Ecological Informatics 93 (2026) 103598

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Assessing the impact of land use and land cover on predicting in-stream total phosphorus using GIS and machine learning models

Hye Won Lee [a,1], Min Kim [b,1], Baehyun Min [c,*], Jung Hyun Choi [a,*]

[a] Department of Environmental Science and Engineering, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03765, Republic of Korea
[b] Severe Storm Research Center, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03765, Republic of Korea
[c] Department of Climate and Energy Systems Engineering, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03765, Republic of Korea

ARTICLE INFO

Keywords:
Convolutional neural networks
Total phosphorus
Land use and land cover
SHAP analysis
IG analysis
Water quality management
Spatial distribution

ABSTRACT

This study presents a spatially explicit framework for predicting in-stream total phosphorus (TP) concentrations by combining a convolutional neural network (CNN) with explainable AI techniques. A CNN–dense neural network (CNN–DNN) model was trained on five-year average in-stream TP data from 608 monitoring stations across the four major river basins of the Republic of Korea, using 22 detailed land use and land cover (LULC) classes, a digital elevation model (DEM), and slope. The model outperformed proportion-only regression, highlighting the added value of spatial LULC representation. Unlike previous TP prediction studies that relied primarily on proportional LULC metrics, our approach preserves the spatial arrangement of LULC patches through CNN-based feature extraction and integrates SHapley Additive exPlanations (SHAP) and Integrated Gradients (IG) to provide spatially explicit attribution of patch-level influences. This enables detection of near-stream clustering effects and landscape configurations that cannot be captured by proportion-only models. Model interpretation with SHAP revealed that residential zones generally reduced TP, while commercial and public utility areas increased it when clustered near streams. Paddy fields consistently exerted the strongest positive influence in agricultural watersheds, modulated by slope and hydrologic connectivity. Forests largely buffered TP, though mixed stands and adjacent industrial parcels occasionally increased nutrient loads. To enhance robustness, IG analysis with three baselines revealed consistent attribution patterns that reinforced SHAP findings. Our results reveal that spatial watershed structure (LULC, DEM, and slope) alone can reproduce long-term TP patterns, suggesting that watershed structure is a primary control on phosphorus loads and that land-use planning can serve as an effective long-term management tool independent of short-term meteorological variation. Scenario-based simulations further showed that forest-to-residential conversions slightly lowered TP, whereas industrial and agricultural conversions—particularly to paddy fields—substantially increased TP, with central clusters producing the largest impacts. Overall, the CNN–SHAP–IG framework provides local interpretability and domain-wide robustness, representing a methodological advancement over proportion-based studies and offering a practical decision-support tool for watershed management and spatially informed land-use planning to mitigate phosphorus pollution.

1. Introduction

Total Phosphorus (TP) strongly influences aquatic ecosystems and has been identified as a key driver of algal blooms in freshwater systems (Cheng et al., 2018; Lee et al., 2022; Shen et al., 2020). Rising TP concentrations are linked to human-induced land use and land cover (LULC) change and altered rainfall regimes under climate change, which accelerate nutrient runoff and intensify bloom events (Ahmadisharaf et al., 2020; Alamdari et al., 2022). Agricultural fertilizer use, urban impervious surfaces, and deforestation increase nutrient transport, while the spatial arrangement and proximity of land uses to water bodies further determine runoff pathways. Thus, understanding the coupled effects of LULC, sediment transport, and climate variability is crucial for addressing water quality challenges. Recent evidence consistently demonstrates that LULC change substantially affects hydrology and surface water quality across varied watershed types, including

* Corresponding author.
E-mail addresses: bhmin01@ewha.ac.kr (B. Min), jchoi@ewha.ac.kr (J.H. Choi).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ecoinf.2026.103598
Received 13 December 2024; Received in revised form 3 January 2026; Accepted 3 January 2026
Available online 7 January 2026

Mediterranean reservoirs, rapidly urbanizing basins, and future climate–land use scenarios (Alamdari et al., 2022; Mayou et al., 2024; Molina-Navarro et al., 2014). Global analyses confirm that LULC configuration shapes river water quality (Clark et al., 2022), while remote-sensing-based reviews provide further insights into LULC–water quality linkages (Cheng et al., 2022). These findings reinforce the central role of LULC patterns in determining nutrient dynamics and highlight the need for spatially explicit approaches.

These complex interactions among LULC, sediment transport, and climate change create nonlinear relationships that pose significant challenges for conventional process-based models, which often struggle to represent feedbacks, thresholds, and multivariate linkages while remaining sensitive to input uncertainties (Zhang et al., 2022; Du and Pechlivanidis, 2025; Höge et al., 2022; Fu et al., 2020; Liu et al., 2020). As a result, machine learning (ML) techniques have become increasingly important for capturing nonlinear and dynamic processes in water quality. ML applications have demonstrated strong potential in predicting nutrient concentrations and modeling ecosystem dynamics, including chlorophyll-a forecasting, algal bloom prediction, and ecological structure modeling (Hirn et al., 2024; Kruk, 2023; Ly et al., 2021). Additional ML/DL research has evaluated stream water quality prediction using ANN, SVM, GMDH, hybrid CNN–LSTM architectures, ensemble learning, and AutoDL methods (Adedeji et al., 2022; Barzegar et al., 2020; Haghiabi et al., 2018; Prasad et al., 2022), while recent reviews summarize advances in DL-based modeling and uncertainty assessment (Rabby et al., 2026; Wai et al., 2022; Zhi et al., 2024).

Over the past decade, ML has been increasingly applied to water quality research, offering new opportunities to predict pollutant concentrations and inform management strategies (Chen et al., 2021; Hanson et al., 2020; Kim et al., 2020; Shen et al., 2020; Shim and Choi, 2024). Case studies have demonstrated its versatility, including forecasting water quality goal achievement in Korean river basins, identifying key drivers of nutrient dynamics in artificial lakes, predicting algal blooms, and assessing feature importance in land-use variables across diverse watersheds (Lee et al., 2022; Ly et al., 2021; Venkateswarlu and Anmala, 2023). These studies consistently highlight the role of LULC in shaping water quality. However, most have simplified watershed complexity by considering only the proportional extent of each land-use type, thereby overlooking spatial heterogeneity—particularly near rivers—that governs pollutant transport. This underscores the need for spatially explicit approaches that account for the distribution and configuration of LULC to improve the accuracy and interpretability of water quality predictions. These limitations indicate that process-based models often fail to incorporate fine-scale spatial patterns such as riparian land-use configuration and upstream–downstream connectivity, which are critical for pollutant routing. Therefore, a spatially explicit CNN is particularly suitable for this study because it directly learns spatial dependencies from gridded LULC inputs without requiring manual parameterization.

Convolutional neural networks (CNNs) have recently gained attention in hydrological and water quality research as powerful tools for capturing spatial heterogeneity and complex interactions among watershed variables. For example, CNN-based models have been applied to hyperspectral imagery to classify water quality grades with high accuracy, supporting monitoring and management decisions (Li et al., 2024). CNNs combined with hydrological and water quality data have also been used to estimate aquatic ecosystem health indices such as TDI, BMI, and fish assemblage metrics, demonstrating their ability to integrate spatio-temporal drivers (Kwon et al., 2024). Furthermore, hybrid CNN–CRNN models have been developed to predict water quality indices (WQI), achieving higher performance than conventional methods by simultaneously accounting for spatial and sequential dependencies (Ehteram et al., 2024). Collectively, these studies highlight the growing utility of CNNs in water quality prediction, complementing traditional machine learning approaches while effectively addressing the challenges posed by spatial complexity. However, ML/DL models face well-known challenges in generalizing to unseen spatial domains, as spatial autocorrelation can inflate performance under random data splits (Bartlett et al., 2025; Karpatne et al., 2017; Kazadi et al., 2024; Pakdehi and Ahmadisharaf, 2025; Pakdehi et al., 2024). Compared to process-based models such as SWAT and HSPF, which are often more generalizable due to their mechanistic basis, ML/DL methods offer advantages in capturing nonlinear spatial patterns and enabling explainable AI techniques that yield spatially interpretable insights.

Therefore, this study develops a CNN–dense neural network (CNN–DNN) model for predicting TP concentrations by explicitly incorporating spatial LULC information. The model utilizes 22 detailed LULC classes, along with digital elevation model (DEM) and slope data, to preserve the spatial arrangement of input features and capture their hierarchical patterns. In contrast to RF and SVM models that operate on flat, tabular summaries of LULC, CNNs employ local receptive fields and shared kernels to learn feature maps directly from two-dimensional LULC images. This enables the model to capture patch geometry, adjacency, and near-stream clustering effects that are inherently lost when LULC is represented only as proportional fractions. (Youssef et al., 2024; Zhang et al., 2021). The spatial extent was set to 1000–3000 m based on prior studies showing strong TP responses at this scale (Wang et al., 2023; Zhang et al., 2021), ensuring both near-stream and catchment-scale influences were represented. In addition, scenario-based analyses were performed to evaluate how variations in both the proportion and spatial configuration of LULC types (e.g., urbanization, deforestation, agricultural intensification) affect TP concentrations. The findings provide insights into how future LULC transitions may influence watershed-scale nutrient dynamics, while positioning the proposed model as a decision-support tool for spatially explicit land-use planning and sustainable watershed management.

The main objectives of this study are fourfold: (a) to develop a coordinate-based CNN–DNN model for predicting TP concentrations across the four major river basins of the Republic of Korea; (b) to interpret model outputs using SHapley Additive exPlanations (SHAP) in order to identify how detailed LULC classes and their spatial distributions influence water quality at site-specific scales; (c) to complement SHAP with Integrated Gradients (IG) analysis, thereby ensuring robust and generalized interpretations of LULC impacts across the entire study domain; and (d) to conduct scenario-based simulations of LULC change to assess how both the composition and spatial configuration of LULC influence TP concentrations, providing insights to support proactive watershed management and spatially informed land-use planning.

This study distinguishes itself from previous TP modeling research by incorporating spatially explicit LULC patches into a CNN–DNN framework, rather than relying solely on proportional LULC metrics. Furthermore, by integrating SHAP and IG, the proposed approach provides spatially interpretable and robust insights into how LULC configurations shape TP dynamics, offering a level of explanatory capability not achievable with traditional models.

2. Materials and methods

2.1. Study region

The Republic of Korea is divided into four major river basins for water quality management: the Han, Geum, Nakdong, and Yeongsan Rivers. The study watershed, located within the Dwa climate zone (humid continental climate with hot summers and dry winters) according to the Köppen–Geiger classification, recorded an average annual precipitation of 1306.3 mm between 1991 and 2020. Approximately 54% of this precipitation occurs during the summer monsoon, leading to critically low water availability during the remainder of the year. This seasonal imbalance drives fluctuating flow regimes and short-term pulses of nutrient transport during heavy rainfall events (Kang et al., 2010).

Agriculture in the Republic of Korea is dominated by paddy rice cultivation and upland crops, with most activities concentrated between May and October. In this context, the main non-point source of TP pollution originates from fertilizer application, whereas pesticide contributions are relatively minor (Lee et al., 2023; Seo et al., 2021). Livestock farming, including swine, poultry, and cattle, also represents a significant source of nutrient loading in specific catchments (Choi et al., 2023). Point sources are densely clustered due to the country's limited land area, with industrial discharges generally contributing higher pollutant loads than domestic wastewater (Kwon and Jo, 2023). Nationwide, land use is characterized by extensive forests (about 60%), followed by agricultural land (20%), urban and industrial areas (15%), and surface water bodies (5%) (Seo et al., 2021). The spatial heterogeneity of these land-use types strongly influences nutrient dynamics across watersheds. Water quality monitoring has frequently reported exceedances of both nitrogen and phosphorus standards, especially in the Nakdong and Geum River Basins, where agricultural and urban interfaces dominate.

Since 2004, the Republic of Korea has implemented a Total Maximum Daily Load (TMDL) policy to regulate water quality. This policy sets basin-specific water quality targets and calculates allowable pollutant loads, with biochemical oxygen demand (BOD) introduced as the first regulatory indicator in 2004 and TP added in 2016 (Kim et al., 2016). Measures such as expanded wastewater treatment, pollutant removal upgrades, and low-impact development practices (e.g., permeable pavements, riparian buffers) have significantly reduced pollutant discharges, with TP loads declining by 45.5% since implementation. Nevertheless, future basin-level targets, such as the 2030 goals for the Han and Nakdong River basins, require further reductions of 27% relative to 2020 levels. Fig. 1 shows the locations of the four major river basins, the 145 TMDL watersheds, and five-year average TP concentrations at 608 monitoring stations. These basin characteristics—including monsoon-driven runoff from paddy fields, steep upland forests that accelerate sediment transport, and the clustering of industrial and urban areas along river corridors—produce spatially heterogeneous TP loading patterns. Because such fine-scale spatial configurations cannot be captured by proportion-only LULC metrics, a spatially explicit CNN is physically well aligned with the watershed processes of the Republic of Korea. The CNN framework is therefore expected to learn localized runoff pathways, near-stream clustering effects, and upstream–downstream connectivity that govern TP transport. (See Fig. 2.)

2.2. Data acquisition, processing, and preparation for ML

The dataset used in this study includes measurements from 608 river water quality monitoring stations located across 145 watersheds in the Republic of Korea, where the TMDL policy has been implemented. Monitoring was conducted at 8-day intervals between March 2018 and November 2022 and water quality data were obtained from the Water Environment Information System (http://water.nier.go.kr/). To ensure data completeness, we selected stations that consistently reported at least one valid TP measurement per month during the active monitoring period (March–November). This corresponds to a minimum of nine observations per year, resulting in at least 45 measurements over the five-year period, which were required to compute a reliable long-term mean. To characterize basin-level temporal variability prior to model training, we examined seasonal TP distributions (spring, summer, and fall) across the four major river basins (Supplementary Fig. S1). All basins exhibited markedly right-skewed distributions, with consistently higher TP in the Nakdong River and comparatively lower values in the Geum and Yeongsan Rivers. These spatio-temporal characteristics justify the use of log transformation to mitigate skewness and stabilize variance in subsequent ML modeling. Seasonal TP distributions and potential outliers were visually inspected using distribution plots (Supplementary Fig. S1), and anomalous values attributable to measurement or recording errors were excluded. Because TP exhibited a right-skewed distribution, TP concentrations were log-transformed prior to ML model training, and all predictor variables were normalized to a 0–1 scale to

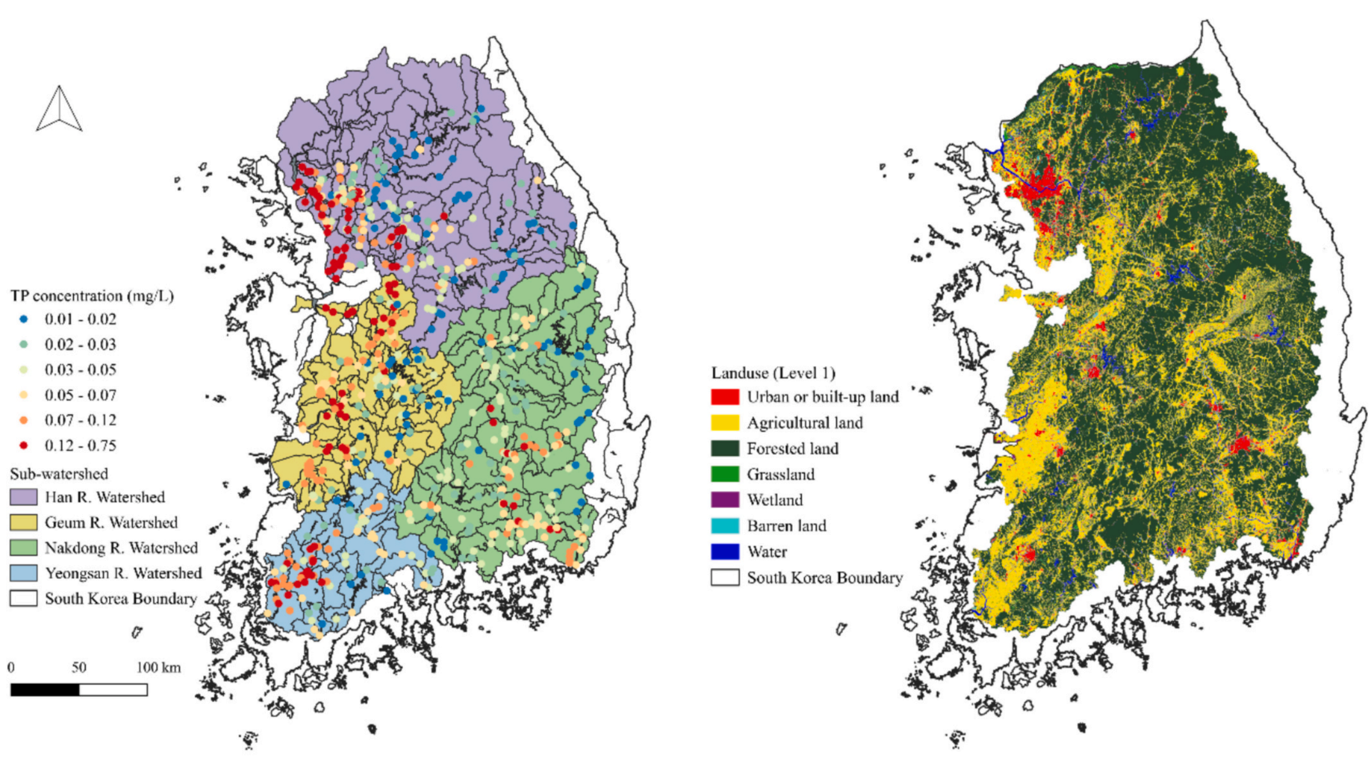

Fig. 1. (a) Spatial distribution of five-year (2018–2022) average TP concentrations at water quality monitoring sites and (b) 2022 LULC map of the study watershed used in the predictive modeling.

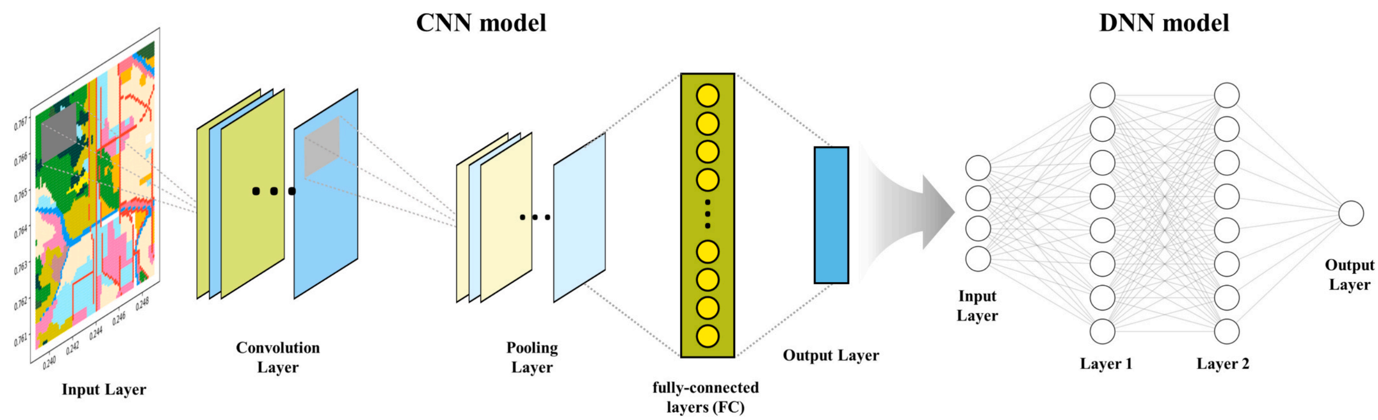

Fig. 2. Layer structure of the CNN–DNN model applied in this study.

support balanced learning. Stations with insufficient temporal coverage that did not meet the minimum data requirement were excluded from the analysis.

The input variables consisted of coordinate-based LULC, DEM, and slope, and the spatial datasets (LULC and DEM) were retrieved from the Environmental Geographic Information Service (https://egis.me.go.kr/) managed by the Ministry of Environment. The slope data were derived from the DEM using QGIS. The LULC dataset represents conditions in 2022, and both the LULC and DEM data were provided at a spatial resolution of 30 m × 30 m. Importantly, the LULC maps used in this study were not newly classified by the authors. Instead, we used the official 22-class national LULC product generated by the Ministry of Environment through a standardized workflow that integrates multispectral satellite imagery (e.g., Landsat, Sentinel), aerial photographs, field surveys, and expert manual verification. This national 22-class product is also the most recent dataset available at this thematic resolution; annual LULC maps for 2018–2021 have not yet been released. Because LULC patterns in Korean watersheds generally remain stable in the absence of large-scale construction or redevelopment, the 2022 LULC map was treated as a representative snapshot of LULC structure for the 2018–2022 TP averaging period. Nevertheless, we acknowledge that recent localized changes in rapidly urbanizing areas may not be fully captured.

The LULC data are categorized into seven broad LULC classes—urban or built-up, agricultural, forested, grassland, wetland, barren land, and water—which are further subdivided into 22 detailed LULC classes, as presented in Table 1 Several subcategories of urban or built-up LULC types require clarification. These include communication facilities (cultural, sports, and leisure facilities), transportation infrastructure (airports, harbors, railways, and roads), and public utilities (environmental infrastructure: wastewater treatment plants, educational institutions: schools and universities, and administrative facilities: government offices). Additionally, the nationwide percentage coverage of each detailed LULC class was calculated to provide a comprehensive understanding of their spatial distribution across the country.

To generate spatial input data for modeling, a 30 m × 30 m base grid covering a 2400 m × 2400 m area centered on each monitoring station was used. This spatial extent was selected based on previous studies that identified strong correlations between LULC and TP concentrations within 1000 to 3000 m from sampling sites (Wang et al., 2023; Zhang et al., 2021). The resolution was chosen to match that of commonly used satellite products, where LULC datasets are derived from sources such as Landsat imagery. The grid cells were categorized into 22 detailed LULC classes, allowing for a more accurate representation of landscape composition and configuration. The nationwide percentage coverage of each LULC class was also calculated to illustrate their distribution across the study region.

In this study, the model inputs capture major structural watershed processes. Detailed LULC classes represent the spatial distribution of point and non-point sources and serve as proxies for anthropogenic intensity and hydrologic connectivity. DEM characterizes topographic gradients that govern flow direction and erosion potential, and slope reflects the energy of overland flow and its capacity to mobilize phosphorus-rich sediments. To align the predictors with the response variable, we paired five-year averaged TP concentrations (2018–2022) with the static 2022 LULC, DEM, and slope layers, thereby emphasizing

Table 1
LULC classification system by the Ministry of Environment, the Republic of Korea.

Broad LULC	Detailed LULC (used in this study)	%	Color	Broad LULC	Detailed LULC (used in this study)	%	Color
Urban or built-up land	Residential	3.2		Forested land	Deciduous forest land	17.3	
	Industrial	0.6			Coniferous forest land	28.2	
	Commercial	0.3			Mixed forest land	17.5	
	Communication	0.0		Grassland	Natural grassland	0.5	
	Transportation	1.0			Artificial grassland	0.1	
	Public utilities	0.5		Wetland	Inland wetland	0.5	
Agricultural land	Paddy fields	14.7			Coastal wetland	0.2	
	Non-irrigated land	8.4		Barren land	Natural barren land	0.1	
	Protected cultivation	0.3			Artificial barren land	1.3	
	Orchard	1.2		Water	Inland water	2.3	
	Other cropland	0.3			Seawater	0.2	

long-term structural controls rather than event-scale variability. This design is consistent with our objective of quantifying how persistent watershed structure shapes chronic TP levels, rather than reproducing individual storm responses. Although these variables summarize long-term watershed structure, several short-term hydrologic and biogeochemical processes—such as rainfall intensity, storm-driven runoff pulses, soil moisture fluctuations, and groundwater–surface water exchanges—are not explicitly captured. In rapidly urbanizing or redeveloping catchments, this assumption may introduce bias if recent land conversions are not reflected in the 2022 LULC map. Meteorological inputs were not used because the five-year averaged TP observations inherently suppress short-term variability, making spatial LULC and topographic controls the dominant predictors in our modeling framework.

Based on these data, a coordinate-based CNN model was developed to predict the five-year average TP concentrations at the monitoring stations. CNNs are particularly well-suited for learning from spatially structured inputs, such as gridded LULC maps, because they preserve spatial patterns and extract hierarchical features that are relevant to environmental processes. After training the CNN model, SHAP values were applied to interpret the model outputs. This approach enabled the quantification of nonlinear and complex relationships between the spatial distribution of LULC and the predicted TP concentrations, providing insights into the relative influence of LULC classes and their spatial configuration on water quality.

2.3. Machine learning framework and explainable AI

2.3.1. Convolutional neural network

CNNs are widely used deep learning models capable of extracting hierarchical spatial features from two-dimensional data structures (Sethy et al., 2020). In this study, CNNs were applied to LULC maps to capture spatial dependencies that cannot be represented by proportion-only predictors. The architecture consisted of convolutional and pooling layers, followed by fully connected layers for prediction. We used ReLU as the activation function and max pooling to reduce dimensionality and mitigate overfitting. Model parameters, including kernel size (3×3), learning rate, and batch size, were optimized during training. Detailed algorithmic workflow, mathematical formulations, and layer-specific descriptions are provided in the Supplementary Material 2.

2.3.2. Explainable AI with SHapley additive exPlanations and integrated gradients

To enhance interpretability of the CNN–DNN model, we employed explainable AI approaches, specifically SHAP and IG. SHAP was used to quantify and visualize the relative influence of detailed LULC classes on TP predictions at each monitoring site, providing spatially localized insights into land-use effects on water quality. Grounded in cooperative game theory, SHAP offers transparent attribution that extends beyond conventional feature importance metrics. To complement SHAP, IG was applied as a gradient-based attribution method, assessing the robustness of explanations by integrating gradients from predefined baselines to the actual inputs. By considering three different baselines (patch-wide, ring-wise, and cluster-based), IG validated the consistency of attributions and reduced reliance on a single interpretability framework. Together, SHAP and IG establish a reliable explainable AI framework that links spatial land use patterns to nutrient dynamics, allowing both researchers and policymakers to better understand not only what the model predicts, but also why those predictions are made.

2.4. Development of DNN-CNN framework

We developed a CNN–DNN model to predict in-stream TP concentrations in the Republic of Korea. All ML simulations in this study were performed in a Python 3.9.2 environment, using TensorFlow 2.8.0 and Keras 2.8.0 for CNN and DNN implementation, respectively.

The architecture of the ML model (CNN–DNN) was designed as follows: (1) The CNN component was provided with spatially explicit, gridded inputs consisting of 22 detailed LULC classes, elevation (DEM), and slope; (2) convolutional operations were conducted using a 3×3 kernel, followed by average pooling with a 2×2 kernel, and this process was repeated eight times; (3) the extracted feature maps were flattened and passed to the DNN component; and (4) TP concentrations were predicted using a DNN consisting of two fully connected layers and one dropout layer. To design the CNN model, linear operators (such as convolution) or nonlinear activation functions in the convolutional layers were used to isolate and identify various features of the data for image analysis (Khosravi et al., 2023). Previously, sigmoid and tanh functions were commonly used as nonlinearity functions; however, the Rectified Linear Unit (ReLU), defined as $f(x) = \max(0, x)$ has become the standard choice in recent applications because it has simpler definitions for both the function and gradient (Saad et al., 2017). In this study, ReLU was used as the activation function in all convolutional layers.

The changes in spatial dimensions across convolutional and pooling layers are summarized in Table S1 (Supplementary Material 3). Setting the padding during the convolution process prevents the output size of the layer from shrinking with depth (O'Shea and Nash, 2015). This study employed zero-padding, and the equation is as follows (Saad et al., 2017):

$$O = 1 + \frac{N + 2P - F}{S}$$

where P is the number of zero-padding layers, N is the image dimension ($N \times N$ dimensions), F is the kernel size ($F \times F$ kernel), and O is the output size. In the pooling process, a 2×2 kernel size was used, and the spatial size of the pooling layer was halved.

Hyperparameter tuning was performed to optimize the performance of the CNN–DNN model. A grid search was conducted to determine the optimal hyperparameter values for each model (Bergstra and Bengio, 2012). The ranges and intervals of the tested hyperparameters are provided in Table S2 (Supplementary Material 3). After the grid search, a set of hyperparameter values with the minimum mean square error (MSE) and mean absolute error (MAE) was selected for each CNN and DNN model, and the final hyperparameters used in the analysis are summarized in Table S3. The model was trained using a total of 608 samples, which were divided into a fixed ratio of 480 samples for training and 128 samples for validation. During training, a single validation set was used to evaluate the generalization performance of the model. The input data were structured as three-dimensional arrays, with the training dataset having a shape of (480, 6561, 26) and the validation dataset having a shape of (128, 6561, 26), corresponding to the number of samples, spatial locations, and input variables, respectively.

Because the train–validation split was random rather than spatially constrained, some degree of spatial autocorrelation may remain between samples. This can lead to slightly optimistic estimates of predictive performance, particularly when using spatially structured predictors such as LULC and topography. However, the primary objective of this study is to evaluate the ability of the CNN–DNN framework to learn and interpret spatial LULC patterns within a nationwide monitoring network, rather than to assess strict spatial transferability across basins.

3. Results

3.1. Model performance and configuration

The prediction performance metrics of the optimized CNN model for estimating TP concentrations are summarized in Table 2 and Fig. 3. In addition to R and R^2, model bias was evaluated using percent bias (PBIAS) for the training dataset, which showed a negative value

Table 2
Prediction performance of the optimized CNN–DNN model.

Dataset	R	R^2	MSE	MAE
Train	0.726	0.521	0.020	0.112
Test	0.711	0.497	0.021	0.119

(−14.3%), indicating a tendency toward underprediction. The results demonstrate that the CNN model achieved a reliable level of accuracy in estimating TP concentrations, comparable to findings reported in previous studies. For instance, Shen et al. (2020) performed regression analysis using an RF model on seasonal total nitrogen (TN) and TP concentrations collected from 62,495 stations across the United States, incorporating 47 environmental variables. They reported the highest prediction accuracy for TP in spring ($R = 0.81$), followed by summer ($R = 0.73$), fall ($R = 0.64$), and winter ($R = 0.56$). Similarly, Dolph et al. (2023) utilized nearly 300 geospatial variables—including soil type, land cover, land use, topography, nutrient inputs, and climate—at a 100 m grid scale for the Upper Mississippi River Basin, USA. By applying an RF model to a publicly available soil phosphorus dataset, they achieved a prediction accuracy of $R^2 = 0.58$ for total soil phosphorus. These prior findings support the robustness of our CNN-based modeling approach in capturing the spatial and environmental complexities associated with TP concentration prediction.

To further evaluate the added value of incorporating spatial configuration, we conducted additional analyses using conventional approaches. A multiple linear regression (MLR) model was developed using only the proportional values of LULC, slope, and DEM at each monitoring site, without considering any spatial arrangement. A non-spatial DNN was also applied with the same set of predictors. As shown in Supplementary Fig. S3, both MLR and DNN exhibited substantially lower predictive performance than the CNN–DNN framework. Although the CNN–DNN achieved only moderate improvements in accuracy compared to these conventional models, its distinctive strength lies in explicitly capturing local spatial configuration. This capability enables spatially explicit scenario analyses—such as clustered urban expansion or forest-to-agricultural conversion—that proportion-only regression models cannot address. To further diagnose model behavior beyond aggregate accuracy metrics, basin-wise residual boxplots and observed–residual scatter plots were analyzed for both the training and test datasets and are presented in the Supplementary Material 5.

3.2. LULC importance analysis

To enable comparison across diverse landscape contexts, we analyzed six representative monitoring stations encompassing urban, agricultural, and forested watersheds (two sites for each LULC type). These stations were selected based on having the highest proportion of a single dominant LULC type within a 2400 m × 2400 m grid centered on each monitoring location (Supplementary Table S4). The input variables used in the CNN model at each site included DEM, slope, and the proportions of 22 LULC classes. Analysis of these six representative stations revealed distinct yet interconnected effects of dominant LULC types on TP concentrations (Figs. 4–6). In the following subsections, we present the SHAP and IG results for each site, highlighting how variations in urban, agricultural, and forested landscapes shape TP dynamics.

In urban watersheds (Points #188 and #435), both sites were characterized by flat terrain (average slope < 1.0°) and over 90% urban land cover. At Point #188, residential land dominated (43.8%) and exhibited moderate negative SHAP contributions, particularly in the northern areas, likely reflecting the mitigating role of sewage infrastructure in this metropolitan satellite city. Public utility areas, despite their small extent, showed localized positive contributions, while commercial and industrial areas exerted weak to moderate influence. Interestingly, barren land as a source of non-point pollution displayed negative SHAP values. Orthophotos confirmed that these areas functioned as managed green spaces and permeable recreational fields, thereby promoting infiltration and reducing runoff (Supplementary Fig. S5). Point #435, by contrast, was more commercialized (22.5% commercial, 45.7% residential, 13.5% transportation). Here, SHAP values for commercial land were stronger and formed connected spatial bands along stream corridors, suggesting intensified effects due to clustering and impervious surfaces. Residential zones remained largely negative but showed some positive edges influenced by adjacent land use effects. Proximity to hydrological features and spatial connectivity clearly amplified commercial impacts.

Agricultural watersheds (Points #453 and #555) consistently exhibited strong positive contributions from paddy fields. At Point #453, paddy fields covered 62.1% of the area and produced uniformly high SHAP values across the watershed, unaffected by distance to the monitoring point. Water bodies acted as negative contributors, while residential land showed weak negative values. Commercial zones, though comprising only 0.3%, had disproportionately high SHAP values near streams, likely exacerbated by lower sewage coverage (78%) compared to the national average (94.5%). Forested land generally had negative SHAP values, though coniferous stands showed more variability. At Point #555, paddy fields were even more dominant (75.2%), but SHAP intensity declined in the southern watershed, suggesting that slope, distance, or hydrological separation moderated the effect. This

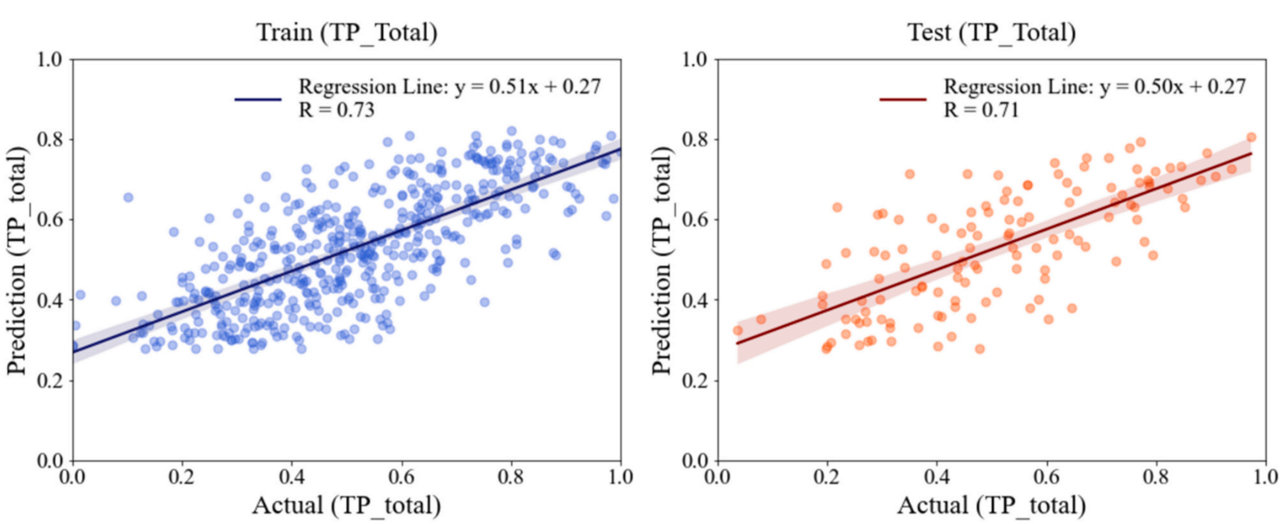

Fig. 3. Relationship between actual and predicted TP concentrations in the training and test sets.

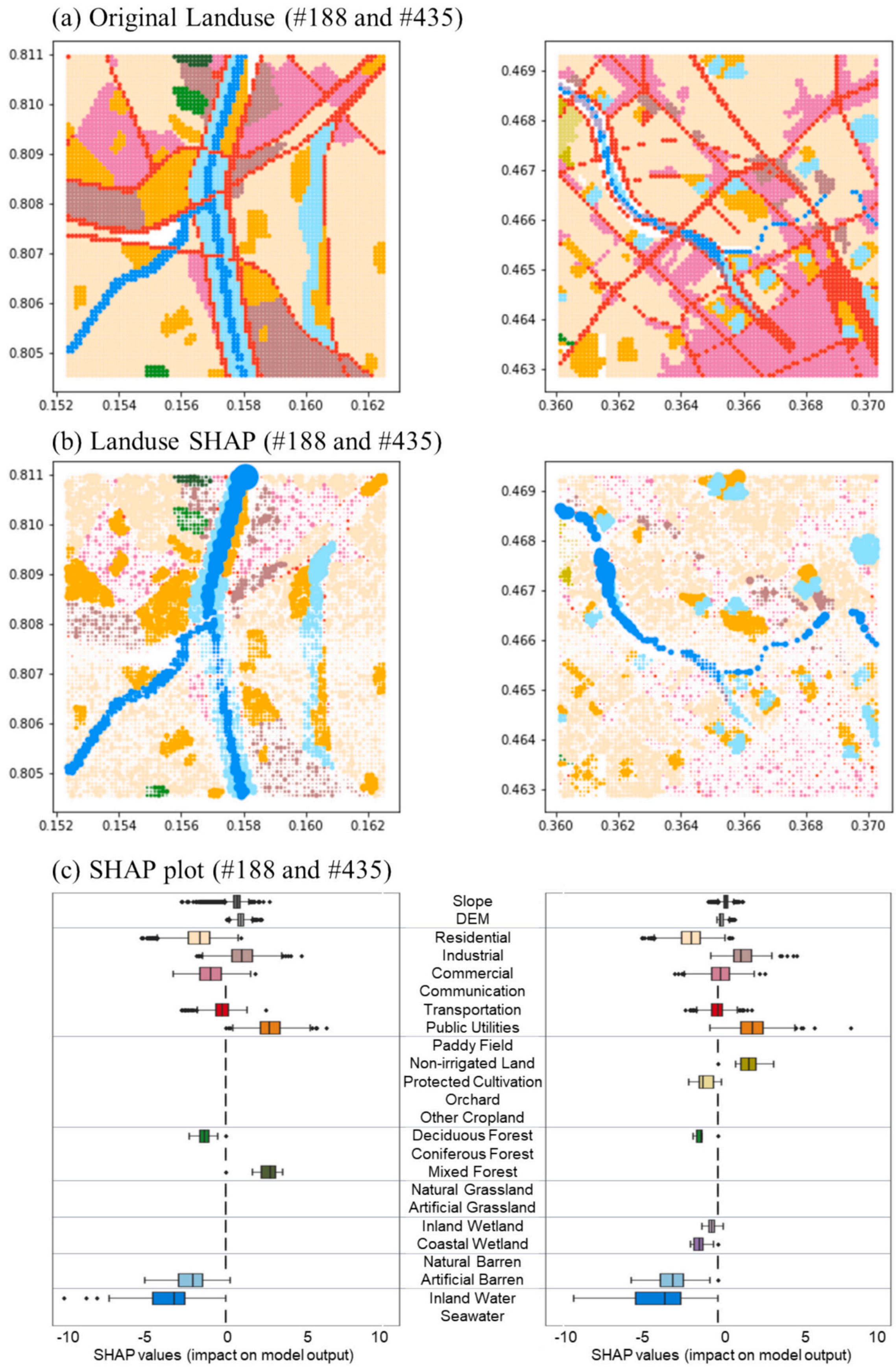

Fig. 4. Results for the two representative urban monitoring stations (#188 and #435), selected for having the highest urban land proportions (>90%) within a 2400 m × 2400 m grid around each site. For both stations, panels show (a) the spatial distribution of detailed LULC classes, (b) spatial SHAP maps indicating location-dependent TP influence, and (c) SHAP dependence plots.

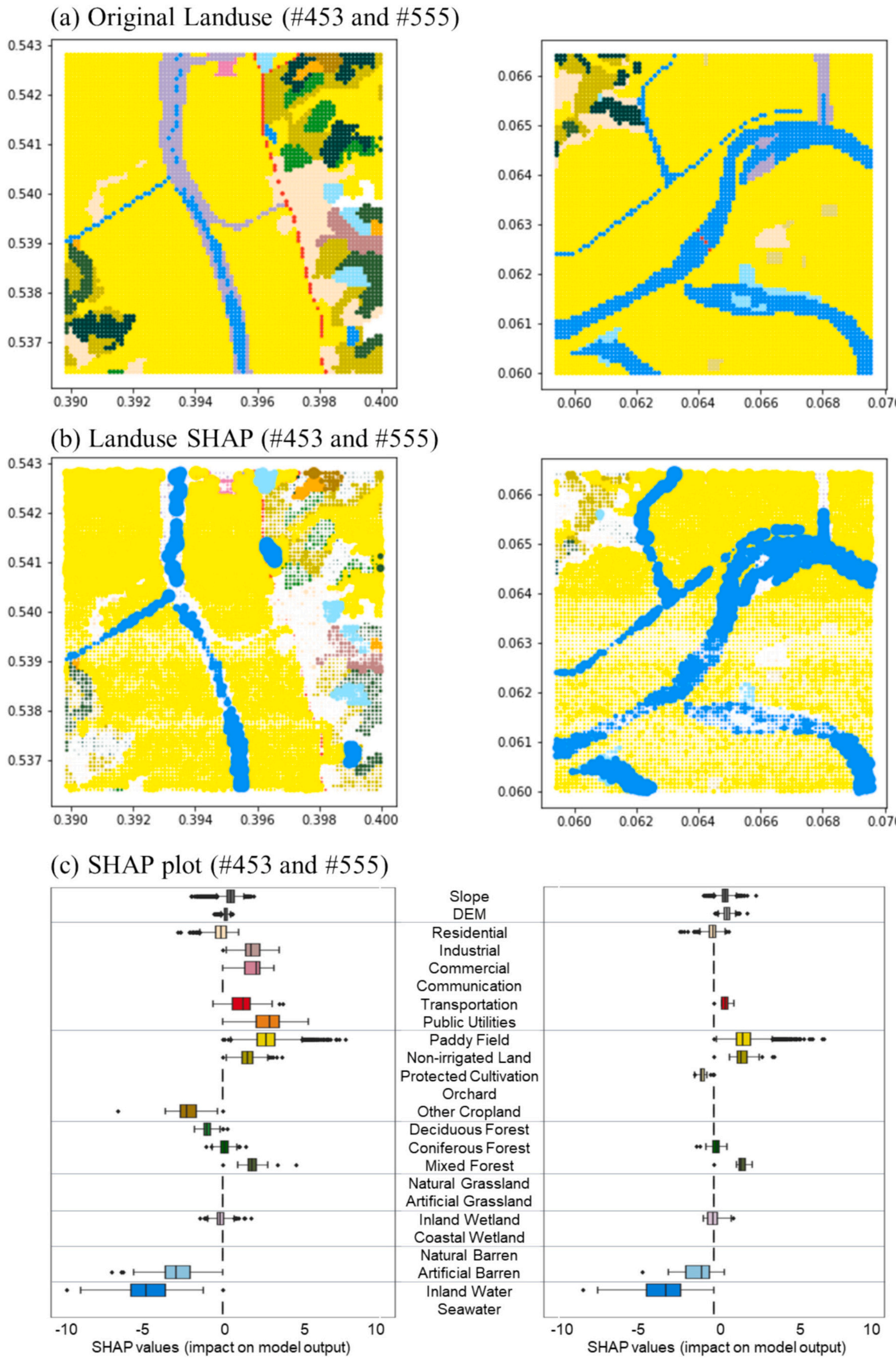

Fig. 5. Results for the two representative agricultural monitoring stations (#453 and #555), selected for having the highest proportions of agricultural land—particularly paddy fields—within a 2400 m × 2400 m grid around each site. For both stations, panels show (a) the spatial distribution of detailed LULC classes, (b) spatial SHAP maps indicating location-dependent TP influence, and (c) SHAP dependence plots.

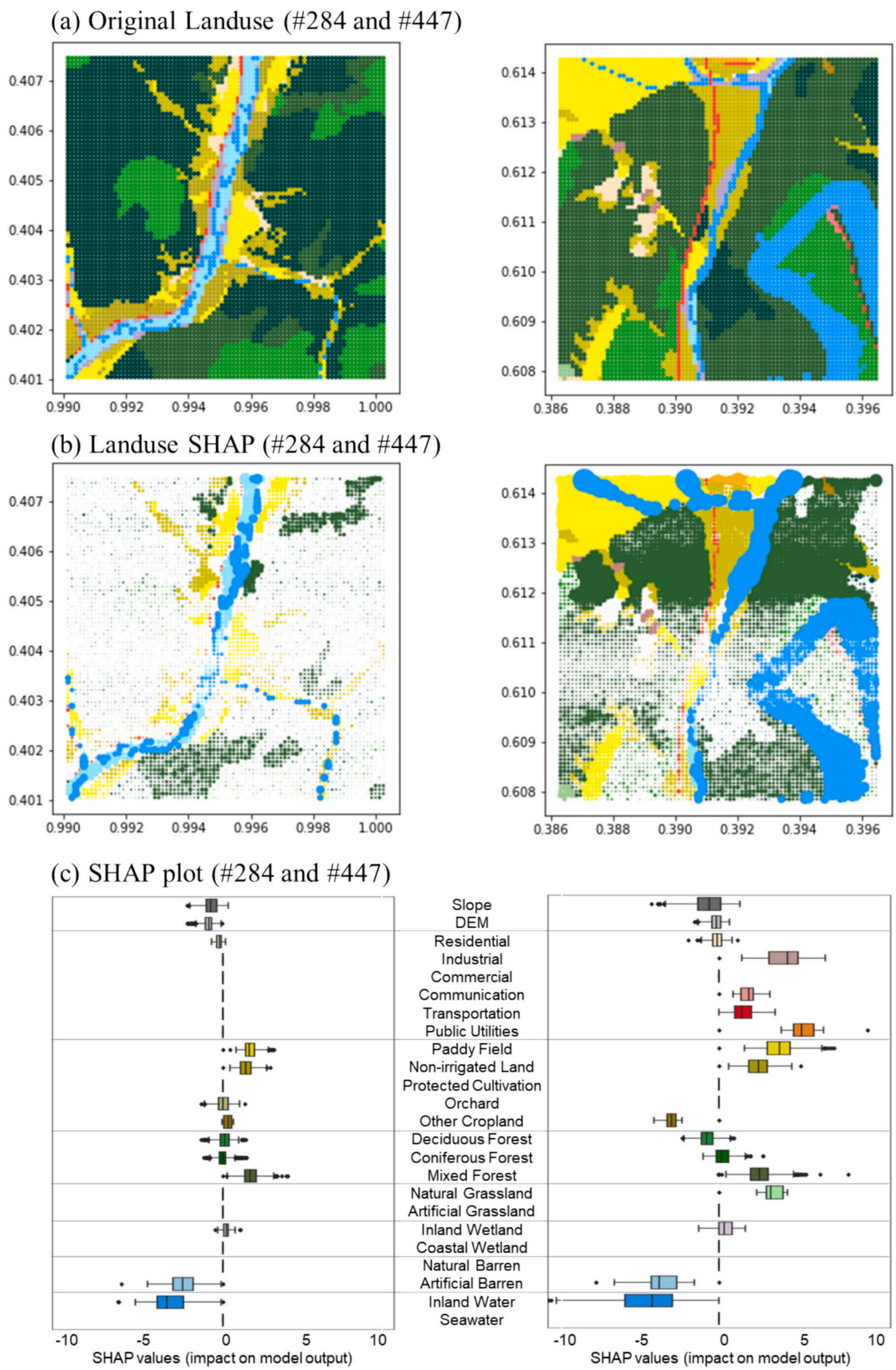

Fig. 6. Results for the two representative forested monitoring stations (#284 and #447), selected for having the highest proportions of forest land within a 2400 m × 2400 m grid surrounding each site. For both stations, panels present (a) the spatial distribution of detailed LULC classes, (b) spatial SHAP maps showing location-dependent TP influence, and (c) SHAP dependence plots.

contrasted with the uniform patterns at Point #453, underscoring the role of terrain and stream connectivity in shaping nutrient contributions.

Forested watersheds (Points #284 and #447) showed lower TP contributions overall, but notable contrasts emerged between sites. Point #284 was dominated by coniferous forest (51.3% of total area, 70% of forest cover) and exhibited generally negative or neutral SHAP values, with only upstream agricultural patches producing localized positive responses. Point #447, in contrast, was dominated by mixed forest (44.1% of total area, 70.1% of forest cover) and displayed greater variability. High positive SHAP values appeared in the upper watershed, where forest, public utility, and industrial parcels clustered near streams, while deciduous stands in the south exhibited negative contributions. Even small industrial (0.4%) and public utility (0.1%) parcels showed disproportionately high SHAP values, again linked to stream proximity.

Across all six representative sites, several consistent patterns emerged. Residential zones tended to buffer TP concentrations when supported by sewage infrastructure, whereas commercial land, especially when clustered near streams, intensified contributions. Paddy fields were the strongest positive drivers in agricultural watersheds, with their influence modulated by hydrological connectivity and slope. Forests generally mitigated TP, though coniferous and mixed stands occasionally showed positive contributions under specific spatial or compositional conditions. Public utility and industrial lands, although small in area, repeatedly appeared as localized hotspots when hydrologically connected. These findings confirm that TP dynamics in freshwater systems are shaped not only by LULC type and proportion but also by spatial arrangement, terrain, and infrastructural context.

3.3. Integrated gradients analysis

We applied IG to quantify how LULC types contributed to the CNN model's predictions of TP concentrations across the study domain. To evaluate attribution sensitivity under different spatial assumptions, we employed three distinct IG baselines—patch-wide, ring-wise, and cluster-based. The patch-wide baseline represents a spatially homogeneous landscape with evenly distributed land-use types; the ring-wise baseline reflects a distance-based gradient that distinguishes near-stream from far-stream influences; and the cluster-based baseline simulates spatially aggregated land-use patterns. Fig. 7 illustrates IG attribution magnitudes for each LULC class across these baselines, with the yellow bar representing the averaged values used as domain-wide importance scores.

The results show that paddy fields consistently produced the highest attribution, confirming their dominant role in driving TP increases. Forest categories followed in the order of mixed > coniferous > deciduous forest, highlighting the importance of forest composition. Inland water also ranked highly, reflecting strong hydrologic connectivity. Artificial barren and residential areas contributed moderately, while urban–industrial uses (public utilities, transportation, industrial, commercial) exhibited relatively low average attributions but can play critical roles when clustered near streams. Other cropland, natural land, coastal wetlands, and seawater showed negligible importance.

Cross-baseline comparison revealed consistent rankings and magnitudes for the leading classes, indicating low baseline sensitivity. Although small variations appeared among some forest subclasses and water-related categories, these shifts were minor and did not alter the dominant patterns. The general agreement across the three baselines—which emphasize homogeneous, distance-gradient, and clustered spatial structures—demonstrates that the major attribution patterns are robust to baseline selection. The averaged values further stabilized these variations, enhancing the robustness and generalizability of the results.

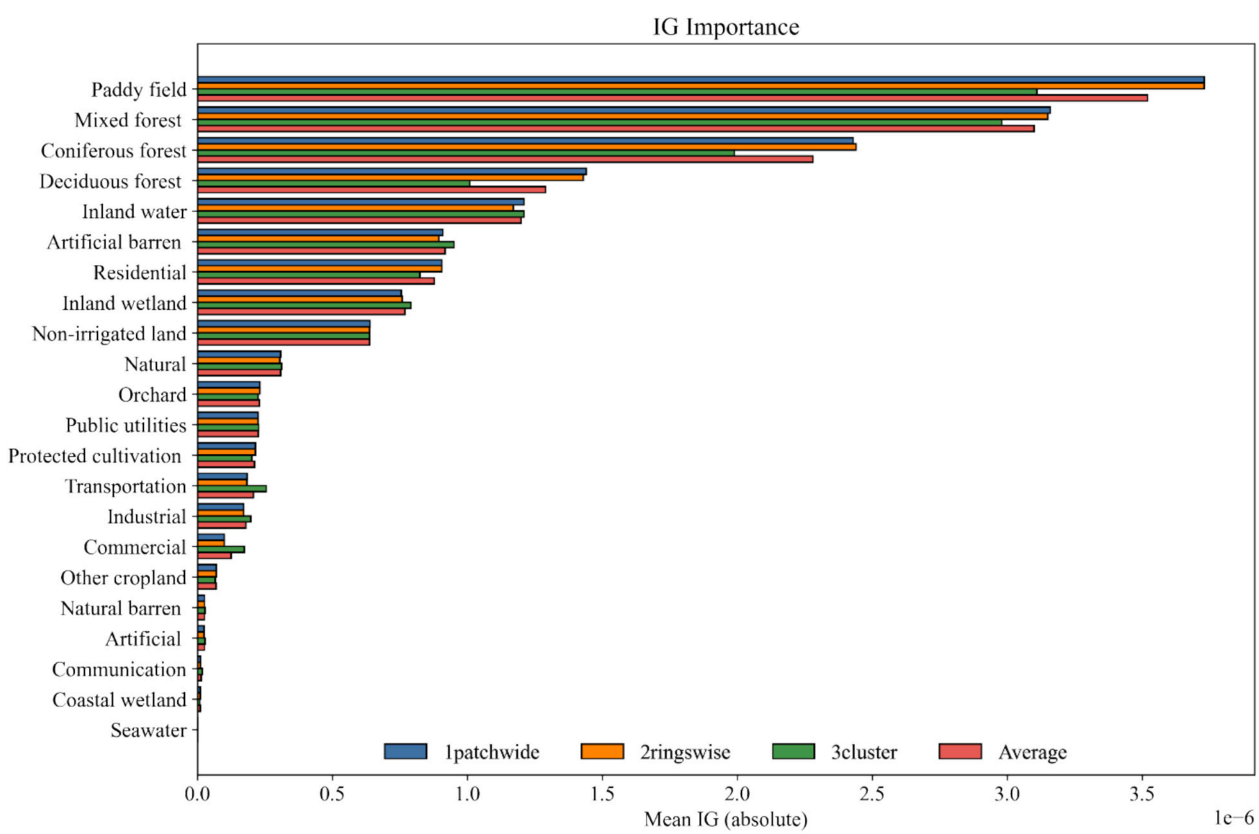

Fig. 7. IG results for all monitoring sites, showing the relative importance of land use variables across three baselines (patch-wide, ring-wise, and cluster-based) and their average.

3.4. Scenario-based analysis of LULC change effects on TP concentration

In this study, land use spatial patterns were systematically modified to evaluate the impact of land use distribution on water quality. Approximately 30% of the forest area was converted into residential, industrial, or agricultural land and reallocated into five spatial clusters: upper-left, upper-right, lower-left, lower-right, and center (UL, UR, LL, LR, and C). These clustered scenarios assumed that new developments typically form contiguous zones rather than being spatially dispersed. Proximity-based selection algorithms in Python were applied to select adjacent grid cells around cluster centers, minimizing spatial dispersion. The modified ASC-format raster files were then used as inputs to the CNN-based TP prediction model.

To clarify the scenario assumptions, we emphasize that only the LULC cells were modified, while all other covariates (DEM and slope) were held constant throughout the analysis. This reflects the assumption that short-term land-use change alters LULC distribution without modifying underlying topography. The CNN–DNN model was not retrained for each scenario; instead, modified LULC rasters were fed directly into the already-trained model to isolate the effect of spatial configuration. The five spatial clusters (UL, UR, LL, LR, and C) were selected not merely for visualization but to mimic plausible development geometries observed in real-world watersheds, such as central infilling, peripheral expansion, and quadrant-based clustering. A conceptual schematic (Fig. 8) was added to illustrate the scenario-generation workflow, including baseline data preparation, cluster selection, LULC pixel replacement, and prediction using the pretrained CNN–DNN model.

At the selected site, the five-year average TP concentration was 0.033 mg/L, exactly matching the model prediction, thereby validating the model's reliability. The LULC composition of the site comprised 1.1% residential, 0.8% other urban, 21.7% agricultural, and 73.5% forest, characterizing it as a forest-dominant watershed. Fig. 9(a) shows the baseline condition, Figs. 9(b)–(f) depict forest-to-residential conversions, Figs. 9(g)–(k) show forest-to-industrial conversions, and Figs. 9 (l)–(p) represent forest-to-agricultural transformations.

Fig. 10 presents the percentage change in TP concentrations relative to the baseline for each cluster. In the forest-to-residential scenarios, most clusters exhibited reductions of 1.1–4.4%, with an average decrease of 2.0%. The only increase was observed in the lower-right cluster (Fig. 9e), which rose by 1.9%. In the forest-to-industrial scenarios, TP concentrations increased in all cases, with an average increase of 9.5%. The central cluster (Fig. 9k) recorded the highest increase at 11.0%, while the lower-left cluster (Fig. 9g) showed the lowest increase at 7.9%. In the forest-to-agricultural conversions, TP concentrations rose to an average of 0.041 mg/L, representing a 23.2% increase relative to baseline. Differences across spatial clusters were minimal in this case, though the central cluster (Fig. 9p) showed slightly higher values.

4. Discussion

4.1. LULC importance analysis

This study confirmed that TP concentrations are influenced not only by LULC proportions but also by their spatial distribution and configuration within watersheds. By integrating SHAP interpretation into a CNN framework, we quantified and visualized the nonlinear and spatially dependent effects of LULC on water quality, providing insights that go beyond proportion-only regression models.

SHAP analysis across urban, agricultural, and forested watersheds demonstrated that TP dynamics are strongly shaped by LULC spatial arrangement, hydrologic connectivity, and topographic features. In urban watersheds, residential zones consistently showed negative SHAP contributions, reflecting the buffering role of wastewater treatment, while commercial and public utility areas exerted stronger positive impacts when clustered along stream corridors. Barren land, often functioning as permeable recreational or green spaces, contributed to TP reduction. In agricultural watersheds, paddy fields were the dominant positive contributors, but their influence varied depending on slope and distance from streams. Point #453 exhibited uniformly high SHAP values across the watershed, whereas Point #555 showed reduced contributions in its southern region. In forested watersheds, coniferous forests generally acted as buffers, but mixed forests and adjacent industrial or public utility parcels near headwaters increased TP levels, suggesting that certain LULC and vegetation combinations can turn forests from sinks into sources.

Compared with MLR models, the CNN–DNN framework consistently achieved higher explanatory power, highlighting the value of incorporating spatial patterns and detailed LULC classifications. This spatially explicit representation marks a methodological improvement over previous TP studies that relied solely on proportional LULC metrics, as it allows the model to capture patch geometry, adjacency, and near-stream clustering effects that proportion-based models inherently overlook.

4.2. Implications of IG-based analysis

The IG analysis provided robust support for generalizing SHAP-based findings to the entire study domain. By incorporating three baselines—patch-wide, ring-wise, and cluster-based—IG confirmed the stability and reliability of the most influential LULC types. Paddy fields consistently emerged as the strongest contributors, forests generally mitigated TP but varied with composition, and inland waters and wetlands acted as critical modulators. Urban–industrial uses, though small in areal proportion, exerted disproportionately high impacts when connected to streams. These results emphasize that effective water quality management requires attention not only to LULC proportions but also to spatial configuration, hydrologic connectivity, and infrastructure context. By combining spatially explicit CNN-derived features with two complementary interpretability methods (SHAP and IG), this study provides one of the first integrated frameworks capable of revealing consistent, location-dependent phosphorus drivers across an entire

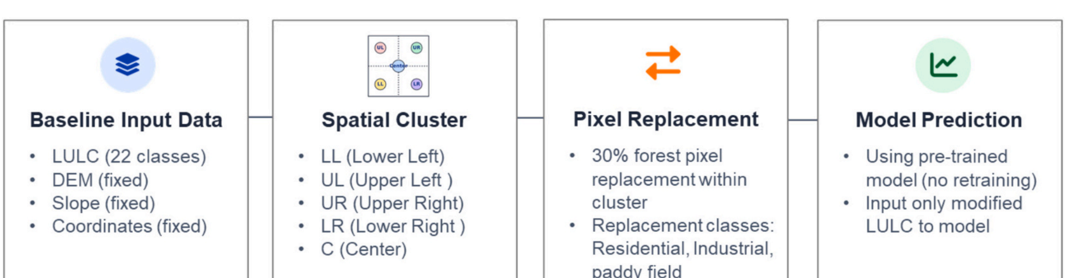

Fig. 8. Framework for scenario-based LULC modification used in this study.

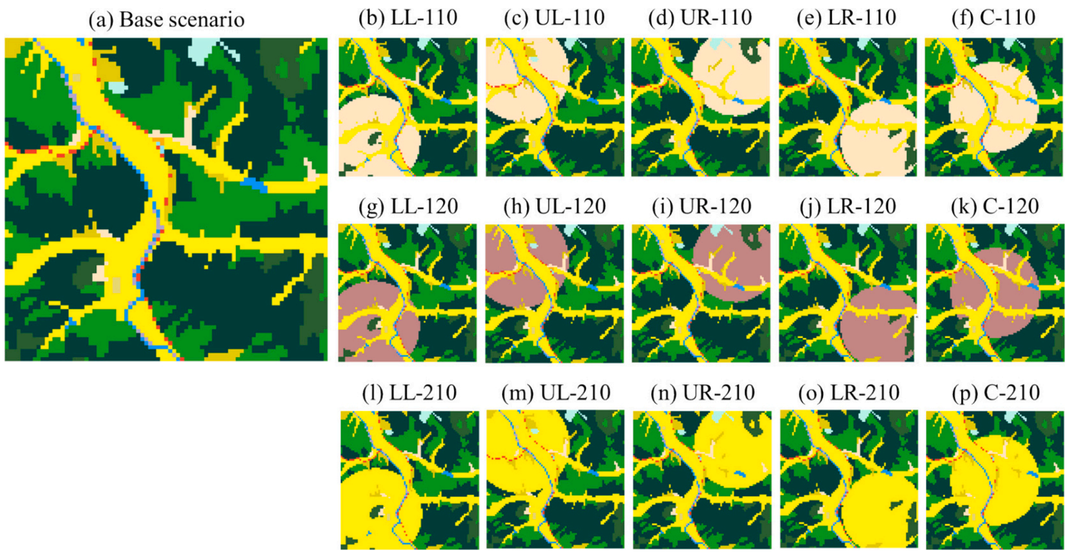

Fig. 9. Land use change scenarios used for TP concentration analysis. (a) Baseline land use; (b–f) forest-to-residential conversions; (g–k) forest-to-industrial conversions; (l–p) forest-to-paddy field conversions.

watershed domain.

4.3. Implications of scenario-based LULC change

Scenario analysis demonstrated that TP responses are determined not only by the areal extent of LULC conversions but also by their spatial arrangement. Forest-to-residential conversions generally reduced TP concentrations slightly, reflecting lower non-point source loads and sewage infrastructure buffering, while forest-to-industrial conversions consistently raised TP levels, with the central cluster showing the highest increase. Forest-to-agricultural conversions, especially to paddy fields, produced the largest increases, averaging 23.2%. These outcomes indicate that slope, distance to streams, and surrounding land uses mediate the magnitude of change. The CNN-based framework therefore provides a practical tool for evaluating realistic development scenarios and supports proactive watershed management and phosphorus reduction strategies. The ability to evaluate both the magnitude and spatial configuration of LULC conversions distinguishes this approach from previous scenario frameworks that relied only on areal changes, demonstrating the added value of spatially explicit LULC patches in predictive modeling.

4.4. Limitations and future directions

The model was trained on five-year average TP data (2018–2022) and static LULC maps, which limits its applicability to time-varying or climate-driven changes and may underrepresent the effects of recent urbanization or land conversion in rapidly changing areas. In addition, only the most recent 22-class national LULC product is currently available at the spatial and thematic resolution required for this study, so fully dynamic, year-by-year LULC maps for 2018–2022 could not be incorporated. As a result, the model implicitly assumes that the 2022 LULC pattern is representative of the multi-year averaging period. Future work could relax this assumption by combining multi-temporal or satellite-derived LULC products with sequence-based or hybrid physics–ML models to jointly capture spatial structure and temporal evolution of watershed conditions.

Point-source discharges, such as effluent volumes from wastewater treatment plants, were not explicitly included, which may have contributed to an underrepresentation of anthropogenic mitigation effects in urban areas. Although the negative SHAP values observed for residential zones likely reflect the indirect influence of wastewater treatment infrastructure, explicitly incorporating WWTP-related variables—such as discharge volume, treatment capacity, and spatial proximity—would improve interpretability and predictive realism. Moreover, the model does not represent stormwater management infrastructure, including dry wells, rain gardens, permeable pavements, and conventional open-channel drainage systems, all of which substantially influence phosphorus retention and transport in urban catchments. Because these engineered features are not distinguished within the national LULC dataset, the CNN framework captures the spatial pattern of urban land cover but cannot differentiate among functionally distinct drainage systems. Incorporating municipal-scale stormwater infrastructure maps would meaningfully enhance urban hydrologic representation in future studies.

Finally, several short-term hydrologic and biogeochemical drivers were not represented. Seasonal variability, event-scale rainfall intensity, storm-generated runoff pulses, soil moisture dynamics, and groundwater–surface water exchanges are absent from the current framework. Because of these limitations, aggregate error metrics for the test dataset were not emphasized in the main text; instead, generalization behavior was evaluated using residual pattern analyses to avoid overinterpreting event-scale variability that lies beyond the scope of long-term watershed assessment. The 30 m resolution of the input spatial data also constrains the representation of sub-grid heterogeneity in mixed-use areas. Future studies should integrate time-series water quality observations, dynamic LULC datasets, hydrologic states, and wastewater infrastructure information, and expand the analysis to additional pollutants such as TN and BOD. Applying this framework across diverse regions would further enhance its generalizability and policy relevance.

4.5. Applicability and transferability of the framework

The applicability of the proposed CNN–SHAP–IG framework extends beyond the Korean basins, provided that several key conditions are met. The ability of the model to reproduce long-term TP patterns using only spatially distributed LULC and topographic variables suggests that structural watershed characteristics dominate phosphorus dynamics, while the relative contribution of point sources is limited at the long-term average scale. Accordingly, the proposed framework is particularly well suited to watersheds where point-source discharges are already regulated and diffuse nutrient inputs associated with land-use configuration play a primary role. First, sufficiently detailed and high-resolution LULC data are required, ideally at 30 m resolution or finer,

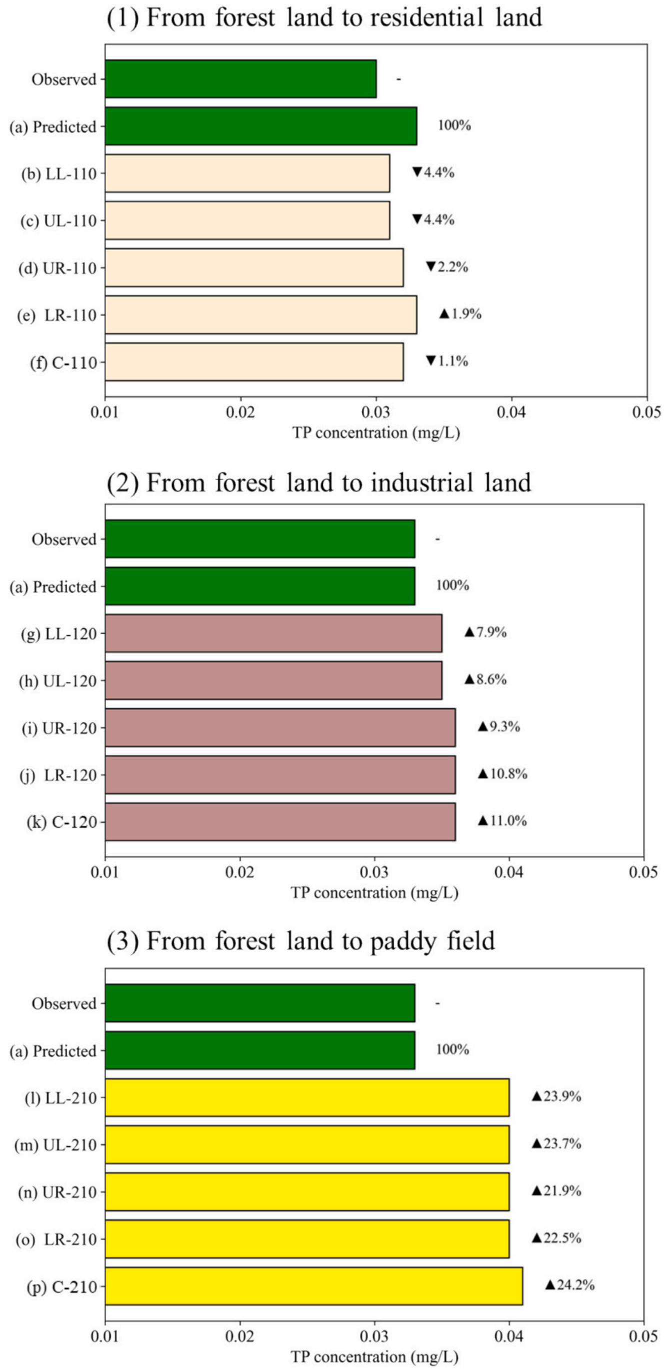

Fig. 10. Changes in TP concentration under land use scenarios: (1) forest-to-residential, (2) forest-to-industrial, and (3) forest-to-paddy field. Bars represent the percentage change from baseline for each spatial cluster.

to preserve patch geometry and spatial configuration. Second, an adequate density of water quality monitoring stations is necessary to capture watershed heterogeneity and support stable model training. Third, the framework is best suited to regions where nutrient dynamics are strongly shaped by land-use arrangement, hydrologic connectivity, and event-driven runoff—conditions commonly found in monsoon-influenced or mixed urban–agricultural systems. Basins dominated by snowmelt processes, arid-climate hydrology, or groundwater-driven nutrient transport may require additional covariates or region-specific model retraining. Finally, transferring the framework to new regions would benefit from incremental calibration using local LULC classifications and management practices. With these conditions satisfied, the model provides a flexible and spatially explicit tool for nutrient prediction and land-use planning in diverse watershed environments.

5. Conclusions

This study integrated SHAP and IG into a CNN–DNN prediction framework to quantify and visualize the impacts of LULC on TP concentrations in freshwater systems. The results confirmed that TP dynamics are shaped not only by the proportion of LULC types but also by their spatial configuration, hydrologic connectivity, and topographic context. SHAP analysis provided detailed, site-specific insights, while IG analysis, supported by three baseline settings, ensured robustness and generalizability at the watershed scale. Across LULC types, paddy fields consistently emerged as the strongest positive drivers, forests generally functioned as buffers with variability by composition, and urban–industrial parcels, though small in area, exerted disproportionate impacts when connected to streams.

Scenario-based simulations further demonstrated the impacts of LULC change are strongly mediated by spatial arrangement. Forest-to-residential conversions slightly reduced TP, whereas forest-to-industrial and forest-to-agricultural conversions led to pronounced increases, with agriculture (especially paddy fields) producing the largest rise. These findings underscore the necessity of considering not only LULC proportions but also their spatial arrangement and contextual factors in watershed management.

Overall, this study demonstrates that coupling deep learning models with explainable AI approaches provides superior predictive and interpretive capacity compared to proportion-only regression models. The proposed framework offers practical value as a decision-support tool for land-use planning and water quality management. Future work should integrate time-series monitoring, dynamic land-use change data, and hydrological modeling to improve temporal applicability and extend the framework to multiple pollutants and diverse geographic contexts.

In contrast to previous TP prediction studies that relied solely on proportional LULC metrics, the spatially explicit CNN–SHAP–IG framework introduced here demonstrates that incorporating LULC patch geometry and spatial configuration substantially improves both predictive accuracy and interpretability. This methodological advancement highlights the importance of spatial patterns in watershed modeling and offers a new direction for data-driven water quality assessment.

CRediT authorship contribution statement

Hye Won Lee: Writing – original draft, Visualization, Investigation, Formal analysis. **Min Kim:** Writing – original draft, Visualization, Methodology, Formal analysis. **Baehyun Min:** Writing – review & editing, Supervision, Conceptualization. **Jung Hyun Choi:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Jung Hyun Choi reports financial support was provided by Ewha Womans University. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2018-NR031078, RS-2025-16072431) and the Brain Korea 21 FOUR Program “REFINANCE: Research and Education Team for an Integrated Approach on Newly Emerging Contaminants

in Multiple Environmental Media" at the Department of Environmental Science & Engineering, Ewha Womans University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2026.103598.

Data availability

The datasets generated and analyzed in this study are openly available on Zenodo at https://zenodo.org/records/17238623 (DOI: 10.5281/zenodo.17238623). Detailed descriptions of data processing and preparation are provided in the Methods section and Appendix.

References

Adedeji, I.C., Ahmadisharaf, E., Sun, Y., 2022. Predicting in-stream water-quality constituents at the watershed scale using machine learning. J. Contam. Hydrol. 251, 104078. https://doi.org/10.1016/j.jconhyd.2022.104078.

Ahmadisharaf, E., Lacher, I.L., Fergus, C., Benham, B.L., Akre, T., Kline, K.S., 2020. Projecting land use change impacts on nutrients, sediment and runoff in multiple spatial scales: business-as-usual vs. stakeholder-informed scenarios. J. Clean. Prod. 257, 120466. https://doi.org/10.1016/j.jclepro.2020.120466.

Alamdari, N., Claggett, P., Sample, D.J., Easton, Z.M., Yazdi, M.N., 2022. Evaluating the joint effects of climate and land use change on runoff and pollutant loading in a rapidly developing watershed. J. Clean. Prod. 330, 129953. https://doi.org/10.1016/j.jclepro.2021.129953.

Bartlett, M.S., Van Blitterswyk, J., Farella, M., Li, J., Smith, C., Parolari, A.J., Krishnamoorthy, L., Mrad, A., 2025. Physically based dimensionless features for pluvial flood mapping with machine learning. Water Resour. Res. 61 (4), e2024WR039086. https://doi.org/10.1029/2024WR039086.

Barzegar, R., Aalami, M.T., Adamowski, J., 2020. Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. Stoch. Env. Res. Risk A. 34, 415–433. https://doi.org/10.1007/s00477-020-01776-2.

Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281–305.

Chen, Z., Xu, H., Jiang, P., Yu, S., Lin, G., Bychkov, I., Hmelnov, A., Ruzhnikov, G., Zhu, N., Liu, Z., 2021. A transfer learning-based LSTM strategy for imputing large-scale consecutive missing data and its application in a water quality prediction system. J. Hydrol. 602, 126573. https://doi.org/10.1016/j.jhydrol.2021.126573.

Cheng, C., Zhang, F., Shi, J., Kung, H.-T., 2022. What is the relationship between land use and surface water quality? A review and prospects from remote sensing perspective. Environ. Sci. Pollut. Res. 29, 56887–56907. https://doi.org/10.1007/s11356-022-21348-x.

Cheng, Y., Li, P., Xu, G., Li, Z., Yu, K., Cheng, S., Zhao, B., Wang, F., 2018. Factors that influence soil total phosphorus sources on dam fields that are part of ecological construction programs on the loess plateau, China. Catena 171, 107–114. https://doi.org/10.1016/j.catena.2018.07.006.

Choi, H., Koh, D.-C., Yoon, Y.Y., 2023. Spatial investigation of water quality and estimation of groundwater pollution along the main stream in the Geum River basin, Korea. Environ. Geochem. Health 45, 6387–6406. https://doi.org/10.1007/s10653-023-01643-3.

Clark, K.E., Bravo, V., Giddings, S.N., Davis, K.A., Pawlak, G., Torres, R., Adelson, A.E., César, C., Boza, X., Collin, R., 2022. Land use and land cover shape river water quality at a continental Caribbean land–ocean interface. Front. Water 4, 737920. https://doi.org/10.3389/frwa.2022.737920.

Dolph, C.L., Cho, S.J., Finlay, J.C., Hansen, A.T., Dalzell, B., 2023. Predicting high resolution total phosphorus concentrations for soils of the upper Mississippi River basin using machine learning. Biogeochemistry 163, 289–310. https://doi.org/10.1007/s10533-023-01029-8.

Du, Y., Pechlivanidis, G.I., 2025. Hybrid approaches enhance hydrological model usability for local streamflow prediction. Commun. Earth Environ. 6, 334. https://doi.org/10.1038/s43247-025-02324-y.

Ehteram, M., Ahmed, A.N., Sherif, M., El-Shafie, A., 2024. An advanced deep learning model for predicting water quality index. Ecol. Indic. 160 (13), 111806. https://doi.org/10.1016/j.ecolind.2024.111806.

Fu, B., Horsburgh, J.S., Jakeman, A.J., Gualtieri, C., Arnold, T., Marshall, L., Rashleigh, B., 2020. Modeling water quality in watersheds: from here to the next generation. Water Resour. Res. 56 (11). https://doi.org/10.1029/2020WR027721 e2020WR027721.

Haghiabi, A.H., Nasrolahi, A.H., Parsaie, A., 2018. Water quality prediction using machine learning methods. Water Qual. Res. J. 53 (1), 3–13. https://doi.org/10.2166/wqrj.2018.025.

Hanson, P.C., Stillman, A.B., Jia, X., Karpatne, A., Dugan, H.A., Carey, C.C., Stachelek, J., Ward, N.K., Zhang, Y., Read, J.S., Kumar, V., 2020. Predicting lake surface water phosphorus dynamics using process-guided machine learning. Ecol. Model. 430, 109136. https://doi.org/10.1016/j.ecolmodel.2020.109136.

Hirn, J., Sanz, V., García, J.E., Goberna, M., Montesinos-Navarro, A., Navarro-Cano, J.A., Sánchez-Martín, R., Valiente-Banuet, A., Verdú, M., 2024. Transfer learning of species co-occurrence patterns between plant communities. Ecol. Inform. 83, 102826. https://doi.org/10.1016/j.ecoinf.2024.102826.

Höge, M., Scheidegger, A., Baity-Jesi, M., Albert, C., Fenicia, F., 2022. Improving hydrologic models for predictions and process understanding using neural ODEs. Hydrol. Earth Syst. Sci. 26 (19), 5085–5102. https://doi.org/10.5194/hess-26-5085-2022.

Kang, M.A., Choi, B.W., Yu, J.J., 2010. Loading characteristics of non-point source pollutants by rainfall: case study with cherry tree plot. J. Eng. Geol. 20, 401–407.

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V., 2017. Theory-guided data science: a new paradigm for scientific discovery from data. IEEE Trans. Knowl. Data Eng. 29 (10), 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168.

Kazadi, A., Doss-Gollin, J., Sebastian, A., Silva, A., 2024. FloodGNN-GRU: a spatio-temporal graph neural network for flood prediction. Environ. Data Sci. 3, e21. https://doi.org/10.1017/eds.2024.19.

Khosravi, K., Rezaie, F., Cooper, J.R., Kalantari, Z., Abolfathi, S., Hatamiafkoueieh, J., 2023. Soil water erosion susceptibility assessment using deep learning algorithms. J. Hydrol. 618, 129229. https://doi.org/10.1016/j.jhydrol.2023.129229.

Kim, E.J., Kim, Y.S., Rhew, D.H., Park, B.K., 2016. The effect of TMDL program on water quality improvement in the Geum River basin, Korea. Desalin. Water Treat. 57 (2), 597–609. https://doi.org/10.1080/19443994.2014.986824.

Kim, S., Alizamir, M., Zounemat-Kermani, M., Kisi, O., Singh, V.P., 2020. Assessing BOD using neural networks and ensemble tree approaches in South Korea. J. Environ. Manag. 270, 110834. https://doi.org/10.1016/j.jenvman.2020.110834.

Kruk, M., 2023. Prediction of environmental factors responsible for chlorophyll-a-induced hypereutrophy using explainable machine learning. Ecol. Inform. 75, 102005. https://doi.org/10.1016/j.ecoinf.2023.102005.

Kwon, H.G., Jo, C.D., 2023. Water quality assessment of the Nam River, Korea, using multivariate statistical analysis and WQI. Int. J. Environ. Sci. Technol. 20, 2487–2502. https://doi.org/10.1007/s13762-023-04756-5.

Kwon, Y.S., Kang, H., Pyo, J., 2024. Estimation of aquatic ecosystem health using deep neural network with nonlinear data mapping. Ecol. Inform. 81, 102588. https://doi.org/10.1016/j.ecoinf.2024.102588.

Lee, H.W., Kim, M., Son, H.W., Min, B., Choi, J.H., 2022. Machine-learning-based water quality management of river with serial impoundments in the Republic of Korea. J. Hydrol. Reg. Stud. 41, 101069. https://doi.org/10.1016/j.ejrh.2022.101069.

Lee, J.W., Park, S.R., Lee, S.W., 2023. Effect of land use on stream water quality and biological conditions in multi-scale watersheds. Water 15 (24), 4210. https://doi.org/10.3390/w15244210.

Li, H., Zhao, H., Wei, C., Cao, M., Zhang, J., Zhang, H., Yuan, D., 2024. Assessing water quality environmental grades using hyperspectral images and a deep learning model: a case study in Jiangsu, China. Ecol. Inform. 84, 102854. https://doi.org/10.1016/j.ecoinf.2024.102854.

Liu, M., Xu, X., Jiang, Y., Huang, Q., Huo, Z., Liu, L., Huang, G., 2020. Responses of crop growth and water productivity to climate change and water-saving in arid region. Sci. Total Environ. 703, 134621. https://doi.org/10.1016/j.scitotenv.2019.134621.

Ly, Q.V., Nguyen, X.C., Lê, N.C., Truong, T.D., Hoang, T.H.T., Park, T.J., Maqbool, T., Pyo, J., Cho, K.H., Lee, K.S., Hur, J., 2021. Application of machine learning for eutrophication analysis and algal bloom prediction in an urban river: a 10-year study of the Han River, South Korea. Sci. Total Environ. 797, 149040. https://doi.org/10.1016/j.scitotenv.2021.149040.

Mayou, L.A., Alamdari, N., Ahmadisharaf, E., Kamali, M., 2024. Impacts of future climate and land-use/land-cover change on urban runoff using fine-scale hydrologic modeling. J. Environ. Manag. 362, 121284. https://doi.org/10.1016/j.jenvman.2024.121284.

Molina-Navarro, E., Trolle, D., Martínez-Pérez, S., Sastre-Merlín, A., Jeppesen, E., 2014. Hydrological and water quality impact assessment of a Mediterranean limno-reservoir under climate change and land use management scenarios. J. Hydrol. 509, 354–366. https://doi.org/10.1016/j.jhydrol.2013.11.053.

O'Shea, K., Nash, R., 2015. An introduction to convolutional neural networks. arXiv. https://doi.org/10.48550/arXiv.1511.08458, 1511.08458.

Pakdehi, M., Ahmadisharaf, E., 2025. Hindcasting maximum water depths in coastal watersheds: the importance of incorporating off-channel data and their uncertainties in machine learning models. Water Resour. Res. 61 (4). https://doi.org/10.1029/2024WR039244 e2024WR039244.

Pakdehi, M., Ahmadisharaf, E., Nazari, B., Cho, E., 2024. Transferability of machine-learning-based modeling frameworks across flood events for hindcasting maximum river water depths in coastal watersheds. Nat. Hazards Earth Syst. Sci. 24 (10), 3537–3555. https://doi.org/10.5194/nhess-24-3537-2024.

Prasad, D.V.V., Venkataramana, L.Y., Kumar, P.S., Prasannamedha, G., Harshana, S., Srividya, S.J., Harrinei, K., Indraganti, S., 2022. Analysis and prediction of water quality using deep learning and auto deep learning techniques. Sci. Total Environ. 821, 153311. https://doi.org/10.1016/j.scitotenv.2022.153311.

Rabby, H.S., Sun, X., Ibna Hafiz, A.M., Yan, Z., Imtiaz, S.U., Nasr Azadani, M., Pakdehi, M., Salou Moumouni, A., Ahmadisharaf, E., Alamdari, N., 2026. Application of machine learning methods in water quality modeling. In: Kumar, A., Zhang, X. (Eds.), Machine Learning and Artificial Intelligence in Toxicology and Environmental Health. Academic Press, pp. 271–309. https://doi.org/10.1016/B978-0-443-30010-3.00016-7.

Saad, S., Cereghetti, G., Feng, Y., Picotti, P., Peter, M., Dechant, R., 2017. Reversible protein aggregation is a protective mechanism to ensure cell cycle restart after stress. Nat. Cell Biol. 19, 1202–1213. https://doi.org/10.1038/ncb3600.

Seo, M., Heo, J., Kim, Y., 2021. Present and potential future critical source areas of nonpoint source pollution: a case of the Nakdong River watershed, South Korea. Environ. Sci. Pollut. Res. 28, 45676–45692. https://doi.org/10.1007/s11356-021-12976-w.

Sethy, P.K., Barpanda, N.K., Rath, A.K., Behera, S.K., 2020. Nitrogen deficiency prediction of rice crop based on convolutional neural network. J. Ambient. Intell. Humaniz. Comput. 11, 5703–5711. https://doi.org/10.1007/s12652-020-01938-8.

Shen, L.Q., Amatulli, G., Sethi, T., Raymond, P., Domisch, S., 2020. Estimating nitrogen and phosphorus concentrations in streams and rivers, within a machine learning framework. Sci. Data 7 (1), 161. https://doi.org/10.1038/s41597-020-0478-7.

Shim, S.H., Choi, J.H., 2024. Building an XGBoost model integrating landscape metrics and meteorological data for non-point source pollution management in the Nakdong River watershed. Ecol. Indic. 165, 112156. https://doi.org/10.1016/j.ecolind.2024.112156.

Venkateswarlu, T., Anmala, J., 2023. Importance of land use factors in the prediction of water quality of the upper Green River watershed, Kentucky, USA, using random forest. Environ. Dev. Sustain. 26, 23961–23984. https://doi.org/10.1007/s10668-023-03630-1.

Wai, K.P., Chia, M.Y., Koo, C.H., Huang, Y.F., Chong, W.C., 2022. Applications of deep learning in water quality management: a state-of-the-art review. J. Hydrol. 613, 128332. https://doi.org/10.1016/j.jhydrol.2022.128332.

Wang, W., Yang, P., Xia, J., Huang, H., Li, J., 2023. Impact of land use on water quality in buffer zones at different scales in the Poyang Lake, middle reaches of the Yangtze River basin. Sci. Total Environ. 896, 165161. https://doi.org/10.1016/j.scitotenv.2023.165161.

Youssef, Y.M., Gemail, K.S., Atia, H.M., Mahdy, M., 2024. Insight into land cover dynamics and water challenges under anthropogenic and climatic changes in the eastern Nile Delta: inference from remote sensing and GIS data. Sci. Total Environ. 913, 169690. https://doi.org/10.1016/j.scitotenv.2023.169690.

Zhang, Z., Zhang, F., Du, J., Chen, D., Zhang, W., 2021. Impacts of land use at multiple buffer scales on seasonal water quality in a reticular river network area. PLoS One 16, e0244606. https://doi.org/10.1371/journal.pone.0244606.

Zhang, Z., Huang, J., Duan, S., Huang, Y., Cai, J., Bian, J., 2022. Use of interpretable machine learning to identify the factors influencing the nonlinear linkage between land use and river water quality in the Chesapeake Bay watershed. Ecol. Indic. 140, 108977. https://doi.org/10.1016/j.ecolind.2022.108977.

Zhi, W., Appling, A.P., Golden, H.E., Podgorski, J., Li, L., 2024. Deep learning for water quality. Nat. Water 2, 228–241. https://doi.org/10.1038/s44221-024-00202-z.