Dąbrowska Ewa, 0000-0002-6704-9541
Gdynia Maritime University, Gdynia, Poland, e.dabrowska{at}wn.umg.edu.pl
Dąbrowski Maciej, 0000-0002-8308-9156
Centre of Informatics Tricity Academic Supercomputer and networK – CI TASK, Gdańsk, Poland, mdabrowski{at}task.gda.pl

Data processing for oil spill domain movement models

Keywords

data processing, mathematical models, quality control, model reliability, oil spill domain movement

Abstract

This chapter reviews various data processing techniques for modelling the movement of oil spills, including data acquisition, quality control, and pre-processing. It highlights the importance of incorporating both physical and environmental factors such as wind, currents, and water temperature, in oil spill trajectory prediction models. It also discusses the challenges associated with data processing, including data availability and uncertainty. It emphasizes the significance of sound data processing practices to ensure effective response planning and mitigation efforts. Finally, by discussing the potential areas of improvement, and model assumptions and limitations, the chapter aims to inspire further research and development in the field, which can lead to constructing more accurate and reliable oil spill movement models.

1. Introduction

Oil spills pose a significant threat to the marine environment and can have a devastating impact on the local ecosystem, wildlife, and human health (Aguilera et al., 2010; Bogalecka, 2020). Accurate prediction of the movement of an oil spill is crucial for effective response planning and mitigation efforts (Bogalecka & Kołowrocki, 2018a, 2018b). To achieve this, complex mathematical models were developed to simulate the movement of oil spills in the water area. However, the accuracy and reliability of these models are highly dependent on the quality and processing of input data.

Data processing plays a critical role in generating accurate and reliable oil spill movement models. The term refers to the various steps involved in preparing and analyzing data to extract meaningful information (Dąbrowski, 2021, 2022). In the context of oil spill movement modelling, this includes data acquisition, quality control, and pre-processing. Data acquisition involves collecting data from various sources, such as satellite imagery, buoys, and ship sensors. Quality control involves ensuring the accuracy and reliability of the data by removing outliers and errors. Pre-processing involves transforming the data into a format that can be used in the model, such as converting raw data into numerical values. This process typically involves organizing and structuring the data in a systematic manner and assigning appropriate numerical values or codes to represent different variables or observations. The purpose is to facilitate data analysis, visualization, and further statistical computations.

Incorporating both physical and environmental factors, such as wind, currents, and water temperature, is essential in oil spill models (Fingas, 2013; French-McCay et al., 2005; Ju et al., 2022; Kołowrocki & Kuligowska, 2018; Reed et al., 1999). These factors significantly influence the movement of oil spills, and their accurate measurement and incorporation into models are critical

Safety and Reliability of Systems and Processes, Summer Safety and Reliability Seminar 2023.

DOI: 10.26408/srsp-2023-04.

for the prediction of oil spill movement. However, there are significant challenges associated with data processing for oil spill modelling. One significant dispute is over the availability of data. Data is often limited, and the cost of data acquisition can be high. Additionally, data quality and uncertainty are other significant challenges that can affect the accuracy of models.

To address these challenges, researchers have developed various data processing techniques. For instance, machine learning algorithms have been applied to improve the accuracy and models' reliability (Kailkhura et al., 2019). These algorithms can learn from historical data and can provide accurate predictions of oil spill movement. Other data processing techniques, such as data assimilation and data fusion, have also been developed to incorporate multiple sources of data into models (Xiong et al., 2020).

The Chapter offers a novel perspective by presenting an original approach and providing an overview of oil spill domain movement models, along with discussions on data collection and refinement, and the assumptions and limitations of these models. It is composed of Introduction, Sections 2–5 and Conclusion.

Section 2 discusses oil spill domain movement models used to predict the movement and fate of oil spills based on various factors, such as wind, currents, tides, and temperature. The section covers different types of models, including agent-based, stochastic, Bayesian inference, Monte Carlo simulation, and Latin hypercube sampling. The third section discusses data collection and refinement for oil spill modelling. It presents various data sources for these models, such as historical hydro-meteorological data, oceanographic data, oil spill data, remote sensing data, GIS data, and chemical properties of the spilled oil (Reed et al., 1999). The section also discusses methods for collecting and refining this data, such as on-site observations, aerial surveillance, satellite imagery, and reporting systems. Section 4 discusses the assumptions and limitations of oil spill models. Model assumptions include those about the environment, steady-state conditions, human intervention, and idealized oil properties. Model limitations include uncertainty, lack of data, sensitivity to inputs, lack of model validation, scale limitations, and model complexity. Finally, the future research perspectives are given.

2. Oil spill domain movement models

Oil spill domain movement models are modelling movement of oil spills in the environment. These models take into account various factors that affect the movement of oil, including wind, currents, tides, and temperature. The purpose of these models is to provide accurate predictions of the movement and fate of oil spills, which can help to inform decision-making in emergency response and cleanup efforts.

There are several types of oil spill domain movement models, including agent-based modelling: Eulerian, Lagrangian, and hybrid particle tracking models (Huby et al., 2022; Ivorra et al., 2021; Madsen et al., 2022; Spaulding, 2017). Agent-based modelling is a type of modelling that focuses on simulating the behaviour of individual agents within a larger system. It can be used to model the behaviour of oil spills, including how they interact with the environment and response efforts, and can provide a more detailed understanding of the dynamics of oil spills. Eulerian models are based on a fixed grid system and focus on the movement of water and oil within each grid cell. Lagrangian models, in contrast, track individual particles of oil as they move through the water. Hybrid models combine elements of both Eulerian and Lagrangian models.

On the other hand, stochastic models can be used to analyze or simulate a wide range of possible scenarios and provide probabilistic estimates of the potential trajectory and fate of an oil spill (Dąbrowska & Kołowrocki, 2019b, 2020b). These models are particularly useful when dealing with complex environments where traditional deterministic models may not provide accurate results (Grabski, 2014; Kołowrocki, 2014; Kołowrocki & Soszyńska-Budny, 2011).

Other modelling techniques that can be used in oil spill response are e.g. Bayesian inference (Das & Goerlandt, 2022). Bayesian inference is a probabilistic approach to modelling that allows for the incorporation of prior knowledge or beliefs about the input parameters. It can be used to update the probability distributions of input parameters as new data becomes available, making it useful in situations where data is limited or uncertain.

In the context of oil spill response, Monte Carlo simulation can be used to estimate the potential trajectory and fate of an oil spill, as well as the associated risks and impacts (Dąbrowska, 2023;

Dąbrowska & Kołowrocki, 2019a, 2020a). By simulating a large number of possible scenarios, decision-makers can gain a better understanding of the range of possible outcomes and the likelihood of each outcome. Monte Carlo simulation can also be used to assess the effectiveness of different response strategies and mitigation measures. By simulating the response to an oil spill under different scenarios, decision-makers can determine which strategies are most effective at reducing the potential risks and impacts of the spill. However, Monte Carlo simulation can be computationally intensive and time-consuming, particularly when simulating a large number of scenarios or when using complex models with many input parameters (Asmussen & Glynn, 2007; Bogalecka & Dąbrowska, 2023; Harris & Van Horn, 1996; Kim et al., 2013; Marseguerra & Zio, 2002; Pocora et al., 2018; Rao & Naikan, 2016). Additionally, the accuracy of the results is heavily dependent on the quality and accuracy of the input data and the assumptions made in the modelling process (Drews et al., 2003; Leigh & Bryant, 2015).

Moreover, Latin hypercube sampling is a statistical technique that can be used to sample input parameters from probability distributions in a more efficient way than traditional Monte Carlo simulation. It allows for a more systematic exploration of the input parameter space while still incorporating randomness and uncertainty.

The use of oil spill domain movement models is essential in reducing the environmental and economic impact of oil spills. However, the accuracy of these models is heavily dependent on the quality of the input data and the processing methods used to prepare the data for modelling.

3. Data collection and refinement

3.1. Data sources and availability

According to the US National Oceanic and Atmospheric Administration (NOAA, 2023) and (Bogalecka, 2020; Dąbrowska, 2021; Fingas, 2016), data sources for oil spill mathematical models considering other influential factors e.g. hydro-meteorological conditions, may include the following items.

- Historical hydro-meteorological data sets: this can be obtained from meteorological agencies or research institutions and may include data on wind speed, wind direction, precipitation, temperature and other factors.
- Oceanographic data: this can include data on water currents, water temperature, salinity, and other variables relevant to oil spill trajectory modelling; this data can be obtained from oceanographic institutions or research centres.
- Oil spill data: historical oil spill data can be used to develop and validate mathematical models. This data can be used to compare the predicted trajectory of an oil spill with observed trajectories from past spills. If the model accurately predicts the behaviour of past spills, it is more likely to accurately predict the behaviour of future spills. Historical oil spill data can also be used to identify patterns and trends in the behaviour of oil spills. For example, data on the movement and fate of oil spills in a particular region can be used to identify areas that are particularly vulnerable to spills or to identify the factors that influence the behaviour of spills in that region. This data can be obtained from various sources, including government agencies, industry reports, and scientific publications.
- Remote sensing data: satellite images and other remote sensing data can be used to monitor and track oil spills in real-time; this data can be obtained from various sources, including government agencies and research institutions.
- Geographic Information System (GIS) data: GIS data can be used to create maps and visualizations of oil spill trajectories and impacts; this data can include information on coastal features, bathymetry, and other environmental factors that can influence oil spill behaviour.
- Chemical properties of the spilled oil: the chemical properties of the spilled oil can influence its behaviour and fate in the marine environment (Reed et al., 1999). Data on the physical and chemical properties of the oil can be obtained from laboratory analysis or industry reports. Field data includes Analytical Spectral Devices (ASD) data and site photos can be used to assist in the production of ground truth image and to assess different oil spill pollution types, seawater and skylight to obtain their radiance. Ground truth in machine learning refers to the reality to be modelled with a machine learning algorithm.

The availability of these data sources may vary depending on the region and the specific circumstances of the oil spill. In some cases, data may need to be collected or synthesized from multiple sources to create a comprehensive dataset for modelling purposes.

According to (Bogalecka, 2020; DeDominicis et al., 2018; EPA, 2021; Fingas & Brown, 2018; Fingas, 2013, 2016; Jiao et al., 2019; McCreight, 2023; NOAA, 2023), identifying oil spills and collecting statistical data on them can involve several methods and sources, including:

- on-site observations: field teams can be dispatched to investigate suspected oil spills and collect data on their location, size, and characteristics; this can involve taking water samples and conducting laboratory analysis to confirm the presence of oil, especially from illegal discharges,
- aerial surveillance: aircraft equipped with cameras or sensors can fly over areas suspected of having oil spills to collect visual or infrared images; these images can be used to estimate the extent and size of the spill,
- satellite imagery: remote sensing techniques can detect oil slicks on the surface of the water and provide information on their location, size, and trajectory; satellites equipped with Synthetic Aperture Radar (SAR) sensors can detect oil spills even at night and in cloudy conditions,
- reporting systems: government agencies and industry organizations may have reporting systems in place to collect information on oil spills; this can include reports from vessels or facilities that may have caused the spill, as well as reports from concerned citizens or other observers.

Once oil spills are identified, statistical data can be collected to characterize their frequency, size, and impacts. This data can be used to inform policy decisions and improve oil spill response strategies.

Statistical data may include:

- the number of oil spills reported per year or per region,
- the size and volume of oil spilled, typically measured in barrels or gallons,
- the location and extent of oil spills, often described using geographic coordinates or map data,
- the environmental and socio-economic impacts of oil spills, such as damage to wildlife, habitats, and fisheries,
- the causes and contributing factors of oil spills, such as equipment failure, human error, or natural disasters, as stated by (EPA, 1999, 2021).

This statistical data can be obtained from various sources, including government agencies, industry reports, and scientific publications.

3.2. Significant components and parameters of oil spill model

Identifying statistically significant components of oil spill models can involve several steps, including ones given below (Dąbrowska & Kołowrocki, 2019a, 2019b, 2020a, 2020b; DeDominicis et al., 2018; Pasquill & Smith, 1983; Pethybridge et al., 2019; Zhang, 2020).

- Model selection: choosing an appropriate model that incorporates relevant influencing factors, e.g. hydro-meteorological conditions, is the first step. This can include models that simulate oil spill behaviour in response to wind, waves, and/or currents, such as simulation models, or the common ones: the ADIOS or the GNOME model (Keramea et. al., 2021; NOAA, 2023).
- Data collection: collecting data on hydro-meteorological conditions and other factors is essential to developing the model. This can include historical weather data, oceanographic data, and satellite data,
- Model calibration: calibrating the model involves adjusting its parameters to fit observed data. This can help to identify which parameters are most important for simulating oil spill behaviour.
- Sensitivity analysis: conducting a sensitivity analysis can help to identify which components of the model are most important in simulating oil spill behaviour. This involves varying the model parameters one at a time (systematically changing each individual parameter) and observing the resulting changes in the model output. By varying the model parameters this way, we can identify which parameters have the greatest impact on the model output and which have little or no impact; this information can be used to refine the model and improve its accuracy, as well as to identify areas where further

research is needed to better understand the behaviour of oil spills in the environment.

- Statistical analysis: using statistical methods to analyze the model output can help to identify which components of the model are statistically significant in simulating oil spill behaviour. This can include analyzing the correlation or conducting regression analysis to identify the most significant variables,
- Model validation: validating the model using observed data can help to confirm the accuracy of the model and identify any areas where it may need improvement. This can involve comparing the model output to observed oil spill behaviour under different hydro-meteorological conditions.

Overall, identifying statistically significant components of oil spill models involves a combination of model development, data collection, and statistical analysis. It requires a thorough understanding of the physical processes involved in oil spill behaviour and the environmental conditions that can influence those (Atlas & Hazen, 2011).

Evaluating unknown parameters of the oil spill central point drift curve can involve several methods, including ones given below (Dąbrowska & Kołowrocki, 2019a, 2019b, 2020a, 2020b; Hadzagic & Michalska, 2011; Harper et al., 2008; Zhang, 2020).

- Curve fitting: one approach is to fit a mathematical function to the observed central point drift curve. This can involve selecting a function that closely matches the observed data and estimating the parameters that minimize the difference between the observed data and the fitted function. Nonlinear regression analysis can be used to estimate the parameters.
- Bayesian inference: Bayesian inference can be used to estimate the unknown parameters of the central point drift curve by combining prior knowledge with observed data. This involves specifying a prior probability distribution for the unknown parameters and updating it based on the observed data.
- Maximum likelihood estimation: maximum likelihood estimation involves finding the values of the unknown parameters that maximize the likelihood of the observed data. This can involve specifying a statistical model for the central point drift curve and using maximum likelihood estimation to estimate the model parameters' values.
- Monte Carlo simulation: Monte Carlo simulation can be used to simulate the central point drift curve under different sets of parameter values. This involves randomly sampling from the parameter distributions and using them to simulate the drift curve. The simulated drift curves can be compared to the observed data to estimate the most likely values of the unknown parameters.

Overall, evaluating unknown parameters of the oil spill central point drift curve involves a combination of statistical analysis and model fitting techniques. It requires careful consideration of the physical processes involved in oil spill behaviour and the uncertainties associated with the available data.

3.3. Quality control and data cleaning

Quality control and data cleaning are vital steps in ensuring the accuracy and reliability of data used in oil spill modelling. Quality control entails verifying the accuracy, completeness, and consistency of the data, aligning it with the expected standards. On the other hand, data cleaning focuses on detecting and rectifying errors and inconsistencies within the data (Van den Broeck et al., 2005).

The following are some common methods used in quality control and data cleaning (Mitchell & Sheehy, 1997; Neutatz et al., 2022):

- data validation: this involves checking the data for completeness, accuracy, and consistency; this can include verifying that the data is in the correct format, that there are no missing values, and that the values fall within expected ranges,
- outlier detection: outliers are data points that are significantly different from other data points and may indicate errors or anomalies in the data; outlier detection involves identifying and either correcting or removing these data points,
- data imputation: missing data can be imputed or estimated based on other available data; this can involve methods such as mean imputation, regression imputation, nearest neighbour imputation or extrapolation,
- duplicates detection: duplicates are records that are identical or nearly identical to each other; detecting and removing duplicates is important to avoid bias in the analysis,

- cross-validation: cross-validation involves testing the model on a subset of the data that was not used in training the machine-learning model; it means that we are using a set of input data and their corresponding output data to teach the model how to make accurate predictions or classifications.

To clarify the last method, if we develop a machine learning model to predict the trajectory of an oil spill based on environmental conditions such as wind speed, ocean currents, and water temperature, we need to train the model using historical data of oil spills and the corresponding environmental conditions. This training process involves presenting the model with a set of input data (e.g., wind speed, ocean currents, water temperature) and their corresponding output data (e.g., the observed trajectory of the oil spill) and adjusting the parameters of the model until it can accurately predict the output data. The training process typically involves dividing the available data into two sets: a training set and a validation set. The training set is used to train the model, while the validation set is used to evaluate the performance of the model on new, unseen data. Once the model has been trained and its parameters optimized, it can be used to make predictions or classifications on new data. The accuracy of the model's predictions on new data can be evaluated using metrics such as mean squared error, mean absolute error, or cross-entropy loss. This can help ensuring that the model is not overfitting the data and is generalizing well to new data.

In general, quality control and data cleaning are essential steps in ensuring the accuracy and reliability of data used in oil spill modelling. They can help to identify and correct errors and inconsistencies in the data and improve the quality of the analysis.

3.4. Data preparation for modelling

Data preparation is a critical step in modelling oil spills (Jiao et al., 2019). The following are some common methods used in data preparation (Maharana et al., 2022):

- data integration: data integration involves combining data from different sources and formats into a single dataset; this can include combining hydro-meteorological data with oil spill data, or integrating data from different sensors or platforms,
- data pre-processing: data pre-processing involves transforming the data to make it suitable for modelling, this can include converting data to a common format, scaling the data to a specific range, or normalizing the data to remove bias,
- feature selection: feature selection involves identifying the most relevant variables or features for the model, this can help to reduce the dimensionality of the data and improve the model's performance,
- data partitioning (for machine-learning models): data partitioning involves dividing the data into training, validation, and testing sets; this can help to evaluate the performance of the model and avoid overfitting,
- data augmentation (for machine-learning models): data augmentation involves creating additional data points by applying transformations to the existing data; this can help to increase the size of the dataset and improve the model's performance.

Generally, data preparation is a critical step in modelling oil spills impacted by e.g. hydro-meteorological conditions. It involves combining and transforming the data to make it suitable for modelling, selecting relevant features, partitioning the data for training and testing, and augmenting the data to increase its size and diversity. These steps can help to ensure the accuracy and reliability of the model and improve its performance in predicting oil spill behaviour.

4. Assumptions and limitations of models

Oil spill modelling techniques can be useful tools for emergency response and contingency planning. However, it is important to understand the assumptions and limitations of the models used.

4.1. Model assumptions

Oil spill modelling relies on a number of assumptions (Dąbrowska, 2023; Neutatz et al., 2022):

- homogeneous environment: some oil spill models assume that the environment in which the oil is spilled is homogeneous, with no significant variations in water currents, winds, waves, or temperature; in reality, the environment is complex and heterogeneous, with many factors influencing the behaviour of spilled oil,

- steady-state conditions: most oil spill models assume steady-state conditions, where the oil is not changing over time; in reality, oil is constantly changing, weathering, and breaking down, which can have a significant impact on its behaviour,
- no human intervention: oil spill models may assume that there is no human intervention in the response during the spill, such as the use of booms or dispersants; however, the use of these interventions can have a significant impact on the behaviour and fate of spilled oil,
- idealized oil properties: oil spill models often assume idealized properties for the spilled oil, such as a constant density and viscosity; in reality, the properties of spilled oil can vary widely depending on the type of oil, the environmental conditions, and the age of the spill.

While these assumptions are necessary for the models to be useful, they can also introduce limitations and uncertainties in their predictions.

4.2. Model limitations

It is important to understand the limitations of the oil spill models (Dąbrowska, 2023; Zio & Aven, 2013; Zhang, 2020):

- uncertainty: oil spill modelling is subject to a high degree of uncertainty, due to the many variables that can affect the behaviour of spilled oil; models can provide useful predictions, but they should always be used in conjunction with other sources of information, such as field observations and remote sensing data,
- lack of data: oil spill models require input data on a range of parameters, such as wind, waves, currents, and oil properties; however, in many cases, there may be limited data available, which can limit the accuracy of the model predictions,
- sensitivity to inputs: oil spill models are highly sensitive to the input parameters used, such as the initial spill volume, the oil properties, and the environmental conditions; small changes in these parameters can have a significant impact on the model's predictions,
- lack of model validation: oil spill models require validation against real-world data to confirm their accuracy; however, there may be limited opportunities to validate models due to the infrequency of oil spills and the logistical challenges of collecting data in the field,
- scale limitations: oil spill models are often limited by the scale at which they operate, for example, a model designed to predict the behaviour of spilled oil in a river may not be applicable to a large oil spill in the ocean; models must be carefully calibrated and validated for the specific scenario in which they are being used,
- model complexity: oil spill models can be highly complex and require significant computational resources; this can limit their applicability in emergency response situations where time is critical; simplified models may be more appropriate in these situations.

While oil spill modelling can be useful for predicting the behaviour and fate of spilled oil in the environment, there are still important limitations to consider.

5. Future research

Future research on oil spill modelling should aim to improve the accuracy and reliability of the models by exploring new and innovative data sources. This could involve the use of remote sensing technologies like radar and satellite imagery, as well as drones and autonomous underwater vehicles. Collaborations with local communities and industries could also provide comprehensive data on the affected environment, such as the types of flora and fauna present, which could inform models and mitigation strategies.

Furthermore, the subsequent studies should focus on identifying the most significant components and parameters of oil spill models to improve their accuracy. Sensitivity analyses could determine which inputs have the greatest influence on the model outputs, and the development of machine learning and artificial intelligence algorithms could identify hidden patterns and relationships in the data.

To ensure the accuracy of oil spill models, further investigation should improve the quality control and data cleaning processes. Automated tools could identify and correct errors in the data, and standardized data formats could ensure consistency across different data sources. Additionally, standardized data pre-processing workflows that take into account the unique characteristics of the data and comprehensive data dictionaries could improve the transparency and reproducibility of the modelling process.

Finally, future research should identify and quantify the assumptions and limitations of oil spill models and develop methods for addressing them. Uncertainty and sensitivity analyses could assess the impact of different assumptions on the model outputs, and the development of alternative modelling approaches could improve the accuracy and reliability of oil spill modelling. Continued research in these areas could lead to more accurate and reliable oil spill models and better strategies for mitigating their environmental impact. Additionally, identifying areas for subsequent studies can help to guide the allocation of resources and funding to address these critical issues.

6. Conclusion

This chapter provides an overview of the data processing concern in oil spill domain movement models, including the steps involved in data processing, the challenges associated with data processing, and best practices for ensuring the accuracy and reliability of oil spill domain movement models.

In summary, data processing is critical for generating precise oil spill movement models. Incorporating physical and environmental factors, such as wind, currents, and water temperature, is essential for the prediction of oil spill movement. However, data availability, quality, and uncertainty are significant challenges that must be addressed to ensure the accuracy of models.

References

Aguilera, F., Méndez, J., Pásaro, E., & Laffon, B. 2010. Review on the effects of exposure to spilled oils on human health. *Journal of Applied Toxicology* 30(4), 291–301.

Asmussen, S. & Glynn, P.W. 2007. *Stochastic simulation – algorithms and analysis. Stochastic modeling and applied probability.* Springer Science & Business Media, New York.

Atlas, R.M. & Hazen, T.C. 2011. Oil biodegradation and bioremediation: A tale of the two worst spills in US history. *Environmental Science & Technology* 45(16), 6709–6715.

Bogalecka, M. 2020. *Consequences of Maritime Critical Infrastructure Accidents. Environmental Impacts. Modeling – Identification – Prediction – Optimization – Mitigation*. Elsevier, Amsterdam – Oxford – Cambridge.

Bogalecka, M. & Dąbrowska, E. 2023. Monte Carlo simulation approach to shipping accidents consequences assessment. *Water* 15, 1824.

Bogalecka, M. & Kołowrocki, K. 2018a. Prediction of critical infrastructure accident losses of chemical releases impacted by climate-weather change. *Proceeding of 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Institute of Electrical and Electronics Engineers, Bangkok, 788–792.

Bogalecka, M. & Kołowrocki K. 2018b. Minimization of critical infrastructure accident losses of chemical releases impacted by climate-weather change. *Proceeding of 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM),* Institute of Electrical and Electronics Engineers, Bangkok, 1657–1661.

Das, T. & Goerlandt, F. 2022. Bayesian inference modeling to rank response technologies in arctic marine oil spills, *Marine Pollution Bulletin,* Part A, 185, 114203.

Dąbrowska, E. 2021. Conception of oil spill trajectory modelling: Karlskrona seaport area as an investigative example. *Proceeding of 5th International Conference on System Reliability and Safety (ICSRS)*, Palermo, Italy, 307–311.

Dąbrowska, E. 2023. Oil discharge trajectory simulation at selected Baltic Sea waterway under variability of hydro-meteorological conditions. *Water* 15, 1957.

Dąbrowska, E. & Kołowrocki, K. 2019a. Monte Carlo simulation applied to oil spill domain at Gdynia port water area determination. *Proceeding of IEEE, International Conference on Information and Digital Technologies (IDT)*, Žilina, Slovakia, 98–102.

Dąbrowska, E. & Kołowrocki, K. 2019b. Stochastic determination of oil spill domain at Gdynia port water area, *Proceeding of International Conference on Information and Digital Technologies*, Žilina, Slovakia, 92–97.

Dąbrowska, E. & Kołowrocki, K. 2020a. Monte Carlo simulation approach to determination of oil spill domains at port and sea waters areas, *The International Journal on Marine Navigation and Safety of Sea Transportation* 14(1), 59–64.

Dąbrowska, E. & Kołowrocki, K. 2020b. Probabilistic approach to determination of oil spill domains at port and sea water areas, *The International Journal on Marine Navigation and Safety of Sea Transportation* 2020, 14 (1), 51–58.

Dąbrowski, M. 2021. Analysis and comparison of data collection and processing tools for network devices MSc Thesis. *Faculty of Electronics, Telecommunications and Informatics*, Gdańsk University of Technology.

Dąbrowski M. 2022. Collecting and processing data of network devices impacting system load in terms of monitoring and warning system implementation. K. Kołowrocki et al. (Eds.), *Safety and Reliability of Systems and Processes, Summer Safety and Reliability*, Gdynia Maritime University, Gdynia, 65–77.

DeDominicis, M., Galmarini, S., Doucet, P.G., Solazzo, E., Cappelletti, A., Klaić, Z.B., Carmichael, G., Hirtl, M., Potempski, S., Trapp, W. & Vira, J., 2018. Model evaluation and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII phase 2. *Atmospheric Environment* 178, 155–172.

Drews, T.O., Braatz, R.D., Alkire, R.C. 2003. Parameter sensitivity analysis of Monte Carlo simulations of copper electrodeposition with multiple additives. *Journal of the Electrochemical Society* 150(11), C807–C812.

EPA. 2021. *Oil Spills Prevention and Preparedness Regulations*, https://www.epa.gov/oil-spills-prevention-and-preparedness-regulations (accessed 29 May 2023).

EPA. 1999. *Understanding Oil Spills and Oil Spill Response. Understanding Oil Spills in Freshwater Environments*. EPA 540–K–99–007.

Fingas, M. & Brown, C.E. 2018. Review of oil spill remote sensing. *Sensors* 18(1), 91.

Fingas, M. 2013. *The Basics of Oil Spill Cleanup*. 3rd ed. CRC Press, Boca Raton – London – New York.

Fingas, M. 2016. *Oil Spill Science and Technology*, 2nd edn. Elsevier, Amsterdam – Boston – Heidelberg – London – New York – Oxford – Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo.

French-McCay, D., Rowe, J., Sankaranarayanan, S., Kim, S. & Aurand, D. 2005. Use of probabilistic trajectory and impact modelling to assess consequences of oil spills with various response strategies. *Proceeding of the 28th Arctic and Marine Oilspill Program (AMOP) Technical Seminar*, Canada, 1134, 253–271.

Grabski, F. 2014. *Semi-Markov Processes: Application in System Reliability and Maintenance*. Elsevier, Amsterdam – Boston – Heidelberg – London – New York – Oxford – Paris – San Francisco – Sydney – Tokyo.

Hadzagic, M. & Michalska, H.A. 2011. Bayesian inference approach for batch trajectory estimation. *Proceeding of 14th International Conference on Information Fusion*, Chicago, USA, 1–8.

Harper, W.V., James, T.R., Eschenbach, T.G., & Slauson, L.V. 2008. Maximum likelihood estimation methodology comparison for the three-parameter weibull distribution with applications to offshore oil spills in the Gulf of Mexico. *JSM Proceeding, Risk Analysis Section, American Statistical Association, Denver, CO. 1231-1238.*

Harris, G. & Van Horn, R. 1996. *Use of Monte Carlo Methods in Environmental Risk Assessment at the INEL: Application and Issues.* Report, 1 June 1996,; Idaho Falls, Idaho. https://digital.library.unt.edu/ark:/67531/metadc671670/m1/1/ (accessed 26 Jun 2023)

Huby, A.A., Sagban, R. & Alubady, R. 2022. Oil spill detection based on machine learning and deep learning: A review. *Proceeding of 5th International Conference on Engineering Technology and its Applications (IICETA)*, Al-Najaf, Iraq, 85–90.

Ivorra, B., Gomez, S., Carrera, J. & Ramos, A.M. 2021. A compositional Eulerian approach for modeling oil spills in the sea. *Ocean Engineering* 242, 110096.

Jiao, Z., Jia, G. & Cai, Y. 2019. A new approach to oil spill detection that combines deep learning with unmanned aerial vehicles. *Computers & Industrial Engineering* 135, 1300–1311.

Ju, X., Li, Z., Dong, B., Meng, X. & Huang, S. 2022. Mathematical physics modelling and prediction of oil spill trajectory for a Catenary Anchor Leg Mooring (CALM) System. *Advances in Mathematical Physics*, 2022, 3909552.

Kailkhura, B., Gallagher, B., Kim, S. Hiszpanski, A. & Han, T.Y.-J. 2019. Reliable and explainable machine-learning methods for accelerated material discovery. *NPJ Computional Materials* 5, 108.

Keramea, P., Spanoudaki, K., Zodiatis, G., Gikas, G. & Sylaios, G. 2021. Oil spill modeling: A critical review on current trends, perspectives, and challenges. *Journal of Marine Science and Engineering* 9(2), 181.

Kim, T., Yang, C.-S., Ouchi, K. & Oh, Y. 2013. Application of the method of moment and Monte-Carlo simulation to extract oil spill areas from synthetic aperture radar images, *2013 OCEANS – San Diego*, 1–4.

Kołowrocki, K. & Soszyńska-Budny, J. 2011. *Reliability and Safety of Complex Technical Systems and Processes: Modeling – Identification – Prediction – Optimization*. Springer, London.

Kołowrocki, K. 2014. *Reliability of Large and Complex Systems*. Elsevier, Amsterdam – Boston – Heidelberg – London – New York – Oxford – Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo.

Kołowrocki, K. & Kuligowska, E. 2018. Operation and climate-weather change impact on maritime ferry safety. S. Haugen et al. (Eds). *Safety and Reliability – Safe Societies in a Changing World*. CRC Press Taylor & Francis Group, London, 849–858.

Leigh, J.W. & Bryant, D. 2015. Monte Carlo strategies for selecting parameter values in simulation experiments. *Systematic Biology* 64(5), 741–751.

Madsen, M., Skov, H. & Potthoff, M. 2022. Combining predicted seabird movements and oil drift using lagrangian agent-based model solutions. M. Mancuso et al. (EDS.). *Marine Pollution – Recent Developments.* Environmental Sciences. IntechOpen 2023, 106956.

Maharana, K., Mondal, S. & Nemade, B. 2022. A review: Data pre-processing and data augmentation techniques, *Global Transitions Proceedings* 3(1), 91–99.

Marseguerra, M. & Zio, E. 2002. *Basics of the Monte Carlo Method with Application to System Reliability*. LiLoLe – Verlag, Hagen.

McCreight, R. 2023. Gauging the impact of satellite & space systems on critical infrastructure [CI]: risk management is neither an enigma nor a mystery for CI systems security. *Journal of Homeland Security and Emergency Management* 20(2), 183–208.

Mitchell, P.L., Sheehy, J.E. 1997. Comparison of predictions and observations to assess model performance: a method of empirical validation, F.W. Penning de Vries (Eds.). *Applications of Systems Approaches at the Field Level. Systems Approaches for Sustainable Agricultural Development* 6. Springer, Dordrecht.

Neutatz, F., Chen, B., Alkhatib, Y., Ye, J. & Abedjan, Z. 2022. Data cleaning and AutoML: would an optimizer choose to clean? *Datenbank Spektrum* 22, 121–130.

NOAA. 2023. *GNOME Suite for Oil Spill Modeling*, https://response.restoration.noaa.gov/oil-and-chemical-spills/oil-spills/response-tools/gnome-suite-oil-spill-modeling.html (accessed 29 May 2023).

NOAA. 2023. *Oil Spills. Oil Types*, https://response.restoration.noaa.gov/oil-and-chemical-spills/oil-spills/oil-types.html (accessed 29 May 2023).

Pasquill, F. & Smith, F.B. 1983. *Atmospheric Diffusion*, 3rd ed. Ellis Horwood Limited, Chichester.

Pethybridge, H.R., Weijerman, M., Perrymann, H., Audzijonyte, A., Porobic, J., McGregor, V., Girardin, R., Bulman, C., Ortega-Cisneros, K., Sinerchia, M., Hutton, T., Lozano-Montes, H., Mori, M., Novaglio, C., Fay, G., Gorton, R. & Fulton, E. 2019. Calibrating process-based marine ecosystem models: an example case using Atlantis, *Ecological Modelling* 412, 108822.

Pocora, A., Purcarea, A.A., Nicolae, F. & Cotorcea, A. 2018. Modelling and simulation of oil spills in coastal waters. *Proceeding of the 2018 IOP Conference Series: Earth Environmental Science* 172, 012012.

Rao, M.S. & Naikan, V.N.A. 2016. Review of simulation approaches in reliability and availability modeling. *International Journal of Performability Engineering* 12(4), 369–388.

Reed, M., Johansen, O., Brandvik, P.J., Daling, P., Lewis, A., Fiocco, R., Mackay, D. & Prentki, R. 1999. Oil spill modeling towards the close of the 20th century: overview of the state of the art. *Spill Science & Technology Bulletin* 5(1), 3–16.

Spaulding, M.L. 2017. State of the art review and future directions in oil spill modeling. *Marine Pollution Bulletin* 115(1–2), 7–19.

Van den Broeck, J., Cunningham, S. A., Eeckels, R., & Herbst, K. 2005. Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS Medicine* 2(10), e267.

Xiong, Z., Sang, J., Sun, X., Zhang, B. & Li, J. 2020. Comparisons of performance using data assimilation and data fusion approaches in acquiring precipitable water vapor: A case study of a western United States of America area. *Water* 12, 2943.

Zhang, Y. 2020. Sensitivity analysis in simulating oil slick using CA model, *Ocean Engineering* 218, 108216.

Zio, E. & Aven, T. 2013. Model output uncertainty in risk assessment. *International Journal of Performability Engineering* 29(5), 475–486.