Plastic Waste Leakage in Urban River Basins of Sri Lanka: A Geospatial and Machine Learning Perspective

K.D.B Aratthanage
Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nugegoda, Sri Lanka
kdara@ou.ac.lk

N.G.P.R Premachandra
Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nugegoda, Sri Lanka
ngpre@ou.ac.lk

J.Kirushika
Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nugegoda, Sri Lanka
jkiru@ou.ac.lk

T.D.T Madhushika
Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nugegoda, Sri Lanka
tdmad@ou.ac.lk

H.P.I Subasighe
Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nugegoda, Sri Lanka
hpsub@ou.ac.lk

N.C. Wanigasuriya
Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nugegoda, Sri Lanka
ncwan@ou.ac.lk

***Abstract*— Plastic waste leakage into the urban river basins has emerged as a critical environmental challenge in Sri Lanka, threatening aquatic ecosystems, public health, and urban resilience. Rapid urbanization, inadequate waste management, and poor drainage infrastructure have accelerated the accumulation and transport of plastic debris into waterways. This study presents a geospatial and machine learning approach to identify and model the key drivers of plastic waste leakage in selected urban river basins. High-resolution geospatial datasets, including land use, population density, slope, rainfall patterns, and proximity to waste disposal sites, were integrated with field-based leakage observations. A Random Forest classifier was employed to predict leakage hotspots, achieving a moderate level of accuracy. Feature importance analysis highlighted waste site proximity, urban density, and hydrological factors as the dominant predictors. The confusion matrix further illustrates the model's strengths in identifying high-risk zones; however, misclassification in low-risk areas suggests potential improvement through the inclusion of additional environmental and socio-economic variables. The findings offer actionable insights for urban planners, waste management authorities, and policymakers to develop targeted interventions for mitigating plastic waste in riverine environments.**

Keywords—Plastic waste leakage, Urban river basins in Sri Lanka, Geospatial analysis, Machine learning, Environmental modeling.

I. INTRODUCTION

Plastic pollution in urban rivers has become a pressing environmental concern across the globe, and Sri Lanka is no exception. Major urban river basins such as the Kelani, Mahaweli, and Gin are increasingly burdened by plastic waste, which worsens water pollution and threatens aquatic ecosystems. While past studies in Sri Lanka have largely focused on microplastic contamination in coastal areas, there is a noticeable lack of research addressing how plastic waste behaves within inland urban rivers [1], [2].

Managing plastic waste in these complex river environments requires more than just traditional monitoring, it calls for the integration of advanced tools like remote sensing (RS), geographic information systems (GIS), and machine learning (ML). Although satellite imagery, particularly from platforms like Sentinel-2, has proven useful in environmental monitoring, its use in tracking plastic waste in urban waterways remains underutilized in the Sri Lankan context. Furthermore, the potential of machine learning to anticipate plastic leakage patterns or detect high-risk zones along rivers has not yet been fully explored [6].

Plastic waste pollution in freshwater systems is a growing global concern, and Sri Lanka is beginning to experience the harsh effects of this issue, particularly in its urban river basins. Rivers such as the Kelani, Mahaweli, and Gin play a vital role in supporting urban populations, providing water for domestic use, agriculture, and industry. However, these rivers are increasingly burdened by mismanaged plastic waste, which accumulates due to rapid urbanization, insufficient waste disposal infrastructure, and lack of public awareness. As plastic leakage continues to rise, it degrades water quality, endangers aquatic life, and contributes to broader ecological and public health risks.

The challenge of managing plastic leakage in river basins is complex due to the dynamic nature of river systems and the diverse sources of plastic waste. A variety of factors, including land use, topography, rainfall patterns, and proximity to waste disposal sites, influence urban river environments. Despite the growing availability of spatial data and advancements in environmental monitoring technologies, there has been a lack of effort to integrate these tools into plastic pollution research in the Sri Lankan context. Remote sensing (RS) and geographic information systems (GIS) have been employed globally to monitor land use and water quality. However, their application in detecting and modeling plastic leakage into rivers is still in its nascent stages locally. Similarly, ML techniques, which are gaining traction in environmental modeling, are still not fully utilized in researching plastic waste pathways and predicting hotspots in Sri Lanka [3], [6].

This lack of integration highlights several research gaps. Firstly, there is minimal application of geospatial techniques for mapping and quantifying plastic waste in urban freshwater systems. Secondly, ML algorithms have not been sufficiently applied to predict leakage patterns or prioritize high-risk zones [7]. Lastly, existing studies tend to isolate variables rather than integrating multiple spatial and environmental layers, which is critical for accurately modeling plastic leakage. Additionally, the absence of field-validated high-resolution data limits the reliability and usefulness of any spatial model that may be developed. To address these issues, this study aims to develop a predictive geospatial framework that models plastic waste leakage in selected urban river basins of Sri Lanka. By incorporating variables such as land use, proximity to dumpsites, elevation, rainfall distribution, and population density, the research will identify key plastic pollution hotspots using machine learning techniques. The integration of GIS, remote sensing data (e.g., Sentinel-2 imagery), and advanced spatial analytics will enhance the understanding of the plastic leakage process and provide practical tools for decision-makers. The outcome of this study is expected to support waste management planning and inform future river conservation strategies.

This study aims to bridge these gaps by using a combined geospatial and data-driven approach to model plastic waste leakage in Sri Lanka's urban river basins. By analyzing variables such as land use, distance to waste disposal sites, rainfall intensity, and terrain features, we propose a spatial model that can effectively identify and predict plastic waste hotspots. The goal is to support better waste management planning and provide actionable insights for policymakers and environmental agencies working to protect Sri Lanka's vital freshwater resources.

II. Related Work

Plastic pollution in rivers has been widely studied, with prior research mainly focusing on microplastic contamination in Sri Lanka's coastal areas [3]. However, studies on urban river basins in Sri Lanka are limited. Globally, geospatial tools and machine learning have been used to map riverine plastic waste. For instance, employed Sentinel-2 imagery with supervised classifiers to detect plastic litter, demonstrating the effectiveness of combining remote sensing with machine learning. These methods have not been applied to Sri Lankan rivers [4]. Boucher et al. developed a GIS-based source-to-sea model for plastic leakage, however, this framework has not been applied in the urban contexts of Sri Lanka. Additionally, while rainfall and hydrology are known to influence plastic transport, these factors remain underutilized in existing predictive models for urban river plastic pollution [5].

III. Literature Review

Plastic pollution in aquatic ecosystems is a growing global concern, with rivers serving as major conduits that transport terrestrial plastic waste to marine environments. Sri Lanka, a rapidly urbanizing island nation, faces increasing challenges related to mismanaged plastic waste entering river systems, particularly in urban basins such as the Kelani and Mahaweli rivers. Existing studies in Sri Lanka have primarily focused on microplastic contamination in coastal and estuarine zones. For instance, Wickramasinghe et al. [1] conducted microplastic sampling from estuaries along the western coast of Sri Lanka, including the Kelani River mouth. Similarly, field studies in Galle City examined the accumulation of plastic debris in inland waterways using CCTV-enabled waste traps, revealing that plastic bottles were the predominant pollutant [2]. These studies provide foundational insights into the presence and composition of plastic waste, but they lack spatial modeling or predictive analytics. Recent work has explored the use of ML in environmental monitoring in Sri Lanka. Liyanage and Manage [3] employed unsupervised learning and pollution indexing to identify sources of industrial pollutants in the Lower Kelani River Basin. However, their focus was on heavy metal contamination, and no plastic-specific modeling was performed. While this demonstrates the potential of ML in water quality assessments, it highlights the absence of similar applications for plastic waste. Internationally, researchers have increasingly leveraged remote sensing and ML to detect plastic in terrestrial and riverine environments. For example, Vo et al. [4] developed an ML-based method to detect riverine litter using Sentinel-2 imagery and supervised classifiers such as Random Forest (RF) and Support Vector Machines (SVM). These studies, although promising, are rarely validated in tropical regions such as Sri Lanka and often lack ground-truth data. Similarly, Boucher et al. [5] proposed a source-to-sea framework for modeling plastic leakage in the Mekong River Basin, integrating hydrology, urban land-use, and waste infrastructure data within GIS environment. No such approach has yet been adapted to Sri Lankan river basins. The lack of labeled datasets for riverine plastic detection remains a significant limitation for ML implementation. Studies in other contexts have used synthetic datasets or crowdsourced annotations to train models [6], but Sri Lanka lacks open-access datasets the capture GPS-tagged plastic waste accumulations or drone imagery of riverbanks. Additionally, although rainfall and topography are recognized as key drivers of plastic leakage [7], these features have not been integrated into any plastic pollution hotspot models in Sri Lanka [8].

While UAV (drone)-based imagery has been applied in Southeast Asia for high-resolution plastic detection [8], there is no evidence of similar applications in Sri Lanka. The combination of UAV data with CNNs or U-Net architectures offers a promising yet unexplored method for detecting and mapping plastic accumulation along riverbanks in the country. Furthermore, current environmental governance lacks spatially enabled decision-support systems for managing plastic waste in urban river catchments. While national policies (e.g., bans on polythene <20 microns) exist, there is a disconnect between policy enforcement and spatial monitoring at the river basin level. In summary, there is a clear research gap in integrating GIS, remote sensing, and machine learning to model plastic waste leakage in Sri Lanka's urban river systems. No existing study provides an end-to-end spatial framework that combines remote sensing, ground-truth data, ML algorithms, and hydrological modeling for predicting plastic leakage hotspots. This gap presents a valuable opportunity for the development of a geospatial ML-based system for plastic pollution assessment in Sri Lanka's urban river basins [9], [12].

A. *Plastic Pollution in Rivers*

Research in Sri Lanka has mostly concentrated on microplastic pollution in coastal and estuarine areas, such as the Kelani and Maha Oya estuaries. While these studies are valuable, they neglect the inland urban river basins, which are often the initial zones where plastic waste enters the water system. There is a significant gap in understanding how plastics leak into rivers within cities, where human activity and poor waste management are more intense.

B. *Geospatial Analysis (GIS/Remote Sensing)*

GIS and Remote Sensing (RS) have been widely used in Sri Lanka for analyzing land use and water quality, often through pollution indices or watershed studies. However, these tools have not yet been applied to map plastic waste hotspots [10], [11]. The current research lacks integration of key spatial layers such as land use/land cover (LULC), proximity to rivers, road networks, and dump sites.

C. *Machine Learning in Environmental Studies)*

ML has found applications in Sri Lanka for predicting floods, water pollution, and other environmental phenomena. Despite this, ML has not been employed to analyze or predict plastic waste hotspots in river ecosystems. There is an opportunity to pioneer this approach in your study by using models such as Random Forest (RF), XGBoost, or Support Vector Machines (SVM). These algorithms can classify regions based on environmental and anthropogenic variables, helping to predict where plastic waste is likely to accumulate and offering data-driven support for mitigation strategies [7].

D. *Plastic Leakage Source-to-Sink Mapping*

Globally, plastic leakage models have been developed for major rivers such as the Mekong and Ganges, tracing the flow of plastic waste from urban sources through rivers to the ocean. In Sri Lanka, no such source-to-sink models currently exist [11].

E. *Labeled Dataset Availability*

One of the major challenges in developing ML models for plastic pollution is the absence of labeled datasets, especially for riverine plastic in Sri Lanka. There is a lack of ground-truth data, such as GPS-tagged plastic waste locations, image datasets, or manually labeled satellite images. Although studies mention the impact of rainfall and topography on general waste dispersion, these variables are rarely integrated into plastic pollution models in Sri Lanka [13], [14], [15], [16]. Monsoonal flows have a significant influence how plastic is transported within river basins.

F. *Drone/UAV-Based Monitoring*

Enhancing global success in using drones/UAVs for environmental monitoring, Sri Lanka has not yet adopted this approach for tracking riverine plastic pollution. This presents a clear research gap [17]. By using UAVs to collect high-resolution images of river basins and applying deep learning techniques such as Convolutional Neural Networks (CNNs) or U-Net, your study can advance real-time plastic detection, especially in areas where satellite imagery may lack sufficient resolution [18], [19]. While Sri Lanka has implemented national policies such as bans on single-use plastics, these initiatives lack spatial decision support systems to guide local-level interventions. Policymakers currently operate without detailed maps or hotspot predictions [20].

IV. METHODOLOGY

This research focuses on urban river basins, specifically the Kelani and Mahaweli Rivers, due to their ecological value and proximity to dense urban areas. Using open-access datasets, satellite imagery, and limited field validation, the study notes limitations in data availability, seasonal variation, and high-resolution ground truth. The adaptable, scalable framework supports future studies and sustainable environmental management in Sri Lanka. The methodology has five stages.

A. Study Area Selection

The research focuses on two major urban river basins in Sri Lanka, the Kelani River Basin and the Mahaweli River Basin, due to their high population density, proximity to urban settlements, and documented issues with plastic waste pollution. These basins are representative of the challenges faced by rapidly urbanizing regions where solid waste management infrastructure is limited. The Kelani River, which flows through the capital city of Colombo, and the Mahaweli River, the longest river in Sri Lanka, both experience substantial urban and industrial runoff, making them ideal candidates for studying the dynamics of plastic waste leakage into urban rivers.

B. Data Collection and Preprocessing

Multiple spatial and non-spatial data sources were used to support this study: Satellite Imagery: Sentinel-2 optical data was accessed via Google Earth Engine for land use/land cover (LULC) classification and to detect surface water boundaries. The high-resolution imagery (10m, 20m, and 60m) enabled accurate mapping of land cover types and water features within the study area. Topographic Data: Shuttle Radar Topography Mission (SRTM) data, providing digital elevation models (DEMs) at a 30m resolution, was used to derive slope and elevation information necessary for understanding the physical landscape's effect on plastic waste movement. Rainfall Data: Monthly precipitation data were retrieved from the CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) dataset. Rainfall intensity data were essential to model the effect of monsoonal rains on plastic waste transport. Waste Sites and Urban Infrastructure collected via OpenStreetMap and local authority databases were used to map the locations of dumpsites, roads, residential zones, and industrial facilities. These datasets provided insight into potential sources of plastic leakage and urban infrastructure that may influence plastic accumulation. For ground Observations a limited set of field observations and drone imagery were used to validate hotspot locations. UAV imagery was particularly useful for identifying visible plastic waste accumulation points in areas where satellite imagery might not have sufficient resolution. All datasets were preprocessed in QGIS, ArcGIS, and Python to ensure spatial alignment and to convert categorical features into numerical layers suitable for machine learning model training.

C. Variable Derivation and GIS Analysis

From the collected data, the following key spatial predictors were derived.

TABLE I. VARIABLE DERIVATION

Factor	*Description*
Distance to waste disposal sites	Proximity to disposal points influences waste accumulation patterns.
Distance to rivers/streams	Closer proximity increases risk of plastic waste entering water bodies.
Land use category	Urban/industrial areas generate more plastic waste than vegetated areas.
Slope and elevation (DEM)	Topography affects plastic waste flow and accumulation.
Proximity to road networks	Roads can act as pathways for plastic waste movement.
Rainfall intensity (monthly avg.)	Heavy rainfall facilitates plastic waste transport via runoff.
Population density	Higher density correlates with increased plastic consumption and mismanagement.

Each variable was rasterized and standardized to a common spatial resolution (e.g., 30m × 30m) to create a unified geospatial data cube, which was then used as input for the machine learning model.

D. Machine Learning Model Development

RF classifier was employed because of its robustness in handling heterogeneous environmental data and its interpretability within spatial modeling tasks. The workflow was structured to ensure transparency and reproducibility and comprised data labeling, feature selection, model training, and hotspot prediction to map plastic-waste risk. The model workflow included the following steps.

TABLE II. MODEL DEVELOPMENT

Steps	*Description*
Labeling	Known plastic accumulation points from field surveys and previous studies were used as positive classes (i.e., locations where plastic waste is present).
Feature Selection	Recursive feature elimination (RFE) was employed to select the most influential predictors from the available spatial variables.
Model Training	A 70/30 train-test split was applied to the data. The model was trained using Scikit-learn in Python, and cross-validation was used to optimize hyperparameters, ensuring the model generalized well to unseen data
Hotspot Prediction	After training, the model was applied to the entire study area to predict plastic waste leakage probability

The output was a raster layer indicating the likelihood of plastic waste accumulation at each pixel.

E. Model Validation and Accuracy Assessment

The model's performance was evaluated using several key parameters. Accuracy measured the overall percentage of correct predictions, while precision assessed the model's ability to correctly identify plastic waste hotspots, and recall evaluated its capacity to capture all relevant plastic waste points. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was used to determine the model's ability to distinguish between plastic waste and non-plastic waste locations. Additionally, a confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives, offering deeper insight into the classification performance. Spatial validation was also performed by comparing predicted hotspots with observed plastic waste clusters from UAV imagery and field visits. This ensured the model's practical relevance and allowed for adjustments in model parameters.

V. RESULTS AND DISCUSSIONS

A. Model Performance

The model achieved an overall accuracy of 69.33%, indicating a moderate level of performance. While this is a reasonable result, it suggests that the model could be further improved, especially in terms of correctly identifying areas with plastic waste leakage. The Confusion Matrix (shown below) reveals the following.

TABLE III. MODEL PERFORMANCE

Results	***Description***
True Positives	Plastic leakage predicted correctly): 3
True Negatives	No plastic leakage predicted correctly): 205
False Positives	Plastic leakage predicted incorrectly): 5
False Negatives	No plastic leakage predicted incorrectly): 87

From this confusion matrix, we can see that the model is better at identifying "No Plastic" locations (with a high true negative count), but it struggles with detecting actual plastic leakage (due to a high number of false negatives) (Fig.1). The AUC-ROC score of 0.5048 indicates that the model has just about a random guessing capability. A higher score (closer to 1) would indicate a better ability to distinguish between plastic and non-plastic locations. The AUC-ROC score suggests that further refinements to the model are needed, especially in handling imbalanced data or feature selection.

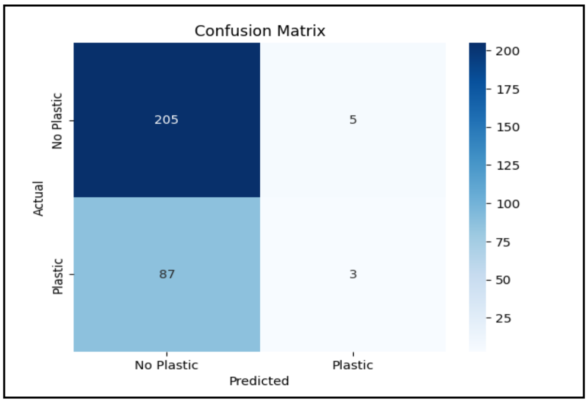

Fig.1. Confusion Matrix

B. Feature Importance

The feature importance plot highlights the most influential variables in predicting plastic leakage. Distance to Waste Sites emerged as the most significant predictor, with an importance score of 0.1288, indicating that areas closer to waste disposal sites are more likely to experience plastic leakage. Other important predictors include population density with a score of 0.1220, elevation at 0.1207, longitude at 0.1206, and slope percentage at 0.1203, all contributing substantially to the model's predictive capability (Fig.2).

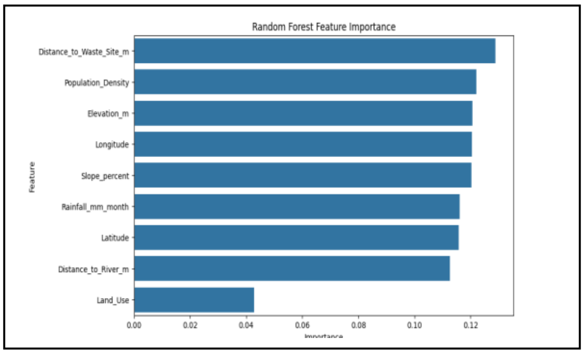

Fig.2. Random Forest Feature Importance

These features reflect environmental and infrastructural factors that contribute to plastic waste accumulation. Land use, with the lowest importance score of 0.0428, had a minimal impact on the model's predictions. This could be due to the coarse categorization of land use in the dataset or a need for better feature engineering.

VI. FUTURE WORK

To enhance the current model and overcome the limitations identified, future research could focus on addressing the issue of class imbalance by employing oversampling techniques or specialized algorithms such as SMOTE. Incorporating temporal data, including seasonal variations in waste accumulation and rainfall patterns, would likely improve the model's ability to capture dynamic changes over time. Utilizing higher-resolution imagery or conducting UAV-based surveys can aid in detecting smaller waste deposits that may have been missed by the existing approach. Exploring advanced machine learning methods, such as XGBoost or deep learning architectures, may further improve prediction accuracy and robustness. Additionally, integrating real-time monitoring systems through sensors or Internet of Things (IoT) technologies could enable continuous data collection and proactive intervention strategies, thereby enhancing plastic waste management in urban river basins.

VII. Conclusion

This study presents a machine learning approach to modeling plastic waste leakage in urban river basins of Sri Lanka, integrating geospatial data and environmental factors. Although the model achieved a moderate accuracy of 69.33%, the analysis also revealed challenges such as limited discriminatory power (AUC-ROC of 0.50), class imbalance, and restricted ground-truth data, which affected precision and recall. Despite these limitations, the study identifies the most influential predictors of plastic leakage proximity to waste sites, population density, elevation, slope, and rainfall offering valuable insights for waste management and policy interventions. The proposed framework is low-cost, adaptable, and scalable, providing a baseline for monitoring riverine plastic waste in other regions facing similar environmental pressures. This work serves as a foundation for more effective waste management strategies, helping urban planners and environmental authorities target areas most at risk of plastic pollution, thereby improving environmental conservation efforts.

References

[1] D. Wickramasinghe, A. Jayakody, and R. Weerasinghe, "Sampling and Analysis of Microplastics in the Coastal Environments of Sri Lanka: Estuaries of the Kelani River to Mahaoya," *Water*, vol. 16, no. 13, pp. 1932, 2024.

[2] S. Jayasinghe and M. Samarawickrama, "Analysing Plastic Pollution in Inland Waterways: A Case Study of Galle City," in *Proc. Forestry and Environment Symposium*, 2024.

[3] K. Liyanage and A. Manage, "Use of Machine Learning and Indexing Techniques for Identifying Industrial Pollutant Sources: A Case Study of the Lower Kelani River Basin," *Water*, vol. 16, no. 19, pp. 2766, 2024.

[4] M. Vo, J. Lebreton, and T. H. Nguyen, "Machine Learning-Based Detection and Mapping of Riverine Litter Using Sentinel-2 Imagery," *Environmental Science and Pollution Research*, vol. 30, pp. 63421–63435, 2023.

[5] J. Boucher and D. Billard, "Mapping Plastic Leakage in the Mekong River Basin Using GIS and Remote Sensing," *Sustainability*, vol. 14, no. 13, pp. 7879, 2022.

[6] D. Tran-Thanh, A. N. Rinasti, K. Gunasekara, A. Chaksan, and M. Tsukiji, "GIS and remote sensing-based approach for monitoring and assessment of plastic leakage and pollution reduction in the lower Mekong River Basin," *Sustainability*, vol. 14, no. 13, Art. no. 7879, Jun. 2022, doi: 10.3390/su14137879.

[7] N. Wijayaweera, L. Gunawardhana, J. Bamunawala, J. Sirisena, L. Rajapakse, C. Samarasuriya Patabendige, and H. Karunaweera, "Use of machine learning and indexing techniques for identifying industrial pollutant sources: a case study of the Lower Kelani River Basin, Sri Lanka," *Water*, vol. 16, no. 19, Art. no. 2766, Sep. 2024, doi: 10.3390/w16192766.

[8] N. P. P. S. Nugawela, A. S. Mahaliyana, and E. P. D. N. Thilakarathne, "Migration, pathways, and management of microplastics (MPs) through remote sensing and GIS techniques," in *Handbook of Microplastic Pollution in the Environment*, 1st ed., CRC Press, 2025.

[9] S. Pandey, P. C. Pandey, R. Costache, A. Arora, P. Ranjan, A. Chamoli, and M. Pandey, "An overview of machine learning applications in disaster waste management," in *Geo-information for Disaster Monitoring and Management*, P. C. Pandey et al., Eds., Cham, Switzerland: Springer, 2024, pp. 469–499. doi:10.1007/978-3-031-51053-3_21.

[10] A. N. Rinasti, I. F. Ibrahim, K. Gunasekara, T. Koottatep, and E. Winijkul, "Developing scenario of plastic waste leakage in the Jakarta hydrology environment using seasonal data conditions and socioeconomic aspects," in *Applied Geography and Geoinformatics for Sustainable Development*, Apr. Nov., Thailand: Springer, 2022, pp. 65–88. doi:10.1007/978-3-031-16217-6_5.

[11] A. N. Zulkifli, H. Z. M. Shafri, K. Hirose, N. M. Noor, K. N. A. Maulud, M. Z. Asmawi, L. Mokraoui, and Y. Ang, "Plastic waste mapping and monitoring using geospatial approaches," in *IOP Conference Series: Earth and Environmental Science*, vol. 1064, 2022, Art. no. 012008. doi: 10.1088/1755-1315/1064/1/012008.

[12] E. Ramos, A. G. Lopes, and F. Mendonça, "Application of Machine Learning in Plastic Waste Detection and Classification: A Systematic Review," *Processes*, vol. 12, no. 8, Art. no. 1632, 2024. doi: 10.3390/pr12081632.

[13] C. Kruse, E. Boyda, S. Chen, K. Karra, T. Bou-Nahra, D. Hammer, J. Mathis, T. Maddalene, J. Jambeck, and F. Laurier, "Satellite monitoring of terrestrial plastic waste," *PLOS ONE*, vol. 18, no. 1, Art. no. e0278997, Jan. 2023. doi: 10.1371/journal.pone.0278997.

[14] M. Kanevski, A. Pozdnukhov, and V. Timonin, "Machine Learning Algorithms for GeoSpatial Data. Applications and Software Tools," in *4th International Congress on Environmental Modelling and Software*, Barcelona, Spain, Jul. 2008.

[15] H. Mojaddadi Rizeei, B. Pradhan, H. Nampak, N. Ahmad, and A. H. B. Ghazali, "Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS," *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 1080–1102, 2017. doi: 10.1080/19475705.2017.1294113.

[16] A. Kovacs-Györi, A. Ristea, C. Havas, M. Mehaffy, H. H. Hochmair, B. Resch, L. Juhasz, A. Lehner, L. Ramasubramanian, and T. Blaschke, "Opportunities and Challenges of Geospatial Analysis for Promoting Urban Livability in the Era of Big Data and Machine Learning," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 12, art. 752, 2020. doi: 10.3390/ijgi9120752.

[17] J. M. Grzybowski and A. R. Pasha, "Geospatial analytics for enhanced environmental planning," in *Proceedings of the 2015 IEEE International Conference on Geospatial Analytics for Disaster Management*, Chennai, India, Dec. 2015, pp. 50–55. doi: 10.1109/GMAP.2015.7367615.

[18] C. Bellinger, M. S. M. Jabbar, O. Zaïane, and A. A. Osornio-Vargas, "A systematic review of data mining and machine learning for air pollution epidemiology," *BMC Public Health*, vol. 17, art. 907, Nov. 2017. doi: 10.1186/s12889-017-4914-3.

[19] M. Zhu, J. Wang, X. Yang, Y. Zhang, L. Zhang, H. Ren, B. Wu, and L. Ye, "A review of the application of machine learning in water quality evaluation," *Eco-Environment & Health*, vol. [Volume], no. [Issue], pp. [Pages], 2022. doi: 10.1016/j.eehl.2022.06.001.

[20] X. Wang, Y. Li, Q. Qiao, A. Tavares, and Y. Liang, "Water Quality Prediction Based on Machine Learning and Comprehensive Weighting Methods," *Entropy*, vol. 25, no. 8, art. no. 1186, Aug. 2023. doi: 10.3390/e25081186.