Brodogradnja Volume 77 Number 1 (2026) 77103

journal homepage: www.brodogradnja.fsb.hr

Brodogradnja

An International Journal of Naval Architecture and Ocean Engineering for Research and Development

Improving anchorage occupancy forecasting with stacked ensemble learning

Dae-han Lee [1] *and Joo-sung Kim* [2,*]

[1] *Graduate School of Maritime Transportation System, Mokpo National Maritime University, Mokpo 58628, Korea*
[2] *Division of Navigation Science, Mokpo National Maritime University, Mokpo 58628, Korea*

ARTICLE INFO

Keywords:
Time-series forecasting
Anchorage occupancy
Stacking ensemble
Hexagonal occupancy
LOA estimation
Port management
Maritime safety
Spatial optimization

ABSTRACT

Anchorage areas are essential for safe and efficient maritime operations. However, conventional forecasting models often underperform in dynamic port conditions, as they rely heavily on historical averages and static assumptions. To address these limitations, this study proposes a forecasting framework for anchorage occupancy. This framework uses stacked ensemble learning, integrating both statistical and machine learning models to enhance predictive accuracy and operational reliability. The proposed approach was applied to occupancy data from the E1 anchorage at Ulsan Port, with performance evaluated across various forecasting models and ensemble strategies. In addition, a hexagon-based occupancy estimation method was implemented to assess spatial efficiency and safety in comparison to the traditional anchor circle method. The results demonstrate that the stacking ensemble model effectively captures complex, nonlinear patterns in vessel traffic and delivers improved forecasting performance. These findings highlight the practical potential of stacking ensemble techniques and spatial modeling innovations in enabling proactive anchorage management, reducing congestion, and enhancing maritime safety in real-world port environments.

1. Introduction

An anchorage area is a designated water space intended to ensure the safe mooring of vessels and to facilitate smooth berthing and cargo operations [1]. As a critical component of maritime infrastructure, it plays a vital role in supporting efficient port operations and the safe arrival and departure of vessels. To perform this function effectively, an anchorage must provide not only sufficient space and depth but also the flexibility to accommodate vessels of varying sizes and operational requirements. This necessitates efficient port management practices, particularly in the strategic allocation of port facilities and anchorage areas to support maritime traffic.

The recent expansion of the global economy has led to a substantial increase in maritime trade. However, existing port infrastructure—especially anchorage areas—has struggled to keep pace with growing demand. This challenge is largely attributable to the historical underemphasis on anchorage areas in port planning and development, despite their essential role in vessel queuing and overall port safety [2]. Given the increasing

* Corresponding author.
E-mail address: jskim@mmu.ac.kr

https://doi.org/10.21278/brod77103

Received 3 June 2025; Received in revised form 28 August 2025; Accepted 28 August 2025
Available online 4 September 2025
ISSN 0007-215X; eISSN 1845-5859

volume of maritime traffic and the imperative need for optimized port operations, the proactive and efficient management of anchorage areas has become critical.

Despite their importance, conventional models for forecasting anchorage occupancy remain inadequate under dynamic port conditions. These conventional approaches typically rely on historical averages and static assumptions about berth utilization and vessel arrival patterns, which limit their adaptability to real-time operational fluctuations. For instance, many existing studies focus on long-term planning based on historical trends [3], which often fail to capture short-term fluctuations. Others employ simulation models that require extensive computation and calibration, reducing their practicality for real-time decision-making [4]. Furthermore, these methods frequently oversimplify spatial constraints and fail to account for critical environmental factors—such as wind speed—or dynamic vessel-specific characteristics, such as variations in turning radius. In some cases, predictive modeling capabilities are entirely absent, leaving these approaches ill-equipped to address the evolving dynamics of vessel traffic [1, 5].

The limitations of conventional forecasting models present several critical operational challenges for port authorities. The persistent shortage of adaptable anchorage capacity increases the risk of maritime accidents, including anchor-dragging incidents during adverse weather conditions, as documented at Ulsan Port [6]. These constraints also hinder efficient port management, making it difficult to strategically allocate scarce anchorage areas to accommodate growing and increasingly complex maritime traffic. Therefore, there is a pressing need for an advanced predictive model capable of accurately forecasting both real-time and future changes in anchorage occupancy, thereby reducing safety risks and optimizing spatial utilization.

Recent studies in Canada [7] and New Zealand [8] have provided valuable insights into anchorage management strategies, particularly regarding space utilization and occupancy prediction. The New Zealand study was conducted in response to escalating concerns about anchoring risks in coastal waters and aimed to establish safe anchorage locations while addressing environmental and navigational hazards. Similarly, the Canadian study focused on the increasing vessel traffic at the Port of Prince Rupert and the associated challenges of accommodating larger ships that require more anchorage capacity. Both studies adopted data-driven approaches to analyze anchorage capacity and vessel behavior, emphasizing the optimization of anchorage capacity to prevent congestion and enhance safety. Key considerations included vessel turning radius to minimize collision risks, seabed composition for secure anchoring, and environmental conditions such as wind to improve operational efficiency. Additionally, both studies highlighted the use of automatic identification system (AIS)-based monitoring as a crucial tool for real-time tracking and predictive occupancy modeling.

The increasing adoption of data-driven and predictive modeling techniques in the maritime sector has expanded opportunities for operational forecasting and optimization. For example, Ozsari [9] utilized artificial neural networks to predict main engine power and emissions for container, cargo, and tanker vessels, demonstrating the potential of machine learning to enhance ship-level operational efficiency. Similarly, Atasayan et al. [10] introduced an offline grey-box modeling approach to predict ship maneuvering performance, addressing the complexities of nonlinear and dynamic behaviors in maritime operations. These studies highlight the practical value of predictive analytics in maritime contexts and suggest the potential applicability of similar modeling techniques to anchorage occupancy forecasting—an area that remains relatively underexplored.

To address the current limitations, this study proposes an advanced predictive model to improve the accuracy of anchorage occupancy forecasting. Specifically, we employ stacked ensemble learning to enhance time-series forecasting performance. By integrating a range of predictive models—including both conventional time-series forecasting models and machine learning algorithms—our approach seeks to:

- Improve predictive accuracy by combining diverse forecasting techniques to capture complex temporal patterns and reduce prediction errors,
- Enhance adaptability to dynamic conditions by incorporating vessel-specific characteristics and external environmental factors,

- Provide robust, data-driven insights for anchorage management, facilitating more efficient allocation of anchorage capacity and enabling proactive responses to occupancy fluctuations, thereby improving safety and optimizing overall port operations.

The remainder of this study is structured as follows: Section 2 provides a comprehensive review of related literature, focusing on time-series forecasting applications in maritime and other sectors, as well as the conceptual foundations and applications of stacked ensemble learning. Section 3 details the proposed approach for estimating vessel occupancy within anchorage areas, including the stacked ensemble methodology, base models—autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), vector autoregressive (VAR), Prophet, extreme gradient boosting (XGBoost), and long short-term memory (LSTM)—meta-model configurations—simple ensemble, Random Forest stacking, XGBoost stacking, and gradient boosting stacking—and performance evaluation metrics. Section 4 outlines the experimental setup, data preprocessing procedures, and the case study context. Section 5 presents the experimental results and compares the forecasting performance of individual models with that of the stacking ensemble models. Section 6 discusses the key findings, practical implications, and limitations of the study. Finally, Section 7 concludes the paper.

2. Literature review

2.1 Overview

A time series is a sequence of data points recorded at successive, typically uniform time intervals, and time-series forecasting involves analyzing these patterns to estimate future values. Time-series data often exhibit features such as trends, seasonality, cyclicity, and randomness. Effectively modeling these characteristics is essential for accurate time-series analysis. In this study, time-series forecasting techniques are applied to predict anchorage occupancy. This method, widely used across various industries, identifies patterns in regularly observed data to generate future estimates. In the context of shipping and port logistics, time-series forecasting has been employed for tasks such as predicting ocean freight rates, container throughput, vessel anomaly detection, and weather conditions. Despite the critical role of anchorage management in ensuring port safety and operational efficiency, the application of time-series forecasting to anchorage occupancy remains largely unexplored and underrepresented in the literature.

To address this gap, this section reviews time-series forecasting applications in maritime logistics, supplemented by examples from other industries where such methods have been successfully implemented. It also outlines the forecasting techniques used in this study—ARIMA, SARIMA, VAR, Prophet, XGBoost, LSTM, and ensemble models—and explains the rationale for adopting a stacked ensemble learning.

2.2 Related works

In freight rate forecasting, Veenstra and Franses [11] used a VAR model to predict dry bulk freight rates, demonstrating its suitability for short-term forecasting. Similarly, Kavussanos and Alizadeh [12] applied ARIMA and VAR models to identify seasonality in bulk carrier markets. Chen et al. [13] examined ARIMA and VAR family models to forecast dry bulk spot rates, while Schramm and Munim [14] utilized ARIMA, ARIMAX (autoregressive integrated moving average with exogenous variables), and VAR models to predict container freight rates. Regarding container throughput forecasting, Peng and Chu [15] conducted a comparative analysis of six univariate models at major Taiwanese ports, highlighting the classical decomposition model and SARIMA for their effectiveness. Rashed et al. [16] demonstrated the superior accuracy of the ARIMAX model by incorporating macroeconomic indicators to forecast short-term throughput at the Port of Antwerp. Farhan and Ong [17] validated the SARIMA model's ability to capture seasonal patterns in major international ports. In the area of maritime surveillance and security, Kim and Shin [18] developed an LSTM-based model to detect abnormal vessel behavior using time-series trajectory data. Their approach effectively enhanced coastal surveillance by identifying anomalies in vessel movements.

While these studies underscore the broad applicability of time-series analysis in maritime logistics, the specific application of such methods to anchorage occupancy forecasting—particularly with attention to dynamic spatial constraints and safety—remains underexplored. Conventional assessments of anchorage

capacity have primarily relied on historical data and static assumptions, limiting their effectiveness in forecasting real-time changes. Fan and Cao [19] developed a seaspace capacity and operational strategy analysis system to estimate the capacities of anchorage areas, berthing spaces, and fairways; however, their model depends on predefined traffic assumptions, reducing its adaptability to real-time conditions. Huang et al. [4] introduced a simulation-based model that incorporates circular packing algorithms to assess anchorage capacity, but they acknowledged the significant computational demands and calibration requirements that limit the model's practicality for real-time decision-making. Lee and Lee [3] concentrated on long-term planning based on historical trends, which may not reflect short-term fluctuations. Kwon et al. [5] proposed an anchorage capacity index but did not incorporate real-time spatial constraints or vessel-specific anchoring behaviors. Park et al. [1] analyzed anchorage density and occupancy rates, yet their work lacked predictive modeling to account for dynamic vessel traffic. Critically, none of these studies explicitly integrate safety considerations with dynamic occupancy forecasting. A limited number of recent efforts have begun to address safety in anchorage planning through optimization. For instance, Oz et al. [20] introduced a multi-objective optimization strategy that accounts for both utilization and safety by proposing novel metrics such as the Arrival Intersection Factor and Departure Intersection Factor to quantify collision risks. More recently, Shin and Yang [21] advanced this perspective by applying deep reinforcement learning for integrated vessel path planning and safe anchorage allocation, highlighting how artificial intelligence-driven decision-making can complement forecasting models in improving overall port safety and efficiency. Their approach demonstrated the ability to identify significantly safer berth locations while maintaining utilization efficiency, emphasizing the importance of integrating safety considerations into anchorage planning. However, these optimization methods are primarily oriented toward planning and allocation under current conditions and do not forecast future occupancy dynamics at high temporal resolution.

In addition to shipping and port logistics, time-series forecasting has been widely applied across numerous other industries. Gifty and Li [22] used LSTM, ARIMA, and XGBoost to predict Google stock price movements, finding that XGBoost yielded the highest accuracy following optimization. Zhang et al. [23] applied a range of models—including LSTM, XGBoost, gradient boosting decision trees, ARIMA, and Prophet—to forecast product sales volumes and reported that XGBoost performed best in controlling overfitting. Swami et al. [24] used LSTM, XGBoost, and ARIMA to predict retail product sales, demonstrating that hyperparameter tuning significantly improved LSTM's accuracy. Chen et al. [25] forecasted commodity prices using LSTM, XGBoost, support vector regression, ARIMA, and Prophet, with LSTM showing the best performance when applied to large datasets. Mukhlis et al. [26] optimized LSTM hyperparameters for agricultural production forecasting and reported superior performance compared to ARIMA. Hadri et al. [27] found that XGBoost delivered the best short-term accuracy and computational efficiency for forecasting electricity loads in smart buildings, outperforming LSTM, ARIMA, SARIMA, and Random Forest. In the maritime sector, Tang et al. [28] employed LSTM combined with signal decomposition techniques to improve short-term load forecasting for trimarans, further illustrating the growing application of machine learning in marine engineering. Su et al. [29] developed a real-time ship equipment fault monitoring system by integrating digital twin technology with Random Forest models, demonstrating the practical value of machine learning-based forecasting.

Complex, large-scale time-series data have been widely used to address time-sensitive challenges such as demand forecasting, production planning, dynamic pricing, and anomaly detection. Although these techniques have been successfully applied across diverse sectors—including agriculture, energy, healthcare, transportation, and economics—no direct application has yet been made to anchorage occupancy forecasting.

To address this gap, we propose a forecasting model based on stacked ensemble learning. This model integrates multiple predictive techniques while incorporating vessel-specific characteristics and environmental factors. The model offers high-resolution temporal predictions. It also enhances spatial accuracy through a hexagon-based occupancy estimation method, improving the practical applicability of anchorage management and supporting more informed decision-making in dynamic port conditions.

2.3 Stacking ensemble

2.3.1 Ensemble learning and categorization

Ensemble learning is a machine learning paradigm that combines multiple models to improve predictive accuracy and generalization performance. By aggregating the outputs of diverse learning algorithms, ensemble methods can reduce variance as well as bias and enhance robustness against overfitting. Ensemble learning techniques are generally classified into three categories [30].

Bootstrap aggregation (bagging) involves training multiple models on randomly resampled subsets of the original dataset (with replacement) and aggregating their predictions—typically through averaging for regression tasks or majority voting for classification tasks. This approach effectively reduces variance and overfitting, with Random Forest being one of the most widely recognized implementations.

Boosting constructs models sequentially, where each new model focuses on correcting the errors of its predecessor by assigning greater weight to misclassified instances. While boosting can significantly enhance predictive accuracy by concentrating on difficult-to-predict samples, it can lead to overfitting if not properly regularized.

In contrast, stacking integrates multiple independently trained base models and combines their outputs using a meta-model, which learns to optimize the combination of individual predictions. Unlike bagging and boosting, stacking allows for the integration of heterogeneous models. Common stacking approaches include direct stacking, cross-validation-based stacking, and blending.

2.3.2 Stacking ensemble applications in time series forecasting

Zhan et al. [31] demonstrated the effectiveness of stacking by proposing a model to improve the accuracy of ocean wave height predictions. Their ensemble integrated four distinct base learners—XGBoost, Light Gradient Boosting Machine, Random Forest, and Adaptive Boosting—using a linear regression meta-learner. Utilizing real-world oceanographic and meteorological data, the stacked ensemble significantly reduced both mean absolute error (MAE) and mean squared error (MSE) compared to the individual models, demonstrating the effectiveness of stacked ensembles in marine environmental forecasting.

Hoque and Sharma [32] proposed an ensemble deep learning model that combines AIS data, spectral clustering, and LSTM for real-time maritime anomaly detection. Ensemble learning has also demonstrated strong capabilities in maritime-specific operational forecasting. For example, Bodunov et al. [33] employed a voting ensemble of tree-based models—including Random Forest, Gradient Boosting Decision Trees, XGBoost, and Extremely Randomized Trees—for real-time vessel destination and estimated time of arrival (ETA) prediction. Their model achieved high accuracy rates of 97 % for destination prediction and 90 % for ETA, effectively addressing the complexities of real-world maritime spatiotemporal streaming data influenced by vessel trajectory, speed, and weather conditions. This demonstrates the robustness of ensemble methods in capturing the intricate, nonlinear dynamics characteristic of maritime traffic forecasting. Kim [34] used a stacking ensemble approach for wind power forecasting, demonstrating improved accuracy over single-model methods. In the automotive sector, Wi [35] applied stacking techniques to optimize the design of electric vehicle motors, enhancing performance while reducing computational costs. Seo [36] integrated anomaly detection into a stacking ensemble framework to improve credit default prediction in imbalanced financial datasets. In healthcare, Gupta et al. [37] employed a stacking ensemble model to predict post-COVID-19 cardiovascular complications, achieving superior accuracy compared to individual models.

Given the limitations of conventional forecasting models in managing the complex, nonlinear, and volatile patterns of anchorage occupancy, this study proposes a stacked ensemble learning. Vessel operation data often exhibit irregular patterns shaped by external environmental conditions—such as weather—and unpredictable fluctuations in demand. To address these complexities, we integrated conventional time-series forecasting models (e.g., ARIMA and VAR) with machine learning algorithms (e.g., XGBoost and LSTM) within a stacking ensemble framework. This approach harnesses the complementary strengths of individual models, reduces the risk of overfitting, and more effectively captures the nonlinear and volatile dynamics inherent in anchorage utilization data.

A significant research gap persists in accurately forecasting anchorage occupancy, particularly given the dynamic and complex interactions between vessel-specific characteristics and external environmental factors. While prior studies have addressed elements of anchorage capacity and operational efficiency, they often rely on static assumptions or lack robust predictive capabilities for capturing real-time, dynamic changes. Existing approaches frequently oversimplify spatial constraints and inadequately account for environmental influences and the heterogeneous behaviors of vessels. To address these critical limitations, this study proposes a novel framework that employs stacked ensemble learning to enhance time-series forecasting of anchorage occupancy. The methodology integrates a range of predictive models to effectively capture the complex, nonlinear, and volatile patterns inherent in anchorage utilization data. Additionally, the framework incorporates a hexagon-based occupancy estimation method that accounts for vessel-specific length overall (*LOA*) and variations in wind speed, offering a more realistic and safety-oriented assessment of occupied space compared to traditional anchor circle methods. This integrated approach aims not only to improve forecasting accuracy but also to provide a practical and scalable solution for optimizing anchorage capacity management and enhancing maritime safety in real-world port operations.

3. Methodology

3.1 Enhancing anchorage capacity calculation through hexagonal occupancy modeling

In conventional anchorage planning, the anchor circle—typically determined by a vessel's *LOA* and the water depth at the anchorage—is commonly used to estimate the area occupied by an individual vessel. When this traditional concept is applied using a circle packing algorithm [4], the spatial arrangement of vessels within the anchorage can be visualized, as illustrated in Figure 1. However, the interstitial spaces between vessels—often referred to as "dead space"—are generally unsuitable for anchoring and thus represent underutilized portions of the anchorage. This inefficiency limits overall spatial utilization. From a geometric perspective, the most efficient way to pack equal-sized circles is through hexagonal close-packing, which provides the highest possible theoretical density of approximately 90.7 % ($\frac{\pi}{2\sqrt{3}}$). In contrast, square-lattice packing of circles achieves a lower density of approximately 78.5 % ($\frac{\pi}{4}$). This inherent difference arises because hexagonal close-packing minimizes interstitial voids between circles, while square-based arrangements leave larger unusable gaps.

Accordingly, this study adopts a hexagon-based occupancy estimation method. Rather than treating dead space as unusable, this method considers it a safety buffer between vessels, thereby providing a more conservative and realistic estimate of occupied area. This approach is designed to enhance anchorage capacity while reducing the risk of vessel-to-vessel collisions, ultimately improving both operational efficiency and safety through a more practical spatial layout for anchorage management.

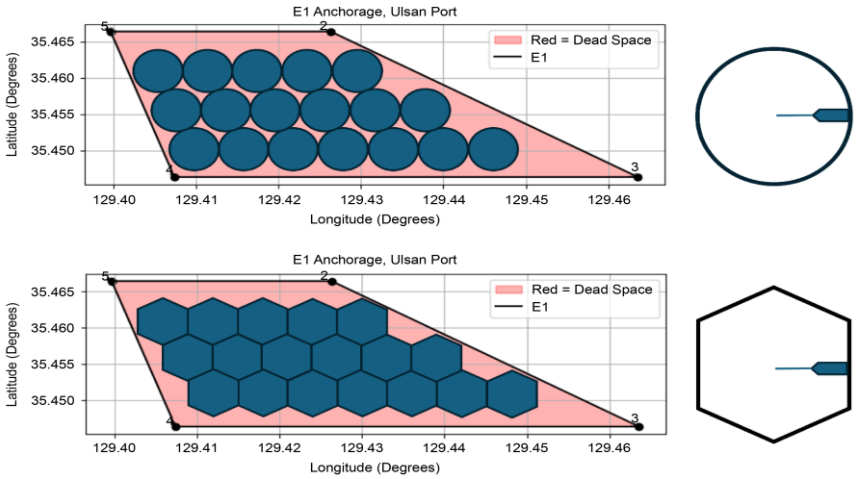

Fig. 1. Comparison of the area occupied by a traditional anchor circle and a hexagon

Based on the specifications of vessel "A" (a chemical tanker with a gross tonnage (*GT*) of 4,060 and a *LOA* of 103.1 m), anchored at E1 Anchorage, occupancy areas were calculated using both the traditional anchor circle method and the proposed hexagonal-based method. The calculations followed the anchorage design standards outlined in the Korean Design Standard (KDS 64 40 10, 2024) [38] issued by the Ministry of Oceans and Fisheries (MOF) (Table 1).

Table 1 Anchorage turning radius with depth and length of the ship (MOF)

Purpose of Use	Seabed or Wind Speed Condition	Turning Radius(m)
Offshore Waiting or Cargo Handling	The seabed is **Good** for anchoring.	$L^{1} + 6D^{2}$
	The seabed is **Bad** for anchoring.	$L + 6D + 30$

[1] *L*: Length overall (m)

[2] *D*: Depth (m)

According to data from the Korea Hydrographic and Oceanographic Agency [39], the average wind speed at Ulsan Port between 2020 and 2024 was approximately 3.25 m/s. Additionally, the Ulsan Port Authority [40] reports that the average water depth at E1 Anchorage is approximately 40 m, with a seabed composed of mud. Based on this information, the turning radius (*R*) for vessel "A" was calculated using its *LOA*.

$$R = L + 6D = 103.1 + 6 \times 40 = 343.1 \text{ m} \tag{1}$$

Based on this turning radius, the traditional anchor circle was estimated to cover ~369,820.78 m^2, while the hexagonal-based occupancy area covered ~407,785.76 m^2. Therefore, the hexagonal method requires ~37,965 m^2 more space, which is an increase of ~10.3 %. Figure 2 presents a comparison between the traditional anchor circle and the hexagonal occupancy area.

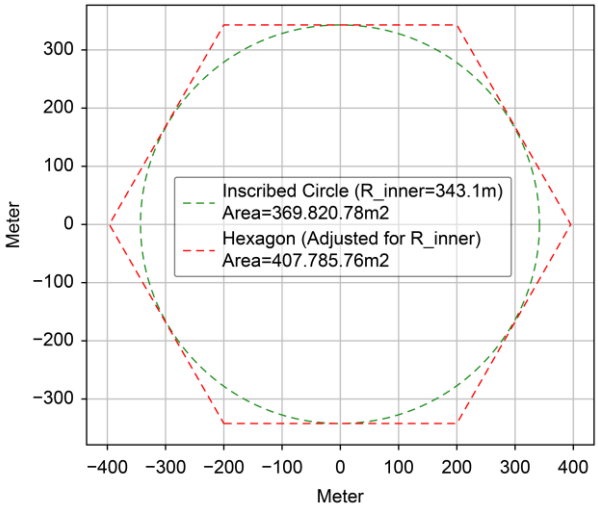

Fig. 2. Hexagon-based occupancy area with adjusted inscribed circle

3.2 Stacked ensemble learning: concept and mechanism

Stacked ensemble learning, also known as stacked generalization, was first introduced by Breiman [41] and has since been extensively developed to enhance predictive accuracy by combining multiple machine learning models. Unlike bagging and boosting, which typically rely on homogeneous learners, stacking employs a diverse set of algorithms to more effectively model complex data patterns. The architecture of a stacking ensemble consists of two primary levels.

1. Base models (level-0 learners): These models are trained independently on the training data and can include various machine learning algorithms. Each base model generates predictions that are subsequently used as input features for the next level.

2. Meta models (level-1 learner): Meta models take the predictions from the base Models as input and learn to combine them optimally. By training on these outputs, the meta-model refines the final prediction and improves generalization performance.

3.3 Base models

To construct a robust and generalizable stacking ensemble, we selected six time-series forecasting models—ARIMA, SARIMA, VAR, Prophet, XGBoost, and LSTM—based on their complementary strengths in capturing diverse temporal patterns and their demonstrated effectiveness in related domains.

Classical statistical models such as ARIMA, SARIMA, and VAR offer strong interpretability and computational efficiency, making them suitable for large-scale experimentation. These models are well-suited for capturing linear dependencies and seasonal structures and have been widely used in contexts such as port throughput forecasting and freight rate analysis [13, 15]. Prophet was selected for its ability to handle multiple seasonalities, calendar effects, and event-driven fluctuations, which are characteristic of port operations [42]. It strikes a reasonable balance between modeling flexibility and computational efficiency, facilitated by its built-in optimization procedures. To capture nonlinear and complex interactions in the data, we included XGBoost, a tree-based gradient boosting algorithm. XGBoost is particularly well-suited for structured data and has shown strong performance in port forecasting tasks [27]. Although computationally intensive, we parallelized its training to reduce runtime. Finally, LSTM networks were included for their capacity to model long-range dependencies in time-series data. LSTM has been widely used in maritime applications, including vessel trajectory prediction [18]. To manage its high computational cost, we implemented mini batch learning and early stopping strategies to reduce training time and prevent overfitting.

By integrating these six models, the proposed ensemble aims to capture both short-term fluctuations and long-term trends, as well as linear and nonlinear temporal patterns. This diverse model pool enhances the predictive robustness of the ensemble, allowing the meta-model to learn optimal combinations of individual predictions based on each model's respective strengths.

In terms of computational cost, the classical statistical models—ARIMA, SARIMA, and VAR—offer fast training times and are well-suited for deployment in resource-constrained resources. In contrast, machine learning models such as XGBoost and LSTM are more computationally intensive. To address this, we optimized their training procedures: XGBoost was trained using parallelized boosting techniques, while LSTM utilized mini batch learning and early stopping to improve training efficiency and prevent overfitting. As the meta-model operates only on the predictions of the base models, its additional computational overhead is minimal.

3.3.1 AutoRegressive Integrated Moving Average (ARIMA)

Developed by Box and Jenkins [43], ARIMA is a widely used model that combines autoregressive (AR), differencing (I), and moving average (MA) data to forecast stationary time-series data. It is simple and interpretable, making it effective for small datasets; however, it struggles to capture seasonality and nonlinear patterns. The ARIMA parameters were manually selected based on autocorrelation and partial autocorrelation analysis, reflecting conventional time-series forecasting modeling practices optimized for the non-seasonal pattern of the data.

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{2}$$

where:

p: the non-seasonal autoregressive (AR) order

d: the non-seasonal differencing (I) order

q: the non-seasonal moving average (MA) order

c: a vector of constant term (intercept)

Y_t: the value of the time series at time t (i.e., the forecasted value)

ε_t: the error term at time t

3.3.2 Seasonal ARIMA (SARIMA) [44]

SARIMA, an extension of the ARIMA introduced by Box and Jenkins [33], incorporates seasonal components to model periodic patterns in time-series data. It is effective for handling seasonality; however, the need for tuning additional seasonal parameters increases the model complexity. The SARIMA model was configured using an automated parameter selection procedure based on the Akaike Information Criterion (AIC), enabling the discovery of optimal seasonal orders and improving robustness against manual misspecification.

Mathematically, a SARIMA$(p, d, q)(P, D, Q)_m$ model is expressed as:

$$\Phi_P(B^m)\phi_p(B)(1-B)^d(1-B^m)^D Y_t = \Theta_Q(B^m)\theta_q(B)\varepsilon_t, \quad (3)$$

where:

P: the seasonal autoregressive (AR) order

D: the seasonal differencing (I) order

Q: the seasonal moving average (MA) order

m: the number of observations per seasonal cycle

Y_t: the value of the time series at time t (i.e., the forecasted value)

B: the backshift operator, where $B^k Y_t = Y_{t-k}$

ε_t: the error term at time t

3.3.3 Vector Autoregression (VAR)

Proposed by Sims [45], VAR models the relationships among multiple time-series variables by treating them as interdependent. It is powerful for multivariate forecasting but becomes computationally inefficient when applied to univariate or high-dimensional data. The lag order of the VAR model was selected using AIC-based model selection. This approach ensures the inclusion of the most informative temporal dependencies while avoiding overfitting.

Mathematically, VAR(p) is defined as:

$$Y_t = c + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_\ell Y_{t-\ell} + \varepsilon_t, \quad (4)$$

where:

Y_t: a multivariate time series vector at time t, representing the values of n different time series variables at that moment.

c: a vector of constant term (intercept)

A_i: coefficient matrices (of size $n \times n$) that capture the linear relationships between the lagged values of all variables in the system and their current values

ℓ: the maximum lag order included in the model

ε_t: a vector of error terms (white noise) at time t, representing the unpredictable part of the current values

3.3.4 Prophet

Developed by Taylor and Letham [42], Prophet is an additive model that decomposes a time series into trends, seasonality, and holiday effects. It is easy to use and robust to missing data and outliers but can underperform on simple, non-seasonal datasets. The Prophet model was implemented using its default configuration, which automatically detects changepoints and decomposes the series into trend and seasonality components, making it especially suitable for operational forecasting without intensive tuning.

The model is represented as:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t, \tag{5}$$

where:

$y(t)$: the forecasted value of the time series at time t

$g(t)$: the trend component, representing non-periodic changes in the time series

$s(t)$: the seasonal component, representing periodic changes (e.g., weekly or yearly cycles)

$h(t)$: the holiday effect, representing impacts from irregular events like holidays

ε_t: the error term at time t, representing the unpredictable part of the current value

3.3.5 Extreme Gradient Boosting (XGBoost)

Developed by Chen [46], XGBoost is a gradient boosting-based ensemble method optimized for speed and accuracy. It achieves high performance on structured data; however, it requires careful hyperparameter tuning and may be overfit on smaller datasets. Key hyperparameters were manually adjusted based on empirical testing and validation performance to balance model complexity and predictive accuracy.

$$L(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega\left(f_k\right), \tag{6}$$

where:

$L(\theta)$: the overall objective function that XGBoost minimizes, combining the loss function and regularization term

$l(y_i, \hat{y}_i)$: the represent the loss function (e.g., squared error for regression and log loss for classification)

$\Omega(f_k)$: the regularization term that penalizes model complexity

3.3.6 Long Short-Term Memory (LSTM)

Proposed by Hochreiter and Schmidhuber [47], LSTM is a type of recurrent neural network specifically designed to capture long-term dependencies in sequential data. It excels at modeling complex temporal patterns but is computationally intensive and typically requires large datasets for effective training. The LSTM architecture was tuned empirically through iterative experiments, with layer sizes, dropout rates, and sequence lengths adjusted to balance training stability and generalization. Early stopping was applied to prevent overfitting.

The core equations for an LSTM cell are:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_{f,} \tag{7}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i, \tag{8}$$

$$\hat{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C, \tag{9}$$

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t, \tag{10}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o, \tag{11}$$

$$h_t = o_t * \tanh\left(C_t\right), \tag{12}$$

where:

f_t: the forget gate activation vector at time t. It determines what information from the previous cell state C_{t-1} should be discarded

i_t: the input gate activation vector at time t. It determines what new information from the current input x_t should be stored in the cell state

o_t: the output gate activation vector at time t. It determines what part of the current cell state C_t should be output to the hidden state h_t

C_t : the cell state vector at time t. This is the long-term memory of the LSTM cell, carrying information through the sequence

h_t : the hidden state vector at time t. This is the output of the LSTM cell, used for predictions and as input to the next time step

W_f, W_i, W_o: weight matrices for the forget, input, candidate cell state, and output gates, respectively. These are learned during training

b_f, b_i, b_C, b_o: bias vectors for the forget, input, candidate cell state, and output gates, respectively. These are also learned during training

3.3.7 Summary of model complexity

To synthesize the computational characteristics of the six base models, their asymptotic complexity is summarized in Table 2. This table provides a theoretical perspective using Big-O notation, highlighting the trade-offs between each model's predictive accuracy and computational cost. This summary serves to provide a holistic view of the operational trade-offs inherent in each model.

Table 2 Computational complexity of the base models

Model	Training Complexity	Inference Complexity	Key Considerations
ARIMA	$O(n \times p^2)$	$O(P)$	Efficient for small datasets; limited in handling seasonality and nonlinearity
SARIMA	$O(n \times (p+q+P+Q)^2)$	$O(p+q+P+Q)$	Additional seasonal parameters increase cost but capture periodic patterns
VAR	$O(n \times k^3)$	$O(k^2)$	Effective for multivariate data but scales poorly with number of variables
Prophet	$O(n \times k)$	$O(k)$	Handles multiple seasonalities and events efficiently
XGBoost	$O(n \times d \times T)$	$O(d \times T)$	Strong nonlinear modeling capacity; higher training cost due to tree ensembles
LSTM	$O(n \times h^2)$	$O(h^2)$	Captures long-term dependencies; computationally intensive and data-hungry

* Notes: n = number of observations, p, q = AR and MA orders, P, Q = seasonal orders, k = number of variables or seasonalities, d = tree depth, T = number of trees, h = number of hidden units.

This theoretical comparison demonstrates that classical statistical models (ARIMA, SARIMA, VAR) are generally lightweight in terms of computation, making them practical for rapid experimentation. In contrast, machine learning (XGBoost) and deep learning models (LSTM) require substantially more resources, reflecting a trade-off between computational cost and their ability to capture nonlinear, complex patterns. Prophet occupies a middle ground, offering flexible modeling with relatively moderate complexity. This diversity among base models justifies their integration into a stacked ensemble, which leverages complementary strengths while mitigating individual limitations.

3.4 Meta models

3.4.1 Simple Ensemble [48]

The Simple ensemble approach combines predictions from multiple base models by averaging their outputs. While this method reduces variance and improves prediction stability, it assigns equal weight to all models, potentially underutilizing the strengths of higher-performing models.

3.4.2 Random Forest Stacking [49, 50]

In this approach, predictions from base models are used as input features for a Random Forest meta-learner. By capitalizing on the robustness and low variance of Random Forest algorithms, this stacking method effectively integrates diverse predictions, enhancing overall forecasting accuracy.

3.4.3 XGBoost Stacking [42, 49]

XGBoost stacking employs an XGBoost model as the meta-learner to sequentially learn from the outputs of the base models. Known for its high performance and flexibility, this method captures complex relationships among predictions but requires careful hyperparameter tuning to avoid overfitting.

3.4.4 Gradient Boosting Stacking [49, 51]

This method uses a gradient boosting algorithm as the meta-learner, iteratively reducing the residual errors from previous base model predictions. It is particularly effective in refining predictive accuracy by capturing subtle nonlinear patterns among base model outputs, making it well-suited for complex time-series forecasting tasks.

3.5 Performance metrics for evaluation

To objectively evaluate model performance, predictive accuracy was evaluated using statistical metrics, including the MAE, RMSE, and MAPE (Table 3).

Table 3 Evaluation parameters for forecasting performance

Criterion	Definition
MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
MAPE	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i - \hat{y}_i}{y_i}\right\rvert$

Figure 3 shows an overview of the methodological framework adopted in this study, which encompasses data preprocessing, model training, and ensemble integration.

4. Experimental setup

4.1 Survey port

Ulsan Port is one of South Korea's largest and most prominent ports, specializing in the handling of liquid cargo. As part of its strategic expansion, the port is actively developing its facilities and infrastructure under the Northeast Asia Oil Hub project, aiming to establish itself as a key crude oil terminal in the region. The anchorage area spans approximately 42.51 km² and can accommodate up to 40 vessels simultaneously. However, a persistent shortage of anchorage capacity has been reported, contributing to an increased risk of maritime accidents—including anchor-dragging incidents during adverse weather conditions—particularly in Anchorage Group E [52].

The port currently operates ten designated anchorages (E1–E3 and M1–M7) within its port limits. These anchorages are situated in open waters and are characterized by the simultaneous anchoring of multiple large vessels in confined areas. Owing to the heightened risk of safety-related incidents, this study focuses specifically on Anchorage Group E at Ulsan Port. The analysis covers the period from 2020 to 2024, utilizing

vessel operation data from the E anchorage area. The dataset was obtained from the Port Management Information System (PORT-MIS) [53], operated by the MOF of the Republic of Korea.

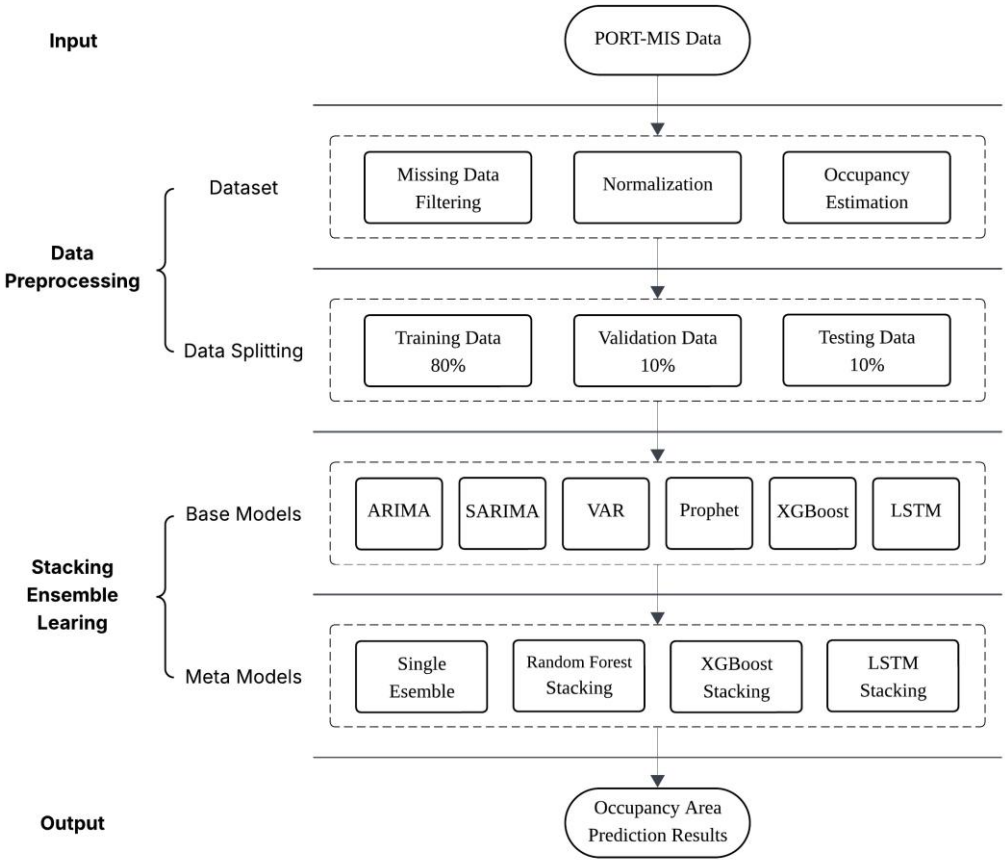

Fig. 3. Overall framework of the stacking ensemble-based anchorage occupancy forecasting model

4.2 Data preparation and preprocessing

4.2.1 Data preprocessing

Using the initial 241,220 arrival and departure records for Ulsan Port obtained from the PORT-MIS, a rigorous data preprocessing and cleaning pipeline was implemented to ensure high-quality inputs for reliable forecasting. Preliminary inspection revealed several common issues in large-scale operational datasets, including missing values, inconsistencies, and outliers. To address these, a multi-stage data cleaning procedure was applied.

First, records with missing values (393 data points) were systematically excluded. Additionally, 434 data points corresponding to vessels that had been berthed for unusually long periods before January 1, 2020, 00:00, were excluded, as this likely represented long-term storage rather than active anchorage events. After these exclusions, 240,393 valid records were retained for analysis. Duplicate entries were meticulously identified and removed to ensure data integrity. The analysis was strictly limited to the period from January 1, 2020, 00:00, to December 31, 2024, 24:00.

To prepare the dataset for statistical analysis and modeling, data consistency was prioritized. Ship type names were translated into English for standardization, while commas were removed and decimal places rounded to ensure numerical uniformity. Vessel type names were further standardized to reduce categorical fragmentation. To simplify data distribution and enhance the model's ability to capture non-linear relationships, several continuous variables were discretized into categorical variables.

The analysis period was categorized monthly, and anchorage occupancy time was calculated using each vessel’s "ARRIVAL" and "DEPARTURE" timestamps. Occupancy durations were then classified into five

categories: <12 h, 12-24 h, 24-48 h, 48-72 h, and >72 h. Records with negative occupancy durations or durations exceeding 72 h were removed to eliminate erroneous entries.

Finally, vessel gross tonnage was categorized into the following intervals: <100, 100–500, 500–1,000, 1,000–5,000, 5,000–10,000, 10,000–30,000, 30,000–150,000, and >150,000 tons.

To ensure robust outlier management beyond basic filtering—particularly for *LOA* estimation—the interquartile range (IQR) method was applied. This approach systematically identified and removed data points falling below $Q1-1.5\times IQR$ or above $Q3+1.5\times IQR$, thereby enhancing the reliability of the *LOA*-to-gross tonnage (*GT*) ratio and improving the accuracy of subsequent *LOA* estimations. This rigorous data cleaning process was critical in reducing data quality issues and establishing a reliable foundation for accurate time-series forecasting.

4.2.2 Basic Statistical analysis of the arrival and departure data of Ulsan Port

1. Vessel type analysis of Ulsan port

An analysis of vessel arrivals and departures at Ulsan Port between 2020 and 2024 revealed that oil product tankers (43.63 %, 38,413 cases) and chemical tankers (25.32 %, 22,287 cases) were the most frequently occurring vessel types, reflecting the port's specialization in liquid cargo handling.

At the E1 anchorage, oil product tankers accounted for the largest share (45.84 %, 22,906 cases), followed by chemical tankers (32.79 %, 16,385 cases), with consistent activity from LPG carriers and general cargo ships. A similar pattern was observed at the E2 anchorage, which was primarily used by oil product tankers (42.79 %, 9,497 cases) and chemical tankers (20.00 %, 4,440 cases). Notably, crude oil tankers utilized E2 more frequently (951 cases) than E1 (164 cases), indicating a preference for the former.

In contrast, E3 anchorage was primarily used by oil product tankers (37.88 %, 6,010 cases) and bulk carriers (23.48 %, 3,725 cases). The presence of various vessel types at E3 suggests that it functions as a multipurpose anchor.

2. Arrival and departure statistics by vessel tonnage categories at Ulsan Port

An analysis of vessel tonnage categories at Ulsan Port from 2020 to 2024 revealed that vessels in the 1,000–30,000 ton range accounted for the majority of arrival and departure activity (80.29 %), with vessels in the 1,000–5,000 ton category being the most frequent (36.96 %).

E1 Anchorage primarily accommodates small to medium-sized vessels under 10,000 tons (99.65 %), particularly those transporting liquid cargo. E2 Anchorage is predominantly used by medium to large-sized vessels, with the 10,000–30,000 ton category comprising 89.25 % of activity. In contrast, E3 Anchorage records the highest activity from large vessels over 30,000 tons (85.65 %), including crude oil tankers and bulk carriers, while also serving a range of mid-sized vessels.

3. Analysis of anchorage occupancy time

Anchorages E1, E2, and E3 collectively exhibited a high proportion of long-term anchorage occupancy exceeding 72 h (65.27 %). E2 and E3 showed a clear pattern of usage by vessels engaged in extended anchorage durations. Conversely, E1 Anchorage demonstrated a bimodal occupancy pattern, with vessels most commonly anchored for either 24–48 h (22.20 %) or more than 72 h (53.68 %).

4.2.3 Predicting *LOA* from *GT* for improved turning radius estimation

Accurate estimation of a vessel's turning radius within an anchorage area requires reliable *LOA* information. However, many records in the PORT-MIS dataset lack this critical attribute. To address this issue, a *LOA*-to-*GT* ratio-based estimation method was applied. A total of 69,810 vessel records containing both *LOA* and *GT* information were collected from the arrival and departure dataset, excluding vessels longer than 400 m, which are not permitted to anchor at Ulsan Port. Outliers were initially removed by filtering records with an *LOA*-to-*GT* ratio less than 10 or greater than 1,500. After this initial refinement, 55,811 valid records remained. To further enhance data reliability, the IQR method was applied. The first quartile (Q1) and third quartile (Q3) were calculated, and data points falling below $Q1–1.5\times IQR$ or above $Q3+1.5\times IQR$ were

classified as outliers and excluded. This multi-stage data refinement process improved the robustness of the *LOA*-to-*GT* ratio, ensuring more accurate *LOA* estimations for turning radius calculations.

Using the refined dataset, several regression models were evaluated to predict *LOA* from *GT*, including linear regression, second-order polynomial regression, Random Forest, and multilayer perceptron (MLP)-based neural networks. The Random Forest regression model achieved the best predictive performance, with the lowest RMSE of 10.34 (Table 4).

Table 4 Comparison of regression model performance for estimating *LOA* from *GT*

Model	Linear regression	Second-order polynomial regression	Random Forest regression	Neural Network (MLP)
RMSE	28.71	20.28	10.34	14.97

Using the Random Forest regression model, *LOA* values were estimated for vessels in the arrival and departure records from the E1, E2, and E3 anchorages at Ulsan Port where *LOA* information was missing. The model utilized *GT* values from the PORT-MIS dataset as independent variables to predict the corresponding *LOA* through a trained regression equation. This process enabled the reliable supplementation of missing *LOA* information, which was subsequently used to calculate vessel turning radii. Figure 4 presents a comparison between the original and predicted *LOA* values based on *GT*, demonstrating the effectiveness of the Random Forest regression model.

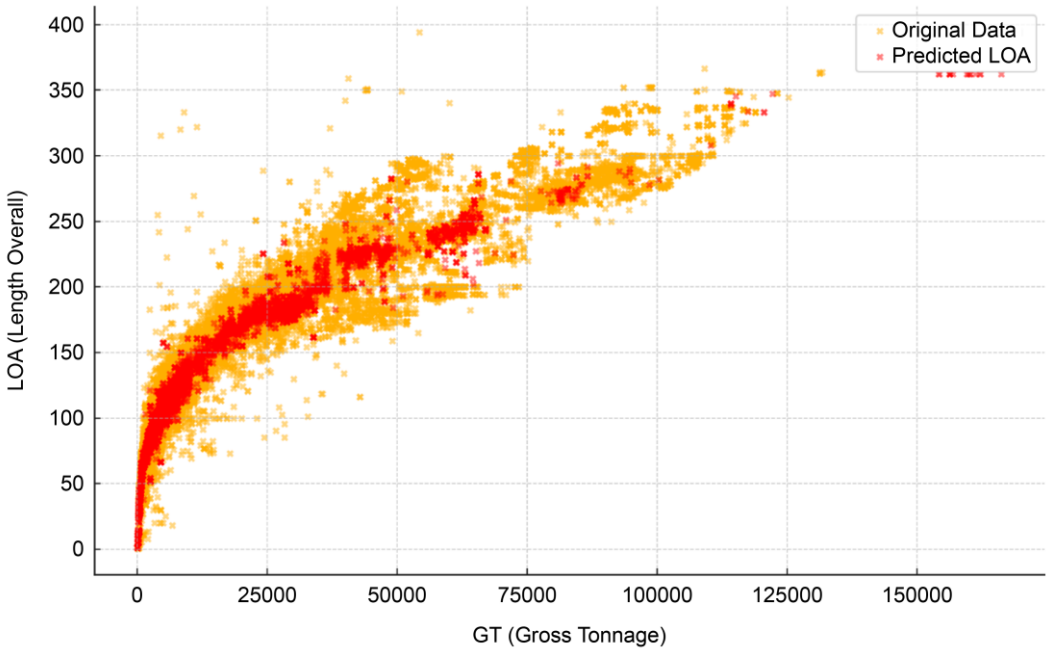

Fig. 4. Comparison of original and predicted *LOA* based on *GT* using random forest regression

4.2.4 Adjustment of turning radius based on external factors

According to the MOF in South Korea, the KDS recommends considering seabed conditions and wind speed when calculating a vessel's turning radius. However, the current standards provide only generalized clearance distances based on broad categorizations of seabed quality and wind conditions, lacking detailed, quantifiable criteria. In contrast, the Permanent International Association of Navigation Congresses (PIANC) [54] offers more granular and systematic guidelines, specifying clearance distances across defined wind speed ranges (see Table 5). In practical anchoring operations, wind speed plays a critical role in determining anchor chain length. As wind speed increases, longer anchor chains are required, resulting in larger turning radii. Therefore, PIANC's wind-based clearance distance standards are widely regarded as both reasonable and operationally relevant for anchorage planning. In this study, seasonal variations in maritime environmental conditions and vessel movement patterns at Ulsan Port were incorporated to improve the

accuracy of occupancy predictions at E1 Anchorage. Vessel turning radii were adjusted using PIANC's wind-speed-based clearance guidelines.

Table 5 Comparative analysis of design criteria

Environmental conditions		**Turning Radius**		
Seabed	Wind (m/s)	MOF	PIANC	Proposed
Good	≤ 10	$L+6D$	$L+5D+0$ m + Safety Clearance*	$L+6D+0$ m
	= 20		$L+5D+60$ m + Safety Clearance	$L+6D+60$ m
	= 30		$L+5D+120$ m + Safety Clearance	$L+6D+120$ m
	≥ 30		$L+5D+180$ m + Safety Clearance	$L+6D+180$ m
Bad	≤ 10	$L+6D+30$ m	$L+5D+30$ m + Safety Clearance	$L+6D+60$ m
	= 20		$L+5D+90$ m + Safety Clearance	$L+6D+120$ m
	= 30		$L+5D+150$ m + Safety Clearance	$L+6D+180$ m
	≥ 30		$L+5D+210$ m + Safety Clearance	$L+6D+240$ m

* A safety clearance which may be 10 % of the *LOA*, with a minimum of 20 m

To systematically reflect the influence of maritime environmental factors, anchorage design standards were revised by combining MOF guidelines with PIANC's wind-based clearance distances. Based on these adjusted standards, the occupancy area of vessel "A" was recalculated under varying wind conditions. The corresponding results are shown in Figure 5. Occupancy areas were estimated using daily maximum wind speed data provided by the Korea Hydrographic and Oceanographic Agency. These wind-adjusted occupancy estimates were then used as input for anchorage occupancy forecasting, allowing the model to reflect more realistic and dynamic maritime conditions.

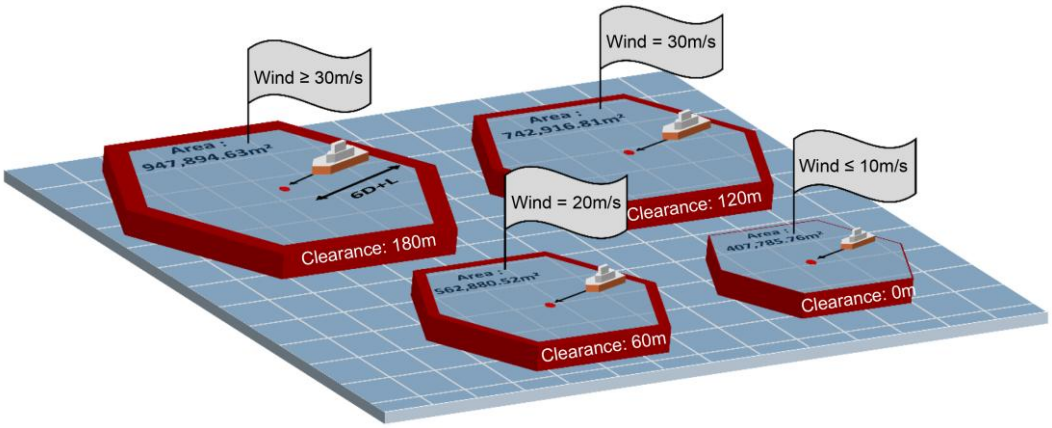

Fig. 5. Effects of wind speed on anchorage occupancy estimation

5. Experiments and results

All computational experiments were conducted in a Google Colab Pro environment to ensure a consistent and high-performance computing environment. The hardware specifications used for training and inference were as follows: Intel(R) Core(TM) i9-10900KF CPU @ 3.70GHz, 128GB of RAM, and an NVIDIA GeForce RTX 3090 GPU.

After collecting data from vessels that utilized the E1 Anchorage at Ulsan Port, the *LOA* of each vessel was estimated using a regression model based on its *GT*, and the corresponding hexagon-based occupancy area was calculated. The daily occupancy area was then derived by aggregating the individual occupancy areas of all vessels anchored at E1 on a given day, serving as the primary variable for analysis.

5.1 Analysis of daily occupancy area and vessel count trends

Figure 6 presents the time-series trends of the total daily occupancy area (in m²) alongside the daily vessel count at E1 Anchorage. The lower panel displays the number of vessels anchored each day, which generally ranged from 20 to 35, with occasional peaks exceeding 40 vessels. The trends confirm a clear positive relationship: an increase in the daily number of vessels correspond with proportional increase in the total occupancy area.

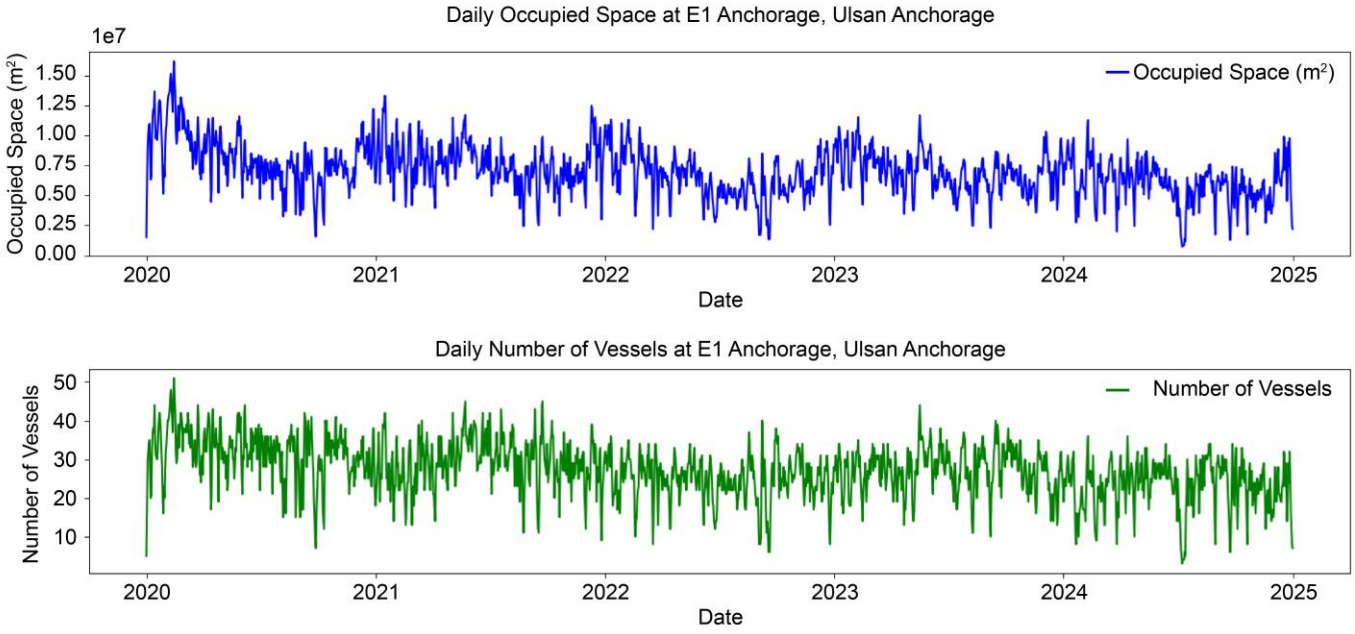

Fig. 6. Trends in daily occupied area and number of vessels at E1 Anchorage, Ulsan Port

A correlation analysis between the daily occupancy area and the number of anchored vessels yielded a correlation coefficient of approximately 0.84, indicating a strong positive association. This suggests that as the number of vessels increases, the total occupied anchorage area expands significantly. These findings indicate that the dataset is well-suited to serve as a foundational input for developing a berth occupancy demand forecasting model.

5.2 Forecasting experiment and results using single models

The dataset was structured using daily variables and consisted of two variables: the date and the cumulative anchorage occupancy area (m²) for all vessels anchored on that day. The dataset was divided into the following ratios to make the collected data suitable for time-series forecasting:

Training Data: January 1, 2020 – June 30, 2023 (~80 % of the data)

Validation Data: July 1, 2023 – March 31, 2024 (~10 % of the data)

Testing Data: April 1, 2024 – December 31, 2024 (~10 % of the data)

During the modeling process, min–max scaling was applied to normalize the variables within the range of 0–1.

This study employed representative models from time-series forecasting, machine learning, and deep learning paradigms to evaluate and compare predictive performance. The six models included in the experiment were ARIMA, SARIMA, VAR, Prophet, XGBoost, and LSTM. Their forecasting outputs were visually assessed by comparing predicted and actual values through time-series plots, as shown in Figure 7.

To ensure robust model evaluation and applicability to unseen data, a fixed-period hold-out validation strategy was adopted while maintaining the chronological sequence of the data. For the ARIMA, SARIMA, and VAR models, a rolling forecast approach was utilized during the validation period. In this method, each model was initially trained on historical data up to a given time point, produced a one-step-ahead prediction, and was subsequently updated by incorporating the newly observed value from the validation set. This

sequential updating process closely simulates a real-time forecasting scenario. In contrast, Prophet, XGBoost, and LSTM models were trained once on the entire training dataset and used to generate forecasts for the entire validation period based on the learned temporal patterns. This consistent validation framework enabled direct comparison of all base models' performances before their integration into the stacked ensemble. The final stacking ensemble models were then evaluated on the same validation dataset, using the predictions generated by the base models as input features.

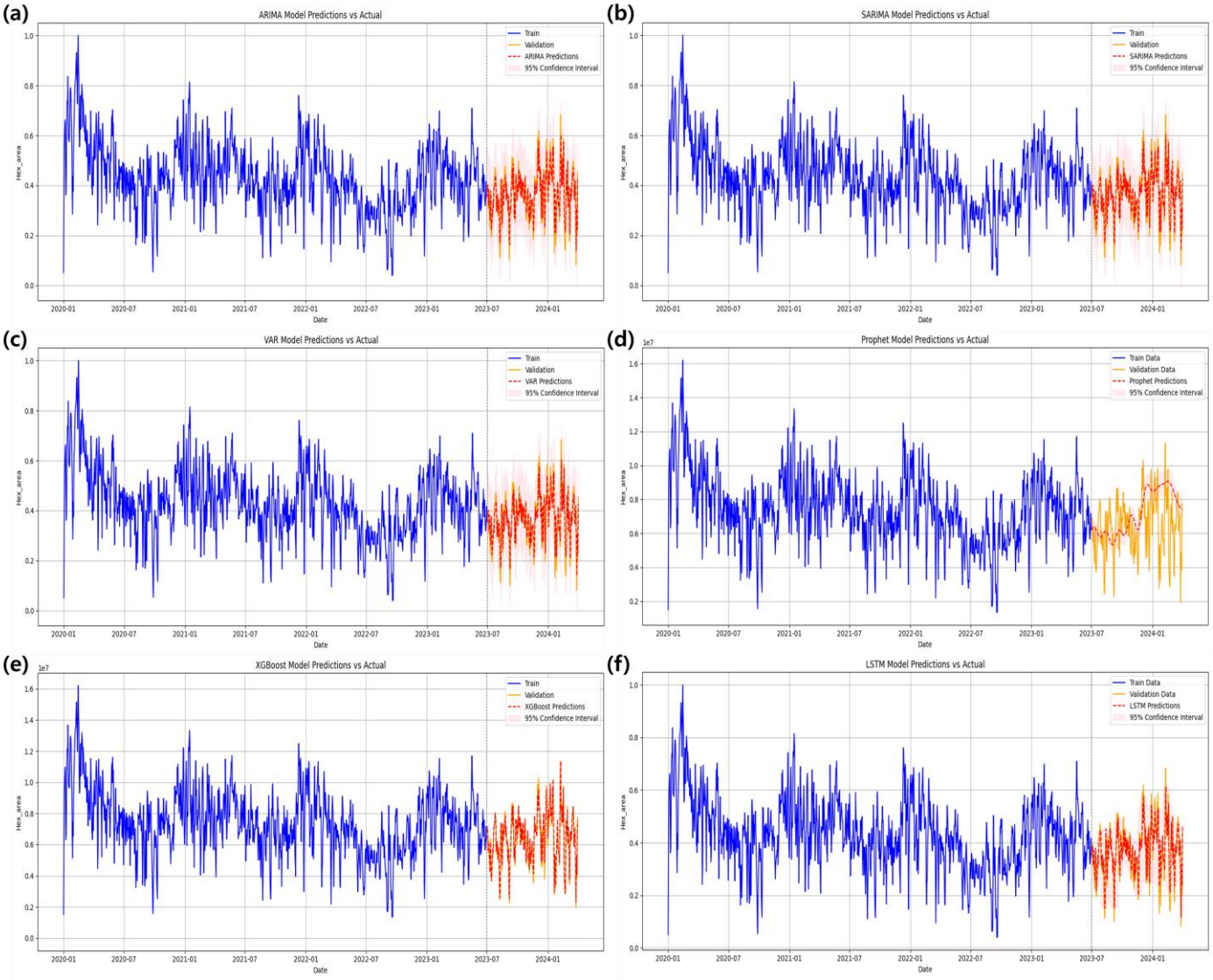

Fig. 7. Forecasting results of six predictive models compared to actual values. Graph comparing a) ARIMA, b) SARIMA, c) VAR, d) Prophet, e) XGBoost, and f) LSTM predictions with actual values with 95 % bootstrap confidence interval

The machine learning-based XGBoost model demonstrated the highest forecasting accuracy among the individual models (Table 6), achieving the lowest values for MAE (0.0132), RMSE (0.0171), and MAPE (3.37 %). The Prophet model also performed well, with a MAPE of 6.79 %. In contrast, conventional time-series forecasting models such as ARIMA, SARIMA, and VAR produced higher MAPE values (~16–17 %), highlighting their limitations in capturing the complex and nonlinear characteristics of anchorage occupancy area data. The deep learning-based LSTM model showed slightly improved performance over conventional models but underperformed relative to the machine learning-based models.

Table 6 Performance parameters for forecasting models

Model	MAE	RMSE	MAPE (%)	MAPE CI 95 % (%)	
				Lower Bound	Upper Bound
ARIMA	0.0511	0.0634	16.80	0.78	88.11
SARIMA	0.0512	0.0634	16.82	0.80	89.65
VAR	0.0513	0.0643	17.33	0.77	93.92
Prophet	0.0229	0.0285	6.79	6.13	15.50
XGBoost	0.0132	0.0171	3.37	18.72	31.76
LSTM	0.0502	0.0624	13.77	19.39	29.44

The analysis of individual model forecasting performance and runtime revealed a clear trade-off between accuracy and computation time. Classical statistical models (ARIMA, SARIMA, VAR) showed a linear increase in runtime as data size grew, making them vulnerable to computational load and resulting in relatively lower accuracy. In contrast, Prophet's accuracy improved significantly with more data, and its runtime of 0.2-0.7 seconds was very short, making it highly suitable for real-time applications. XGBoost emerged as a practical alternative, showing the most consistently low errors across all ranges and demonstrating excellent computational efficiency. The deep learning model, LSTM, showed improved performance when sufficient data was secured, but it required the longest computation time, exceeding 1844 seconds, indicating the greatest computational burden when scaling data (Table 7).

Table 7 Runtime comparison of base models based on training data volume

	ARIMA		SARIMA		VAR		Prophet		XGBoost		LSTM	
Data Ratio (%)	MAPE (%)	Run (s)	MAPE (%)	Run (s)	MAPE (%)	Run (s)	MAPE (%)	Run (s)	MAPE (%)	Run (s)	MAPE (%)	Run (s)
25	16.64	283.8	16.87	197.1	18.52	2.4	44.31	0.2	5.44	0.2	24.46	73.0
50	16.88	395.1	16.98	326.3	18.38	4.9	43.17	0.3	4.61	0.3	24.81	114.6
75	16.75	518.5	16.86	418.1	17.64	1.9	29.57	0.6	3.88	0.3	24.51	155.8
100	16.8	628.6	16.82	530.6	17.33	4.4	6.79	0.7	3.37	0.4	13.77	1844.1

5.3 Forecasting results using stacking ensemble models

To further enhance predictive accuracy, a stacking ensemble model was developed, integrating the strengths of multiple forecasting models. Four ensemble strategies were tested: simple ensemble, Random Forest stacking, XGBoost stacking, and gradient boosting stacking. The forecasting results for each meta model were visualized and compared (Figure 8).

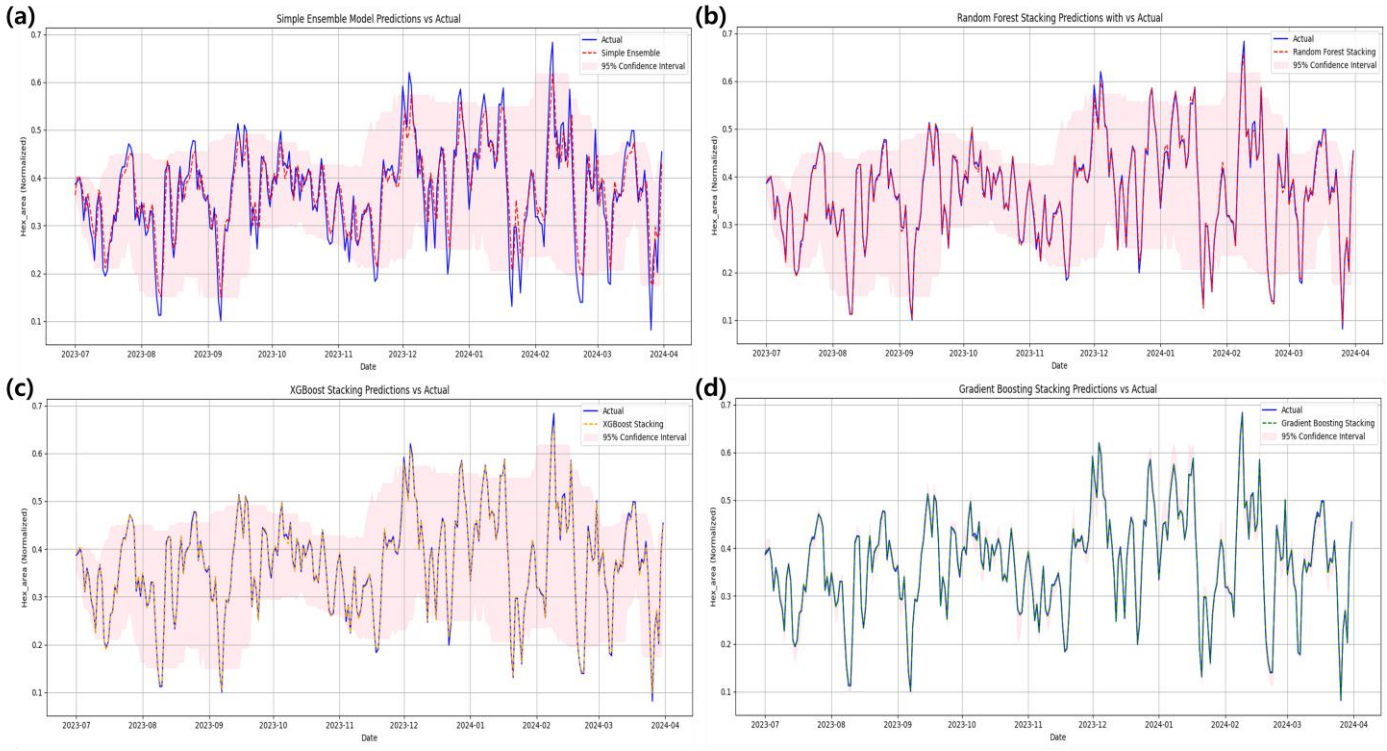

Fig. 8. Forecasting results of stacking ensemble models compared to actual values. Graph comparing (a) simple ensemble, (b) random forest stacking, (c) XGBoost stacking, and (d) gradient boost stacking predictions with actual values with 95 % bootstrap confidence interval

Forecasting performance results are summarized in Table 8. The gradient boosting stacking model demonstrated the best performance, with an MAE of 0.0021, RMSE of 0.0026, and MAPE of 0.59 %. Both the XGBoost and Random Forest stacking models also recorded low error rates, demonstrating strong predictive capabilities. Overall, the stacking ensemble models consistently outperformed individual models, confirming their superior forecasting performance. The simple ensemble method, while effective, showed relatively higher error rates compared to the other stacking models.

Table 8 Performance parameters of stacking ensemble models

Model	MAE	RMSE	MAPE (%)	MAPE CI 95 % (%)	
				Lower Bound	Upper Bound
Simple ensemble	0.0369	0.0473	12.37	30.51	41.73
Random Forest stacking	0.0056	0.0075	1.63	1.30	1.66
XGBoost stacking	0.0037	0.0051	1.13	0.93	1.27
Gradient boosting stacking	0.0021	0.0026	0.59	0.49	0.59

The analysis of ensemble forecasting strategies further highlighted the trade-off between accuracy and computation time (Table 9). Random Forest stacking exhibited stable accuracy with moderate runtime growth. Gradient Boosting stacking consistently achieved the highest accuracy with runtime remaining in the range of 0.22–0.59 s, indicating a favorable trade-off. Notably, XGBoost stacking demonstrated excellent computational scalability, as runtime decreased with larger data sizes while maintaining competitive accuracy.

Overall, all stacking models scaled sub-linearly with data size, showing robustness in computational efficiency. Compared to the simple ensemble, stacking models required greater computation but provided markedly superior predictive accuracy, with Gradient Boosting delivering the most accurate forecasts and XGBoost offering the most balanced trade-off between accuracy and runtime.

Table 9 Runtime comparison of meta models based on training data volume

	Simple ensemble		**Random Forest stacking**		**XGBoost stacking**		**Gradient boosting stacking**	
Data Ratio (%)	MAPE (%)	Run (s)	MAPE (%)	Run (s)	MAPE (%)	Run (s)	MAPE (%)	Run (s)
25	11.85	0.004	1.45	0.426	0.63	0.887	0.22	0.434
50	10.46	0.003	1.34	0.428	0.66	0.074	0.25	0.225
75	10.95	0.003	1.60	0.235	0.87	0.130	0.44	0.267
100	12.37	0.003	1.63	0.269	1.13	0.123	0.59	0.319

5.4 Feature importance analysis

To evaluate the contribution of each base model within the stacking ensemble, a feature importance analysis was conducted. As shown in Figure 9, the Prophet model had the most significant influence on the final predictions, followed by relatively high contributions from XGBoost and VAR. In contrast, LSTM, ARIMA, and SARIMA contributed minimally to the ensemble's performance.

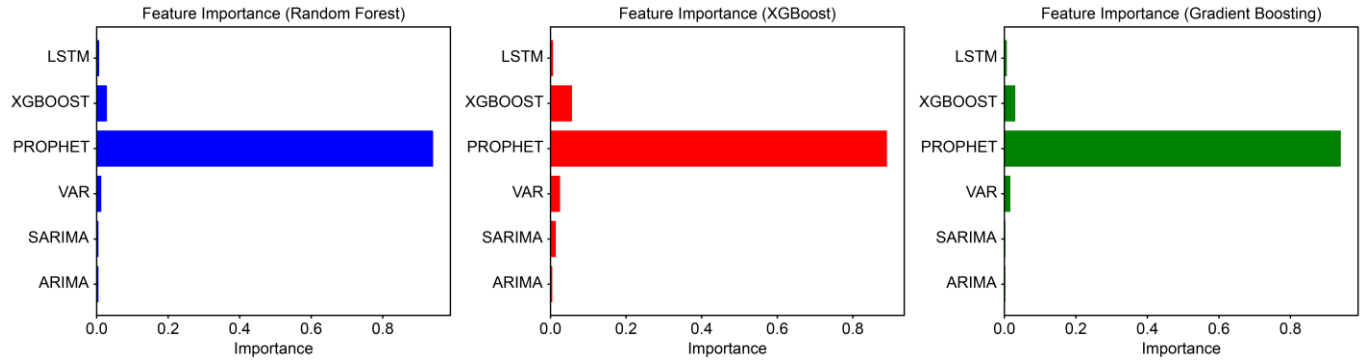

Fig. 9. Feature importance analysis of base models in meta-learning

These results confirm that the predictive powers of the machine learning models, and Prophet were more heavily weighted within the ensemble model.

5.5 Statistical Significance Testing

To statistically validate the performance improvement of the ensemble model, two complementary tests were conducted: the paired t-test and the Diebold-Mariano (DM) test. These tests compared the prediction errors of the best-performing single model—XGBoost—with those of the best-performing ensemble model—Gradient Boosting Stacking. As shown in Table 10, the paired t-test, which evaluates whether the mean prediction errors of two dependent models differ significantly, yielded a test statistic of –10.1854 with a p-value < 0.0001. This result indicates that the Gradient Boosting Stacking model produced significantly lower mean absolute errors than the XGBoost model.

In addition, the Diebold-Mariano test, which is specifically designed to compare the forecasting accuracy of time-series models, produced a test statistic of –6.0602 with a p-value < 0.0001. This result further confirms that the performance gains observed in the Gradient Boosting Stacking model are statistically significant and unlikely to be due to random variation. Together, these statistical tests provide robust evidence that the Gradient Boosting Stacking ensemble model offers a significant improvement in forecasting performance over the best single-model baseline.

Table 10 Statistical significance test results comparing Gradient Boosting Stacking and XGBoost models

Test Type	Statistic	p-value	Interpretation
Paired t-test	-10.1854	< 0.0001	The difference in mean prediction errors is statistically significant ($p < 0.05$)
Diebol-Mariano Test	-6.0602	< 0.0001	The difference in forecast accuracy is statistically significant ($p < 0.05$)

6. Discussion

This study addresses the research question of whether a stacked ensemble approach can overcome the limitations of conventional time-series forecasting models. The aim is to improve the accuracy of anchorage occupancy forecasting. Specifically, the study evaluated the effectiveness of stacking ensemble models in integrating predictions from both conventional statistical and machine learning models. While also accounting for environmental conditions and vessel-specific characteristics within real-world port operations.

These results indicate several noteworthy findings. First, although deep learning models such as LSTM are known for their capacity to model complex temporal dependencies, the LSTM model in this study underperformed, achieving a MAPE of 13.77 %. This result is unexpected and warrants further analysis. While LSTM is effective at capturing long-range sequential patterns, its underperformance may be attributed to several factors: (a) the dataset, although substantial, may not be large enough to fully utilize LSTM's high model complexity; (b) anchorage occupancy dynamics may be more influenced by discrete, event-driven dynamics and spatial interactions, rather than the purely sequential dependencies LSTM models are optimized to learn; and (c) real-world port operation data may contain high levels of noise or abrupt changes that hinder LSTM's ability to generalize effectively. These challenges suggest that while LSTM is powerful in certain contexts, it may be less suited to the unique, multifaceted nature of anchorage occupancy forecasting.

Moreover, the simple ensemble approach—where model predictions were averaged equally—did not achieve the desired level of forecasting accuracy (MAPE=12.37 %). This result highlights a key limitation of naive ensemble strategies: equal weighting of all models fails to account for performance variability among individual models, thereby diluting the benefits of ensemble learning.

However, the proposed stacked ensemble approach significantly improved the prediction accuracy compared to individual models. Among the standalone models, XGBoost (MAPE=3.37 %) and Prophet (MAPE=6.79 %) demonstrated superior performance, effectively capturing the nonlinear and complex temporal patterns characteristic of anchorage occupancy data. Conversely, conventional time-series forecasting models such as ARIMA, SARIMA, and VAR exhibited higher error rates, reinforcing their limitations in modeling the intricate and dynamic nature of anchorage usage.

To confirm the statistical significance of the performance improvement, both a paired t-test and the Diebold-Mariano test were conducted, comparing the XGBoost model with the Gradient Boosting Stacking ensemble. The paired t-test revealed a statistically significant difference in mean absolute errors ($t = -10.1854$, $p < 0.0001$), while the Diebold-Mariano test confirmed a significant improvement in forecasting accuracy ($DM = -6.0602$, $p < 0.0001$). These results provide strong statistical evidence that the ensemble model's superior performance is robust and not attributable to random variation.

Among the stacking ensemble models, the Gradient Boosting Stacking ensemble achieved the highest accuracy (MAPE = 0.59 %), while the Random Forest and XGBoost stacking models also demonstrated strong predictive performance. These findings demonstrate that integrating diverse models within an ensemble framework significantly enhances overall forecasting accuracy.

A feature importance analysis further highlighted the key contributions of the Prophet, XGBoost, and VAR models within the ensemble. Prophet's prominent role underscores the relevance of underlying trends and seasonality in anchorage occupancy data—components that Prophet is particularly well-equipped to model. XGBoost's substantial contribution reflects its strength in capturing complex, nonlinear relationships and adapting to dynamic fluctuations, such as abrupt shifts in demand or operational events. VAR's significant role indicates the importance of multivariate dependencies, particularly the strong correlation between daily occupied space and vessel count, which VAR effectively modeled.

Conversely, the relatively lower contributions of ARIMA, SARIMA, and LSTM suggest that, while each has merits in specific contexts, their individual strengths were less influential in capturing the multifaceted dynamics of anchorage occupancy compared to other models in the ensemble. The complementary strengths of the base models enabled the stacked ensemble to exploit their unique capabilities. This integration resulted in superior predictive accuracy.

Additionally, a simulation using real-world data from the E1 anchorage at Ulsan Port revealed that the hexagonal-based occupancy estimation method required approximately 10.3 % more space than the traditional anchor circle method. This increase represents a more conservative yet operationally safer approach to anchorage management.

This study demonstrates significant improvements in forecasting accuracy and proposes a more practical method for spatial anchorage management. However, several challenges remain in translating these findings into real-world implementation. Integrating disparate real-time data streams from various port systems and ensuring their interoperability remains a significant technical hurdle. Furthermore, deploying and maintaining sophisticated stacking ensemble models—which require considerable computational resources for training and real-time inference—demands robust IT infrastructure. Such resources may not be readily available within many port authorities. Additionally, adapting existing port regulations and operational protocols to accommodate new methodologies, such as the hexagon-based occupancy estimation method, may require coordinated policy reform and active stakeholder collaboration.

This study has several limitations. First, the anchorage occupancy data were limited to the E1 anchorage of Ulsan Port—a major liquid cargo hub characterized by persistent shortage of anchorage capacity and elevated safety risks in Group E. While this focus enabled an in-depth case study and effective validation of the proposed methodologies under real-world conditions, it inherently restricts the generalizability of the findings. Although the model demonstrated strong performance for Ulsan Port, caution should be exercised when extrapolating the results to other ports with different operational profiles. Moreover, the *LOA* estimation method was validated only under the specific operational context of Ulsan Port and would require additional validation and calibration before being applied elsewhere. The simulations also did not fully capture detailed vessel interactions or the nuanced dynamics of real-world anchoring practices. These gaps may limit the operational applicability of the findings. While the hexagon-based occupancy estimation method improved spatial representation and safety margins, further research is needed to incorporate actual anchoring behaviors and variability in vessel turning radii under dynamic environmental conditions.

Furthermore, to better assess the generalizability of the proposed framework beyond a single port, future studies could incorporate simulation-based validation using synthetic datasets representing diverse port characteristics. Such simulations would allow systematic testing under varying traffic densities, anchorage configurations, and regulatory environments, complementing real-world case studies and helping to identify context-specific limitations.

Another important limitation involves the computational cost and scalability of the proposed framework. Although computational efficiency was addressed through strategies such as parallelized training for XGBoost and mini batch learning with early stopping for LSTM, a comprehensive analysis of runtime performance was not conducted. In particular, quantitative scalability testing across different data volumes or system complexities was not conducted. Such testing is essential for deploying the framework in larger and more complex port environments.

This study also focused mainly on generating accurate point forecasts for anchorage occupancy. However, it did not address prediction intervals or provide a comprehensive assessment of forecast uncertainty. While point forecasts are essential for direct operational planning, understanding the uncertainty associated with these predictions is equally important for informed decision-making, risk assessment, and contingency planning in dynamic port environments. This omission represents another limitation of the present study.

7. Conclusion

This study aimed to enhance anchorage occupancy forecasting under dynamic port conditions. To achieve this, a stacked ensemble learning approach was proposed, integrating statistical and machine learning models. The ensemble method outperformed individual forecasting models. Among them, the Gradient Boosting Stacking model achieved the best performance (MAPE = 0.59 %). Statistical validation using the paired t-test and Diebold-Mariano test confirmed that the ensemble model was superior to the best-performing single model, XGBoost.

Feature importance analysis revealed that Prophet, XGBoost, and VAR contributed most significantly to the ensemble's performance. These models effectively captured seasonality, nonlinear patterns, and multivariate dependencies within the anchorage occupancy data. Additionally, a simulation-based comparison showed that the hexagon-based occupancy estimation method required approximately 10.3 % more area than the traditional anchor circle method. This approach provides a more conservative yet operationally safer alternative for spatial planning in port operations.

Despite these advancements, this study has several limitations. Key limitations include:

- Reliance on data from a single port, potentially limiting generalizability to other port environments.
- Focus on point forecasts without incorporating prediction intervals or formal uncertainty quantification.
- Lack of quantified computational scalability, with no analysis of performance across varying system sizes or data volumes.
- Simplified spatial simulations that do not fully capture dynamic vessel interactions or specific maneuvering behaviors.
- Uncertainty in *LOA* estimations without explicit error propagation analysis.
- Application of PIANC standards without validation under localized Korean operational and regulatory conditions.

To enhance the robustness and applicability of the proposed framework, future research should first explore its adaptation to a wider range of ports to assess generalizability across diverse maritime environments. Incorporating uncertainty quantification and statistical validation will be critical for producing reliable prediction intervals and improving decision-making confidence. Enhancing the spatial modeling component—through the inclusion of realistic anchoring behaviors, variable turning radii, and vessel-specific interaction dynamics—will further refine the accuracy of the hexagon-based occupancy estimation approach. In addition, simulation-based experiments using synthetic datasets that mimic heterogeneous port conditions can provide a controlled environment for testing model robustness, stress-testing under extreme scenarios, and identifying potential limitations prior to real-world deployment. In terms of implementation, research should investigate scalable computational strategies such as cloud-based infrastructures or hardware acceleration to support real-time application in operational settings. Given the limited performance of the LSTM model in this study, alternative deep learning architectures—such as Transformer-based models—can offer improved forecasting capabilities under highly dynamic settings. Finally, collaboration with port authorities for pilot implementation will be essential to evaluate the practical utility of the proposed system in improving berth allocation, reducing congestion, and supporting data-driven anchorage management.

By addressing these avenues, the proposed framework holds strong potential to evolve into a robust, scalable, and deployable decision-support tool for anchorage management in complex and uncertain port conditions.

ACKNOWLEDGMENTS

The authors thank Joo Sung Kim of Mokpo National Maritime University for his invaluable assistance throughout this study.

FUNDING

This research was supported by the National University Development Project through the National Research Foundation of Korea, funded by the Ministry of Education.

REFERENCES

[1] Park, J. M., Yun, G. H., Kang, M. K., Lee, Y. S., 2021. Improvement plan for Ulsan anchorage based on adequacy evaluation criteria. *Journal of the Korean Society of Marine Environment and Safety*, 27, 247-255. https://doi.org/10.7837/kosomes.2021.27.2.247

[2] Park, J. M., Yun, G. H., Jeon, H. D., Kong, G. Y., 2016. The proper capacity of anchorage in Ulsan Port with reference to the anchorage operating rate. *Journal of the Korean Society of Marine Environment and Safety,* 22, 380-388. https://doi.org/10.7837/kosomes.2016.22.5.380

[3] Lee, C.- H., Lee, H. H., 2014. A study on expansion of anchorage according to increased trading volume at Pyeongtaek Port. *Journal of the Korean Society of Marine Environment and Safety*, 20, 663-670. https://doi.org/10.7837/kosomes.2014.20.6.663

[4] Huang, S. Y., Hsu, W. J., He, Y., Song, T., Charles, D. S., Rong, Y., Chuanyu, C., Stuti, N., 2009. Anchorage capacity analysis using simulation. In *Proceedings of the International Conference on Harbor, Maritime & Multimodal Logistics Modeling and Simulation* (HMS 2009), 23-25 September, Puerto de la Cruz, Spain, 100-105.

[5] Kwon, S.-C., Yu, Y.-U., Park, J.-M., Lee, Y.-S., 2019. A basic study on the demand analysis of waiting anchorage using anchorage capacity index. *Journal of the Korean Society of Marine Environment and Safety*, 25, 519-526. https://doi.org/10.7837/kosomes.2019.25.5.519

[6] Korean Maritime Safety Tribunal, Status of Marine Accidents, 2025. https://kmst.go.kr/ (accessed 22th July 2025).

[7] Prince Rupert Port Authority, 2020. Marine navigational and anchorage areas risk assessment: Final report. *Dillon Consulting Limited*.

[8] Port and Harbour Marine Safety Code, 2022. Establishing and managing anchorages under the Port and Harbour Marine Safety Code. *New Zealand Government*.

[9] Ozsari, I., 2023. Predicting main engine power and emissions for container, cargo, and tanker ships with artificial neural network analysis, *Brodogradnja*, 74(2), 77-94. https://doi.org/10.21278/brod74204

[10] Atasayan, E., Milanov, E., Alkan A., 2024. Application of an offline grey box method for predicting the manoeuvring performance, *Brodogradnja,* 75(3), 75304. https://doi.org/10.21278/brod75304

[11] Veenstra, A., Franses, P., 1997. A co-integration approach to forecasting freight rates in the dry Bulk shipping sector. *Transportation Research Part A: Policy and Practice*, 31, 447-458. https://doi.org/10.1016/S0965-8564(97)00002-5

[12] Kavussanos, M., Alizadeh, A., 2002. Seasonality patterns in tanker spot freight rate markets. *Economic Modelling,* 19, 747-782. https://doi.org/10.1016/S0264-9993(01)00078-5

[13] Chen, S., Meersman, H., Voorde, E., 2012. Forecasting spot rates at main routes in the dry bulk market. *Maritime Economics & Logistics*, 14, 498-537. https://doi.org/10.1057/mel.2012.18

[14] Schramm, H. J., Munim, Z., 2021. Container freight rate forecasting with improved accuracy by integrating soft facts from practitioners. *Research in Transportation Business & Management,* 41, 100662. https://doi.org/10.1016/j.rtbm.2021.100662

[15] Peng, W.Y., Chu, C.-W., A comparison of univariate methods for forecasting container throughput volumes. *Mathematical and Computer Modelling* 2009, *50*, 1045-1057. https://doi.org/10.1016/j.mcm.2009.05.027

[16] Rashed, Y., Meersman, H., Voorde, E., Vanelslander, T., 2017. Short-term forecast of container throughout: An ARIMA-intervention model for the port of Antwerp. *Maritime Economics & Logistics,* 19, 749-764. https://doi.org/10.1057/mel.2016.8

[17] Farhan, J., Ong, G., 2018. Forecasting seasonal container throughput at international ports using SARIMA models. *Maritime Economics & Logistics*, 20, 131-148. https://doi.org/10.1057/mel.2016.13

[18] Kim, H.- S., Shin, J.- M., 2022. Development of an intelligent coastal surveillance system using deep learning: LSTM model. *Korean Journal of Military Art and Science*, 78, 471-496.

[19] Fan, H. S., Cao, J. M., 2000. Sea space capacity and operation strategy analysis system. *Transport Planning and Technology*, 24, 49-63. https://doi.org/10.1080/03081060008717660

[20] Oz, D., Aksakalli, V., Alkaya, A.F., Aydogdu, V., 2015. An anchorage planning strategy with safety and utilization considerations. *Computers & Operations Research*, 62, 12-22. https://doi.org/10.1016/j.cor.2015.04.006

[21] Shin, G.-H., Yang, H., 2025. Deep reinforcement learning for integrated vessel path planning with safe anchorage allocation, *Brodogradnja*, 76(3), 76305. https://doi.org/10.21278/brod76305

[22] Gifty, A., Li, Y., 2024. A comparative analysis of LSTM, ARIMA, and XGBoost algorithms in predicting stock price direction. *Engineering and Technology Journal*, 9(8), 4978-4986. https://doi.org/10.47191/etj/v9i08.50

[23] Zhang, L., Bian, W., Qu, W., Tuo, L., Wang, Y., 2021. Time series forecast of sales volume based on XGBoost. *Journal of Physics: Conference Series*, 1873, 012067. https://doi.org/10.1088/1742-6596/1873/1/012067

[24] Swami, D., Shah, A. D., Ray, S. K. B., 2020. Predicting future sales of retail products using machine learning. arXiv:2008.07779. https://doi.org/10.48550/arXiv.2008.07779

[25] Chen, Z., Goh, H. S., Sin, K. L., Lim, K., Chung, N. K., Liew, X. Y., 2021. Automated agriculture commodity price prediction system with machine learning techniques. *Advances in Science, Technology and Engineering Systems Journal,* 6, 376-384. https://doi.org/10.25046/aj060442

[26] Mukhlis, M., Kustiyo, A., Suharso, A., 2021. Peramalan Produksi Pertanian Menggunakan model long short-term memory. *Bina Insani ICT Journal,* 8, 22-32. https://doi.org/10.51211/biict.v8i1.1492

[27] Hadri, S., Naitmalek, Y., Najib, M., Bakhouya, M., Fakhri, Y., Elaroussi, M., 2019. A comparative study of predictive approaches for load forecasting in smart buildings. *Procedia Computer Science,* 160, 173-180. https://doi.org/10.1016/j.procs.2019.09.458

[28] Tang, H., Zhu, R., Wan, Qian., Ren, Deyuan., 2025. Short-term prediction of trimaran load based on data driven technology, *Brodogradnja*, 76(1), 76101. https://doi.org/10.21278/brod76101

[29] Su, S., Miao, Z., Zhao, Y., Song, N., 2024. Digitally twin driven ship cooling pump fault monitoring system and application case, *Brodogradnja*, 75(4), 75403. https://doi.org/10.21278/brod75403

[30] Zheng, H., Sherazi, S. W. A., Lee, J. Y., 2021. A stacking ensemble prediction model for the occurrences of major adverse cardiovascular events in patients with acute coronary syndrome on imbalanced data. *IEEE Access,* 9, 113692-113704. https://doi.org/10.1109/ACCESS.2021.3099795

[31] Zhan, Y. Zhang, H., Li, J., Li, G., 2022. Prediction method for ocean wave height based on stacking ensemble learning model. *Journal of Marine Science and Engineering*, 10, 1150. https://doi.org/10.3390/jmse10081150

[32] Hoque, X., Sharma, S., 2019. Ensembled deep learning approach for maritime anomaly detection system. *In Proceedings of ICETIT 2019: Lecture Notes in Electrical Engineering*, 605, 862-869. https://doi.org/10.1007/978-3-030-30577-2_77

[33] Bodunov, O., Schmidt, F., Martin, A., Brito, A., Fetzer, C., 2018. Real-time Destination and ETA Prediction for Maritime Traffic. *Proceedings of the 12th ACM international conference on distributed and event-based systems,* 198-201. https://doi.org/10.1145/3210284.3220502

[34] Kim, H. C., 2024. A study on application of stacking ensemble machine learning algorithm for improving wind power prediction accuracy. *Master Thesis*, Jeju National University, Jeju, Republic of Korea.

[35] Wi, C. H., 2023. Optimal design of permanent magnet assisted synchronous reluctance motor for electric vehicle drive using surrogate model based on stacking ensemble. *Master Thesis*, University of Ulsan, Ulsan, Republic of Korea.

[36] Seo, J. Y., 2024. Enhancing credit default prediction model using stacking ensemble and anomaly detection techniques: a case study on lending club P2P loan data, *Master Thesis*, Chonnam National University, Gwangju, Republic of Korea.

[37] Gupta, A., Jain, V., Singh, A., 2021. Stacking ensemble-based intelligent machine learning model for predicting post-COVID-19 complications. *New Generation Computing,* 40, 987-1007. https://doi.org/10.1007/s00354-021-00144-0

[38] Ministry of Oceans and Fisheries, 2024. Guidelines of port and harbor design, Korea Design Standard 64 40 10. 14.

[39] Korea Hydrographic and Oceanographic Agency, 2021. WIND SPEED, Report of Real-time Korea Oceanogaraphic Observations.

[40] Ulsan Regional Office of Oceans and Fisheries, Ulsan Port Facilities Status, 2025. https://ulsan.mof.go.kr/ (accessed 13th January).

[41] Breiman, L., 1996. Stacked regressions. *Machine Learning*, 24, 49-64. https://doi.org/10.1023/A:1018046112532

[42] Taylor, S. J., Letham, B., 2018. Forecasting at scale. *The American Statistician*, 72, 37-45. https://doi.org/10.1080/00031305.2017.1380080

[43] Box, G. E. P., Jenkins, G. M., 1970. Time series analysis: forecasting and control. Holden-Day, San Francisco.

[44] Samal, K. K. R, Babu, K. S., Das, S., Acharya, A., 2019. Time series based air pollution forecasting using SARIMA and the Prophet model. In the *Proceedings of the 2019 International Conference on Information Technology and Computer Communications* (ITCC 2019), Singapore, 16-18 August, 80-85. https://doi.org/10.1145/3355402.3355417

[45] Sims, C., 1980. Macroeconomics and reality. *Econometrica*, *48*, 1-48. 10.2307/1912017. https://doi.org/10.2307/1912017

[46] Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD '16), San Francisco, CA, USA, 13-17 August 2016; Association for Computing Machinery: New York, NY, USA, 785-794. 785-794. https://doi.org/10.1145/2939672.2939785

[47] Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Computation*, 9, 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[48] Opitz, D., Maclin, R., 1999. Popular ensemble methods: an empirical study. *Journal of Artificial Intelligence Research*, 11, 169-198. https://doi.org/10.1613/jair.614

[49] Breiman, L., 2001. Random forests. *Machine Learning*, 45, 5-32. https://doi.org/10.1023/A:1010933404324

[50] Wolpert, D., 1992. Stacked generalization. *Neural Networks*, 5, 241-259. https://doi.org/10.1016/S0893-6080(05)80023-1

[51] Friedman, J., 2001. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29, 1189-1232. https://doi.org/10.1214/aos/1013203451

[52] Lee, Y.-S., 2014. A study on the anchoring safety assessment of E-group anchorage in Ulsan Port. *Journal of the Korean Society of Marine Environment and Safety*, 20, 172-178. https://doi.org/10.7837/kosomes.2014.20.2.172

[53] Port-MIS, Port Management Information System, 2025. https://portmis.go.kr/ (accessed 13th January 2025).

[54] McBride, M., Briggs, M., Groenveld, R., Boll, M., Cao, L., Cockrill, D., et al., 2014. PIANC Report No. 121 - 2014, Harbour Approach Channels - Design Guidelines.