From Conventional to Machine Learning Methods for Maritime Risk Assessment

A. Rawson[1], M. Brito[1], Z. Sabeur[2] & L. Tran-Thanh[3]
[1] University of Southampton, Southampton, UK
[2] Bournemouth University, Bournemouth, UK
[3] University of Warwick, Warwick, UK

ABSTRACT: Within the last thirty years, the range and complexity of methodologies proposed to assess maritime risk have increased significantly. Techniques such as expert judgement, incident analysis, geometric models, domain analysis and Bayesian Networks amongst many others have become dominant within both the literature and industry. On top of this, advances in machine learning algorithms and big data have opened opportunities for new methods which might overcome some limitations of conventional approaches. Yet, determining the suitability or validity of one technique over another is challenging as it requires a systematic multicriteria approach to compare the inputs, assumptions, methodologies and results of each method. Within this paper, such an approach is proposed and tested within an isolated waterway in order to justify the proposed advantages of a machine learning approach to maritime risk assessment and should serve as inspiration for future work.

1 INTRODUCTION

Accidents to navigating vessels have the potential to result in loss of life, environmental pollution and economic losses. To better understand where and why these accidents occur, a significant body of literature that might be described as maritime risk analysis has developed, particularly within the last decade [23]. These methods seek to apply quantitative methods to measure the likelihood or consequence of hazards such as collisions, groundings and allisions from occurring. These risks have an inherent spatial component and mapping where these risks are highest can help decision makers in both allocation of risk control measures and marine spatial planning.

The multitude and variety of these methods is significant, with reviews undertaken by several authors [4, 21–23, 28]. A key challenge when comparing these methods is that each have their own assumptions and limitations that could introduce biases and therefore it would be unreasonable to assume that any one method works better than others in all situations. Whilst this makes evaluation between established models difficult, it also makes judgements on the suitability and validity of novel methods similarly challenging.

One such novel approach is that of the use of machine learning (ML) techniques for maritime risk assessment. ML might be described as a subset of artificial intelligence whereby algorithms improve through experience rather than being explicitly programmed. These models can be supervised, whereby the model is constructed on data containing both input and outputs, or unsupervised, whereby structure is sought on unlabelled data. Few have sought to apply ML to ship navigation [9, 36] and even fewer have attempted to use these methods to assess navigation safety [5, 19]. Whilst some have

argued that such methods have numerous advantages over traditional risk assessment techniques [15], such benefits have not been demonstrated within the maritime domain.

A possible solution to this challenge is through a systematic evaluation of different methods within a common framework, using set criteria against which each method can be judged. In [28], a comparison is made between 20 models against standardised criteria, but without implementing these techniques, evaluation is more challenging, particularly where the models are proprietary.

This paper sets out to develop such a framework by comparing 10 different maritime risk analysis methods and identifying suitable criteria that could be used for an evaluation. We make a number of contributions; firstly, we conduct a systematic and applied evaluation of a selection of the most widely researched maritime risk models, in order to highlight their methodological strengths and weaknesses. Secondly, we introduce four ML techniques and how they could be utilised for predicting the likelihood of accidents, with some high-level implementations and a discussion of opportunities for greater application of these techniques. Thirdly, we propose a list of criteria through which these methods can be directly compared, proposing further work for a multi-criteria evaluation of maritime risk models. Whilst the evaluation requires further work, we make a number of observations on the different techniques to provide initial feedback on the capability of ML for maritime risk assessment.

1.1 Case Study

To achieve these aims, we utilise a case study of the waterway between Washington State (United States), Vancouver Island (Canada) and British Columbia (Canada). This waterway is known as the Puget Sound or Salish Sea, and extends from the Pacific Ocean, through the Strait of Juan de Fuca, before heading north through the San Juan Islands towards Vancouver, or south through Admiralty Inlet towards Seattle (Figure 1). This area is notable for several reasons. Firstly, it has a significant volume of traffic, of all types, including cargo and tanker traffic bound for various ports and terminals, significant recreational and fishing fleets, and major ferry routes. Secondly, traffic within the area is managed by Traffic Separation Schemes (TSS), pilotage districts, escort towage and a cooperative VTS between the United States and Canada. Thirdly, the area has been extensively studied in other maritime risk studies, most notably Vessel Traffic Risk Assessment (VTRA) [39].

Vessel traffic data from the Automatic Identification System (AIS) was obtained from the MarineCadastre for June 2018 covering the waterway. AIS is an automatic ship reporting system that transmits dynamic (positional, speed and course) and static (ship type and size) information that can be collected to produce high spatial-temporal resolution datasets. Furthermore, incident data was available from the US and Canadian Coastguards for the years 2002-2014.

Figure 1. Study Area with TSS overlaid.

2 CONVENTIONAL METHODS

Six broad conventional maritime risk analysis methods were identified from the literature and are discussed below.

2.1 Risk Matrices and Expert Judgement

At an operational level, most decisions on maritime safety are made using risk matrices. Such an approach is also recommended for the screening stage of the Formal Safety Assessment [18]. A list of hazards are identified and a group of experts or stakeholders score the likelihood and consequence against set criteria to produce a risk score. Within the study area, we might score three hazards as Table 1, noting that the navigational complexity of the waterway is such that groundings are more likely, but would have lower consequences than collisions.

Table 1. Simple hazard table with 5x5 Matrix

ID	Hazard	Likelihood (1-5)	Consequence (1-5)	Risk (1-25)
1	Collision	3	3	9
2	Grounding	4	2	8
3	Allision	3	2	6

Such a method enables the inclusion of non-modelled issues [2] and may be suitable in situations where there is little quantitative data. However, such an approach has received significant criticism regarding the limitations and bias of expert prediction [37, 38] or the inherent properties of the matrices [17]. Further, only a single score is provided per hazard and therefore does not reflect the distribution of risk across the study area. As such, it is a highly simplistic method of risk assessment, but a useful means of risk evaluation [28].

2.2 Incident Rates and Analysis

By comparing the number of accidents against some unit of exposure (such as time or distance), the relative risk of incidents between locations and situations can be compared [3]. Figure 2 compares the number of accidents against the number of hours of exposure at 1nm resolution in the study area. A key challenge is the relative sparsity of accident data that prevents high-resolution output with most locations having zero incidents, which might be incorrectly interpreted as zero risk. Other criticisms include the under-reporting of accidents [14] and the assumed static relationship between accidents and traffic [30]. One method to overcome this is to calculate an accident rate and use the statistical relationship between accidents and traffic to estimate the accident rate (Figure 2). Whilst this increases coverage of the risk map, it loses the influence of spatial factors that might elevate risk in certain locations, becoming highly sensitive to traffic volume.
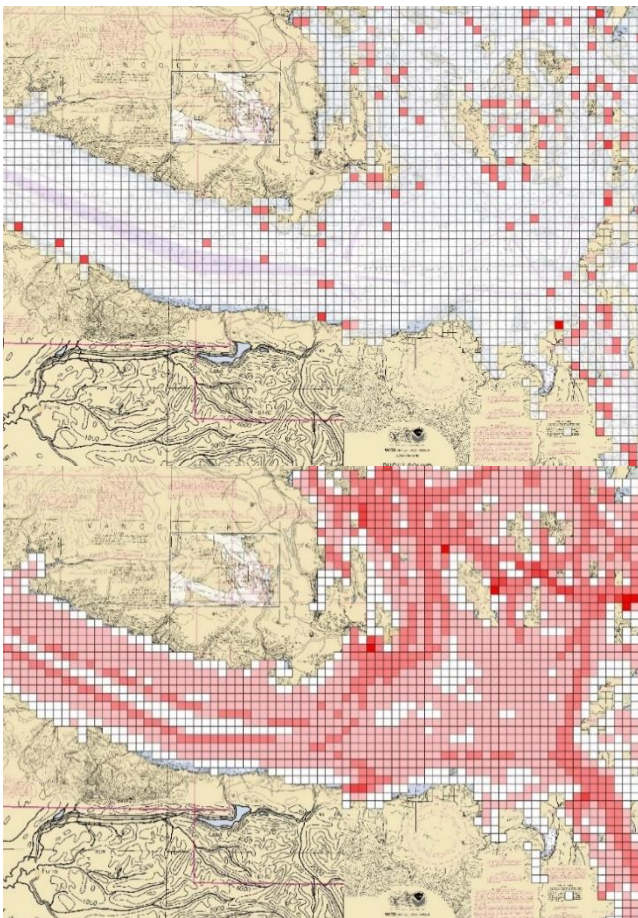

Figure 2. Calculated Incident Rates (top) and estimated incident rate (bottom).

2.3 Weighted Overlay Analysis

A further method to include the influence of other spatial factors is through a weighted overlay model. This approach can be summarised that risk is the product of the scores of likelihood factors (L) and their weightings (w) with the scores of consequence factors (C) and their respective weightings.

$$Risk = \sum_{i}^{NL} w_i^L L_i \sum_{i}^{NC} w_i^C C_i \qquad (1)$$

In Figure 3, grounding risk is estimated using a set of scoring criteria and weightings for traffic volume, proximity to shore, proximity to traffic lanes and ports. Higher risk areas are shown to the east with more complex navigation around the San Juan islands than in the middle of the Strait. Whilst such an approach enables inclusion of other risk factors and the production of high spatial resolution risk maps, the choice of weightings and factors are to some extent arbitrary, subjective and lacking in treatment of uncertainties.
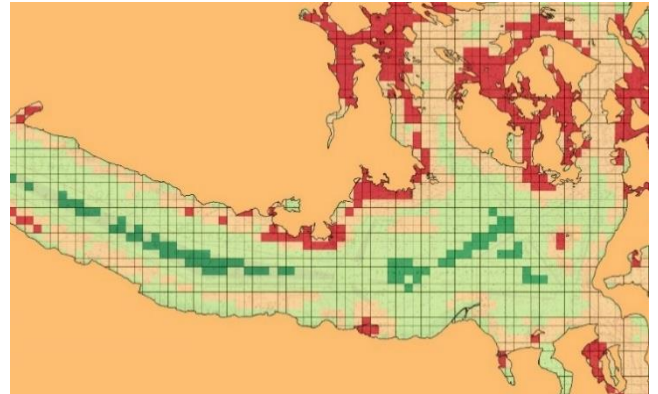

Figure 3. Weighted Overlay Analysis.

2.4 Geometric Method

A geometric method is one that aggregates vessel traffic into routes, with known distributions and frequencies, before performing mathematical functions to calculate accident candidates. Whilst variations exist [22, 27] the work of [29] has been particularly influential and has been adopted by IALA's IWRAP Risk Modelling Tool [8]. An IWRAP model was developed for the study area with the traffic legs representing the major routes and the shoreline inputted as a grounding hazard. From this, the risk of collision and grounding can be calculated by vessel type and location (Figure 4).
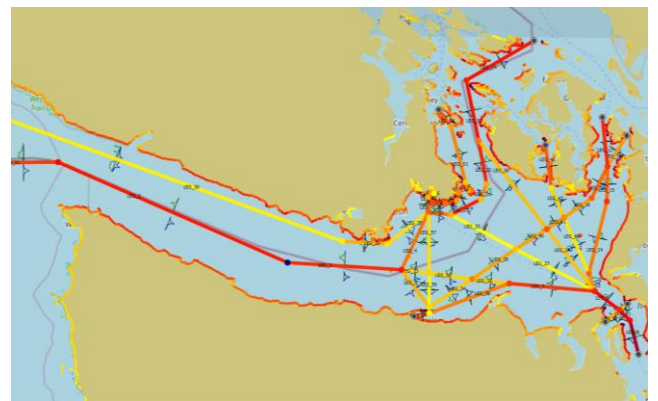

Figure 4. IWRAP Model Results.

Geometric methods have been widely discussed in the literature and have attracted numerous criticisms. Firstly, aggregation undermines the individual

behaviours of vessels, particularly where vessel transits are non-linear, such as in a pilot boarding area. Secondly, movements are averaged over 24 hours and don't reflect variations in risk throughout the day, such as where tidal heights dictate channel access, or recreational and fishing activities are diurnal. Furthermore, where the data collection is limited to short periods of time or conducted in quieter places, establishing representative traffic distributions may not be possible [24]. Fourthly, the method omits some hazard types such as drifting vessel collisions [1]. Fifthly, the choice of legs and other input parameters is subjective and depending on the expertise of the analyst [28]. Sixthly, the results are highly sensitive to the causation probability [27] which might be chosen with little evidential basis. Finally, some have questioned the underlying assumption that risk is directly related to traffic flow [25, 30].

2.5 *Domain Analysis*

Ship domain models construct a region of safe water that a master wants to keep clear of other vessels or fixed objects [12], and by measuring the frequency and types of encounters between those vessel domains, an indicative measure of collision risk is provided [6]. Whilst a multitude of domain designs have been proposed [35], we implement the model proposed by [40] that is dynamic given vessel size and speed. Figure 5 shows the frequency of domain encounters across the study area, with the majority clustered in the key ports and harbours of Victoria, Port Angeles and Anacortes, rather than in the Straits to the west.
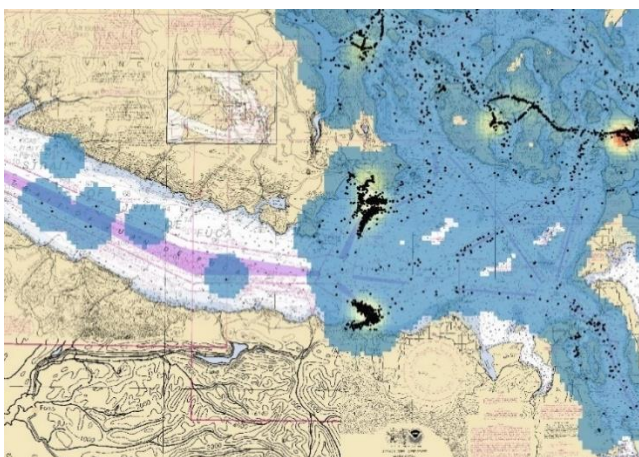

Figure 5. Domain Analysis Encounters.

Whilst the inclusion of a temporal dimension within domain models overcome some limitations of static methods, they have also received criticism. Some have questioned both the reliability and validity of such methods [11] and also the statistical relationship between encounters and collision frequency [31]. Furthermore, the interactions between vessels have been shown to vary by many other factors such as waterway geometry [16] that are not captured in existing domain models. Finally, domain models are often limited to one-vs-one collision situations, omitting the influence of other nearby vessels on developing collision situations [6].

2.6 *Bayesian Networks*

Bayesian Networks are a technique for graphically representing a joint probability distribution of a selected set of variables and many have argued that they are well suited to maritime risk analysis [13, 41]. In particular, they enable the inclusion of expert judgment, a far greater number of conditions that cannot be easily quantified and can be employed in situations where there is little historical data, such as autonomous vessels or Arctic shipping. Furthermore, the impact of risk controls can be tested by interfering with specific elements within the model [26]. Finally, uncertainties can be reflected within the model. Figure 6 shows a highly simplified Bayesian Network for predicting grounding risk, compared to others proposed in the literature [26].
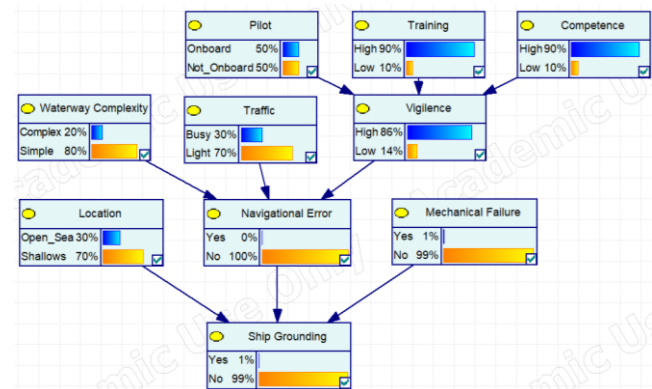

Figure 6. Simplified Bayesian Network.

Bayesian Networks are not without their criticisms [16, 41]. Firstly, the lack of data makes determining priors challenging. Secondly, the reliance on expert judgment might introduce biases [37, 38] and limits the frequency and scope of such assessments, given the time commitments necessary on experts and stakeholders.

2.7 *Summary of Conventional Methods*

The six methods described above are broad but encapsulate a significant portion of the academic and industry techniques used for maritime risk assessment. In general, the different methods consistently identify higher risk to the east of the study area in the constrained and busy waters of the San Juan Islands. Whilst we have provided a high-level introduction to each method and identified some key criticisms in each case, conventional methods of maritime risk assessment have more fundamental challenges. At a workshop of EMSA [7], it was noted that existing methods used by risk assessment projects had a high cost, used proprietary models of consultants, were time-consuming and often failed to communicate their uncertainties. Given these criticisms, it may be that alternative methods may be more suitable for maritime risk assessment.

3 MACHINE LEARNING METHODS

ML consists of numerous approaches and applications, and here we discuss how three broad

categories could be applied to maritime risk assessment: namely supervised ML, unsupervised ML and deep learning techniques.

3.1 *Supervised Machine Learning*

Predicting whether an accident will occur in a certain location or time extent can be framed as a form of supervised ML. We seek the function $y=f(x)$ that describes the relationship between a response variable y, in this case accident occurrence, with a set of explanatory factors x, such as traffic volume, depth or ship size as vectors in the form $(x_1,x_2,…,x_n)$. In order to learn this relationship, a dataset is split into a training dataset, through which the model is developed, and then tested on a set-aside dataset. The accuracy of such models can be gauged through numerous metrics, such as classification accuracy (correct predictions to incorrect predictions) or prediction value error in the case of regression.

Within the literature, there are relatively few applications of ML for maritime risk assessment [5, 19]. Where these methods have been used, it is more common for the training dataset to consist of a static list of vessels [19] or port state control inspection outcomes [33]. In such a context, a model is developed that seeks to predict whether a vessel has an accident given the characteristics of the vessel, such as age, flag state or type. However, there is much greater scope to advance these methods for spatial or real-time maritime risk assessment using vessel traffic data and spatial-temporal risk factors.
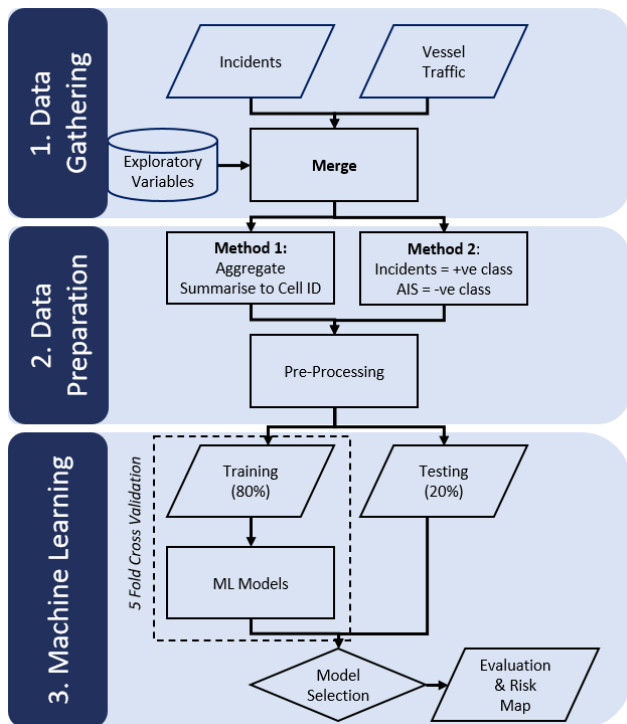
Figure 7. Maritime risk ML framework.

Figure 7 shows a framework through which such methods could be applied to predict vessel traffic risk. In stage 1, AIS, incident and other exploratory datasets are combined as input features and labels. In stage 2, the dataset is either labelled as positive or negative (accident occurred or didn't occur) as a classification problem, or the accident frequency is calculated as a regression problem by aggregating the datasets. In stage 3, the data is split into a training and testing dataset, with the ML model developed on the former and tested on the latter, before the trained model is deployed as a risk analysis tool.

In this example, the Random Forest ML algorithm is used which is an ensemble of decision trees using subsets of the training data, with the final prediction the average prediction of the individual trees. In the first case, the study area was subdivided into grid cells and the traffic volume and depth of water in each cell used as input features and the number of groundings used as the label. The dataset is split into a training and testing dataset with the ratio of 80% to 20%, which once trained, achieved an R2 of 0.58 and a Mean Squared Error of 0.004 on the testing set. Figure 8 shows the results, highlighting the waters around the San Juan Islands to the east, where traffic is concentrated into small and shallow channels, locations where most grounding have historically occurred.

Secondly, the same framework is applied but the individual vessel traffic positions and historical groundings are used as the positive and negative classes respectively. By using features such as vessel size (length and draught), depth of water, distance from shore and vessel traffic density, the probability of ship accident can be predicted. In this case, an accuracy of 98.8% is achieved on the test dataset, with 15 of the 26 groundings and 427,054 of 432,154 non-groundings correctly predicted. This indicates a high recall (0.58) but a low precision (0.003), suggesting the model can differentiate groundings, but at the expense of a number of false positives. Figure 9 shows the predicted transit risk for an Aframax tanker approaching Anacortes, with the risk of grounding being predicted to increase as it passes through the Guemes Channel.
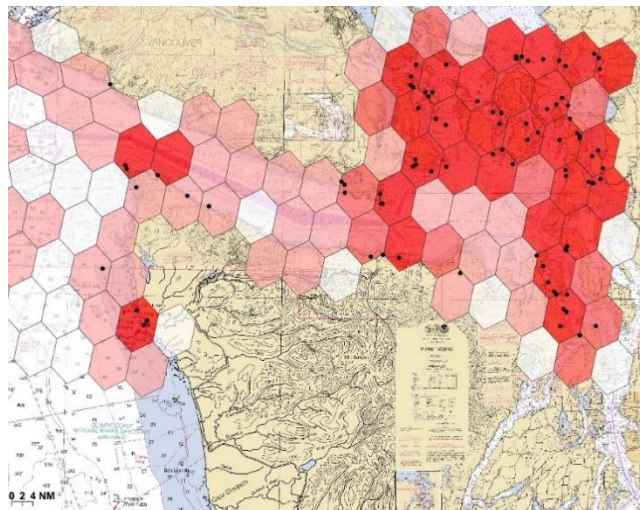
Figure 8: Area Grounding probability using Random Forest.

In this work, only a limited number of features are utilised but there is scope to improve the predictive capability through inclusion of weather, waterway geometry, ship characteristics and a plethora of other risk factors. However, the results show good potential; firstly, the results exceed the predictive power that depth or traffic volume alone could have

achieved. Secondly, in Figure 8, the model is able to predict accident risk in locations where no accidents have occurred, automatically and at far higher resolution than through human input. Thirdly, in Figure 9, even without local historical accident data, the model has learnt to distinguish the factors associated with other ship groundings and made a prediction for each individual vessel position.
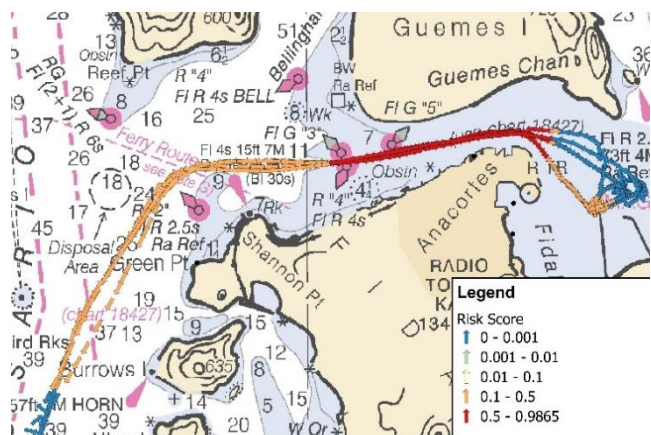

Figure 9. Transit Grounding probability using Random Forest.

3.2 *Unsupervised Methods*

In contrast to supervised methods, unsupervised methods seek to identify undetected patterns in unlabelled data, in this case identifying anomalies or clusters amongst a dataset. By learning the "normal" behaviour of vessels, abnormal transits might be interpreted as at risk. A significant body of work has developed for navigation anomaly detection [33].

Figure 10 shows the results of DBSCAN (a Density Based Clustering algorithm) used to detect positional anomalies (using latitude and longitude of vessel traffic) and behavioural anomalies (using the course and speed). Unlike K-Means, DBSCAN does not require the user to specify the number of clusters, will identify irregular clusters and will automatically identify outliers. Firstly, positional anomalies are shown, clustering the westbound and eastbound lanes, but highlighting vessels that deviate from those lanes as anomalous. Secondly, vessel behaviour is clustered, highlighting vessels which are transiting abnormally fast or slow, or making unconventional course changes. It is notable that in both cases, the vessels crossing the separation zone between the two traffic lanes are identified as anomalous.

In this example, and common with much of the literature on anomaly detection, is the degree to which an anomalous transit is inherently riskier. For example, either transiting slower or departing the traffic lane could be a response to a developing hazardous situation, and therefore would be safer than a "normal transit". Unsupervised clustering can be expanded to measure risk response of vessels specifically to a hazard, such as the actions taken to avoid the paths of hurricanes [32].
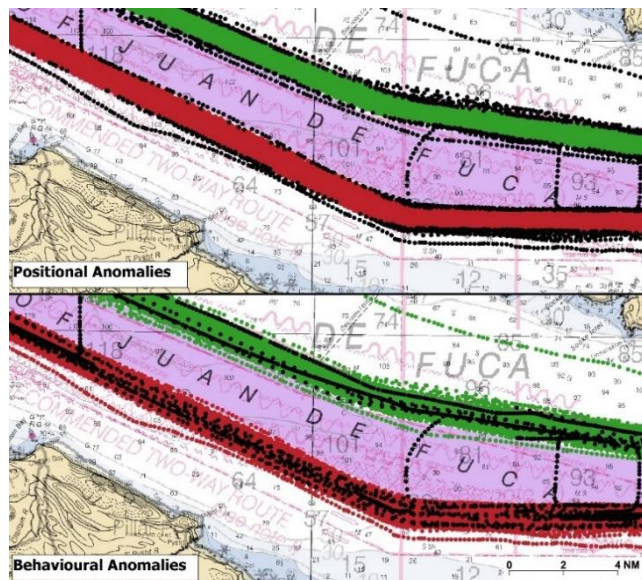

Figure 10. DBSCAN Anomaly detection.

3.3 *Opportunities in Deep Learning: CNNs and RNNs*

Convolutional Neural Networks (CNNs) are a type of deep learning neural network that has achieved prominence in image-based feature extraction or classification, including within the maritime domain specifically as a means for autonomous vessel navigation [20]. Unlike conventional neural networks, CNNs can capture the spatial and temporal dependencies of the dataset. CNNs consist of three types of layers: convolutional layers use learnable filters that move across the width and height of the input layer to extract high level features. Multiple convolutions may be used to extract successively more intricate representations of features. The pooling layer then reduces the convolution layer through downsampling such as extracting the maximum or average value. Finally, a regular neural network fully connected layer takes the flattened output of the previous layers and applies weights to learn the output labels.

A key advantage of CNNs are the ability to consider the relationships among neighbouring spatial datasets, an ability which conventional methods omit. However, limitations include the requirement for far greater computational resources, complexity in defining the model architecture and requirement for significant volumes of training data. Notably, the complex structure of CNNs makes them liable to overfitting with limited training data. As a result, the application of CNNs in maritime risk assessment is relatively unexplored compared to other disciplines. One example is the study by [42] who trained a CNN on qualitatively assessed traffic pictures to predict the risk of an accident on unseen traffic situations.

Recurrent Neural Networks (RNNs) are a class of neural networks with loops which allow previous outputs to be used as inputs, making them especially useful in timeseries data and the fields of natural language processing. Long Short-Term Memory networks (LSTMs address the long-term dependency problems of RNNs, enabling better retention of information over long periods through the use of gates. Most commonly within the maritime domain,

these methods are used for predicting the future path of a vessel based on its past behaviour (see for example [10]). By combining other features and more complex architectures, applications could be diverse, such as early warning of groundings using route prediction [34]. This makes them particularly useful for dynamic and real-time assessments. However, RNNs further increase the complexity, data requirements and technical expertise required to implement such methods and as a result their application is limited.

4 CONCLUSIONS AND PROPOSED CRITERIA

In this paper, a review of six dominant conventional methods for maritime risk assessment are provided and four ML methods are introduced. Some have proposed that ML methods have inherent advantages over conventional methods as they make no assumptions between dependent and exploratory variables [19], particularly relevant given that maritime accident causation is a highly complex issue with numerous interacting factors [42]. Furthermore, such techniques are better able to leverage multi-dimensional datasets [19], particularly through the combination of vessel traffic, accident and other datasets, which is being seen as an important emerging area of research [21]. Others have noted how automated risk methods can provide greater decision-support tools to waterway managers [5], which would not be possible with conventional methods which are laborious to set-up [7]. In this paper, a high-level implementation of several of these techniques is presented, demonstrating the potential strengths of their application.

Table 2: Proposed Criteria.

ID	Criteria	Description
1	Competency Req.	Whether the method can be implemented by non-technical or technical personnel.
2	Adoption Level	How widespread this method is within industry/academia.
3	Computational Req.	To what extent specialist software or hardware is required to calculate the results.
4	Transparency	Whether the model is black-box or has clear inputs, methods and outputs.
5	Data Req.	Volume and types of data required as inputs into the model.
6	Subjectivity	Relative ratio of expert/qualitative and quantitative inputs.
7	Spatial Representation	The spatial scale or resolution of study, regional to localised outputs.
8	Uncertainty	Degree to which uncertainty are identified or treated in the model.
9	Suitability for Strategic Assessment	How suitable the methods are for spatial risk modelling – risk between areas.
10	Suitability for Dynamic Assessment	How suitable the methods are for real-time risk modelling – risk between ship transits.

Work to develop ML methods is ongoing by many authors, but in order to demonstrate whether such techniques are more advantageous than existing techniques requires criteria against which to compare model properties. In [28], a summary matrix of model properties is introduced including applicability, resource requirements, skill requirements and whether it is quantitative or qualitative. We propose that other criteria should be included in any evaluation of risk models, and as such propose Table 2. It is notable that we have not included criteria that represent the validity or accuracy of the model results, given that different models will inevitably provide differences in risk scores [11].

For example, if we compare a risk matrix and a CNN, we can see that given the former can be done by hand or in excel, it requires significantly less technical skill, computational power or input data. Furthermore, risk matrices are more widely adopted and if expert input is properly recorded, more transparent than a neural network. However, a CNN being largely data-driven can achieve far higher resolution outputs that might make it more suitable for performing strategic assessments.

It is proposed that through implementing each of the aforementioned techniques within a defined study area and scoring the models against the criteria presented in Table 2, the proposed benefits of ML techniques over conventional methods could be identified, and further work is being undertaken to realise this.

REFERENCES

1. Altan, Y.C.: Collision diameter for maritime accidents considering the drifting of vessels. Ocean Engineering. 187, 106158 (2019). https://doi.org/10.1016/j.oceaneng.2019.106158.
2. Aven, T., Zio, E.: Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliability Engineering & System Safety. 96, 1, 64–74 (2011). https://doi.org/10.1016/j.ress.2010.06.001.
3. Bye, R.J., Almklov, P.G.: Normalization of maritime accident data using AIS. Marine Policy. 109, 103675 (2019). https://doi.org/10.1016/j.marpol.2019.103675.
4. Chen, P., Huang, Y., Mou, J., van Gelder, P.H.A.J.M.: Probabilistic risk analysis for ship-ship collision: State-of-the-art. Safety Science. 117, 108–122 (2019). https://doi.org/10.1016/j.ssci.2019.04.014.
5. Dorsey, L.C., Wang, B., Grabowski, M., Merrick, J., Harrald, J.R.: Self healing databases for predictive risk analytics in safety-critical systems. Journal of Loss Prevention in the Process Industries. 63, 104014 (2020). https://doi.org/10.1016/j.jlp.2019.104014.
6. Du, L., Goerlandt, F., Kujala, P.: Review and analysis of methods for assessing maritime waterway risk based on non-accident critical events detected from AIS data. Reliability Engineering & System Safety. 200, 106933 (2020). https://doi.org/10.1016/j.ress.2020.106933.
7. EMSA: Joint Workshop on Risk Assessment and Response Planning in Europe. , London (2018).
8. Friis-Hansen, P.: IWRAP MK II: Working Document: Basic Modelling Principles for Prediction of Collision and Grounding Frequencies, https://www.iala-

aism.org/wiki/iwrap/images/2/2b/IWRAP_Theory.pdf, last accessed 2020/12/15.

9. Fujino, I., Claramunt, C., Boudraa, A.-O.: Extracting Courses of Vessels from AIS Data and Real-Time Warning Against Off-Course. In: Proceedings of the 2nd International Conference on Big Data Research. pp. 62–69 Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3291801.3291823.

10. Gao, M., Shi, G., Li, S.: Online Prediction of Ship Behavior with Automatic Identification System Sensor Data Using Bidirectional Long Short-Term Memory Recurrent Neural Network. Sensors. 18, 12, (2018). https://doi.org/10.3390/s18124211.

11. Goerlandt, F., Kujala, P.: On the reliability and validity of ship–ship collision risk analysis in light of different perspectives on risk. Safety Science. 62, 348–365 (2014). https://doi.org/10.1016/j.ssci.2013.09.010.

12. Goodwin, E.M.: A Statistical Study of Ship Domains. Journal of Navigation. 28, 3, 328–344 (1975). https://doi.org/10.1017/S0373463300041230.

13. Hänninen, M.: Bayesian networks for maritime traffic accident prevention: Benefits and challenges. Accident Analysis & Prevention. 73, 305–312 (2014). https://doi.org/10.1016/j.aap.2014.09.017.

14. Hassel, M., Asbjørnslett, B.E., Hole, L.P.: Underreporting of maritime accidents to vessel accident databases. Accident Analysis & Prevention. 43, 6, 2053–2063 (2011). https://doi.org/10.1016/j.aap.2011.05.027.

15. Hegde, J., Rokseth, B.: Applications of machine learning methods for engineering risk assessment – A review. Safety Science. 122, 104492 (2020). https://doi.org/10.1016/j.ssci.2019.09.015.

16. Hörteborn, A., Ringsberg, J.W., Svanberg, M., Holm, H.: A Revisit of the Definition of the Ship Domain based on AIS Analysis. Journal of Navigation. 72, 3, 777–794 (2019). https://doi.org/10.1017/S0373463318000978.

17. Hubbard, D.W.: The Failure of Risk Management: Why It's Broken and How to Fix It. Wiley (2020).

18. IMO: Revised Guidelines for Formal Safety Assessment (FSA) For Use in the IMO Rule-Making Process. (2018).

19. Jin, M., Shi, W., Yuen, K.F., Xiao, Y., Li, K.X.: Oil tanker risks on the marine environment: An empirical study and policy implications. Marine Policy. 108, 103655 (2019). https://doi.org/10.1016/j.marpol.2019.103655.

20. Kim, H., Koo, J., Kim, D., Park, B., Jo, Y., Myung, H., Lee, D.: Vision-Based Real-Time Obstacle Segmentation Algorithm for Autonomous Surface Vehicle. IEEE Access. 7, 179420–179428 (2019). https://doi.org/10.1109/ACCESS.2019.2959312.

21. Kulkarni, K., Goerlandt, F., Li, J., Banda, O.V., Kujala, P.: Preventing shipping accidents: Past, present, and future of waterway risk management with Baltic Sea focus. Safety Science. 129, 104798 (2020). https://doi.org/10.1016/j.ssci.2020.104798.

22. Li, S., Meng, Q., Qu, X.: An overview of maritime waterway quantitative risk assessment models. Risk Anal. 32, 3, 496–512 (2012). https://doi.org/10.1111/j.1539-6924.2011.01697.x.

23. Lim, G.J., Cho, J., Bora, S., Biobaku, T., Parsaei, H.: Models and computational algorithms for maritime risk analysis: a review. Annals of Operations Research. 271, 2, 765–786 (2018). https://doi.org/10.1007/s10479-018-2768-4.

24. Liu, Z., Wu, Z., Zheng, Z.: A novel framework for regional collision risk identification based on AIS data. Applied Ocean Research. 89, 261–272 (2019). https://doi.org/10.1016/j.apor.2019.05.020.

25. Mazaheri, A., Montewka, J., Kotilainen, P., Edvard Sormunen, O.-V., Kujala, P.: Assessing Grounding Frequency using Ship Traffic and Waterway Complexity. Journal of Navigation. 68, 1, 89–106 (2015). https://doi.org/10.1017/S0373463314000502.

26. Mazaheri, A., Montewka, J., Kujala, P.: Towards an evidence-based probabilistic risk model for ship-grounding accidents. Safety Science. 86, 195–210 (2016). https://doi.org/10.1016/j.ssci.2016.03.002.

27. Mazaheri, A., Ylitalo, J.: Comments on Geometrical Modeling of Ship Grounding. Presented at the 5th International Conference on Collision and Grounding of Ships (2010). https://doi.org/10.13140/2.1.3359.3284.

28. OpenRisk: OpenRisk Guideline for Regional Risk Management to Improve European Pollution Preparedness and Response at Sea, https://portal.helcom.fi/meetings/OPENRISK%20WS%203-2018-527/Related%20Information/OPENRISK%20Guide_Final_13_6_18.pdf, last accessed 2020/12/15.

29. Pedersen, P.: Collision and grounding mechanics. The Danish Society of Naval Architects and Marine Engineers. 125–157 (1995).

30. Rawson, A.: An Analysis of Vessel Traffic Flow Before and After the Grounding of the MV Rena, 2011. In: Weintrit, A. (ed.) Marine Navigation. pp. 203–209 CRC Press (2017). https://doi.org/10.1201/9781315099132-24.

31. Rawson, A., Brito, M.: A critique of the use of domain analysis for spatial collision risk assessment. Ocean Engineering. 219, 108259 (2021). https://doi.org/10.1016/j.oceaneng.2020.108259.

32. Rawson, A.D., Brito, M.: Modelling of ship navigation in extreme weather events using machine learning. In: Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment And Management Conference. Research Publlishing (2020).

33. Riveiro, M., Pallotta, G., Vespe, M.: Maritime anomaly detection: A review. WIREs Data Mining and Knowledge Discovery. 8, 5, e1266 (2018). https://doi.org/10.1002/widm.1266.

34. Suo, Y., Chen, W., Claramunt, C., Yang, S.: A Ship Trajectory Prediction Framework Based on a Recurrent Neural Network. Sensors. 20, 18, (2020). https://doi.org/10.3390/s20185133.

35. Szlapczynski, R., Szlapczynska, J.: Review of ship safety domains: Models and applications. Ocean Engineering. 145, 277–289 (2017). https://doi.org/10.1016/j.oceaneng.2017.09.020.

36. Tang, L., Tang, Y., Zhang, K., Du, L., Wang, M.: Prediction of Grades of Ship Collision Accidents Based on Random Forests and Bayesian Networks. In: 2019 5th International Conference on Transportation Information and Safety (ICTIS). pp. 1377–1381 (2019). https://doi.org/10.1109/ICTIS.2019.8883590.

37. Tetlock, P.E.: Expert Political Judgment: How Good Is It? How Can We Know? Princeton University Press (2005).

38. Tversky, A., Kahneman, D.: Judgment under Uncertainty: Heuristics and Biases. Science. 185, 4157, 1124–1131 (1974). https://doi.org/10.1126/science.185.4157.1124.

39. Van Dorp, J., Merrick, J.: VTRA 2010 Final Report. George Washington University (2014).

40. Wang, N.: An Intelligent Spatial Collision Risk Based on the Quaternion Ship Domain. Journal of Navigation. 63, 4, 733–749 (2010). https://doi.org/10.1017/S0373463310000202.

41. Zhang, G., Thai, V.V.: Expert elicitation and Bayesian Network modeling for shipping accidents: A literature review. Safety Science. 87, 53–62 (2016). https://doi.org/10.1016/j.ssci.2016.03.019.

42. Zhang, W., Feng, X., Goerlandt, F., Liu, Q.: Towards a Convolutional Neural Network model for classifying regional ship collision risk levels for waterway risk analysis. Reliability Engineering & System Safety. 204, 107127 (2020). https://doi.org/10.1016/j.ress.2020.107127.