Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMPARATIVE MACHINE LEARNING MODELS FOR DATA-DRIVEN VESSEL TURNAROUND TIME PREDICTION AT TANGER MED PORT

MOSTAFA AL UAHABI[1], HICHAM ATTARIUAS[2]

[1] PhD student, Department of System Engineering, Faculty of Sciences, Abdelmalek Essaadi University, Tetouan, Morocco

[2] Professor, Department of System Engineering, Faculty of Sciences, Abdelmalek Essaadi University, Tetouan, Morocco

E-mail: [1]mostafa.aluahabi@etu.uae.ac.ma, [2]hattarius@uae.ac.ma

ABSTRACT

Maritime shipping underpins global trade, handling more than 90% of global goods movement and links markets across continents. Ports sit at the core of this system and their effectiveness greatly impacts the smooth functioning of worldwide supply chains. Even though larger ships have boosted efficiency per call, they require more docking time. This strain on terminal infrastructure makes Vessel Turnaround Time (VTT) a crucial metric for container terminals (TCs). Existing research predicting VTT often suffers from limited timeframes and a lack of comprehensive variables. This research creates and contrasts eight machine learning regression models to predict VTT at the second TC in Tanger Med Port, representing one of the initial studies to forecast VTT at this specific location using 12 years of Terminal Operating System data. Following a detailed preprocessing phase and feature selection using RFECV, seven significant predictive variables were isolated from an initial set of 114, concentrating on cargo handling volume, crane productivity and operational delays within the terminal. Hyperparameters for each model underwent tuning with a grid search incorporating 5-Fold cross-validation. The best-performing model was BPNN (R^2 = 0.9741, RMSE = 2.40 h, MAE = 1.61 h). These findings illustrate that a streamlined, data-driven feature set markedly enhances VTT prediction accuracy, offering terminal operators a dependable decision-support tool for operational planning.

Keywords: *Vessel Turnaround Time, Machine Learning, Data-Driven Prediction, Feature Selection, Tanger Med Port.*

1. INTRODUCTION

Maritime transport is at the heart of world trade. For ages it has powered the forces of globalization. It plays the role as an artery of commerce across the world and ships goods across the oceans to provide an efficient and cost-effective way to move them. 90% of trade by volume across the world is handled by the shipping industry [1]. Shipping is one of the key players for world trade carrying out mass movement of goods to promote global trade, linking economies and markets to a global network.

Ports assume a crucial and prominent role as they are important nodes in the global supply chain, facilitating trade between countries and affecting global supply chains performance, driving trade and stimulating economic growth. Hence the fortunes of ports have long been tied to their ability to provide services and handle the ever-increasing number of ships and amount of cargo that passes through them.

Container terminals (CT), part of ports, sort and organise large quantities of goods that have travelled on several different vessels enabling logistics and trade flows to flow. Vessels have increased in size because greater carrying capacity improves economy but this means they remain in port for longer with greater strain placed on port infrastructure resources[2].

The time a vessel stays in port—its Vessel Turnaround Time (VTT), calculated from its arrival in an anchorage until the time of departure —is taken as one of the clearest indicators of performance and productivity of a container terminal [1] . Efficient ports are adept at quickly turning ships around, even when handling high volumes, which highlights the superiority of their operational capabilities [3].

The Port of Tanger-Mediterranean is located at the crossroads of the busiest trade routes in the world, situated directly on the Strait of Gibraltar; this port holds significance as one of the largest ports in the Mediterranean and Africa, playing a twin role as a competitive node within global supply chains and as an engine for Moroccan logistical and industrial competitiveness. This importance is not matched, however, by the research attention the port has received: no prior study has examined VTT at Tanger Med specifically, apart from our earlier contribution [4], in which VTT is addressed only as a secondary concern.

Although the importance of port efficiency is increasingly recognized, research on VTT—an indicator that remains nonetheless central to terminal performance—continues to lag on several fronts. Compared to other industry applications, VTT problems have not yet received adequate attention from data-driven ML models. Existing studies often rely on datasets that are insufficiently detailed and weakly operational, partly because terminal operators are reluctant to share such data for reasons of confidentiality and operational security; these datasets also tend to cover restricted time periods with a limited number of variables, while studies conducted in specific port contexts exhibit uncertain transferability to other environments. Weather conditions likewise remain insufficiently integrated into existing models [5]. Furthermore, whether relying on classical statistical methods or on machine learning, studies rarely compare a sufficient number of competing models, limiting the systematic evaluation of their relative performance. Finally, the literature tends to focus more on identifying the factors influencing VTT than on developing a complete predictive framework [6].

In current practice, turnaround-time estimates rely largely on operator experience, historical averages and deterministic scheduling rules, or on estimates reported by terminal systems; on our dataset, the terminal's own estimates reach an R^2 of 0.92, which sets the reference that the models developed here must improve upon. This study therefore takes the maritime traffic data captured by the terminal operating system (TOS) and builds a complete regression framework for VTT prediction, combining systematic preprocessing, feature selection and a comparison of several model families. Eight machine-learning models were used to build the regression models, each configured through grid search and cross-validation, and the analysis additionally examines feature importances and their correlations. The best model reaches an R^2 of 0.97, against 0.92 for the terminal's own estimates. It is this systematic comparison that allows the best-performing model to be justified rather than merely reported, and hence to be proposed as a predictive tool reliable enough to inform terminal operations.

The remaining parts of this article are organised as follows. Section 2 reviews the literature on vessel turnaround time prediction and on modelling port performance. Section 3 introduces the Port of Tanger Med and the background of the case study. Section 4 sets out the data, covering the research flow, the data source and the preprocessing pipeline. Section 5 presents the exploratory data analysis, the feature selection and the machine learning models for regression tasks, together with their hyperparameter optimisation and assessment metrics. Section 6 reports the prediction results and discusses them, and Section 7 concludes with the main outcomes of the study, its limitations and directions for future research.

2. LITERATURE REVIEW

VTT is closely tied to the berth allocation problem, in which the central objective is to allocate a specific berthing time and position to each vessel, while minimizing handling time, subject to constraints of space and time [7], [8]. Effective management of VTT reflects well-organized vessel scheduling that enables the swift handling of vessel operations, minimizes delays, and ensures optimal resource utilization. As such, VTT directly affects a port's operational efficiency, its competitiveness and economic viability. Achieving this level of efficiency requires ports to continuously strengthen their planning and operational capabilities through the deployment of innovative equipment and state-of-the-art technology aimed at optimizing terminal logistics processes.

Significant research studies have examined the sources of uncertainty in vessel operation timing. In [5], the authors developed a VTT prediction approach based on the XGBoost algorithm, achieving a 23% reduction in MAE and a 24% reduction in RMSE, while also deriving policy recommendations on how port efficiency and green shipping operations can be improved. Yoon et al. [9] presents a comparative study predicting vessel dwell times at Busan New Port using several machine learning models, while Kourounioti et al. [10] proposes a data-mining-based framework to forecast ship arrivals and dwell time at Yining container terminal.

The ground for this research was set earlier in [11], where regression models were used to measure VTT at Klang Port, finding high and significant correlations between VTT and crane allocation and container volume. A related study [12] focuses specifically on petroleum terminal operations, looking for controllable operational factors to reduce the level of VTT, whereas Jayaprakash et al. [13] develop non-linear regression models assuming that service time is a better measure of overall port performance than VTT.

Beyond these foundational contributions, numerous studies have investigated the specific factors driving VTT. In [14], the association between VTT and total moves and discharge/load is emphasized. Smith [15] uses AIS data to investigate the determinants of dwell time at major U.S. container ports, emphasizing that anticipated cargo volume and schedule adherence are major factors. In [16], classification models are developed to predict ship stay time in Brazilian ports, with random forest achieving accuracy and F1 scores above 73% and revealing location and cargo as the main factors. The authors in [17] conduct a spatial-network investigation of the determinants behind VTT at 2,300 ports worldwide, finding that efficiency increases with connectivity and GDP per capita, and declines with delays and larger vessel size. In [18] , port revenue is linked to cargo throughput and VTT, demonstrating their combined significance at Lagos Port Complex.

More broadly, efficient ports are consistently characterized by faster VTT, shaped jointly by port facilities and vessel characteristics. Reference [19] analyzes data from New Mangalore Port to identify operational factors affecting turnaround time, emphasizing the need to optimize both operations and infrastructure. In [20], a tree-based machine learning system is developed for predicting VTT at the Port of Bordeaux using port call records and generic vessel features, achieving error rates below 10%. Kokila et al. [21] examine VTT at Cochin Port, highlighting inefficiencies such as pre-berthing detention and resource-related delays, while Moon et al. [22] emphasize that reducing port time improves efficiency by lowering costs, CO_2 emissions, and other externalities.

The stacking regression method in [6] combines tree-based learners to predict vessel service time, achieving a reduction in MAE of 29.7% and in RMSE of 31.9% relative to existing Estimated Departure Time (EDT) benchmarks. Tanker terminals have received relatively less attention; in [23], a dynamic approach considering cargo-level operations, supported by AIS data, is presented to predict berth duration at the Port of Singapore, achieving prediction accuracy of up to 98.81%. More recently, Park et al. [24] proposed a VTT forecasting framework based on operation indicators derived from multi-class queuing at Busan Port, yielding up to 28% lower MAPE than existing approaches.

The body of work reviewed above evidently depicts a methodological shift—from classical statistical methods and early regression models toward tree-based and ensemble machine learning approaches—alongside growing evidence of the impact of vessel characteristics, cargo, handling resources, and planning and scheduling practices on VTT [25]. Despite this progress, the literature examined above exhibits several notable limitations. First, many studies rely on datasets that are insufficiently detailed and weakly operational, often covering restricted time periods with a limited number of explanatory variables.

Second, the use of ML models remains typically confined to a small number of competing models, with a significant portion of the literature relying on classification approaches that represent VTT as discrete categories rather than as the continuous variable it inherently is. Third, environmental factors—particularly weather-related delays—remain largely unexplored, even though they may plausibly affect vessel operations. Fourth, the focus tends to fall more on identifying the factors influencing VTT than on developing a complete predictive framework. Finally, no prior study has addressed VTT at Tanger Med specifically, with the exception of one study that touches on the topic only as a secondary aspect.

3. TANGER MED PORT COMPLEX

Positioned on the Strait of Gibraltar, a heavily trafficked sea route, Tanger Med Port Complex connects with 180 ports across more than 70 destinations around the world, offering direct access to the main shipping routes with minimal route deviation. This maritime traffic flow positions the port as a worldwide transshipment hub, ranking it as the leading container port in the Mediterranean / African regions. Operational since 2007, it has leveraged its strategic location, state-of-the-art port infrastructure and fully digitized management system to grow into a world-class center for container transshipment. Ranked 19th in the world [26] and 6th in the 2025 CPPI [27], Tanger Med saw its throughput reach 10.24 million TEUs [28]. These figures reflect its excellence and continuous technological advancement due to digitalization.

The port complex houses a container terminal, Rail freight terminal, Hydrocarbon terminal, General cargo terminal, vehicle terminal, Passenger and Ro-Ro terminal and a Logistics Zone. Upgraded infrastructure, state-of-the-art handling equipment and optimized maritime operations reduce the vessel waiting and maneuvering time and strengthen the port's competitiveness in world shipping trade.

That said, non-productive time which occurs during a ship's call —the idle periods that arise while cargo is being processed— poses a huge challenge for shipping companies, which incur higher terminal costs when such delays occur, in addition to disturbing the natural downstream flow of the supply chain. This underscores the need for accurate turnaround time prediction as a means of minimizing avoidable delays. Thus, ports must increasingly turn to advanced data-driven approaches to reduce idle time and optimize overall terminal operations.

Our study focuses specifically on the second container terminal (TC2) of the port. Based on historical port-call data stretching over 12 years, the aim is to create approaches to improve the prediction accuracy of VTT, thereby helping to improve the operational efficiency of the terminal. After all, efficient ports are those that can turn ships around more quickly even under high cargo volumes [3], and improving VTT prediction can bring actual improvements to the operational competitiveness of Tanger Med port as perceived by its customers, as well as help the port better position itself in the ongoing competition among ports.

4. DATASET AND PREPROCESSING

This section describes the dataset and its preparation: the research flow, the operational records extracted from the terminal's operating system, the target variable, and the preprocessing pipeline applied to these records.

4.1. Research Flow

The flowchart in Figure 1 illustrates the process followed to develop the VTT prediction ML models, starting with input data from the TOS of TC2. Preprocessing techniques (data integration, cleaning, transformation, and encoding) were then executed to ensure high data quality. Next, feature extraction and selection were carried out to identify suitable features for the ML model's VTT prediction.

A range of eight distinct types of ML models was considered for our regression models -covering tree based, ensemble, kernel-based, non-parametric, neural network, and meta-ensemble approaches. By tuning different hyperparameters through grid search, each model was optimized to achieve the best possible results. The best-performing model was then evaluated in detail against the actual dataset drawn from terminal operations.

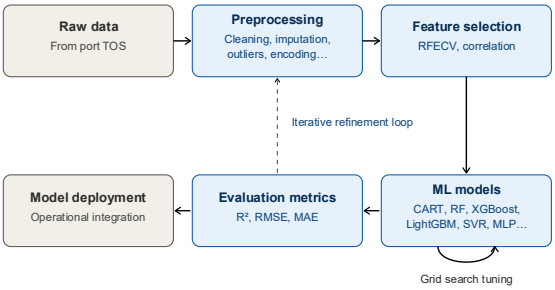


Figure 1. Research Flow For VTT Prediction.

4.2. Input Data

The dataset used in this study is both sufficiently detailed and directly linked to the terminal's operational processes, capturing the full range of aspects that characterise container terminal activity. Obtaining such detailed information from port authorities is, however, notoriously difficult: terminal operators are often reluctant to share operational data, citing concerns over confidentiality and operational security [29]. Against this backdrop, the dataset mobilised here stands out from those typically found in the literature, spanning an extended time period and encompassing a sufficiently rich set of variables to comprehensively capture terminal activity.

The data used in this study come from the TOS of TC2, obtained with the authorisation of the terminal management and under a confidentiality agreement. They consist of operational records of vessel calls, providing a comprehensive view of vessel movements at the terminal over a 12-year period. The dataset comprises 13,796 rows and 114 attributes, organised into five categories of features: vessel and voyage information, container and cargo operations, terminal operations and scheduling, user and system information, and miscellaneous attributes. Crucial attributes include loaded and unloaded cargo volumes, actual arrival and departure times (ATA–ATD), and estimated time of arrival, with each row representing a single call of a vessel to the terminal.

Weather data were also integrated into the dataset to examine its potential influence on analysis and prediction. Weather may plausibly affect vessel operations at the port, such as strong winds which can sometimes cause cranes to cease working for an extended period of time and rain during the

loading/unloading of dry cargo. Accordingly, there was reason to test whether including weather data could bring added explanatory power to the models used below. This hypothesis is examined in the experiment section. The first step was a comprehensive preprocessing phase, aimed at refining and structuring the dataset, checking its quality, and ensuring that the training data supplied to the models were reliable, thereby laying the groundwork for a robust process to predict VTT.

4.3. Data Preprocessing

Data preprocessing is a critical step and the most fundamental one of a data-driven or ML application, since the quality of the initial data used directly determines the precision and reliability of the resulting predictions. It encompasses a set of techniques used to improve the quality of raw data and adapt it into data of a more manageable and analytically useful format.

The preprocessing workflow was adapted and applied step by step, as described previously in [4]. The steps followed in the present analysis are summarized below:

- Data Integration: The data had been merged and synchronised; Three redundant attributes were removed and the duplicate tuples had already been handled.
- Data Cleaning: All irrelevant features, like free-text fields, mostly null or almost empty columns, features that have near-zero variance, and categorical features with high cardinality, were removed, so that only usable features remained. Figure 2 shows the number of non-null values for the columns containing null values.
- Missing Data Handling: Features with greater than 90% missing values were discarded. Remaining missing values were imputed according to feature type, for categorical columns, using new-category imputation and most frequent category imputation; and through K-NN Imputer, for numerical columns.
- Outlier Detection: IQR, Z-score, One-Class SVM (OCSVM), and DBSCAN methods have been evaluated. DBSCAN was selected as the most reliable and effective method for identifying and removing outliers, owing to its efficiency, and suitability for large-scale datasets. This method removed 1,329 outliers, reducing the dataset from 13,796 to 12,467 vessel calls.
- Data Normalization: Numerical features were rescaled through z-score normalization, yielding features with zero mean and standard deviation of one, thereby giving features a unified metric scale, but preserving all other distributional properties of our data.
- Categorical Data Encoding: One-hot encoding was applied to categorical variables, using N-1 binary indicators for each feature. Sparse representations were used to reduce memory overhead.

After preprocessing, the dataset consisted of 40 features and 12,467 instances. With the data prepared and cleaned, the next step involves exploring which features are the most relevant through exploratory data analysis, in order to train the ML comparative models using these features.

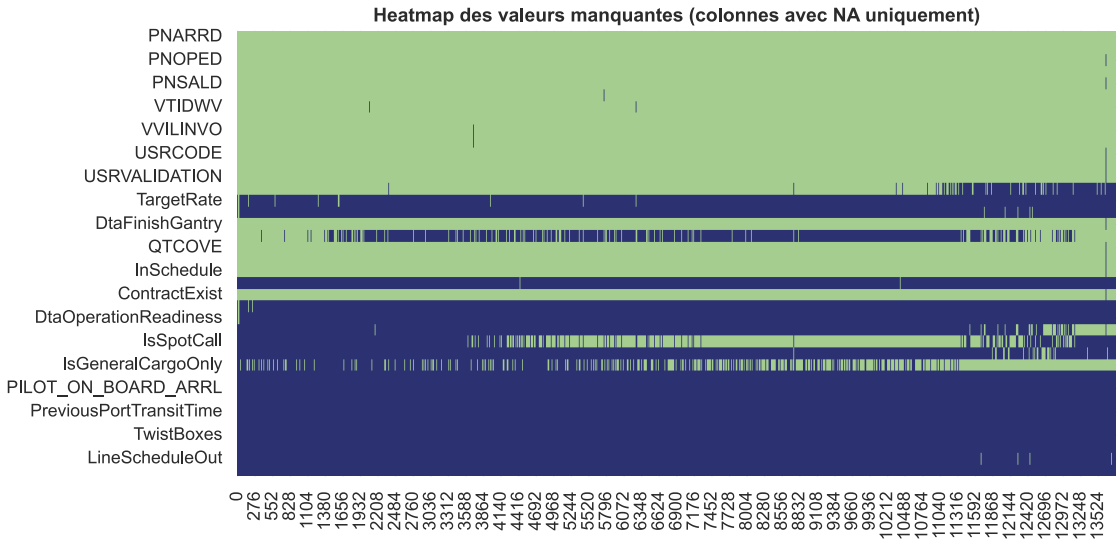


Figure 2. Heatmap Of Missing Values In The Dataset.

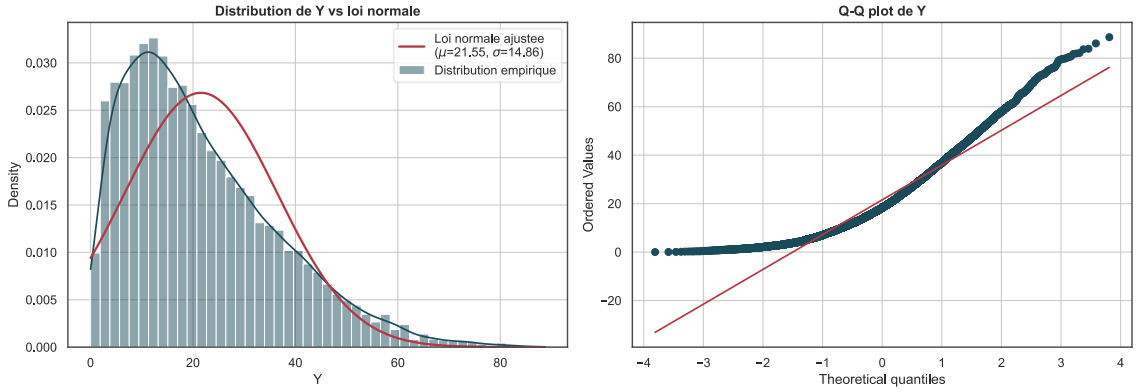


Figure 3. Distribution Of VTT Values

5. MACHINE LEARNING MODELS FOR REGRESSION TASKS

5.1. Feature Engineering

The process of feature engineering involves creating, selecting, or transforming input variables in order to improve the predictive outcomes of a model. Categorical feature encoding was already carried out in the preceding stage, while other features are assessed according to the magnitude of their impact on VTT. To identify the key aspects of the prediction issue, VTT —the ground truth of the regression task —was calculated as the difference between Actual Departure Time (ADT) and Actual Arrival Time (ATA). Descriptive statistics for this target variable are presented in Table 1.

Table 1. Descriptive Statistics For The Target Variable

Count	Mean	Median	Maximum	Minimum	Std. Deviation
12,467	21.55	18.29	88.67	0.17	14.86

The distributions of the VTT values are shown in Figure 3 . Most of the turnaround times lay in the range of 5 to 52 hours. Values that fall outside this range occur less common. The distribution is close to Gaussian with little positive skewness with a high density around 18 hours when observed closely. It is also consistent to the statistical information provided in Table 1 above.

There are several different methods to find features that are relevant to affecting VTT and those that are not. In this study, relevant features were selected through a correlation-based screening step (along with the previously discussed simple low/zero-variance filter that we include in the preprocessing) as well as RFECV (Recursive Feature Elimination with Cross-Validation).

A correlation matrix heatmap of all continuous features was plotted in Figure 5. to show the relationships between these features. The correlation coefficient ranges between -1 and 1, indicating how two variables are linearly related; in the case of a perfect positive relationship, the correlation is 1, and it is 0 when there is no relationship; -1 indicates a perfect negative relationship. Features with a correlation above approximately 0.90 (indicating that they are strongly linearly correlated and convey redundant information) were examined for potential information leakage and removed, along with any feature showing an implausibly strong correlation with the target variable itself, which would suggest data leakage rather than a genuine predictor.

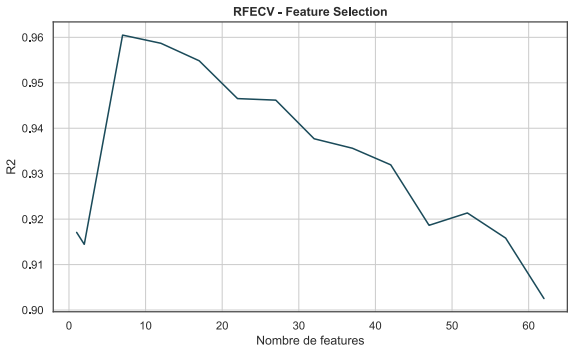


Figure 4. Number Of Features Selected

RFECV was applied to the original preprocessed dataset, using 5-fold CV and an Extra Trees Regressor (ETR) as the base estimator. The relationship between the number of features and model performance is illustrated in Figure 4; in this case the 7 most important predictors performed the best and adding further features did not result in any additional gain. The relative importance values of these features ranged between 0.08 and 0.25 (Figure 6) with VPNMOV, QTPARK, VPNCMO, VPGCOT, QTTMNL, QTCOVE and NbCrane being the most important ones.

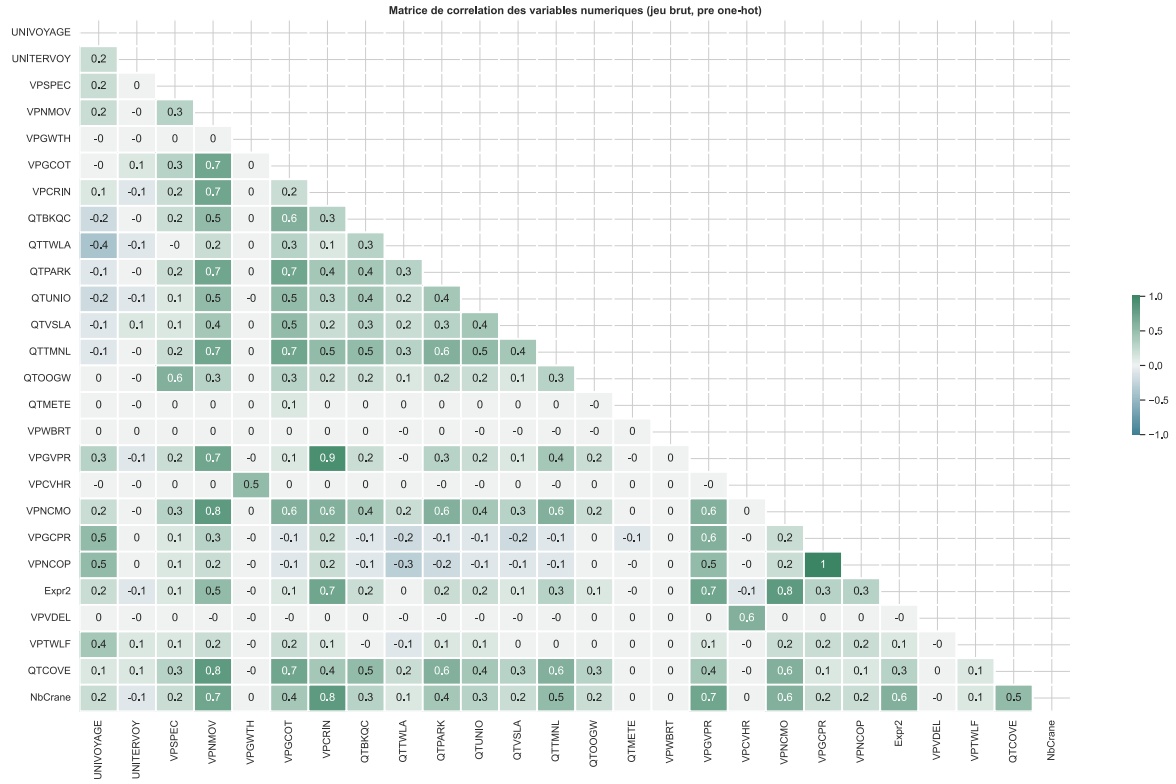


Figure 5. Feature Correlation Matrix.

5.2. Predictive Models

Machine learning for regression tasks typically comprises three main stages: data preprocessing, algorithmic modeling, and model refinement [25]. After the preprocessing and feature engineering discussed above, a clean set of relevant features for predicting VTT at Tanger Med Port was successfully established. The next step consisted of choosing suitable ML algorithms to construct regression models and evaluate their comparative performance, with hyperparameter optimization to fine-tune each model and achieve optimal results.

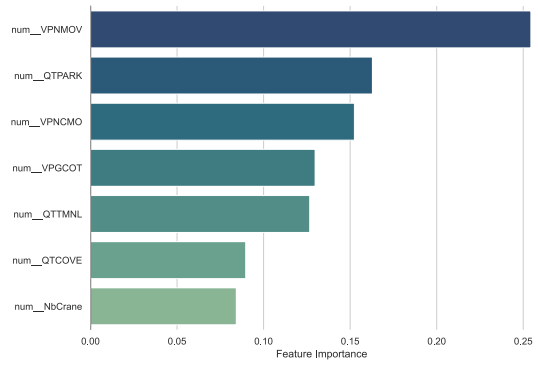


Figure 6. Importance Of The Selected Features

Models built on decision trees, such as XGBoost and Random Forest (RF), are widely used in predictive analytics [30] and can outperform neural networks when working with tabular datasets specifically [31]. However, in order to conduct a thorough comparative analysis and avoid limiting the study to a single family of algorithms, this study applies a broader range of regression techniques spanning tree-based, ensemble, kernel-based, non-parametric, neural network, and meta-ensemble approaches. Below are the models used along with a brief explanation of each (Table 2).

All models were implemented in Python 3.10.12 (scikit-learn 1.3.2, XGBoost 2.0.3, LightGBM 4.1.0), and experiments were run on a single workstation (Windows 11 Pro, Intel Core i7-12700H, 32 GB RAM). The random seed was fixed at 42 for every stochastic component and applied identically across all models, so that the comparison rests on identical data partitions. The data were split 80/20 using that same seed. The training set was used for hyperparameter tuning by grid search with 5-fold cross-validation, and the test set was used once for the final evaluation. The Stacking model combines RF, XGBoost, LightGBM and SVR as base learners with a Ridge meta-model trained on out-of-fold predictions.

5.3 Hyperparameter Optimization

Using a combined technique of Grid Search and K-fold CV, the models were enhanced for more accurate future outcome prediction and data generalizability. The Grid Search technique entails defining a range of hyperparameters to systematically investigate across, until the optimal combination for each model emerges. Complementary to this is K-Fold Cross-Validation, which partitions the dataset into K subsets. One of these subsets serves as the validation set, while the remaining K-1 are used for training. The results are then averaged across all folds to ensure reliability [32]. This procedure prevents models from overfitting to specific parts of the data, providing a robust evaluation of model performance.

In this study, 5-Fold CV was used during model training. As each algorithm involves different hyperparameters, specific ranges and intervals were defined individually for tuning.

5.4 Model Assessment Metrics

Three standard metrics were used to evaluate regression performance in predicting VTT: R^2 (the coefficient of determination), and MAE (Mean Absolute Error), RMSE (Root Mean Squared Error). A better model has a greater R^2 value (more close to 1), whereas smaller RMSE and MAE values indicate more accurate estimation. Of these metrics, R^2 is considered the most important for evaluating the overall quality of a prediction model.

6. RESULTS AND DISCUSSION

Before presenting the comparison results, it is useful to review the operational meaning of the seven predictors selected in the RFECV process. Namely, there are three functional groups of variables: container volume and planned cargo movements (VPNMOV, VPNCMO, and QTCOVE); crane productivity and operation time (NbCrane and VPGCOT); and waiting time and delays attributable to the terminal (QTPARK and QTTMNL). Overall, all the selected predictors correspond to the operational workload and delay dynamics created by each vessel call, and not the static scheduling metadata, which plausibly explains the predictive power of such predictors for VTT

A weather-related predictor was also considered among the candidate variables, but its inclusion did not improve predictive accuracy. This is consistent with the results of Stepec et al., who experimented with weather data, in particular wind and precipitation, combined with historical port call records at the Port of Bordeaux, and found that presenting these features to the predictive model did not result in an increased accuracy [20]. In the context of Tanger Med, this may be explained by the predominance of cargo-handling workload and terminal scheduling over weather-related disturbances in accounting for VTT. The contribution of weather to explaining VTT in this dataset may therefore be smaller than that of other predictors, and may vary across cargo and operational contexts.

Table 2. Overview Of The Machine Learning Models Used For VTT Prediction.

Model	Family	Key Principle
CART	Decision tree	Divides the input space step by step into uniform regions, splitting on the variable that most reduces prediction error at each node.
Random Forest (RF)	Ensemble (bagging)	Aggregates predictions from a large number of decision trees, each grown from a bootstrap sample and a randomly restricted set of features, to reduce variance.
XGBoost	Ensemble (boosting)	Adds trees in sequence, each one trained to patch the errors accumulated by earlier trees, using regularization to limit overfitting.
LightGBM	Ensemble (boosting)	Refines gradient boosting through histogram-based splitting and leaf-wise tree growth, improving training speed on large datasets.
SVR (Support Vector Regression)	Kernel-based	Fits a function that stays within a set margin of the observed targets, relying on kernel functions to capture non-linear patterns.
k-Nearest Neighbors (k-NN)	Non-parametric	Predicts a target value as the average outcome of the k most similar observations in the feature space, without assuming a fixed functional form.
Backpropagation Neural Network (BPNN)	Neural network	A feedforward network with one or more hidden layers, trained via backpropagation to iteratively adjust weights and capture non-linear feature interactions.
Stacking Regression	Meta-ensemble	Pools predictions across multiple diverse base learners, feeding them into a second layer that figures out the optimal way to blend them

Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

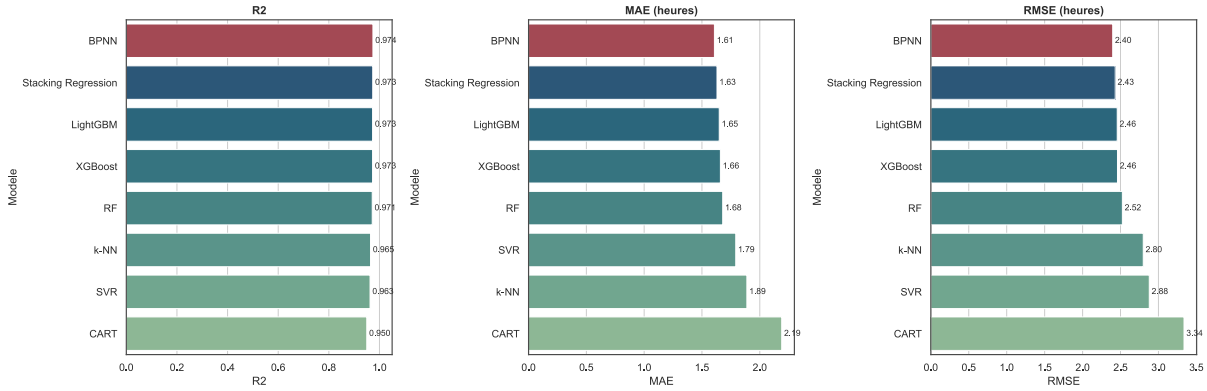


Figure 7. Comparative Model Performance For VTT Prediction.

Table 3 contains a comparative analysis of the eight regression models estimated on the hold-out test set, using seven selected features. The models are ranked based on the R^2 score, accompanied by RMSE and MAE. Figure 7 illustrates these results graphically*Figure 1*.

Table 3. Comparative Performance Of The Eight Regression Models For VTT Prediction.

Model	R^2	RMSE (h)	MAE (h)	Training time
BPNN	0.9741	2.396	1.608	2.84
Stacking Regression	0.9734	2.430	1.630	14.33
LightGBM	0.9728	2.456	1.649	0.29
XGBoost	0.9728	2.460	1.659	0.51
Random Forest	0.9713	2.524	1.679	0.77
k-NN	0.9647	2.799	1.888	0.03
SVR	0.9627	2.878	1.792	2.09
CART	0.9499	3.336	2.190	0.04

All eight models achieve an R^2 greater than 0.94, which means that the selected feature set through the preprocessing and RFECV has enough predictive power for the VTT. In addition, this result was obtained with the use of only seven predictors: VPNMOV, VPGCOT, QTPARK, QTTMNL, VPNCMO, QTCOVE, and NbCrane, out of 114 raw attributes included in the TOS extract. This represents a significant improvement over the prediction of the VTT made by the terminal itself, and is in line with the improvements reported in the literature on VTT prediction using ML algorithms[5]. The fact that a parsimonious seven-feature model outperforms the terminal's own estimation process underscores the value of a systematic, data-driven feature selection procedure over the heuristic or experience-based reasoning typically underlying terminal-reported estimates. Beyond the gain in predictive accuracy, this result is also practically meaningful in terms of prediction speed, since the reduced feature set makes the model considerably cheaper to maintain and deploy operationally.

The BPNN achieves the best overall performance, marginally outperforming Stacking Regression and the two gradient-boosting models, LightGBM and XGBoost. Its superior performance underscores its efficacy in capturing the complex, non-linear relationships inherent in terminal operations, justifying the additional computational cost relative to faster alternatives like LightGBM. This finding contrasts with the general observation that tree-based models often outperform neural networks on tabular data [31], although the differences here are small: 0.0007 in R^2 between the BPNN and Stacking Regression, and 0.0013 between the BPNN and LightGBM. However, the comparison should be interpreted cautiously unless the studies' data, target definitions, and evaluation procedures are sufficiently aligned. Thus, the contribution lies not in claiming that neural networks are generally superior for this task, but in demonstrating the predictive value of a carefully preprocessed, feature-selected, and data-driven feature set.

The stacking Regression ranks second (R^2 = 0.9734) but takes 5-28 times longer to train than the top models, with a negligible accuracy gain which makes the model rather impractical for operation. LightGBM and XGBoost take the third and fourth places respectively (R^2 = 0.9728 each) and train in under a second, making them the most preferable models when accuracy should be balanced with speed and deployability.

The Random Forest shows a rather good result (R^2 = 0.9713) although it uses considerably simpler ensembling approach than the top models, yet still trains significantly slower than LightGBM or

XGBoost without any accuracy gain. On the other hand, the k-NN and SVR rank further down the list, probably due to the weaker capability to capture the underlying structure of the dataset for the seven numeric predictors. Finally, CART produces the weakest result (R^2 = 0.9499), which proves the advantages of the ensemble and boosting approaches over a single decision tree.

Table 4. Optimal Hyperparameters Adopted For BPNN.

Hyperparameter	Search space	Optimal value
Hidden layer sizes	(64, 32); (128, 64); (128, 128); (256, 128)	(128, 64)
Alpha (L2 regularization)	0.00001; 0.0001; 0.001; 0.01	0.0001
Learning rate (initial)	0.0001; 0.001; 0.01; 0.1	0.01

The grid search was performed for all eight models with their particular sets of candidate hyperparameters. As the best performing model turned out to be BPNN, the configuration of this particular model alone is described here, to avoid unnecessary repetition. Using the 5-Fold CV, the best configuration that has been selected, as presented in Table 4, is a two-hidden-layer architecture, with light regularization strength and a moderate initial learning rate. This relatively deep architecture, paired with light regularization and moderate learning rate, shows that the interaction between the selected predictors and VTT benefits from a network with sufficient capacity to model non-linear interactions, whereas the low alpha value implies a low overfitting risk.

Three limitations bound these results. The first is scope: the models were built on calls at a single terminal, TC2, so the accuracy reported here has yet to be tested on another terminal. Models developed on one container terminal are specific to it [9] : ports are not homogeneous in their layout, operational details and available data, so models must be calibrated to each port [25] . The second concerns the information used. The seven retained variables exclude the season and congestion indicators. Weather was tested through a single delay variable reported by the TOS and brought no gain, although its effect may depend on equipment and cargo type [33]. AIS records, which describe the approach trajectory, were not used. The third is data access. Terminal operators own their operational records, which are by nature not shared and confidential [23]. Many ports withhold detailed operational data to protect their competitive position, and obtaining such data remains the main hurdle in applying machine learning to port studies [25]. The pipeline therefore cannot be re-run by third parties, and confidentiality prevents external replication.

Future research should focus on three directions. First, applying the pipeline to other container terminals at Tanger Med and to ports elsewhere would show whether the seven-predictor structure is transferable or whether terminal-specific variable sets are needed. Second, incorporating seasonal and congestion indicators—such as the number of vessels in port at the time of arrival [5] —could improve predictions, provided these data are available when the prediction is made. Third, integrating the predictions into berth-allocation and yard-planning models, and extending the approach across terminals, could make call-duration estimates more comparable.

7. CONCLUSION

This study set out to determine whether vessel turnaround time at a container terminal can be predicted accurately from the operational records the terminal already produces, and which of eight machine learning models does so best. Using twelve years of terminal operating system data from TC2 at Tanger Med, preprocessing and RFECV-based feature selection reduced 114 raw attributes to seven predictors, and the BPNN reached R^2 = 0.9741, RMSE = 2.396 h and MAE = 1.608 h on the hold-out test set. Because the differences between the models are small, the more consequential result is that a seven-variable model matches the accuracy of considerably more expensive alternatives.

The broader significance of this finding lies in what it displaces: inefficient port operations generate costs for shipping companies and unproductive days for vessels, whereas shorter and more predictable turnaround times reduce vessel idling, costs and emissions and strengthen port competitiveness [22]. The ability to anticipate the duration of a call before it completes changes the information available to berth and crane planning, and allows arrival and departure estimates to be updated as a call progresses. At the level of a single terminal, that means better allocation of quay cranes and berths; beyond it, it contributes to the competitiveness of the port complex and to the reliability of the schedules that carriers publish to shippers. For a hub handling over 11 million TEU a year [34] , even modest gains in planning accuracy accumulate across thousands of calls. Because the study draws on a twelve-year span covering periods of growth in vessel size and call

volumes, the framework also offers a template for terminals that hold comparable historical records but lack the analytical capacity to exploit them.

Several applications extend beyond this study. The same pipeline could be applied to the other container terminals of Tanger Med and to other ports, testing whether the seven-predictor structure transfers or whether terminal-specific variable sets are required. Coupling the approach to berth allocation and yard planning models would let predicted call durations replace the fixed handling-time assumptions currently used, and applying it across terminals would put call-duration estimates on a comparable basis, enabling benchmarking.

Realising that potential depends on access to data. The obstacle is the confidentiality constraints under which terminal operating data are held. Progress in this field therefore requires two things: terminals and port authorities willing to share data with researchers under confidentiality terms that protect commercial interests, and researchers who report their studies in enough detail for the results to be tested. Extending the approach to additional terminals for multi-port validation, and piloting it inside terminal operating systems, would be the natural next step.

REFERENCES:

[1] UNCTAD, *Review of maritime transport 2025* . United Nation Publications, 2025.

[2] W. Sheikh, K. R. Hasan, H. Islam, and M. Billah, "'Does bigger mean slower?' Investigating the relationship between vessel characteristics and turnaround time," *Asian Transport Studies*, vol. 11, p. 100165, 2025, doi: 10.1016/j.eastsj.2025.100165.

[3] C. Comtois, B. Wiegmans, P. Witte, and B. Slack, "Ships time in port," *International Journal of Shipping and Transport Logistics*, vol. 10, no. 1, p. 45, 2018, doi: 10.1504/IJSTL.2018.10008534.

[4] M. Al Uahabi, H. Attariuas, M. Saleh, and M. Chentouf, "An Adaptive Data Preprocessing Framework for Improved Learning: A Case Study of Tangier Container Terminal," *Journal of Computer Science*, vol. 20, no. 3, 2024, doi: 10.3844/jcssp.2024.265.275.

[5] Z. Chu, R. Yan, and S. Wang, "Vessel turnaround time prediction: A machine learning approach," 2024, doi: 10.1016/j.ocecoaman.2024.107021.

[6] R. Yan, Z. Chu, L. Wu, and S. Wang, "Predicting vessel service time: A data-driven approach," *Advanced Engineering Informatics*, vol. 62, Oct. 2024, doi: 10.1016/j.aei.2024.102718.

[7] P. W. A. van Leeuwen and R. Dekker, "Vessel arrival planning with full information," *International Transactions in Operational Research*, vol. 34, no. 1, pp. 372–396, Jan. 2025, doi: 10.1111/itor.70108.

[8] R. Raeesi, N. Sahebjamnia, and S. A. Mansouri, "The synergistic effect of operational research and big data analytics in greening container terminal operations: A review and future directions," Nov. 01, 2023, *Elsevier B.V.* doi: 10.1016/j.ejor.2022.11.054.

[9] J. H. Yoon, S. W. Kim, J. S. Jo, and J. M. Park, "A Comparative Study of Machine Learning Models for Predicting Vessel Dwell Time Estimation at a Terminal in the Busan New Port," *J. Mar. Sci. Eng.*, vol. 11, no. 10, Oct. 2023, doi: 10.3390/jmse11101846.

[10] I. Kourounioti and A. Polydoropoulou, "Identification of container dwell time determinants using aggregate data," *International Journal of Transport Economics*, vol. 44, no. 4, pp. 567–588, Dec. 2017, doi: 10.19272/201706704004.

[11] K. Mokhtar and M. Z. Shah, "A regression model for vessel turnaround time," 2006.

[12] N. Ming, M. S.-E. I. S. on, and undefined 2008, "Petroleum terminal's operation processes on vessel turnaround time," *academia.eduNS Ming, MZ ShahEASTS International Symposium on Sustainable Transportation, 2008•academia.edu*, 2008.

[13] O. G. Karuppanan. Jayaprakash, "Measurement of Port Performance Utilising Service Time of Vessels.," *International Journal of Civil Engineering & Building Materials*, 2012.

[14] K. Loke, M. Othman, … A. S.-O. J. of M., and undefined 2014, "Analysis of variables of vessel calls in a container terminal," *scirp.orgKB Loke, MR Othman, AH Saharuddin, MN FadzilOpen Journal of Marine Science, 2014•scirp.org*, 2014.

[15] D. Smith, "Big Data Insights into Container Vessel Dwell Times," *https://doi.org/10.1177/03611981211015248*, vol. 2675, no. 10, pp. 1222–1235, Jun. 2021, doi: 10.1177/03611981211015248.

[16] L. R. Abreu, I. S. F. Maciel, J. S. Alves, L. C. Braga, and H. L. J. Pontes, "A decision tree model for the prediction of the stay time of ships in Brazilian ports," *Eng. Appl. Artif. Intell.*, vol. 117, p. 105634, Jan. 2023, doi: 10.1016/J.ENGAPPAI.2022.105634.

[17] C. Ducruet and H. Itoh, "Spatial Network Analysis of Container Port Operations: The Case of Ship Turnaround Times," *Netw. Spat. Econ.*, vol. 22, no. 4, pp. 883–902, Dec. 2022, doi: 10.1007/S11067-022-09570-Z/METRICS.

[18] S. O. Buhari, S. O. Buhari, O. Ndikom, and N. Theophilus, "An Assessment of the Relationship Among Cargo-Throughput, Vessel Turnaround Time and Port-Revenue in Nigeria. (A Study of Lagos Port Complex)," *Journal of Advance Research in Business, Management and Accounting (ISSN: 2456-3544)*, vol. 3, no. 7, pp. 01–09, Jul. 2017, doi: 10.53555/nnbma.v3i7.54.

[19] Dayananda Shetty Karnoji, Vijayanand Gurudev, and Dwarakish Gowdagere Siddaramaiah, "Factors affecting the Vessel Turnaround time in a Seaport," HYDRO-2020 International Conference, Mar. 2020.

[20] D. Stepec, T. Martincic, F. Klein, D. Vladusic, and J. P. Costa, "Machine Learning based System for Vessel Turnaround Time Prediction," *Proceedings - IEEE International Conference on Mobile Data Management*, vol. 2020-June, pp. 258–263, Jun. 2020, doi: 10.1109/MDM48529.2020.00060.

[21] Kokila A. V. and Abijath V., "Reduction of Turnaround Time for Vessels at Cochin Port Trust," 2017.

[22] D. S. H. Moon and J. K. Woo, "The impact of port operations on efficient ship operation from both economic and environmental perspectives," *Maritime Policy & Management*, vol. 41, no. 5, pp. 444–461, Jul. 2014, doi: 10.1080/03088839.2014.931607.

[23] D. Zhai, X. Fu, X. F. Yin, H. Xu, and W. Zhang, "Predicting Berth Stay for Tanker Terminals: A Systematic and Dynamic Approach," Apr. 2022.

[24] D. Park, T. Noh, Y. Koo, H. Lee, D. Kim, and H. Bae, "Time-series approach to vessel turnaround time forecasting using queuing-based operation indicators," *Advanced Engineering Informatics*, vol. 69, p. 103974, Jan. 2026, doi: 10.1016/J.AEI.2025.103974.

[25] S. Filom, A. M. Amiri, and S. Razavi, "Applications of machine learning methods in port operations – A systematic literature review," *Transp. Res. E Logist. Transp. Rev.*, vol. 161, p. 102722, May 2022, doi: 10.1016/J.TRE.2022.102722.

[26] Tanger Med Port Authority, "TANGER MED JOINS THE GLOBAL TOP 20 AND RANKS 19th OUT OF 500 CONTAINER PORTS WORLDWIDETANGER MED JOINS THE GLOBAL TOP 20 AND RANKS 19th OUT OF 500 CONTAINER PORTS WORLDWIDE," https://www.tangermedport.com/en/tanger-med-joins-the-global-top-20-and-ranks-19th-out-of-500-container-ports-worldwide/.

[27] *The Container Port Performance Index 2025: A Comparable Assessment of Performance based on Vessel Time in Port*. Washington, DC: World Bank, 2026. doi: 10.1596/45008.

[28] Container Mag, "Tanger Med Port surpasses 10m TEU in2024," https://container-mag.com/2025/01/24/tanger-med-port-surpasses-10m-teu-in-2024/.

[29] S. Nikghadam, K. F. Molkenboer, L. Tavasszy, and J. Rezaei, "Information sharing to mitigate delays in port: the case of the Port of Rotterdam," *Maritime Economics & Logistics*, vol. 25, no. 3, pp. 576–601, Sep. 2023, doi: 10.1057/s41278-021-00203-9.

[30] C. Lei *et al.*, "Can manual and automated trucks coexist efficiently? A data-driven evaluation of mixed operation mode in retrofitted container terminals," *Comput. Ind. Eng.*, vol. 210, p. 111539, Dec. 2025, doi: 10.1016/j.cie.2025.111539.

[31] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why Do Tree-Based Models Still Outperform Deep Learning on Typical Tabular Data?," in *Advances in Neural Information Processing Systems 35*, San Diego, California, USA: Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2022, pp. 507–520. doi: 10.52202/068431-0037.

[32] Z.-H. Zhou, *Machine Learning*. Singapore: Springer Singapore, 2021. doi: 10.1007/978-981-15-1967-3.

[33] T. Brian. Mapotsi, "Factors affecting vessel turnaround time at the port of Richards Bay dry bulk terminal.," 2019. Accessed: Feb. 03, 2024. [Online]. Available: https://researchspace.ukzn.ac.za/handle/10413/20584

[34] *RAPPORT ANNUEL 2025*. Tanger Med Special Agency , 2025.