DELAY PREDICTION IN CONTAINER TERMINALS: A COMPARISON OF MACHINE LEARNING METHODS

Claudia Pani, University of Cagliari/University of Antwerp, claudia.pani@unica.it
Massimo Cannas, University of Cagliari, massimo.cannas@unica.it
Paolo Fadda, University of Cagliari, fadda@unica.it
Gianfranco Fancello, University of Cagliari, fancello@unica.it
Luca Frigau, University of Cagliari, frigau@unica.it
Francesco Mola, University of Cagliari, mola@unica.it

ABSTRACT

One of the most important issues in Container Terminal (CT) management is to manage adequately late arrivals. In fact, despite contractual obligations to notify the Estimated Time of Arrival (ETA) 24 hours before arrival, often ship operators have to revise it due to unexpected events such as weather conditions, delay in previous port and so on.
In a daily planning scenario, this causes a series of inconveniences directly associated with the resource allocation problem. Terminal operators need to increase the accuracy of incoming demand in order to allocate more efficiently human resources, equipment and spatial resources required to satisfy the predicted demand. For planners the decision-making processes related with demand uncertainty may sometimes be highly complex without the support of suitable methodological tools.
Specific models should be adopted, in a daily planning scenario, to provide a useful support tool in CTs and to help mitigating the consequences of late arrivals. In this study, using data collected in a Mediterranean Transhipment Container Terminal, we illustrate a data mining approach for predicting the level of daily alarm related to late arrivals. First, we defined three levels of daily alarm ranking the delay of arrivals. Then we obtained an estimate of the alarm level using three different Machine Learning models (Naive Bayes, Decision Trees and Random Forests) and we compared their predictive power on a test data set.

Keywords: delay forecasting, container terminal, machine learning

INTRODUCTION

A container terminal is a complex infrastructure where the most advanced techniques and technologies are adopted in order to maximize efficiency, intended as the maximum productivity at the least cost of the terminal's core-business. A variety of complex, inter-related decisions are taken each day to achieve this aim. In practice the terminal strives to minimize vessel waiting time, vessel berthing time and the resources required to complete handling operations. (Murty et al., 2005, Salido et al., 2011).

Essentially, container terminal activities include seaside operations, internal logistics and landside operations. Because the different operations carried out in terminals are strongly interrelated, there is a need not only to maximize efficiency of each operation, but also to ensure proper coordination, hence to solve integrated decision-making problems (Stahlbock and Voß, 2007). Vis and De Koster (2003), Stahlbock and Voß (2008) provide an interesting overview of the classification of decision problems in a container terminal on the basis of the five main logistic processes: arrival of the ship, loading/unloading of the ship, transport of containers from quayside to yard and viceversa, stack of containers in the yard and transport of containers outside the terminal with other modalities.

This research work focuses on the vessel arrival process. Shipping lines have the contractual obligation of notifying their ETA (Estimated Time of Arrival) at predetermined time intervals. However, even the last ETA, 24 hours prior to expected arrival, often has to be updated, due to unexpected events such as adverse sea/weather conditions, delays at previous port, etc. Thus, on a daily level the uncertainty surrounding the time of arrival of a vessel in port persists. This causes a series of inconveniences directly associated with:

1. allocation of berthing space. Berths are allocated so as to reduce vessel loading/unloading times and the distance from the origin/destination container yards (Zhen et al., 2011). In the event of ship delay, the berthing space has to be reallocated. As the containers to be loaded onto the vessel have already been moved to the stacking yard, a remarshaling plan is necessary to minimize berthing time (Salido et al., 2011). Vessels delay could, for instance, make additional handling operations within the terminal necessary, resulting in higher costs;

2. allocation of the resources required for handling operations. It is necessary to allocate the resources (human and mechanical) required to satisfy the actual demand, which are often over/under estimated at the planning stage. The handling equipment also needs to be available and in perfect working order so as to avoid a shortage of basic resources when required to cope with increased workloads. (Gambardella et al., 1996, Fancello et al., 2011).

Clearly, in similar circumstances the ability of predicting vessel delay, and more generally the time of arrival in port, is the key for achieving efficient planning of resources (human, space and equipment), at least in the short term. However, very few works in the scientific literature deal explicitly with this problem but mostly focus on container flows prediction in and out of container terminals (seaside and landside) over a daily time horizon (Murty et al., 2005; Fung, 2002; Sideris et al., 2002; Gambardella et al., 1996).

The papers that most closely deal with the topic addressed here are Fancello et al. (2011), Fadda et al. (2012) and Lu et al. (2008).

Fancello et al. (2011) and Fadda et al. (2013) attempted to predict delay of ship arrivals in a Mediterranean port, and they found that neural networks and regression trees (CART)

perform best. In particular with the CART model the probability of univocally determining the work-shift of vessels arrival is around 75%. Another possible route to punctual predictions is forecasting the delay levels over a temporal interval, which requires the prior definition of delay levels. Lu et al. (2008) use this approach for airport data. They built a model that attempt to predict the level of daily alarm due to aircraft delay, in one of the most important hubs in China.
In this work, following Lu's approach, we defined via cluster analysis three levels of daily alarm related to the delay of arrivals at the studied port, using data recorded in a Mediterranean TCT in 2010 and 2011. Then we compared three techniques (Naive Bayes, Decision Trees and Random Forest) to obtain an estimate of the alarm level for each day from January 2012 to June 2012.
The remainder of the paper is organized as follows. In the second section the data collected from the port authority are briefly described. In the third section, the variables considered as most appropriate to measure the phenomenon of delay are selected and used to build an index ranking the delay at the daily level. In the fourth section some machine learning methods for predicting the delay level are compared before discussing the results. The fifth section introduces briefly the practical implications of the models. Ultimately, in the sixth section some conclusions are drawn.

DATA

We collected data on vessel arrivals from January 2010 to June 2012 in a TCT in the Mediterranean[1]. The database includes all mother and feeder arrivals at the container terminal over a period of 30 months.
Considering more than 2000 arrivals collected, the available variables can be divided into 4 main classes:

1. Vessel features (length, draft, gross tonnage, capacity, vector type). These variables are indicators of the vessel physical structure and directly affect the duration of handling operations and sailing times. The length also provides important information concerning berth occupancy. In particular the variable "vector type" classifies vessels as mother or feeder.

2. Vessel service (shipping lines, ports rotation, sailing direction, previous port). The shipping lines provide the service on several routes. Sometimes the routes have one direction (standard) or several directions (eastbound/westbound - Tyrrhenian bound/Levant bound). So, vessel service depends on ports' rotation and sailing direction and provides important information about the organization/availability of previous port.

3. Vessel containers (number and type of containers to be processed). These variables provide an indication of the number and type of containers to be loaded

[1] We are grateful to the CICT (Cagliari International Container Terminal) for their cooperation and valuable research support

onto/unloaded from each vessel arriving in port. These variables are strongly related with the loading/unloading operation times.

4. Vessel position (Estimated Time of Arrival, Actual time of Arrival, Berthing/Unberthing Time, Start/End operations time). These variables give an indication about the position of the vessel from the time it arrives at the pilot point to unberthing time.

RANKING THE DELAY AT THE DAILY LEVEL

In this section we use the variables described in section 2 to obtain a synthetic index ranking the delay over a fixed time interval. In order to do this we first select the variables measuring delay severity and then we aggregate the arrivals over the chosen time interval. In maritime transportation there is no standard index for ranking delay severity.
On the basis of a period of observation at the studied terminal, and from a preliminary exploratory data analysis, we set the time interval equal to one day and then we selected, for each day of activity the following variables:

1. proportion of delayed vessels;
2. total delay hours, calculated as the difference between the Estimated Time of Arrival and the Actual Time of Arrival;
3. total length of the delayed vessels;
4. total number of mother vessels delayed;
5. total number of delayed containers (import, export and restows).

The chosen variables can be considered the “dimensions” over which the severity of the late arrivals can be measured, from the point of view of port operators. In fact they are the main indices related to delay complexity management in ports.
The first variable represents the proportion of delayed ships with respect to the total number of vessels arriving on a given day. Ships arriving up to 15 minutes before and no later than 15 minutes after the notified ETA are considered to be on time. The threshold was set at 15 minutes for operational reasons, as a delay of a quarter of an hour does not cause any disruptions. The second variable considers the total time of delayed arrivals on a given day. It is expressed in hours and is calculated as the difference between the Estimated Time of Arrival and the Actual Time of Arrival. The third variable indicates the total length of vessels not arriving on time and provides important information concerning vessel dimensions and berth occupancy. The fourth specifies the total number of mother ships arriving late on a given day and considers the objective physical/structural differences between mother and feeder vessels. It also gives an indication of the different dynamics characterizing the services provided by the container terminal to cope with the delay: generally mother vessels have priority over feeder vessels. Instead the last variable indicates the total number of

containers to be handled for each delayed arrival on a given day. It is strongly related with vessel berthing time.

At the end of this process a data set was obtained containing information for 645 days of operations. The index ranking the delay can be obtained by partitioning the records and ordering the final groups with respect to the chosen dimensions. In particular, we use cluster analysis to partition the records into groups maximizing some measures of internal homogeneity and external heterogeneity, so that the profiles of objects in the same cluster are very similar, whereas the profiles of objects in different clusters are quite distinct. Cluster analysis techniques can be classified as hierarchical or partitioning. In the hierarchical method the number of clusters is not fixed a priori. Instead, a series of partitions takes place, which may run from *n* clusters containing a single object to a single cluster containing all objects. The process proceeds sequentially, starting from single original records and yielding a nested arrangement of records in groups. A commonly used approach in hierarchical clustering is Ward's method. In Ward's method (Ward, 1963), at each stage the algorithm merges the two clusters that result in the least increase in "information loss", usually measured by the within-clusters variance. Other merging criteria are possible, for example in the complete linkage method, at each stage the algorithm merges the two clusters having the least distance between the two most distant observations, one for each cluster. By contrast, partitional algorithms, such as the k-means algorithm (Hartigan and Wong, 1979), require the prior choice of the number of groups. The algorithm randomly chooses a set of initial centres, and then assigns each record to the group showing the least distance from its centre.

In order to choose the "best" partition among the output of a hierarchical cluster analysis (or among several outputs of a partitional algorithm) a number of graphical procedures and numerical indices have been developed. In particular several authors have proposed cluster validity indices, using different approaches. Following Theodoridis and Koutroubas (1999), these indices can be classified as external (based on previous knowledge about data), internal (based on the information intrinsic to the data alone) and relative indices (based on comparison of different clustering schema). Especially among internal validation indices it is possible to distinguish some specific indices (the stability measures) that work very well when data are highly correlated (Brock et al., 2008). The stability measures are based on comparison between clusters achieved using all variables of data and clusters achieved removing the variables, one at a time.

In our application, we used both k-means and Ward's method to cluster the daily records on the basis of delay dimensions. The two partition methods substantially overlap (only six records have different classification), so in the following we always refer to Ward's solution.

From the literature review conducted, we chose 6 internal indices, of which 3 are stability measures, to check which partition could be the optimal one.

The three stability measures used are: the Average Proportion of Non-overlap (APN), the Average Distance between Means (ADM) and the Figure Of Merit (FOM). The APN measures the average proportion of observations that are not included into same cluster, considering clusters achieved on all data and on data with a variable removed. The ADM measures the average distance between cluster centers (calculated as the mean of observations of the cluster), considering clusters achieved on all data and on data with a variable removed. The FOM measures the average of variance of the observations in the

removed variable. For all these three stability measures small values correspond better performances.
The other three internal indices used are: the Connectivity, the Dunn index and the Silhouette width. The Connectivity measures for each observation the number of own nearest neighbours not belonging to same cluster. The Dunn index measures the ratio of the smallest distance between observations not in the same cluster and the maximum distance between observations in cluster. The Silhouette width is the average of each observation of the Silhouette value. The Silhouette value measures the normalized difference between two the average distances: the first one is the average distance between a single observation and all other observations in the same cluster, and the second one is the average distance between an observation and the observations in the nearest neighboring cluster. In order to correspond better performances, Silhouette width and Dunn index should be maximized, instead Connectivity minimized.

Table I shows the values of cluster validation indices for Ward's method and highlights in bold the optimal scores.

Table I – Values of cluster validation indices for partitioning in 3, 4, 5 or 6 clusters for the Ward's method

Indices	3 clusters	4 clusters	5 clusters	6 clusters	Best partition
APN	**0.0524**	0.0864	0.1080	0.1099	3
ADM	**0.1897**	0.2531	0.3332	0.3969	3
FOM	0.5952	0.5608	0.5459	**0.5416**	6
Connectivity	**6.9091**	24.9250	37.0095	37.1095	3
Dunn	**0.1575**	0.0759	0.0670	0.0670	3
Silhouette	0.6755	0.6884	**0.6895**	0.6829	5

As four indices out of six suggest, we used a three cluster solution. Table II shows the mean values of each variable for the three cluster solutions using Ward's method.

Table II – Cluster means and standard deviations

Variables	Cluster 1	Cluster 2	Cluster 3
Total number of mother ships delayed	0 ± 0	0.49 ± 0.50	1.53 ± 0.56
Proportion of delayed vessels	0.02 ± 0.08	0.69 ± 0.27	0.80 ± 0.19
Total length of delayed vessels	9.29 ± 34.22	231.0 ± 84	533.0 ± 123
Total number of delayed containers	15.13 ± 70.5	572.0 ± 277.6	1421.3 ± 393.3
Total delay hours	0.03 ± 0.20	1.03 ± 0.98	3.04 ± 1.83

The clusters can be naturally sorted by increasing values of the five dimensions: we can label the three resulting groups as "low", "medium" and "high", respectively.
With respect to the delay level, the first class is characterized by vessels that arrive essentially on time. The second class comprises vessels with a less than average delay (around 1 hour). Vessels in the third class are characterized by the highest values of delay. In this class the delay ranges from 1 to 5 hours, with an average of around 3 hours.
Using these labels we can rank the delay of each day in the data set.

FORECASTING DELAY LEVEL USING MACHINE LEARNING ALGORITHMS

In this section we present some machine learning models for predicting the severity of late arrivals in a day, from the point of view of port operators. Predictors were selected heuristically, i.e. we tried different subsets of predictors and finally retained the subset yielding the best results in terms of predictive power. This subset includes, for each day:

1. total number of vessel arrivals;
2. quota of mother vessels scheduled for the day;
3. total vessel length;
4. total number of containers to be handled (import, export and restows).

Variables related to the vessel service did not provide useful information in predicting delay levels, and we excluded them from the final models. As for the machine learning algorithms involved we compared the performance of CART, Random Forest and Naive Bayes algorithms. We briefly describe these algorithms herein.
The acronym CART stands for Classification And Regression Trees, a technique developed by Breiman et al. (1984) and based on a binary recursive procedure. Starting from a single node containing all observations. At each stage the algorithm selects the variable and the split value of the variable that reduces to the greatest extent the variance (or more generally a measure of impurity) of the target variable in the new nodes. A node is terminal when it reaches a minimum number of observations or when the impurity measure can no longer be significantly improved. Predictions can be obtained by examining the distribution of the target variable in the terminal nodes. The tree is developed in two phases: tree growing and tree pruning. In the first phase a large tree is built by recursive binary splits of the explanatory variables. In the second phase the tree size is reduced in order to facilitate its interpretation and to generalize its ability to recognize patterns outside the training data set.
Random Forests (Breiman et al.) is a machine learning algorithm that uses a set of correlated trees to reduce overfitting, a well-known problem with tree models. A random forest performs parallel calculations: the data are used for training a large number of trees and the final prediction is obtained from the single trees via majority rule, i.e., the predicted outcome is the modal outcome over all predictions. In order to avoid overfitting, each tree uses a random subset of the overall set of predictors. When the size of the random set is large, the trees are likely to be very similar to each other and the final average would be very similar to that of a normal tree. Thus reducing the size of the random set of predictors makes the trees have different behaviour and reduce the dependence on the trained data so that the final fitted model is more amenable to generalizations.
Finally, the Naive Bayes algorithm is a simple classifier for predicting a nominal or ordinal variable C via estimation of the conditional probability of C given the observed value of the predictors: $P(C \mid F_1, ..., F_n)$. Using Bayes theorem and assuming independence of predictors

the latter probability can be written as $\prod_{1}^{n} P(F_i | C)P(C)$ and this, in turn, allows for efficient estimation even when *n* is large since the estimation problem is limited in a one dimensional context. The NaïveBayes algorithm behaves well even when the independence assumption is not accurate, in particular when the target is not exact class probabilities but class belonging (Hand et al. 2001).

We trained the three algorithms described above with the port data in the period January 2010 to December 2011, leaving the data from January 2012 to June 2012 (25% of total records) for testing the predictive performance of the algorithms on new instances. All algorithms were run using R version 2.15. The usefulness of the algorithms is directly related to their predictive power on test data, as the latter indicates the ability of the algorithms to generalize to new situations. Indeed, once a machine learning algorithm has been trained and tested successfully, the planner can use it to obtain a delay prevision for another day by simply substituting the input values for the day for which the prevision is required.

In general, as the training dataset becomes larger, we expect classifiers to have similar performances, even if the learning rates can be quite different. However, in real data applications the size of the dataset is fixed and it is fundamental to test different algorithms and to compare the results. In Table III we show the three common performance metrics for evaluating predictive power of a classifier: the percentage of misclassified instances, the kappa statistics and the weighted kappa. They show for each algorithm how accurate the delay alarm level prediction is. The percentage of misclassified instances is simply the percentage of incorrectly classified delay levels. The kappa statistic is defined as *k=(Pr(a)-Pr(e))/(1-Pr(e))* where *Pr(a)* is the relative frequency of agreement between predicted and observed levels, while *Pr(e)* is the probability of agreement by chance. In contrast with the raw rate of misclassified instances the kappa statistics is a more clean performance metrics since it takes into account the chance agreement between predicted and observed levels. Finally, the weighted kappa is similar to the kappa statistic but it grades agreement depending on the "distance" among categories. In particular it should be used a suitable metrics in order to give less weight to agreement as categories are further apart.

Table III – Performance metrics of Machine Learning algorithms

Methods	% of misclassified instances	kappa statistic	weighted kappa
Pruned Tree	33.66	0.131	0.111
Random Forests	**29.70**	**0.350**	**0.260**
Naive Bayes	32.67	0.154	0.133

From Table III it is clear that CART and Naive Bayes have quite similar performance while Random Forests outperforms competitors on the port dataset.

Below we illustrate some outputs from each algorithm used.

In Tables IV-VII we show, for each delay level, the conditional mean and standard deviation of each predictor variable, given the delay levels. These values are the key components of

Naive Bayes Prediction: combined with the a priori probability of the delay level they yield the final predicted probabilities.

Figure 1 shows the final pruned tree used for predicting delay level. The classification tree is composed of 7 nodes, four of which are terminal nodes. In order of importance, the main discriminating variables are: total vessels length, quota of mother vessels scheduled for the day and total number of containers to be handled. In each terminal node we indicate the absolute frequencies of the delay levels.

Finally, in Figure 2 we show the "importance plot" of the variables used by the Random Forest for predicting delay levels. This is a plot showing the importance of each predictor in the forest and it can be obtained by measuring the overall reduction in the complexity measure due to this predictor, considering all splits and all trees in the forest.

The plot clearly shows that the four variables do not have the same predictive power: in particular it seems that the variables capturing vessels "content" and "size" of the day are more important determinants of delay than those related to the scheduled movements.

In an operational context, in fact, the number of containers to be handled and the length of the incoming ships are the most important variables affecting resource allocation: berth allocation, crane intensity, human resources allocation, yard space allocation, etc.

Table IV – Naive Bayes model table: total vessel arrivals

Delay level	Mean	Standard Deviation
Low	1.65	0.85
Medium	2.32	1.12
High	2.52	1.21

Table V – Naive Bayes model table: total mother vessels

Delay level	Mean	Standard Deviation
Low	0.79	0.71
Medium	0.61	0.69
High	1.65	0.78

Table VI – Naive Bayes model table: total containers

Delay level	Mean	Standard Deviation
Low	834.15	537.30
Medium	960.08	626.31
High	1399.48	662.48

Table VII – Naive Bayes model table: total length

Delay level	Mean	Standard Deviation
Low	333.30	180.53
Medium	411.32	229.46
High	559.83	247.05

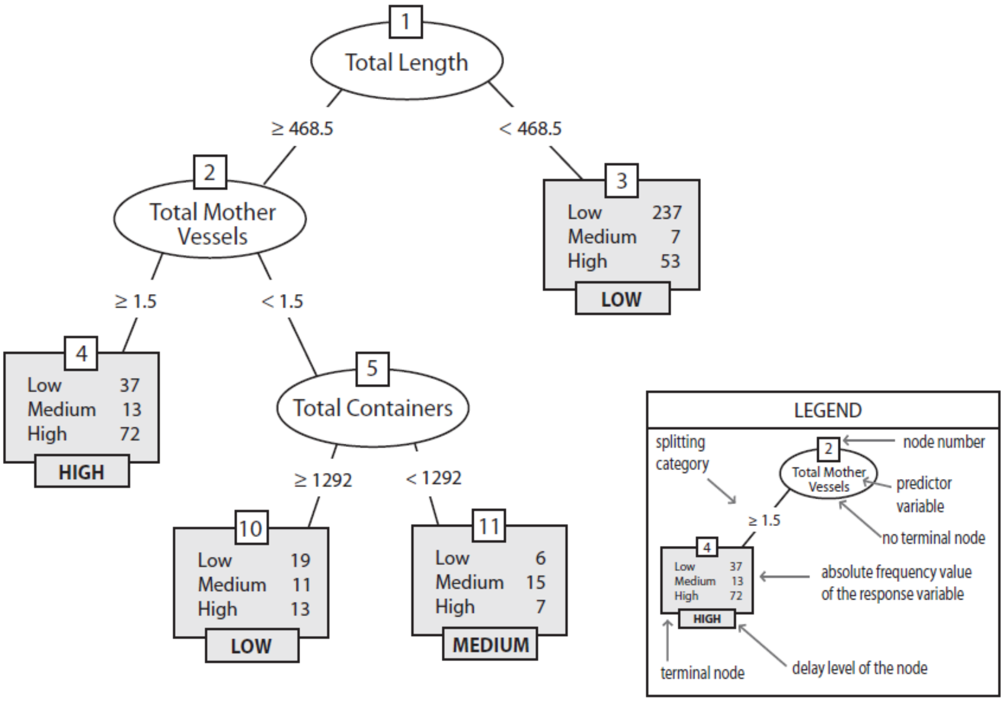

Figure 1 – Classification Tree plot.

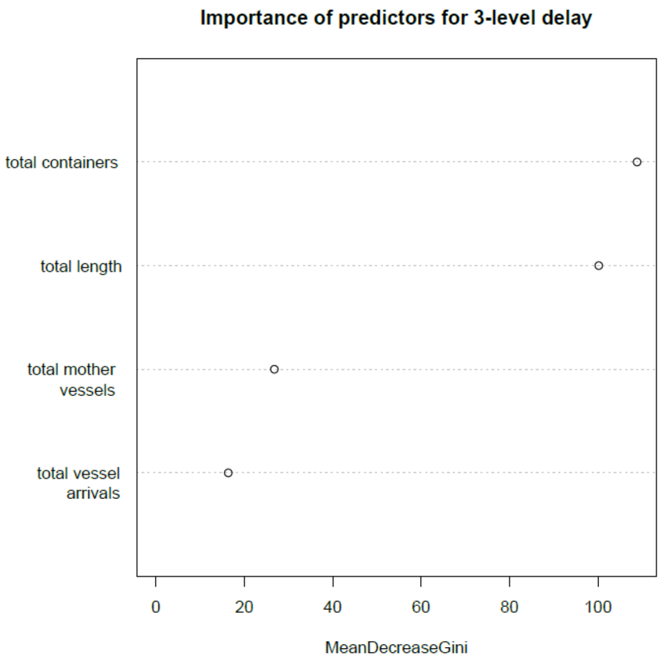

Figure 2 – Importance of predictors for the Random Forest algorithm.

PRACTICAL IMPLICATION

From a research/policy perspective, this research falls in the framework of a broader project aimed at developing a Decision Support System (DSS) for port management and for policy makers that incorporates two different modules:

1. delay alarm level forecasting, over a specific time interval;
2. specific forecasting of vessel arrival time[2].

On the basis of the alarm level severity it would be useful for the planner to be able to estimate specifically the arrival time of each vessel. The proposed models and consequently the DSS are able to learn from experience, following the well-known Data Mining paradigm "learning from data".

From a practical perspective, predicting vessel delay 24 hours in advance means that the relative demand for each work shift can be determined with greater accuracy. Moreover, the basic resources (human, mechanical and spatial) to satisfy that demand can be allocated more efficiently. This is a major issue in a container terminal where the cost, in particular that of manpower, is relatively high. Because of the uncertainty of demand, the schedule for each day assigns 'fixed' workers to one specific shift and 'flexible' workers to a shift to be decided during daily scheduling. Contractual terms for flexible shifts specify that a flexible worker will be assigned to a shift only 24 hours in advance. When demand requires additional manpower in a daily planning scenario, external workers are also hired (Fancello et al., 2011). The berth space and the handling equipment also need to be available and in perfect working order to satisfy the predicted demand. In a similar context, the DSS can be a useful support tool for planners in the short term to help them reducing the costs resulting from over/underestimation of resources. This could maximize terminal efficiency and hence competitiveness.

CONCLUSION

The demand uncertainty problem places a constraint on planning effectiveness in container terminals where the decision-making processes needs to be constantly adjusted and updated. Generally, Estimated Time of Arrival (ETA) is not respected by vessels because of unexpected events. In this work we applied a Data Mining strategy for forecasting the delay level in a Mediterranean transhipment container terminal, on a daily time horizon. The advantage of the Machine Learning algorithms used is that they can be easily "updated" to new instances by simply repeating the training process on new data, as soon as they become available. We proposed a classification of daily alarm levels related to delayed vessel arrivals in a Mediterranean port, from the point of view of port authorities. Then, three different Machine Learning models (Naive Bayes, Decision Trees and Random Forests) for predicting the alarm level for each day are tested. Three measures of agreement between the predicted and the observed levels of delay are used to compare the algorithms. The best

[2] The approach developed for direct punctual prediction is detailed in Fadda et al. (2013).

results are obtained with the Random Forest algorithm that gave a relative absolute error of 29%.

The results obtained provide a basis for furthering the research work, which will focus on refining the model with new variables (for example weather conditions which were not available for the data analyzed here) and identifying alternative approaches. For instance, a possible solution is to build a time-dependent measure of delay, $D(t)$, and exploit the possible relation of $D(t)$ with $D(t-k)$ for $k=1,2,\dots,n$, using time series models. If port activities are performed continuously, these models may perform well since the past status is likely to be a good predictor of the present status.

REFERENCES

Breiman, L., J. Friedman, R. Olshen and C. Stone (1984). Classification and Regression Trees. Wadsworth Belmont, CA.

Brock, G., V. Pihur, S. Datta and S. Datta (2008). clValid: An R package for cluster validation. Journal of Statistical Software, 25.

Fadda, P., G. Fancello, L. Frigau, F. Mola and C. Pani (2013). A Data Mining Approach to Forecast Vessel Arrival in a Transhipment Container Terminal. In press.

Fancello, G., C. Pani, M. Pisano, P. Serra, P. Zuddas and P. Fadda (2011). Prediction of arrival times and human resources allocation for container terminal. Maritime Economics & Logistics, 13, 142 – 173.

Fancello, G., C. Pani, P. Serra, P. Zuddas and P. Fadda (2010). Previsione degli arrivi delle navi ed ottimizzazione delle risorse umane in un terminal hub container. Scenari di riferimento per i porti container italiani nel sistema euro-mediterraneo. Applicazioni di modelli e metodi per la previsione della domanda e per la valutazione dell'offerta”, ISBN 978-88-568-3214-3, Franco Angeli, Milano.

Fung, M.K. (2002). Forecasting in Hong Kong’s container throughput: An error-correction model. Journal of Forecasting, 21, 69–80.

Gambardella, L.M., G. Bontempi, E. Taillard, D. Romanengo, G. Raso and P. Piermari (1996). Simulation and forecasting in intermodal container terminal. In: Proceedings of the 8th European Simulation Symposium. Society for Computer Simulation International.

Gambardella, L. M., A. E. Rizzoli and M. Zaffalon (1998). Simulation and planning of an intermodal container terminal. Simulation, 71, 107–116.

Hand, D.J., and Yu, K. (2001). Idiot's Bayes - not so stupid after all?. International Statistical Review, 69, 385-399.

Hartigan, J. A. and M. A. Wong (1979). A K-means clustering algorithm. Applied Statistics, 28, 100–108.

Lu, Z., J. Wang and G. Zheng (2008). A new method to alarm large scale of flights delay based on machine learning. Proceeding of The International Symposium on Knowledge Acquisition and Modeling (KAM 2008). Wuhan, China, IEEE CS Press, 589-592.

Murty, K.G., J. Liu, Y. Wan and R. Linn (2005). A decision support system for operations in a container terminal. Decision Support Systems, 39, 309– 332.

Rendón, E., I. Abundez, A. Arizmendi and E. M. Quiroz (2011). Internal versus External cluster validation indexes. International journal of computers and communications, 5.

Salido, M. A., M. R. Molins and F. Barber (2012). A decision support system for managing combinatorial problems in container terminals. Knowledge-Based Systems 29, 63–74.

Sideris A. C., M. P. Boilé and L. N. Spasovic (2002). Using On-Line Information To Estimate Container Movements For Day-To-Day Marine Terminal Operations. 81st Annual Meeting of the Transportation Research Board, Washington.

Theodoridis S. And K. Koutroubas (1999). Pattern Recognition, Academic Press.

Zhen L., L.H. Lee and E. P. Chew (2011). A decision model for berth allocation under uncertainty. European Journal of Operational Research 212, 54–68.