IJISEM
INTERNATIONAL JOURNAL OF INNOVATIONS IN SCIENCE ENGINEERING AND MANAGEMENT

International Journal of Innovations In Science Engineering And Management

Check for updates

OPEN ACCESS

Volume: 5

Issue: 4

Month: November

Year: 2026

ISSN: 2583-7117

Published: 21.11.2026

Citation:

Dr. Shubhangi Kamble, Sara Ansingkar, Saloni Tiwari, Sairaj Patil, Darshan Shete "OceanTwin AI: A Maritime Intelligence Platform for Smart Ports" International Journal of Innovations In Science Engineering And Management, vol.5, no.4, 2026, pp. 349-353.

DOI:

10.69968/ijisem.2026v5i4349-354

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

OceanTwin AI: A Maritime Intelligence Platform for Smart Ports

Dr. Shubhangi Kamble[1], Sara Ansingkar[1], Saloni Tiwari[1], Sairaj Patil[1], Darshan Shete[1]

[1] *Computer Science–Artificial Intelligence Vishwakarma Institute of Technology Pune, India*

Abstract

The current maritime operations are being affected more and more by such factors as port congestion, weather disruption, sediments, fuel efficiency, and sustainability guidelines. It has become a problem in current maritime operations to include in their system the technologies of telemetry simulation, machine learning, and predictive analytics because of different modules that the current maritime systems contain. In this paper, we propose OceanTwin AI, the framework of maritime intelligence, which integrates the mentioned technologies in maritime operations through digital twin concept. The architecture of OceanTwin AI framework is 3 tiered where the frontend, API gateway, and machine learning services are implemented in React, Node JS and Python respectively. Our framework includes such functionality as vessel tracking, port congestion, sediment risks, fuel efficiency, sustainability, and alerts. We performed evaluation of our framework using 3 physics inspired synthetic datasets and predictive models including Linear regression, Random forest and XGBoost. Furthermore, we evaluated our framework using 5-fold cross-validation and Scenario based digital twin simulations.

Keywords: *Maritime Intelligence; Smart Ports; Digital Twin; Machine Learning; Predictive Analytics; Telemetry Simulation; Sustainability Analytics; Explainable AI; Decision Support Systems.*

INTRODUCTION

- Modern maritime transportation covers about 80% of all world trading [4] and is an integral part of international logistics. At the same time, there are such problems that appear in modern shipping as congestion of ships, weather conditions, fuel consumption, sediments, and environmental standards. Such problems may result in decreased effectiveness of the process and additional expenses.
- Nowadays, due to the advancements in the areas of digital twins and artificial intelligence, it became possible to develop intelligent maritime management systems which will be able to monitor operations, predict disruptions, and help decision-making. However, most of the solutions developed for this field of research are focused on particular aspects and do not provide any integrated environment with visualizations and simulations.
- Thus, the aim of this paper is to propose OceanTwin AI – maritime intelligence platform with telemetry simulation, machine learning services, and scenario analysis based on the concept of digital twin. The proposed platform offers such features as vessel tracking, congestion in the port prediction, sediment risk prediction, fuel consumption prediction, sustainability analysis, and alerting.

Main contributions of this paper include:

- Development of integrated maritime intelligence plat- form with simulation and predictive analytics.
- Comparison of various machine learning models for delay prediction, fuel estimation, and sediment risk assessment.
- Application of explainable AI, cross-validation, and scenario-based digital twin experiments.
- Reproducible prototype showing the possibility of intelligent decision support for smart-port application.

LITERATURE SURVEY

De Bari et al. [1] offered a decoupled-layer architecture for seaport digital twins, stressing the importance of scalable and standardized digital infrastructure. According to their approach, sensor networks, communication infrastructure, and cloud-based data layers are separated, which increases modularity and interoperability. Importantly, the authors stressed that digital twins be-come much more useful if combined with artificial intelligence and used for predictive decision making.

Zeneli and Marinova [2] researched the use of digital twinning in port logistics and proved that the use of virtual replications of the physical system is efficient for simulation, monitoring, and forecast-ing of performance. The authors found out that the digital twin platform can provide predictive information several hours ahead of time.

The use of artificial intelligence and IoT technologies for digital twins was also explored by Badea et al. [3]. They have shown that continuous sensor data analysis and automatic data processing allow increasing visibility and the effectiveness of decisions made on the basis of this information. Fragmented data architecture was mentioned as one of the main obstacles of implementing largescale port digital twins.

The machine learning techniques have been used to solve a variety of maritime problems. Predictive models for estimating vessel delays were suggested by Heilig and Vo [5], and Zhang et al. [6] researched the problem of predicting port congestion based on operational maritime data.

Recently, AI solutions for sustainability optimization in maritime transportation have been developed. For instance, machine learning models have been used for the prediction of fuel consumption and emission.

SYSTEM ARCHITECTURE

The OceanTwin AI architecture comprises a three-level approach to incorporating telemetry simulation, machine learning, and maritime decision support into one platform. The architecture includes the presentation layer, the application layer, and the machine learning inferencing layer.

Overall Architecture

The presentation layer is developed with React technology and offers a dashboard that helps to track vessels, port activities, analyze the routes, sediment risks, and sustain-ability performance. User's requests go via the Node.js API gateway that deals with authentication, authorization, data orchestration, and communication with machine learning services. The illustration of the architecture of OceanTwin AI can be seen in Figure 1.

Intelligence layer uses Python microservices built on Flask framework, which hosts machine learning algorithms for prediction of vessel delays, fuel efficiency estimation, and sediment risk assessment. The telemetry simulation engine continuously refreshes vessel and port states, and a digital twin inspired simulation engine allows testing disruption scenarios.

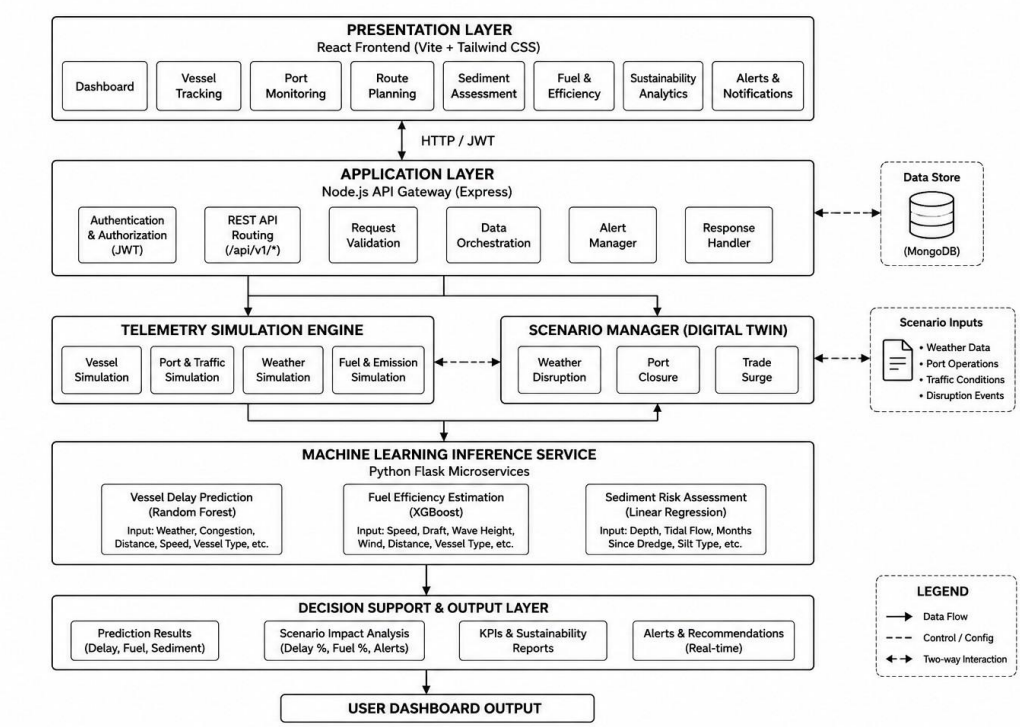


Figure 1. OceanTwin AI three-tier architecture.

Frontend Layer

The frontend part is developed with React, Vite, and Tailwind CSS technologies. The frontend is a command center offering the visualized vessel activity, port activity, predictive analytics, sustainability data, and other alerts.

Backend Layer

The backend system is developed on Node.js and Express. The backend works as an API gateway that will perform requests, authentication based on JSON Web Token (JWT), role based authorization, and interfacing among frontend components and machine learning services. Telemetry updates, alerting, and scenarios execution are performed on the backend side.

Inference Layer

The ML inference layer is developed on Python, Scikit-learn and XGBoost. This layer provides prediction capabilities for vessel delay time prediction, fuel consumption prediction and sediment prediction models. These models can be used through REST APIs from the frontend via the gateway API.

Telemetry and Digital Twin Engine

A telemetry simulation engine updates the status of vessels, fuel consumption, emission and port environment data periodically. The digital twin inspired engine facilitates scenario-based experiment through weather disruption, port shutdown and cargo rush.

METHODOLOGY

This research work will concentrate on building an OceanTwin AI, which will be an intelligent maritime platform built to perform maritime operations by analyzing, visualizing, and optimizing the same using telemetry simulation, machine learning services, and scenario analysis, which is based on digital twin concept, in three layers architecture. The three-layer architecture consists of a presentation layer, an application gateway layer, and machine learning inference layer.

Telemetry Simulation

The system works on the basis of simulation in the back-end that constantly updates the vessel information at set intervals of time. In each update session, all vessels are taken one step closer towards their respective destinations, along with updating the amount of fuel used and their emission status along with the operational status of the vessel. Alongside this, the port queue information and the berthing status are updated according to the live situation. When the vessel is nearing its destination, the arrival notification is automatically generated.

Digital Twin Scenarios

For assessing operational resilience, the platform uses three pre-set disruption scenarios. These scenarios are aimed at creating specific disruptions and their effects on the system. The first one is related to the weather disruption scenario where there is reduced speed of vessels and weather alert generation. Port closure disruption increases congestion and increased queue times at selected ports. The cargo surge disruption simulates an increase in the cargo demand and loading delay. The activation of the disruption scenario has cascading effects on the entire platform.

Machine Learning Models

Three machine learning models form the intelligence layer of OceanTwin AI.

1) **Vessel Delay Prediction:** The vessel delay prediction module estimates expected delays using sailing distance, ves-sel speed, weather severity, and port congestion as input features. A Random Forest regression model was selected due to its ability to capture non-linear relationships while maintaining stable predictive performance.
2) **Sediment Risk Assessment:** The sediment risk module predicts sediment accretion risk using channel depth, sedimen-tation indicators, and environmental factors. A Linear Regres-sion model was selected because of its interpretability and suitability for modeling the approximately linear relationships present in the generated dataset.
3) **Fuel Efficiency Prediction:** The fuel efficiency module estimates fuel consumption under varying operating conditions using vessel speed, draft, wave height, wind speed, and weather severity as inputs. An XGBoost regression model was employed due to its strong predictive performance on tabular datasets and ability to model complex interactions among features.

Dataset and Experimental Setup

Three synthetic datasets inspired by physics experiments were created for the purpose of evaluating the machine learning aspects of the platform. The first dataset included 2,000 samples, the second dataset included 1,500 samples, and the third dataset had 5,000 samples. The datasets were randomly divided in the proportion 80% training to 20% testing using a constant random state (random state = 42).

The purpose of creating the datasets was to simulate realistic operating conditions of the maritime industry, but at the same time, to enable controlled experimentation. Holdout validation and 5-fold cross-validation were conducted.

Evaluation Methodology

Evaluation of the models was performed through the use of a range of evaluation metrics such as the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and coefficient of determination

(R^2). For the purpose of comparison in order to evaluate the performance, experiments were performed on three different models such as Linear Regression, Random Forest and XGBoost. Furthermore, feature importance analysis and SHAP-based explainability methodology were used in order to improve model transparency and interpretation of results.

Apart from model evaluation, we also performed the digital twin scenarios experiment in order to understand the impact of disruptions such as weather disruption, port closure, and increased cargo volume on the system. The impact of these disturbances was quantified through specific metrics such as delay time, increased fuel consumption and alert generation.

EXPERIMENTAL RESULTS

The machine learning modules of OceanTwin AI were tested using hold-out test set and 5-Fold Cross Validation techniques. The testing was done on physics-based synthetic maritime data sets created in a manner similar to real-life scenarios in the maritime environment. For measuring the effectiveness of these machine learning modules, several performance metrics were used which included MAE, RMSE, MAPE, and coefficient of determination (R^2). Besides, several comparative experiments were conducted using three different models, Linear Regression, Random Forest, and XGBoost, for comparing their performance with that of the machine learning modules of OceanTwin AI.

Baseline Model Comparison

Figure 2 provides a thorough comparison of the performance of various machine learning models for each of the three prediction tasks. When it comes to the problem of vessel delay prediction, the Random Forest model was the one that provided stable performance along with high interpretability which is valuable for practical application. As for the fuel consumption prediction problem, XGBoost model showed excellent results and was used in the final model. Although the ensemble models provided better scores for the sediment prediction task, the Linear Regression model was selected because of its high interpretability and good fit with linear patterns in the data.

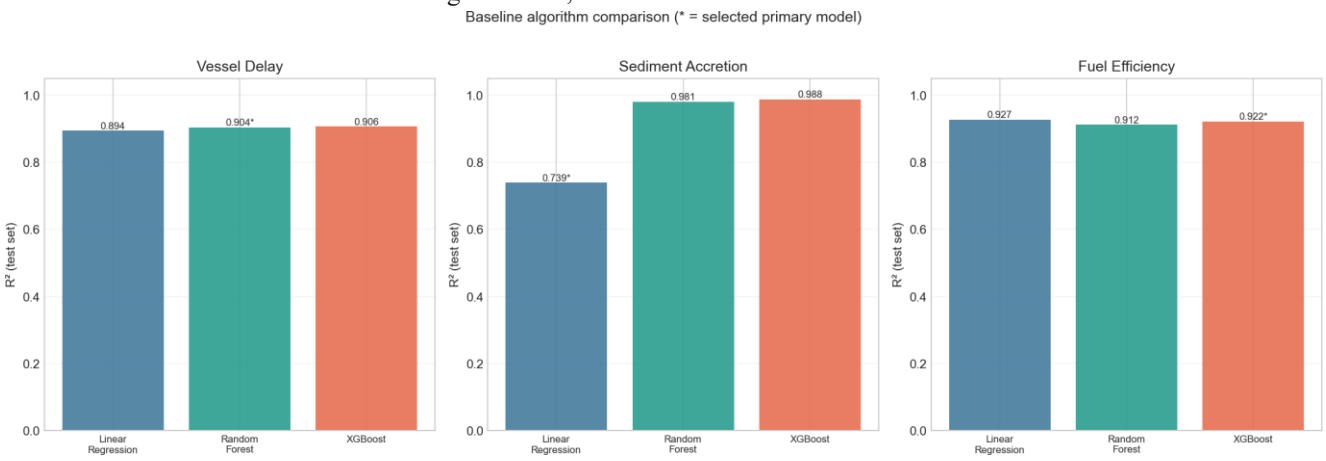


Figure 2. Comparison of Linear Regression, Random Forest, and XGBoost across maritime prediction tasks.

Final Model Performance

Further evaluation was done for the selected machine learning models through testing on held-out test data. The model of fuel efficiency performed better among the three models in terms of accuracy with an R^2 score of 0.922. On the other hand, the model of vessel delays also exhibited good predictive capabilities with an R^2 of 0.904. The model of sediment risks recorded lower R^2 but still had minimal prediction errors and enhanced interpretation. All models performed similarly during five-fold cross-validation.

Table 1 Final Model Performance

Model	MAE	RMSE	MAPE (%)	R2
Delay Prediction	1.82	2.25	17.42	0.904
Sediment Risk	0.15	0.21	21.73	0.739
Fuel Prediction	0.41	0.53	10.53	0.922

Feature Importance and Explainability

To increase the level of transparency of the prediction models, feature importance analysis was conducted. In the case of the vessel delay model, one can see that the feature importance of the congestion variable is 61.9%, which makes it the most important feature. The weather severity variable accounts for 33.6% of the overall feature importance. Such results correspond to the expectations in the maritime industry, as both congestion and weather severity are known factors of vessel delays. The results were confirmed via further SHAP analysis.

Digital Twin Scenario Evaluation

The digital twin experiments were performed with the help of three pre-defined scenarios of possible disruptions in order to analyze the influence of different kinds of disturbances on the process of maritime operations. In case of the weather disruption scenario implementation, there were delays

increasing by 22.7%. As we could see, delays increased by 42.3% during the port closure scenario implementation, which showed a great disturbance in flow. In addition, delays increased by 19.1% because of cargo surge scenario, which meant the influence of sudden increase in the volume of cargo.

Out of all three scenarios, the port closure scenario caused the biggest operational impact and generated the biggest increase in delays as well as the biggest number of alerts. This fact shows the vulnerability of the port system in case of temporary blocking of the entrance to the port.

In conclusion, we can state that such an approach to testing scenarios is quite effective and useful in predicting possible disturbances in advance.

Table 2 Scenario Evaluation Results

Scenario	Delay Increase	Fuel Increase	Alerts
Weather Disruption	22.7%	12.3%	19
Port Closure	42.3%	8.4%	22
Cargo Surge	19.1%	6.3%	5

DISCUSSION

In particular, the fuel efficiency model stood out in terms of being the best performing predictor of the set of models under consideration, which provided the highest predictive performance compared to all other models. Similarly, the vessel delay model demonstrated high prediction capabilities. In spite of the fact that the sediment risk model was associated with the lowest R^2, it provided low error and high interpretability at the same time. With regard to the feature importance analysis, it was found that the most important factors causing vessel delays were port congestion and weather severity.

From the platform perspective, it should be stressed that it successfully incorporated several modules such as vessel tracking, port analysis, sediment risk analysis, route planning, and predictive analytics into one coherent interface. In turn, both the telemetry engine and the digital twin scenarios worked consistently for all of these modules and there was a fallback solution in case of machine learning services unavailability.

It is necessary to mention that the machine learning models used were estimated based on the physics-inspired synthetic datasets and not on real operational maritime data. Moreover, the digital twin scenarios represent simulations of disruptions and have not been validated with any real-world maritime events yet. Future research should involve the use of real vessel telemetry data, AIS data, and operational port datasets.

CONCLUSION

The OceanTwin AI platform is an innovative maritime intelligence solution that combines several sophisticated technologies into one cohesive system. Namely, it involves telemetry simulation, machine learning services, and scenario analysis inspired by the concept of digital twins into one single architecture. The platform supports a variety of applications, such as vessel tracking, monitoring of the port congestion situation, evaluating the risks of sediment accumulation, predicting fuel efficiency, assessing sustainability, simulating disruption, among others.

The experiments demonstrated the predictive stability of the system, explainable behavior of models and consistency of simulations. Thus, the combination of simulation methods and machine learning technologies is proved to be a useful approach for intelligent decision-making in the environment of smart ports.

For the next stage, it is intended to use real-life maritime telemetry data, AIS data and operational port data.

Acknowledgment

It is with profound and sincere appreciation that the authors express their gratitude towards their Project Guide, Prof. Shubhangi Kamble, for the valuable guidance, encouragement, and feedback which she has provided during the process of this project. The authors are equally grateful to their Project Coordinator as well as the Department of Computer Science-Artificial Intelligence at Vishwakarma Institute of Technology, Pune, who have helped provide the resources needed for this project.

REFERENCES

[1] A. De Bari et al., “A decoupled layers architecture for seaport digital twins,” IEEE Access, vol. 12, 2024.

[2] J. M. Zeneli and D. Marinova, “Digital twins in port logistics: simu-lation, monitoring and predictive performance,” International Journal of Logistics Research and Applications, 2024.

[3] V. Badea et al., “Artificial intelligence and IoT integration in smart port digital twin frameworks,” IEEE Transactions on Intelligent Transporta-tion Systems, 2025.

[4] International Maritime Organization, “Initial IMO GHG Strategy,” IMO Publications, 2018.

[5] L. Heilig and S. Voß, “A scientific approach to vessel delay prediction,” Computers & Industrial Engineering, vol. 131, pp. 651–666, 2019.

[6] D. Zhang, C. Zhang, and H. Chen, “Research on port congestion and vessel delay prediction,” Journal of Marine Science and Engineering, vol. 9, no. 4, 2021.

[7] OceanTwin AI Project Repository, “Maritime Intelligence Platform Source Code,” unpublished internal documentation, 2026.