# Quantifying the effect of noise, vibration and motion on human performance in ship collision and grounding risk assessment

**Jakub Montewka**[a*], **Floris Goerlandt**[a], **Gemma Innes-Jones**[b], **Douglas Owen**[b], **Yasmine Hifi**[c], **Markus Porthin**[d]
[a] Aalto University, Department of Applied Mechanics, Marine Technology, Research Group on Maritime Risk and Safety, Espoo, Finland
b – Lloyd's Register, EMEA, Bristol, UK
c – Brookes Bell R&D, Glasgow, UK
d - VTT Technical Research Centre of Finland, Espoo, Finland

**Abstract:** Risk-based design (RBD) methodology for ships is a relatively new and a fast developing discipline. However, quantification of human error contribution to the risk of collision or grounding within RBD has not been considered before. This paper introduces probabilistic models linking the effect of ship motion, vibration and noise with risk through the mediating agent of a crewmember. The models utilize the concept of Attention Management, which combines the theories described by Dynamic Adaptability Model, Cognitive Control Model and Malleable Attentional Resources Theory. To model the risk, an uncertainty-based approach is taken, under which the available background knowledge is systematically translated into a coherent network and the evidential uncertainty is qualitatively assessed.

The obtained results are promising as the models are responsive to changes in the GDF nodes as expected. The models may be used as intended by naval architects and vessel designers, to facilitate risk-based ship design.

**Keywords:** Risk-Based Ship Design, Bayesian Belief Networks, Risk Assessment, Collision Probability, Grounding Probability.

## 1. INTRODUCTION

In the Risk-Based Ship Design (RBSD) methodology the assessment of the risk level of a new ship is conducted in the early design stage, where a design modification is easy and cost-effective [1]. In this approach, risk is evaluated alongside conventional design performance measures like sufficient strength and stability, low resistance, cargo carrying capacity, propulsion and maneuvering capability. Risk is thus treated as a design objective rather than a constraint imposed by prescriptive safety rules.

In line with an increased focus on the human element in the maritime domain, it is possible to focus on human performance in the early design phase. This paper presents advances to the RBSD methodology, focusing on the effect of specific performance-shaping factors (PSFs) known as global design factors (GDF) on human performance and ultimately on ship collision and grounding risk. These GDFs comprise ship motion, noise and whole-body vibrations (WBV). Considerable scientific attention has been paid to developing risk models for ship collision and a variety of approaches have been presented quantifying the specific effect of GDFs on human error and collision risk has not been considered before, for the review of the existing methods see for example [2]–[5]. Quantifying this effect is however challenging as common human error quantification frameworks do not readily account for the specific effect of the GDFs, see for example [6]–[10].

A workable approach for human performance risk modelling has emerged from the literature on the effects of exposure to ship motion, noise and vibration GDFs focussing on attention management, [11]. It is based on three theories: the Dynamic Adaptability Model, - [12] - Cognitive Control Model - [13] - and Malleable Attentional Resources Theory - [14].

These foundations are used as a guide for constructing the risk model, which is developed using Bayesian Belief Networks (BBNs). The methodology applied here for the risk assessment stems from an uncertainty-based risk perspective, where the assessment is used to transform the available background knowledge in a coherent framework to assess uncertainty about the occurrence of a collision or grounding event. Special attention is given to the evidential uncertainties underlying the BBNs structure and probabilities, which are considered alongside sensitivities. BBNs, as probabilistic tools are capable of representing background knowledge about the collision and grounding phenomena, the quantification of associated uncertainties, efficient reasoning and updating in light of new evidence, see for example [2], [15], [16]. Two models are presented in this paper. They are stand-alone models that behave in response to GDF inputs as intended. The results are comparable with other existing models, and the data from the marine industry. The models can enable comparative assessment of ship designs, which is their primary intention.

The remainder of the paper is organized as follows: Section 2 introduces the adopted risk perspective, upon which general structure of the model is developed, as presented in Section 3. Section 4 elaborates on the human performance model. Section 5 introduces and discusses two risk models developed here, whereas Section 6 concludes.

## 2. ADOPTED RISK PERSPECTIVE

In this paper we adopted an uncertainty perspective of risk, where risk is seen as follows, [17]:

$$R \sim C\&U \tag{1}$$

This means that risk assessment is an expression of an assessor's uncertainty (U) about the occurrence of events and the associated consequences (C). Following this perspective, risk assessment can always be performed, as the risk model is seen as a tool to describe and convey uncertainties rather than a tool to uncover the truth. For this purpose, the risk model encompasses the events and the consequences of the events when they become true.

The models presented here focuses on two events namely ship-ship encounter and ship-ground encounter. The consequences of the above events are ship-ship collision and ship running aground respectively. It is assumed, that these accidents stem from the inability of a navigator to perform evasive action, when exposed to an encounter with another ship or shallow waters. This inability results from the reduced performance, which has been affected by the GDFs. Thus, the model links the effect of GDFs on human performance with the probability of an accident.

Finally, the risk model presented in this paper delivers the probability of an accident given encounter, accounting for the available background knowledge about the analysed phenomena. The background knowledge in terms of available data, models, theories and expert judgement, is assessed qualitatively, which makes it possible to elaborate on the evidential uncertainty of the risk model. Such information coupled together with the results of the sensitivity analysis of the model provides information about the outcome uncertainty of the model. This way of analysing the uncertainty allows an analyst determining the areas of the domain under study, which needs further research in order to improve the overall performance of the risk model.

The adopted risk perspective allows defining the plethora of consequences and associated uncertainties. This makes it possible to expand the model with desired consequences, which can be very specific, depending on the ship type under analysis. For instance, in case of a RoPax ship, the societal impact may be of interest, thus the risk in this case will be expressed through the probability of a number of fatalities, see for example [2]. Whereas, in case of a tanker, the environmental impact of an accident may be a key issue, thus the risk is expressed as the probability of oil spill of certain size [18]. However, detailed quantification of the specific consequences of an accident is out of the scope of this paper.

## 3. GENERAL STRUCTURE OF THE RISK MODEL

To describe the process through which exposure to GDFs causally affects the probability of the specified unwanted outcomes, a causal pathway was developed through the mediating agent of the crewmember. Importantly, the causal chain represents the effects of GDFs exposure on human performance in a way that could be developed and elaborated in the risk model. It served to do three things:

1. Represent the mechanism by which GDFs exposure impacts collision and grounding risk.
2. Describe the overall topography of the final model.
3. Facilitate the identification of nodes.

GDFs can be considered a type of PSF, where PSFs are an aspect of the human's individual characteristics, environment, organisation, or task that specifically decrements or improves human performance, thus increasing or decreasing the likelihood of human error respectively, [19]. While there are many other PSFs that can affect human behaviour – for instance training, experience, competence, time available, workload, job design, manning, ergonomics of the equipment and procedures - these are excluded from the collision and grounding risk models as they are not affected by exposure to GDFs. All the excluded PSFs are implicitly assumed to remain constant within the model.

Other potentially relevant factors, which are not considered, are the long terms effects of GDFs on the crew performance. For example, we do not consider the hearing loss due to long-term noise exposure either individually or in combination with other GDF effects. In practical terms only the effect of GDF-affected human performance on the possible occurrence of collision and grounding in combination with the safety critical task being performed are considered.

In the presented model the inputs and outputs of the model are predetermined. The GDFs form the three inputs: ship motion, noise and vibration. The effect of the latter is considered through WBV. The unwanted outcomes form the two outputs: collision, grounding. In reality, crew exposure to GDFs is likely to result in a plethora of effects on human performance and subsequent outcomes. Likewise, the unwanted outcomes are likely to have numerous causal inputs, which may include GDFs. However, to remain within the scope of our study the causal representation is limited to describing only those mechanisms that can describe the relationships between the predetermined inputs and outputs, as depicted in Figure 1.

Two main paths linking GDF exposure to human behaviour, and subsequently to collision and grounding, have been identified:

- Path 1: Stressor effects. Exposure to a GDF acts as a stressor and can affect the perceptual, cognitive and physical capabilities of an individual (e.g. attention management), which can subsequently impair the performance of the individual (i.e. the actual behaviour produced).
- Path 2: Physical effects. Exposure to a GDF can have specific and direct effects on the behaviour produced. For example, Ship motion can result in Motion-Induced Interruptions (MII). MII does not affect the underlying human capabilities of balance or fine motor control, but it exceeds the ability of the human to compensate and produce the intended behaviour. Similarly, WBV can directly impact the actual behaviour produced.

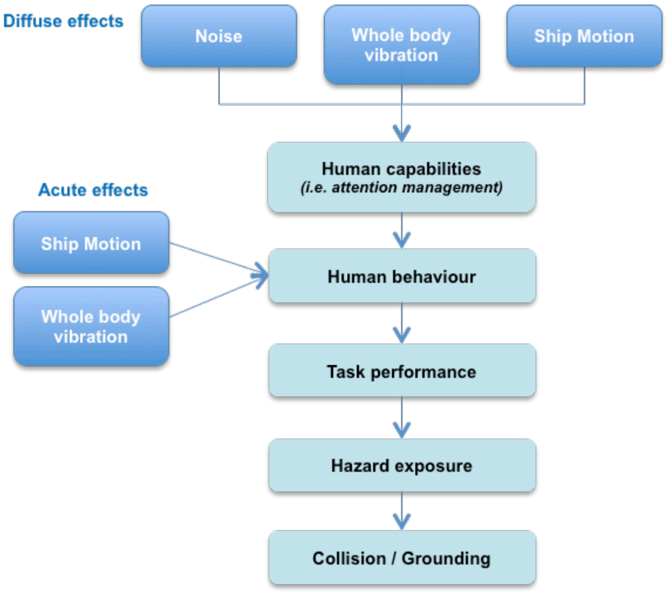
These two paths show how GDF exposure affects human behaviour, which in turn influences the performance of safety critical tasks. It is the outcomes of an individual's actions and behaviour that determine the success or failure of a safety critical task. Insufficient performance of the safety critical tasks associated with maintaining safe vessel navigation and avoiding collision or grounding create an antecedent for the unwanted outcome. However, insufficient task performance alone does not determine whether or not a collision or grounding occurs; the vessel must also be exposed to the collision or grounding hazard, as follows:

- For a collision to occur, another vessel must be on a collision course.
- For a grounding to occur, the ship must be in shallow water.

This causal mechanism makes the following assumptions:

- While we recognise that individuals have differing cognitive and physical abilities, it is assumed that all individuals have the same basic set of capabilities (i.e. all individuals can manage their attention, irrespective of the extent of this capability).
- Human behaviour is influenced by diffuse and acute effects of GDF exposure as represented by the paths in Figure 1.
- The crew perform safety critical tasks related to collision and grounding.
- Tasks are appropriate, processes and procedures are optimised, and are undertaken by a competent operator.
- Safety critical tasks must be performed correctly to maintain safe vessel operation.
- Safety critical tasks manage the exposure of the vessel to the collision and grounding hazard.
- While it is recognised that interaction effects between GDFs within each pathway are likely to exist, these are excluded from the model, as the literature does not provide any information describing this interaction [20]. While no representation is included of any potential interaction effects from the exposure to multiple GDFs acting through a single pathway (i.e. stressor or physical effects), a cumulative effect is represented for the presence of GDFs acting both through stressor and physical effects of GDFs simultaneously.

**Figure 1: Causal chain describing the relationship between crew GDF exposure and unwanted outcomes**



## 4. HUMAN PERFORMANCE MODEL

It was found that the data on the specific GDF effects of ship motion (with the exception of MII), noise, WBV on human performance are sparse and in many, but not all, cases generated under very specific, often non-marine, conditions. Data shows that there is certainly evidence for GDFs having some effect on human performance. However, the direct effects of GDF exposure on human performance tend to be weak, whereas secondary effects acting through another mechanism (e.g. fatigue, Motion Induced Sickness (MIS)) tend to be stronger and more pervasive. Specifically there are some data that describe the:

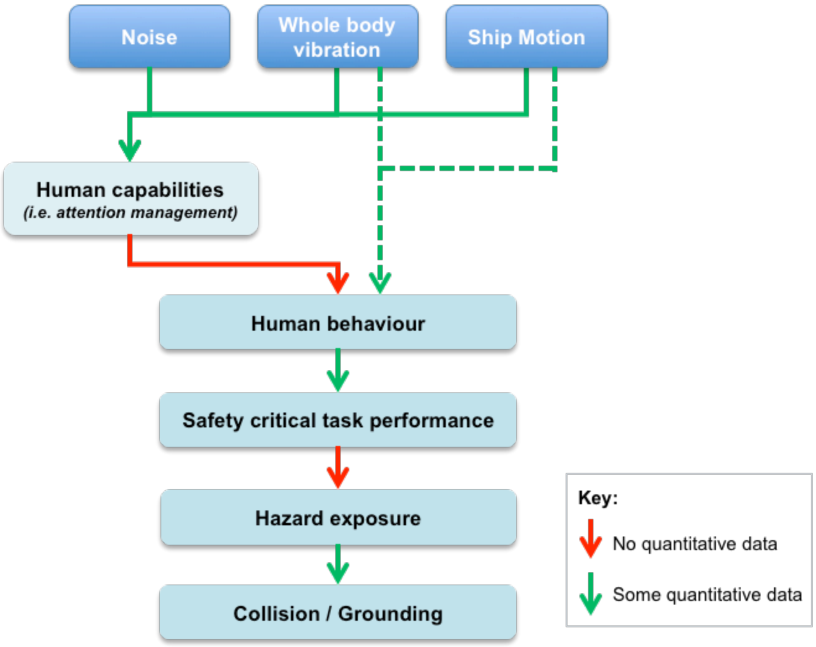
- Impact of GDFs on specific human capabilities, [21].
- Impact of GDFs on specific human behaviours, [22].
- Impact of errors on task performance, [23].

However, there is very little data about the link between the following components:

- Degraded human capabilities and collision or grounding related performance
- Degraded task performance and exposure to the collision / grounding hazard

Figure 2 demonstrates the links in the causal chain for which some quantitative data are available (in green) and the links for which there is no data (in red). In addition to this gap, a given level of exposure to GDFs of certain intensity or duration may not affect all individuals equally. For example, while a given frequency and amplitude of ship motion may be generally MIS-inducing, individual experiences may range from significant nausea to no negative affects whatsoever, depending on their underlying susceptibility to MIS and the degree to which they have acclimatized. Moreover, with the possible exception of secondary effects on human performance caused by fatigue, attributable to sleep disruption, a holistic view could not readily be derived directly from the individual findings. As such, our approach was guided by the relevant theoretical models available in the scientific literature.

**Figure 2: Supporting data for links in the causal chain.**



## 4.1. Theoretical models of human performance

The approach taken here to describe a mechanism that accounts for the impact of stressors on human performance, has been based on the principles of attention management, [24]. It combines the principles from three theoretical models:

1. Dynamic Adaptability Model (DAM), [12].
2. Cognitive Control Model (CMM), [13].
3. Malleable Attentional Resources Theory (MART), [14].

Under the DAM paradigm, GDFs are seen as types of physical stressor that affect human capabilities associated with maintaining a desired level of task performance either directly or indirectly (e.g. via fatigue). When exposed to GDFs, CCM describes humans compensating through the effortful direction of more cognitive resources at the task, typically at the cost of performance in other areas. Despite the sophisticated, and potentially subconscious, strategies humans have at their disposal, there is a limit to how much an individual can compensate without experiencing degradation in primary or secondary task performance. In addition, the extent to which human can compensate for task demands is not fixed. MART describes this compensatory capability changing as a function of task demands and associated arousal an individual experiences – attentional resources available vary as a function of load. When humans are in a state of under-load (i.e. bored) their pool of attentional resources is relatively small and will increase proportionately with the demands placed on them. However, there is a limit to how much the pool of attentional resources can grow. When task demands exceed the pool of attentional resources available (either transiently or when the upper attentional resource limit is exceeded), performance can breakdown and errors may be made.

Generally, task performance is only expected to degrade and become insufficient when compensatory mechanisms have failed. However, the literature does not allow prediction of how and when (chronologically) an operator would fail, under what conditions of GDF exposure, and what the specific effect on behaviour (i.e. type of error) would be.

In the risk models presented here, the main task around which the models revolve is to perform the accident evasive action. This task is complex and distributed in time, but it can be decomposed into three major phases:

1. detection (D),
2. assessment (A),
3. action (Act).

These three phases (DAAct) reflect the basic cognitive functions of observation, interpretation and planning, and execution, see for example [25], [26].

In terms of risk modelling, an approach based on attention management theory allows representation of the effect of GDF exposure as a stressor that sits either above or below the threshold of attentional capacity for any given task. If the stressor exceeds the attentional capacity then a negative effect is expected, whereas no negative effect on human performance would result if the stressor can be managed within the available attentional capacity.

Representing ship motion, noise and WBV GDFs as stressors interacting with an individual's attention management capabilities provides an evidence-based mechanism for human performance that has been used to develop the risk models presented here.

### 4.2. Integration of Human Reliability Assessment in the Risk model

Due to the limitations in data on the effects of GDF exposure on human performance, one cannot find precise values in the scientific literature. Hence probabilistic representation of the human performance component in the risk model was potentially problematic. A solution was found in Human Reliability Analysis (HRA) techniques. While HRA techniques do not typically cover the specific GDFs or the maritime environment, the human error probabilities (HEPs) generated by HRA allow sensible bounds to be determined. While imperfect, this approach at least allows calibration of probabilities in the risk model against established generic human error probability values.

The HRA method Nuclear Action Reliability Assessment (NARA) was selected to provide the HEPs associated within collision and grounding model, [27]. NARA is a third generation HRA method that, while nuclear industry focussed, uses a broad range of industries in the CORE-DATA dataset underlying the HEP calculation, arguably making it more suitable for navigation tasks performed on the bridge. NARA was adopted to enhance the accuracy of the risk model through the generation of validated (albeit non-marine specific) HEPs associated with task characteristics that are compatible with tasks performed by the Officer of the Watch (OOW) and helmsman. NARA also provided baseline error rates for a given Generic Task Type (GTT) unaffected by GDFs. This allowed probabilistic estimation of the effect of GDF exposure on HEPs.

However, NARA was not used to represent the direct physical effect exposure some GDFs may have as physical aspects of task performance. This is out the scope the intended application of the NARA method. The probability of insufficient human performance resulting from physical effects of GDF exposure was estimated based on judgement alone.

NARA categorises the factors that negatively influence human performance as one of eighteen Error Producing Conditions (EPCs). The EPC that best represented the causal mechanism from GDF exposure to human performance was EPC No. 15: *Poor Environment*. This EPC represents the stressor effect of GDF exposure on attention management capability. The potential strength of effect of this EPC was set using the Assessed Proportion of Affect (APOA) variable. The APOA level was set based on the application of the NARA methodology to subjectively determine an appropriate value, nominally between 0 (no effect) and 1 (maximum effect). However, based on the guidance available for NARA, it was decided to cap the maximum APOA associated with the EPC to 0.1.

**Table 1: Maximum APOA Value Caps for the EPC associated with Attention Management**

| *GDF Effect Path* | *NARA EPC Reference* | *NARA EPC Description* | *Maximum APOA Cap* | *Justification* |
|---|---|---|---|---|
| Attention Management | EPC No. 15 | Poor environment | 0.1 (low) | The task environment is generally benign even in the presence of GDFs. An APOA of 1.0 would represent an extreme physical environment in which crew are required to wear PPE. |

NARA was also used to determine a baseline human error rate (Nominal HEP) to set the task performance HEP unaffected by GDF exposure. One of the limitations of the application of NARA in this context was highlighted by the selection of the task from the predefined list within NARA from which the HEP was established to represent task performance associated with the hazard exposure. To limit the complexity of the model, a single GTT was sought to represent all relevant navigational tasks performed by the OOW that are important in managing collision or grounding risk. The GTT that is most analogous is:
*Task C1 – Simple response to alarms/indications providing clear indication of situation (Simple diagnosis required) Response might be direct execution of simple actions or initiating other actions separately assessed. (Nominal HEP = 0.0005)*

A second GTT was identified to account for possibility that a helmsman may also be present. In this case the helmsman is steering the ship based on verbal instructions communicated by the OOW. The GTT that is most analogous is:
*Task D1 - Verbal communication of safety critical data*

While having a helmsman present may introduce the possibility of a miscommunication error with the OOW, NARA also recognises a mitigating effect of a team. The NARA Human Performance Limiting Value for 'Actions taken by a team of operators' was used to cap the potential error rate at 1E-4 for the condition where a helmsman is present. The same value is taken for the probability of potential error of not performing evasive action by another ship involved in the encounter.

The NARA calculation allows inclusion of multiple EPCs and an Extended Time Factor (ETF). In this risk model for collision and grounding, GDFs are represented using only one EPC and there is little justification to include the ETF. Thus, the HEP was calculated based on the following formula:

$$\mathrm{HEP} = \mathrm{GTT} \times [(\mathrm{EPC}-1) \times (\mathrm{APOA} + 1] \tag{2}$$

## 5. RISK MODEL DEVELOPMENT

In this section we briefly present the process of risk model development, which is seen as translation of background knowledge into a coherent structure, which allows inference and decision-making under uncertainty. Also the results of the quantitative uncertainty assessment are presented in this section. Two risk models are developed, as depicted in figures 3 and 4, for collision and grounding respectively.

### 5.1. Risk model quantification

The model starts with the evaluation whether or not the levels of GDFs exceed the threshold. The GDFs may have either an acute or diffuse effect. In case of an acute effect, the threshold refers to the GDFs level at which an individual may be unable to physically compensate for GDF exposure and perform actions as intended. In case of diffuse effect, the threshold refers to the amount of motion, vibration or noise an individual can endure before it acts as a stressor (with a corresponding stress response). The exact value of the threshold will vary between individuals and it is dependent upon previous experience, exposure duration and sensitivity. If the thresholds are not exceeded, then the attention of a navigator is not affected, otherwise the attention management capability is degraded. Attention management is the supervisory human capability that directs, allocates and regulates the attentional resources required to perform various tasks. This high-level supervisory capability manages lower-level tasks such as perception, cognition, decision-making, memory, fine motor control and locomotion.

Representing GDF exposure effects on safety behaviour via the attention management path provides a structure compatible with the introduction of an Error Producing Conditions (EPC) using NARA.
In the presented model a variable called *C1 - Detection, Assessment and execution of simple actions* (NARA Generic Task Type – GTT No. C1) determines whether or not the performance of detection, assessment and actions required to avoid collision or grounding is sufficient or insufficient. This is an all-encompassing definition including whatever tasks are required to maintain situational awareness and to respond appropriately to avoid collision or grounding.
The probability of insufficient 'Detection, assessment and execution of simple actions' is calculated based on the integration of GDF Physical Effect and Attention Management Capability nodes. The calculation of Attention Management Capability of insufficient performance is performed using the NARA calculation.

The GDF Physical Effect on the probability of insufficient safety behaviour is assumed to be very weak (p=0.001) when Attention Management Capability (AMC) is *normal*. However, this effect is represented as being more significant in combination with *a degraded* AMC (p=0.0011), thus representing an additional drain on cognitive resources compensating for physical task disruption. These values were estimated using judgement.
To set the probability of *insufficient* performance unaffected by GDF exposure to reflect a baseline error rate the NARA GTT value for Task C1 of 0.0005 is used. The HEP of 0.0006 was calculated using NARA given exposure to EPC No. 15 representing the effect of GDF exposure via the AMC node.

If evasive manoeuvres are performed with the presence of helmsman, which requires appropriate communication of information between the OOW and helmsman, the node *D1 - verbal communication of safety critical data* is evaluated. This node (NARA GTT No. D1) determines whether or not the verbal communication of safety critical data to the helmsman required to avoid collisions is sufficient or insufficient. This node is only active if the *Helmsman present* node is set to *Yes* to reflect the fact that the OOW is not controlling the vessel directly. The NARA GTT value for Task D1 of 0.006 is used to set the probability of *insufficient* performance unaffected by GDF exposure to reflect a baseline error rate. The HEP of 0.0072 was calculated using NARA given exposure to EPC No. 15 representing the effect of GDF exposure via the *Attention Management Capability* node. It is assumed that *Verbal communication of safety critical data* is unaffected by the *GDF Physical Effect*.

The probability that OOW successfully executes actions required to manoeuvre the vessel to take evasive action is assessed in the node *Evasive Action*. It is assumed that if *Detection, assessment and execution of simple actions* and *Verbal communication of safety critical data*, where applicable, are performed sufficiently then *Evasive Action* will be executed. If *Helmsman present* is *Yes* then the HEP for *Evasive Action* being in a state *Not executed* is capped at 1E-4 in line with the NARA Human Performance Limiting Value for 'Actions taken by a team of operators'.

*Technical failure* node quantifies the probability of the relevant systems not functioning as a result of lack of maintenance or poor maintenance caused by the GDFs. This node was included in recognition of the importance of maintenance in sustaining the functionality of vessel equipment such that it performs as it is designed to. Errors during maintenance on systems that provide the manoeuvring capability of the vessel can limit the vessel's response to control inputs associated with evasive action, hence affecting the probability of an unwanted outcome. *Technical failure* node determines whether or not maintenance actions performed on equipment that provides the vessel's manoeuvring capability has been completed successfully or not. The probability of insufficient *Maintenance Task Performance* is calculated based on the integration of *GDF Stressor Technical* and *Attention*

*Management Capability-2* nodes. To set the probability of *insufficient* performance unaffected by GDF exposure to reflect a baseline error rate the NARA GTT value for Task C1 of 0.0005 is used. The HEP of 0.0006 was calculated using NARA given exposure to EPC No. 15 representing the effect of GDF exposure via the *Attention Management Capability-2* node.

The GDF Physical Effect on the probability of insufficient safety behaviour is not anticipated. This comes from an assumption about lack of preventive maintenance carried out if the levels of GDFs are above thresholds. This is in line with most of the operations guidelines of ships, where crews abstain from certain tasks if the weather conditions do not allow for safe performance of these tasks.

*Evasive action of another ship* is a node which accounts for the behaviour of OOW on a ship that is encountered by the own vessel. The probability for this node being in a state *Not executed* is capped at 1E-4 in line with the NARA Human Performance Limiting Value for ‘Actions taken by a team of operators’. The last node of the model is *Collision,* and it exists in the state *Yes*, if and only if:

- *Evasive Action of Another Ship=No, and Evasive Action=No, Technical Failure=Yes or No,*
- *Evasive Action of Another Ship=No, Evasive Action=Yes* and *Technical Failure=Yes.*

Otherwise, the node *Collision* is set to its state *No*.

## 5.2. Sensitivity, uncertainty and importance assessment of the risk model

The sensitivity analysis is performed to identify the essential variables that have the highest impact on the outcome of the model. For this purpose, every conditional and prior probability in the BBNs is systematically varied in turn while keeping the others unchanged. This allows the effects on the output probabilities computed from the network to be examined. To determine the sensitivity of an output variable to a given parameter of the model a sensitivity function is estimated for each single variable. This function describes outcome of the model as a function of the parameter *z*, which takes the following form:

$$\boldsymbol{z = p(Y = y_i | \pi)} \tag{3}$$

Where $y_i$ is one state of a network variable Y, and π is a combination of states for *Y*’s parent nodes. The sensitivity values were obtained using dedicated tools, which are implemented in the software package GeNIe, used for the development of the models presented here. Based on the findings from the sensitivity analysis, the following can be concluded:

- Both models are highly sensitive to the following parameters: *Maintenance Task Performance*, *C1 - Detection, Assessment and execution of simple actions and D1 - verbal communication of safety critical data*.
- The model assessing the probability of collisions is also sensitive to *Evasive action of another ship and Helmsman present.* However, the effect that these parameters have on the output is significantly lower than the effects of *C1* and *D1*, as specified above.
- The model assessing the probability of grounding is sensitive to *Helmsman present.*
- The remaining nodes have very low sensitivity values, meaning that their effects on the models outputs are rather minor.

Secondly, the evidential uncertainty assessment has been carried out on the mot sensitive model parameters. To rank the uncertainty, the following qualitative scoring system is applied, see [28]:

*Significant uncertainty*

One or more of the following conditions are met:

- The phenomena involved are not well understood; models are non-existent or known/believed to give poor predictions.
- The assumptions made represent strong simplifications.
- Data are not available, or are unreliable.
- There is lack of agreement/consensus among experts.

*Moderate uncertainty*

Conditions between those characterising significant and minor uncertainty, e.g.:

- The phenomena involved are well understood, but the models used are considered simple/crude.
- Some reliable data are available.

*Minor uncertainty*

All of the following conditions are met:

- The phenomena involved are well understood; the models used are known to give predictions with the required accuracy.
- The assumptions made are seen as very reasonable.
- Much reliable data are available.
- There is broad agreement among experts.

**Table 2: The qualitative assessment of evidential uncertainty for models assessing the probability of an accident**

| ***Model parameter*** | ***Justification for the evidential uncertainty score*** | ***Evidential uncertainty score*** |
|---|---|---|
| Maintenance Task Performance<br>C1 - Detection, Assessment and execution of simple actions | This node represents the performance of navigation tasks critical in collision or grounding avoidance and provides a structure compatible with the introduction of a NARA GTT, potentially affected by EPC No. 15 via 'Attention Management Capability' and GDF Physical effects | Moderate |
| D1 - verbal communication of safety critical data | This node represents the communication of vessel manoeuvring instructions critical in collision or grounding avoidance with a helmsman present and introduction of a NARA GTT, potentially affected by EPC No. 15 via 'Attention Management Capability'. | Moderate |
| Evasive action of another ship | The node represents the performance of navigation tasks critical in accident avoidance on board other ship. It is quantified based on NARA. | Moderate |
| Helmsman present | At the moment this node is quantified fully based on judgement. However, more detailed assessment is possible, by performing survey among shipping companies. | Moderate |

Finally, by combining the results of sensitive and uncertainty assessment, the parameter importance ranking is carried out. It allows for screening the model for both uncertain and sensitive parameters. For the models presented here, three parameters have high importance score, namely *Maintenance Task Performance*, *C1-Detection, Assessment and Execution of Simple Action, D1 – verbal communication on safety critical data.* This means, that states of these parameters should be carefully selected for any analysis. Moreover, when making the comparison between two designs of a ship, these two parameters shall not be changed, otherwise, their effect may outshine the effect of GDFs, which is much weaker.

**Table 3: The qualitative assessment of model parameters importance for models assessing the probability of an accident.**

| ***Model parameter*** | ***Evidential uncertainty score*** | ***Sensitivity score*** | ***Importance score*** |
|---|---|---|---|
| Maintenance Task Performance | Moderate | High | High |
| C1 - Detection, Assessment and execution of simple actions | Moderate | High | High |
| D1 - verbal communication of safety critical data | Moderate | High | High |
| Evasive action of another ship | Moderate | Moderate | Moderate |
| Helmsman present | Moderate | Moderate | Moderate |

## 6. CONCLUSION

The models presented in this paper offer a novel, evidence-based approach to modelling risk of ship-ship collision and grounding. They provide a flexible framework that could readily be extended to encompass the actions of third parties and mechanical failures in the future. The flexibility to extend the model's application is provided by the causal mechanism represented within the model that describes occurrence of an accident as the result of insufficient performance of an individual when exposed to hazardous situation.

The models focus on modelling improper performance in critical situations. This is also compatible with the general conceptualisation of human error within the Human Factors (HF) domain and its relationship to task performance.

As expected, the paucity of data on GDF effects presented a particular challenge. However, attention management theory successfully provided a means to represent the mechanism by which ship motion, noise and WBV affect cognitive performance. Due to the supervisory role attention management has in human cognition, this approach may also be readily generalizable. However, caution is required to ensure that this framework can be applied to model other performance shaping factors outside of the GDFs. It should be noted that DAM, CCM and MART are theories that best fit the current scientific data on GDFs, but more research is required to understand whether the integration of these three models into a combined model of 'attention management theory' is a robust and valid representation of human performance within different complex environments.

The integration of HRA (specifically NARA) to support the calculation of HEPs within the risk models has clear positive and negatives. On the one hand, it provided a facility to generate 'reasonable' HEPs using a well-known method, which would not have been possible otherwise. On the other hand, the application of NARA to physical tasks associated with physical effect of the GDFs on *Detection, Assessment and Execution of Simple Actions*, is stretching its application to, and perhaps beyond, its limit.

Despite the limitations and the paucity in data supporting certain hypotheses, the application of BBNs as a modelling tools, allows for clear representation of the modelled problem and comprehensive distribution of all the recognised uncertainties. By adopting BBNs and performing the importance analysis, we learned that the crucial elements of the models are the nodes, where the human error probabilities are quantified. Whereas the detailed quantification of the levels of GDFs associated with a given ship design or their effect on the attention management capability is less important.

The inherent feature of BBNs, two-ways reasoning, allows not only forward propagation of the evidences resulting in an outcome, but also the back propagation of the evidences, and estimation of the input variables, given a selected state of the output is possible.

Finally, comparative assessment of vessel designs based on manipulation of the GDF input nodes is possible in principle. The models are responsive to changes in the GDF nodes as expected. The models may be used by naval architects, vessel designers, and vessel system designers as intended, provided access to HF expertise is available to assist with application and interpretation. It is important to recognise the relevance of human factors input during its eventual application. HF provides the understanding of the complexities of human behaviour in operational settings, its interdependencies and interactions.

**Figure 3: Risk model for ship-ship collision.**

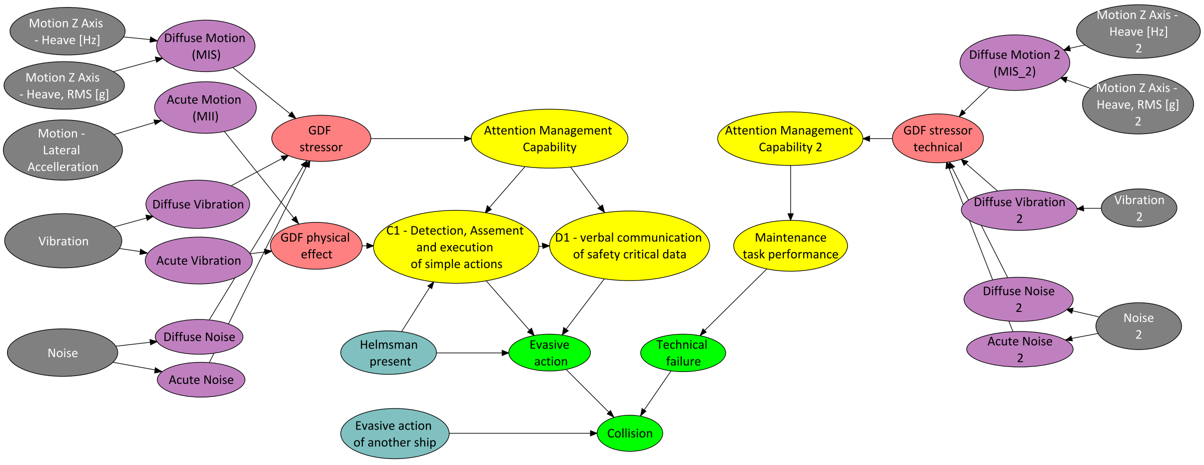
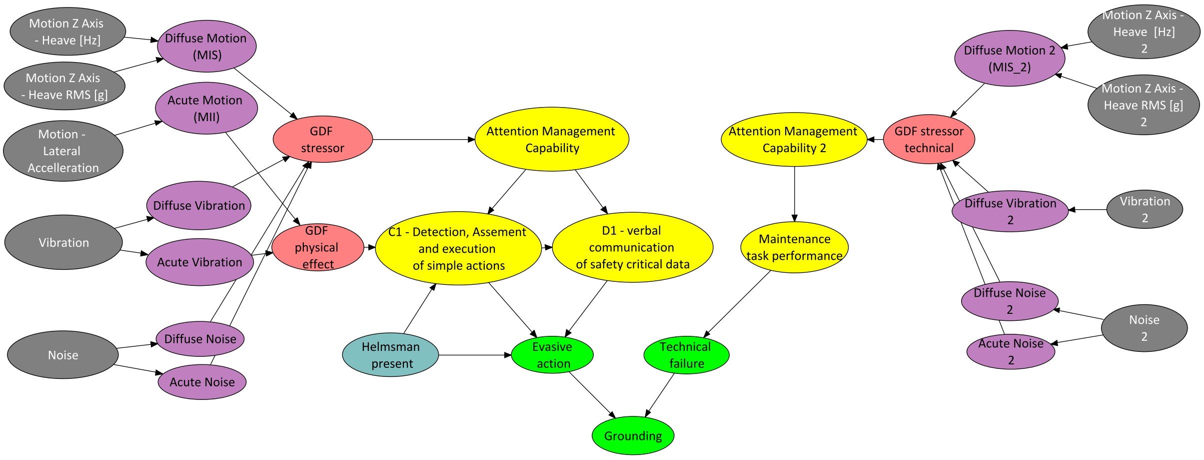


**Figure 4: Risk model for ship-ship grounding.**



## Acknowledgements

The authors appreciate the financial contributions of the EU, as this research was co-funded by the FAROS project (2012–2015). The probabilistic models introduced in this paper were created using the GeNie modeling environment developed at the Decision Systems Laboratory, University of Pittsburgh, available from http:// genie.sis.pitt.edu/.

## References

[1] A. Papanikolaou, Ed., *Risk-Based Ship Design: Methods, Tools and Applications*. Springer, 2009, p. 376.

[2] J. Montewka, S. Ehlers, F. Goerlandt, T. Hinz, K. Tabri, and P. Kujala, "A framework for risk assessment for maritime transportation systems—A case study for open sea collisions involving RoPax vessels," *Reliab. Eng. Syst. Saf.*, vol. 124, pp. 142–157, 2014.

[3] P. T. Pedersen, "Review and application of ship collision and grounding analysis procedures," *Mar. Struct.*, vol. 23, no. 3, pp. 241–262, Jul. 2010.

[4] A. Mazaheri, J. Montewka, and P. Kujala, "Modeling the risk of ship grounding—a literature review from a risk management perspective," *WMU J. Marit. Aff.*, vol. 12, Dec. 2013.

[5] S. Li, Q. Meng, and X. Qu, "An Overview of Maritime Waterway Quantitative Risk Assessment Models," *Risk Anal.*, vol. 32, no. 3, pp. 496–512, 2012.

[6] DNV, "Formal Safety Assessment - Large Passanger Ships, ANNEX II: Risk Assesment - Large Passenger Ships - Navigation," 2003.

[7] M. Hänninen and P. Kujala, "Influences of variables on ship collision probability in a Bayesian belief network model," *Reliab. Eng. Syst. Saf.*, vol. 102, no. 0, pp. 27–40, Jun. 2012.

[8] M. C. Leva, P. Friis-Hansen, E. S. Ravn, and A. Lepsoe, "SAFEDOR: A practical approach to model the action of an officer of the watch in collision scenarios," in *European safety and reliability conference; Safety and reliability for managing risk; ESREL 2006*, 2006, pp. 2795–2804.

[9] M. R. Martins and M. C. Maturana, "Application of Bayesian Belief networks to the human reliability analysis of an oil tanker operation focusing on collision accidents," *Reliab. Eng. Syst. Saf.*, vol. 110, pp. 89–109, Feb. 2013.

[10] M. C. Kim, P. H. Seong, and E. Hollnagel, "A probabilistic approach for determining the control mode in CREAM," *Reliab. Eng. Syst. Saf.*, vol. 91, no. 2, pp. 191–199, Feb. 2006.

[11] S. Kivimaa, A. Rantanen, T. Nyman, D. Owen, T. Garner, and B. Davies, "Ship motions, vibration and noise influence on crew performance and well-being studies in FAROS project," in *Transport Research Arena Conference TRA*, 2014.

[12] P. A. Hancock, "A dynamic model of stress and sustained attention," *Hum. Factors*, vol. 31, no. 5, pp. 519–537, 1989.

[13] G. Robert and J. Hockey, "Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework," *Biol. Psychol.*, vol. 45, no. 1, pp. 73–93, 1997.

[14] M. S. Young and N. A. Stanton, "Malleable attentional resources theory: a new explanation for the effects of mental underload on performance.," *Hum. Factors*, vol. 44, no. 3, pp. 365–75, Jan. 2002.

[15] A. Darwiche, *Modeling and Reasoning with Bayesian Networks*, 1st ed. Cambridge University Press, 2009.

[16] J. Montewka, F. Goerlandt, and P. Kujala, "On a systematic perspective on risk for Formal Safety Assessment (FSA)," *Reliab. Eng. Syst. Saf.*, vol. 127, no. July, pp. 77–85, Apr. 2014.

[17] T. Aven and O. Renn, "On risk defined as an event where the outcome is uncertain," *J. Risk Res.*, vol. 12, no. 1, pp. 1–11, Jan. 2009.

[18] F. Goerlandt and J. Montewka, "A probabilistic model for accidental cargo oil outflow from product tankers in a ship-ship collision," *Mar. Pollut. Bull.*, vol. 79, no. 1–2, pp. 130–144, 2014.

[19] H. S. Blackman, D. I. Gertman, and R. L. Boring, "Human error quantification using performance shaping factors in the SPAR-H method," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2008, vol. 52, no. 21, pp. 1733–1737.

[20] J. Montewka, "Summarizing literature review (Human Factors in Risk-Based Ship Design Methodology, Project no 314817)," 2014.

[21] D. M. Jones, W. J. Macken, and N. A. Mosdell, "The role of habituation in the disruption of recall performance by irrelevant sound," *Br. J. Psychol.*, vol. 88, no. 4, pp. 549–564, Nov. 1997.

[22] P. Crossland and K. J. N. . Rich, "A method for deriving MII criteria," in *Human factors in ship design and operation; RINA international conference*, 2000.

[23] G. E. Conway, J. L. Szalma, and P. A. Hancock, "A quantitative meta-analytic examination of whole-body vibration effects on human performance.," *Ergonomics*, vol. 50, no. 2, pp. 228–45, Feb. 2007.

[24] D. Owen, S. Pozzi, and J. Montewka, "Validation results and amendments made." p. 18, 2014.

[25] E. Hollnagel, *Cognitive Reliability and Error Analysis Method (CREAM)*. Elsevier Science, 1998, p. 302.

[26] X. He, Y. Wang, Z. Shen, and X. Huang, "A simplified CREAM prospective quantification process and its application," *Reliab. Eng. Syst. Saf.*, vol. 93, no. 2, pp. 298–306, Feb. 2008.

[27] C. Spitzer, U. Schmocker, and V. N. Dang, Eds., *Probabilistic Safety Assessment and Management*. London: Springer London, 2004.

[28] R. Flage and T. Aven, "Expressing and communicating uncertainty in relation to quantitative risk analysis," *Reliab. Risk Anal. Theory Appl.*, vol. 2, no. 13, pp. 9–18, 2009.