FLOODSTAND-deliverable:

Uncertainty bounds on time to capsize models

Authors	Qi Chen
Organisation	SSRC
Revision	2.0
Deliverable No.	D4.5
Date	02 February 2012

<table>
<tr><td colspan="2">Document identification sheet</td></tr>
<tr><td colspan="2">FLOODSTAND Integrated Flooding Control and Standard for Stability and Crises Management FP7-RTD- 218532</td></tr>
<tr><td>Title: Uncertainty bounds on time to capsize models</td><td rowspan="2">Other report identifications:</td></tr>
<tr><td>Investigating partners:

Authors: Qi Chen (SSRC)

Reviewed by: Andrzej Jasionowski</td></tr>
<tr><td>[ ] Outline
[ ] Draft
[x] Final
Version number: 2.0
Revision date: 02 February 2012
Next version due:
Number of pages: 74</td><td>[x] A deliverable
[ ] Part of a deliverable
[ ] Cover document for a part of a deliverable
[ ] Deliverable cover document
[ ] Other
Deliverable number: D4.5

Work Package: WP4
Deliverable due at month: 30</td></tr>
<tr><td rowspan="2">Accessibility:
[x] Public
[ ] Restricted
[ ] Confidential (consortium only)
[ ] Internal (accessibility defined for the final version)</td><td>Available from:
Distributed to: Public
Disclosees when restricted:</td></tr>
<tr><td>Comments:</td></tr>
<tr><td colspan="2">Abstract:
This report presents a new predictive probabilistic approach for characterizing the stochastic process of ship capsize by deploying Bayesian Inference techniques combined with Markov Chain Monte Carlo algorithm for Bayesian computation. This application is based on benchmark data performed in Task 4.1 and other existing experimental data available. A new inference methodology allowing the quantification of uncertainty associated with the prediction of the stability deterioration process is elicited.</td></tr>
</table>

Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 218532. The financial support is gratefully appreciated.

CONTENTS

Page

1. EXECUTIVE SUMMARY

The purpose of this task is to establish a standard uncertainty analysis scheme which is to be applied for the methods that have been put forward in Tasks 4.2, 4.3 and 4.4 for modelling of the stochastic process of ship capsize after flooding occurrence.

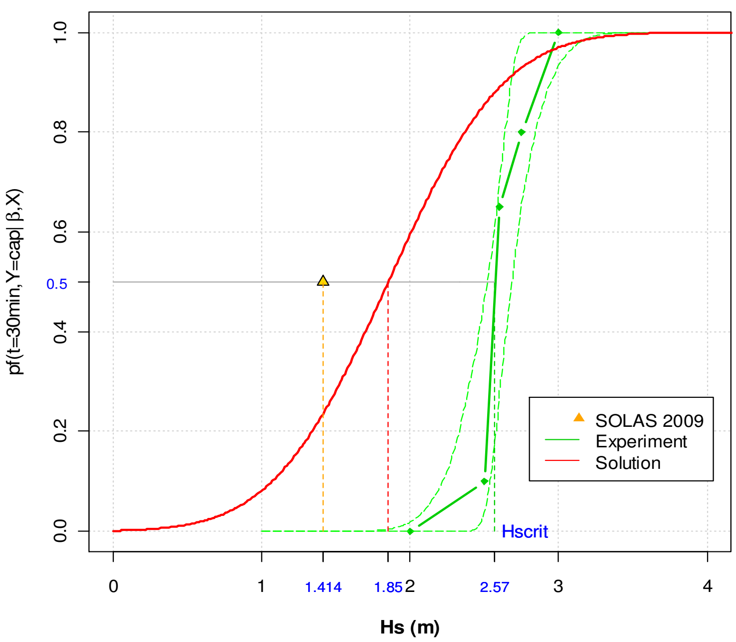
The assessment of ship stability deterioration process can be performed through stability standards calculation (i.e. standards in SOLAS), analytical model, performance-based numerical simulations and benchmarking model experiments. The discrepancies among these predictive applications have been observed even for identical damage scenarios of a specific ship. Thus the variation of the predictions from the physical model tests as well as the spread of the results of different approaches needs to be quantified and ultimately uncertainty bounds should be assigned to identify deviations.

For better describing the physical phenomenon of ship capsizing, the experimental data as documented in Task 4.1, has been given greater assent as one of the most reliable sources of information for the proposed method, which allows for quantifying inherent uncertainties associated with the survivability assessment process.

The solution used for explicit and continuous uncertainty quantification is as follows:

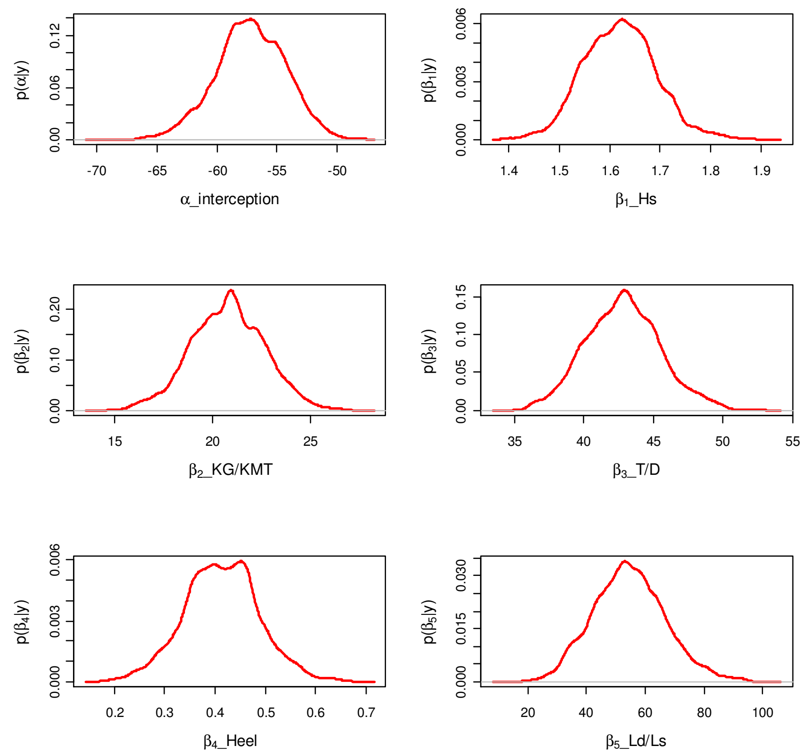
$$P(t_0, Y = cap) = \Phi(-56.54 + 1.64H_s + 21.18\frac{KG}{KMT} + 43.42\frac{T}{D} + 0.43Heel + 38.83\frac{L_d}{L_s})$$

This model provides the probability that a ship will capsize within 30 minutes when she is exposed to a specific flooding extent, given the continuous information of H_s, KG/KMT, T/D, $Heel$ and L_d/L_s. With respect to the damage case P6-7.1.0 as tested in Task 4.1, predictions of the deterioration process of ship stability derived from both theoretical methods and model experiments are shown as below:

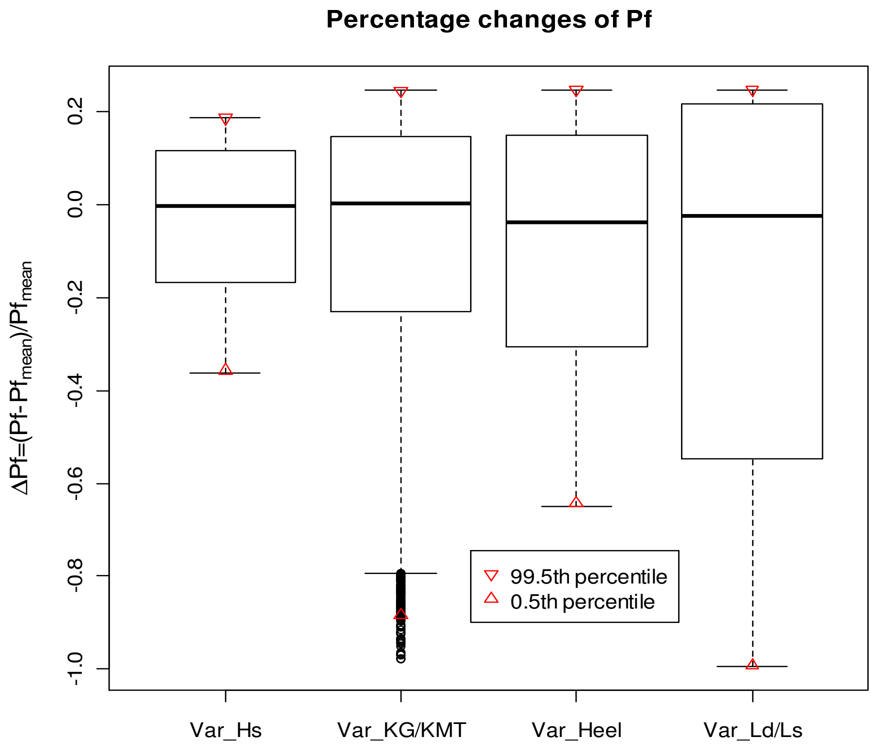


With the established model, it is possible to assess the uncertainty in the model output $P(t = 30min, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$, as governed by the uncertainties in continuous input parameters ($\boldsymbol{X}$= H_s, KG/KMT, T/D, *Heel*, L_d/L_s). The uncertainty in $\boldsymbol{X}$ is specified as the probability density function. The uncertainty in $P(t_0, Y = cap \,|\, \boldsymbol{\beta}, \boldsymbol{X})$ is calculated by propagating uncertainties in $\boldsymbol{X}$. With the identified Probit regression model for Bayesian data analysis, model coefficients $\boldsymbol{\beta}$ are regarded as the sensitivity indicator describing the impact of $\boldsymbol{X}$ on the outcome $P(t_0, Y = cap \,|\, \boldsymbol{\beta}, \boldsymbol{X})$.

Finding solution for unknown parameters $\boldsymbol{\beta}$ is the most critical step in quantifying the uncertainty propagation. Explicit uncertainty quantification is performed by establishing uncertainty bounds for each sensitivity coefficient. Bayesian inference and MCMC algorithm are adopted to approximate the posterior distributions of regressor indices $P(\boldsymbol{\beta}|y)$, as shown in the figure below.



Sensitivity study is subsequently conducted by ranking the contributions of individual inputs on the uncertainty in the output ΔP_f, as shown below.



It appears that the extent of flooding is the most critical information needed for assessment of criticality of flooding situation.

2. INTRODUCTION

The occurrence of ship flooding event gives rise to a major threat to life of all onboard, the environment as well as the property. Therefore, a fast and rational measurement of ship survivability following flooding is crucial for facilitating informed and quick decision making for possible aversion of the ultimate consequences.

Although new regulations governing state of ship stability, namely SOLAS2009 CH II provide with probabilistic models, these do not have sufficient resolution to reflect upon typical loss circumstances.

Therefore, as it has been explicitly explained in the deliverables of Task 4.2, 4.3, and 4.4, "time to capsize" can be considered as efficient and intuitive means to quantify ship survivability.

In order to achieve an optimal solution, the following principles should be adhered to:

- A predictive probabilistic model should be proposed with dominant input variables which are capable of describing key accidents characteristics included. The underlying uncertainty in the model outcome needs to be estimated easily and simultaneously so that a desired (relatively conservative) survivability level can be defined.
- Key influential variables linking to the ship behaviour at damaged conditions should be included and well-defined in the model. These refer to sea environment, intact conditions of the ship just before accidents, and the extent of flooding, etc. All the information should be ultimately cooperated with the survivability measure "time to capsize" and presented in probabilistic manner.
- The model should be derived through a plausible inference process using the existing experimental data assembled from a series of benchmarking tests conducted during several research projects (e.g. HARDER (HARDER, 2003), SAFEDOR (GL, 2002)), which are fully or partially supported by the European Commission. Only by doing so, it can offer the flexibility and reliability to assess any independent flooding case.
- As the aforementioned entails a methodology that is able to perform probabilistic model development and, importantly, to quantify the associated uncertainty, this can be readily employed for addressing similar time-critical accidental events in a probabilistic framework.

The report is structured as follows:

- The significance of this work concerning ship survivability assessment under the new probabilistic framework is highlighted in the introduction chapter.
- The decision-making procedure on the selection of appropriate multivariate models for model training is depicted.
- A unique model estimation methodology based on Bayesian inference and Markov Chain Monte Carlo algorithm is disclosed in detail. This seems as a tailored solution for addressing the issue of uncertainty quantification.
- The established model is deployed for measuring the rate of ship stability loss when the ship is exposed to a specific flooding event within given time. Two steps including model selection and model validation give a tangible representation of the proposed methodology.
- Extensive uncertainty analysis and sensitivity studies are conducted for better understanding of the contributions of individual inputs in the model to the uncertainty in the output. A solution is put forward for minimizing the uncertainty in the model output through quantifying uncertainties in a set of input parameters in the model.

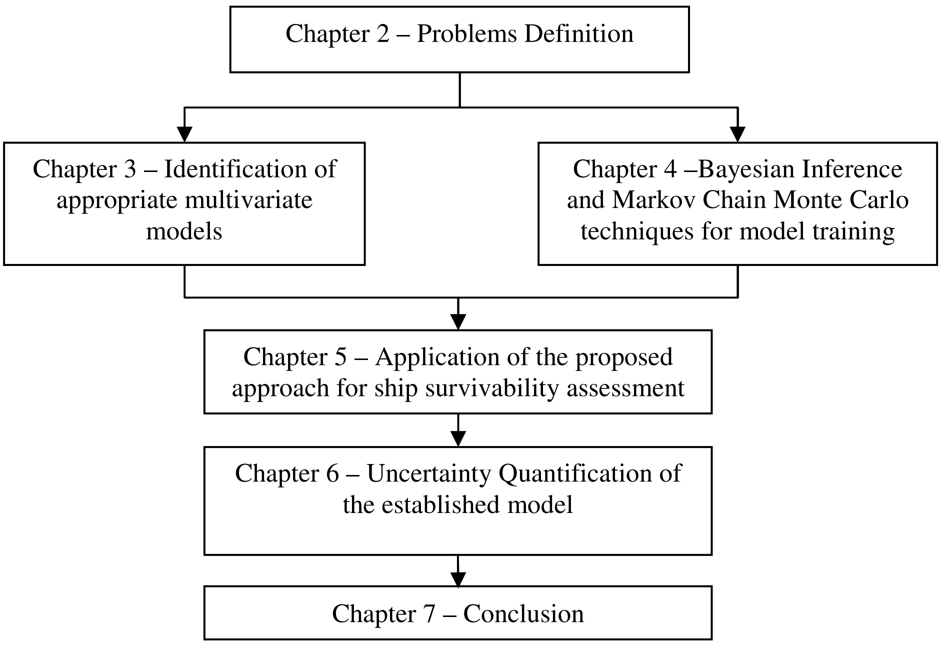


Figure 1: *Structure of the Report*

3. MODEL IDENTIFICATION

For the purpose of transforming the available information into interpretable and manageable knowledge, a proper mathematical model needs to be identified so that the collected benchmarking data can be properly utilized to assist uncertainty quantification. Traditionally, simple statistical approaches (e.g. liner models) are adopted to explain the correlations between the various variables. However, considering the characteristics of the data at hand, it is important to bear in mind: i) the output variable is binary (i.e. capsize or not capsize within a given period); ii) there is more than a handful of variables influencing the status of the output variable; iii) the output measure needs to be presented in a probabilistic manner to resonate well with the existing probabilistic safety framework in the maritime industry.

Considering the above, it becomes apparent that more sophisticated modelling techniques are needed in this case. Regarding this, the field of multivariate data analysis can be considered as a promising aspect to be further investigated. A classification of the various multivariate statistical techniques is presented in *Multivariate Data Analysis* (Hair, 2009). As it is revealed, the selection of the appropriate multivariate techniques for a specific research problem should be made on the basis of the research objectives and the nature of the data.

This chapter attempts to describe a decision procedure of selecting the appropriate multivariate statistical models for further training. The procedure comprises two steps:

- The first step deals with model preparation, which starts with a definition of the research problem and put forward a conceptual model describing the relationships to be examined.
- The second step addresses the identification and definition of specific multivariate models to be deployed for further training.

3.1 Preparation of Model Building up

The starting point for a multivariate analysis is to define the research problem and analysis objectives in conceptual terms before specifying any variables or measures. Considering the subject to be investigated, it is obvious that a dependent relationship between ship behaviour when subject to flooding and a list of important influential variables is sought to be established. These variables (independent) refer to:

- Loading conditions (e.g. draft, centre of gravity)
- Watertight architecture (e.g. subdivision arrangement, and watertight door status)
- Damage characteristics (e.g. location, extend)
- Sea environment (e.g. wave height and direction, wind)
- Etc.

Regarding the output (dependent) variable, an appropriate measure would be the indication of whether the ship capsizes within a given time interval. In this context, the outcome can be best described as a binary response, which is denoted by either 0 or 1 in the experiment.

3.2 Identification of Multivariate Models

Having the research objective specified, it is important to realize the measurement characteristics (format) of both dependent and independent variables recorded in the available data. In this respect, all the independent variables are measure metrically (i.e., quantitative data) while, in contrast, the dependent variable is nonmetric (i.e., qualitative data). Hence, on the basis of flow chart as shown in Figure 2, both multiple discriminant analysis and logistic regression techniques can be promising solutions for addressing the identified problems.

Nevertheless, comparing discriminant analysis and logistic regression techniques, the latter offer a platform where all types of independent variables (metric and nonmetric) can be accommodated. More importantly, the logistic regression technique is specifically designed to deal with binary dependent variables (Yes/No, True/False, Capsize/Survive, etc.). As a result, logistic regression technique has been selected for establish such a relationship between a single dependent variable of ship response and a list of different influential independent variables. The general form is shown in equation 1.

$$\underset{\text{(binary nonmetric)}}{Y} = \underset{\text{(nonmetric and metric)}}{x_1 + x_2 + x_3 + \cdots + x_n} \qquad (1)$$

Due to the categorical response $\boldsymbol{Y}$ is defined as a binary random variable (i.e. either survival or capsizing denoted by 1 or 0), such binary outcomes can be described by expressing them from a probabilistic perspective. In this case, the distribution of Y is specified by probability $P(Y = 1) = \pi$ of capsize and $P(Y = 0) = 1 - \pi$ of survive. This implies $\mu = E(Y) = \pi$. Hence, a reformulation of equation (1) is performed by considering the probability π as the output (dependent) variable, which is shown in equation (2).

$$\pi(x) = P(Y = 1|\boldsymbol{X}) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n = \alpha + \beta\boldsymbol{X} \qquad (2)$$

Where: i) the response variable $\boldsymbol{Y}$ has a binomial distribution
ii) Explanatory variables are represented by the design matrix $\boldsymbol{X}$
iii) $\boldsymbol{\alpha}$ is the intercept
iv) $\boldsymbol{\beta}$ represents a set of regression coefficients for all predictor variables and $\boldsymbol{\beta X}$ is defined as the linear predictor

As can be noted, equation (2) is a simple linear probability model. However, there is an inherent difficulty arising from structure defect to restrict the estimated probability π to be within the interval [0, 1]. In this situation, it would be very desirable to have a transformation function to connect the linear predictors on the left hand side with a probability function on the left hand size.

As a result, the aforementioned difficulties entail the interested model to be equipped with the features of the Generalized Linear Model (GLM) in statistic terms. A GLM allows the linear predictors to be correlated with the response variable via a link function (Agresti, 2007). The link function can be written as $g(.)$ which relates a function of the expectation to the linear predictors. A typical GLM is defined in equation (3).

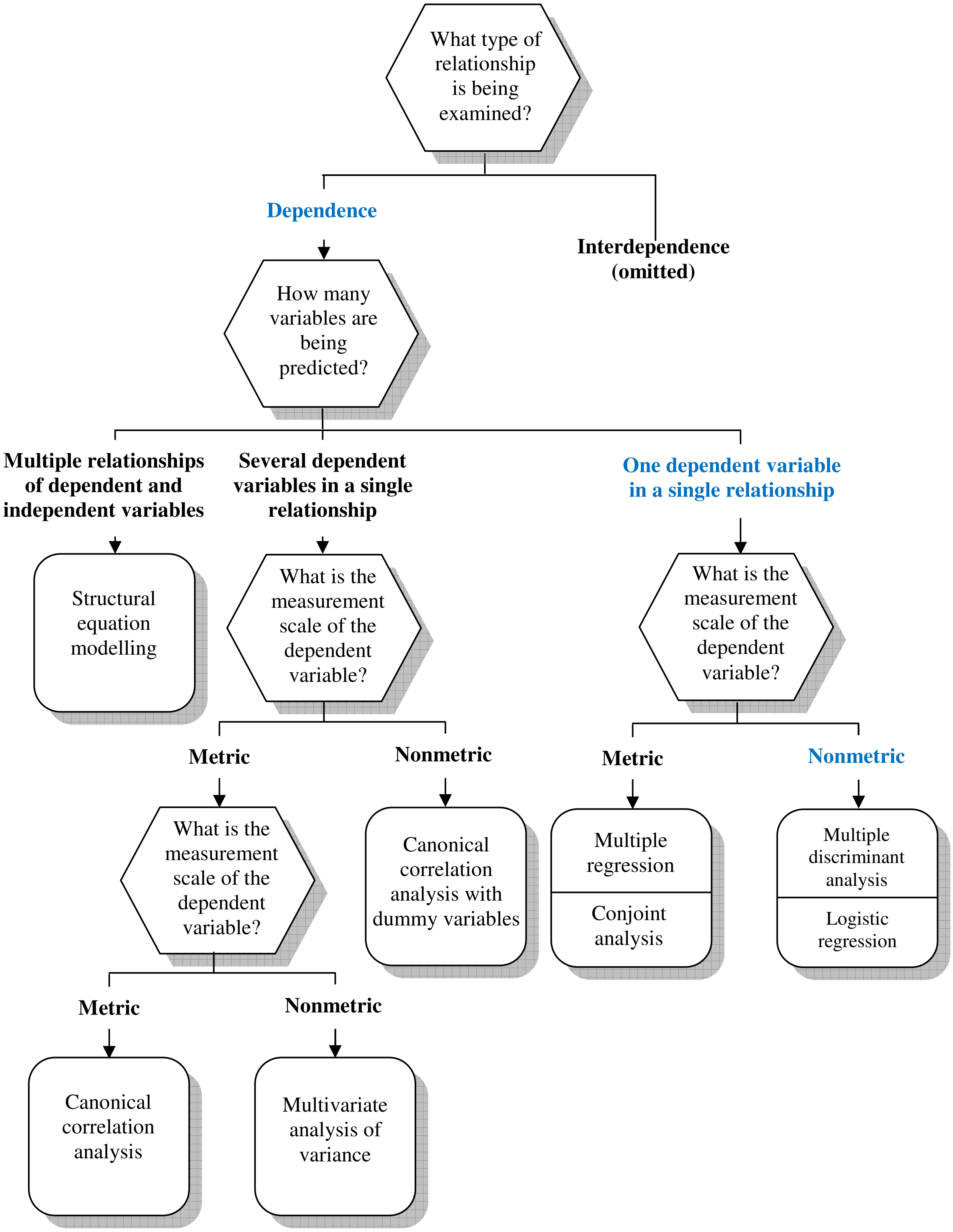


Figure 2 *Selecting a Multivariate Technique*

$$g[E(Y|X)] = g[\pi(x)] = \alpha + \beta X \tag{3}$$

To ensure that the predicted probabilities π fall into the 0-1 interval, it is often modelling with a cumulative probability distribution:

$$\pi = \int_{-\infty}^{t} f(s).ds$$

Where $f(s) \geq 0$ and $\int_{-\infty}^{\infty} f(s).ds = 1$. The selection of a proper probability density function $f(s)$ will impact on the type of GLMs to be used. In considering the defined question with binary outcomes, two common GLMs of probit model and logit model fulfil the requirements of providing necessary link functions. The probit model will be mainly discussed from this context, where $f(s)$ is assumed to be normal, with mean zero and variance σ^2, therefore:

$$\pi = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] ds = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{4}$$

Where Φ denotes the cumulative probability function for the standard normal distribution N (0, 1), thus the probit link function $g[\pi(x)]$ exists as an inverse function Φ^{-1}(.).Hence, in GLM form,

$$\Phi^{-1}[\pi(x)] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n = \alpha + \boldsymbol{\beta X} \tag{5}$$

Consequently, the ultimate selected multivariate model is depicted in equation (6):

$$\pi(x) = P(Y = 1|\boldsymbol{\beta}, \boldsymbol{X}) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi(\alpha + \boldsymbol{\beta X}) \tag{6}$$

Overall, the identified model offers the following key characteristics and benefits:

- The response curve for $\pi(x)$ [or for $1 - \pi(x)$, when $\beta < 0$] has the appearance of the normal CDF with mean $\mu = -\alpha/\beta$ and standard deviation $\sigma = 1/|\beta|$, which has a sigmoid profile as shown in Figure 3. This flexibility offers a unique platform for presenting the response object, "Ship behaviour" after damage based on a set of key random independent variables describing accident characteristics.
- The model can be easily extended to include more variables simultaneously in the linear predictors (i.e.$\beta\boldsymbol{X}$). This provides a convenient and interpretable means to perform sensitivity study the subsequent model adjustment and validation.

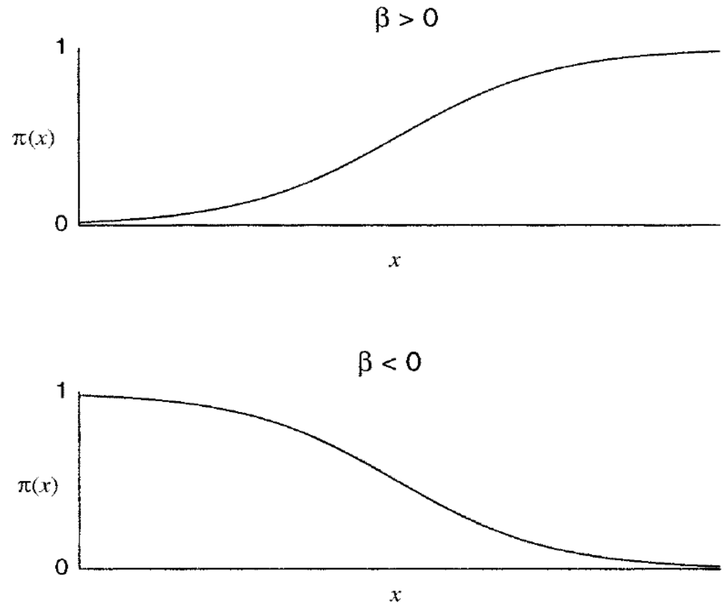


Figure 3 *Probit Regression Function*

4. BINARY REGRESSION MODEL ESTIMATION

The identified Probit model is a useful data analysis tool to tackle complex multivariate data analysis problems. This chapter focuses on detailing appropriate model estimation techniques. Classical model training technique (such as Maximum likelihood estimation) can be easily applied in this case. However, considering the one of the key objectives of this task is to quantify uncertainties associated with the estimations/predictions, a unique model training technique, Bayesian inference, is adopted for this study. As Bayesian inference technique draws the latest development in advance mathematical simulation techniques, (i.e. Markov Chain Monte Carlo (MCMC)), it has been increasingly recognized as a powerful means for inference making through probability distribution updating, model training.

Following a brief discussion of the underlying considerations for selecting Bayesian inference technique rather than maximum-likelihood approach, this chapter will elaborate in details on the Bayesian method and its corresponding application for Probit model training. Emphasis will be placed on how MCMC algorithm can be employed for approximating the posterior distribution of the interested variables (e.g. model coefficients in this case).

4.1 The Underlying Consideration

The methods based on maximum likelihood are proper to work out the parameters of generalized linear models. With the identified binary response regression model presented in equation (6), this dependent relationship can be established with ease once the model coefficients (i.e.α, β) to be estimated. The implementation of the maximum likelihood approach is generally based on Newton-Raphson algorithm to iteratively converge towards the maximum value of the likelihood function. For instance, an iterative weighted least squares procedure as depicted in (Charnes et al., 1976) is frequently applied for maximum likelihood estimation of model parameters.

On the other hand, the fundamental of Bayesian theory came first between the late 17^{th} and early 18^{th} centuries, but practical Bayesian analysis has only recently become available. This availability is mostly attributed to the MCMC methods and the greatly improved computing technologies in the last 20 years. A major limitation towards more widespread implementation of Bayesian approaches is that obtaining the posterior distribution often requires the integration of high-dimensional functions. This can be computationally very difficult, but the idea of MCMC simulation is in a sense to bypass the mathematical operations rather than to implement them.

Comparing maximum-likelihood and Bayesian inference approaches, it is very important to understand the merits and limitations of the each technique.

- The maximum-likelihood approach is a powerful and well-accepted model training technique. However, the quality of the outcome is heavily dependent on the size of the sampled data. In the case of large samples, the method can be very precise. However, the maximum likelihood estimation can be heavily biased with small numbers of sampled data. In the respect, it is important to bear in mind that the amount of experiment data can be collected regarding ship survivability is always restricted by the allocated resources (in terms of time and fund). Hence, the objective of reflecting the status of the ship survivability of thousands of ships sailing worldwide by using only a very limited experimental data in the maritime industry can be surely considered as a small sample size problem. And, moreover, it is expected that such difficulties will remain to be a challenge in the foreseeable future.

- In contrast, Bayesian inference approach is much less sensitive to the changes of the sample size. This is because the well-founded Bayes' theorem can naturally combine the evidence from both collected data and prior information. A major barrier for the wider adoption of Bayesian method in the past is due to the computation complexity (e.g. random sampling). However, such a barrier can be easily overcome with today's computation capability.

On the basis of the foregoing, it seems that Bayesian methods suit better to the intended study. It is acknowledged that practical Bayesian analysis sometimes is more complex. For instance, the size of the parameter space can become extremely large for problems involving multiple parameters, such as $(\boldsymbol{\theta} = \alpha, \beta_1, \ldots, \beta_n)$, hence the subsequent computation is massive, which make it infeasible to make the exact inference. Nevertheless, with the profoundly increased computing power in the last 20 years, the MCMC will be put forward to solve the previously intractable problems.

4.2 Bayes' rule

A key feature of Bayesian statistics is that Bayes' theorem synthesizes two separate sources of information about the interested parameters. The first source is the sampled data, expressed formally by the likelihood function. The second is the prior distribution, which represents additional information that is available beforehand. Figure 4 shows a schematic representation of this process. The result of combining the prior information and data in this way is the posterior distribution, from which inferences about parameters can be derived. The computation process is explained using Bayes' equation (Lee, 2004),

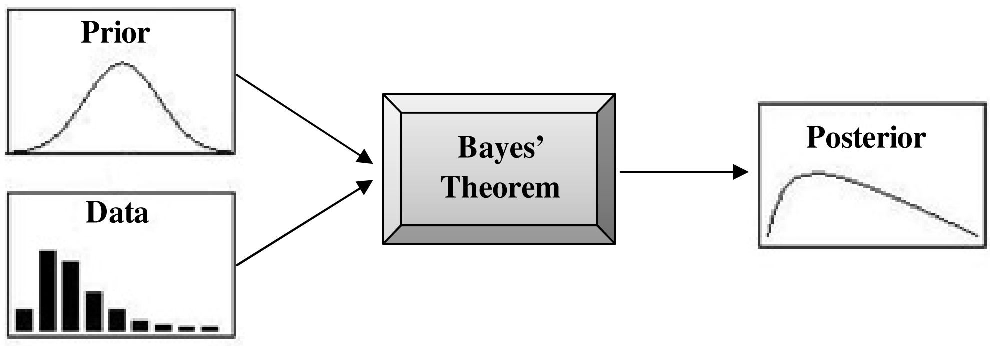


Figure 4 *Synthesis of information by Bayes' theorem*

$$P(\theta|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\theta)P(\theta)}{P(\boldsymbol{y})} = \frac{P(\boldsymbol{y}|\theta)P(\theta)}{\sum P(\boldsymbol{y}|\theta)P(\theta)} \tag{7}$$

Where θ can be either a single or a set of unknown parameters and **y** is the observed data. Then:

- The likelihood function $P(\boldsymbol{y}|\theta)$ is the conditional probability of data **y** depending on parameter θ, which is also the fundamental to frequentist inference.
- The prior distribution $P(\theta)$ is used only in the Bayesian approach. The assignment of the values should be based on the additional information that is available.
- If the range of possible values of θ is assumed to be discrete, summation can be performed.

In practice, Bayes' equation (7) can be simplified to,

$$P(\theta|\boldsymbol{y}) \propto P(\boldsymbol{y}|\theta)P(\theta) \tag{8}$$

Where the proportionality symbol $\propto$ expresses the fact that the product of the likelihood function and the prior distribution on the right hand side of the equation (8) must be scaled to integrate to 1 over the range of plausible θ values for it to be a proper probability distribution. The scaled product $P(\theta|y)$ is defined as the posterior distribution for θ given the data, and expresses what is now known about θ based on both the sample data and prior information. A better illustration of how Bayes' theorem works is Figure 5.

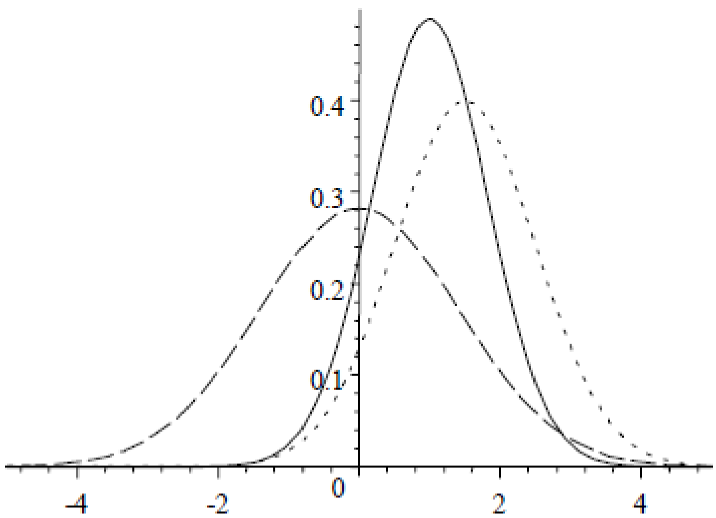


Figure 5 *Example of a triplot. Prior density (dashed), likelihood (dotted) and posterior density (solid)*

4.3 Markov Chain Monte Carlo simulation for Posterior Approximation

4.3.1 Metropolis-Hastings Algorithm

For a continuous parameter space θ, Bayes' formula in equation (7) for the posterior distribution is

$$P(\theta|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\theta)P(\theta)}{\int P(\boldsymbol{y}|\theta)P(\theta)d\theta} \tag{9}$$

This equation is often hard to calculate due to the integral in the denominator. Particularly if there are n unknown parameters, such as $(\boldsymbol{\theta} = \alpha, \beta_1, \dots, \beta_{n-1})$, then the denominator involves integration over n-dimensional parameter space which becomes intractable for large values of *n*. In this case a numerical method for calculating complex integrals and hence making inference about $\boldsymbol{\theta}$ is required. The method is called MCMC combines two methods: Monte Carlo integration and Markov chain sampling. As suggested by its name, there are two important features associated with MCMC techniques:

- MCMC has the feature of Monte Carlo integration method which simplifies a continuous distribution by taking discrete samples. It is useful when a continuous distribution is too complex to integrate explicitly but can readily be sampled (Gelman, 2004). For instance, there is a continuous distribution of $P(\boldsymbol{\theta})$ in Figure 6(a) which can be approximated by a histogram of the discrete samples in Figure 6(b). The larger number of samples leads to a closer approximation to the continuous distribution. Hence, getting back to the original problem is to evaluate the posterior distribution of model coefficients, if a histogram is a reasonable approximation to $P(\beta|\boldsymbol{y})$ for each coefficient, then any inference about β can be made by simply using the sampled values. Moreover, in Figure 6, it is able to sample θ from $P(\boldsymbol{\theta})$ directly, and hence numerous independent and identically distributed random θ can be generated, that is simple to obtain Monte Carlo approximations to posterior quantities. However, drawing samples from $P(\boldsymbol{\theta})$ is not always achievable because it may have a complex, or even unknown form.
- The other feature of MCMC methods is to allow drawing samples from the target density $P(\theta)$ follows a Markov chain. This method has the Markov property (Bartlett, 1978), which entails the next sample in the chain is dependent only on the previous sample. So in this case, a chain of samples $\left(\theta^{(1)}, \dots, \theta^{(n)}\right)$ can be built up after specifying a starting value $\theta^{(0)}$ (see Figure 7). Two of the most popular MCMC algorithms: the Metropolis-Hastings and the Gibbs samplers can be used for generating Markov chains to approximate the target probability distribution. The former will be put forward for further applications

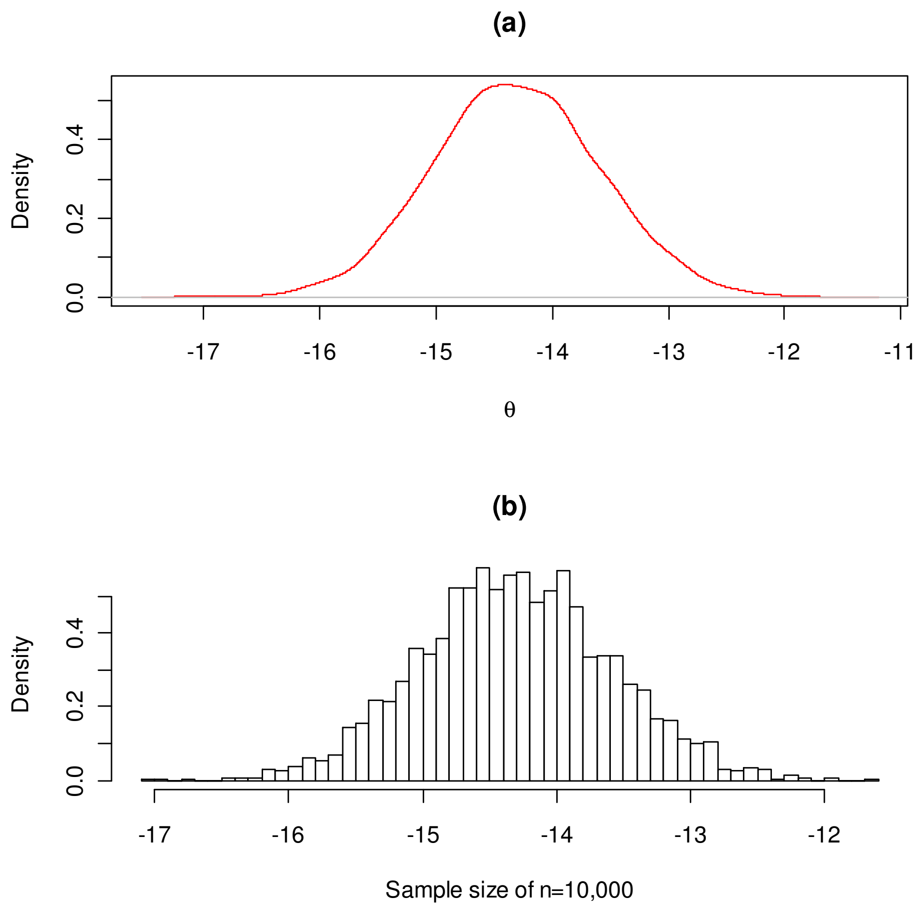


Figure 6 *(a) a continuous distribution of* $P(\boldsymbol{\theta})$*; (b) Approximating* $P(\boldsymbol{\theta})$ *using discrete random samples, for sample size of 10,000 and bin width of 0.1.*

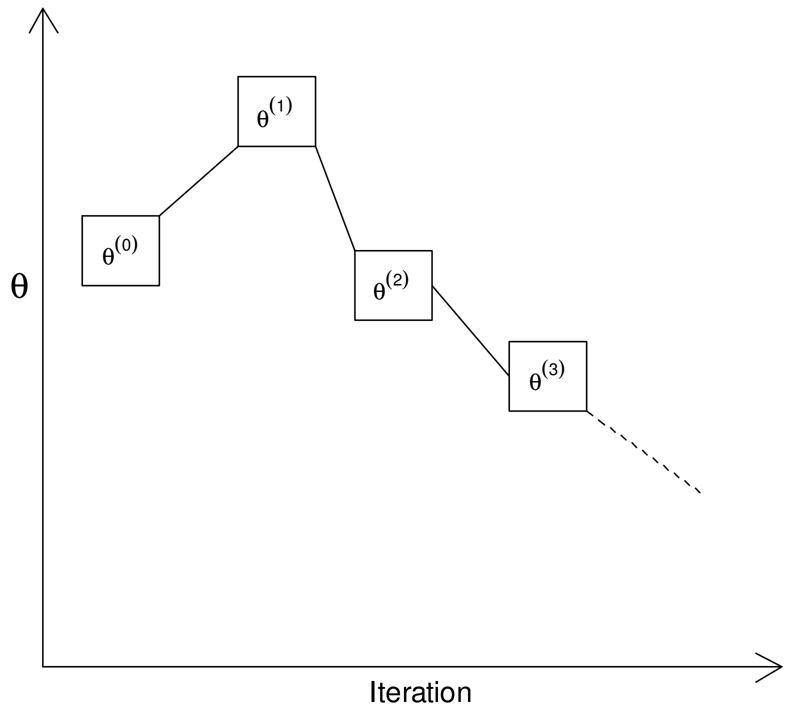


Figure 7 *A simple example of a Markov chain*

This section attempts to describe the Metropolis-Hasting algorithm is a generic method for approximating the posterior distribution of a model coefficient. $P(\boldsymbol{\beta}|y) = P(\boldsymbol{\theta}|y)$ is a joint posterior distribution because more than one predictor variables (i.e. coefficients $\alpha, \beta_1, \dots, \beta_n$) need to be approximated simultaneously. In this case, an iteration of the Metropolis-Hasting sampling is a complete cycle through every unknown parameter $\boldsymbol{\beta}^{(i)} = \left\{\alpha^{(i)}, \beta_1^{(i)}, \dots, \beta_n^{(i)}\right\}$. This sampler works by randomly proposing a new value $\boldsymbol{\beta}^*$. If $P(\boldsymbol{\beta}^*|y) > P\left(\boldsymbol{\beta}^{(i)}\middle|y\right)$ then the next value in the chain becomes the proposed value $\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^*$. If $P(\boldsymbol{\beta}^*|y) < P\left(\boldsymbol{\beta}^{(i)}\middle|y\right)$ then it seems $\boldsymbol{\beta}^*$should not necessarily be included, the previous value is retained $\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)}$. The decision of including $\boldsymbol{\beta}^*$ or not should be based on a comparison of $P(\boldsymbol{\beta}^*|y)$ to $P\left(\boldsymbol{\beta}^{(i)}\middle|y\right)$. Fortunately, this comparison can be made using an acceptance ratio γ even if $P(\boldsymbol{\beta}|y)$ is not achievable:

$$\gamma = \frac{P(\boldsymbol{\beta}^*|y)}{P(\boldsymbol{\beta}^{(i)}|y)} = \frac{P(y|\boldsymbol{\beta}^*)P(\boldsymbol{\beta}^*)}{P(y|\boldsymbol{\beta}^{(i)})P(\boldsymbol{\beta}^{(i)})} \tag{10}$$

The way of creating proposal values $\boldsymbol{\beta}^*$ is to add a random variable to the current values $\boldsymbol{\beta}^{(i)}$ using a proposal density $Q\left(\boldsymbol{\beta}^*\middle|\boldsymbol{\beta}^{(i)}\right)$. Usually $\boldsymbol{Q}$ can be taken considered as a symmetric attribute from a uniform distribution or a standard normal distribution. So that proposals closer to the current value are more likely. In practice, given $\boldsymbol{\beta}^{(i)}$, the Metropolis-Hasting algorithm generates a value $\boldsymbol{\beta}^{(i+1)}$as follows (Hoff, 2009):

1. Sample $\boldsymbol{\beta}^* \sim Q\left(\boldsymbol{\beta}^*\middle|\boldsymbol{\beta}^{(i)}\right)$;
2. Compute the acceptance ratio

$$\gamma = \frac{P(\boldsymbol{\beta}^*|y)}{P(\boldsymbol{\beta}^{(i)}|y)} = \frac{P(y|\boldsymbol{\beta}^*)P(\boldsymbol{\beta}^*)}{P(y|\boldsymbol{\beta}^{(i)})P(\boldsymbol{\beta}^{(i)})} \times \frac{Q\left(\boldsymbol{\beta}^{(i)}\middle|\boldsymbol{\beta}^*\right)}{Q(\boldsymbol{\beta}^*|\boldsymbol{\beta}^{(i)})} \tag{11}$$

3. Let $U \sim \text{uniform}[0,1]$, and setting

$$\boldsymbol{\beta}^{(i+1)} = \begin{cases} \boldsymbol{\beta}^*, & \text{if } U < \gamma \\ \boldsymbol{\beta}^{(i)}, & \text{otherwise} \end{cases}$$

For simulating the posterior probability of each coefficient in equation (5), $P(y|\boldsymbol{\beta})$ in equation (11) is the log-likelihood function of the probit model:

$$\ln \mathcal{L}(\boldsymbol{\beta}) = \sum \left(\boldsymbol{y} \ln \Phi(\boldsymbol{X}'\boldsymbol{\beta}) + (1 - \boldsymbol{y}) \ln\left(1 - \Phi(\boldsymbol{X}'\boldsymbol{\beta})\right)\right) \tag{12}$$

where $\boldsymbol{y}$ is the observed data of experiment. In step 2, the prior distribution is chosen from a multivariate normal distribution $P(\boldsymbol{\beta}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where covariance matrix $\boldsymbol{\Sigma}$ of multiple coefficients can be approximated by an inverse of the Hessian matrix $\boldsymbol{\Sigma} = \boldsymbol{\mathcal{H}}^{-1}$ (Yuen, 2010). Quasi-Newton methods are normally used to produce the inverse Hessian $\Delta\beta = -\mathcal{H}^{-1} \cdot f'(\beta)$. If the proposal density $\boldsymbol{Q}$ is symmetric (i.e. a multivariate normal), then γ in equation (11) is simplified to a likelihood ratio.

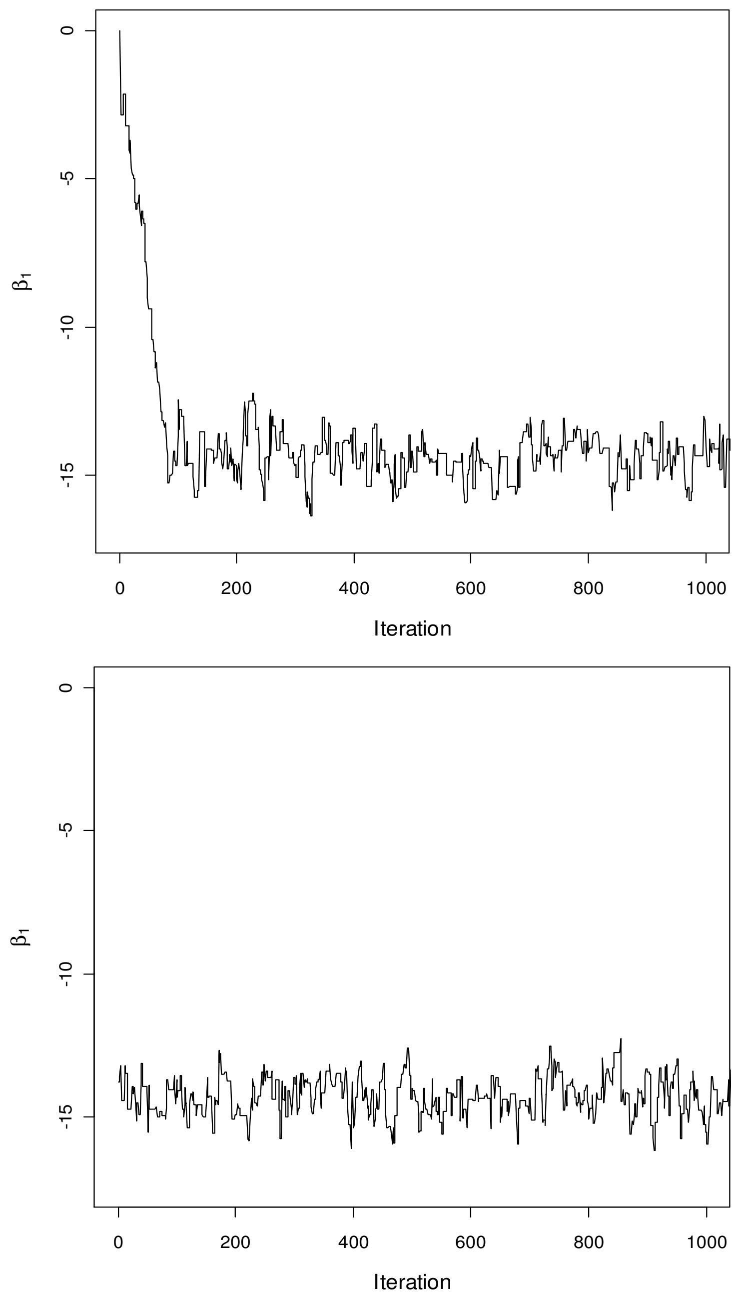


Figure 8 *Markov chain samples using Metropolis-Hastings sampling for* β_1

An example of Metropolis-Hasting samplings for β_1 is shown in Figure 8. The likelihood of $P(\beta_1 \mid \boldsymbol{y})$ is calculated complying with equation (12). The first 1000 Markov chain samples for β_1 are illustrated. Evidently the starting value $\beta_1^{(0)} = 0$ at the top plot is far from a reasonable estimate, the lack of knowledge about $\beta_1^{(0)}$ leads to such poor estimates at initial steps. The early estimates are not allowed for any inference for β_1 and usually to be defined as burn-in period. For example the bottom plot discards the first 500 iterations and shows an immediate convergence of the simulation to the correct value for β_1 graphically.

4.3.2 Estimation of Uncertainty Bounds

The previous section make obvious that the Markov chain samples provide the complete distributions when the chains have correctly converged, thus making inferences of unknown parameters are feasible. Figure 9 shows a histogram of β_1 using Metropolis-Hastings sampling. The first 1000 samples are shown in Figure 8. This histogram is based on discarding the first 500 samples (burn-in) and using a sample size of 10,000. A bell-shaped distribution is displayed and it is possible that the posterior distribution about β_1 is approximately Normal. Hence, various summary of β_1 can be computed. For example, the mean of $\beta_1$is estimated using the mean of the sampled value, as given in equation (13).

$$\overline{\beta} = \frac{1}{N}\sum_{i=1}^{N}\beta^{(i)} \tag{13}$$

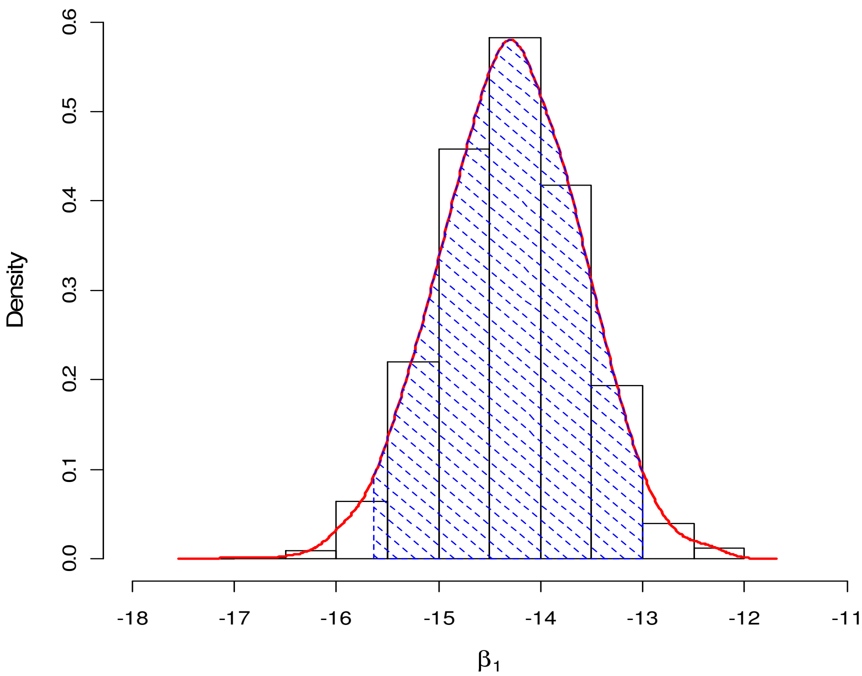


Figure 9 *Histograms of Markov chain samples using Metropolis-Hastings sampling for* β_1

By doing so, a likely range for $\boldsymbol{\beta}$ using a posterior interval can be determined. The $\alpha\%$ posterior interval contains the central $\alpha\%$ of the sampled values, and so the interval contains the true estimate with probability α. The lower limit of the interval is the $(50-\alpha/2)$th percentile of the Markov chain samples, and the upper limit is the $(50+\alpha/2)$th percentile. So a 95% posterior interval goes from the 2.5th to 97.5th percentile. As shown in Figure 9, using our sampled β_1 the empirical mean is -14.30 and standard deviation is 0.68 for the sample size *N*=10,000. Meanwhile, a 95% posterior interval for β_1 is from -15.64 to -12.99. It means that the posterior probability $P(\beta_1\,|\boldsymbol{y})$ lies in such an interval is 0.95. In the case of the probit model shown in equation (6), the model coefficients are random variables. So once the empirical distribution of each model coefficient can be simulated by Markov chain samples, the corresponding posterior probability $P(Y=1|\boldsymbol{\beta},\boldsymbol{X})$ will also be obtained. Therefore it will be straightforward to quantify the uncertainty boundary through the quantile estimation.

Based on a detailed account of algorithms behind Bayesian and MCMC methods, the next section elaborates on an application of Bayesian methods in the maritime industry addressing ship survivability assessment after damage. The given case study presents an example of Bayesian analysis applied for generalized linear model with a probit link function.

4.4 Bayesian Analyses for a Binary Regression Model

4.4.1 An Example of Ship Survivability Prediction

As Bayesian methods enable plausible inferences based on the assembled data of experiments, it is reasonable to be deployed for predicting the stochastic ship behaviour follow flooding with relevant data. In the knowledge that the interested phenomenon is physically influenced by a list of variables as stressed previously (e.g. loading condition, damage characteristics, sea environment), the following example will focus on training predictive model (i.e. the probability of capsize) by using only a single influential variable (i.e. sea environment measured by Significant Wave Height H_s) to illustrate the applicability of the proposed approach.

On the basis of the identified Probit model, the question at hand can be easily formulated as shown in equation (14), in which a series of predictor variables $(\boldsymbol{X}=x_1,\dots x_n)$ can be included simultaneously. For the time being, only a single variable H_s is considered for model training. Hence, $\boldsymbol{X}=x=H_s$.

$$P(t=30min, Y=capsize|\boldsymbol{\beta},\boldsymbol{X})=\Phi(\alpha+\boldsymbol{\beta}\boldsymbol{X}) \tag{14}$$

The next immediate task is to collect evidence so that Bayesian methods can be deployed for estimating the proposed model. The data collected from physical model experiments under Task 4.1, (Rask, 2010), is used and tabulated in Table 1. In total, there are 83 measured records. Having the data, the target distribution $P(\boldsymbol{\beta}|y)$ is approximated using 10,000 iterations with a burn-in sample size of 500. Commonly the M-H sampling algorithm is considered as convergence if the acceptance rate stands between 20 and 50%. In this case, $\boldsymbol{\beta}^*$ is accepted as $\boldsymbol{\beta}^{(i+1)}$ for 28.5% of all the iterations. (i.e. 2996 times out of total 10500 iterations)

Table 1 *M/V Estonia, experimental test matrix, t=30 minutes*

Theoretical H_s	No. of Capsize	No. of tests	Rate of Capsize
2	0	3	0
2.5	2	20	0.10
2.6	13	20	0.65
2.75	16	20	0.80
3	20	20	1

The estimated coefficients are summarized in Table 2. The value of β is positive indicating a positive correlation between the probability for the ship to capsize and the significant wave height. Figure 10 plots the MCMC approximations to the marginal posterior densities of the two model coefficients α and β. The 99% confidence interval of each parameter is based on 0.5% and 99.5% posterior quantiles of $P(\boldsymbol{\beta}|y)$.

Table 2 *Fitting a ship response model (after flooding) to the experimental data*

Parameter	Mean	SD	99% posterior interval
α	-21.63	4.631	-34.833, -11.228
β for H_s	8.295	1.769	4.334, 13.368

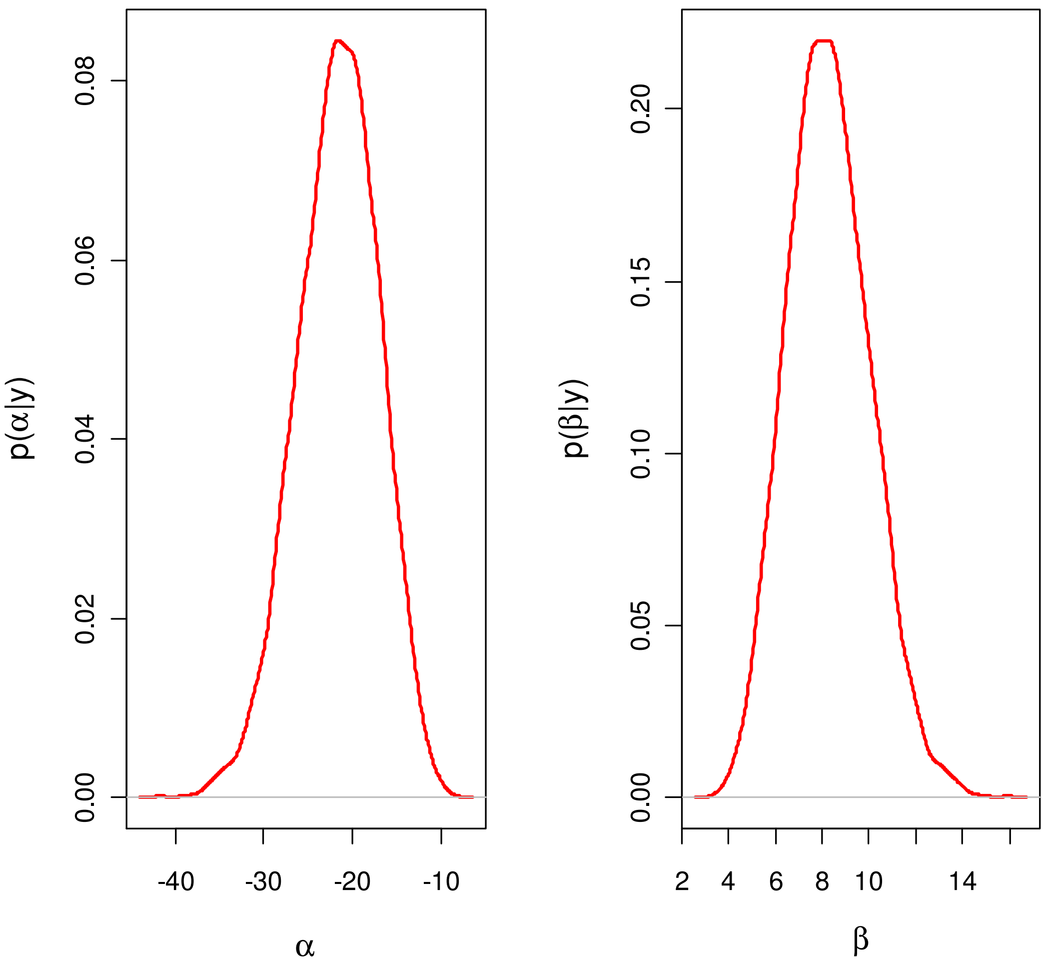


Figure 10 *MCMC approximations to the posterior distributions of α and β*

With the simulated distribution of the model coefficients $P(\boldsymbol{\beta}|y)$, the corresponding posterior probability of the output variable, probability of capsize $P(Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ can be estimated easily for a given sea state. A continuous distribution of the rate of capsizing is attained, as shown in Figure 11, if a range of sea states are considered. It is clear that the depicted Bayesian distribution, cooperated with 99% confidence interval have a good agreement with the experimental results.

Despite the fact that the estimated Probit model includes only a single variable , the estimated probit model using Bayesian techniques still can be regarded as a promising alternative for fast predicting ship survivability. Currently the data applied for model training is based on tailored experiments for M/V Estonia only. In order to avoid a common problem that the developed model may only represent the best 'fit' to specific damage situations, additional experimental data measured in other important projects (i.e. HARDER, SAFEDOR) must be taken into account for database expansion. All of the estimations in Chapter 5 are in accordance with the refined data. Moreover, the next chapter intends to include more influencing variables for model selection. After comparing the results of the new model to other traditional methods, it can be seen that Bayesian probabilistic methods are ideal for applications in the maritime field, for proposing the numerous types of models and quantifying the uncertainties associated with the investigated problems.

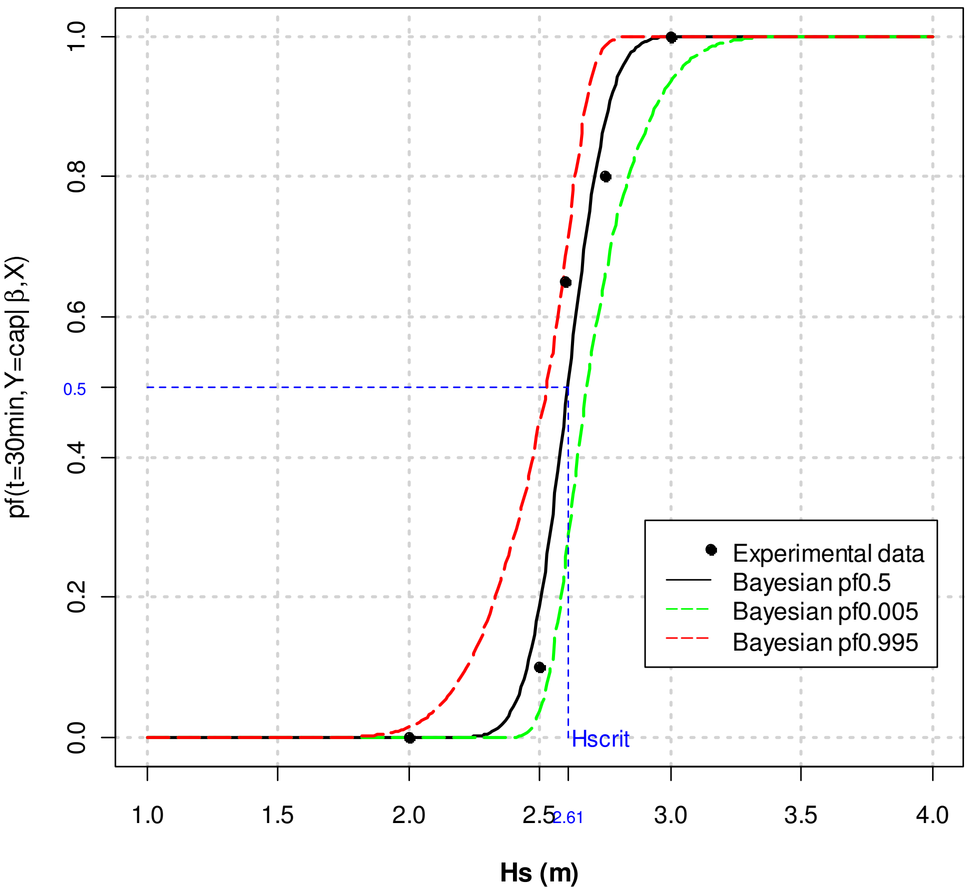


Figure 11 *M/V Estonia, posterior distribution of rate of capsizing simulated by Bayesian techniques (Applying tailored experimental data for model training)*

4.4.2 Ship Survival Time Assessment

As it is shown in Figure 11, the estimated model can be deployed to predict the rate of capsizing for a specific flooding case. The rate of changes has a sigmoid shape, which varies with the encountered sea states. Nevertheless, it is important to realise that such survivability prediction is made based on the experiment period lasting 30 minutes. Hence, it is necessary to set up the relationship between the time and the consequent survivability for given damage conditions.

Regarding this, the theory of the Bernoulli trial process is applicable (Jasionowski, 2006). Suppose that the probability of capsizing p_f is constant for a given damage j, consequently, the probability that the **n^th** test is a case of "capsize" can be measured by equation (15). The number of trials can be determined from $n = \frac{t_{cap}}{t_0=30min}$, where t_{cap} (in minutes) is the cumulative time needed for a damaged ship to capsize. Thus the probability of capsizing within t_{cap} can be evaluated.

$$F(t_{cap}|H_s, j) = 1 - (1-p_f)^n = 1 - (1-p_f)^{\frac{t_{cap}}{t_0}} \tag{15}$$

As a result, the matching probability density function of "time to capsize" can be formulated as shown in equation (16).

$$g\ (t_{cap}|H_s, j) = \frac{\partial F_{t_{cap}}}{\partial t_{cap}} = -\ln(1-p_f)\cdot(1-p_f)^{\frac{t_{cap}}{t_0}}\cdot {t_0}^{-1} \tag{16}$$

Concerning the experimental results as tabulated in Table 1, the case at $H_s = 2.6$ m where $p_f = 0.65$ measured within 30 minutes, has been select to demonstrate how the ship survivability assessment can be extended to any time period. Following 20 independent survivability tests, the histogram of time to capsize is plotted in Figure 12. Proceed from a comparison, probability density function based on equation (16) for the same damage situation is also considered, where $p_{f\_Bayes} = 0.477$ is estimated through Bayesian methods in section 4.4.1.

It should be pointed out that no initial transient flooding has been modelled during the experiments. The wave tests start from an equilibrium stage after the damaged compartments are flooded. Thus a reasonable deviation can be found regarding the density distributions in the first 600s. Subsequently, corresponding cumulative distributions $F(t_{cap}|H_s)$ are illustrated in Figure 12, where the CDF curve provides the probability of the cumulative amount of time that it takes the ship sink/capsize. Meanwhile, Bayesian estimates as a 99% confidence interval of $F(t_{cap}|H_s)$ present a good way to quantify the uncertainties associated with the measurements.

In the same way, Figure 13 and Figure 14 show the cases when the experiments performed at $H_s = 2.75$ m with $p_f = 0.8$ and $H_s = 3.0$ m with $p_f = 1$. Apart from that, the computed rates p_{f_Bayes} are 0.878 and 0.999 through Bayesian methods. It can be seen that all of measured probabilities of capsizing within 30 minutes lie within the assigned 99% uncertainty bounds.

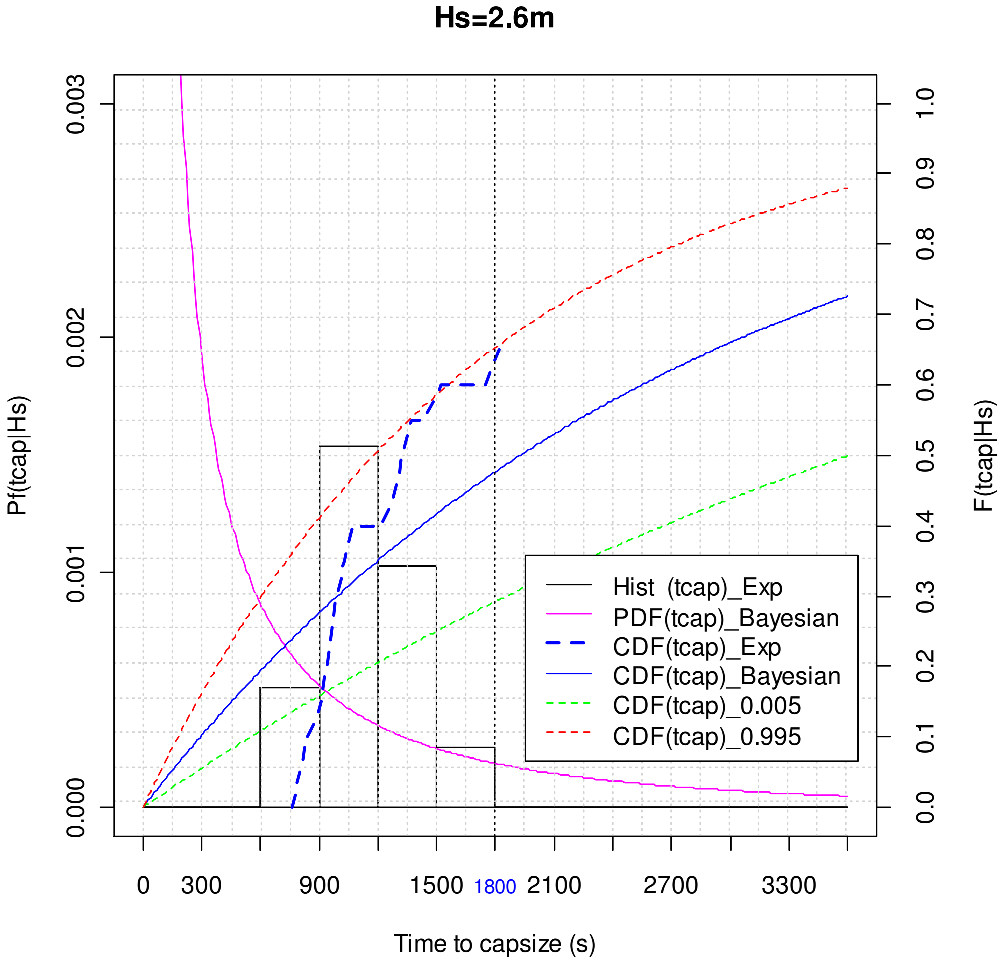


Figure 12 *M/V Estonia, Histogram of probability density function of time to capsizing given the damage case tested at* H_s =2.6m *with* $p_f = 0.65$. *Comparison with Bayesian density; Cumulative probability of time to capsize* (Ft_{cap}). *Comparison with Bayesian CDF. The evaluation time of experiment = 30 minutes.*

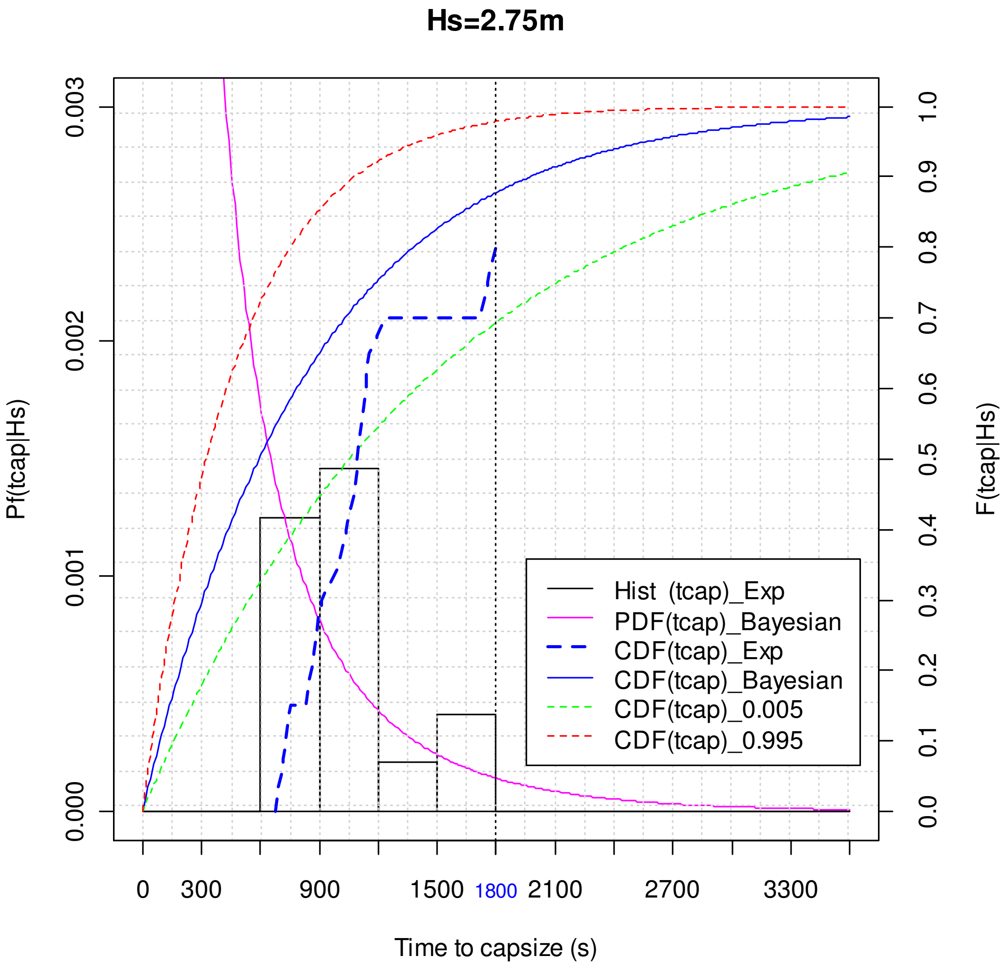


Figure 13 *M/V Estonia, Histogram of probability density function of time to capsizing given the damage case tested at Hs* =2.75m *with* $p_f = 0.8$. *Comparison with Bayesian density; Cumulative probability of time to capsize* (Ft_{cap}). *Comparison with Bayesian CDF. The evaluation time of experiment = 30 minutes.*

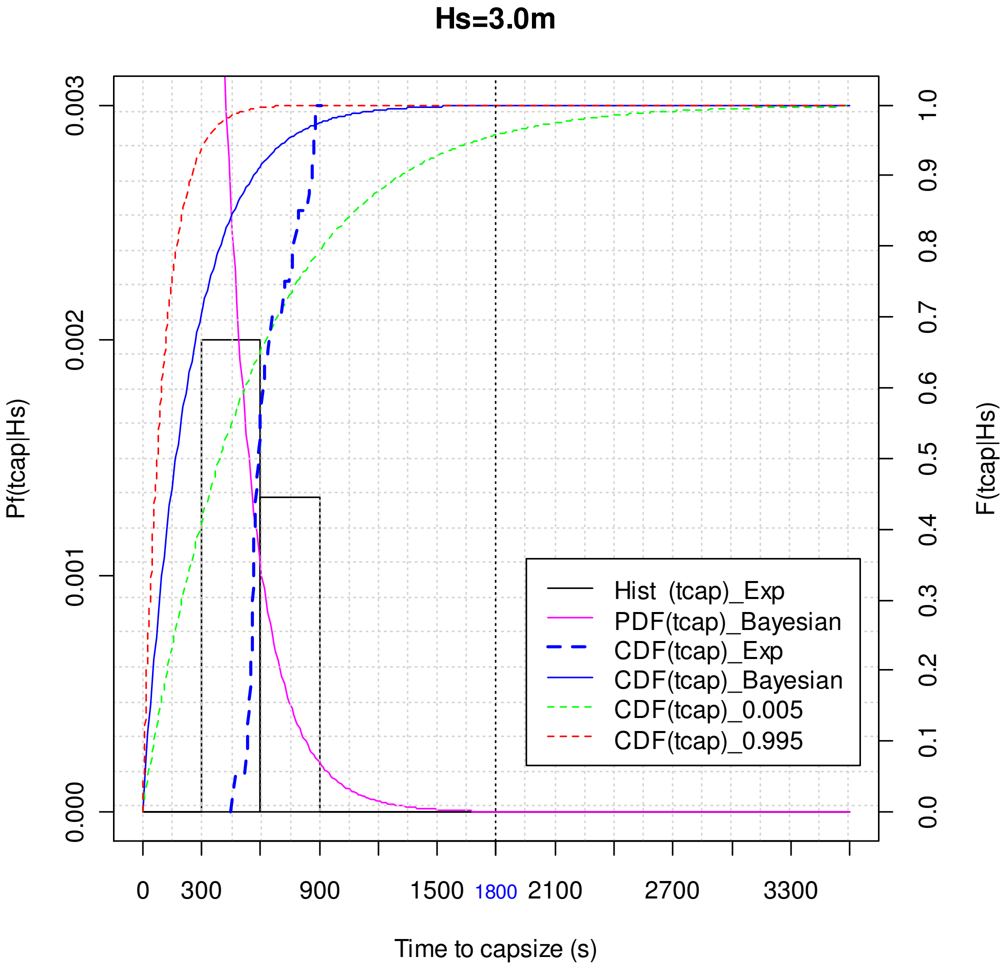


Figure 14 *M/V Estonia, Histogram of probability density function of time to capsizing given the damage case tested at Hs* =3.0m *with* $p_f = 1$. *Comparison with Bayesian density; Cumulative probability of time to capsize* (Ft_{cap}). *Comparison with Bayesian CDF. The evaluation time of experiment = 30 minutes.*

5. MODEL APPLICATION

With the explained Probit model and Bayesian inference techniques, this chapter details the application of the proposed new methodology for making inference on ship survivability following flooding. The application procedure consists of two main steps:

- Model selection starts with an identification of the dominant variables to be included in the proposed models. And then based on the collected experimental data, different predictive models are estimated by employing the techniques as described in Chapter 3 and 4. It should be noted that the considered predictor variables should be detailed enough to capture the ship behaviour from the instance of flooding up to the ultimate capsize/sink (if it occurs). Furthermore, it is also important to keep the list compact so as to maintain a well balance between the number of key variables influencing ship survivability and the quality of the predictability of the model.

- Validation of the models proposed in the above step must be performed. The relevant studies are discussed in two ways: i) at the damage scenario level, the estimated results are compared with both the experiments and conventional methods (i.e. damage stability clauses in SOLAS 2009); ii) at a ship level, the established model is applied to predict survivability of new ships, which are subjected to possible different damage scenarios. In this case, it is noteworthy that these flooding cases are likely to be different from the cases used for training the model. The desired goal in this step is to have a model can be used for effectively and scientifically predicting survivability of the RoPax fleet as a whole.

5.1 Model Selection

5.1.1 Dominant Variables Identification

Ship survivability is affected by a large amount of physical variables (e.g. ship internal watertight arrangements, loading conditions, the hull damage characteristics, and external sea environment). In this context, sensitivity analysis is deemed desirable to assess the significance of each variable so that those relatively more important variables can be identified scientifically. Nevertheless, one need to appreciate that the amount of effort required (in terms of time and resources) is tremendous. On the other hand, it is necessary to understand that the ship survivability is a subject that has been studied in a great detail since the past century. Hence, it is deemed appropriate to make use of the existing resources and knowledge.

On the basis of the research work carried out in HARDER and the subsequent updated probabilistic rules in SOLAS concerning ship damage stability, it is noted that GZ particulars are recognised as the bridge to link the various physical variables (e.g. loading condition, subdivision arrangement, and damage characteristic) to the survivability. This is indeed an important step forward to address ship damaged stability probabilistically. Nevertheless, the existing formulations also have the following challenges need to be further investigated:

- The importance of sea environment to the behaviour of damaged ships is not explicitly included. This raises the concern over the quality and realistically of the prediction using the current formulations.
- For the purpose of accommodating the probabilistic assessment of the dynamic ship survivability at operational stage, the computational effort needed for assessing the GZ particulars of the considered flooding cases is tremendous. Hence, it is not straightforward to employ these formulations for decision support during operation, particularly under emergency situations.

In response to the aforementioned challenges, a series of key variables divided into three categories is identified in Table 3. All of them are separated to configure two different predictive models. The following sections will go into detail about the estimation and validation of these proposed models.

Table 3 *Predictor variables for model training*

Category	Variable	Model
Sea environment	H_s	1 & 2
Initial loading	KG/KM_T	2
	T/D	2
Damage attributes	GZ_{max}	1
	Range	1
	$Heel$	2
	L_d/L_s	2

5.1.2 Data Collection

Having the identified variables, it is important to collect pertinent data. In this respect, it is noted that the size of the data set is always restricted by both allocated time and budget for each study. For the sake of statistical inference, it is generally agreed that the sample size has direct impact on the quality of the outcome. With smaller samples, using statistical techniques may easily lead to either 1) too little statistical power for the test to realistically identify significant results, or 2) the results are too “over fitting” to the data, however they are artificially good because they fit the sample but not enough to provide generalised information (Hair, 2009). Hence, this study attempts to collect as much reliable data as possible so as to assure the quality of the trained model.

Fortunately, a series of benchmarking experiments (in accordance with Directive 2003/25/EC (EC, 2003)) on survivability assessment of RoPax vessels in the projects partially funded by the European Commission (i.e. HARDER (Tuzcu and Tagg, 2001), SAFEDOR (Chen et al., 2009), FLOODSTAND) and more recently by the European Maritime Safety Agency (EMSA) has been made available. In total, this represents 756 runs of the repetitive wave test on a number of specifically developed scaled ship models.

5.1.3 Model Estimation using Variables including GZ particulars

Based on the current formula of computing survival factor in SOLAS 2009, the variables of damage GZ particulars are the key means to link ship survivability to those influential variables (e.g. loading conditions, ship subdivision).

In order to keep consistency of the assessment, GZ properties are included in the model at this time. Moreover, the significant sea state also plays an important role. Hence, concerning a particular damage case, three governing variables ($\boldsymbol{X} = x_1, x_2, x_3$) that consist of the damage GZ_{max}, the related Range in EQ, and the significant sea state (H_s) are considered as the prime parameters for model estimation. In such case, four model coefficients must be estimated and each target distribution $P(\boldsymbol{\beta}|y)$ is approximated by Metropolis-Hastings samples. All of assembled data (756 tests), as summarised in Table 4, are utilised for model training. The MCMC method simulates 10,000 times with the burn-in of the first 500 samples. $\boldsymbol{\beta}^*$ is accepted as $\boldsymbol{\beta}^{(i+1)}$ for 37.1% of the iterations, which satisfies the required acceptance rate.

Table 4 *Assembled experimental data for model training (with damage GZ particulars)*

Project	No. of tests	Damage case	Measured Hs	Static damage stability	
				GZ_{max}	Range
HARDER	425	PRR01_case 2	1.5 - 6.25	0.30	15.81
		PRR01_case 3		0.19	16.31
		PRR01_case 5		0.33	18.79
		PRR01_case 6		0.20	13.93
		PRR01_case 7		0.10	9.85
		PRR01_case 9		0.32	16.62
		PRR01_case15		0.32	17.62
		PRR01_case16		0.17	12.57
		PRR01_case17		0.27	13.98
		PRR01_case18		0.16	10.15
		PRR01_case21		0.53	23.88
		PRR01_case23		0.38	19.24
		PRR01_case24		0.26	15.64
		PRR01_case25		0.14	11
		PRR01_case26		0.41	20.4
		PRR01_case27		0.28	16
SAFEDOR	248	Pentalina	1.5 - 2.5	0.20	25
FLOOSTAND	83	M/V Estonia	2.0 - 3.0	0.09	9.545

It is noteworthy that 83 repetitive wave tests for a RoPax model of M/V Estonia (in scale 1:40) have been undertaken in Task4.1. A port-side damage (DS/P6-7.1.0) was modelled as shown in Figure 15. Some correlated hydrostatics and stability information is tabulated in Table 5. Furthermore, Figure 16 gives the computed damaged GZ curve.

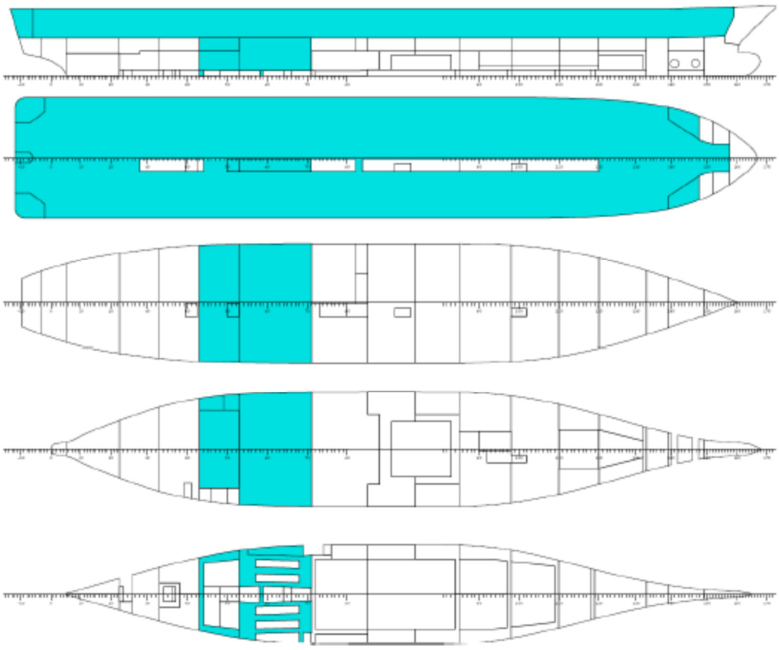


Figure 15 *M/V Estonia, DS/P6-7.1.0 (2-compartment damage)*

Table 5 *M/V Estonia, hydrostatics and stability information*

Damage Case	Initial condition			Damaged stability			
	PER	T	Trim	KG	GZ_{max}	Range	Heel
DS/P6-7.1.0	0.95	5.39 m	0.435m by aft	10.62 m	0.09m	9.545 deg	1.95 deg to port

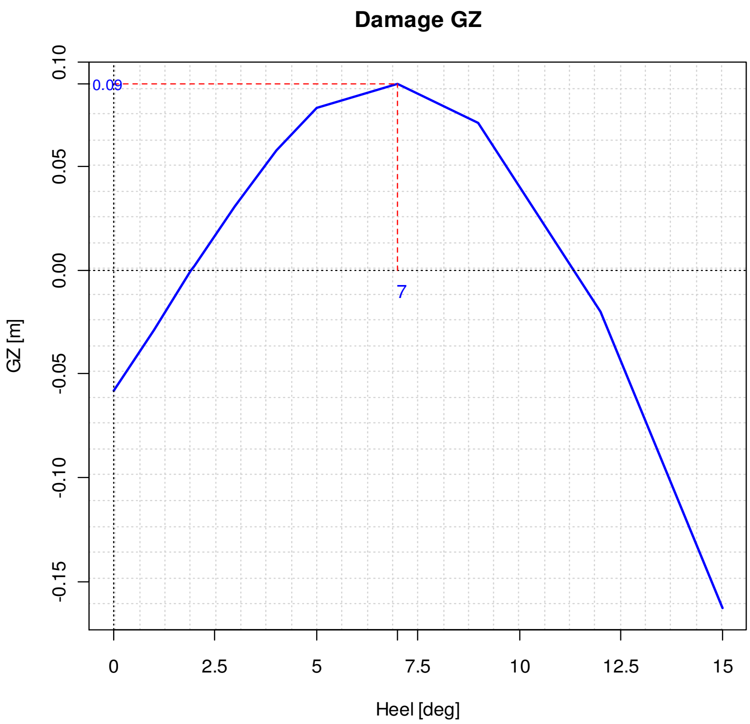


Figure 16 *M/V Estonia, Damage GZ curve, GZ_{max}=0.09m where Range =9.545deg*

The subsequently simulated model coefficients are summarized in Table 6. The MCMC approximations to the posterior densities $P(\boldsymbol{\beta}|y)$ for all the coefficients $(\alpha, \beta_1, \beta_2, \beta_3)$ are depicted graphically in Figure 17. These distributions seem to interpret all these three variables have much influence on the ship survivability following flooding, since the 99% quantile-based posterior intervals for all these coefficients do not contain zero. The estimated magnitudes β_1 and β_3 are positive. In contrast, the negative sign of β_2 indicates that GZ_{max} is inversely proportional to the rate of ship capsizing, which agrees well with the phenomenon observed in reality.

Table 6 *Summary of Metropolis-Hastings samples of model coefficients (with damage GZ particulars)*

Parameter	Mean	SD	99% posterior interval
α (interception)	-3.8844	0.2182	-4.4896, -3.3411
β_1 for H_s	1.3655	0.0551	1.2267, 1.5268
β_2 for GZ_{max}	-14.3076	0.7115	-16.1117, -12.4717
β_3 for *Range*	0.1565	0.0121	0.1247, 0.1887

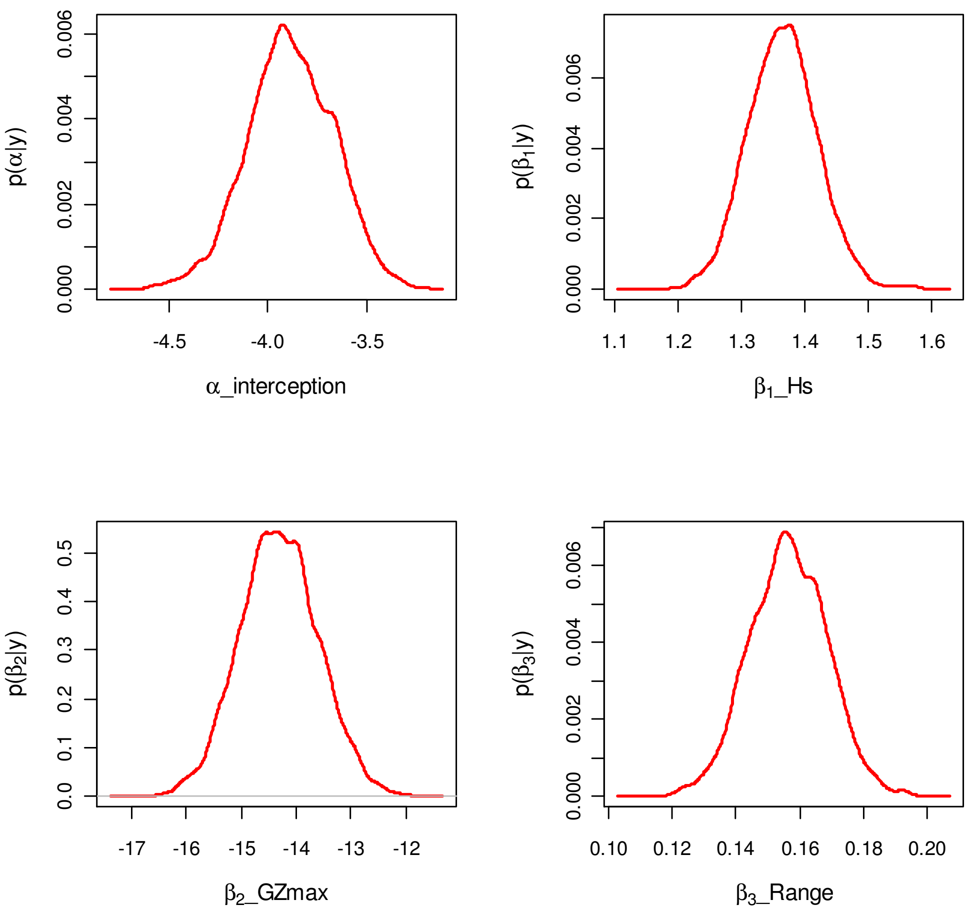


Figure 17 *MCMC approximations to posterior distributions of model coefficients (with damage GZ particulars)*

5.1.4 Model Estimation using Variables excluding GZ particulars

As it is noted in the traditional approach for addressing ship survivability assessment, damage GZ particulars have been widely used as the principle parameters for estimation. However, during the real operation, the response time for a damaged ship in a seaway is very limited. This demands a really fast assessment of ship survivability for decision support in emergency situations. In reality, the computation of damage stability is very time-consuming as it requires information of ship subdivision arrangement, and the status of openings which must be linked with the instantaneous of the statues of watertight doors on board. Such computational defect may constrain the applicability of damage GZ particulars in practical.

In such circumstance, a possible way out is to reformulate the predictive model with other key variables instead of GZ_{max} and Range. Some easy accessible parameters at operational stage are more preferable. In order to strive for regression model parsimony, heeling angle at the damaged equilibrium and the extent of flooding are selected for describing the damage attributes. Beyond this operational information, initial conditions at the design stage must be considered, such as the design draught and the centre of gravity KG of the damaged ship.

Attempting to maximize the utility of the limited data set for model training, non-dimensional measurement of the aforementioned variables are adopted for processing the data analysis regardless the size of the ship. In this way, the ratios between draught and depth, KG and KM_T, damage length and the subdivision length of the ship are considered respectively. The pertinent experiment data is tabulated in Table 7.

Table 7 *Assembled experimental data for model training (without damage GZ particulars)*

Damage case	Measured Hs	Intact Condition		Damage Condition	
		KG/KMT	T/D	Heel	Ld/Ls
PRR01_case 2	1.5 - 6.25	0.8306	0.6944	3.17	0.0466
PRR01_case 3		0.8669	0.6944	4.07	0.0466
PRR01_case 5		0.7962	0.6944	2.54	0.0466
PRR01_case 6		0.8413	0.6944	3.23	0.0466
PRR01_case 7		0.8782	0.6944	4.14	0.0466
PRR01_case 9		0.7747	0.6944	2.36	0.0466
PRR01_case15		0.8483	0.6389	3.72	0.0466
PRR01_case16		0.8898	0.6389	5.32	0.0466
PRR01_case17		0.7990	0.7500	2.59	0.0466
PRR01_case18		0.8443	0.7500	3.34	0.0466
PRR01_case21		0.7860	0.6944	1.76	0.0466
PRR01_case23		0.8306	0.6944	2.25	0.0466
PRR01_case24		0.8669	0.6944	2.89	0.0466
PRR01_case25		0.9093	0.6944	4.33	0.0466
PRR01_case26		0.7962	0.6944	1.65	0.0466
PRR01_case27		0.8413	0.6944	2.08	0.0466
Pentalina	1.5 - 2.5	0.8533	0.7860	1.00	0.0717
M/V Estonia	2.0 - 3.0	0.8949	0.7046	1.95	0.0518

On the basis of the applied predictive model $P(Y = capsize|\boldsymbol{\beta}, \boldsymbol{X}) = \Phi(\alpha + \boldsymbol{\beta X})$, five influencing variables $\boldsymbol{X}$ have been fitted in, which comprise sea environment H_s, initial KG/KMT, T/D, damaged attributes Heel in EQ and L_d/L_s. In current study, all these input variables are assumed independent and thus there is no interaction form $(x_i \cdot x_j)$ in the model. By using the same MCMC algorithm for approximating the target posterior distribution of each regressor index $P(\boldsymbol{\beta}|y)$, a summary of simulated model coefficients are listed in Table 8. The Metropolis-Hastings acceptance rate is 26.7% for this case, which satisfies the requirement.

As it is shown in Figure 18, the variations (within 99% quantile-based posterior interval) in the obtained coefficients do not cover zero, which implies the significance of the chosen variables to assess the response of damaged ships. Furthermore, as the estimated $\boldsymbol{\beta}$ are positive, they indicate positive correlations. This implies:

- Considering a ship with a specific draught, the larger KG, the poorer damage stability the ship has.
- As the ratio between draught and depth provides information on residual freeboard at design stage, the higher ratio leads to the smaller amount of freeboard left. Thus, it presents a higher rate of losses.
- Similarly, a higher ratio between damage length and subdivision length means the larger size of the damage and the lower survivability the ship has.

Table 8 *Summary of Metropolis-Hastings samples of model coefficients (without damage GZ particulars)*

Parameter	Mean	SD	99% posterior interval
α (interception)	-57.2016	2.9094	-65.3078, -50.2090
β_1 for H_s	1.6203	0.0717	1.4390, 1.8264
β_2 for KG/KMT	20.7519	1.8823	15.9061, 25.7977
β_3 for T/D	42.8016	2.7077	36.3450, 49.9021
β_4 for $Heel$	0.4126	0.0742	0.2158, 0.6193
β_5 for L_d/L_s	54.6001	12.3074	26.4065, 90.7295

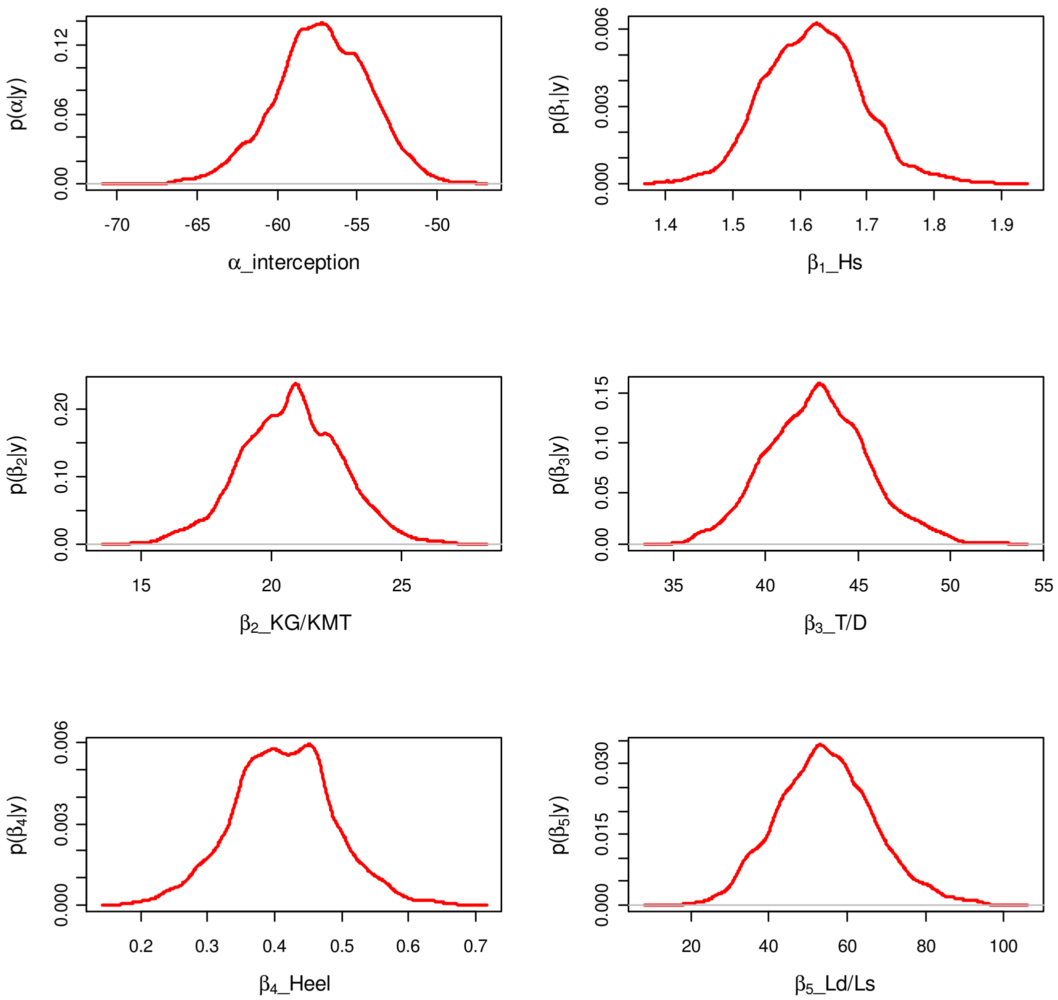


Figure 18 *MCMC approximations to posterior distributions of model coefficients (without damage GZ particulars)*

5.2 Model Validation

The two aforementioned models have been selected for comparison with the benchmark data and the conventional methods (as entailed in SOLAS2009). In the following context, the model allowing for damage GZ particulars as input variables is defined as Model 1. Model 2 is to denote the estimated model using other accessible operational parameters (see Table 3). A detailed discussion of comparison findings is given later in each part.

5.2.1 Result Comparison with Experiments

In this section, the estimated models are put forward to make predictions, which will be compared with the findings of experiments data. The investigation is performed in two aspects: survivability prediction and ship survival time assessment.

5.2.1.1 Model fitted with damage GZ particulars (Model 1)

With the estimated Model 1, in which the coefficients and their variations are tabulated in Table 6, the probability distributions of the model coefficients $P(\boldsymbol{\beta}|y)$ are known. As a result, the rate of capsizing $p_f(Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ can be computed directly once the values of predictor variables $\boldsymbol{X}$ are assigned. In the flooding case considered for the “M/V Estonia” model, the survivability prediction is depicted as Figure 19, where the inputs are: $\boldsymbol{X} = (H_s, 0.09m, 9.545deg)$.

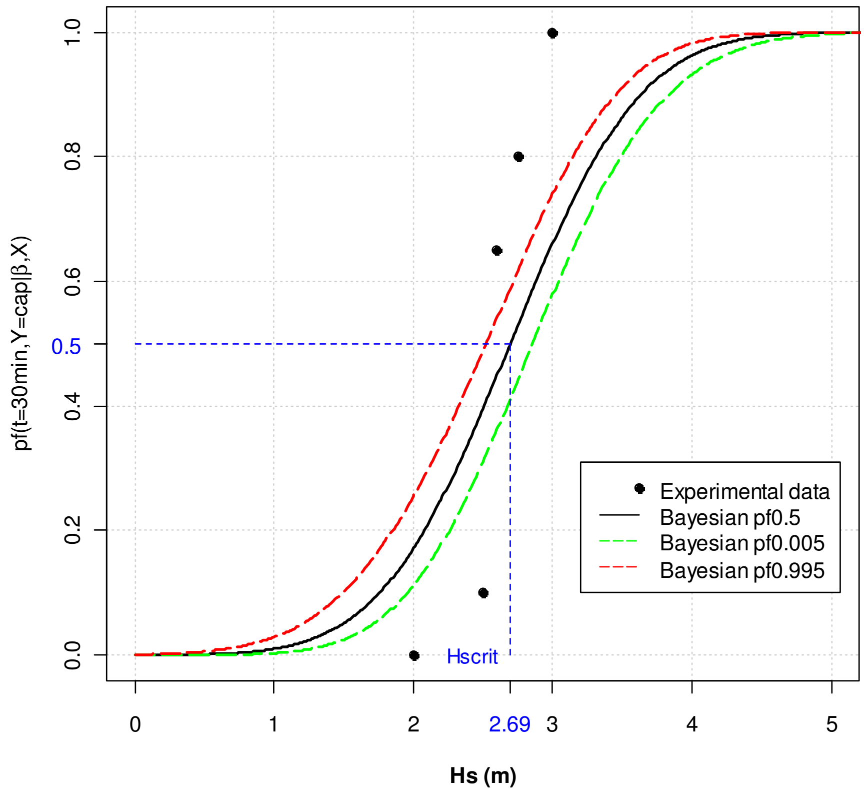


Figure 19 *M/V Estonia, posterior distribution of rate of capsizing*
$p_f(t = 30min, Y = cap|\boldsymbol{\beta}, \boldsymbol{X} = H_s, GZ_{max}, Range)$
(Applying the training data set in Table 4)

Comparing with the experimental data, it can be seen from Figure 19 that Model 1 underestimates the ship survivability when $H_s < 2.6m$, and vice versa. The discrepancy can be interpreted in two ways.

- Firstly, the data set for model training involves data from other benchmarking experiments assembled to date (see Table 4). Hence, the current estimation for this particulars damage case can be best described as a compromise among all the observations to depict the survivability of RoPax vessels in general. In this context, it is important to point out that the expansion of the training data set is essential to make the developed model that is capable of describing the population (world fleet) as a whole.
- Secondly, the values of the input parameters (i.e. damage GZ particulars) are difficult to be obtained from experiments. In this case, other accessible operation parameters are preferred for model estimation for the sake of a fast prediction. Hence, the next section provides a demonstration on the prediction using other input variables.

Other than the tested damage case on the model M/V Estonia, the remaining observations as shown in Table 4 can be also given parallel estimations. At this time, the mean critical sea state (Hs_{crit}) is deployed as an indicator to compare the survivability between the real measurements and the predictions of the models. To define Hs_{crit}, it stands for the significant wave height at which the vessel has 50% rate of capsizing within $t = 30$ minutes following flooding.

According to Figure 19, the critical sea state $Hs_{crit} = 2.69$ can be easily measured once the probability distribution of ship capsizing $p_f(Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ is plotted against the encountered sea states. Through Figure 20, it can be observed that the predicted critical sea states are very much comparable with that of the situations identified during the experiments as almost half out of the total 18 cases are blow the diagonal line. Meanwhile, the upper confidence bound (e.g. 99%) of each estimated mean Hs_{crit} is also assigned (i.e. denoted by red dots) Nevertheless, it is notable that the tested "Estonia" case, as indicated in Figure 19, overestimates the survivability in higher sea states. Hence, it is difficult for the model to produce conservative estimations comparing to the actual measurements.

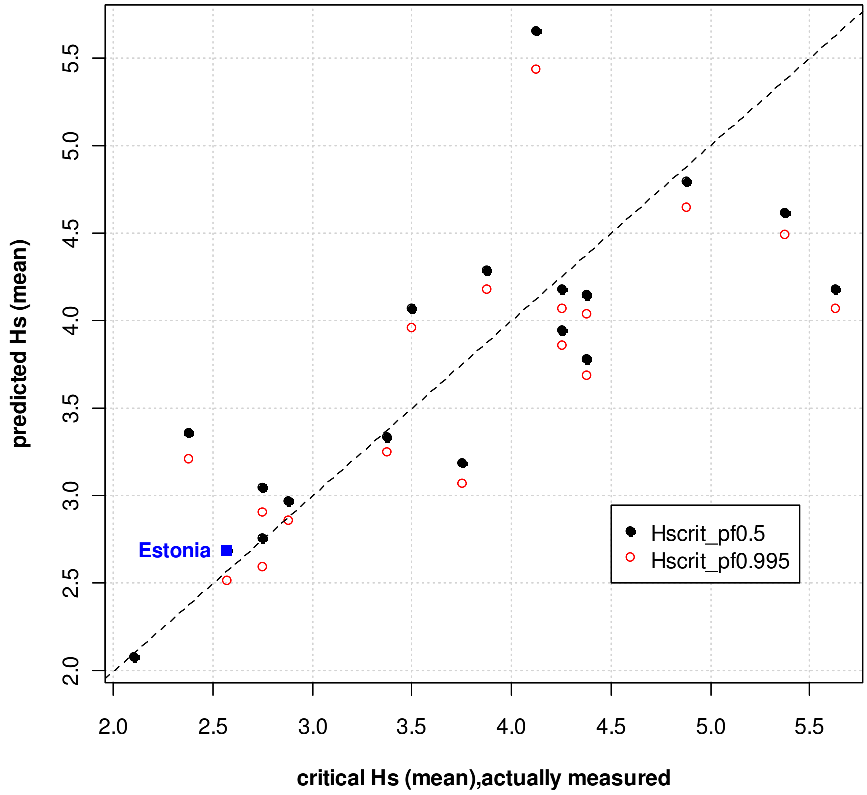
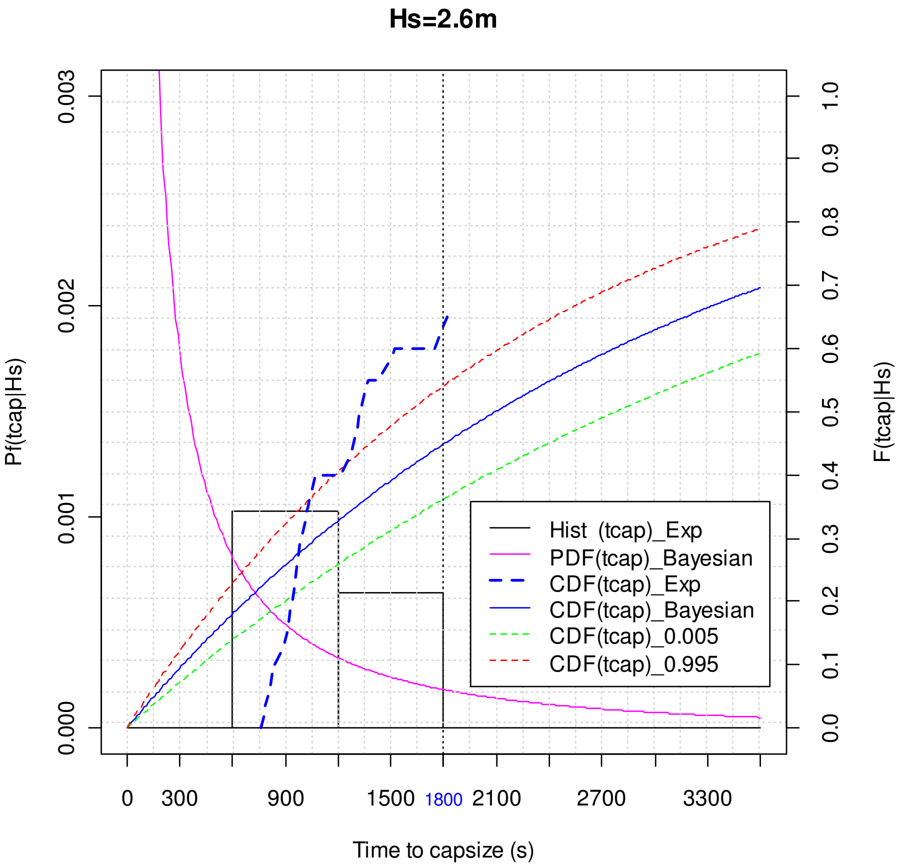
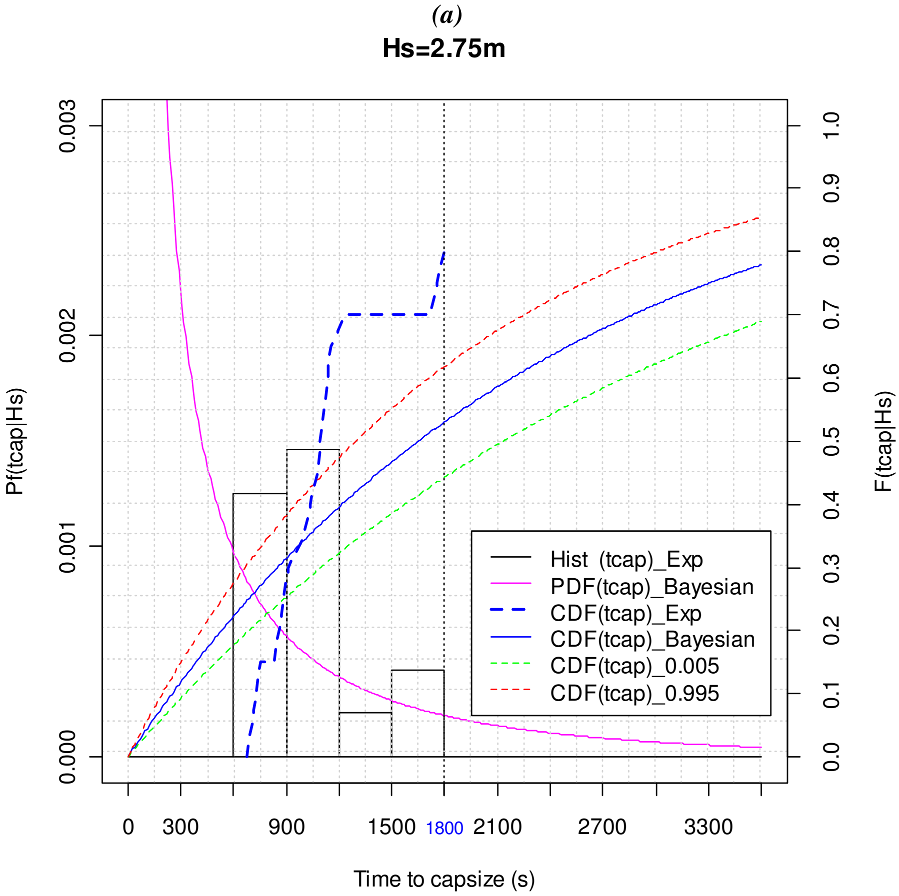


Figure 20 *Comparison of mean critical sea states* $(p_f = 0.5)$*between Model 1 prediction and the experimental measurement*

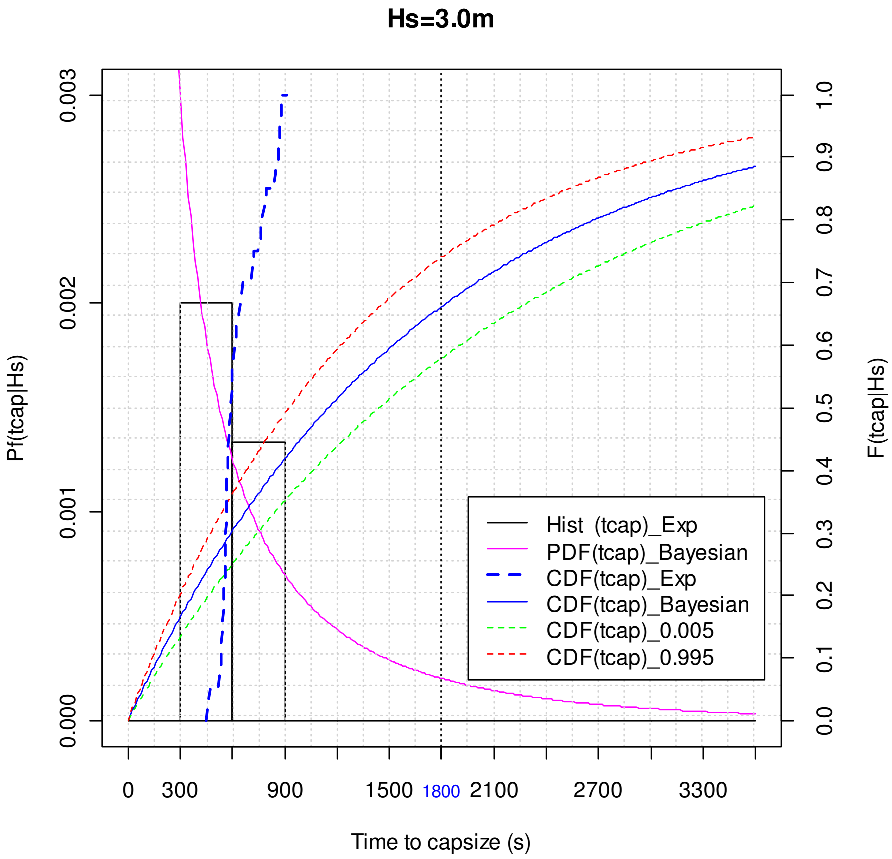
As far as the survival time assessment is concerned, the principle of generating a cumulative distribution of the probability for "time to capsize" is based on a Bernoulli trial process (See section 4.4.2). As the predicted rates of capsizing $p_f(Y = cap|\beta, X)$ are less than the physical records when $\mathrm{H_s} \geq 2.6\mathrm{m}$, it is expected the estimated ship vulnerability to flooding to be smaller than the measured values within a given time interval (e.g. 30 minutes). Figure 21 exhibits such described phenomenon.


Hs=2.6m
Pf(tcap|Hs)
F(tcap|Hs)
0.000
0.001
0.002
0.003
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
300
900
1500
1800
2100
2700
3300
Time to capsize (s)
Hist (tcap)_Exp
PDF(tcap)_Bayesian
CDF(tcap)_Exp
CDF(tcap)_Bayesian
CDF(tcap)_0.005
CDF(tcap)_0.995

(a)


Hs=2.75m
Pf(tcap|Hs)
F(tcap|Hs)
0.000
0.001
0.002
0.003
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
300
900
1500
1800
2100
2700
3300
Time to capsize (s)
Hist (tcap)_Exp
PDF(tcap)_Bayesian
CDF(tcap)_Exp
CDF(tcap)_Bayesian
CDF(tcap)_0.005
CDF(tcap)_0.995

(b)



(c)

Figure 21 *M/V Estonia, Cumulative probability of time to capsize* $\left(Ft_{cap}|H_s\right)$.
The evaluation time of experiment = 30 minutes.
a) H_s =2.6m with $p_f = 0.65$ and $p_{f\_Bayes} = 0.449$
b) H_s =2.75m with $p_f = 0.8$ and $p_{f\_Bayes} = 0.530$
c) H_s =*3.0*m with $p_f = 1$ and $p_{f\_Bayes} = 0.662$

5.2.1.2 Model fitted without damage GZ particulars (Model 2)

In terms of Model 2, the sampled weights of each variable have been summarized in Table 8. Similar to the examination for Model 1, a conditional probability distribution of $p_f(Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ given $P(\boldsymbol{\beta}|y)$ and $\boldsymbol{X}$ can be generated. Model 2 differs from Model 1 in that it takes operational information of heel angle and damage extent into account. Both variables are easier to be measured than damage GZ particulars.

Referring to the deliverable of Task 4.1, the average of recorded heel angle is 4.25 degree towards the port during the static measurement before the wave test. Moreover, the damage was caused to the model in accordance with SOLAS damage opening standard, which requests the damage length is 0.03L+3m (i.e. equals 7.12m for this case). So the input variables of Model 2 can be fixed by also using other intact information $\boldsymbol{X} = (H_s, 0.9088, 0.7046, 4.25deg, 0.0518)$. Consequently, it is revealed that the estimated survivability of Model 2 achieves a better fit with experimental data than that of Model 1, as illustrated in Figure 22.

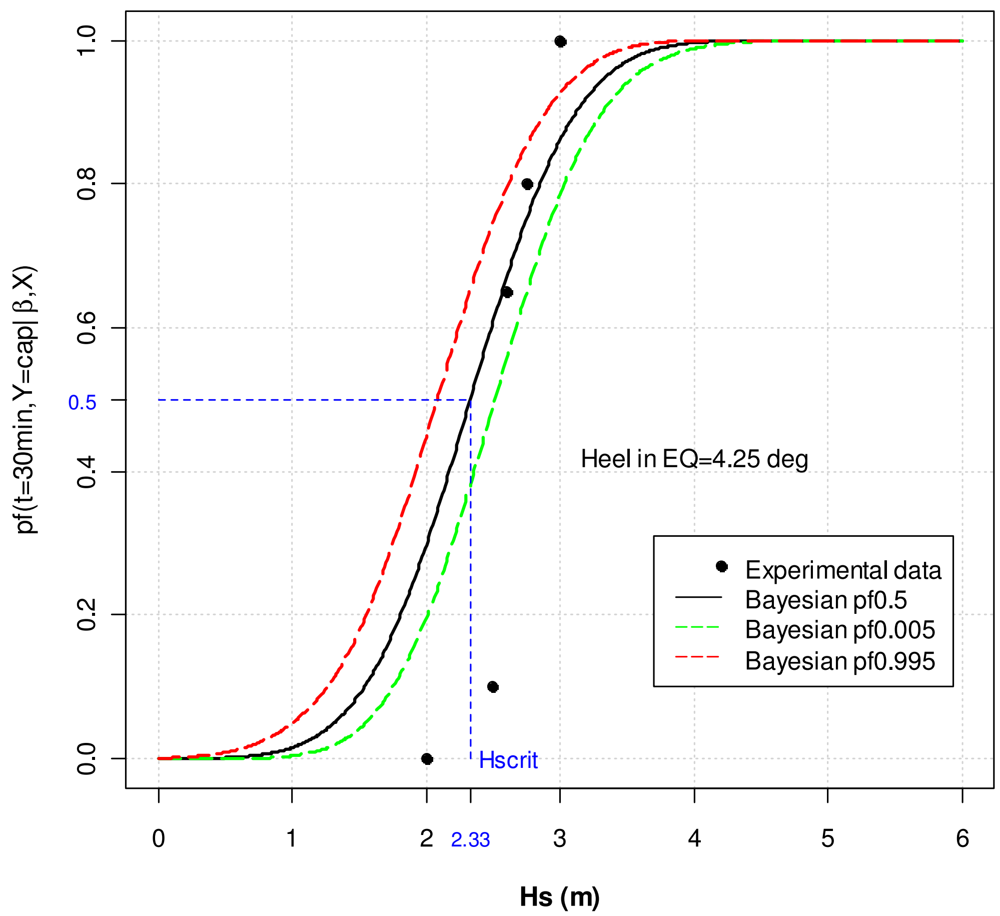


Figure 22 *M/V Estonia, posterior distribution of rate of capsizing*
$p_f(t = 30min, Y = cap|\boldsymbol{\beta}, \boldsymbol{X} = H_s, KG/KMT, T/D, Heel, L_d/L_s)$
(Applying the training data set in Table 7)

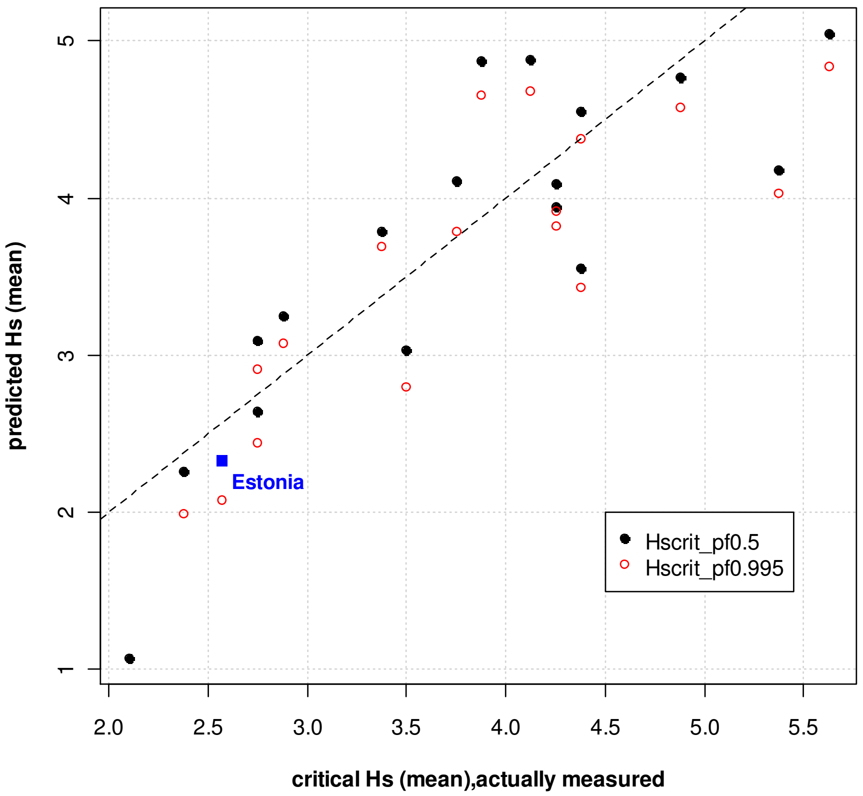
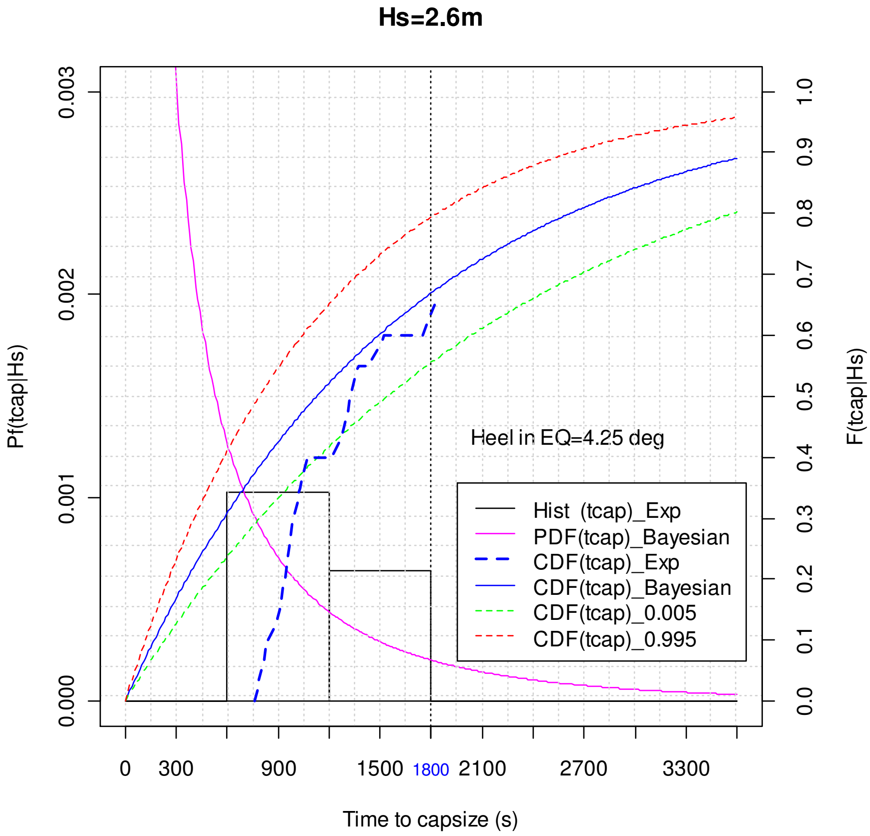
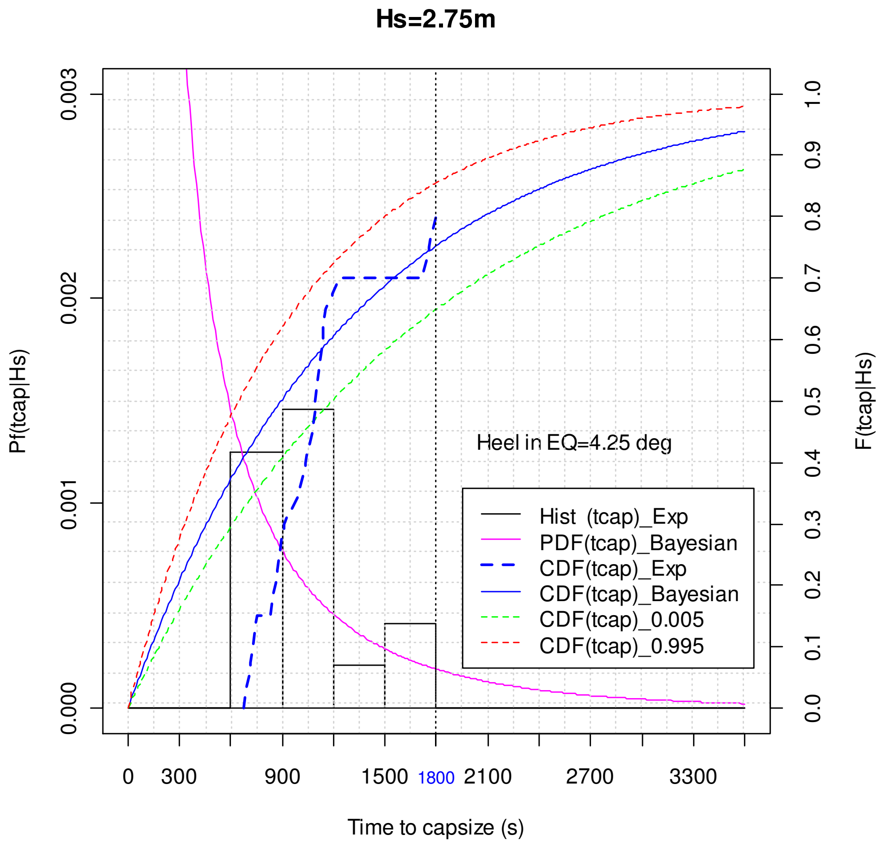


Figure 23 *Comparison of mean critical sea states* $(p_f = 0.5)$*between Model 2 prediction and the experimental measurement*

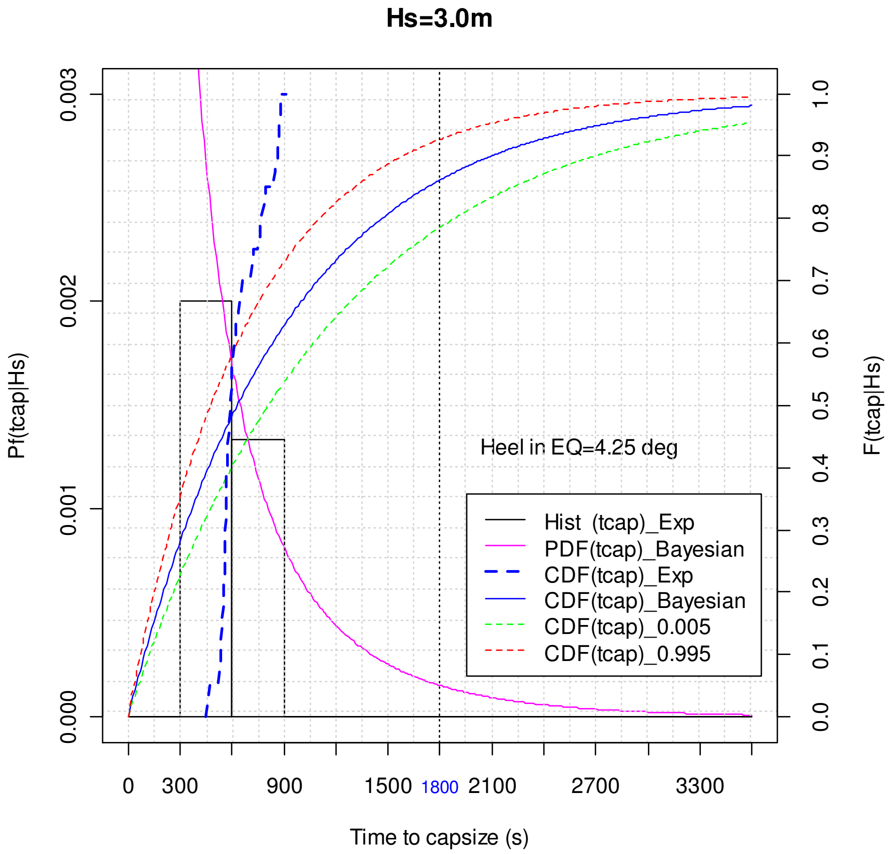
In light of the revealed relationship in Figure 23, it seems the predicted critical sea states of Model 2 are still comparable with that of the actual situations despite altered predictor variables have been applied. In contrast to a similar evaluation in Figure 20, results of Model 2 under-predict more cases than Model 1 when comparing with the survivability observed during experiments. In this case, apparently the estimated survival sea state for the tested “Estonia” case is less than the actual measurement. As a result, considering the survival time assessment, the predicted vulnerability for this case is not underestimated any more. It is visible that the computed probability of time to capsize at higher sea states is more close to the real observation as depicted in Figure 24.


Hs=2.6m
Pf(tcap|Hs)
F(tcap|Hs)
Heel in EQ=4.25 deg
Hist (tcap)_Exp
PDF(tcap)_Bayesian
CDF(tcap)_Exp
CDF(tcap)_Bayesian
CDF(tcap)_0.005
CDF(tcap)_0.995
0.000
0.001
0.002
0.003
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
300
900
1500
1800
2100
2700
3300
Time to capsize (s)

(a)


Hs=2.75m
Pf(tcap|Hs)
F(tcap|Hs)
Heel in EQ=4.25 deg
Hist (tcap)_Exp
PDF(tcap)_Bayesian
CDF(tcap)_Exp
CDF(tcap)_Bayesian
CDF(tcap)_0.005
CDF(tcap)_0.995
0.000
0.001
0.002
0.003
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
300
900
1500
1800
2100
2700
3300
Time to capsize (s)

(b)



(c)

Figure 24 *M/V Estonia, Cumulative probability of time to capsize* $\left(Ft_{cap}|H_s\right)$.
The evaluation time of experiment = 30 minutes.
a) H_s =2.6m with $p_f = 0.65$ and $p_{f\_Bayes} = 0.670$
b) H_s =2.75m with $p_f = 0.8$ and $p_{f\_Bayes} = 0.753$
c) H_s =*3.0*m with $p_f = 0.1$ and $p_{f\_Bayes} = 0.862$

On the basis of the comparison results with experimental data, it can be seen that the selection of model variables have significant impact on the prediction outcome. Despite the damage GZ curve has long been regarded as the means for measuring residual stability, the computational complexity is an inevitable obstacle to be applied at operational stage for decision support in emergencies. As an alternative, Model 2 is trained with other more accessible operational variables. The comparisons as presented from Figure 22 to Figure 24, indicates that Model 2 can be also a promising solution. Regarding the tested case to the "Estonia" model, Model 2 has better predictive performance than Model 1.

5.2.2 Result Comparison with SOLAS2009

In order to gain better understanding of the predictive performance of the trained models, this section attempts to compare the output differences between the obtained model and the standard calculation in SOLAS 2009.The critical sea state (Hs_{crit}) is still deployed as the indicator of the prediction performance of ship survivability.

According to the recent survival "s" factor as developed in project HARDER (Tuzcu and Rusaas, 2003), it accounts for the probability of survival of a damaged vessel after flooding the compartment or group of compartments under consideration. The formula is given as equation (17)

$$s_{\text{final,i}} = K \cdot \left[\frac{GZ_{max}}{0.12} \cdot \frac{Range}{16}\right]^{\frac{1}{4}} \tag{17}$$

Where GZ_{max} is not to be taken as more than 0.12m and Range is not greater than 16 degrees.

Such a relationship has been derived from two correlations as shown in Equation (18) and (19). The former represents the relationship between GZ properties and the critical sea state Hs_{crit} , which are established through physical model experiments. The latter gives the cumulative probability distribution $F_{Hs_{collision}}$for wave height recorded at the instant of collision (Heimvik, 2001).

$$Hs_{crit} = 4 \cdot \left(\frac{GZ_{max}}{0.12} \cdot \frac{Range}{16}\right) \tag{18}$$

$$F_{Hs_{collision}} = e^{-e^{0.16-1.2 \cdot Hs_{collision}}} \tag{19}$$

Based on the correlations presented above, it appears that the current formula for Hs_{crit} has a direct impact on the assigned probability "s". Hs_{crit} is an intuitive measure to describe the ability of the damaged ship to withstand the encountered sea states. Thus comparisons of the computed Hs_{crit} between both the established models and that from equation (18) are performed.

5.2.2.1 Model fitted with damage GZ particulars (Model 1)

It can been seen in Figure 25, the distribution of the points indicates that Model 1 might overestimate the survival sea state Hs_{crit} for half of the cases, particularly for the chosen damage case on the "Estonia" model. In the meantime, it is important to notice that equation (18) is valid only for significant wave heights of less than 4 meters, as it is explicitly implied in equation (19). Hence, the points at the upper right corner indicate a greater survivability of all these cases have been estimated by both methods.

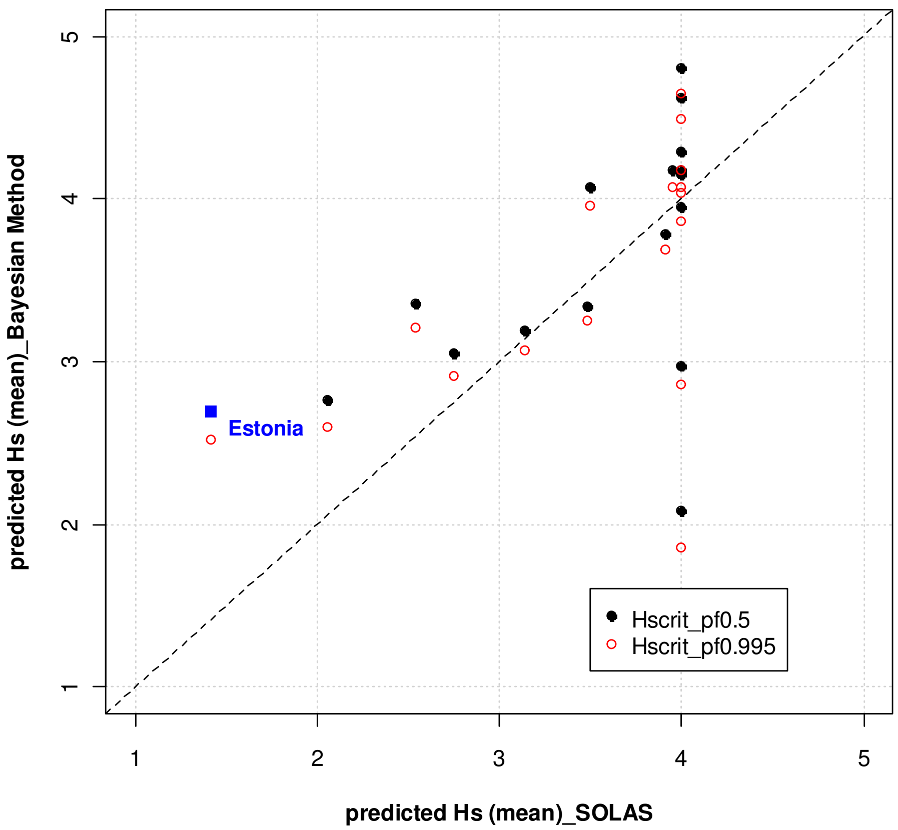


Figure 25 *Comparison of theoretical critical sea states* $(p_f = 0.5)$ *between Model 1 and SOLAS2009*

5.2.2.2 Model fitted without damage GZ particulars (Model 2)

A similar comparison can be performed between Model 2 and SOLAS estimations concerning Hs_{crit}. As shown in Figure 26, it demonstrates a poor correlation in four cases including the "Estonia" case, where Model 2 overestimates the survival sea state. On this basis, it is difficult to draw a definite conclusion. Nevertheless, a comparison of the theoretical estimation (through SOLAS) and experimentally derived $\mathrm{Hs}_{\mathrm{crit}}$ provides some interesting insight. As depicted in Figure 27, the actual survivability is not underestimated by the current formulations in Equation (18) (adopted in regulation 7-2.3 of IMO MSC.216 (82)) for some of the experimental cases. Hence, such evidence suggests that the prediction performance of Model 2 is comparable with the results of model tests as described in Figure 23.

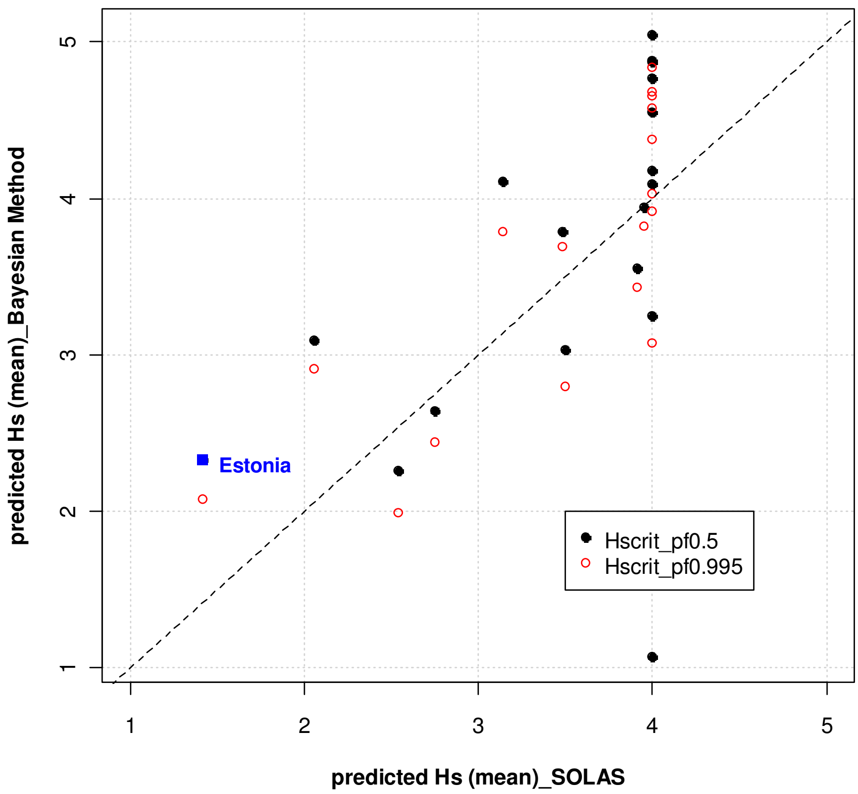


Figure 26 *Comparison of theoretical critical sea states* $(p_f = 0.5)$ *between Model 2 and SOLAS2009*

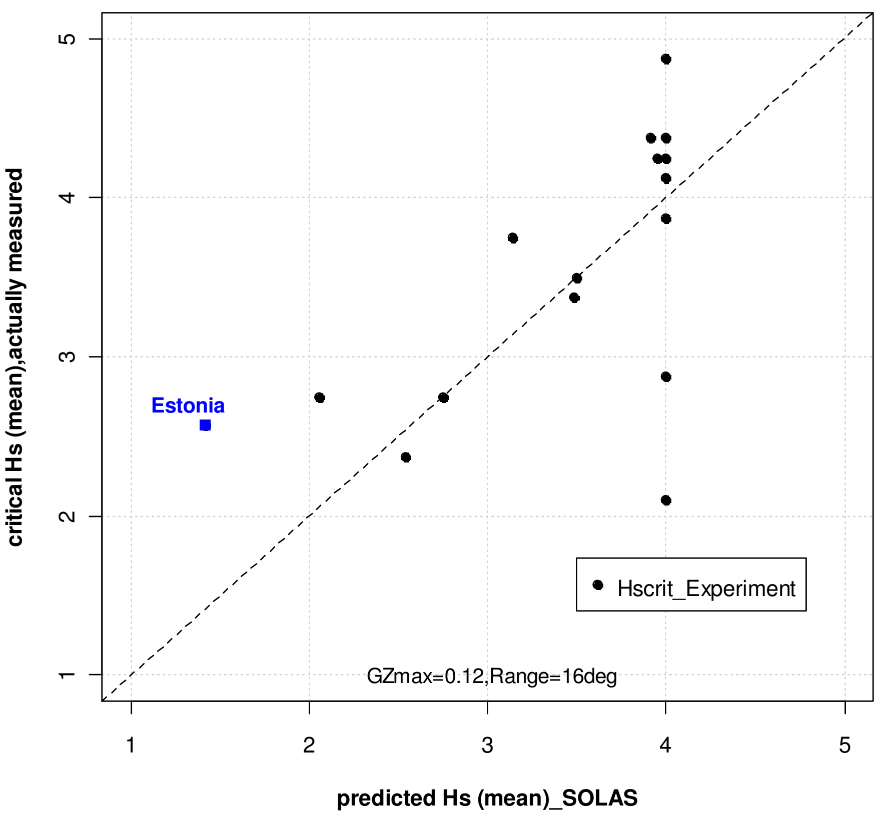


Figure 27 *Comparison of critical sea states* $(p_f = 0.5)$ *between experimental measurement and SOLAS2009*

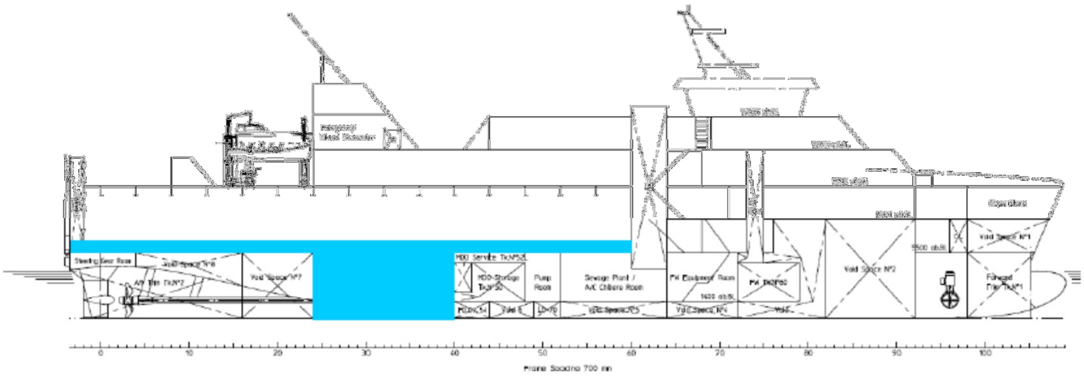
5.3 Model Testing with HSVA data

As it is described previously, the generalizability of the proposed models should be examined with some new damage cases, which are not involved in model training. On this basis, the recent physical experiments undertaken by HSVA in a research study funded by the EMSA (Project No: EMSA/OP/09/2008) are put forward for further investigation. Particular attention should be paid to Model 2 as it is more desirable to have a survivability assessment model with more accessible variables.

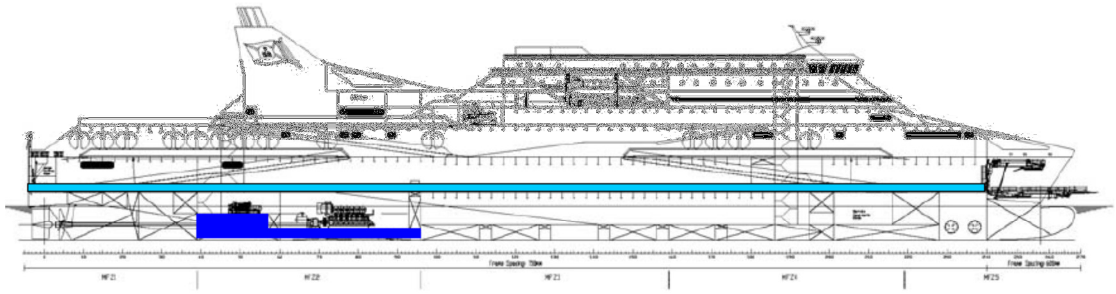
Two RoPax vessels were designed according to the new probabilistic SOLAS 2009 damage stability standard (Ludwig, 2009a) (Ludwig, 2009b). The physical model tests have been performed in accordance with the Directive 2003/25/EC. Table 9 and Figure 28 outline the information of the two tested damage cases which are used for examining Model 2.

Table 9 *Assembled experimental data for model testing*

Damage case	Measured Hs [m]	Intact Condition		Damage Condition	
		KG/KMT	T/D	Heel [deg]	Ld/Ls
EMSA1_D1	2.63 – 4.36	0.8259	0.7436	-1.0	0.065
EMSA2_D1	3.73 - 4.99	0.7596	0.7391	2.5	0.04345



(a)*An 80m RoPax vessel*



(b) *An 200m RoPax vessel*

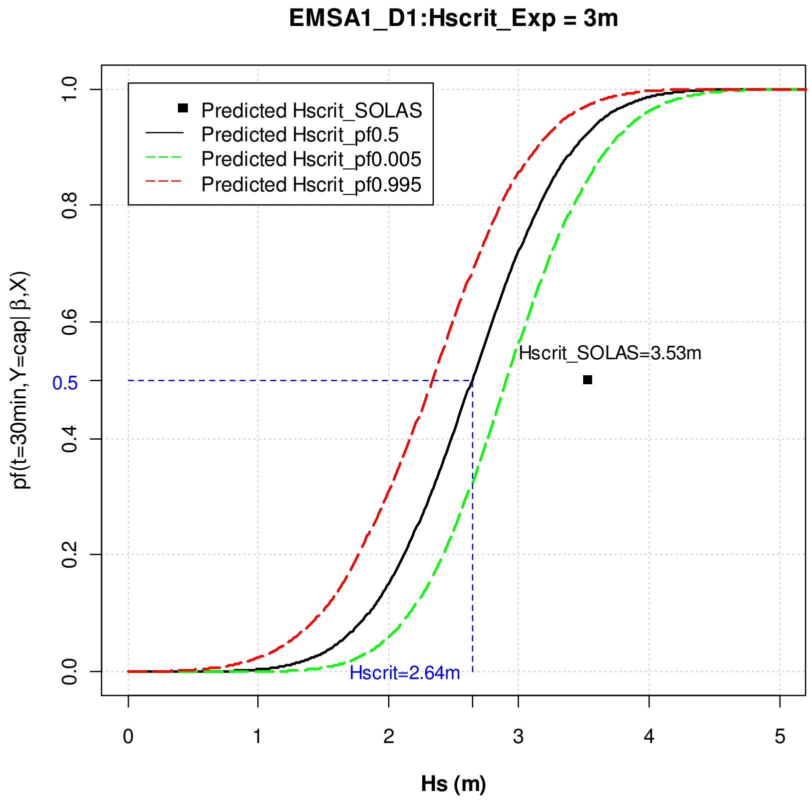
Figure 28 *Damage cases for physical model experiment*
(a) EMSA1_D1; (b) EMSA2_D1

Given values of predictor variables $\boldsymbol{X}$ for each damage case, the correlated posterior distribution of rate of capsizing $p_f(Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ can be evaluated by Model 2 directly as shown in Figure 29. Investigations of the predicted survivability are justified through the mean critical sea state Hs_{crit} .

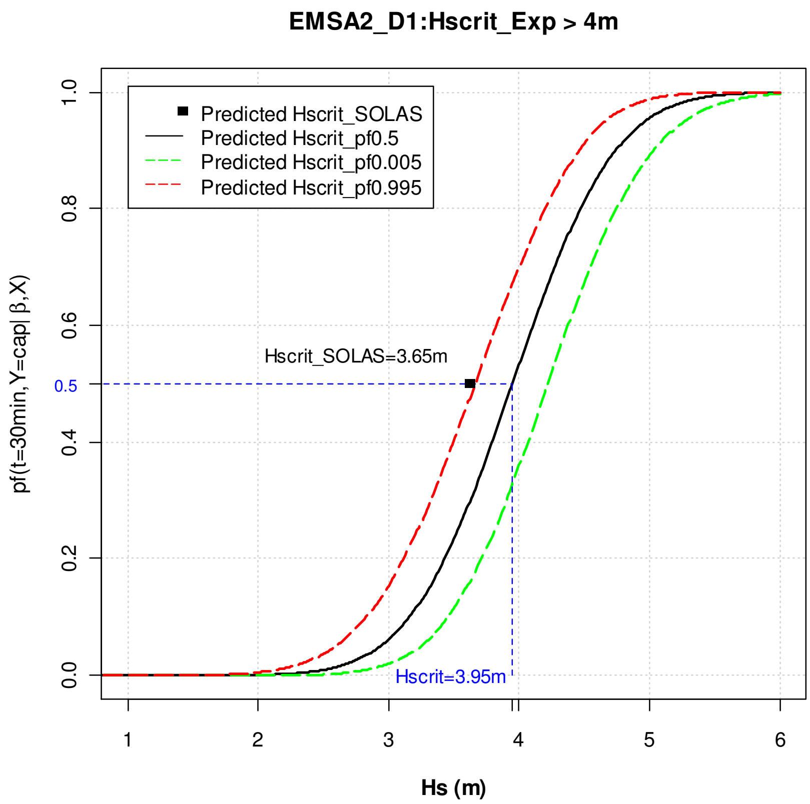
For the first damage case (denoted as EMSA1_D1), the experimental measurement of Hs_{crit} stands at 3 meters. In contrast, Model 2 suggests the value to be 2.64 meters. Moreover, the result from equation (18) (implied by SOLAS2009) is 3.53 meters, which represents a significant overestimation. It seems the current formulations of Hs_{crit} cannot always produce a conservative prediction of ship survivability.

Concerning the second damage case (EMSA2_D1), the observed survival sea state from experiments is more than 4 meters, which indicates a much higher survivability. In addition, both theoretical methods underestimate the survivability as the established survival sea state is below 4 meters. After testing the proposed model based on the two sampled cases, Model 2 can be regarded as a promising option for fast and accurate flooding predictions. Hence, the immediate question needs to be answered is how reliable the predicted survivability is by using the estimated model (i.e. Model 2). A reasonable explanation is a matter of uncertainty quantification of the response output $P(t = 30min, Y = capsize|\boldsymbol{\beta}, \boldsymbol{X})$.

As can be seen from Figure 20, 23, 25, 26 27, and 29, different assessment approaches present comparable but inconsistent estimations for the benchmarking tested cases. The deviations of the predictions need to be quantified in this case. An effective way is to ensure the proposed method provides more conservative predictions than others. This can be achieved by shifting the computed Hs_{crit} to the right hand side of the comparison diagram and below the diagonal line as shown in Figure 31. The following chapter emphasizes on the need for uncertainty quantification.



(a)



(b)

Figure 29 *EMSA1_D1 & EMSA2_D1, posterior distribution of rate of capsizing within the test duration* $t_0 = 30$ *minute*

$$p_f(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X} = H_s, KG/KMT, T/D, Heel, L_d/L_s)$$

6. UNCERTAINTY QUANTIFICATION

Because of the existed deviations among different assessment approaches, the ultimate goal is to define a proper method which can guarantee a comparable and relatively conservative estimation of the survivability for the majority of considered flooding cases. In this respect, a predictive model must be estimated. Attempting to achieve a balance between the prediction accuracy and the computational effort, a fast and accurate analytical model is more desirable. It now appears that most of the models used for dynamic survivability assessment are computer programs to better describe the physic phenomenon of ship capsizing, thus regression analysis can be applied to produce an analytical expression. In this way, based only on a few dominant input variables, the more complex computer model can be represented. Apart from getting a fast estimation, the uncertainty associated with the model response (output) can be calculated by propagating the uncertainty in the model inputs. General procedures of uncertainty analysis and sensitivity study constitute the core of this chapter for discussion.

6.1 Probabilistic Uncertainty Analysis

Typically uncertainty is classified into aleatory uncertainty and epistemic uncertainty (Hacking, 1975). Aleatory uncertainty is also called stochastic uncertainty which arises from natural variability. Epistemic uncertainty is defined as a knowledge-based uncertainty that represents the lack of knowledge. The most important distinction between these two types of uncertainty, at a practical level, is that the latter can be reduced by further study. With respect to the first principle models undertaking the performance-based survivability assessment, the inherent uncertainties in modelling are commonly divided into two major groups as the parameter uncertainty and the model uncertainty. For instance, the values of input parameters used in models may not be known accurately, and the parameters used in a model may be subjected to natural variability. On the other hand, the model uncertainty arises from the fact that any model inevitably is a simplification of the reality it is designed to represent, precise information is unavailable. As Katherine Laskey put in her lecture notes on probability in Artificial Intelligence: “All models are wrong, but some are useful” (Laskey, 1994). Being aware of these, the treatment of both the group of parameter and model uncertainty must be allowed for in reducing the uncertainties caused by vague input parameters and imperfect models.

The general procedure of uncertainty analysis follows the propagation of the uncertainty through a model as shown in Figure 30. Probabilistic treatment of uncertainties is the most widely used in this process (Abrahamsson, 2002). In this example, the uncertainty in the model output variable G is derived according to the propagation of the uncertain variables f_1, f_2 and f_3 through the model *function* (f_1, f_2, f_3). The parameter uncertainty displayed at the top level is specified as probability density functions. The first subjected problem is to identify which parameter should be included in the model, and then to propose a measure for assigning the probability distribution of each model input to characterize its inherent variability. Moreover, with respect to the model uncertainty, the fairly common measure of treatment is to make use of several parallel models to enhance credibility in the results.

As with the schematically described uncertainty propagation, the uncertainty analysis is composed of two steps in this context. First, a predictive regression model has been identified in Chapter 3 which accepts the uncertainty transmission while the information of model inputs is available. Second, the probability distribution of each model input has been simulated through Bayesian inference as clarified in Chapter 4.

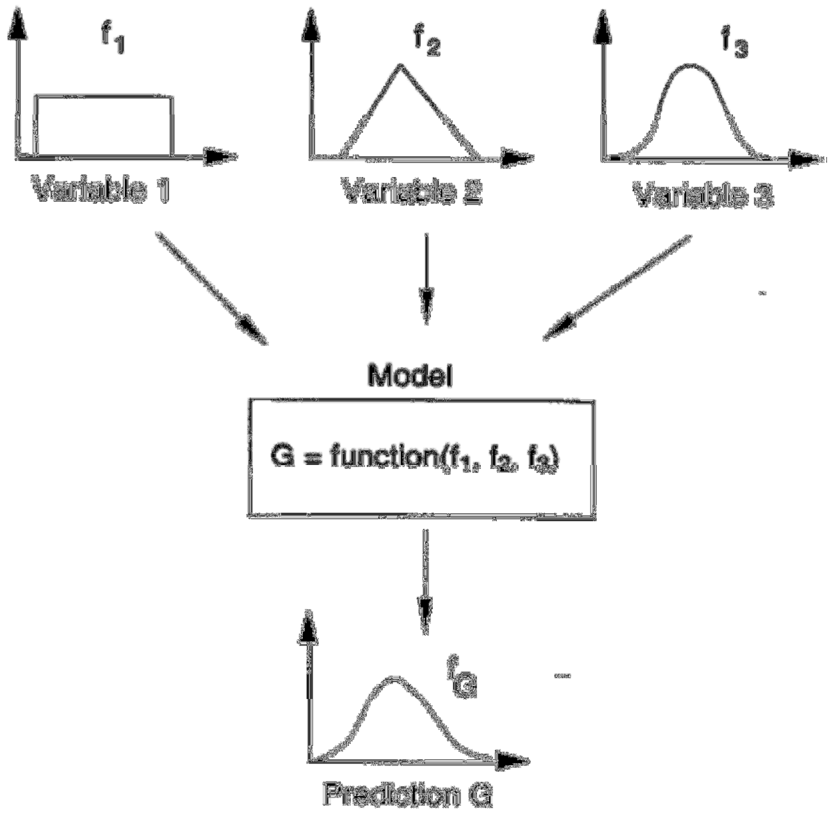


Figure 30 *Propagation of uncertainty through a model*

Since the underlying uncertainty of the model response (output) needs to be quantified, it implies uncertainty bounds of the obtained results have to be estimated on the basis of an acceptable level. By doing so, an acceptable lower bound of the survivability prediction can be regarded as a conservative solution. And therefore, the epistemic type of uncertainties seems to be mitigated mostly when considerable conservatism built into the construct of the conditional probability of ship capsizing in given environment.

In the current study, additional attention is paid to the established Model 2 which uses more accessible operational parameters to characterize the instantaneous status of a ship subject to damage. Based on the discussions in the previous sections, quantifying uncertainty of the model response $P(t_0, Y = capsize \,|\, \boldsymbol{\beta}, \boldsymbol{X})$ needs to be translated into measuring the variations of the component of the predictor variables $\boldsymbol{X}$. The impact of each input variable on the model response is reflected through the corresponding coefficient $\boldsymbol{\beta}$ assessed. In this situation, it is noted that the variation in the model response given input variables is connected explicitly with the variation in the model coefficients. Thus, it is vital to understand that uncertainty quantification on the interested response variable should be sought through establishing uncertainty bounds associated with each model coefficient.

According to Bayesian inference methods and MCMC sampling algorithm, the target probability distributions of model coefficients $P(\boldsymbol{\beta}|y)$ can be approximated. Following this, assigning uncertainty bounds of them is just a matter of quantile estimation as explained in section 4.3.2. The sampled mean value of $\boldsymbol{\beta}$ can be utilized in equation (20) to derive a relationship for survivability prediction as given in equation (21).

$$P(t_0, Y = capsize \,\big|\, \boldsymbol{\beta}, \boldsymbol{X}) = \Phi(\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n) \tag{20}$$

Depending on the estimated results shown in Table 8, Model 2 is defined as

$$P(t_0, Y = cap \,\big|\, \boldsymbol{\beta}, \boldsymbol{X}) = \Phi(-57.2 + 1.62H_s + 20.75\frac{KG}{KMT} + 42.8\frac{T}{D} + 0.41Heel + 54.6\frac{L_d}{L_s}) \tag{21}$$

Regarding the result comparisons in section 5.2, the uncertainty boundary of model response (rate of capsizing within 30 minutes) has been assigned by using the 99% posterior interval of all model coefficients. In theory, a conservative assessment can be achieved once the upper limit of the estimated ship vulnerability to flooding is obtained. However, in comparison to the experimentally and theoretically derived results, the proposed model might sometimes provide over-estimation of the actual survivability. Thus the degree of the confidence of the model response needs to be enhanced. In order to meet such requirements, the following issues concerning the model should be addressed:

- The chosen predictor variables to represent the stochastic damage characteristics have to the dominant ones.
- The variation of the values of predictor variables in the assembled data set should be flexible enough to reflect a wide range of possible damage scenarios

On the basis of the foregoing, it appears that the proposed model needs to be further investigated and refined in the light of the underlying parameter uncertainty. Nevertheless, model selection still needs to restrict the regression model (Probit) to those critical influential variables as the quality of model may be reduced significant if too many variables are taken into account simultaneously. It is noted that the input variable L_d/L_s in Model 2 may not be adequate enough to describe a specific hull breach, which is normally characterized by a group of variables including location *x*, length λ, penetration *b* and height *h*. Fortunately, the Bayesian inference technique combined with MCMC method is a robust manner for processing data analysis. In other words, modelling an acceptable relationship on the basis of the available data and knowledge still can be achieved. On the other hand, due to the benchmarking experiments are performed according to the Directive 2003/25/EC, thus the modelled flooding extent (up to two zone damage) is limited. Lacking of variations of the influential variables has a great effect on the assembled data set. As a result, the computed model coefficients may deviate from the actual situations. In order to appreciate the input variable that results in the greatest impact on the model output, the last part of this chapter presents a particular sensitivity study of the rate of ship stability loss in given flooding cases with reference to the different input parameters pertaining to Model 2.

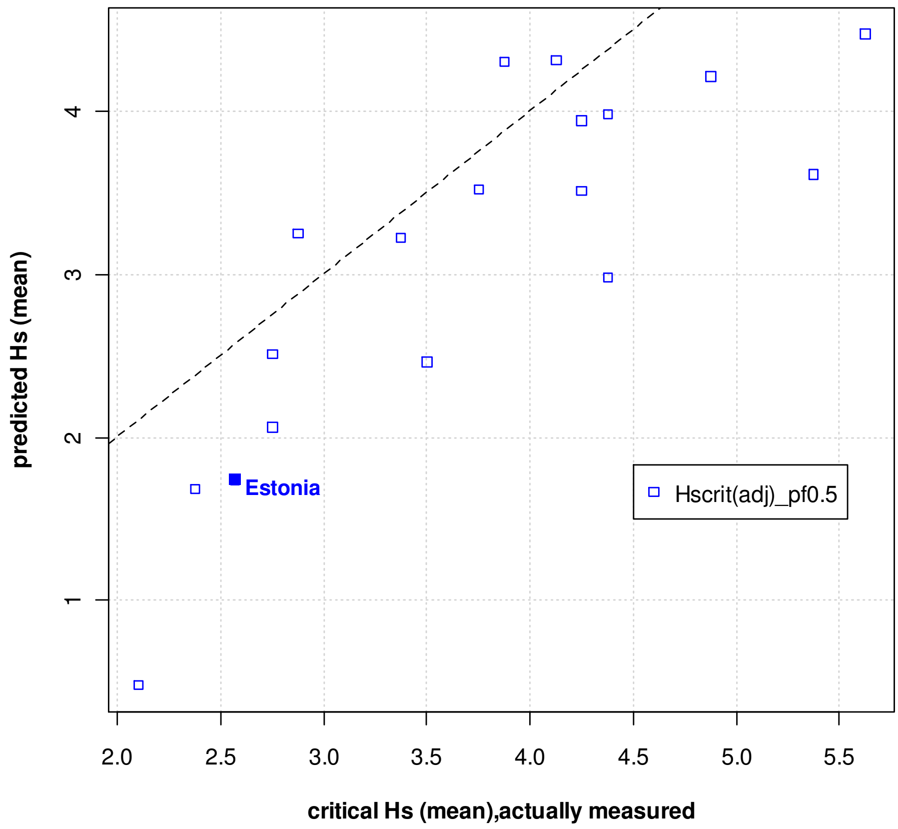
The difficulties raised above disclose that the established Model 2 inevitably subjects to the parameter uncertainty. For the purpose of approaching a conservative prediction while with high reliability, a practical way is to adjust the variation of model coefficients within their uncertainty bounds until an acceptable level of the survivability prediction is reached.

With the simulated $P(\boldsymbol{\beta}|y)$ as outlined in Figure 18, it is reasonable to assume the posterior distribution about each β_i is approximately normal. In this case, tuning the percentile of model coefficients is a reasonable attempt to build up a conservative relationship for survivability assessment. As it is illustrated in Table 10, 55% posterior quantiles of $P(\boldsymbol{\beta}|y)$ have been selected to replace the present sample mean (50% quantile) for establishing a new relationship.

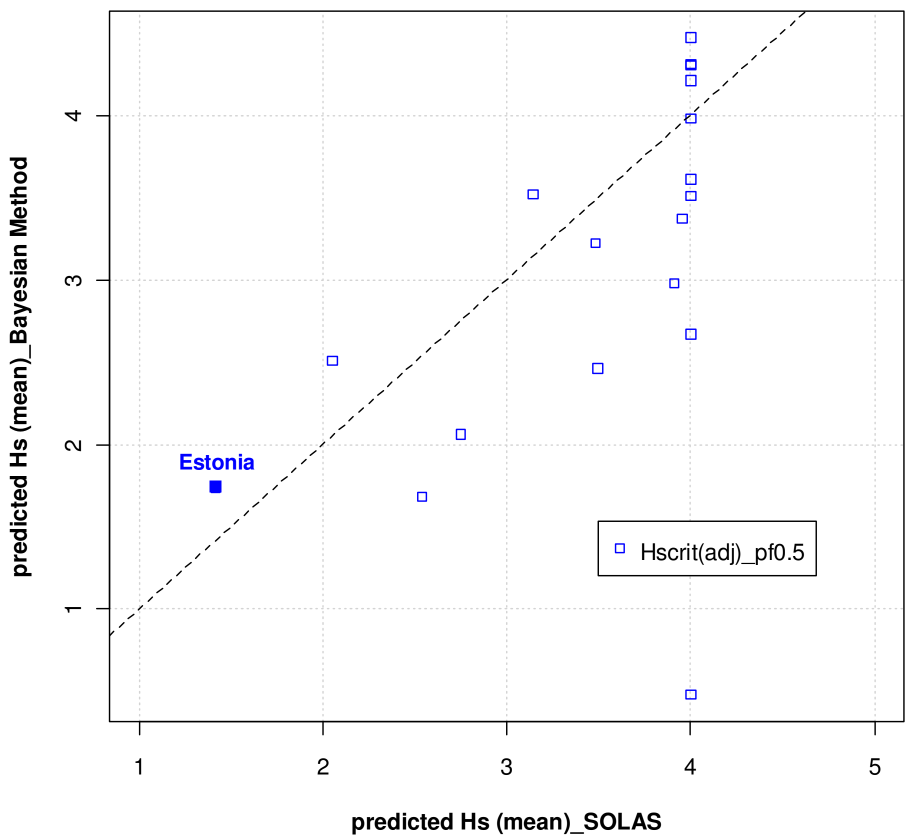
Table 10 *Adjustment of Model 2 coefficients*

Parameter	Mean	SD	0.55 quantile
α (interception)	-57.2016	2.9094	-56.836
β_1 for H_s	1.6203	0.0717	1.6293
β_2 for *KG/KMT*	20.7519	1.8823	20.9884
β_3 for *T/D*	42.8016	2.7077	43.1418
β_4 for *Heel*	0.4126	0.0742	0.4219
β_5 for L_d/L_s	54.6001	12.3074	56.1467

By substituting the equation (21) with the adjusted model coefficients, the distributions of the rate of capsizing against the encountered sea states $p_f(Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ can be estimated for all those tested cases again. Subsequently, the correlated mean sea states Hs_{crit} can be compared with the experimental and the theoretical values as shown in Figure 31. In contrast to Figure 23 and Figure 26, it is apparent that the obtained results are conservative following the slight tuning of the coefficients.



(a)



(b)

Figure 31 *Comparisons of mean critical sea states after model adjustment*

As presented before, the method of quantifying the inherent uncertainty in the established model itself is achievable. Once an acceptable conservative relationship is confirmed, from an operational point of view, an instantaneous prediction of the ship vulnerability to flooding against time can be addressed easily. Moreover, from a design point of view, assigning the probability of survival accommodates for all possible hull damages can be considered. The following part takes the RoPax vessel "M/V Estonia" as an example to demonstrate the trends.

In the latest probabilistic damage stability standards of SOLAS 2009, the factor "s" is defined as a measure of the probability of survival of a damaged ship in waves which allows for all the feasible flooding cases along the ship length. In light of this, "waves" represents a range of sea conditions that might be encountered whilst suffering a collision incident, thus "s" has the following expression:

$$s = \int_{0}^{H_s} dH_s \cdot f_{H_s|coll}(H_s) \cdot F_{surv}(H_s) \tag{22}$$

Where $f_{H_s|coll}(H_s)$ denotes the probability density distribution for sea states expected to be encountered during a collision incident. Its integral solution $F_{H_s|coll}(H_s)$ is approximated by equation (19). $F_{surv}(H_s)$ represents the probability that the ship will survive in that sea state for a given a flooding case.

The proposed relationship in equation (20) gives the probability that the ship will capsize for a specific damage within given time ($t_0 = 30$ minutes). So the corresponding survivability is expected to decrease with increasing sea state, and to vary from 1 to 0 as stated below.

$$F_{surv}(H_s) = 1 - P(t_0, Y = cap \,|\, \boldsymbol{\beta}, x = H_s, given\ other\ X) \tag{23}$$

Comparing with the recent formula adopted in SOLAS 2009, $F_{surv}(H_s)$ for given 30 minutes is simplified as:

$$F_{surv}(H_s) = \begin{cases} 1, & H_s \le Hs_{crit} \\ 0, & H_s > Hs_{crit} \end{cases} \tag{24}$$

Where the mean critical sea state Hs_{crit} is given in equation (18) proposed in HARDER project. In this case, the integral (22) can be solved in a simple way as below:

$$s = \int_{0}^{H_{scrit}} dH_s \cdot f_{H_s|coll}(H_s) = F_{H_s|coll}(H_{scrit}) \tag{25}$$

Depends on the above assumption, the probability "s" has been assigned as the cumulative probability distribution for the wave heights recorded during collisions. Consequently, two questions are put forward of 1) is the assumption behind equation (24) adequate to approximate the observable survivability? And 2) since all experimental tests on survivability for given a flooding case were preformed for 30 minutes, the current probability "s" is actually modelled as "$s(t = 30min)$". In this case, is such evaluation time sufficient to define the ship survivability for longer time? A series of discussion and comparison have been made in SSRC and proved that equation (25) has favourable response to elicit a rational survivability assessment.

At this time, the assigned probability "s" for a specific flooding only depends on the input parameter H_{scrit}. As shown in equation (18), H_{scrit} is formulated by the GZ curve properties at the damage equilibrium stage, which derived based on the benchmarking tests performed in project HARDER. After that, the probability "s" can be transferred as equation (17). Explicitly, this recent formula of "s-factor" is not linked with the sea environment encountered at the instance of collision H_s directly, but it uses an intermediate measure H_{scrit} to the damaged stability of the ship. Hence, the accuracy of the relationship given in equation (18) has a great effect on the assigned probability "s". For the purpose of keeping the knowledge-based uncertainty to a minimum, a new model assessing H_{scrit} has been proposed in Task 4.2 by modifying parameters of equation (18) to GZ_{max} = 0.25m and Range = 25 deg.

Other than the conventional method for evaluating the factor "s", regarding the established model in equation (21), there is no H_{scrit} to be considered any more. Currently, straightforward input information of the sea conditions during a collision is needed. Thus the transition in $F_{surv}(H_s)$ from 1 to 0 can be identified and the probability "s" for a specific flooding within 30 minutes can be assigned as Figure 32.

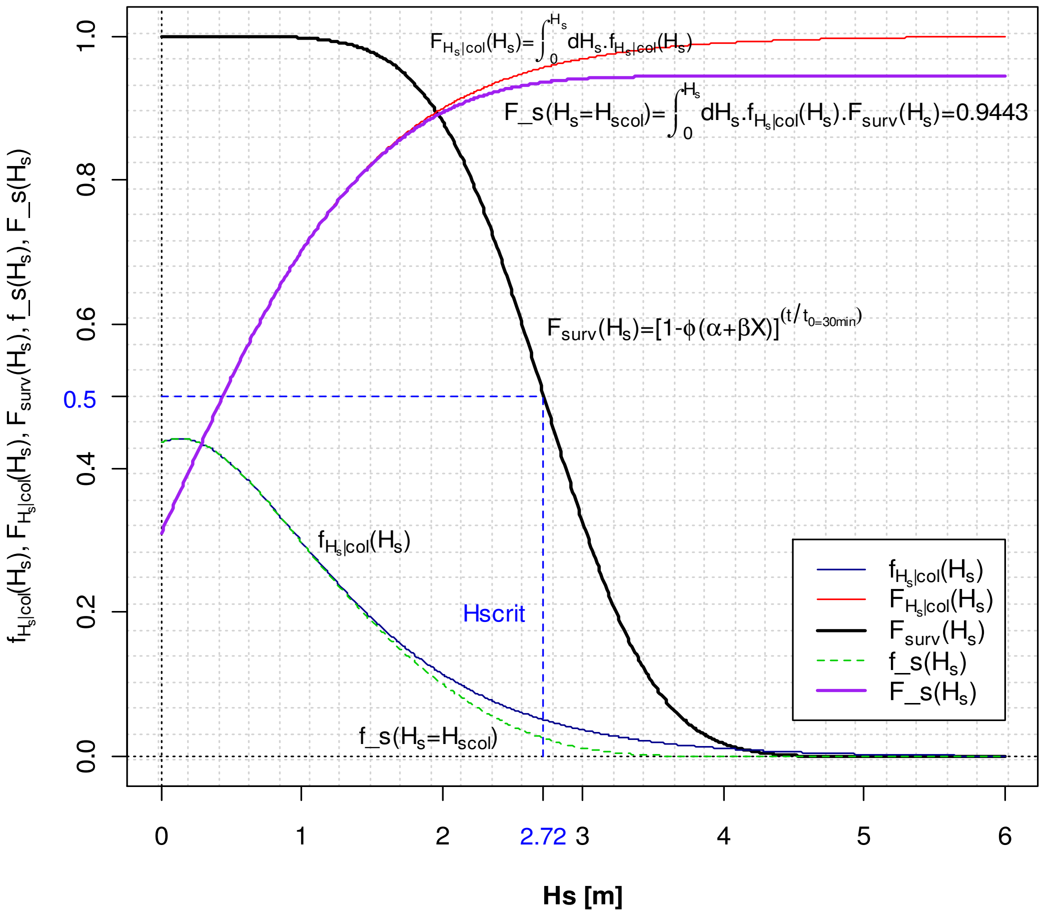


Figure 32 *The process of assigning the probability s for a specific flooding case*

Based on the expression of equation (22), it appears that for a specific flooding case, the information required for assigning the probability “s” within given time actually comes from two aspects: 1) the density distribution for a range of sea states supposed to be encountered during a collision, 2) the probability of survival of the damaged ship corresponding to each considered sea state. A more clear and logical way to explain such relationship is made in Figure 32, it can be seen that for a flooding case where the measured $H_{scrit} = 2.72m$, the probability “s” determined as an averaged survivability below $H_s = 4m$ is equal to 0.9443, based on equation (15) for $t = 30$ minutes.

With the demonstration of the new method (Model 2) as a suitable way of assessing the probability “s”, the remaining part provides the comparison between the new estimated “s” from Model 2 and the one measured complying with the probabilistic damage stability standards of SOLAS 2009. The latter computation is performed through the ship design software NAPA. Focusing on the initial condition of “M/V Estonia” under the deepest draught (see Table 4), at this time, there are 1200 feasible flooding cases extended up to 5 zones’ damages along the ship length. Regarding the original established Mode 2 equation (21), Figure 33 shows the difference when the new “s” compared with SOLAS’ results.

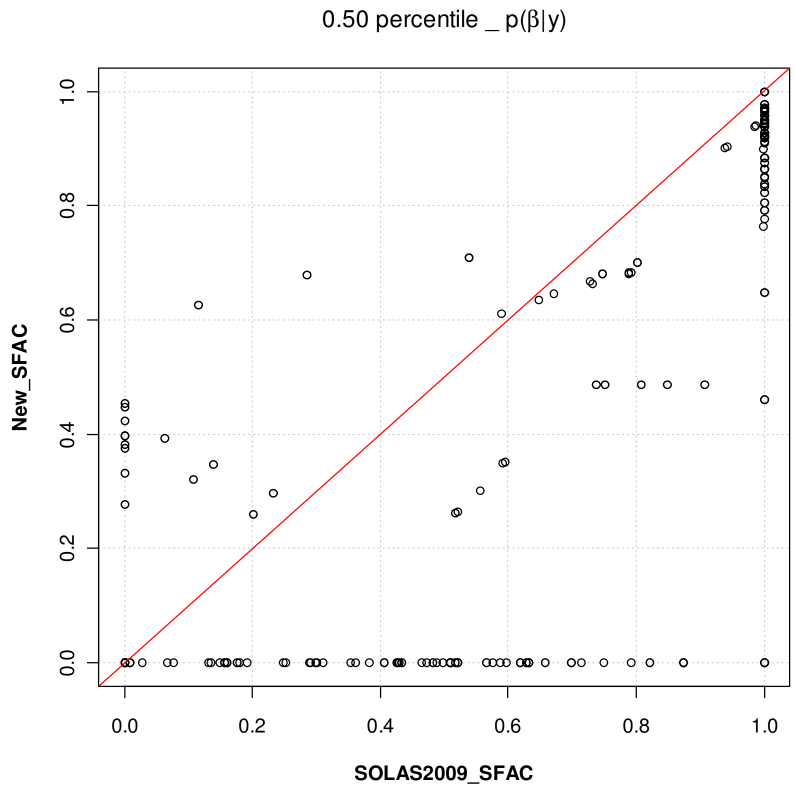


Figure 33 *Comparisons of “s-factor” between the original Model 2 (50th percentile of* $\boldsymbol{\beta}$*) and SOLAS’s results for 1200 expected flooding cases*

Table 11 *Summary of "s-factor" comparison for Figur33*

"s-factor" comparison	No. of case	% out of 1200 cases
New > SOLAS 2009	39	0.0325
Top Left (SOLAS<0.2 &New >0.2)	29	0.0242
Bottom Right (SOLAS>0.2 & New <0.2)	83	0.0692

When comparing the original Model 2 with SOLAS 2009 estimated "s-factors" for the total 1200 flooding cases, the results depicted in Figure 33 are against the observations in Figure 26. Regarding the expressed survivability of the RoPax vessel "M/V Estonia", Figure 26 using Hs_{crit} as a simple measure indicates that Model 2 produces an overestimation in contrast to SOLAS results. However, referring to the summary of Table 11, it is obvious that Model 2 is stricter than the recent standard of SOLAS 2009, as only 39 flooding cases have been over-predicted.

At the beginning of this section, it has been explained that quantifying the uncertainty in Model 2 can be achieved by establishing uncertainty bounds on each model coefficient. Thus to ensure Model 2 is capable of providing conservative predictions with high reliability, a pragmatic solution is to adjust slightly the variations of model coefficients within their uncertainty band. Table 10 presents the adjusted Model 2 after tuning all coefficients from the sampled mean to 55% quantile of $P(\boldsymbol{\beta}|y)$. By doing so, a new comparison of "s-factors" can be updated between the adjusted Model 2 and SOLAS's. Apparently, both Figure 34 and Table 12 disclose considerable conservatism has been built into Model 2 when all $\boldsymbol{\beta}$ adjusted simultaneously.

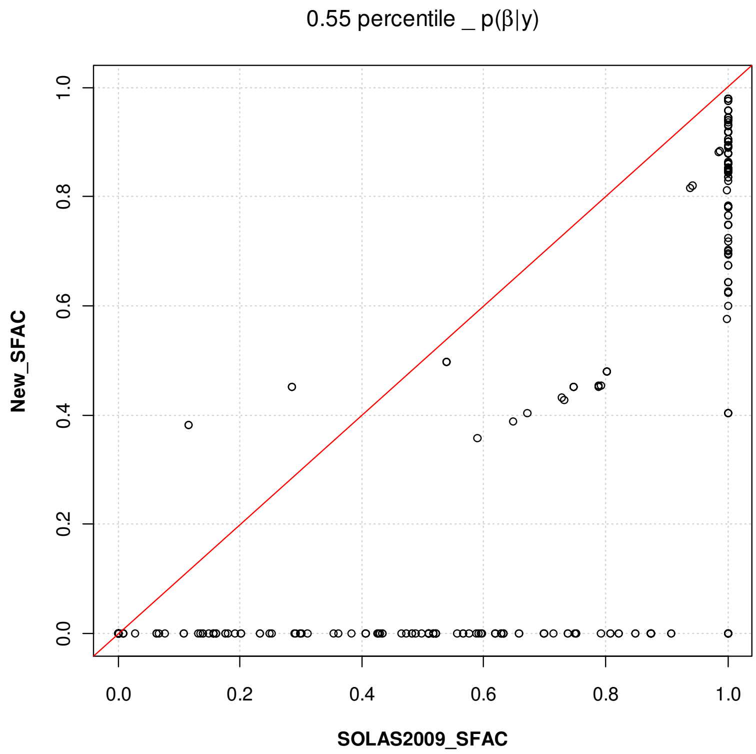


Figure 34 *Comparisons of "s-factor" between the adjusted Model 2 (55th percentile of $\boldsymbol{\beta}$) and SOLAS's results for 1200 expected flooding cases*

Table 12 *Summary of "s-factor" comparison for Figur34*

"s-factor" comparison	No. of case	% out of 1200 cases
New > SOLAS 2009	4	0.0033
Top Left (SOLAS<0.2 &New >0.2)	2	0.0167
Bottom Right (SOLAS>0.2 & New <0.2)	107	0.0892

In accordance with the observations from Figure 34, it is noted that the adjusted Model 2 is too conservative to be applied in practice. More efforts should be made to investigate the hidden problems with the original Model 2. As mentioned in the foregoing, there is still much room for improving the selection of variables. Presently, the variable L_d/L_s is adopted for describing the character of hull breaches. The nature of this variable must be clarified. 1) The data of damage length L_d is assembled for Model 2 training. However, those assessed 1200 flooding cases are described by flooding extent along with the subdivision boundaries. 2) The damage length of the benchmark tested cases spans up to 2 compartments according to the guidelines in the Annex of the Stockholm Agreement (Directive 2003/25/EC). Nevertheless, SOLAS 2009 accommodates for all feasible damage lengths. Thus the variation of flooding extent during model testing is too limited comparing with SOLAS 2009. Obviously, these two methods seem not comparable in some cases. Especially for the damages have large flooding extent.

In light of the problems defined above, it is apparent that the current expected value of $\beta_{L_d/L_s} = 54.6001$ is not appropriate to reflect the actual situation. Hence, its related uncertainty needs to be minimized. A further model adjustment is needed against this coefficient. As a result, β_{L_d/L_s} is tuned by following the approximated normal distribution $(\mu = 54.6001, \sigma = 12.3074)$. It turns out that the estimated "s-factors" by making use of Model 2 towards the actual situation while decreasing the value of β_{L_d/L_s}. In order to achieve a balance that the model is relatively conservative and to reflect the reality, an attempt has been made by using 0.1 quantile of β_{L_d/L_s} and taking 0.59 quantile of the remaining parameters. Table 13 lists the readjusted $\boldsymbol{\beta}$ and Figure 35 outlines the updated "s-factors" estimated through the readjusted Model 2.

Table 13 *Readjustment of Model 2 coefficients*

Parameter	Mean	SD	Tuning $\boldsymbol{\beta}$'s quantile	
α (interception)	-57.2016	2.9094	0.59	-56.5396
β_1 for H_s	1.6203	0.0717	0.59	1.6366
β_2 for KG/KMT	20.7519	1.8823	0.59	21.1802
β_3 for T/D	42.8016	2.7077	0.59	43.4177
β_4 for $Heel$	0.4126	0.0742	0.59	0.4295
β_5 for L_d/L_s	54.6001	12.3074	0.10	38.8275

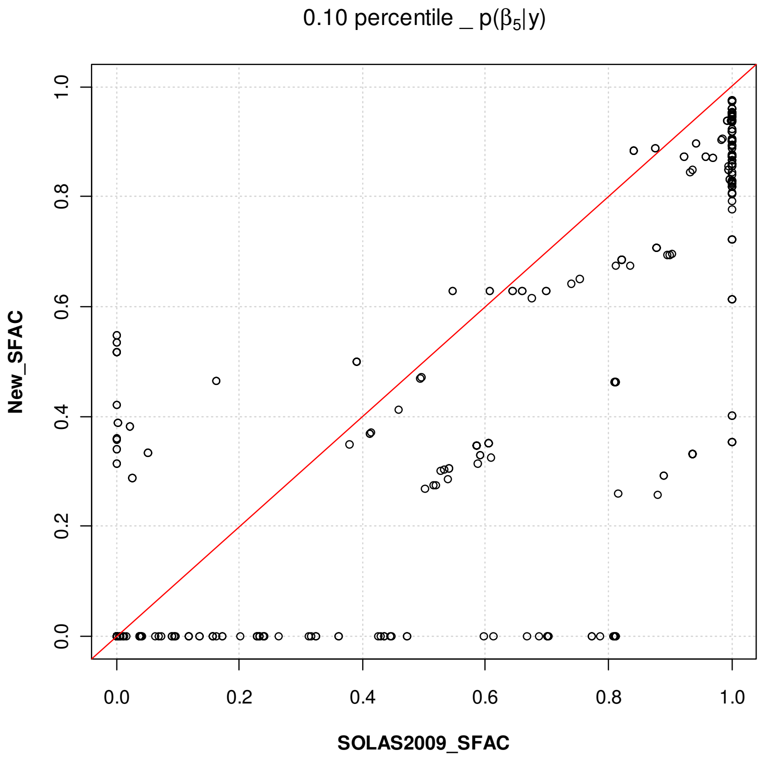


Figure 35 *Comparisons of "s-factor" between the adjusted Model 2 (10^{th} percentile of β_5) and SOLAS's results for 1200 expected flooding cases*

Table 14 *Summary of "s-factor" comparison for Figur35*

"s-factor" comparison	No. of case	% out of 1200 cases
New > SOLAS 2009	44	0.0367
Top Left (SOLAS<0.2 &New >0.2)	30	0.0250
Bottom Right (SOLAS>0.2 & New <0.2)	42	0.0350

For the RoPax vessel "M/V Estonia" given a specific loading as the deepest draught (DS) occurred, Table 14 represents the survivability of almost 4% flooding cases are overestimated by employing the adjusted Model 2. Further interpretation of such modified model is still requested, especially for those cases with "s" locate at the top left (non- conservatism) and bottom right (lesser survivability) of Figure 35.

This section reiterates uncertainty quantification on the established model (Model 2) for survivability prediction is essential. Uncertainty existed in the model output is mainly caused by the uncertainty of the model inputs. By this way, the problem is finally converted to quantify the uncertainty in each model coefficient which accounts for the impact of each variable on the model output. On the basis of analyzing the rationality of current variables and the flexibility of the assembled data, the most uncertain variable has been pointed out as the flooding extent. Then a pragmatic solution is put forward as tuning the variation of the pertinent coefficients. Comparisons of estimated "s" between the new model and the stability standard SOLAS 2009 (Figure 33 to Figure 35) suggest that the model after adjustment is able to meet the target while taking a compromise of the estimation between conservative and reliability.

6.2 Probabilistic Sensitivity Study

Model 2 is derived from Bayesian data analysis for a binary regression model. However, when using such regression analysis to produce a model with an analytical expression, it is imperative to make sure that the variables used in the model are the ones of the most interested for the uncertainty analysis, and that the model is not used outside the parameter range defined by the regression analysis. Being aware of this, a systematic and thorough sensitivity study is needed for understanding how do individual inputs pertaining to Model 2 contribute to the uncertainty in the output.

Based on an extensive discussion in the previous section, it is exposed that some inherent parameter uncertainties existed in Model 2, especially the extent of flooding during crises seems to be the most critical knowledge-based (Epistemic) uncertainty prevailed. After tuning the variation of the model coefficients within its range, the latest Model 2 complying with Table 13 is adjusted as

$$P(t_0, Y = cap \,|\, \boldsymbol{\beta}, \boldsymbol{X}) = \Phi(-56.54 + 1.64H_s + 21.18\frac{KG}{KMT} + 43.42 + 0.43Heel + 38.83\frac{L_d}{L_s}) \qquad (26)$$

Referring to Chapter 4 of Deliverable 4.2, it described a set of external parameters affecting ship stability after flooding could be divided into three parts, (a) the hull breach, characterised by set $\Omega = \{x,\lambda,b,h\}$of its location, length, penetration and height, (b) ship draught T and (c) the environment expected in a collision $H_s|coll$. Corresponding to the input variables included in equation (26) above, the variations of the sea environment H_s , the initial loading at a specific operational draught (KG/KMT) and the damage attributes ($Heel$, L_d/L_s) are going to be used for understanding how changes in each of them influence the rate of ship stability loss in an actual casualty case.

In subsequent parts of this section, first, a global sensitivity study is conducted to rank the impacts of model inputs on the output since a wide spread of all the inputs is required. In such a case, all the data collected in a series of physical model tests are referred to Table 7. Second, due to the input variable L_d/L_s in equation (26) is the only datum employed to represent the extent of flooding. In reality, joint information $\Omega = \{x,\lambda,b,h\}$ charactering the hull breach should be allowed for. A local sensitivity study is performed to identify the importance of considered variables (H_s, $Heel$, Ω) in Model 2 at a specific draught for the uncertainty analysis. In contrast to the study at a ship level, the data applied at this time come from the numerical simulation as given in Table 6 and Table 7 of Deliverable 4.2. Therefore, the randomness of the extent of flooding based on Monte Carlo sampling has been simulated.

6.2.1 A Ship Level Investigation

Given the updated Model 2 in equation (26), all the model coefficients are considered to be constant. The next step is to define the range of variation for all the model inputs.

- *The sea environment:* In accordance with sea states performed in Task 4.1, the variation of the sea environment H_s is assumed to follow a uniform distribution on the interval from 2m to 3m.
- *The centre of gravity:* The initial loading condition at a specific operation draught is represented by a ratio of the centre of gravity (KG) over the transverse metacentre (KMT), the variation of such non-dimensional input is assumed to follow a normal distribution, the relevant mean and standard deviation are sampled from the experimental data set outlined in Table 7.
- *The heel angle at the damaged equilibrium stage:* Since all feasible flooding extents are considered in compliance with SOLAS 2009, according to the discussed 1200 flooding cases on MV Estonia generated by NAPA in the prior section 6.1, the floating position of the ship for each damage can be achieved through a damage hydrostatics calculation. In this study, heel angles corresponding to flooding cases up to 4-zone damages are selected and the deviation of such variable is assumed based on a truncated normal distribution.
- *The extent of flooding:* First, suppose that the damage length L_d is the single parameter to be used for charactering the extent of flooding. Second, in order to avoid the variation in damage lengths is limited by the model testing, resembling the collection of heel angles, the variation of the flooding extend is assumed to have a truncated normal distribution. All feasible sizes of flooding up to 4-zone damages are taken into account.

In response to the foregoing assumptions, the values of all the model inputs are simulated using a sample size of 10,000. The range of variation for each of them is clearly expressed in Figure 36. Next, substituting the equation (26) with the sampled values of one of these four model inputs, meanwhile, the rest parameters keep using sample means as the inputs. Thus, the variation of the rate of ship capsizing $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ against the assessed input variable x_i can be estimated in turn. Table 15 gives a summary of the sampled input values coming from Figure 36. Both the largest observations of the heel angle and the extent of flooding L_d/L_s are 12.568 deg and 0.258 respectively.

Table 15 *Summary of the sampled values of various model inputs*

	Mean	SD	0.50%	50%	99.50%
Hs	2.4989	-	2.0057	2.4965	2.9964
KG/KMT	0.8430	0.0379	0.7450	0.8436	0.9375
Heel	3.3210	2.1085	0.0390	3.0851	9.8228
Ld/Ls	0.0892	0.0425	0.0036	0.0876	0.2069

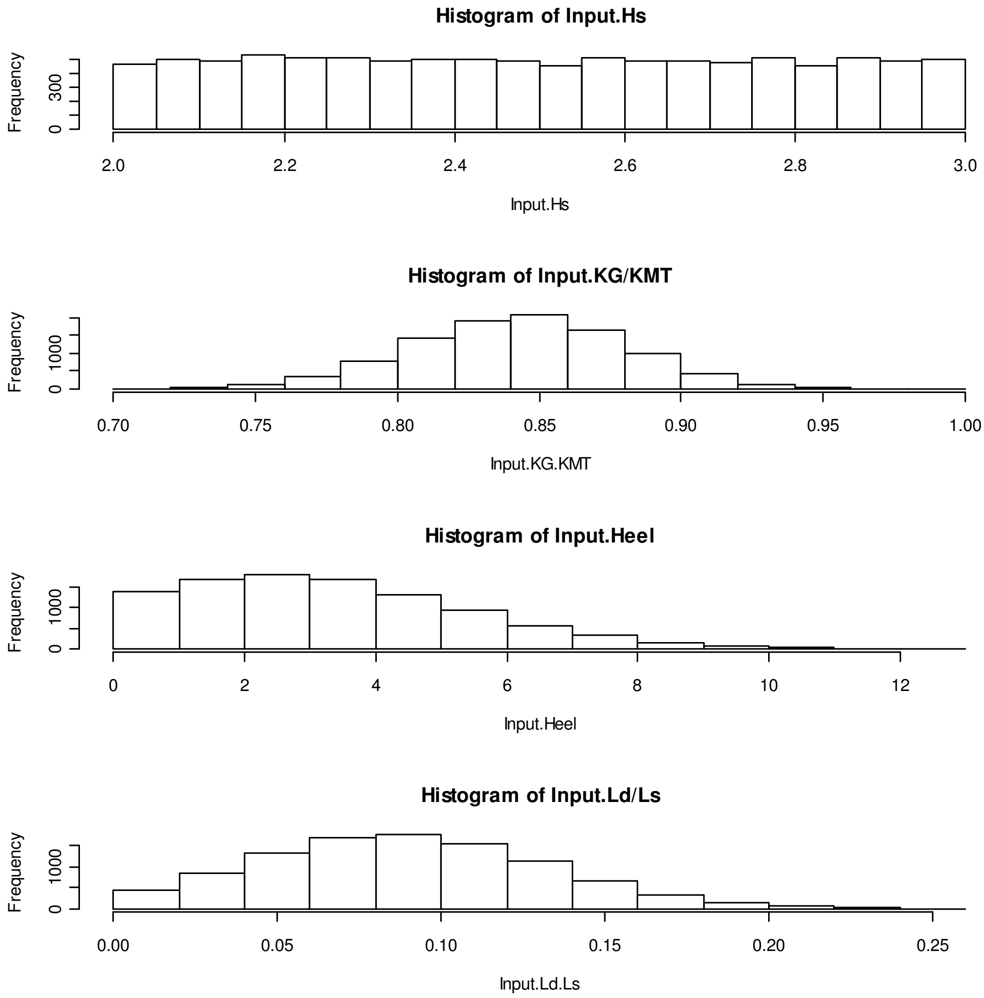


Figure 36 *Histograms of model input values, using a sample size of 10,000*

Based on solving the adjusted Model 2 in equation (26), 10,000 outputs $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ are obtained for describing the rate of capsizing within given time period varies with the interested model input changes. The measured effects of each input ($\boldsymbol{X}$= H_s, KG/KMT, $Heel$, L_d/L_s) on the model output are depicted by a boxplot as shown in Figure 37. It can be seen that the "box" for each variable indicates the output $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ ranges between 25% and 75% quantile. The whiskers (e.g. T-shaped lines) represent the highest and lowest datum within 1.5 × IQR (IQR=75% – 25% quantile) of the upper and lower quantile defined for the "box". Focusing on the variation in the model output which is affected by the changes of KG/KMT at a specific operational draught, the points beyond the extreme of the whiskers indicate the smallest observation away from the lower whisker. A wider box and whisker signifies a greater uncertainty in the model output given the variation in a particular input. According to such features, with respect to Figure 37, it is obvious that the changes in the extent of flooding is ranked to have the greatest influence on ship survivability, followed by the changes in the angle of heel, the centre of gravity and the sea environment.

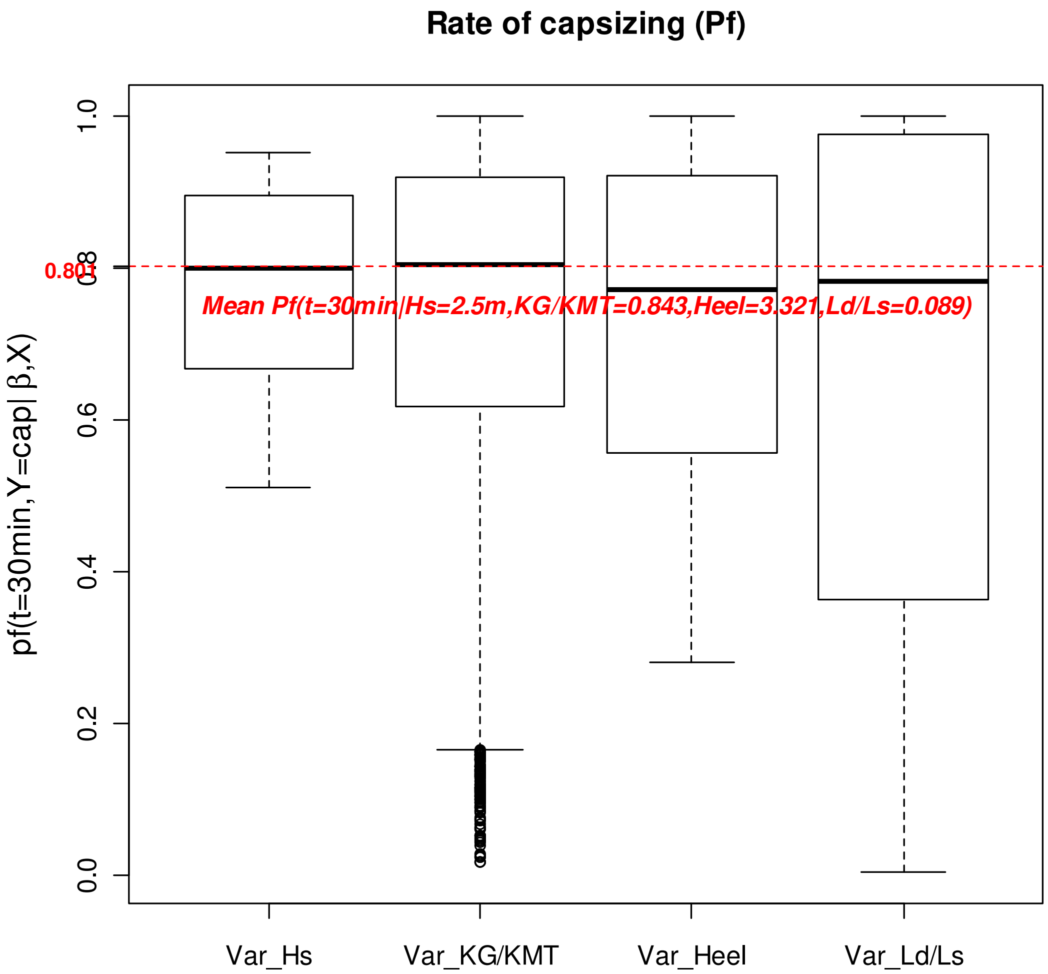


Figure 37 *Boxplot of the results of the model output* $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ *vary with each model input (*$\boldsymbol{X}$*=* H_s*,* KG/KMT*, Heel,* L_d/L_s*) changes*

The significance of the various inputs to the model response depicted above (Figure 37) agrees with the findings outlined in the executive summary of Deliverable 4.2 where it highlights the extent of flooding, affecting the GZ-curve properties, seems to be one of the most critical information needed for confident assessment of criticality of flooding situation. Thus the lack of knowledge in determining the extent of flooding experienced during crises may lead to the greatest uncertainty in the survivability prediction. Table 16 summaries the key values represented in the boxplot. The mean value of the rate of capsizing pointed out as $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X}) = 0.801$ is measured given the condition of using sample means of all inputs variables.

Table 16 *Summary of the variation (uncertainty) in the model output*

*Var_*$P_f(t_0,Y=cap|\boldsymbol{\beta},\boldsymbol{X})$		*Var_Hs*	*Var_KG/KMT*	*Var_Heel*	*Var_Ld/Ls*
P_f	Min.	0.51160	0.01710	0.28080	0.00442
	1st Qu.	0.66780	0.61720	0.55560	0.36409
	Median	0.79990	0.80460	0.77160	0.78271
	Mean	0.77660	0.74544	0.72740	0.65781
	3rd Qu.	0.89560	0.91842	0.92060	0.97496
	Max.	0.95210	0.99997	1.00000	1.00000

According to Figure 37, the percentage change of $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ (ΔP_f) as revealed in Figure 38 is a more intuitive way to express a change in the model output compared to its measured mean value (P_f=0.801). At this time, the percentiles 0.5%, 99.5% are used to show the range of 99% uncertainty in the ΔP_f contributed by the different variations in the inputs of Model 2. The corresponding key values are summarized in Table 17. By doing so, this boxplot can be regarded as a quantitative means of measuring the significance of several input variables to the model output.

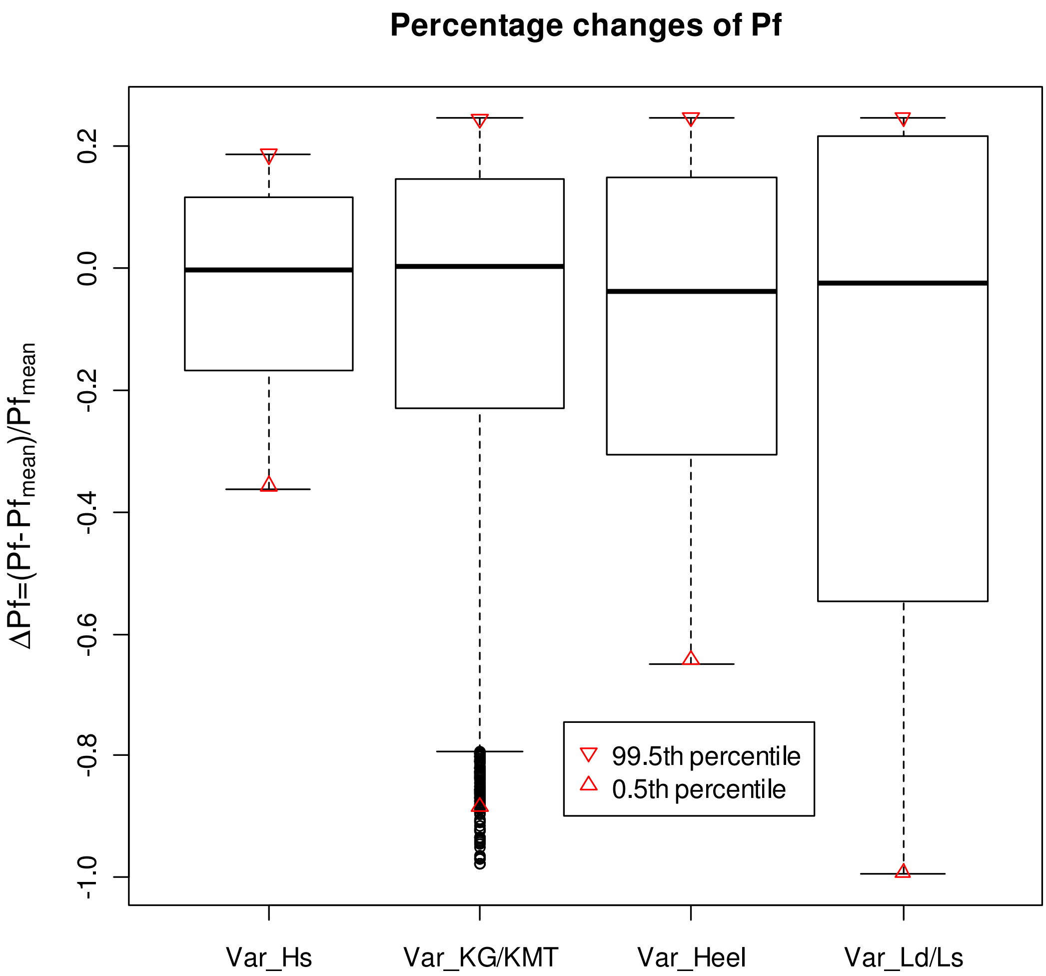


Figure 38 *Sensitivity analyses of the input variables on percentage changes in* $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$

Table 17 *Summary of the variation (uncertainty) in the percentage change of the model output (ΔPf)*

ΔPf	*Var_Hs*	*Var_KG/KMT*	*Var_Heel*	*Var_Ld/Ls*
Mean	-0.03059	-0.06944	-0.09200	-0.17883
SD	0.16313	0.26528	0.26684	0.42505
99.50%	0.18781	0.24558	0.24817	0.24835
50%	-0.00141	0.00442	-0.03679	-0.02291
0.50%	-0.35683	-0.86338	-0.64260	-0.99180

6.2.2 A Scenario Level Investigation

The prior global sensitivity study suggests that the variable L_d/L_s plays the most significant role in estimating the rate of ship stability loss during crises. However, considering a real hull breach which is normally characterised by a set of parameters $\Omega = \{x,\lambda,b,h\}$ of its location, length, penetration and height. Apparently, the damage length $\lambda = L_d$ is insufficient to disclose the impact of the extent of flooding on the ship stability after damage. In this way, a new predictive model (relationship) is desired to be established that allows examining the sensitivity of the model output $P(t_0, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$ to a group of rearranged input parameters as given in Table 18. Noticeably, a comprehensive description of the flooding extent including four variables is available.

Table 18 *Input variables for a local sensitivity study*

Category	Variable	
Sea environment	H_s	
Damage attributes	Heel in EQ	$Heel$
	Location	x_i/L_s
	Length	L_d/L_s
	Penetration	$y/0.5B$
	Height	z/H

As described in section 6.1, the most straightforward aspect of uncertainty analysis is uncertainty propagation. Abstractly, this process begins with a mathematical model of the measurement $Y = f(\boldsymbol{X})$, where $\boldsymbol{X}$ is a vector of the input parameters. Uncertainty analysis aims to assess the uncertainty in Y that is driven by the uncertainties in $\boldsymbol{X}$. The uncertainty in $\boldsymbol{X}$ is usually specified as probability density function. The uncertainty in Y is then calculated by propagating the uncertainty in $\boldsymbol{X}$ through the model $Y = f(\boldsymbol{X})$. Being aware of this, a binary regression model is written as $P(Y = cap|\boldsymbol{\beta}, \boldsymbol{X}) = \Phi(\alpha + \boldsymbol{\beta X})$ has been identified in this study. It looks straightforward to solve this uncertainty propagation formula. However, the biggest difficulty is in the calculation of model coefficients $\boldsymbol{\beta}$ of input variables (also called sensitivity coefficients). Thus, estimation of model coefficients is deemed as the crucial issue in this localized sensitivity study. The entire process will go along with the regression model development consisting of data preparation, model estimation and model validation.

First, for achieving the exhaustive information of the extent of flooding as a result of damage, in contrast to the elicitation of Model 2, the data collected presently for a new model training come from the numerical simulation as reported in Table 6 and Table 7 of Deliverable 4.2. So the dataset of sensitivity study now consists of 1) 200 runs of repetitive simulation for a range of sea states that have a fixed damage opening (i.e. P6-7.1.0) including transient stages of flooding. 2) 1100 runs of Monte Carlo-based simulation for a range of sea states that allow for random changes in the damage extent. Since the initial loading of the typical RoPax vessel “MV Estonia” defined in simulations complies with the condition performed in physical experiments, therefore the variation of the KG/KMT is not to be considered any more. The factors characterizing a hull breach are defined non-dimensional. The damage location x_i denotes the length from the stern to the midpoint of the damage opening in the longitudinal direction. The damage penetration y is measured from the centreline of the ship to the side where suffered damage. The damage height z determines the vertical distance from the base line of the ship to the top of the damage, and the parameter H indicates the height up to the top of the car deck. In this study the hull is assumed water tight up to Deck 4 i.e. H = 13.40m above the base line. All the mentioned terms are clearly remarked in Figure 39.

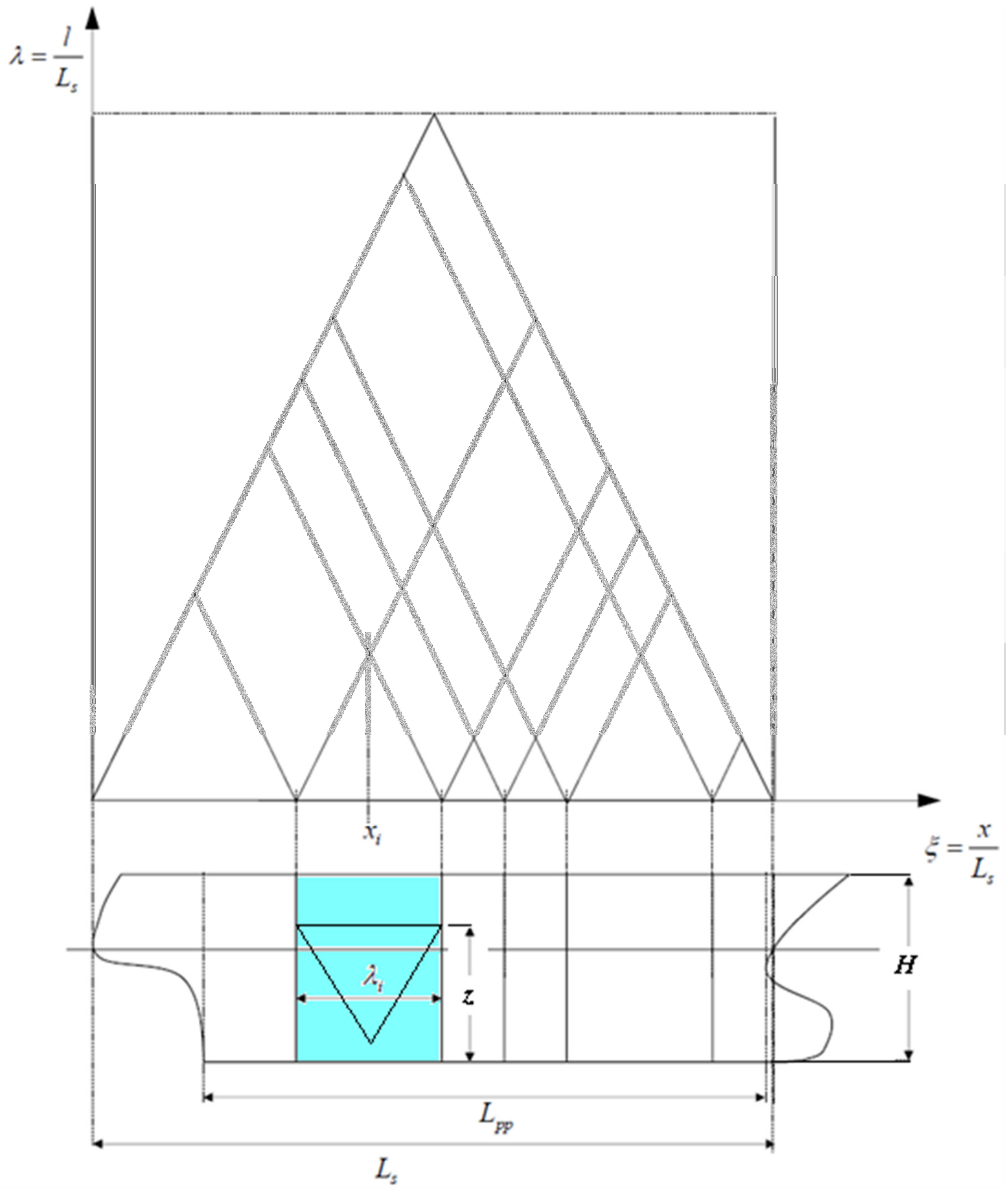


Figure 39 *The extent of flooding as a result of damage*

Second, depending on a set of input variables clarified above, the regression model used for the local sensitivity study is reset as:

$$P(t_0, Y = cap \,|\, \boldsymbol{\beta}, \boldsymbol{X}) = \Phi(\alpha + \beta_1 H_s + \beta_2 \frac{x_i}{L_s} + \beta_3 \frac{L_d}{L_s} + \beta_4 \frac{y}{0.5B} + \beta_5 \frac{z}{H} + \beta_6 Heel\,) \qquad (27)$$

As described in Chapter 4, the estimation of sensitivity coefficients $\boldsymbol{\beta}$ in equation (27) is based on Bayes' theorem. It is reasonable to suppose that prior distributions of these coefficients are approximately Normal. Synthesizing disparate information from numerical simulations and the prior knowledge, the target posterior distributions of $\boldsymbol{\beta}$ corresponding to each input variable can be approximated by MCMC method (i.e. Metropolis-Hastings algorithm). When 10,000 iterative samplings are performed, uncertainties in $\boldsymbol{\beta}$ are specified by probability distributions as showed in Figure 40. A parallel summary of simulated results are outlined in Table 19.

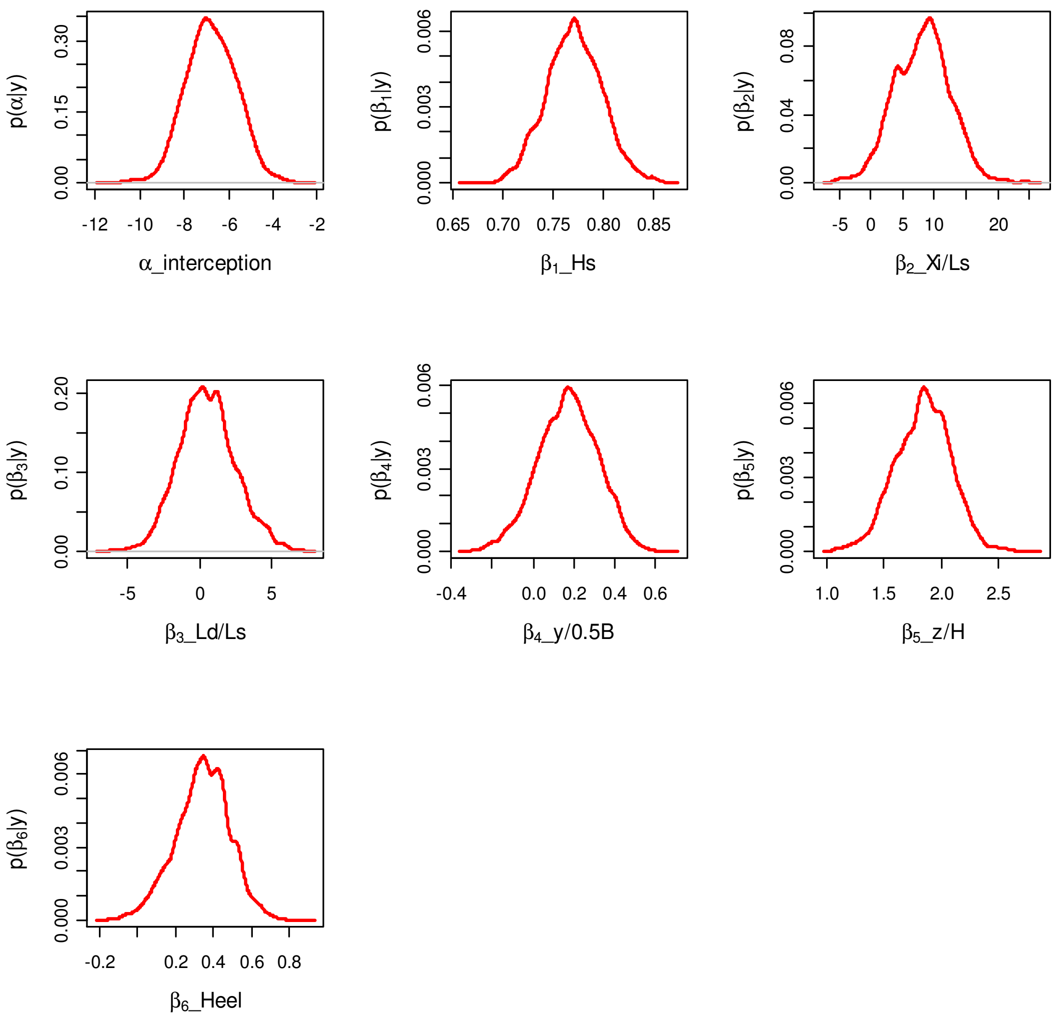


Figure 40 *MCMC approximations to posterior distributions of* $\boldsymbol{\beta}$ *(Sensitivity coefficients)*

Table 19 *Summary of Metropolis-Hastings samples of* $\boldsymbol{\beta}$ *(sensitivity coefficients)*

Parameter	Mean	SD	99% posterior interval
α (interception)	-6.771	1.1048	-10.4096, -3.5973
β_1 for H_s	0.771	0.0272	0.698, 0.852
β_2 for x_i/L_s	8.0716	4.3566	-5.2479, 23.0596
β_3 for L_d/L_s	0.4958	1.9635	-5.7976, 6.7729
β_4 for $y/0.5B$	0.1709	0.1445	-0.2682,0.5681
β_5 for z/H	1.8439	0.2407	1.1004, 2.5931
β_6 for *Heel*	0.3428	0.1411	-0.1021, 0.7517

An intuitive analysis of the sensitivity of the rate of capsizing within given time $P(t_0, Y = cap \,|\, \boldsymbol{\beta}, \boldsymbol{X})$ with reference to several input parameters is achieved regarding the observation of posterior probability distributions depicted in Figure 40. As can be seen that the uncertainty (within 99% quantile-based posterior interval) in the coefficient β_3 contains zero, it indicates the model output is not sensitive to its related parameterL_d/L_s. Such finding may reasonably be explained by the source of data used in the calculation of sensitivity coefficients. Since the prepared dataset represents the results of numerical simulation, while the ship is assumed to be exposed to a set of hull breaches leading to the flooding extent same as P6-7.1.0. Figure 41 demonstrates 100 random damage extents based on MC sampling which are investigated for a range of sea states. Clearly, the flooding domain which the sampled damage openings result in is no more than two damage zones. Under such circumstance, the damage length L_d seems insignificant in the model at this time. Likewise the coefficient β_4 $(y/0.5B)$ corresponding to the damage penetration is the second factor having minor effect.

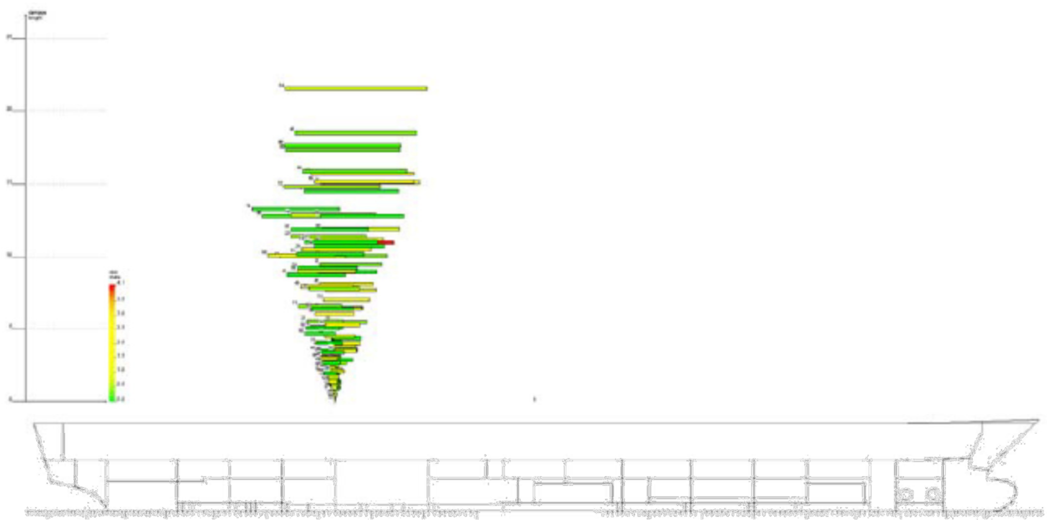


Figure 41 *A set of 100 hull breaches leading to flooding extent P6-7.1.0 on MV Estonia. The bars represent location and length of hull breach.*

Considering the sensitivity of all the unknown coefficients $\boldsymbol{\beta}$, the next step is to simplify the model as equation (27) through excluding irrelevant input parameters to strive for model parsimony. The first attempt is made to create a new relationship without the parameter as the damage length L_d/L_s. The revised range of variation in other coefficients $\boldsymbol{\beta}$ is outlined in Figure 41. Apparently, the impact of eliminating such parameter on the model output is unimportant since the changes in the rest effects $\boldsymbol{\beta}$ are not obvious. It turns out again that the damage penetration β_3 $(y/0.5B)$ may not be considered in the model.

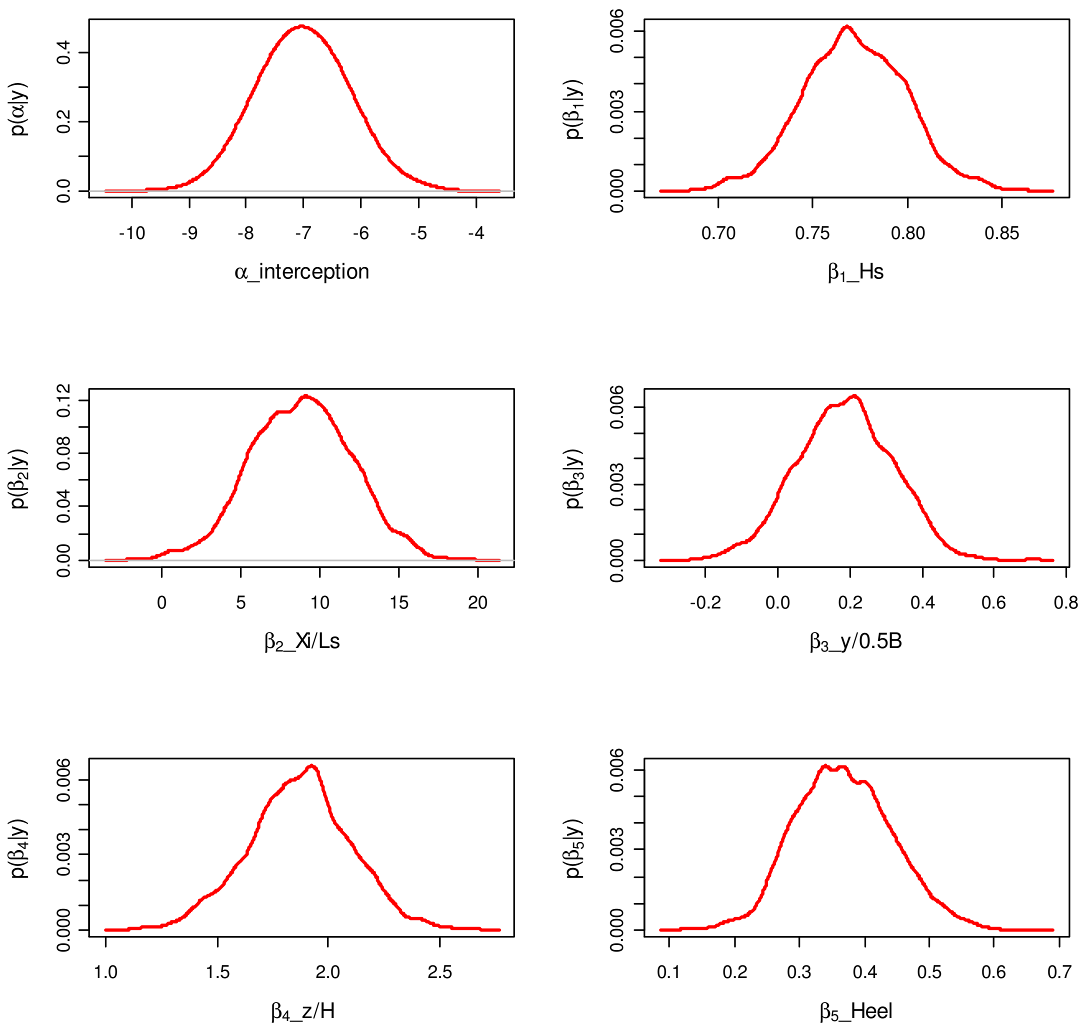


Figure 41 *MCMC approximations to posterior distributions of* $\boldsymbol{\beta}$ *(Exclude the parameter* L_d/L_s *)*

Accordingly, the second attempt is to rewrite the model including four input variables as show by equation (28).

$$P(t_0, Y = cap \,|\, \boldsymbol{\beta}, \boldsymbol{X}) = \Phi(\alpha + \beta_1 H_s + \beta_2 \frac{x_i}{L_s} + \beta_3 \frac{z}{H} + \beta_4 Heel\,) \qquad (28)$$

By the same token, uncertainties associated with remaining input parameters are measured regardless of the damage length and the damage penetration. See Figure 42, the posterior densities $P(\boldsymbol{\beta}|y)$ interpret all these assessed variables are dominant, as the 99% quantile-based posterior intervals for all the coefficients do not contain zero. The positive magnitudes of $\boldsymbol{\beta}$ simulated indicate that the increase in the value of each input parameter will likely increase the rate of capsizing within given time after flooding. On the basis of Table 20, a new regression model used for uncertainty propagation is estimated.

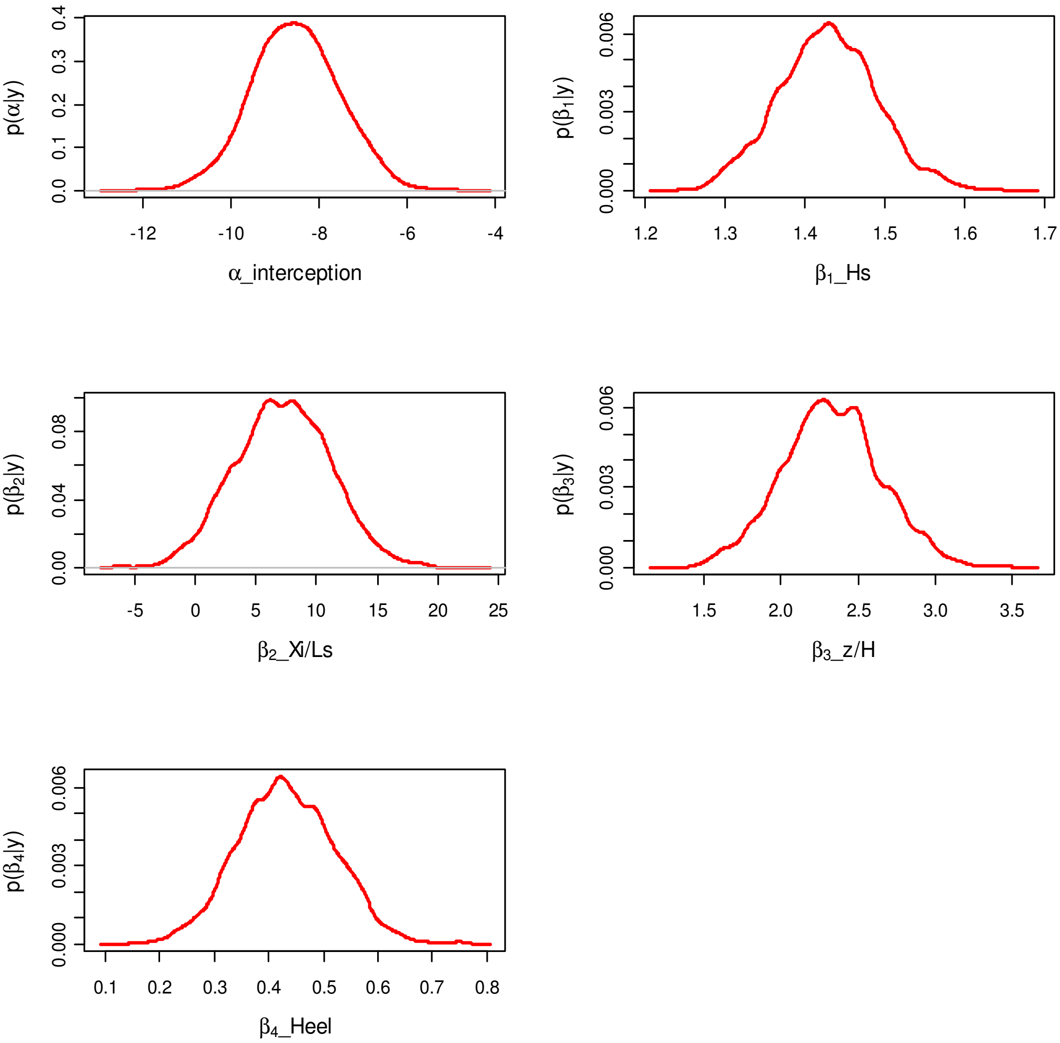


Figure 42 *MCMC approximations to posterior distributions of* $\boldsymbol{\beta}$ *(Exclude parameters* L_d/L_s *,* $y/0.5B$*)*

Table 20 *Summary of Metropolis-Hastings samples of* $\boldsymbol{\beta}$ *(regarding Figure 42)*

Parameter	Mean	SD	99% posterior interval
α (interception)	-8.5323	0.9733	-11.7929, -5.3017
β_1 for H_s	1.4283	0.0614	1.258,1.626
β_2 for x_i/L_s	7.231	3.9117	-4.4368,18.8246
β_3 for z/H	2.3296	0.3189	1.4807, 3.3613
β_4 for *Heel*	0.4322	0.0889	0.1715, 0.7505

Third, a further validation study on the developed model is undertaken. As reported in deliverables of Task 4.1 and Task 4.2, according to physical model tests and numerical simulations, the measured rate of capsizing for a range of sea states is presented in Figure 43, when the ship is exposed to a specific flooding case P6-7.1.0. The derived probability to capsize $p_f(t = 30min, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$

followed by the aforementioned sensitivity analysis is also included. Because the proposed regression model is straightforward for uncertainty propagation, 99% uncertainty bounds of the rate of ship stability loss within 30 minutes are illustrated simultaneously. It is noticeable that the results achieved from different methods are comparable. The most distinguished feature of the established regression model is to quantify the uncertainty in assessment of ship survivability explicitly.

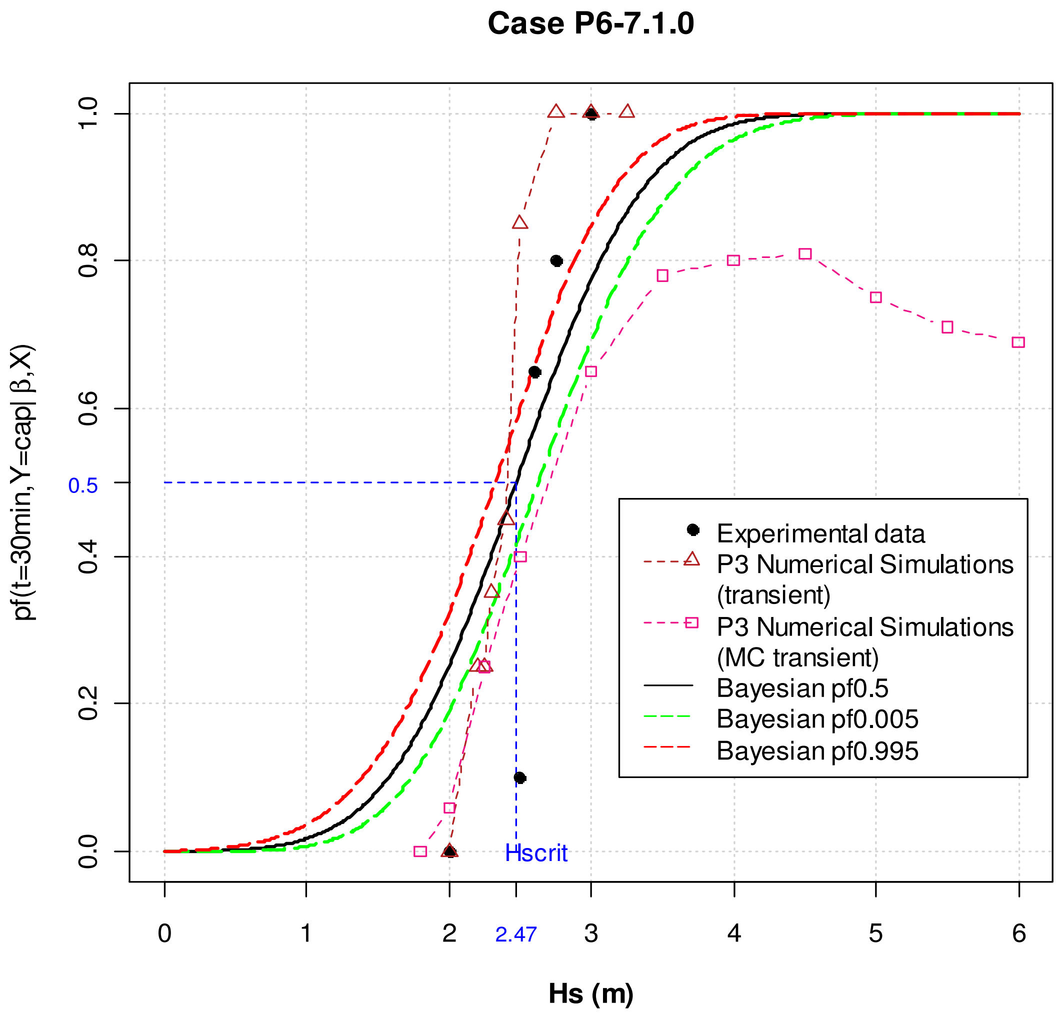


Figure 43 *M/V Estonia, posterior distribution of the rate of capsizing in a flooding case P6-7.1.0*

$$p_f(t = 30min, Y = cap|\boldsymbol{\beta}, \boldsymbol{X} = H_s, x_i/L_s\,, z/H\,, Heel)$$

This section conducts extensive sensitivity studies on the measurement of contributions of input variables to the model outcome. The importance of each predictor variable in a model can be evaluated if its related sensitivity coefficient is known. Subsequently, model simplification should be realized without consideration of irrelevant inputs. Since variations in all the coefficients are specified by probability density functions, uncertainties in the inputs $\boldsymbol{X}$ are easily propagated to the uncertainty in the output $p_f(t = 30min, Y = cap|\boldsymbol{\beta}, \boldsymbol{X})$. In this way, two aspects pertaining to this proposed methodology in assessment of ship survivability are disclosed as which allowing for systematic sensitivity analysis and uncertainty propagation. Further work is still needed on the development of predictive models based on expanding the data generated from numerical simulations. Greater changes with input parameters in a model are expected for performing a thorough sensitivity study at a ship level. Eventually, a simplified regression model is adequate to represent the more complex computer model. Therefore, a fast and accurate assessment of ship survivability can be completed for design optimization or decision support in emergency situations.

7. CONCLUSION

Attempting to get a fast and rational survivability assessment for further deployment in decision support in emergencies, this report has focused on establishing the predictive regression models for the survivability assessment of the damaged RoPax vessels within given time, concurrently the uncertainty bounds on time to capsize models have been assigned in a probabilistic manner. In short, a pragmatic approach to assign quantitatively the uncertainty impact has been presented.

A systematic description of the development of the probabilistic predictive model has been conducted through model identification, model estimation and model application as clarified in Chapter 3, 4 and 5, respectively. Extensive uncertainty analysis and sensitivity studies have been performed in Chapter 6 for better understanding of the contributions of individual inputs in the model to the uncertainty in the output. It has been mentioned that the uncertainty analysis in the process of the ship survivability assessment is important. With respect to the first principle models, the inherent uncertainties in modeling are commonly deriving from both the parameter uncertainty and the model uncertainty. The treatment of such two groups of uncertainties is significant to reduce the uncertainty caused by vague input information and imperfect models.

For better describing the physical phenomenon of ship capsizing, the performance-based experimental observations and the numerical simulation results are deemed as one of the most reliable sources of information for the predictive model estimation. In this context, the applied data is related to RoPax vessel only rather than other ship types. Nevertheless, identical methodology can be adopted for allowing the damage stability assessment of cruise ships. As discussed in D4.2 and the previous Chapter 6, the uncertainty due to a lack of precise input information, such as the extent of flooding, or sea state is far more critical to the projection on what will happen, than the uncertainties underlying the established model itself for either RoPax or cruise ships. However, a series of experiments for RoPax and cruise ships was conducted in the project GOALDS, the available data is to be studied separately based on different ship types for mitigating the model uncertainty. More effort should be devoted to the data preparation on cruise ships for further explicit uncertainty quantification in the assessment of ship survivability.

8. REFERENCES

ABRAHAMSSON, M. 2002. *Uncertainty in quantitative risk analysis : characterisation and methods of treatment,* Lund, Sweden, Dept. of Fire Safety Engineering, Lund University.

AGRESTI, A. 2007. *An introduction to categorical data analysis,* Hoboken, N.J. ; Chichester, Wiley-Interscience.

BARTLETT, M. S. 1978. *An introduction to stochastic processes with special reference to methods and applications,* Cambridge; London; New York, Cambridge University Press.

CHARNES, A., FROME, E. L. & YU, P. L. 1976. The Equivalence of Generalized Least Squares and Maximum Likelihood Estimates in the Exponential Family. *Journal of the American Statistical Association,* 71**,** 169-171.

CHEN, Q., TSAKALAKIS, N. & JASIONOWSKI, A. 2009. Physical Model Experiments *SAFEDOR, D5.6.3.*

EC 2003. DIRECTIVE 2003/25/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on specific stability requirements for ro-ro passenger ships.

GELMAN, A. 2004. *Bayesian data analysis,* Boca Raton, Fla., Chapman & Hall/CRC.

GL. 2002. *SAFEDOR - design, operation and regulation for safety, Integrated Project 516278 in 6th Framework programme of the European Commission* [Online]. Germanischer Lloyd. Available: http://www.safedor.org/index.htm [Accessed 14 April 2010].

HACKING, I. 1975. *The emergence of probability : a philosophical study of early ideas about probability, induction and statistical inference,* Cambridge; New York, Cambridge University Press.

HAIR, J. F. 2009. *Multivariate data analysis,* Upper Saddle River, N.J., Pearson Education International ; [Harlow] : Pearson Education Ltd.

HARDER 2003. HARDER - Harmonisation of Rules and Design Rationale, U Contact No. GDRB-CT-1998-00028, Final Technical Report.

HEIMVIK, N. 2001. Wave Height Distribution According to Damage Statistics. *HARDER No: GRD-1999-10721,2011-11-29,Report No: 3-33-W-2001-01-0.*

HOFF, P. D. 2009. *A first course in Bayesian statistical methods,* New York ; London, Springer.

JASIONOWSKI, A. 2006. Fast and accurate flooding prediction - analytical model. *SAFEDOR, D2.1.3.* .

LASKEY, K. 1994. Model confidence (Lecture notes) Summer Institue on Probability in AI, George Mason University.

LEE, P. M. 2004. *Bayesian Statistics : An Introduction,* London, Hodder Arnold.

LUDWIG, N. 2009a. Damage Stability Tests with the Model of an 80 m RoPax Vessel. *HSVA Report No. S590a/09, Hamburg. Hamburgische Schiffbau-Versuchsanstalt GmbH (HSVA), Hamburg.*

LUDWIG, N. 2009b. Damage Stability Tests with the Model of a 200 m RoPax Vessel. *HSVA Report No. S590b/09, Hamburg. Hamburgische Schiffbau-Versuchsanstalt GmbH (HSVA), Hamburg.*

RASK, I. 2010. Benchmark data on time to capsize for a free drifting model. *FLOODSTAND, D4.1a, SSPA Report No.: 4007 4581-1 rev1.*

TUZCU, C. & RUSAAS, S. 2003. Recommended new harmonized probabilistic damage stability regulations - factor s. *HARDER No: GRD1-1999-10721, 2003-05-31,Report No:7-72-D-2003-03-0.*

TUZCU, C. & TAGG, R. 2001. Model Tests Results - PRR01. *HARDER No: GRD1-1999-10721, 2011-06-15, Report No: 3-31-D-2001-01-0.*

YUEN, K.-V. 2010. *Bayesian methods for structural dynamics and civil engineering,* Oxford, Wiley-Blackwell.