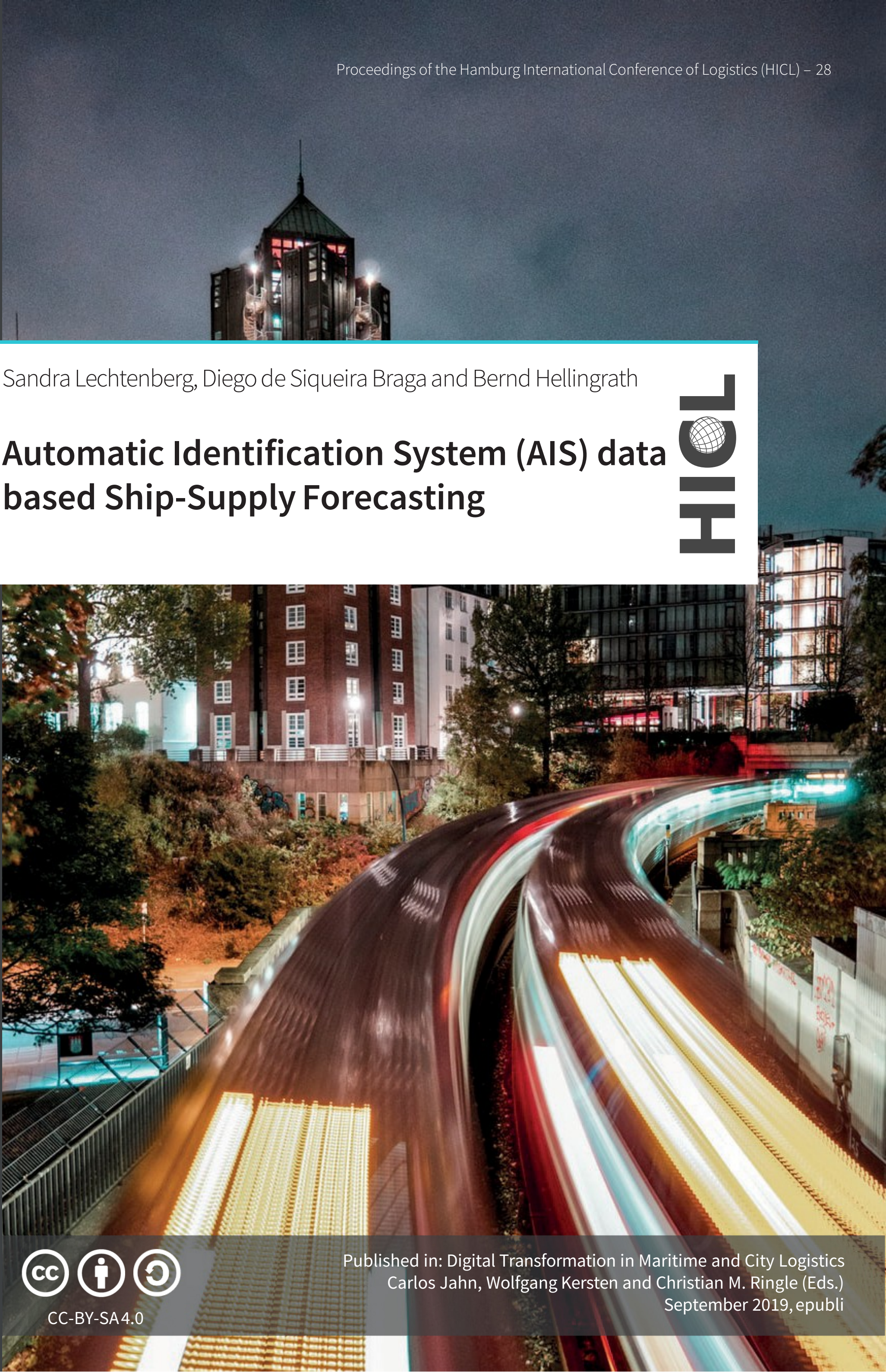
Proceedings of the Hamburg International Conference of Logistics (HICL) – 28

Sandra Lechtenberg, Diego de Siqueira Braga and Bernd Hellingrath

Automatic Identification System (AIS) data based Ship-Supply Forecasting



Published in: Digital Transformation in Maritime and City Logistics
Carlos Jahn, Wolfgang Kersten and Christian M. Ringle (Eds.)
September 2019, epubli

Automatic Identification System (AIS) data based Ship-Supply Forecasting

Sandra Lechtenberg[1], Diego de Siqueira Braga[1] and Bernd Hellingrath[1]

1 – University of Münster

Purpose: The bulk cargo shipping industry is characterized by high cost pressure. Chartering vessels at low prices is important to increase the margin of transporting cargo. This paper proposes a three-step, AI-based methodology to support this by forecasting the number of available ships in a region at a certain time.

Methodology: Resulting from discussions with experts, this work proposes a three-step process to forecast ship numbers. It implements, compares and evaluates different AI approaches for each step based on sample AIS data: Markov decision process, extreme gradient boosting, artificial neural network and support vector machine.

Findings: Forecasting ship numbers is done in three steps: Predicting the (1) next unknown destination, (2) estimated time of arrival and (3) anchor time for each ship. The proposed prediction approach utilizes Markov decision processes for step (1) and extreme gradient boosting for step (2) and (3).

Originality: The paper proposes a novel method to forecast the number of ships in a certain region. It predicts the anchor time of each ship with an MAE of 5 days and therefore gives a good estimation, i.e. the results of this method can support ship operators in their decision-making.

Keywords: AIS data, Ship-supply Forecasting, Dry Bulk Cargo, Artificial Intelligence

First received: 19.May.2019 **Revised:** 27.May.2019 **Accepted:** 11.Jun.2019

1 Introduction

Maritime transport is by far the most used mode to transport goods worldwide. It is believed that more than 90% of the world's goods are transported by sea (Grote, et al., 2016). Indeed, the seaborne trade has grown by 4 percent in 2017 and 10.7 billion tons have been transported. Especially containerized and dry bulk cargo shipping is growing with the latter one constituting almost half of the total dry cargo shipments (United Nations Conference on Trade and Development (UNCTAD, 2018). Dry bulk cargo is solid raw material which is transported mostly unprocessed and in unpackaged large quantities. It can be easily stowed in a single hold with little risk of cargo damage. Moreover, transporting commodities in bulk provides economies of scale. Most of transported dry bulk cargo is comprised of iron ore, coal and grain but also goods such as agricultural products, cement or forest and steel products are transported in bulk (Rodrigue and Browne, 2002). Ship operators charter available ships to serve as carriers and transport dry bulk cargo from A to B. In the last years, there is a decrease in the world fleet growth each year. However, the supply of vessels still increases faster and is by far greater than the demand. This situation of unbalanced demand and supply puts pressure on freight rates leading to continuously decreasing earnings for ship operators (UNCTAD, 2017; 2018). Because of this situation, it is more important for ship operators to know the future demand and supply as precisely as possible. Especially knowing areas, where many vessels will be available for booking in a certain period, helps to secure cargo to be transported from that area early and then hire ships for a rate, which is cheap due to a high number of competitors. Despite the advantage that can be generated by knowing the number of ships in a certain region, there is a lack of research aiming at ship-supply forecasting. Hence, this paper aims at proposing a way to support ship owners by providing information regarding the expected number of ships in a region at a certain time. The remaining paper is structured as follows: Section 2 introduces general knowledge about forecasting in the maritime industry, automatic identification system (AIS) data as well as reasons for utilizing artificial intelligence (AI) based forecasting approaches. The developed ship-supply forecasting methodology is presented in section 3 and followed by a discussion. Section 4 subsumes the paper by highlighting the main results, discussing important limitations and providing an outlook on future research possibilities.

2 AIS data based forecasting in the maritime industry

Forecasting in the maritime industry plays a central role for all included parties such as harbors, manufacturers, and ocean carrier companies. Without short and long-term forecasting of e.g. demand, the allocation of resources, capacity planning and making investment decisions in case of potentially required upscaling, are nearly impossible (Mensah and Anim, 2016). Hyndman and Athanasopoulos (2018) define five generic steps that have to be processed in order to perform forecasting: (1) Problem definition, (2) Gathering information, (3) Preliminary exploratory analysis, (4) Choosing and fitting models and (5) Evaluating a forecast model. While the step of the problem definition might seem trivial at first sight, it is possibly the most difficult step in the forecasting process as it requires a thorough and deep understanding of how the forecast will be used, what the requirements are and how it fits within the structure and processes of the organization (Hyndman and Athanasopoulos, 2018).

In order to forecast something, information such as statistical data, accumulated expertise of employees or historical data is needed. While historical internal data like actual demand occurrence can be gathered automatically over time, external quantitative data has to be acquired and qualitative data has to be accumulated in some way. Independently from the type of available data, information should be thoroughly analyzed before selecting a forecasting method. This is necessary because various models differ in their applicability. Also the selected model should be able to capture the genuine patterns that can be found in historical data but should not replicate occurrences that happened in the past but are not likely to happen again (Hyndman and Athanasopoulos, 2018). AIS data is one information source, which can be of high value for forecasting in the maritime industry. This data is transmitted by AIS transceivers which are installed on vessels and which automatically broadcast information, such as their position, speed, and navigational status. This information is received by other ships, terrestrial receiver stations e.g. by coastal authorities and by satellites (also referred to as S-AIS). As all ships over 299 gross tons (GT) are obliged to be equipped with such a transceiver since December 2004, the major amount of ships interesting to bulk dry cargo shippers carry one on board (Zorbas, et al., 2015).

The information transmitted is threefold: The (1) dynamic broadcast information contains navigational information, which is updated and transmitted automatically every 2 to 10 seconds. The (2) Voyage related information such as the declared destination and estimated time of arrival (ETA) of a trip and (3) static vessel information containing e.g. the ship identifier, name and type, are entered by the vessel's crew and transmitted every 6 minutes, regardless of the vessel's movement status. As regulations require a substantial portion of ships to transmit AIS data, the amount of data collected over time offers significant analytic potential. Besides fleet and cargo tracking, AIS data is mainly used for maritime security, collision avoidance, fishing monitoring as well as for search and rescue (Weintrit and Neumann, 2013).

Mao, et al. (2018) do not directly focus on using AIS data for forecasting but present the construction of an AIS-based database that can serve as an input for further analyses based on AIS data. By classifying vessels and focusing on AIS data sent by fishing vessels, Mazzarella, et al. (2014) are able to automatically detect fishing areas. Pallotta, Vespe and Bryan (2013) aim at increasing situational awareness in the maritime industry by better understanding maritime traffic patterns. They use an unsupervised and incremental learning approach that derives characteristics of ports and off shore platforms as well as spatial and temporal distribution of routes from AIS data. Their results can form a basis to allow for anomaly detection, i.e. ships that deviate from the identified route patterns. Similarly, Nguyen, et al. (2018) develop a multi-task deep learning framework for vessel monitoring in order to reconstruct taken routes, identify vessel types and detect abnormal vessel behaviors. Anomaly detection is also one of the activities related to how AIS data is used for knowledge discovery in the maritime domain discovered by Alessandrini, et al. (2016). Their survey of recent Joint Research Centre (JRC) activities also identifies the mapping of maritime routes or fishing activities as well as monitoring shipping activities in the arctic or falsification of AIS data, i.e. the verification of trustworthiness of

AIS data, as relevant fields. Another application possibility is presented by Ambjörn (2008) who use AIS data to identify which ships have been close to oil in the Baltic or North Sea over a period of time and are therefore likely to be responsible for such an oil spill.

Regarding the forecast of ship-supply so far - to the best of our knowledge - no research has investigated possibilities to predict the number of available ships in a certain region of interest. While there certainly are sources that use AIS data either to estimate the position of one ship in the future or to identify certain route patterns, the achieved results do not give an indication about a general ship availability in the future. For example, Xiao, et al. (2017) use a density-based spatial clustering of applications with noise (DBSCAN) algorithm to extract waterway patterns and predict maritime traffic 5, 30 or 60 minutes ahead - a time horizon which is not long enough to allow for early cargo offer securing. Similarly, the identification of routes typically taken by ships, as e.g. presented by Mazzarella, Arguedas and Vespe (2015), is capable of increasing maritime situational awareness in general but does not provide information dedicated to a specific situation at a certain place and time of interest.

However, what becomes apparent when looking at sources forecasting based on AIS data, is that most of them apply non-traditional forecasting techniques. In this paper, traditional forecasting is mainly seen as quantitative methods and statistical techniques. They objectively predict the demand based on past patterns and relationships. This means, those techniques need historical data for their predictions and are not able to identify systematic changes. It is important to emphasize that the quality of accuracy mainly depends on the target value that is forecasted. Some aspects can be predicted very exact like for instance the sunset times for the next year, whereas other factors are uncertain to forecast. These are for example exchange rates or stock prices. In general, it is difficult for traditional techniques to manage a huge amount of past data in a way to identify the right patterns and relationships of features. Moreover, often not all of the required data of the past years exist or is available. Also, the weak reference of historical data to current activities is a limitation (Bursa, 2008; Byrne, 2012).

As AIS data contains information sent by a huge amount of ships, its size gets too big for such more traditional, statistical forecasting techniques quite fast. Hence, authors tend to rely on more advanced techniques from the field of so-called AI. There is no commonly accepted definition of what AI is or what methods belong to it, but it is generally described as "computational systems that perform tasks commonly viewed as requiring intelligence" (Poole and Mackworth, 2017). AI-based techniques are capable of processing more data and identifying feature-output relations that remain hidden to both a human observer and most statistical techniques. Hence, they seem suitable to be utilized when forecasting something based on AIS data.

3 AIS data based Ship-Supply Forecasting

3.1 Conceptualization of the forecasting process

The objective of this paper is to develop a method that is capable of supporting decision-makers of ship operators by providing better information about the available ship-supply. Based on expert feedback it has been decided to forecast the availability of ships based on regions not specific countries or ports. For deciding on where to operate ships, i.e. where to secure cargo, it is sufficient to know which regions will be crowded. As moving a ship within one region is not very costly, the information on where one ship exactly is, is not relevant to decision makers. Therefore, the world map has been separated into regions as defined by the cooperating industry expert. As the destination port stated in the AIS voyage data is entered by the crew by hand and thus not standardized, a matching procedure has been implemented in order to assign global positioning system (GPS) coordinates to a stated destination port and subsequently a destination region. Destinations are mainly stated in two ways: They can be stated as the port name that can be matched directly or via regular expressions or stated as the UN Code abbreviation for ports. This United Nations (UN) code format consists of five letters, the first two resembling the country the port is located in, and the latter three abbreviating the city/location of the port. Thus, a comprehensive reference table containing over 6000 ports with their respective UN code abbreviation has been utilized.

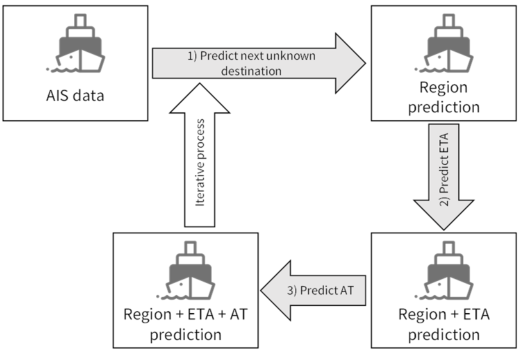


Figure 1: Forecasting process

Forecasting is done on an individual ship level, i.e. it is predicted for each ship at what date it will be at which destination region. To do so, the forecasting process has been divided into several prediction steps as depicted in Figure 1 and described in the following:

(1) Next unknown destination region prediction: The first step is the prediction of the next unknown region. For each ship, it is predicted what the next unknown destination region will be, which the ship will head to after the current destination that is stated in the AIS broadcast. Since the prediction outcome is based on the different 44 regions from the world map, the outcome is categorical. For instance, for ship XY the outcome of the prediction could be region 44.

(2) Estimated time of arrival (ETA) prediction: In the next step, it is necessary to estimate how long it will take each ship to arrive at their destination region. In this case, the prediction outcome is quantitative. For example, ship XY needs 9 days to arrive at the destination region 44.

(3) Anchor time (AT) prediction: In the third step, it is essential to predict for each ship how long it will stay at the destination region until the next trip will begin. Again, the output is quantitative. For instance, ship XY will stay for the next 4 days in the destination region until it will begin a new trip.

After the iterative forecasting process is done, the results for each ship should contain information about the start region and the destination region. Furthermore, the time measures ETA and AT should be included. Additionally, the vessel id should be stated in order to assign each trip to a vessel. Finally, the results are aggregated to provide the relevant information about how many ships are in which region at a certain point of time.

3.2 Selection of suitable forecasting methods

To handle the classification and prediction steps defined in section 3.1 different methods have been compared. Such a comparison is mainly dependent on the available data. The right data preparation is one of the main factors for a high accuracy (Carbonneau, Laframboise and Vahidov, 2008). The case at hand contains prediction and classification problems with an available set of already labeled example data. Therefore, supervised algorithms are the best-fitted ones. Caruana and Niculescu-Mizil (2006) compare supervised learning algorithms using different performance metrics. Their results show that boosted tree algorithms, support vector machines (SVM) and artificial neural networks (ANN) perform best, which is why these three haven been chosen to be tested with the problem of ship-supply forecasting. Based on the data at hand, Markov decision process has also been selected as an alternative method to the machine learning approaches. The reasons for this additional selection is that especially for the region prediction the number of next unknown destinations is limited and Markov decision processes are suitable to depict the situation of selecting an action, i.e. a new region, based on the current state of a ship.

XGBoost is a scalable system implementing the gradient decision tree boosting approach based on Friedman (2001) and is widely used by data scientist, e.g. in machine learning challenges. With XGBoost simple, weak decision tree models are used as a basis. New models are created to predict errors of earlier ones and this way to iteratively improve a final model with marginally modified parameter settings. Its major contributions are among others a sparsity-aware algorithm for parallel tree learning and the ability to handle instance weights in approximate tree learning (Chen and Guestrin, 2016; Reinstein, 2017). SVMs as well as ANNs are quite common supervised machine learning techniques. The first aims at identifying a hyperplane, which best separates the given data points based on their features to then classify new data points according to this hyperplane. The hyperplane forms a boundary separating the data points with the biggest possible distance to them (Hearst, 1998; Russel and Norvig, 2017). ANNs are built after biological networks such as the human brain. The approach is able to detect hidden relationships within the input data. An ANN typically consists of one input layer, several hidden layers and one output layer each consisting of several neurons, which are connected to each other. Each neuron possesses an activation function, which determines whether the neuron is triggered by the former layer’s signals, i.e. the input data. The triggered neurons process the data based on their activation function. The connections between each neurons have a certain weight and the learning process is based on placing adjusting these weights depending on the error of the output produced by the neurons (Tu, 1996; Poole and Mackworth, 2017). In contrast to the other compared approaches, a Markov decision process is no machine learning approach. It is a mathematical framework applied for modeling decision making on a stochastic background. It is based on a discrete time stochastic process, consisting of the current state and possible actions that are to be performed in order to get to the next state. Thus, a Markov decision process consists of a set of possible world states and a set of possible actions (Sutton and Barto, 2017).

For each of the selected forecasting methods and each of the forecasting steps, a prototype has been implemented. The available AIS data has been divided into 90% training and 10% test data. Having executed all prototypes, some performance measures have been calculated to compare the

different methods (cf. Table 1) and Based on the results from the testing phase, it has been decided to use a hybrid solution for the final implementation. For the next unknown region prediction, the Markov decision method was used. For the regressions of ETA and AT, XGBoost was applied. Thus, for the final solution these two methods have been combined into one iterative forecasting process. The step-based forecasting allows for selecting the best approach for each step and hence a hybrid solution leads to the overall best results. As e.g. the accuracy of the ETA forecast depends on whether the correct next unknown region has been predicted, the ship-supply forecasting process will naturally lead to better results when selecting the best-fitted approach for each step.
Table 2).
The results show that the Markov decision process is best suited to predict the next unknown destination region achieving a prediction accuracy of 98%. Also, either about 15 ships too much or too less were predicted for each destination region of the testing data. XGBoost achieves the lowest mean absolute error (MAE) for both ETA and AT prediction, while the root mean square error (RMSE) is about the same for all tested approaches. For each trip of the testing data, on average four days too much or too little were predicted for the ETA. For the AT prediction, with XGBoost roughly four to five days too much or too little were predicted.

Table 1: Measures of next unknown region prediction

	Accuracy	F1	MAE	RMSE
Markov	0.974	-	15.81	24.09
XGBoost	0.494	0.472	666.47	1840.03
ANN	0.519	0.29	1445.95	3719.95
SVM	0.502	0.361	832.6	2028.23

Based on the results from the testing phase, it has been decided to use a hybrid solution for the final implementation. For the next unknown region prediction, the Markov decision method was used. For the regressions of ETA and AT, XGBoost was applied. Thus, for the final solution these two methods have been combined into one iterative forecasting process. The step-based forecasting allows for selecting the best approach for each step and hence a hybrid solution leads to the overall best results. As e.g. the accuracy of the ETA forecast depends on whether the correct next unknown region has been predicted, the ship-supply forecasting process will naturally lead to better results when selecting the best-fitted approach for each step.

Table 2: Measures of ETA and anchor time prediction

	ETA: MAE	ETA: RMSE	AT: MAE	AT: RMSE
Markov	4.9	14.21	5.8	13.63
XGBoost	3.96	14.7	4.48	14.1
ANN	4.86	14.48	4.89	13.84
SVM	5.45	14.47	5.39	13.74

3.3 Visualization and utilization of forecasting method

The results generated by the hybrid, three-step and AI-based algorithm are presented in a Table to the decision maker (Table 3 shows an example of how such an export looks like). This Table provides the number of ships for each destination region along each day of the forecasting horizon. By this, the decision maker resp. ship operator gets a brief and extensive overview of all ships e.g. for the next 30 days.

Based on this Table the decision-maker can estimate if it is promising to accept a cargo offer in a certain region. For example, there will be comparatively many ships in region 1 around July 23. Hence, it will likely lead to low rates for hiring a ship. If the decision-maker is now able to secure cargo in that region for the respective time, it will most likely generate more earnings than in regions or times with a smaller level of ship-supply.

Table 3: Example Ship-Supply forecast export

Region	**22.07.18**	**23.07.18**	**…**	**20.08.18**
Region 1	7	18	…	1
Region 2	0	0	…	12
Region 3	0	15	…	0
…	…	…	…	…
Region n	3	1	…	0

More sophisticated or graphical visualizations are possible as well for example by using tools such as Kibana, which allow to generate heat maps according to the number of available ships etc. However, expert feedback was given that a simple list with numbers is preferred as it reflects the results more detailed and accurately. Moreover, domain knowledge can be better used to interpret the numbers regarding their validity and significance.

The entire forecasting process was developed in the script language R. It was divided into five different R scripts in order to deploy the entire concept. An overview script, which combines all other scripts, reads all necessary input data, starts the data preparation function and later on, the iterative forecasting process. Lastly, the forecasting results are stored in the target Table of the database. There are three scripts containing the functions necessary for the forecasting process: one for the data preparation, one for the Markov decision models and one for both XGboost models (ETA and AT). The fifth script includes all necessary packages and dependencies and is sources by the overview script to provide them. All that is needed to run and use the forecasting process is a database storing the necessary AIS input data and providing Tables for storing the generated output as well as the presented scripts, which can e.g. be stored on a server to be run from there. The next step could be to integrate a job to start the forecasting process regularly in a time interval as desired. Once the forecasts are calculated, ship operators can use it as an additional source of information to base their decisions on.

4 Conclusion

Overall, the paper proposed a three step, AI-based method to forecast the number of ships in a certain region at a time of interest. The ship-supply forecasting method has been conceptualized on the foundation of available literature as well as expert feedback. Based on predefined maritime regions as well as the estimated time of arrival per ship, it has become possible to forecast ship availability as far as the time horizon of the existing input data allows.

While the objective of the paper has been fulfilled, there are certain limitations, which should be kept in mind, as well as possibilities for future development to enhance the method and its results. First of all, the set of techniques evaluated can be extended. A number of approaches has been sought which are appropriate to the problem as well as the data and are therefore promising, but as the set was not exhaustive, it cannot be guaranteed that no other approach leads to equally good or even better forecasts. Moreover, the time horizon of utilized AIS data has been limited and a test with an extended data set, spanning over a longer time horizon, would surely increase the meaningfulness of the forecasting results. Moreover, it is necessary to always keep the accuracy of the predictions in mind when using them for decision-support. The real number of available ships will differ from the predicted numbers and hence have influence on the rate at which ships can be booked. Nonetheless, a deviation of 4 to 5 days in ETA or AT is not major compared to the days it takes a ship to travel from one region to another. Hence, the difference between expected and actually available ships should not be big enough to not use them as a support for deciding where to secure cargo. Even if the number of ships differs, the rates at which ships can be rented will not change dramatically if the predicted amount is roughly as expected.

Regarding future research possibilities, especially the integration of further information to improve the forecasting quality is of high importance. First, expert knowledge could be integrated e.g. in the form of rules. The main purpose could be to remove errors in the forecasting results. For example, explicit knowledge about ports just serving as maintenance or refueling points could be incorporated this way. Other interesting aspects are region relationships or time specifications. Explicit knowledge about the relationships between regions could be useful to avoid forecasting trips from some region to another that would never happen in reality. Based on these rules, the Markov probability matrix could be adjusted to avoid impractical trips. Additionally, time specifications could contain information such as, how long a vessel needs at least to go from one region to the destination region or what the maximum anchor time of a vessel could be. The new information enables the results to be checked and adjusted in case erroneous predictions and also so the adjustment of the learning model. Another way to improve the forecasting results is to integrate known seasonality patterns. As seasonality is one of the greatest uncertainties in dry bulk cargo shipping and leading to e.g. freight rate volatility, the integration of its patterns can result in a higher prediction accuracy. Aspects that could be used to depict seasonality are commodity seasonality, weather data or general maritime traffic patterns.

Even when keeping the limitations and possible improvements in mind, the proposed method is a good starting point for generating valuable information, which can support ship operators in their daily business and help to generated more revenue.

Acknowledgments

We would like to thank 24Vision.Solutions[1] for their cooperation, continuous support as well as their valuable input and feedback. Moreover, we would like to thank all students, who directly or indirectly contributed to this project, for their dedicated work - in particular the team of the project seminar "CargoInShip": Fabian Lutze, Moritz Mersmann, Liliia Mustafina, Raphael Patrick Prager, Maurice Straube, Ekky Wilmasara and Moritz Witte.

[1] https://www.24vision.solutions/

References

Alessandrini, A., Alvarez, M., Greidanus, H., Gammieri, V., Arguedas, V. F., Mazzarella, F., Santamaria, C., Stasolla, M., Tarchi, D. and Vespe, M., 2016. Mining Vessel Tracking Data for Maritime Domain Applications. In: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). Barcelona, Spain: IEEE, pp. 361–367.

Ambjörn, C., 2008. Seatrack Web forecasts and backtracking of oil spills - an efficient tool to find illegal spills using AIS. IEEE/OES US/EU-Baltic International Symposium, pp. 1–9.

Bursa, K., 2008. How to effectively manage demand with demand sensing and shaping using point of sales data. Journal of Business Forecasting, 27(4), pp. 26–28.

Byrne, R. F., 2012. Beyond Traditional Time-Series. Using Demand Sensing to Improve Forecasts in Volatile Times. Journal of Business Forecasting, 31(2), pp. 13–19.

Carbonneau, R., Laframboise, K. and Vahidov, R., 2008. Application of machine learning techniques for supply chain demand forecasting. European Journal of Operational Research, [e-journal] 184(3), pp. 1140–1154. http://dx.doi.org/10.1016/j.ejor.2006.12.004.

Caruana, R. and Niculescu-Mizil, A., 2006. An Empirical Comparison of Supervised Learning Algorithms. In: Proceedings of the 23rd International Conference on Machine Learning. Pittsburgh, PA.

Chen, T. and Guestrin, C., 2016. XGBoost. In: B. Krishnapuram, M. Shah, A. Smola, C. Aggarwal, D. Shen, and R. Rastogi. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD ‘16. the 22nd ACM SIGKDD International Conference. San Francisco, California, USA, pp. 785–794.

Friedman, J. H., 2001. Greedy Function Approximation. A Gradient Boosting Machine. Annals of statistics, pp. 1189–1232.

Grote, M., Mazurek, N., Gräbsch, C., Zeilinger, J., Le Floch, S., Wahrendorf, D.-S. and Höfer, T., 2016. Dry bulk cargo shipping - An overlooked threat to the marine environment? Marine pollution bulletin, 110(1), pp. 511–519.

Hearst, M. A., 1998. Support vector machines. IEEE Intelligent Systems and their applications, 13(4), pp. 18–28.

Hyndman, R. J. and Athanasopoulos, G., 2018. Forecasting. Principles & Practice. 2nd ed. Melbourne, Australia: OTexts.

Mao, S., Tu, E., Zhang, G., Rachmawati, L., Rajabally, E. and Huang, G.-B., 2018. An Automatic Identification System Database for Maritime Trajecotry Prediction. Proceedings of ELM-2016, pp. 241–257.

Mazzarella, F., Arguedas, V. F. and Vespe, M., 2015. Knowledge-based vessel position prediction using historical AIS data. In: 2015 Sensor Data Fusion: Trends, Solutions, Applications (SDF). Bonn, Germany: IEEE, pp. 1–6.

Mazzarella, F., Vespe, M., Damalas, D. and Osio, G., 2014. Discovering vessel activities at sea using AIS data: Mapping of fishing footprints. 17th International Conference on Information Fusion, pp. 1–7.

Mensah, J. and Anim, S. K., 2016. DEMAND FORECASTING IN THE MARITIME INDUSTRY, A CASE OF MAERSKLINE GHANA. Archives of Business Research, [e-journal] 4(1). http://dx.doi.org/10.14738/abr.41.1841.

Nguyen, D., Vadaine, R., Hajduch, G., Garello, R. and Fablet, R., 2018. A Multi-task Deep Learning Architecture for Maritime Surveillance using AIS Data Streams. IEEE 5th International Conference on Data Science and Advanced Analytics, pp. 331–340.

Pallotta, G., Vespe, M. and Bryan, K., 2013. Vessel Pattern Knowledge Discovery from AIS Data: A Framework for Anomaly Detection and Route Prediction. Entropy, 15(12), pp. 2218–2245.

Poole, D. L. and Mackworth, A. K., 2017. Artificial intelligence. Foundations of computational agents. Cambridge: Cambridge University Press.

Reinstein, I., 2017. XGBoost, a Top Machine Learning Method on Kaggle, Explained. [online] Available at: <https://www.kdnuggets.com/2017/10/xgboost-top-machine-learning-method-kaggle-explained.html> [Accessed 24 April 2019].

Rodrigue, J.-P. and Browne, M., 2002. International maritime freight transport and logistics. In: R. Knowles, J. Shaw, and I. Docherty, eds. 2002. Transport Geographies: An Introduction: Blackwell Publishing, pp. 156–178.

Sutton, R. S. and Barto, A. G., 2017. Reinforcement Learning. An Introduction. 2nd ed. Cambridge, Massachusetts: MIT Press.

Tu, J. V., 1996. Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medial Outcomes. Journal of Clinical Epidemiology, 49(11), pp. 1225–1231.

United Nations Conference on Trade and Development (UNCTAD), 2017. Review of Maritime Transport 2017. New York and Geneva.

United Nations Conference on Trade and Development (UNCTAD), 2018. Review of Maritime Transport 2018. New York and Geneva.

Weintrit, A. and Neumann, T., 2013. Marine navigation and safety of sea transportation. Maritime transport & shipping. London: CRC Press.

Xiao, Z., Ponnambalam, L., Fu, X. and Zhang, W., 2017. Maritime Traffic Probabilistic Forecasting Based on Vessels' Waterway Patterns and Motion Behaviors. EEE Transactions on Intelligent Transportation Systems, 18(11), pp. 3122–3134.

Zorbas, N., Zissis, D., Tserpes, K. and Anagnostopoulos, D., 2015. Predicting Object Trajectories From High-Speed Streaming Data. IEEE Trustcom/BigDataSE/ISPA, 2, pp. 229–234.