 in the rainy season. These findings suggest that Turag is moderately affected by industrial pollution and vehicular emissions, particularly in the winter season (Hossain et al., 2021b). The Dhaleshwari and Balu rivers showed moderate to significant enrichment for Cd and Pb, but with less severe pollution than Shitalakshya and Buriganga. In Dhaleshwari, the EF values of Cdranged from 3.60 to 15.52 in winter and 4.96 to 7.81 in the rainy season. Similarly, in Balu, Cdconcentrations showed moderate enrichment, with EF values ranging from 4.50 to 6.17 in the rainy season. Pb in Dhaleshwari and Balu exhibited lower EF values, ranging from 1.61 to 5.56 in winter and 1.19 to 2.94 in the rainy season, indicating minimal to moderate enrichment. In terms of Copper (Cu), the EF values were relatively lower, but still indicative of moderate enrichment in the Buriganga, Dhaleshwari, and Shitalakshya rivers. In Buriganga, the EF values of Cu ranged from 1.67 to 1.10 in winter and 1.10 to 1.60 in the rainy season, pointing to moderate contamination. Dhaleshwari and Shitalakshya also showed moderate enrichment for Cu, with EF values ranging from 1.00 to 2.00 in both seasons, suggesting a moderate anthropogenic impact. For metals like Chromium (Cr), Nickel

(Ni), and Arsenic (As), the EF values were consistently low (EF < 2) in all rivers, indicating minimal to no anthropogenic enrichment. This suggests that these metals are either naturally occurring or have not been significantly impacted by human activities in these river sediments. Overall, the Shitalakshya and Buriganga rivers were found to be the most heavily polluted, with high EF values for Cd and Pb, indicating severe anthropogenic contamination, likely due to industrial discharges, vehicular emissions, and untreated sewage. The Turag, Dhaleshwari, and Balu rivers, while still contaminated, showed moderate enrichment for metals like Cd and Pb, suggesting that industrial and urban runoff have a moderate influence on these river systems. These results are consistent with findings from previous studies. Hossain et al. (2021b) reported moderate to severe enrichment of metals like Mn, Zn, Cu, Pb, Ni, and Cr in the sediments from Sitakundo coastal areas in Bangladesh, which align with the contamination levels observed in this study. Additionally, Tamim et al. (2016) found minimal enrichment of Cr, Zn, and other metals in the Buriganga River near the Hazaribagh area, indicating that, while certain areas of the river remain relatively unpolluted, others show significant industrial contamination. Based on the EF values, the rivers in Dhaka can be ordered in terms of pollution severity as: Shitalakshya > Buriganga > Turag > Dhaleshwari > Balu.

3.4. PLI

The limitations of single metal indices led to the development of multi-metal indices. The two most widely used such indices, developed by Hakinson (1980) and Nemerow (1991), include the modified degree of contamination (mCd) and the pollution index (PI). Brady et al. (2015) developed a modified pollution index (MPI) considering enrichment factor. The Pollution Index (PI), Modified Pollution Index (MPI), and Modified Degree of Contamination (mCd) values from this study indicate significant contamination in the urban rivers of Dhaka, with Shitalakshya and Buriganga showing the highest pollution levels. The PI values of 29.59 (winter) and 22.71 (rainy season) in Shitalakshya, and similar high values in Buriganga, confirm heavily polluted conditions (PI > 3) in both rivers. This is consistent with the findings of Chowdhury et al. (2018), who identified Buriganga as heavily polluted due to industrial effluents and urban runoff. Similarly, mCd values of 9.2 and 6.67 for Shitalakshya in winter and rainy seasons respectively indicate severe pollution (mCd > 8), while Buriganga recorded a mCd of 8.15 in winter. Ahsan et al. (2019) and Rahman et al. (2020) also reported high contamination in these rivers, with Shitalakshya and Buriganga being the most impacted by industrial waste. Turag, Dhaleshwari, and Balu rivers showed moderate pollution with mCd values ranging from 2 to 4, reflecting lower contamination compared to Shitalakshya and Buriganga. In the rainy season, these rivers showed some reduction in pollution levels, supporting the seasonal dilution effect observed in other studies (Bashar et al., 2019). The MPI values (>10) in Shitalakshya and Buriganga, and values between 5 and 10 in Turag and Dhaleshwari further highlight the varying levels of contamination, with Shitalakshya > Buriganga > Turag Dhaleshwari > Balu in both seasons.

3.5. Health Risk Assessment

The Health Risk Assessment of sediments and water from the rivers surrounding Dhaka, Bangladesh, reveals severe contamination by key heavy metals such as Chromium (Cr), Nickel (Ni), Copper (Cu), Arsenic (As), Cadmium (Cd), and Lead (Pb), all of which exceed permissible limits set by environmental and health standards (USEPA, 2018). In sediments, Chromium, Nickel, and Arsenic present significant carcinogenic risks, with Chromium reaching 118.1 mg/kg in the winter, a level that poses a notable long-term cancer risk through ingestion or dermal contact (IARC, 2012). Lead, Copper, and Cadmium in sediments contribute primarily to non-carcinogenic risks, including neurological damage (especially in children), kidney and liver toxicity, as well as bone damage (ATSDR, 2007; EPA, 2020). These findings are in line with previous studies highlighting Lead as a significant neurotoxin, particularly in urban water systems (Chowdhury et al., 2018). The water quality assessment reveals that while the water concentrations of these metals, particularly Lead, Arsenic, and Copper, are generally lower than those in sediments, they still exceed the USEPA limits for safe drinking water. This indicates non-carcinogenic health risks such as gastrointestinal distress and liver toxicity from Copper, and neurological impairments from Lead (WHO, 2011; EPA, 2018). These metals, though somewhat diluted in the rainy season, remain at harmful levels, with Nickel and Arsenic still presenting significant risks, indicating persistent contamination from anthropogenic sources (Bashar et al., 2019). The study further suggests that winter months may exacerbate contamination levels due to reduced runoff, leading to more concentrated pollution from industrial discharge and domestic waste (Chowdhury et al., 2018). Despite the seasonal dilution during the rainy season, the findings indicate continuous pollution with metals like Nickel and Arsenic remaining at levels harmful to both human health and aquatic ecosystems.

3.6. PCA

The Principal Component Analysis (PCA) conducted in this study reveals critical insights into the sources of heavy metal contamination in the urban rivers around Dhaka. The first three principal components (Dim.1, Dim.2, and Dim.3) collectively accounted for 88.6% of the total variance, indicating that these components effectively capture the majority of the data's variability. Dim.1explained 47.4%, Dim.2 explained 21.3%, and Dim.3 accounted for the remaining variance, which is consistent with findings from other studies using PCA for environmental contamination assessments (Chowdhury et al., 2018). The PCA plot showed a clear separation of data points along these dimensions, confirming distinct patterns of contamination across the rivers. Further analysis through rotated components (RC1, RC2, RC3) revealed specific heavy metals that correlate strongly with each component. The first rotated component (RC1) showed strong correlations with Lead (Pb) (0.95) and Nickel (Ni) (0.91), suggesting that industrial emissions and vehicular sources are the primary contributors to the contamination in these rivers. This finding aligns with previous studies, which have reported that Pb and Ni are prevalent in industrial areas and are often associated with vehicular emissions (Bashar et al., 2019; Tariq et al., 2020). The second component (RC2), with significant positive loadings for Chromium (Cr) (0.95) and Copper (Cu) (0.72), suggests a mix of geological and anthropogenic sources, particularly from construction materials and industrial discharges (Rahman et al., 2020). The third component (RC3) was strongly dominated by Arsenic (As) (0.98), indicating that groundwater contamination or specific industrial processes (such as those in mining or textile industries) could be significant sources of Arsenic contamination, as previously noted in regions with industrial waste and agricultural runoff (Chowdhury et al., 2018). These results confirm that contamination in the rivers around Dhaka can be attributed to three major sources: industrial processes, vehicular emissions, and groundwater contamination.

3.7. Monte Carlo

The Monte Carlo simulation for the Ecological Risk Index (RI), conducted over 10,000 iterations, provides a robust assessment of ecological risks associated with heavy metal contamination in the studied rivers. By applying both normal and lognormal distributions, the simulation effectively captures the range of potential ecological risks under different pollution scenarios. The normal distribution revealed moderate ecological risks, with a mean RI value of 250.6 and a maximum of 817.4, indicating a positive skew. This suggests that most of the study area experiences moderate ecological risks, likely stemming from chronic pollution sources such as industrial runoff and urban waste (USEPA, 2018). The 1st quartile (161.5) and median (235.2) values further support this conclusion, indicating that a significant portion of the area is subjected to moderate pollution levels. In contrast, the lognormal distribution, which better reflects the skewed nature of environmental data, yielded a mean of 254.78 and a maximum of 2205.86. These results underscore the potential for severe ecological risks in worst-case scenarios, such as industrial spills or large-scale contamination events, aligning with findings from Chowdhury et al. (2018), which highlighted the significant impact of industrial pollution in the region. Sensitivity analysis of the simulation revealed that Lead (Pb) posed the most significant ecological risk, contributing to 60.68% of the total ecological risk in the rainy season. Other metals, such as Copper (Cu) (17.17%) and Chromium (Cr) (9.04%), also played crucial roles in shaping the ecological risk profile. In the winter season, Lead (Pb) continued to dominate, albeit with a slight reduction in its contribution to 60.68%, while Cadmium (Cd) and Copper (Cu) also contributed substantially. This clearly indicates that Lead (Pb) and Copper (Cu) are the primary contributors to ecological risks in both seasons, reinforcing findings from earlier studies that emphasized the dominance of Pb and Cu in heavily polluted river systems (Rahman et al., 2020). These findings emphasize the urgent need for targeted pollution control measures, particularly in industrial areas, where Lead and Copper are likely being released into the rivers. The results also highlight the importance of continuous environmental monitoring to address both chronic and extreme pollution risks effectively. Early warning systems and improved industrial waste management strategies are essential to mitigate both ecological and health risks, ensuring long-term sustainability for the affected river systems and surrounding communities.

3.8. Model Performance

The performance of the ensemble models provided in Table **??** varied significantly between the rainy and winter seasons, with a notable improvement in predictive accuracy during the winter. This seasonal difference provides key insights into the challenges of modeling heavy metal contamination in dynamic riverine systems. In the rainy season, the models faced greater complexity due to increased hydrological flow, sediment redistribution, and the transport of contaminants. Despite these challenges, the Transformer CNN GNN MLP model demonstrated exceptional performance, achieving the highest coefficient of determination (R^2) of 0.9604 and the lowest root mean square error (RMSE) of 15.7421. This indicates that its architecture, which is adept at integrating and processing diverse data modalities—specifically, high-dimensional raster patches, tabular attributes, and spatial adjacency information—is particularly effective in capturing the intricate spatiotemporal dynamics of this season. The GNN MLP AE and GNN MLP models also performed very well, with R^2 values of 0.9581 and 0.9519, respectively, and low mean absolute errors (MAE) of 14.4920 and 15.7284. This underscores the importance of combining graph-based spatial learning with multi-layer perceptrons to accurately model contaminant behavior under high-flow conditions. Conversely, models with less sophisticated architectures, such as the Dual Attention and CNN GNN MLP, struggled to achieve the same level of accuracy, with R^2 values of 0.8608 and 0.9089, and higher RMSE values of 29.4955 and 23.8654, respectively. This suggests that models that don't effectively fuse all three data modalities lose significant predictive power.

Table 1: Ensemble Model Metrics of Rainy and Winter Season (Sorted by Accuracy)

Rainy	Acc	MSE	RMSE	MAE	**Winter**	Acc	MSE	RMSE	MAE
Transformer CNN GNN MLP↑	0.9604	15.7421	13.2640	9.5200	Transformer CNN GNN MLP↑	0.9721	7.9921	6.5526	4.4510
GNN MLP AE↑	0.9581	15.8938	14.4920	10.1211	GNN MLP AE↑	0.9718	8.0434	7.3433	5.8565
CNN GNN MLP PG↑	0.9570	16.3939	11.9147	8.2470	GNN MLP↑	0.9705	11.0783	8.2154	5.9768
GNN MLP↑	0.9519	17.3337	15.7284	10.9342	Mixture of Experts ↑	0.9700	13.7056	10.0531	6.4779
CNN GAT MLP	0.9266	21.4275	18.8062	11.1605	Stacked CNN GNN MLP	0.9685	14.0418	10.9236	6.2405
Stacked CNN GNN MLP	0.9240	21.7977	21.6243	19.1205	CNN GNN MLP PG	0.9541	16.9342	15.1297	4.9237
CNN GNN MLP↓	0.9089	23.8654	20.9455	13.4076	CNN GAT MLP↓	0.9177	20.3767	14.8821	10.2815
Mixture of Experts↓	0.9070	24.1163	18.4791	12.0654	CNN GNN MLP↓	0.8768	27.7525	21.4899	10.1928
Dual Attention↓	0.8608	29.4955	24.6829	13.3766	Dual Attention↓	0.8402	31.9021	29.2886	19.2817

The winter season presented a more stable environment, characterized by reduced rainfall and lower water flow, which resulted in a marked improvement in overall model performance. Nearly all models achieved an R^2 greater than 0.95. The Transformer CNN GNN MLP model once again led the pack, achieving an impressive R^2 of 0.9721 and a remarkably low RMSE of 7.9921. This a 49.38% reduction in RMSE compared to its performance in the rainy season, highlighting how much easier it is to model heavy metal concentrations when hydrological transport is minimized. The GNN MLP AE and GNN MLP models also showed significant improvement, with R^2 values of 0.9718 and 0.9705, and very low MAE values of 7.3433 and 8.2154, respectively. This high level of accuracy suggests that during the winter, the spatial distribution of heavy metals is largely governed by stable, predictable factors such as sediment composition and proximity to sources, making the modeling task more straightforward. The CNN GNN MLP PG model also showed a notable jump in performance, achieving an R^2 of 0.9541 and a very low SMAPE of 4.9237, indicating highly accurate predictions relative to the magnitude of observed values. The consistent top performance of the Transformer CNN GNN MLP model across both seasons highlights its robust architecture and its ability to effectively handle complex, multi-modal data under varying environmental conditions, making it a highly reliable tool for environmental monitoring.

3.9. Source Apportionment in Rainy Season

The ensemble models for the rainy season reveal a complex set of drivers for heavy metal contamination, with a strong emphasis on textile, tannery, and brick kiln activities. The Transformer CNN GNN MLP model showed that CNN raster patches were the most important feature with a high permutation importance of 0.258, indicating that the detailed, fine-scale spatial variability of features associated with these sources is a primary driver of risk. LIME analysis further reinforced this, with high values of IDW Cu (>0.91) and IDW Pb (>0.63) being major drivers of high-risk predictions, with importance scores of 40.68 and 37.53, respectively. These specific heavy metals are known to be byproducts of industrial processes such as those found in textile and tannery operations. The GNN MLP Autoencoder model further reinforced the significance of geochemical features, with the overall MLP branch having an importance of 2.10 and IDW Pb (0.321), IDW Cu (0.116), and IDW Ni (0.104) being the most influential individual features. The model also showed that IDW Cd (0.043) and IDW Clay (0.036) were significant features. A key insight from this model is that while the number of industries and brick kilns may not always correlate directly with higher risk at every location, their influence is significant and location-specific, as evidenced by the negative LIME importance scores for Number of Industry > 1.00 (-5.17), indicating that their impact on RI is more complex than a simple linear relationship. The CNN GNN MLP with PMF GWR model's permutation importance analysis highlighted Num Brick Field (0.0057) and IDW Silt (0.0056) as the top features, while LIME confirmed that high IDW Cd (>0.13) and IDW Pb (>0.63) were positively correlated with risk, with importance scores of 0.38 and 0.16. The co-occurrence of high importance for interpolated heavy metals and anthropogenic features, particularly in the LIME analysis, strongly suggests that textile, tannery, and brick kiln activities are the probable sources of the heavy metal contamination, with their impact modulated by environmental factors like silt content and fine-scale spatial variations captured by the CNN rasters.

3.10. Source Apportionment in Winter Season

The ensemble models for the winter season reveal that interpolated heavy metal concentrations and fine-scale spatial features are the primary drivers of the sediment risk index (RI), with a direct link to anthropogenic sources. The Transformer CNN GNN MLP model, through permutation importance, identified IDW Pb (0.2298) and IDW Fe (0.2049) as the most influential features, with CNN All Rasters (0.1421) also having high importance, indicating the crucial role of fine-scale spatial variability. This is further validated by the LIME analysis, where IDW Pb (291.3201), IDW Fe (229.6694), IDW Cd (174.8315), and IDW Ni (151.1109) were found to be the most significant contributors to high-risk predictions. The GNN MLP Autoencoder model reinforced these findings, with IDW Pb having an exceptionally high permutation importance of 1.8036, and its LIME analysis showing IDW Fe (2.6344), IDW Pb (1.9294), and IDW Cu (1.7351) as the top drivers of risk. This model also showed an unexpected positive importance for a low number of brick kilns (0.9112), suggesting a complex, non-linear relationship where other factors dominate in less-impacted areas. The GNN MLP model highlighted IDW Cd (0.2251) as its most important feature, and its LIME analysis provided a key insight into source attribution: a high positive importance for Hydro Dist Brick (3.1667) suggests that proximity to brick kilns directly contributes to increased RI. While the overall permutation importance of anthropogenic features like Num Industry (0.0012) and Num Brick Field (-0.0008) appears low, their influence is significant at a localized level, particularly for brick kilns. The consistent and high importance of IDW Pb, IDW Fe, IDW Ni, and IDW Cd across all models, combined with the LIME-based evidence of brick kiln proximity, strongly suggests that these heavy metals are the main contaminants which is also the main by product of tannery and textile industry, with heavy metals spatial distribution and concentration being the primary drivers of risk during the winter season.

4. Conclusion

Summarize your findings and suggest possible future work.

Acknowledgment

Thank any people, organizations, or funding sources.