

Supplementing Fault Trees calculations with neural networks

Dr Victor Bolbot*

Maritime Safety Research Centre, NAOME, University of Strathclyde, United Kingdom. E-mail: victor.bolbot@strath.ac.uk

Dr Christos Gkerekos*

Maritime Safety Research Centre, NAOME, University of Strathclyde, United Kingdom. E-mail: christos.gkerekos@strath.ac.uk

Prof Gerasimos Theotokatos

Maritime Safety Research Centre, NAOME, University of Strathclyde, United Kingdom. E-mail: gerasimos.theotokatos@strath.ac.uk

The use of artificial intelligence algorithms is rapidly gaining ground in engineering applications, including safety engineering. In this paper, we investigate the possibility of using neural networks to supplement fault trees in the safety analysis for the estimation of reliability and importance metrics. For this aim, we employ data from an existing fault tree that models cruise ships blackouts to train a neural network that uses base-event probabilities as input and outputs the estimated top-event probability/frequency. This is done to reduce computational time, as the fault tree model has an extensive number of basic events and is thus computationally demanding. The information that is used as input to the Fault Tree is randomly sampled from a Sobol sequence and is used to estimate the top event probability. The resulting data cloud that corresponds to the fault tree's input-output pairs, is used to train the neural network. The two models, i.e. the probabilistic and the neural network model, are compared to each other in terms of accuracy and computational cost correlated with the number of sampling points that is used. The Fault Tree is developed in Matlab/Simulink and the neural network in Python. For case where the Neural Network is trained using 10,000 points, a 350 times decrease in computational cost is observed compared to the fault tree model, while the mean absolute percentage error (MAPE) remains at under 15%. Based on the results, recommendations for the application and future improvement of the artificial intelligent algorithms in the specific context are made.

Keywords: Neural Networks, Fault Tree, Blackouts, Machine Learning, Safety Engineering, Metamodeling

1. Introduction

We live in a world where exceedingly complex systems are being constantly developed. Cyber-physical systems (CPSs) constitute an example of such complex systems, effectively integrating the cyber with the physical part (Bolbot et al. 2019). The safety of CPSs is ensured using a number of approaches such as safety analysis techniques, verification and testing techniques, real-time monitoring systems, operational procedures, etc (Bolbot et al. 2019).

During the safety analysis and real-time safety monitoring, it is important to estimate safety metrics for the CPSs to identify the critical components and suggest the relevant safety control actions (Papadopoulos and McDermid 2001). Under ideal conditions, the CPSs safety analysis should be completed in a very short time to allow swift safety reconfiguration during real-time, e.g. change of course, or reconfiguration to another component and smooth design decisions. However, the inherent complexity of the CPSs introduces computational challenges for the estimation of the safety metrics (Bolbot et al. 2019, Pereira and Thomas 2020). This can constitute a barrier to the wider adoption of

the CPSs. Therefore, there is an immediate need to consider ways to overcome this problem.

The development of mathematics have equipped the safety engineers with new tools that can be used for ensuring the safety of CPSs such as Artificial Intelligence algorithms, Machine Learning techniques, and so on (Hegde and Rokseth 2020, Hughes, Van Gulijk, and El Rashidy 2019, Schwarz et al. 2020). These new mathematical tools have demonstrated good properties with respect to calculation accuracy as well as speed performance, provided they can have adequate data. Therefore, the aim of this paper is to investigate the potential use of Machine Learning techniques for supplementing the calculations implemented during safety analysis. For the sake of the analysis a Diesel-Electric Propulsion system and relevant developed Fault Tree is being considered.

The paper is organised as follows. In the next section the followed methodology is elaborated. Then the information about the input and the investigated system is provided. After that, the results of substituting the Fault Tree are being presented and discussed. Finally, the main findings of the study are being summarised in the conclusion section.

2. Methodology

The overview of the methodology

The overview of the followed methodology is provided in Fig. 1. For the analysis, a Diesel-Electric Propulsion (DEP) system and its frequency of blackout is being considered. An already developed Fault Tree of blackout for Diesel-Electric Propulsion system is being used for generation of the sampled measures during step 1. Then the sampled estimates are being used to develop a Neural Network metamodel of the fault tree in step 2. This procedure is repeated using different numbers of samples. Once the results are assembled, they are analysed and relevant conclusions are being made regarding the methodology and relevant future steps.

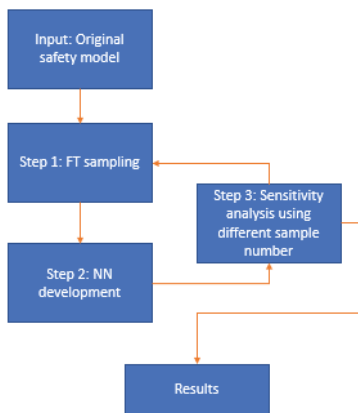


Fig. 1 Methodology overview.

Input description

For the methodology, a safety model of a system with its basic events is required. It can be a Fault Tree (FT), a Bayesian Network (ISO 2009) or Boolean-logic Driven Markovian processes (Bouissou and Bon 2003), or any other complex safety model. In the present study, the Fault Tree has been employed for predicting the probability of the Diesel-Electric propulsion system blackout. This is briefly described in Section 3.

Step 1: FT Sampling

During the first Step, the data points which are required for the development of the Neural Networks are generated. A plethora of methods could be employed herein for the generation of the datasets, such as random sampling (Metropolis et al. 1953), Latin hypercube sampling (McKay, Beckman, and Conover 1979), or Sobol sampling

(Sobol' 1967). The preference is given to Sobol sampling, as this type of sampling has demonstrated some good results in problems of a similar nature (Burhenne, Jacob, and Henze 2011, Kucherenko, Albrecht, and Saltelli 2015, Qian and Mahdi 2020, Bolbot and Theotokatos 2021).

Hence, the failure rates λ which are used as input to the FT are being sampled according to the following formula eq.(1):

$$\lambda_{i,j} = 10^{-3-s_{i,j}6} \quad \text{eq.(1)}$$

Where $s_{i,j}$ is Sobol sample i for component j . Since $s_{i,j}$ varies uniformly between 0 and 1, the $\lambda_{j,s}$ follows log uniform distribution in the range between 10^{-3} per hour and 10^{-9} per hour. The generated $\lambda_{i,j}$ are used as input to the FT and the frequency of blackout is estimated as output. The inputs ($\lambda_{i,j}$) and output (blackout frequency) are used then in the next section for the development of Neural Networks.

Step 2: NN Development

Artificial Neural Networks (ANN) are computing systems that are based on the same working principles as biological nervous systems. They are based on an interconnected group of connected units (neurons) where each connection between these units transmits a signal from one to another, as long as the linear combination of the inputs exceeds some threshold (Russell and Norvig 2002). The receiving unit can further process the signal and consequently pass it on to the next unit. In the case of regression, the output layer of the NN, i.e. its last layer, outputs the value predicted by the network given its specific inputs (Fig. 2).

Key parameters that affect the computation needs and the performance of ANNs are the number of hidden layers and the number of neurons per hidden layers. Other parameters include the, per layer, dropout rate and the learning rate.

Dropout is applied in ANNs with more than one hidden layers to reduce overfitting of models and improve their robustness. This is done by randomly dropping neurons during training based on a probability dictated by the dropout rate. The ANN of Fig. 2 is depicted in Fig. 3, including the effects of dropout.

The tuning is applied through a variation of the HyperBand algorithm (Li et al. 2017), as provided through the TensorFlow/Keras platform. Key parameters include the maximum number of epochs to train the model and the maximum number of full algorithm iterations.

In this case, the ANN input is taken from step 1 of the methodology and the search space for optimal hyperparameters included the information presented in Table 1.

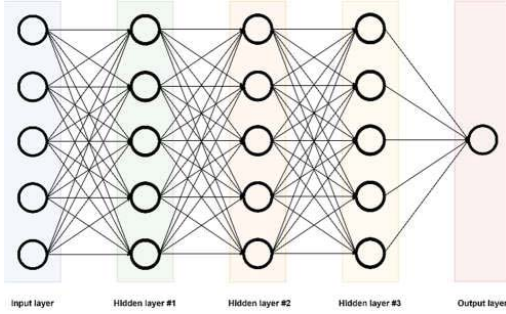


Fig. 2 Typical 3 hidden layer ANN architecture

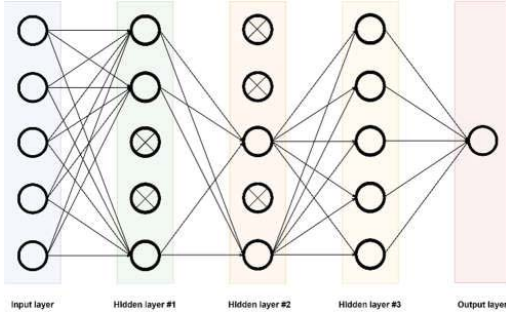


Fig. 3 Typical 3 hidden layer ANN architecture with dropout applied. Dropout rate is 40% after the first hidden layer, 60% after the second hidden layer, and 0% after the third hidden layer.

Table 1 The hyperparameters information.

Hidden layers	4 - 8
Nodes per hidden layer	40 - 157
Dropout rate per layer	0 - 0.9
Learning rate	1E-4 - 1E-2

Step 3: Sensitivity analysis using different sample numbers

The FT sampling (step 1) and ANN development (step 2) are repeated for 100, 1000, 10000 totally samples of failure rates. The estimations of the blackout frequency according to the Fault Tree and Neural networks are then compared using a number of metrics such as Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE) and Mean Squared Logarithmic Error (MSLE). The formulas of these error metrics are presented in eq. (2) to (4).

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\% \quad \text{eq.(2)}$$

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \text{eq.(3)}$$

$$MSLE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (\ln(1 + y_i) - \ln(1 + \hat{y}_i))^2 \quad \text{eq.(4)}$$

Conclusions and findings of the study are derived based on the generated observations.

3. Selected case study and original Fault Tree

Diesel-Electric Propulsion (DEP) systems have been widely used for propulsion and power generation of a number of ships including LNG carriers, icebreakers, drilling units, naval vessels and cruise ships (Hansen and Wendt 2015). The DEP can be viewed as a complex system consisting of a number of subsystems including physical components, hardware and software, which interact with each other to ensure the continuous power generation and distribution, covering the ship power demand (Ådnanes 2003, Rokseth, Utne, and Vinnem 2017) and satisfying the definition of CPSs.

In this analysis, a DEP of cruise ship vessel as presented in (Bolbot, Trivyza, et al. 2020, Bolbot et al. 2021) has been employed. The DEP system includes 6 Diesel Generators (DGs), 3 Azipod propulsors, 3 Switchboards, 4 Bow Thrusters, 2 separate engine rooms as shown in Fig. 4. More details on the investigated system can be found in (Bolbot, Trivyza, et al. 2020, Bolbot et al. 2021).

The system blackout frequency is estimated using Combinatory Approach to Safety Analysis (CASA). The CASA involves integration of System-Theoretic Process Analysis (STPA), Event Sequence Identification and Fault Tree Analysis for the development of an enhanced Fault Tree. More details on CASA can be found in (Bolbot, Theotokatos, et al. 2020, Bolbot et al. 2021). CASA employs a set of refinement rules and probabilistic calculus for the estimation of top event failure rate.

The developed Fault Tree used 142 failure rates as input to the analysis alongside other design and failure parameters. The Fault Tree is too large to be presented in a paper format and required approximately 7 seconds on a conventional desktop computer to provide a single estimation of blackout failure rate in the MATLAB/Simulink environment. The MATLAB/Simulink environment has been used, as it the necessary flexibility for the development of the Fault Tree and the research activities. This computational time is adequate for estimating only a single metric, however, the implementation of criticality analysis would require an additional 142x7=17 min. More details on the developed Fault Tree can be found in (Bolbot, Theotokatos, et al. 2020, Bolbot et al. 2021).

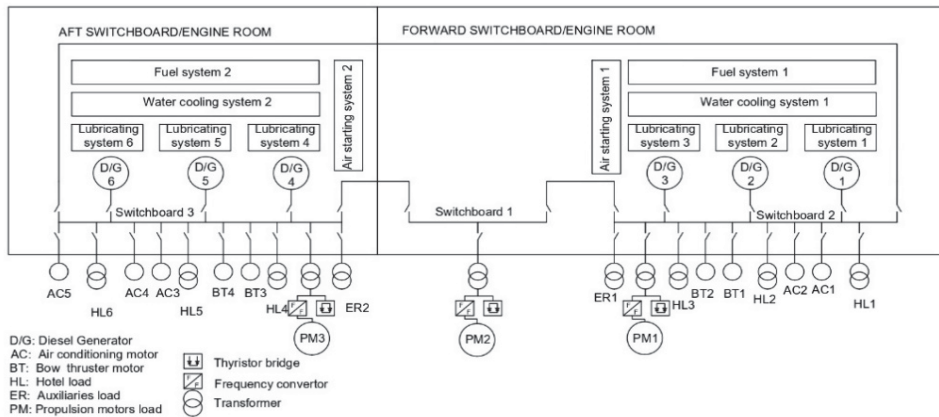


Fig. 4 Investigated system layout

4. Results and discussion

The results of the analysis for the specific case study are presented in Table 2 and in Fig. 5 to 8. As previously mentioned, three cases were evaluated, that of 100, 1000 and 10,000 data points. With 100 points used for training the ANN model accuracy is not sufficient for the needs of the problem, with MAPE reaching almost 1500%. The use of 1000 points, leads to significant improvement in accuracy (MAPE=35%), but only with 10000 it can be considered as tolerable (15%).

The ANN struggles to capture more accurately the FT with more data points added, as the spread of value according to the FT is quite large, extending over a region where the smallest and the largest value have difference of 6 orders of magnitude (Fig. 7 and Figure 8). As the considered region is extended, more datapoints are considered. Even if the outliers are excluded, then the overall accuracy does not change significantly (Figure 8). Yet, it is highly unlikely that values of failure rate resulting in outlier blackout frequency of 0.5 or 100 per hour will ever be encountered in design or during operation.

As it can be observed in Table 2, as the number of data points were considered, the more time was required to develop the ANN model. Still, in all cases, the estimations for a single blackout frequency were obtained at a negligible time cost, of about 2ms, far less than 7 seconds required in the initial model on a typical personal computer. Therefore, it can be argued that the use of ANN allowed us to develop a model of sufficient accuracy, but demanding far less computation resources.

A limitation of the current study is that we considered only the variation in the failure rates that are being used as input. The potential changes in FT structures and sensitivity to

them were not examined. It is anticipated though that during the Fault Tree changes in response to changes in the system, the relevance of the developed ANN will be dependent on the importance of the imposed changes. If the changes do not influence the critical components, then the FT can be reused, if not, it needs to be updated. Potentially, this issue can be faced by training a set of FTs which correspond to specific system configurations.

Whilst the finally achieved accuracy is deemed as tolerable, further improvements are deemed as necessary for the ANN to achieve accuracy required for an application in a safety critical system. A potential future research could focus on identification of clusters of input values to the ANN which provide the estimate according to the set target accuracy. In this way, the use of hybrid computational methods would be pursued, where if the input value belongs to the cluster of input values resulting in accurate ANN estimation, then ANN is used. For values outside these clusters, then the actual Fault Tree could be employed. This will inevitably lead to higher computational costs, but still would result in computation cost reduction. An alternative could be a set of ANN, each of which achieves an accurate estimate of the values in the pre-set clusters. All these are constitute suggestion for future research.

Table 2 The estimation of error metrics with varying the sample size.

	100 points model	1,000 points model	10,000 points model
Training time (10 runs, averaged)	77 seconds ≈ 1min	261 seconds ≈ 4min	3443 seconds ≈ 1 hour
RMSE	0.097	0.030	0.017
MAPE	1448.6%	35.26%	14.55%
MSLE	0.00748	0.00068	0.00015

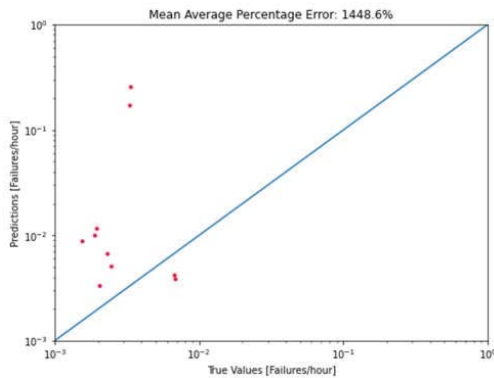


Fig. 5 Results for 100 data points (90% training - 10% testing)

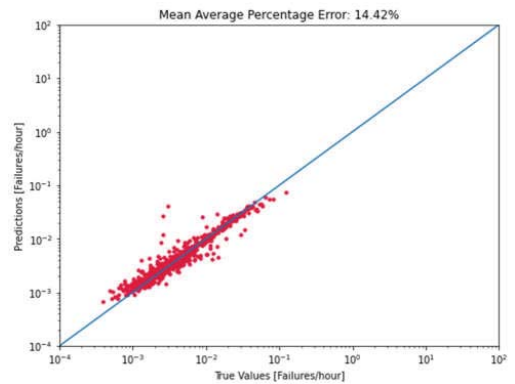


Fig. 8 Results for 10,000 data points (90% training - 10% testing) with polishing outliers

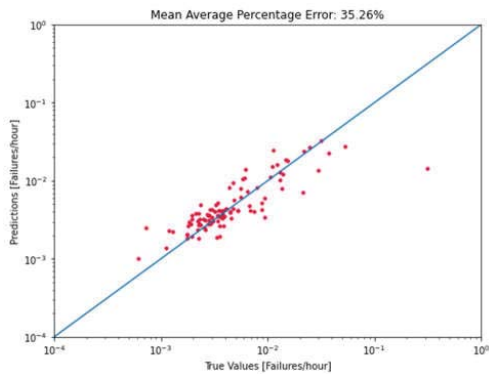


Fig. 6 Results for 1,000 data points (90% training - 10% testing)

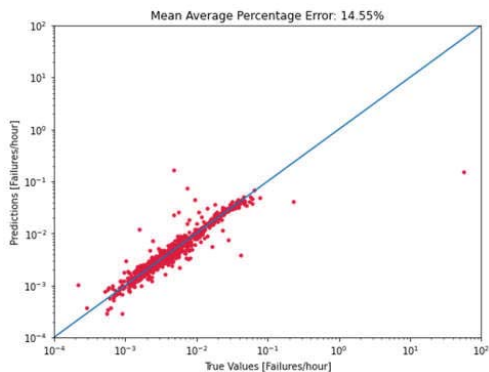


Fig. 7 Results for 10,000 data points (90% training - 10% testing)

5. Conclusions

In this paper a methodology for developing an ANN in the place of a safety model, in specific a Fault Tree, has been suggested. The methodology was applied to a Fault Tree of blackout in a DEP system.

The main findings of the study are as follows

- The developed ANN demonstrated tolerable accuracy when 10,000 training points were involved for 142 failure rates where the minimum and maximum value had a difference of six orders of magnitude.
- The use of ANN allowed to reduce significantly the computational cost of the initial Fault Tree.
- Further ways to ensure better accuracy of the ANN need to be investigated.

A future research could investigate the robustness of the developed ANN model in response to system changes and on how to further improve the model accuracy.

Acknowledgement

The research here has been funded through the AUTOSHIP project (AUTOSHIP 2019) which is funded by the European Union's Horizon 2020 research and innovation programme under agreement No 815012. The authors also greatly acknowledge the funding from DNV and RCCL for the MSRC establishment and operation. The opinions expressed herein are those of the authors and should not be construed to reflect the views of DNV, RCCL, AUTOSHIP partners or the acknowledged individuals and their associated organisations.

References

- Ådnanes, Alf Kåre. 2003. *Maritime electrical installations and diesel electric propulsion*.
- AUTOSHIP. 2019. "Autonomous Shipping Initiative for European Waters." <https://www.autoship-project.eu/>.

- Bolbot, Victor, and Gerasimos Theotokatos. 2021. "Automatically generating collision scenarios for testing ship collision avoidance system using sampling techniques." LEAC.
- Bolbot, Victor, Gerasimos Theotokatos, Evangelos Boulougouris, George Psaros, and Rainer Hamann. 2020. "A novel method for safety analysis of Cyber-Physical Systems - Application to a ship exhaust gas scrubber system." *Safety* 6 (2):26.
- Bolbot, Victor, Gerasimos Theotokatos, Evangelos Boulougouris, George Psaros, and Rainer Hamann. 2021. "A Combinatorial Safety Analysis of Cruise Ship Diesel-Electric Propulsion Plant Blackout." *Safety* 7 (2). doi: 10.3390/safety7020038.
- Bolbot, Victor, Gerasimos Theotokatos, Luminita Manuela Bujorianu, Evangelos Boulougouris, and Dracos Vassalos. 2019. "Vulnerabilities and safety assurance methods in Cyber-Physical Systems: A comprehensive review." *Reliability Engineering & System Safety* 182:179-193. doi: <https://doi.org/10.1016/j.res.2018.09.004>.
- Bolbot, Victor, Nikoletta L. Trivyza, Gerasimos Theotokatos, Evangelos Boulougouris, Athanasios Rentizelas, and Dracos Vassalos. 2020. "Cruise ships power plant optimisation and comparative analysis." *Energy* 196 (MOSES special issue):117061. doi: <https://doi.org/10.1016/j.energy.2020.117061>.
- Bouissou, Marc, and Jean-Louis Bon. 2003. "A new formalism that combines advantages of fault-trees and Markov models: Boolean logic driven Markov processes." *Reliability Engineering & System Safety* 82 (2):149-163. doi: [http://dx.doi.org/10.1016/S0951-8320\(03\)00143-1](http://dx.doi.org/10.1016/S0951-8320(03)00143-1).
- Burhenne, Sebastian, Dirk Jacob, and Gregor P Henze. 2011. "Sampling based on Sobol' sequences for Monte Carlo techniques applied to building simulations." Proc. Int. Conf. Build. Simulat.
- Hansen, J. F., and F. Wendt. 2015. "History and State of the Art in Commercial Electric Ship Propulsion, Integrated Power Systems, and Future Trends." *Proceedings of the IEEE* 103 (12):2229-2242. doi: 10.1109/JPROC.2015.2458990.
- Hegde, Jeevith, and Borge Rokseth. 2020. "Applications of machine learning methods for engineering risk assessment—A review." *Safety science* 122:104492.
- Hughes, Peter, Coen Van Gulijk, and Rawia El Rashidy. 2019. "An interactive machine-learning method to obtain safety information from free text." 29th European Safety and Reliability Conference.
- ISO. 2009. Risk management — Risk assessment techniques. In *ISO 31010*. Switzerland, Geneva: International Organization for Standardization.
- Kucherenko, Sergei, Daniel Albrecht, and Andrea Saltelli. 2015. "Exploring multi-dimensional spaces: A comparison of Latin hypercube and quasi Monte Carlo sampling techniques." *arXiv preprint arXiv:1505.02350*.
- Li, Lisha, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. 2017. "Hyperband: A novel bandit-based approach to hyperparameter optimization." *The Journal of Machine Learning Research* 18 (1):6765-6816.
- McKay, M. D., R. J. Beckman, and W. J. Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* 21 (2):239-245. doi: 10.2307/1268522.
- Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. "Equation of State Calculations by Fast Computing Machines." *The Journal of Chemical Physics* 21 (6):1087-1092. doi: 10.1063/1.1699114.
- Papadopoulos, Yiannis, and JA McDermid. 2001. "Automated safety monitoring: A review and classification of methods." *International journal of COMADEM* 4 (4):14-32.
- Pereira, Ana, and Carsten Thomas. 2020. "Challenges of Machine Learning Applied to Safety-Critical Cyber-Physical Systems." *Machine Learning and Knowledge Extraction* 2 (4):579-602.
- Qian, George, and Adam Mahdi. 2020. "Sensitivity analysis methods in the biomedical sciences." *Mathematical biosciences* 323:108306.
- Rokseth, Borge, Ingrid Bouwer Utne, and Jan Erik Vinnem. 2017. "A systems approach to risk analysis of maritime operations." *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 231 (1):53-68. doi: 10.1177/1748006X16682606.
- Russell, Stuart, and Peter Norvig. 2002. "Artificial intelligence: a modern approach."
- Schwarz, M, P Schepers, J Van Boggelen, R Loendersloot, and T Tinga. 2020. "Application of an Unvalidated Process Model to Define Operational Functional Failures." 30th European Safety and Reliability Conference (ESREL) and the 15th Probabilistic Safety Assessment and Management Conference (PSAM) 2020.
- Sobol', Ilya Meerovich. 1967. "On the distribution of points in a cube and the approximate evaluation of integrals." *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* 7 (4):784-802.