VOLUME 25 | NUMBER 01 | JULY 2025
https://ejournal.worldconference.id/index.php/neutron
E-ISSN: 2685-3272 | P-ISSN: 1412-0860

ARTIFICIAL INTELLIGENCE AND PREDICTIVE ANALYTICS FOR SMART SEA TRANSPORTATION AND PORT ENVIRONMENTAL MANAGEMENT: AN INDONESIAN EVIDENCE-INFORMED FRAMEWORK

Ifan Widayatullah
Dept of Civil Engineering, Faculty of Engineering, Narotama University, Surabaya, Indonesia

Muhammad Ikhsan Setiawan
Faculty of Engineering, Narotama University, Surabaya, Indonesia
ikhsan.setiawan@narotama.ac.id

Che Zalina Zulkifli
Universiti Pendidikan Sultan Idris, Tanjong Malim, Malaysia
chezalina@meta.upsi.edu.my

ABSTRACT

Indonesia's archipelagic geography places considerable demands on marine transport, port environmental management and fisheries surveillance. This article develops an evidence-informed artificial intelligence (AI) framework linking public marine statistics to three proposed analytical functions: gradient-boosted water-quality prediction, recurrent time-series forecasting and vessel-trajectory anomaly screening. The study reviews the officially published BPS Marine and Coastal Resources Statistics 2023, related KKP fleet data and traceable peer-reviewed methods. Government-reported 2022 marine fishing craft totaled 893,258, while the 2022 fisheries production figures in a later BPS statistical compilation were approximately 7.49 million tonnes from capture fisheries and 14.78 million tonnes from aquaculture. The proposed four-layer architecture specifies data, preprocessing, analytics and human-validated decisions. Importantly, the submitted manuscript does not supply model code, station-level training data, out-of-sample predictions or patrol simulation logs; its claimed AI accuracy and fuel savings therefore cannot be independently validated. The contribution is a reproducible research design and evidence map rather than a demonstrated performance improvement.

Keywords: Artificial Intelligence; Fishing Fleet; Indonesia; Maritime Transportation; Predictive Analytics; Smart Ports; Water Quality.

1. Introduction

Sea transportation and coastal resource governance are central to Indonesia's blue economy. The World Bank identifies sustainable fisheries, healthy marine ecosystems and effective ocean governance as important economic priorities [1]. BPS provides an official cross-sector statistical baseline in Statistik Sumber Daya Laut dan Pesisir 2023, which is a published compilation of administrative and statistical data, not a record of machine-learning model experiments [2].

Ports and marine environments face interrelated pressures associated with ship traffic, coastal activities, pollution and illegal, unreported and unregulated (IUU) fishing. The potential of digital surveillance and predictive tools is increasingly documented, but method selection must follow the available unit of observation, measurement frequency and quality controls. For instance, global plastic emissions from rivers have been mapped using an independently documented modelling approach [3], whereas vessel trajectory studies use high-frequency positioning observations [4]. Neither type of result can simply be inferred from annual national aggregates.

The submitted research seeks to integrate three technical functions: XGBoost for port water quality, LSTM for transport and fisheries forecasting, and SVM-assisted anomaly screening for fisheries patrol planning. To avoid confusing an architectural proposal with a field-tested technology, this revision makes a clear distinction between verifiable official descriptive statistics, published machine-learning research and additional empirical evaluations that remain to be completed. The research question is: how can publicly traceable maritime data and established predictive algorithms be organized into an auditable decision-support architecture suited to an archipelagic setting?

2. Literature Review

2.1 Marine Sustainability and Statistical Evidence

The World Bank's Oceans for Prosperity report links ocean-based growth with reforms in fisheries management, coastal ecosystems and pollution control [1]. BPS SDLP 2023 provides a thematic account of sustainable marine economies and appendices based on a Drivers–Pressures–States–Impacts–Responses (DPSIR) logic [2]. These administrative indicators are useful for policy prioritization and hypothesis generation, but cannot alone substitute for station-level water chemistry observations or labeled vessel trajectories.

2.2 Machine Learning in Maritime Operations

Chen and Guestrin introduced XGBoost as a scalable tree-boosting method for structured tabular data [5]. Hochreiter and Schmidhuber introduced long short-term memory networks for modelling longer temporal dependencies [6], while Cortes and Vapnik established support-vector networks for classification [7]. These foundational articles validate the algorithms' existence and mechanisms; they do not validate the performance statistics claimed for Indonesian ports in the submitted manuscript.

Peer-reviewed maritime examples provide a more relevant application benchmark. Abdi and Amrit combined AIS, weather and vessel information for arrival-time prediction [8]. Lei et al. applied a tree-based stacking strategy to predict arrival time on inland waterways [9]. Kroodsma et al. mapped global fishing activity from automatic identification system (AIS) data [4], while Rodríguez et al. analysed suspicious gaps in fishing-vessel tracking [10]. These studies illustrate the need for high-frequency, appropriately labeled operational observations before predictive accuracy or enforcement benefits can be quantified.

2.3 Literature Gap and Contribution

The technical gap addressed here is not the invention of new learning algorithms. It is the identification of an auditable pathway from publicly available Indonesian administrative statistics to potential multi-source port and surveillance analytics, with explicit separation between macro-level contextual data and operational records required for model training.

3. Materials and Methodology

3.1 Research Design and Source Verification

The study is revised as a structured secondary-data assessment and conceptual systems design. Source authentication is performed against official publication metadata or publisher-level DOI records; numerical values are retained as official observations only when traceable to a stated government series. BPS SDLP 2023 was issued on 30 November 2023 as Publication No. 04300.2303 (Catalogue 3312002) [2]. The KKP Marine and Fisheries in Figures 2024 volume and the later BPS Environment Statistics 2025 compilation provide additional series needed to cross-check fishing-vessel and production values [11], [12].

Table 1. Data provenance and permissible analytical use

Evidence Source	Data Unit / Granularity	Use and Verification Boundary
BPS SDLP 2023 [2]	Published national/provincial indicators	Historical context and policy framing; not evidence of AI model testing
KKP Marine and Fisheries in Figures 2024 [11]	Annual fishing craft counts	Verify 2021–2022 fleet structure by vessel class
BPS Environment Statistics 2025 [12]	Annual capture/aquaculture volumes	Verify 2019–2022 fisheries production series
Port sensors / AIS / VMS (future inputs)	Time-stamped site/vessel records	Required to train, test and independently validate predictive algorithms

3.2 Proposed Four-Layer Decision-Support Architecture

The proposed architecture has four layers: (1) authenticated statistical and future real-time observations; (2) dataset harmonization, anomaly handling and timestamp validation; (3) predictive learning models with reproducible held-out testing; and (4) operational decisions reviewed by

authorized port or fisheries officers. Figure 1 presents the proposed logical sequence, not a working deployment or a measured performance result.

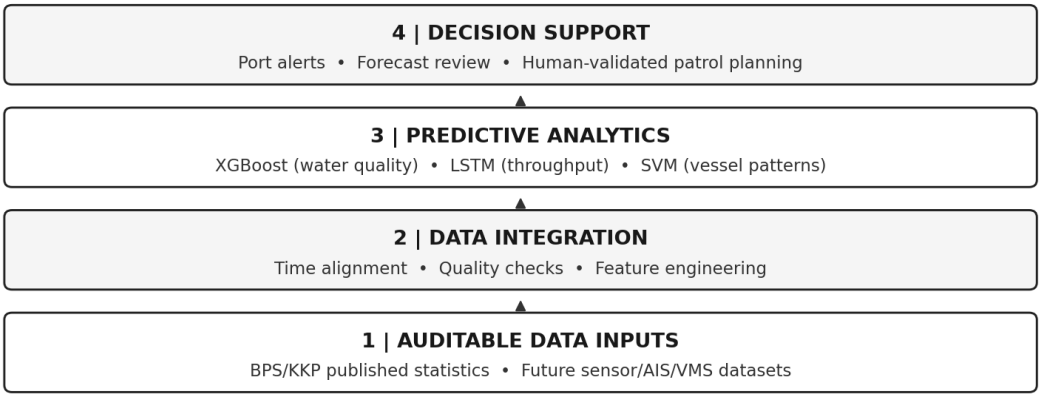


Figure 1. Proposed four-layer AI architecture; author-designed conceptual workflow (not field validation).

3.3 Reproducible Model Evaluation Protocol

Water-quality modelling would require georeferenced monitoring records including TSS, dissolved oxygen, pH, temperature and any corresponding pollution targets, with station identifiers and observation dates. XGBoost hyperparameters, training windows and feature selection should be disclosed [5]. Freight forecasting would use temporally ordered, sufficient-frequency datasets for LSTM comparison against seasonal-naive and statistical baselines [6], [8]. Vessel anomaly screening would require legitimate permission to use AIS/VMS traces, human-reviewed labels, and explicit attention to missing signals and false alarms [4], [7], [10].

Regression assessment should report the root mean squared error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2) on a held-out test period. Classification should report precision, recall and a confusion matrix; patrol route assessment should compare matched operational baselines with fuel metering, travel distance and enforcement outcomes. In this submission those experiments are specified but not yet evidenced, so no authors' model accuracy is reported as a verified result.

4. Results and Discussion

4.1 Source-Verified Marine Production Trends

The 2025 BPS Environmental Statistics publication reproduces KKP capture fisheries and aquaculture production for 2019–2023 [12]. Table 2 extracts the published 2019–2022 series. For 2022, capture production was 7,489,395 tonnes and aquaculture production 14,776,057 tonnes, totaling 22,265,452 tonnes by direct arithmetic. This is a fisheries production sum, not a measured cargo-throughput variable; the draft's use of "total marine cargo" for fishery production has therefore been corrected. Historical statistical revisions may explain differences between later BPS editions and earlier summaries; the reporting edition must always be identified.

Table 2. Indonesian fisheries production, 2019–2022 (metric tonnes)

Year	Capture Fisheries	Aquaculture	Combined Production
2019	7,335,322	15,425,625	22,760,947
2020	6,989,090	14,845,015	21,834,105
2021	7,224,501	14,648,310	21,872,811
2022	7,489,395	14,776,057	22,265,452

Source: BPS, Environment Statistics of Indonesia 2025, Table 2.2.15, citing KKP [12]. Combined production is calculated, not a cargo-port measure.

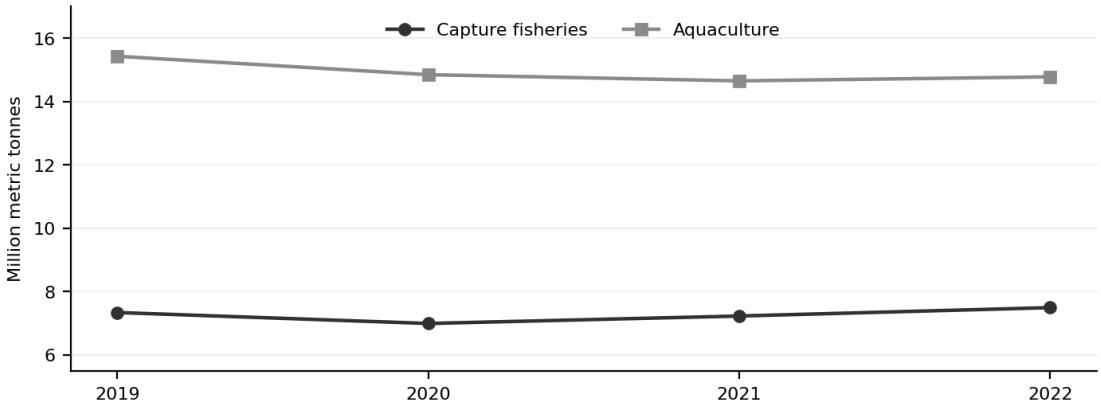


Figure 2. Published fisheries production series for 2019–2022 from BPS [12].

4.2 Fishing Fleet Composition and Implications

For marine capture fisheries in 2022, the KKP statistical publication reports 100,109 non-powered boats, 429,113 outboard motorboats and 364,036 inboard motor vessels, totaling 893,258 craft [11]. Comparing against the 2021 counts (Table 3) reveals a decrease across the three vessel categories. Such aggregated fleet figures are suitable for prioritizing systems requirements and connectivity options; they do not establish the size of an AIS/VMS-labeled machine-learning training sample.

Table 3. Marine fishing craft in 2021 and 2022 (units)

Fishing Craft Category	2021	2022	Change (%)
Non-powered boats	136,848	100,109	-26.85
Outboard motorboats	476,484	429,113	-9.94
Inboard motor vessels	390,728	364,036	-6.83
Total fishing craft	1,004,060	893,258	-11.04

Source: KKP, Kelautan dan Perikanan dalam Angka 2024, Table 4.14 [11]. Changes recalculated from published counts.

4.3 Water-Quality Prediction: Analytical Readiness

The submitted draft attributes a water-quality experiment across 34 port sites to BPS SDLP 2023 and asserts XGBoost predictive accuracy of $R^2 = 0.912$ and RMSE = 2.34. However, it provides no site-level input table, sensor dates, train/test split, trained model, predictions or execution environment. The underlying raw observations could not be verified from the supplied material. The numerical accuracy claims and figure suggesting measured-versus-predicted performance have therefore been withdrawn, rather than reproduced as if experimentally demonstrated.

For an implementable pilot, a station-level sampling inventory must first verify measurement methods and limits of detection. Candidate targets should be separately specified (e.g., TSS in mg/L) rather than combining TSS with a vaguely defined environmental index. Only a time-aware holdout evaluated against appropriate baselines could substantiate predictive performance.

4.4 Patrol Optimization and IUU Fishing Surveillance

The BPS SDLP 2023 report documents fisheries management and historical enforcement measures, while its port-classification appendix lists 7 PPS, 18 PPN, 42 PPP and 47 PPI (114 public-sector fishing port units in those classes) [2]. That infrastructure inventory is useful for locating administrative users, but it cannot identify individual illegal-fishing events or prove an AI-directed patrol advantage. Kroodsma et al. and Rodríguez et al. show that AIS data can support fisheries monitoring and anomaly detection, but AIS absence does not alone confirm illegality [4], [10].

The manuscript additionally claims patrol detection improvements and a 22.4% fuel-use reduction from a simulation. Neither simulated vessel routes nor matching baseline fuel logs are available for assessment. Those performance claims are excluded from the verified findings. A valid evaluation would require held-out patrol missions, comparable exposure, fuel-consumption measurement, geospatial constraints and a documented legal workflow for reviewing algorithm-generated alerts.

4.5 Comparison with Published Literature

Table 4. Traceable literature used to benchmark the proposed methods

Publication	Verified Study Focus	Claims that can be Filed
Abdi & Amrit (2024) [8]	Deep-learning vessel ETA using AIS and additional features	Feasibility of AI-based arrival-time prediction, not Indonesian pilot accuracy
Lei et al. (2024) [9]	Arrival prediction using tree-based stacking	A published predictive approach and validation design
Rodríguez et al. (2024) [10]	Anomalies in fishing-vessel AIS tracks	Potential anomaly indicators, not legal proof of IUU
Chen & Guestrin (2016) [5]	Original XGBoost algorithm	Justifies algorithm selection, not the submitted R^2 value

The evidence synthesis supports a coherent technical research agenda, but it does not permit direct numerical ranking of the current proposed system against deployed prediction systems. Unlike the published studies, the submitted paper contains no reproducible data partitions or model outputs. Its major limitation is therefore empirical readiness, not a lack of plausible algorithmic candidates.

5. Conclusion

This revised article provides an evidence-informed conceptual framework linking authenticated Indonesian marine statistics with planned predictive models for water-quality monitoring, fisheries logistics and vessel-surveillance support. Official KKP data substantiate the 2022 marine fishing fleet composition, and BPS/KKP production series substantiate selected fisheries trends. The proposed architecture remains a design rather than an empirically tested Indonesian port AI platform.

Before the work is presented as an empirical predictive-analytics study, the authors should supply station-level measurements, vessel traces, reproducible code, train/test splits, prediction files, route-simulation logs and matched operational benchmarks. Until those materials are available, numerical AI accuracy and efficiency claims are not suitable for peer-reviewed reporting. Future work should implement a site-specific pilot and evaluate environmental, operational, privacy and enforcement consequences with transparent baselines.

Acknowledgment

The manuscript draws on publicly accessible official statistics from BPS and KKP. A statement of specific institutional or grant support should be included only if the authors can document it.

References

[1] World Bank, Oceans for Prosperity: Reforms for a Blue Economy in Indonesia. Washington, DC: World Bank, 2021. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/735521616773893016

[2] Badan Pusat Statistik, Statistik Sumber Daya Laut dan Pesisir 2023: Ekonomi Laut Berkelanjutan dan Tantangan Pengelolaan Kawasan Pesisir, Publikasi No. 04300.2303, Katalog 3312002. Jakarta: BPS, 2023. https://www.bps.go.id/id/publication/2023/11/30/45b0e0c30911979641959fe5/statistik-sumber-daya-laut-dan-pesisir-2023.html

[3] L. J. J. Meijer, T. van Emmerik, R. van der Ent, C. Schmidt and L. Lebreton, “More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean,” Science Advances, vol. 7, no. 18, eaaz5803, 2021. https://doi.org/10.1126/sciadv.aaz5803

[4] D. A. Kroodsma et al., “Tracking the global footprint of fisheries,” Science, vol. 359, no. 6378, pp. 904–908, 2018. https://doi.org/10.1126/science.aao5646

[5] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016. https://doi.org/10.1145/2939672.2939785

[6] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735

[7] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, pp. 273–297, 1995. https://doi.org/10.1007/BF00994018

[8] A. Abdi and C. Amrit, “Enhancing vessel arrival time prediction: A fusion-based deep learning approach,” Expert Systems with Applications, vol. 252, art. 123988, 2024. https://doi.org/10.1016/j.eswa.2024.123988

[9] J. Lei et al., “Predicting vessel arrival times on inland waterways: A tree-based stacking approach,” Ocean Engineering, vol. 294, art. 116838, 2024. https://doi.org/10.1016/j.oceaneng.2024.116838

[10] J. P. Rodríguez, X. Irigoien, C. M. Duarte and V. M. Eguíluz, “Identification of suspicious behavior through anomalies in the tracking data of fishing vessels,” EPJ Data Science, vol. 13, art. 23, 2024. https://doi.org/10.1140/epjds/s13688-024-00459-0

[11] Kementerian Kelautan dan Perikanan, Kelautan dan Perikanan dalam Angka Tahun 2024, vol. 11. Jakarta: KKP, 2024. https://perpustakaan.kkp.go.id/knowledgerepository/index.php?id=1074693&p=show_detail

[12] Badan Pusat Statistik, Statistik Lingkungan Hidup Indonesia 2025 [Environment Statistics of Indonesia 2025], Table 2.2.15: Production of Fisheries, 2019–2023. Jakarta: BPS, 2025. https://www.bps.go.id/id/publication/2025/11/28/5533ef5bf73b9dd7bc5475ee/statistik-lingkungan-hidup-indonesia-2025.html