Fusing Satellite Remote Sensing and ARGO Float Data for Enhanced Monitoring of Microplastic Concentrations in the West Pacific (2018–2020)

Ting On Chan*, Zhijie Chang, Wing Sum Yan

School of Geography and Planning, Sun Yat -sen University, 510275 Guangzhou, China –chantingon@mail.sysu.edu.cn, (changzhj, yanws5)@mail2.sysu.edu.cn,

Commission Ⅲ

Keywords: Microplastics, West Pacific, Remote Sensing, Machine Learning, Dispersion Mechanisms

Abstract:

With the continuous intensification of marine plastic pollution, monitoring the transport and dispersion of microplastics has become a critical global concern. However, predicting microplastic concentrations remains highly challenging due to the lack of direct satellite signatures and the complex non-linear physical mechanisms governing their dispersion. This study develops an interpretable machine learning framework to monitor surface microplastic concentrations in the West Pacific from 2018 to 2020. We profoundly integrated Japanese AOMI in-situ microplastic observations, ERA5 meteorological/wave reanalysis, and Euro-Argo subsurface profile data utilizing a 3D Inverse Distance Weighting (3D-IDW) spatiotemporal interpolation algorithm. A Random Forest (RF) model was subsequently trained, achieving robust predictive accuracy (R^2 = 0.64, 0.76, and 0.87 for 2018, 2019, and 2020, respectively). Crucially, we incorporated SHapley Additive exPlanations (SHAP) to overcome the "black-box" limitations of traditional ensemble models. The SHAP analysis explicitly revealed a distinct, year-by-year regime shift in dominant environmental drivers: microplastic distribution was primarily governed by stable hydrographic and biological conditions in 2018; by dynamic wave forcing (e.g., long-period swells and Stokes drift) in 2019; and by extreme meteorological events (e.g., typhoon-induced terrestrial flushing) in 2020. Ultimately, this physics-informed framework successfully elucidates the dynamic transition of microplastic transport mechanisms between hydrographic–biological dominance and meteorological–physical forcing, providing vital scientific support for targeted pollution mitigation and coastal resilience planning.

1. Introduction

With accelerating globalization and urbanization, marine plastic pollution has intensified, making the transport and dispersion of microplastics a critical concern. To track their distribution, this study integrates Japanese AOMI (Aqua/Marine Ocean Microplastics) observations with satellite-derived wave parameters and develops a machine-learning framework to quantify how wave conditions influenced microplastic dispersion in the Japan-adjacent western North Pacific during 2018–2020, along with the Array for Real-Time Geostrophic Oceanography (ARGO) data.

Results reveal a year-by-year regime shift in dominant drivers. In 2018, hydrographic control prevailed: microplastics concentrated within warm, fresh, low-oxygen, high-biomass water masses, with biofouling enhancing aggregation. In 2019, long-period swell governed vertical mixing, while strong winds (negative SHAP values) promoted horizontal dispersion and reduced concentrations. In 2020, extreme rainfall—and typhoon events—delivered substantial terrestrial plastic inputs and, via intense turbulent mixing, produced more uniform distributions. Overall, transport and aggregation mechanisms transitioned from hydrographic–biological dominance toward meteorological–physical forcing.

Algorithmic advances are central to intelligent monitoring of dynamic coasts and adjacent ocean ecosystems. Methods have progressed from statistical approaches to machine learning and bespoke deep learning. Remote sensing quantifies coastal change and socio-ecological risks (Mishra et al., 2022); UAV imagery with water indices enables accurate shoreline extraction, with WI often performing best. Machine learning reduces data demands and boosts cross-sensor accuracy (Bayram et al., 2017), while semantic segmentation attains strong performance with limited labels (Zhong et al., 2022). Ensemble deep models such as WaterNet improve water segmentation under challenging illumination and water color (Erdem et al., 2021). These tools scale from human-pressured coasts (Venice; Fogarin et al., 2023) to basin-scale analyses (Black Sea; Görmüş et al., 2021), revealing erosion patterns, rates, and drivers.

Persistent uncertainties remain as indicated by shoreline accuracy declines with tidal range and complex foreshore morphology, exceeding 20 m on high-energy macrotidal beaches (Vos et al., 2023). ML can standardize shoreline indicators across tidal states to enable globally comparable datasets (McAllister et al., 2022). Extending this toolkit to marine plastics, multimodal fusion of satellite, UAV, and in-situ observations detects debris slicks, river plumes, and convergence fronts; physics-informed ML links wind, waves, and currents to microplastic transport; and domain-adapted segmentation distinguishes plastics from foam and sunglint. These integrations sharpen source-to-sink mapping, identify pathways and hotspots, and support targeted mitigation and coastal resilience planning.

Furthermore, understanding the interactions between microplastics and marine phytoplankton presents a critical monitoring dimension. Chlorophyll-a serves as a key biogeochemical proxy for detecting these biologically active water masses, and advanced machine learning models, such as gradient boosting decision trees, have proven highly effective in downscaling high-resolution chlorophyll-a concentrations from remote sensing observations (Zhang et al., 2023). Integrating such biological parameters is essential for capturing the bio-physical coupling that drives microplastic aggregation.

In this study, ERA5 wave data are sourced from ECMWF’s global reanalysis, providing hourly, 0.25°-resolution ocean–atmosphere parameters and widely regarded as one of the most reliable long-term records of global sea state. From the Copernicus Climate Data Store (CDS), we selected the ERA5 hourly single-level product, retrieving significant wave height, mean wave direction, mean wave period, and surface turbulent stress. The spatial domain was set to 0–40°N, 120–160°E, and the study years were 2018, 2019, and 2020. Data were batch-downloaded as NetCDF files for local processing.

Euro-Argo is a web-based tool that allows users to conveniently filter, visualize, and download Argo profile data (in NetCDF or CSV format) based on conditions such as time range, region, type of Argo float (Core Argo, Deep Argo, BGC Argo, etc.), observed parameters, and data quality. By selecting the desired time range and study area on the website and submitting a request, users can download the corresponding Argo profile data files in .csv format.

AOMI plastic data, released by Japan’s Ministry of the Environment, compile long-term monitoring of drifting microplastics/plastic pollution from nationwide shoreline and offshore stations. The dataset includes concentration, sampling time, and geolocation metadata. Using the AOMI web selection tool, we obtained annual trawl survey points for the corresponding years in CSV format. Where necessary, we harmonized units, screened for duplicates and outliers, and georeferenced records to enable fusion with ERA5 and AOMI layers.

2. Study area

This study focuses on two western Pacific sectors north of the Philippine Sea and south of central Japan （Figure 1）. The region has a complex marine environment, strongly influenced by the Kuroshio Extension, intermediate water mass exchange, and the pronounced topographic effects of the Ryukyu Islands. The predicted number of plastic particles is obtained using a Random Forest (RF) model (with input features including ERA5 precipitation, wind speed, wind direction, wave height, wave direction, wave period, and turbulent stress data, as well as ARGO temperature, salinity, dissolved oxygen, and BBP700 data). The reference set is selected from the dataset acquired by AOMI.

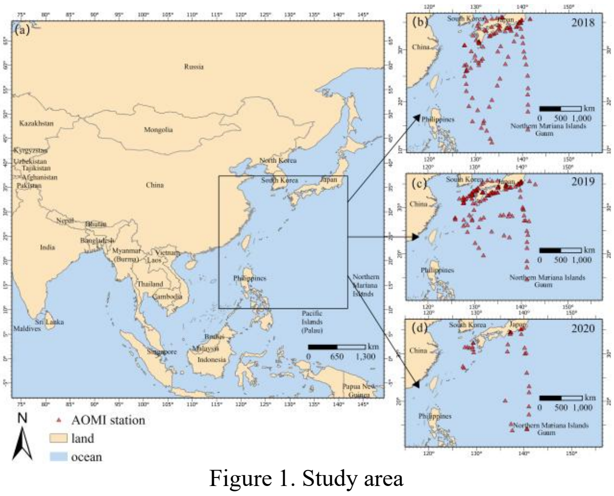


Figure 1. Study area

3. Method

The foundation of this framework relies on the profound integration of three distinct datasets. In-situ microplastic particle counts are derived from the AOMI dataset, serving as the target variable for model prediction. To accurately capture the dynamic marine environment, high-resolution meteorological and wave parameters (including wind speed, wave period, wind stress magnitude, and total precipitation) are extracted from the ERA5 hourly reanalysis data. Simultaneously, subsurface hydrological and biogeochemical profile data (encompassing temperature, salinity/PSAL, dissolved oxygen/DOX2, and particulate backscattering coefficient/BBP700) are acquired from the Euro-Argo float network.

A primary challenge in marine remote sensing and multi-source data fusion is the spatial and temporal mismatch between satellite/float observations and shipborne sampling stations. To address this issue, a 3D Inverse Distance Weighting (3D-IDW) interpolation algorithm is introduced. The equations are as follows:

$$V_P = \frac{\sum_{i=1}^{k} w_i V_i}{\sum_{i=1}^{k} w_i} \tag{1}$$

$$w_i = \frac{1}{d\left(P, O_i\right)^p} \tag{2}$$

Where V_P is the interpolated value of the target point p, k is the number of known observation points involved in interpolation, V_i is the actual observation value of the *i-th* observation point, W_i is the weight of the *i-th* observation point in interpolation, $d(P,O_i)$ is the spatiotemporal distance between the target sampling point P and the *i-th* observation point O, p is a distance decay power function.

This method mathematically weights the spatially scattered Argo profile data based on geographical proximity (longitude and latitude) and temporal closeness (sampling date), ensuring the accurate reconstruction of the corresponding spatiotemporal environmental background field for each specific microplastic sampling event. Then, we divide the interpolated data into training and validation sets, input them into the RF model for training and prediction, calculate the R^2 between the predicted values and the validation set data, and verify the accuracy of the model's predictions. The equation for R^2 is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \tag{3}$$

Where y_i is the true observation value of the *i-th* sample, $\hat{y}_i$ is the model's predicted value for the *i-th* sample, $\overline{y}$ is the mean of all true observations.

To enhance the interpretability of the model's predictions, this study introduces SHAP (SHapley Additive exPlanations) analysis. Initially, a Random Forest regression model is trained using the fused multi-source dataset to predict microplastic concentrations via decision tree ensembles and a majority voting/averaging mechanism. The model's outputs are subsequently fed into the SHAP explainer. By calculating the marginal contribution of each feature, this framework not only explicitly quantifies the feature importance of various environmental variables but also extracts the underlying non-linear physical mechanisms driving microplastic dispersion. The overall workflow is shown in Figure 2.

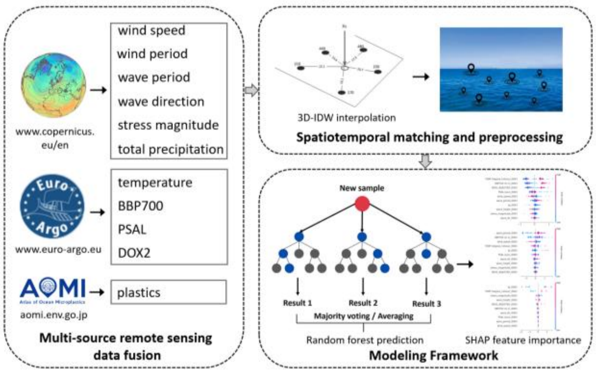


Figure 2. Workflow

4. Results

4.1 Model Validation and Performance Analysis

Figure 3 illustrates the spatial distribution of both the Random Forest (RF) predicted points and the in-situ observation points, with point colors denoting microplastic concentrations. The 2018 sampling points exhibited the most extensive spatial coverage, stretching from the coastal waters of Japan to the low-latitude open ocean. As evident from the model prediction map Figure3(b), the RF model demonstrated reliable predictive accuracy, capturing a spatial gradient consistent with the ground truth observations Figure 3(a). Specifically, highly polluted areas near the southern coast of Japan (the Kuroshio Current region) are highlighted by orange-to-red points, whereas the colors gradually transition to low-concentration greens and blues moving southward into the deep, open ocean further from the continent.

The 2019 sampling points displayed a distinct linear trajectory characteristic of ship tracks. The model successfully predicted the high-concentration zones (yellow/orange) near the Japanese coast. However, for the low-concentration dark blue points near 24°N, 140°E, the model yielded light blue predictions, indicating an overestimation. This reflects the inherent "regression to the mean" phenomenon typical of Random Forest algorithms. The RF model tends to be conservative when predicting extremely low anomalous values, causing a few data points to deviate from the 1:1 diagonal line. Overall, however, the model performed robustly and successfully captured the broader spatial variation trends of microplastics.

The reduced number of data points in 2020 is likely attributable to severe logistical constraints during field surveys, which confined all test sites to nearshore waters. The observation of highly localized concentration hotspots near the coast further corroborates our SHAP analysis conclusion: elevated sea surface temperatures increased water vapor generation, and the resulting intense precipitation substantially flushed terrestrial plastic debris into adjacent coastal waters. Due to these spatial sampling constraints, the microplastic concentrations in 2020 exhibited more geographically homogeneous patterns. Consequently, when evaluated on test data sharing this localized spatial distribution, the model achieved higher predictive accuracy for 2020 than for 2018 and 2019.

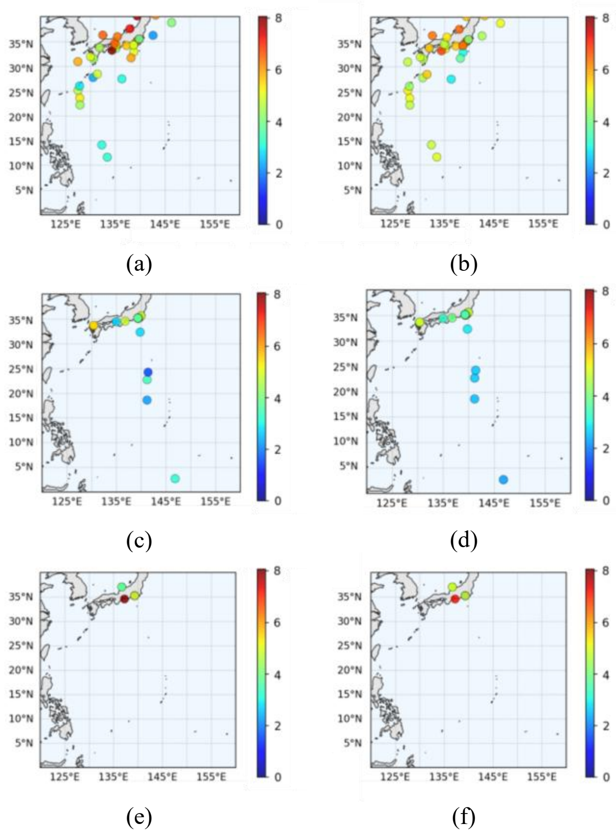


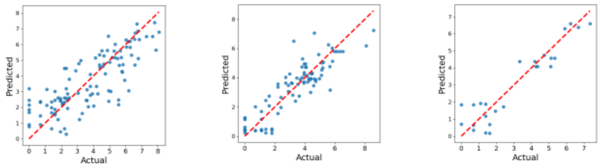
Figure 3. Spatial distribution map of sampling points and prediction points: (a) and (b) are for2018; (c) and (d) are for 2019; (e) and (f) are for 2020.

As seen in Table 1 and Figure 4, the variations observed in the independent test set metrics (R^2 scores and MAE) perfectly corroborate the evolution of the physical mechanisms discussed previously. In 2018, the moderate test set R^2 scores (0.6380) and relatively high MAE (1.0044) reflect the inherent difficulty in predicting microplastic distribution when it is governed by submesoscale hydrological features and biological activities. In 2019, the remarkably high consistency between the OOB R^2 scores (0.7637) and the test set R^2 scores (0.7670) confirms that the prevailing swell-driven dispersion mechanism operated on a broader, more deterministic meteorological scale, thereby endowing the model with superior generalization capabilities.

In 2020, an anomalously high test set R^2 scores (0.8761) and a low MAE (0.6452) were observed. This is likely a statistical artifact primarily resulting from the restricted spatial coverage and reduced sample size caused by pandemic-related limitations. Because the constrained test set in 2020 predominantly captured the strong and homogeneous signals of extreme terrestrial flushing events triggered by typhoons, the data became mathematically easier to fit. Nevertheless, the stable OOB score (0.7473) in 2020 reinforces the conclusion that, despite data scarcity, the underlying model remains well-calibrated and grounded in a solid physical foundation.

Table 1. Model performance for different years

Year	OOB R^2	Testset R^2	Testset MAE
2018	0.7314	0.6380	1.0044
2019	0.7637	0.7670	0.7607
2020	0.7473	0.8761	0.6452



(a) (b) (c)

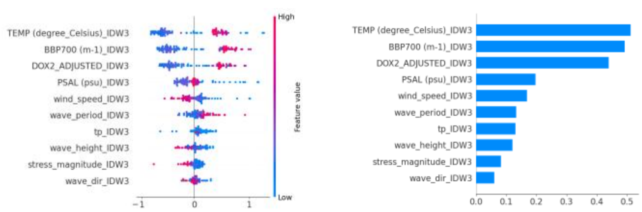
Figure 4. Regression results obtained using the Random Forest (our prediction based on ERA5 and ARGO) and the ground true (AOMI) in the test region: (a) 2018; (b) 2019; (c) 2020.

4.2 SHAP Analysis

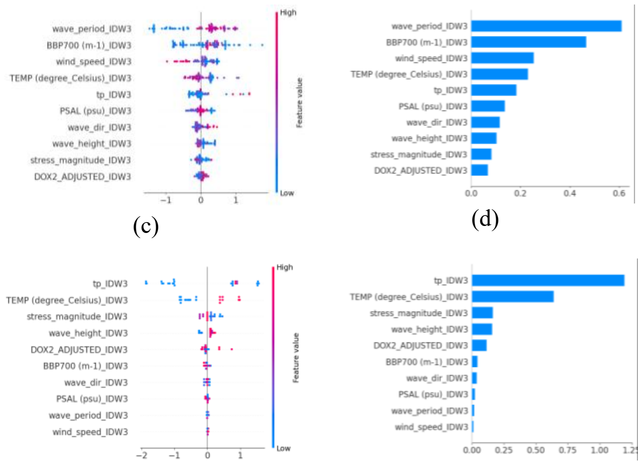
4.2.1 Analysis of SHAP Feature Importance: In 2018, as seen in Figure 5(a) and 5(b), atmospheric forcing (wind, waves, rain) had a relatively weak impact on microplastic distribution, which was primarily controlled by internal oceanic processes. Microplastics accumulated within a water mass characterized by being warm, fresh, low-oxygen, and high-biomass; the formation of this water mass is closely linked to Kuroshio branches and shelf–basin interactions. Microplastics may have entered via the Luzon Strait and rivers on the East China Sea shelf, and were subsequently transported into the study area under the combined influence of secondary Kuroshio circulations and topography-induced upwelling; biofouling further promoted the aggregation and sinking of microplastics.

In 2019, as seen in Figure 5(c) and 5(d), wave period emerged as the most influential factor, indicating that long-period swells from the open ocean (rather than local wind waves) played a dominant role. The open topography of this region favors the propagation and energy retention of remotely generated swells, whose sustained vertical mixing energy significantly affects the vertical distribution of microplastics. The SHAP values for wind speed were negative (higher wind speeds corresponding to lower concentrations), suggesting that while strong winds can enhance turbulent mixing, their primary effect here was to promote horizontal dispersion, leading to reduced microplastic concentrations over gently sloping terrains.

In 2020, as seen in Figure 5(e) and 5(f), precipitation became the dominant factor, closely tied to the frequent typhoon events that year. As a common typhoon passageway, the study area experienced heavy rainfall and typhoon systems that flushed large amounts of accumulated plastic debris from southwestern Japan, the Ryukyu Islands, and the East China Sea coast into the ocean, causing a sharp increase in total microplastics. Typhoon-induced intense turbulent mixing was particularly pronounced over complex topography along the continental slope and around islands, rapidly homogenizing newly introduced microplastics in the surface layer. From 2018 to 2020, microplastic controls shifted from Kuroshio-driven internal processes to open-ocean swells and finally typhoon-related precipitation and mixing that rapidly injected and redistributed coastal plastics.



(a) (b)



(c) (d)

(e) (f)

Figure 5. SHAP summary and feature importance: (a) and (b) are for 2018; (c) and (d) are for 2019; (e) and (f) are for 2020.

4.2.2 Analysis of Single-Feature SHAP Dependence: In 2018, as seen in Figure 6(a) and 6(b), temperature exhibited an extremely high positive peak around 24.5°C, followed by a deep negative trough between 24.55°C and 24.65°C, and a return to stable positive values above 24.7°C. This pattern suggests that microplastics tend to accumulate at frontal zones where distinct water masses converge. The high SHAP values at 24.5°C and > 24.7°C represent the two boundaries of the front, characterized by massive debris accumulation. Conversely, the intermediate range (24.55°C–24.65°C) likely corresponds to a specific rapid and divergent current, preventing microplastics from lingering. Regarding dissolved oxygen (DOX2), values below 215 μmol/kg yielded predominantly negative SHAP values (suppressing accumulation), reaching a minimum between 210 and 215 μmol/kg. When DOX2 exceeded 215 μmol/kg, SHAP values surged and remained at high positive levels (promoting accumulation). High-oxygen waters (DOX2 > 215 μmol/kg) typically indicate well-ventilated Subtropical Mode Water or specific Kuroshio branches that have recently subducted from the surface. The physical characteristics of downwelling and convergence inherent in these water masses transport surface microplastics into the subsurface layer.

In 2019, as seen in Figure 6(b) and 6(c), wave period significantly influenced microplastic distribution: short periods (< 4.5 s) yielded pronounced negative SHAP values, whereas long periods (> 5.5 s) showed significant positive SHAP values. Short-period wind waves, generated by local instantaneous wind fields, possess high wave steepness and are highly susceptible to wave breaking and near-surface turbulence. This intense vertical mixing process drives floating microplastics into deeper waters, leading to a sharp decline in surface sampling concentrations. In contrast, long-period swells, propagating from distant storms, feature smooth wave surfaces, minimal breaking, and substantial Stokes drift, facilitating the surface transport and accumulation of microplastics.

The particulate backscattering coefficient (BBP700) demonstrated a non-linear impact: extremely low values (< 0.0008 m^{-1}) corresponded to exceptionally high positive SHAP values; intermediate values (around 0.0010 m^{-1}) fell into a deep negative trough; and high values (> 0.0014 m^{-1}) recovered to mild positive values. The high accumulation at very low BBP700 indicates nutrient-depleted, low-biomass regions that act as "ocean garbage patches" driven by hydrodynamic convergence. The negative contribution at intermediate values represents biologically active zones, where abundant biomass leads to rapid biofouling on microplastic surfaces, increasing their density and causing them to sink out of the surface layer. The positive contribution at high values corresponds to highly turbid coastal waters, where intense terrestrial runoff discharges both abundant natural suspended particles and massive amounts of urban microplastic waste directly into the sea.

In 2020, as seen in Figure 6(e) and 6(f), when total precipitation (tp) was 0 mm, SHAP values fluctuated, indicating that microplastic concentrations were governed by other environmental factors during dry periods. However, when tp > 0 mm, all SHAP values turned positive (concentrated between 0.5 and 1.0). This strongly demonstrates that the terrestrial flushing effect triggered by rainfall overshadowed conventional ocean hydrodynamics. Furthermore, temperatures < 26.8°C yielded exclusively negative SHAP values, while temperatures > 27.3°C resulted in entirely positive SHAP values. High sea temperatures lead to higher concentrations of microplastics, which are typically associated with frequent convective weather events under high sea temperatures resulting in higher rainfall intensity and frequency.

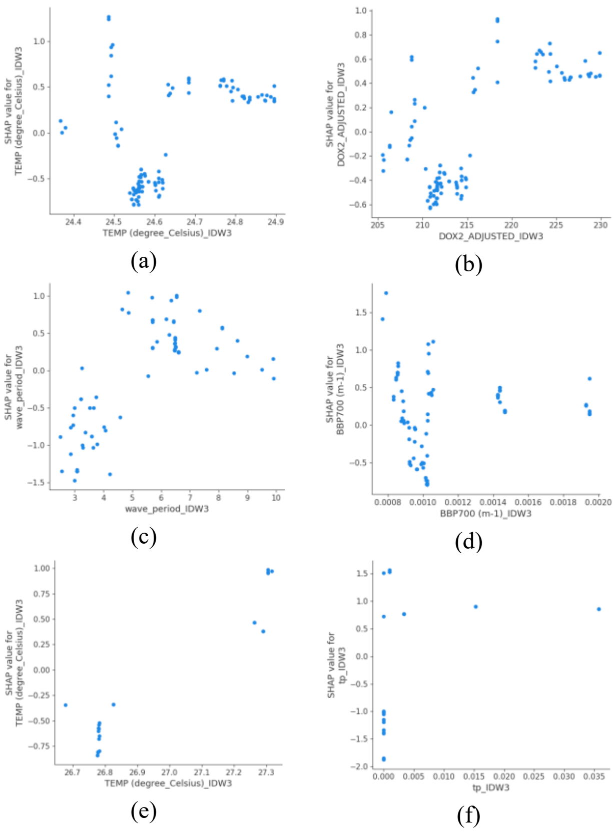


Figure 6. Analysis of Single-Feature SHAP Dependence: (a) and (b) are for 2018; (c) and (d) are for 2019; (e) and (f) are for 2020.

4.3 Analysis of environmental impact factors

In this section, we analyze the dominant environmental factors across different years to investigate their influence on microplastic concentrations, thereby providing a physicochemical explanation for the observed variations in microplastic abundance.

Figure 7(a) illustrates the variations in microplastic concentrations. The violin plot for 2018 is relatively broad and continuous, while for 2019—a year dominated by wave period—it exhibits a pear-like shape. The medians for both years are approximately 3.5 (in log scale). This indicates that under typical hydrological or wave-forcing conditions, the spatial distribution of surface microplastics transitions relatively gradually. The distribution encompasses clean oceanic background regions (low values) as well as polluted zones slowly accumulated by ocean currents and fronts, overall presenting a continuous gradient. Conversely, 2020 displays a distinct bimodal distribution. The high-concentration peak represents sampling points located within terrestrial rainstorm runoff plumes; typhoons flushed concentrated solid waste into the ocean, leading to dramatic surges in localized surface concentrations. The low-concentration peak corresponds to sampling stations that missed these plumes but were nonetheless subjected to severe winds and waves. Intense wind-driven vertical mixing forced the sparsely distributed floating plastics into deeper waters, thereby rendering the sea surface exceptionally clean.

Figure 7(b) presents the variations in sea temperature within the study area from 2018 to 2020. In 2018, sea temperatures were concentrated around 25°C. In 2019, the temperature range spanned from 18°C to 26°C, with a median significantly lower than those of 2018 and 2020. In 2020, sea temperatures were anomalously high, primarily concentrating between 26°C and 29°C.

Because the sampling vessel in 2018 navigated primarily within an exceptionally stable warm water mass, the hydrological conditions remained steady. Consequently, temperature and dissolved oxygen emerged as the dominant factors influencing microplastic concentrations. In contrast, the substantial rise in average sea temperatures in 2020 triggered a marked increase in severe convective weather events, which transported massive volumes of water vapor to terrestrial regions. In July 2020, Japan experienced record-breaking prolonged heavy rainfall, flushing vast quantities of terrestrial microplastics into the ocean. Therefore, precipitation and sea temperature became the dominant drivers of microplastic concentrations in 2020.

Figure 7(c) illustrates the variations in wave period. The median and overall distribution of the wave period in 2019 were notably higher than those in 2018 and 2020. Furthermore, extreme outliers corresponding to long-period swells of nearly 12 seconds were observed in 2019. This indicates that the marine region experienced more frequent and pronounced long-period swells during that year. These long-period swells provided robust Stokes drift, which dominated the aggregation of microplastics in 2019.

Figure 7(d) presents the annual distribution of precipitation. A counterintuitive phenomenon can be observed in the total precipitation distribution: although 2020 was profoundly affected by extreme weather events and terrestrial runoff, the absolute precipitation recorded at the offshore sampling coordinates was extremely low (<1 mm). In fact, the highest instantaneous rainfall was recorded in 2018. This anomaly suggests that, for safety reasons, the survey vessels proactively avoided the centers of extreme storms. Therefore, the trace amounts of precipitation detected offshore in 2020 do not represent the actual magnitude of terrestrial runoff; rather, they serve as a highly sensitive meteorological proxy. These trace amounts indicate the outer peripheries of large cyclonic systems—the very systems that triggered the severe inland rainfall responsible for flushing massive amounts of terrestrial debris into the ocean.

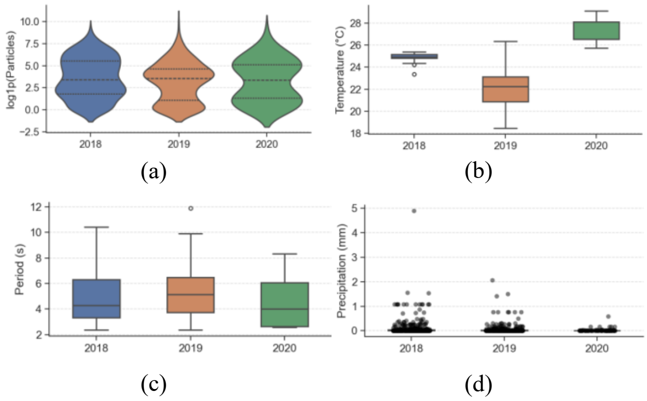


Figure 7. Environmental indicator distribution map: (a) microplastic concentration (b) sea temperature; (c) wave period (d) total precipitation

5. Conclusion

In this study, we developed an interpretable machine learning framework to monitor and analyze the dispersion mechanisms of marine microplastics in the West Pacific from 2018 to 2020. By profoundly integrating AOMI in-situ observations, ERA5 meteorological reanalysis, and Euro-Argo subsurface profile data using 3D-IDW spatiotemporal interpolation algorithm, we successfully constructed and rigorously validated a robust RF predictive model. Crucially, the incorporation of SHAP analysis overcame the inherent "black-box" limitations of traditional ensemble learning, enabling the quantitative extraction of non-linear physical mechanisms driving microplastic transport.

Our analyses revealed a distinct, year-by-year regime shift in the dominant environmental drivers. In 2018, microplastic distributions were primarily governed by stable hydrographic and biological conditions, with particles accumulating in warm, low-oxygen, and high-biomass water masses associated with Kuroshio branches. In 2019, dynamic wave forcing took precedence; long-period swells governed surface accumulation via robust Stokes drift, while short-period wind waves enhanced vertical mixing. In 2020, extreme meteorological events overshadowed conventional oceanic hydrodynamics. Anomalously high sea temperatures coupled with record-breaking precipitation (typhoon events) flushed massive volumes of terrestrial plastic debris into nearshore environments, creating highly localized concentration hotspots.

Despite the spatial sampling constraints in 2020 imposed by severe logistical limitations—which resulted in a geographically restricted dataset but correspondingly high model evaluation metrics due to spatial homogeneity—our physics-informed RF model consistently demonstrated reliable predictive accuracy across varying environmental states. Ultimately, this research underscores that marine microplastic transport dynamically transitions between hydrographic–biological dominance and meteorological–physical forcing. Future research will aim to integrate higher-resolution multimodal satellite imagery (e.g., optical and SAR data) and expand the spatiotemporal domain to further refine source-to-sink trajectories, thereby providing critical scientific support for targeted pollution mitigation and coastal resilience planning.

Acknowledgements

This research was funded by the Science and Technology Program of Guangzhou, China (202201011686)

References

Allen, S., Allen, D., Phoenix, V.R., Le Roux, G., Durántez Jiménez, P., Simonneau, A., Binet, S. and Galop, D., 2019. Atmospheric transport and deposition of microplastics in a remote mountain catchment. *Nature Geoscience*, 12(5), 339–344.

Andrady, A.L., 2011. Microplastics in the marine environment. *Marine Pollution Bulletin*, 62(8), 1596–1605.

Barnes, D.K.A., Galgani, F., Thompson, R.C. and Barlaz, M., 2009. Accumulation and fragmentation of plastic debris in global environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1526), 1985–1998.

Chen, O., 2025. AI-Driven Framework Development for Predictive Classification of Microplastic Concentration of Aquatic Systems in the United States. *2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI)*, , 1–6.

Cui, B., Qi, H., Liu, M., Liu, M., Huang, W., Huang, P., Wang, C., Zheng, X., Ke, H. and Cai, M., 2025. Multi-scenario simulation of future marine microplastic distribution under data scarcity: A deep learning approach. *Water Research*, 286, 124233.

Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R. and Law, K.L., 2015. Plastic waste inputs from land into the ocean. *Science*, 347(6223), 768–771.

Khanam, M.M., Uddin, M.K. and Kazi, J.U., 2025. Advances in machine learning for the detection and characterization of microplastics in the environment. *Frontiers in Environmental Science*, 13.

Kim, S. and Kim, D.-H., 2024. Short-term buoyant microplastic transport patterns driven by wave evolution, breaking, and orbital motion in coast. *Marine Pollution Bulletin*, 201, 116248.

Zhang, S., Liu, N., Luo, M., Jiang, T., Chan, T.O., Yau, C.S.T. and Sun, Y., 2023. Downscaling Sentinel-3 Chlorophyll-a Concentration for Inland Lakes Based on Multivariate Analysis and Gradient Boosting Decision Trees Regression. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16, 7850–7865.

Zhang, L., Zhao, W., Yan, R., Yu, X., Barceló, D. and Sui, Q., 2024. Microplastics in different municipal solid waste treatment and disposal systems: Do they pose environmental risks? *Water Research*, 255, 121443.

Zhen, Y., Wang, L., Sun, H. and Liu, C., 2023. Prediction of microplastic abundance in surface water of the ocean and influencing factors based on ensemble learning. *Environmental Pollution*, 331, 121834.

Zhu, Y., Hang Yeung, C. and Lam, E.Y., 2021. Microplastic pollution monitoring with holographic classification and deep learning. *Journal of Physics: Photonics,* 3(2), 024013.