

# Reducing Uncertainty in Subdivision Optimization

Romanas Puisa, Nikolaos Tsakalakis and Dracos Vassalos

*Ship Stability Research Centre, University of Strathclyde, Glasgow G20 0TL, UK*

**Abstract:** Design of watertight subdivision inherently involves its optimization with the objective to increase the index “A” above its minimum required value. In view of a big popularity of probabilistic search methods such as genetic algorithms, this task is intrinsically time consuming. Thus, even when an optimal subdivision layout (i.e. topology) is determined, it can be found that the optimal bulkhead positions can be a great challenge time-wise, often forcing designers to satisfy with suboptimal solutions. The fundamental reason why this happens is that the nature of the optimized function (e.g., index “A” as a function of bulkhead positions) is unknown and hence it has no effect upon the choice of optimization strategy, which therefore reflects subjective but not factual preferences. In this paper we study the nature of functional dependency between the subdivision index and bulkhead positions, as a simplest case, and indicate pertinent optimization strategies that consequently reduce the optimization time. In our study we use a cruise ship model to demonstrate the application results of our findings.

**Key words:** Damage stability, optimization, watertight subdivision, index “A”, logistic regression, cruise ship, approximation, surrogate function.

## 1. Introduction

Ship stability in a damaged condition is one of the safety critical functions that a designed ship has to provide. Ship subdivision into watertight compartments is a traditional approach to secure a needed level of damage survivability. The attained probabilistic subdivision index “A” of the international convention for the Safety of Life at Sea (SOLAS) 2009 reflects the level of damage survivability and hence its calculation has been a routine task for naval architects. Naturally, design of watertight subdivision involves optimization of the index “A” with the objective to increase it above the minimum required value (denoted as  $R$  [1]) and keep maximising it further as long as it is cost effective. What is cost effective is rather subjective and conditional upon resources and time available. Apparently the both are limited and a quick delivery of the sufficient ship subdivision, i.e.  $A > R$ , is of primary interest. This is surely possible when past subdivision designs are re-used, however even in this case modifications are inevitable such as those brought

by changes in hull dimensions, the number and sizes of tanks, the engine room size, and so on. As a result, the existing level of the damage survivability is likely to be affected and the subdivision optimization towards  $A > R$  becomes necessary. Formally, we face with the following optimization problem, dropping other design objective functions (e.g., life cycle profitability, environmental impact) for simplicity.

$$A(x) \rightarrow \max$$

subject to

$$\begin{aligned} A(x) &> R, \\ h(x) &= 0, \\ g(x) &\leq 0 \end{aligned} \quad (1)$$

where  $A(x)$  is the subdivision index “A” as a function of design variable vector  $x$  (bulkhead positions, number of bulkheads, number of tanks etc.),  $h(x)$  and  $g(x)$  are vectors of equality and inequality constraints that can be additionally imposed. This optimization problem has to be always solved, regardless the presence of various methods that determine a close-to-optimal transversal subdivision. Thus for example, Karaszewski and Pawłowski [2] concluded that the optimal position of the transverse bulkheads is defined by a uniform distribution of the local indices

---

**Corresponding author:** Romanas Puisa, research fellow, Ph.D., research fields: ship design, design optimization and risk analysis. E-mail: r.puisa@strath.ac.uk.

of subdivision. The relatively recent research study by Cabaj [3] allows linking flooding survivability of SOLAS 2009 with positions and lengths of compartments similar like the classical method of floodable length curves.

To this end, we need to select an optimization method to efficiently solve Eq. (1). Apparently, a choice of the most relevant method has to be driven by the mathematical structure of the problem. In particular, we are interested in the topology (structure of the function landscape) of the optimized function. Thus, for regular (continuous and differentiable) and preferably convex functions of design variables, gradient-based methods is the right choice. These deterministic methods guarantee optimal solutions in a few iterations. Their efficiency stems from the fact that they utilise knowledge about the function topology explicitly. In particular, a direction to the optimum and the step size are determined by respectively calculating the gradient vector and the Hessian matrix (related to the local curvature) of the optimized function. In case of discrete (e.g. due to noise) or/and discontinuous (e.g., there is no solution for certain design variable values) and hence not differentiable functions, which can be also multimodal (multiple local optima), probabilistic methods are preferable. The probabilistic methods (e.g. evolutionary algorithms) are intrinsically robust within a broad range of optimization problems, but the cost for that is greatly lesser efficiency. They are less efficient because they deliberately neglect the function topology, ranking alternative solutions based on function values only. Hence, the search for optima is blind, whereas its success is left to chance. As a result, probabilistic methods guarantee an improvement only with the probability. Apparently, this probability is bigger when the problem is simpler and the time allowed for optimization is more. Thus, this probability theoretically approaches one (for any problem) for the infinite search time.

With the above in mind, we need to be certain of whether the optimization method we apply is really

relevant. Surely, the preference falls on some deterministic approach, but its applicability has to be still verified. This can be done through examination of the topology of optimized functions, classifying the functions into convex or non-convex and identifying other properties such as the number of function extrema. This kind of analysis is not generally complex if explicit mathematical formulations of those functions are known and can be written in a closed form. However, in practice this knowledge is often limited. In such situations we can describe a functional relationship only implicitly and more often in the form of a “black box” with only known input and output. This has conventionally been the case with the index “A”. Its general form as a sum of its components is given as follows.

$$A(x) = \sum_{i=1}^{ND} \underbrace{p_i(x_{p_i})s_i(x_{s_i})}_{\text{component } A_i} \quad (2)$$

where  $ND$  is the total number of damage cases,  $p_i(x_{p_i})$  is the probability of flooding of a given compartment or group of compartments  $i$ ,  $x_{p_i}$  is a set of design variables ( $x_{p_i} \subseteq x$ ) that affect  $p_i$ ,  $s_i(x_{s_i})$  is the conditional probability of surviving flooding a given compartment or group of compartments  $i$ ,  $x_{s_i}$  is a set of design variables ( $x_{s_i} \subseteq x$ ) that affect  $s_i$ . Note, equality  $\sum_i p_i = 1$  must hold. Thus, bearing in mind Eqs. (1) and (2), maximization of  $A(x)$  implies maximization of components  $A_i$ .

We are fortunate enough to know that the subdivision index “A” is an aggregative function of linearly combined sub-functions  $A_i$ . Hence, the topology of the aggregative function can be determined via topology analysis of its sub-functions. The knowledge about this will enable us to estimate the complexity of optimization problem in Eq. (1) and select a relevant optimization strategy, as a result. This frames the actual paper, which is organized as follows. The following section addresses the topology analysis of components  $A_i$ , the sequential sections address the problem complexity, the choice of the most relevant optimization strategy and its application to subdivision

optimization. The final section concludes the paper.

## 2. Topology Analysis of Index “A” Function

An explicit form of  $p_i(x_{p_i})$  is known and found in Refs. [1, 4]. The so-called  $p$ -factor,  $p_i$ , is calculated on the basis of probability density distributions for relative damage location, penetration and length, given (a) longitudinal positions of transversal watertight bulkheads that define flooded compartment(s) and (b) the minimum breadth of flooding wing compartment(s). Fig. 1 shows how  $p$ -factor varies with positions of transversal bulkheads confining a damaged compartment. Apparently, the relationship is convex and the maximum  $p$ -factor value is when the bulkheads are farthest apart.

The ranges for  $x_1$  and  $x_2$  in Fig. 1 are taken from a simplified cruise ship subdivision with 22 transversal bulkheads shown in Fig. 2. We assume that the function topologies will not essentially be different in more complex subdivisions (e.g., with numerous longitudinal and stepped transfers bulkheads) and also ship types.

Note, the  $p$ -factor function for multiple compartment damages will have the same topological characteristics as for single compartment damage, but it will be additionally affected by intermediate (between  $x_1$  and  $x_2$ ) bulkhead positions as well [4]. The  $p$ -factor function for multiple compartment damages is of high dimensional and hence cannot be plotted for visual analysis. To this end, the  $p$ -factor function is seemingly convex, continuous and differentiable.

The missing link in understanding the topology of Eq. (2) is the  $s$ -factor,  $s_i$ , which is the second element of component  $A_i$ . Even though its explicit form is found in the regulations [1], that is

$$s_i = K \left( \frac{GZ_{max}}{0.12} \frac{Range}{16} \right)^{\frac{1}{4}}. \quad (3)$$

In Eq. (3),  $K$  is a constant,  $GZ_{max}$  is a maximum value of the positive righting lever,  $Range$  is the range of positive righting levers, both determined through hydrostatic stability calculations.

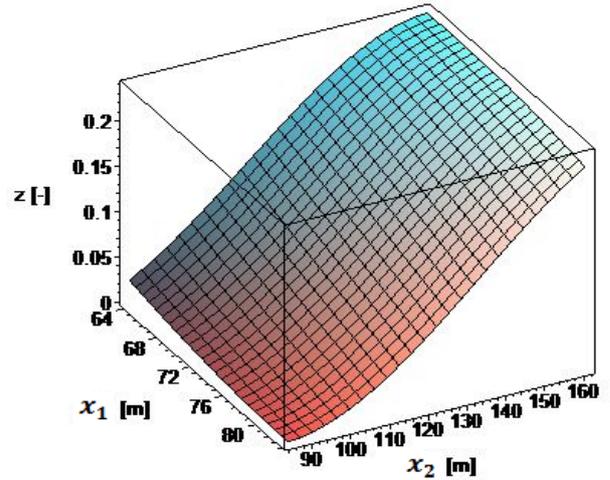


Fig. 1 The  $p$ -factor for single compartment, whose both ends are inside the ship length. Variables  $x_1$  and  $x_2$  are measured in meters from the aft terminal of the ship, hence defining the length and longitudinal position of the compartment.

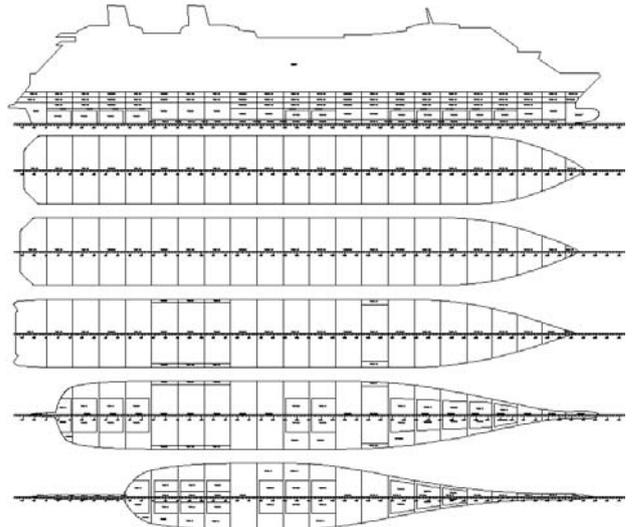


Fig. 2 A reference model of a cruise ship<sup>1</sup> with a simplified subdivision that is usually used in primary damage stability calculations. The required subdivision index is  $R = 0.83$ .

The  $s$ -factor relation to design variables such as bulkheads positions  $x_1$  and  $x_2$  is not necessary obvious. From the physical point of view, the  $s$ -factor must be a function of the size and the location of a flooded compartment (or a group of adjacent compartments). That is, the size of a flooded

<sup>1</sup> Main particulars of the ship: length between particulars  $L_{bp} = 293$  m, length overall  $L_{oa} = 323$  m, breadth  $B = 36.8$  m, design draught  $T = 8.3$  m.

compartment determines the volume of water that reduces the displacement of the damaged ship. In turn, the position of the flooded compartment tends to govern a combined effect of the heel and the trim on the ship's stability. Thus, rather than using the coordinates of transverse watertight bulkheads  $x_1$  and  $x_2$ , we use the size and the position of a flooded compartment they define. That is,

$$\lambda = (x_2 - x_1), \chi = x_1 + \frac{\lambda}{2} \quad (4)$$

where  $\lambda$  is the size and  $\chi$  is the position of the damage compartment or the group of adjacent compartments being flooded. Thus, Fig. 3a shows the contour plots of the  $s$ -factor function across different damage cases that involve one, two, three, four and five adjacent compartments (or zones) filled with water. The contour plots indicate, and this is also confirmed in Fig. 3b, that the  $s$ -factor function follows a sigmoidal shape, which is a continuous surface with large plateau of 1/0 values. From the physical point of view, the plateaux define areas of stability ( $s_i = 1$ ) or instability ( $s_i = 0$ ) for the damaged ship in term of the position and the size of the damage.

To this end, behavior of the  $s$ -factor is well defined. It ranges from 0 to 1 and resembles a Bernoulli trial with two dominating outcomes: survived  $s_i = 1$  (success) and not survived  $s_i = 0$  (failure). In statistics, this particular behavior of the dependent variable is well captured via the logistic regression (LR)<sup>2</sup> [5], which has the following model.

$$y(x) = \frac{1}{1 + \exp(-\beta(x))} \quad (5)$$

where  $\beta(x)$  is usually a linear model with the intercept defined as

$$\beta(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (6)$$

with  $\beta_i$  being regression coefficients. In LR  $y(x) \in [0,1]$  represents the probability of a particular outcome, given a set of explanatory variables  $x$ . This is exactly

what the  $s$ -factor represents: the probability of surviving a damage case described by design variables  $\lambda$  and  $\chi$  and  $b$ ; the latter being the minimum breadth of flooded wing compartment(s) when corresponding longitudinal bulkheads are present. The relevancy of LR is also backed by the fact that logistic function  $y(x)$  in Eq. (5) is also sigmoidal. Thus, Eq. (5) can be rewritten as a regression model for the  $s$ -factor as

$$\tilde{s}(\boldsymbol{\beta}, \lambda, \chi, b) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 \lambda - \beta_2 \chi - \beta_3 b)} \quad (7)$$

where  $\boldsymbol{\beta}$  is the vector of regression coefficients  $\beta_i$ .

Figs. 3b and 3c also shows application of LR on sample data for the  $s$ -factor. Specifically and interestingly, the approximation is very accurate for one and two compartment (zone) damages, and reasonably accurate for three, four and five compartment damages. Thus, the closer sample points to the sigmoidal shape, the more accurate regression results are. It's important to note that regardless of the level of inconsistency between sample/calculated and regression points in Fig. 3, the expected value of the  $s$ -factor,  $E(s)$ , in both data sets is the same. That is, for each damage case model Eq. (8) holds when the number of samples is sufficiently large.

$$\underbrace{\frac{1}{N} \sum_{i=1}^N \tilde{s}(\boldsymbol{\beta}, \lambda, \chi, b)}_{E(s) \text{ of regression data}} = \underbrace{\frac{1}{N} \sum_{i=1}^N s_i}_{E(s) \text{ of sample data}} \quad (8)$$

where  $N$  is the number of samples. Eq. (8) also holds for  $N = 57$ , which is the number of samples used in the plots of Fig. 3. Hence, the use of LR for approximating the  $s$ -factor can also be justified this way.

To this end, the analysis of the  $p$ -factor has shown the function behind it to be continuous, differentiable and likely convex. In turn, the analysis of the  $s$ -factor function confirms its strong sensitivity to the size and position of the damaged compartment or a group of adjacent compartments involved in the damage case. As for the  $s$ -factor function topology, it signifies a continuous sigmoidal surface in 3D of which shape changes with the number of adjacent compartments

<sup>2</sup> Note, one of the properties of LR is that it can accommodate heterogeneous variables in the model (6). That is, some variables can be continuous, other discrete or/and categorical. Thus for example, the compartment size and position can be given in frames (discrete), whereas  $b$  can be given in meters (continuous).

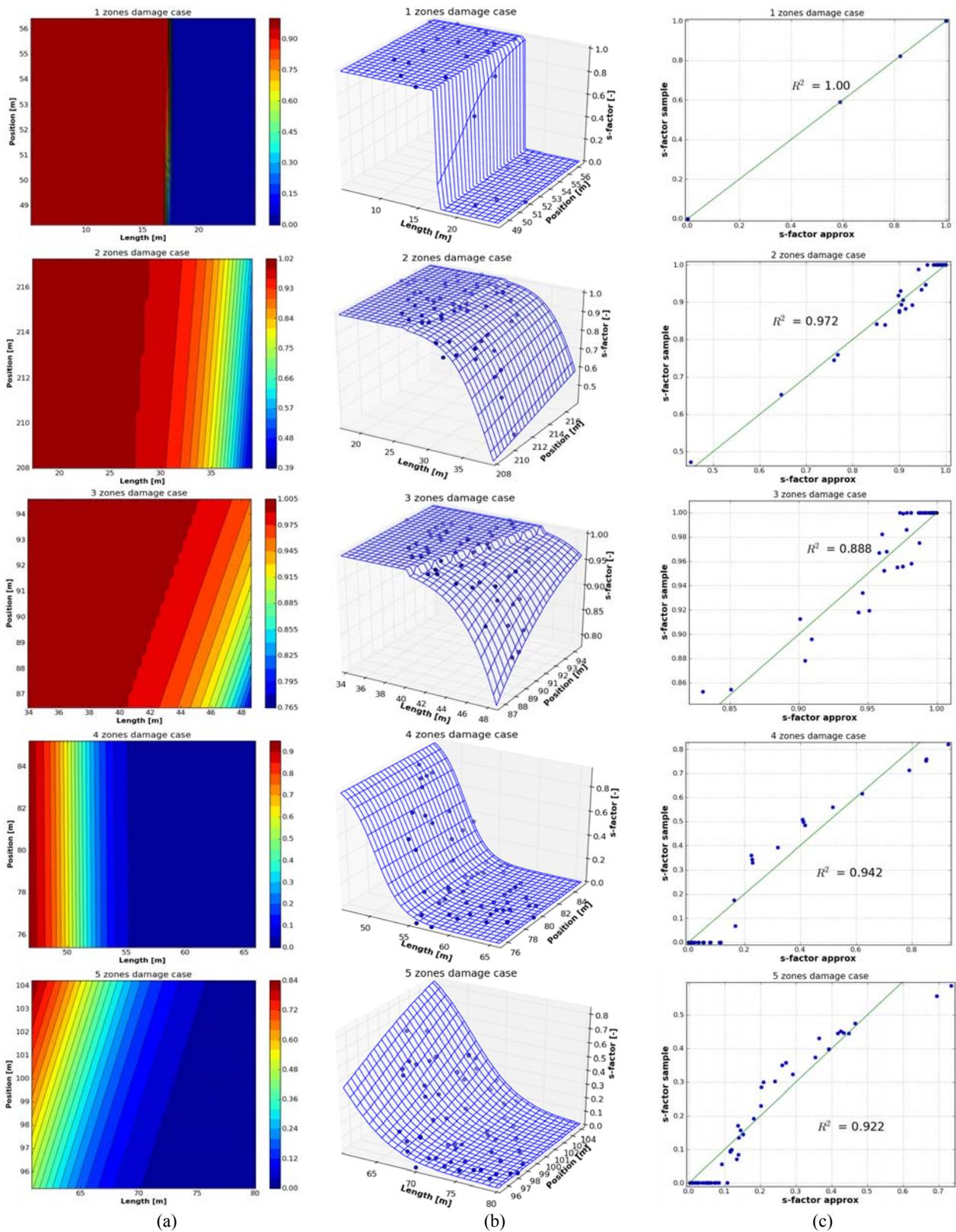


Fig. 3 The variation of the  $s$ -factor function with the length and the position across multiple-zone damage cases: (a) a contour plot of the  $s$ -factor; (b) a 3D surface plots with sample points being approximated by the surface; (c) sampled (observed) vs. approximated (predicted) data plots, which are complementary to (b).

involved in the damage case. The  $s$ -factor surface may exhibit large plateaux with 0/1 values where the function derivatives vanish. Therefore, the  $s$ -factor function may also be classified as a flat or locally constant function. Additionally, the  $s$ -factor function is non-convex, which stems from its sigmoidal shape.

Components  $A_i$  of the subdivision index “A”, as defined in Eq. (2), are hence intersections of convex and non-convex sets that correspond to the  $p$ - and  $s$ -factor functions, respectively. As a result, components  $A_i$  are non-convex functions of higher dimension. Consequently, the subdivision index “A” being a linear combination of non-convex functions is also a non-convex function of high dimension. For the strict convexity of a function implies the existence of just one maximum (or minimum) point, the non-convexity invalidates this theorem. Hence, the subdivision optimization problem such as studied in this work can have more than one global maximum.

### 3. Problem Complexity and Selection of Optimization Strategy

In this section we aim to formalize the complexity of the optimization problem. That is, bearing in mind the character of the  $p$ -factor and  $s$ -factor functions, it becomes obvious that the nonlinearity and irregularity in resulting functions of components  $A_i$  is predominantly driven by the presence of large plateaux in  $s$ -factor functions. Thus, in case of some  $s$ -factors equal one, corresponding components  $A_i$  represent a sum of differentiable and convex functions of the  $p$ -factors, which are hence easy to optimize. As  $s$ -factor decrease, components  $A_i$  become irregular until they cease contributing to the subdivision index “A” when corresponding  $s$ -factors become zero. On this basis, we could define the complexity of the optimization problem to be proportional to the expected  $s$ -factor over all damage cases,  $E(s)$ . The complexity vanishes when the expected  $s$ -factor value is either zero or one, thus forming a bell-shaped curve like the one in Fig. 4.

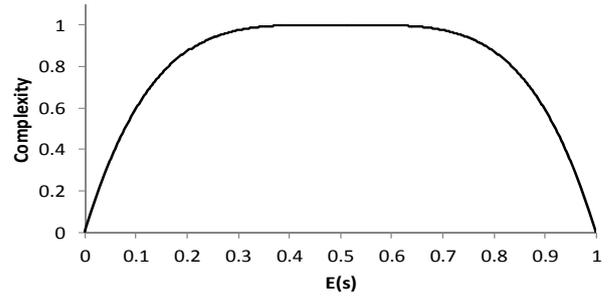


Fig. 4 Optimization complexity vs. expected survivability.

This brings us to the conclusion that the initial (prior to optimization) value of the index “A”, which by definition is the expected survivability, indicates how difficult the optimization of a ship subdivision is going to be. Specifically, the higher initial index “A”, the easier optimization algorithm will be able to improve the design, and vice versa. Certainty, various topological characteristics of the bell-shaped curve in Fig. 4 such as the inception the complexity descent and others will vary across different subdivision designs. However, this subject is beyond the scope of this paper.

As for the choice of an optimization method, deterministic gradient-based methods are definitely not applicable, unless the ship survives all the damages with the probability equal one. In this case the optimization of the index “A” would not be needed. Hence, we have no choice but to employ probabilistic (or stochastic) optimization methods that have been shown [6] to be suitable for this class of problems. Interestingly, probabilistic optimization methods and in particular genetic algorithms (GAs) [7] have been mainly applied in order to improve the index “A” [8-10]. However, the reason why GAs have been so popular is not because of anticipated problem complexity, but due to GA simplicity and hence convenience. This becomes obvious just observing the way the method has been used. That is, due to its stochastic nature, it should let run for an extended period of time<sup>3</sup> to arrive at solutions being close to the global optimum. However, since each evaluation of index “A” takes ca. 3.5 min for

<sup>3</sup> A reasonable number of the index “A” evaluations can be 1,000, which is a population of 10 individuals for 100 generations, neglecting the effect of mutation and crossover.

Ro-Ro ships [9] and up to 20 min for cruise ships, to perform just 1,000 “A” evaluations it would take 58 and 333 hours, respectively. As this might be prohibitive in practice, such an extended optimization is not likely to be performed, satisfying with only some minor improvement. For all probabilistic methods in average having a similar performance [11], the same would also apply in case of using any other probabilistic method.

It stands to reason that the successful optimization of the index “A” requires an extensive exploration of design space. For this to happen, the runtime must be significantly reduced, to allow for numerous optimization runs. Such a reduction is only possible if we replace the time consuming hydrostatic calculations by approximate calculations. In other words, we need to find a surrogate function for the function behind the index “A” that is easy to implement and quick to evaluate. Such a surrogate function, which is based on the regression model from the preceding section, can be proposed to be as follows:

$$\tilde{A}(x) = \sum_{i=1}^{ND} p_i(x_{p_i}) \tilde{s}(\boldsymbol{\beta}, \lambda, \chi, b). \quad (9)$$

Note the  $s$ -factor in Eq. (9) represents the regression model of Eq. (7), whereas the  $p$ -factor is calculated according to the formulae described by the regulations [1, 4]. Interestingly to note, the  $p$ -factor can be also quite accurately approximated using the logistic regression model of Eq. (5) with independent variables corresponding to positions of all the  $p$ -factor affecting bulk heads. Thus for example, Fig. 5 shows a comparison of regression data with sample data for the  $p$ -factor for 5 zones damage case. The plot in Fig. 5 shows a good match even for 5 zones damage case, which involves 6 design variables (No. of zones + 1) affecting the  $p$ -factor. Thus, for fewer zones damage cases, the match should be even better.

Regardless the way the  $p$ -factor is estimated, it does not change the proposed optimization strategy at the heart of which is the use of the surrogate Eq. (9). The next section illustrates an application of the surrogate index “A” function to subdivision optimization.

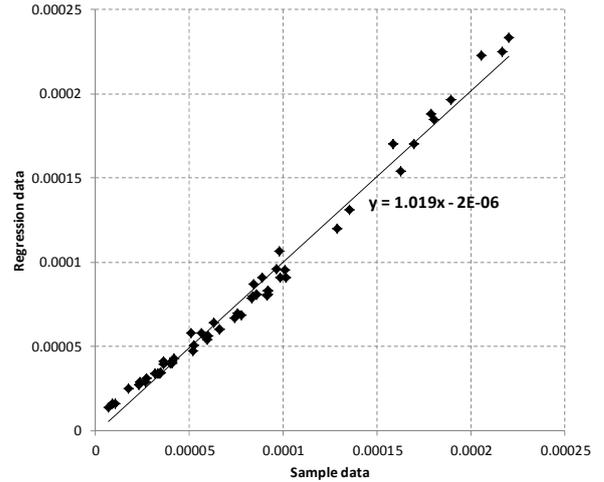


Fig. 5 Regression vs. sample data of the  $p$ -factor for 5 zones damage case. Note, the slope  $\approx 1$  (close to one-to-one match) and the intercept is negligible.

#### 4. Surrogate Optimization of Index “A”

In this section we apply the surrogate index “A” Eq. (9) to optimize the subdivision shown in Fig. 2. First we summarize the process of deriving the regression-based surrogate function.

As for any other regression, we need to provide a data set based on which the regression coefficients can be estimated. The data can be sampled using some randomized sampling mechanism that should also be effective, in view of tedious calculations of the index “A”. We recommend the Latin hypercube sampling (LHS) [12] as an efficient sampling method that uniformly covers the design space. The number of samples in LHS is not required to increase with the number of variables that are subject to variation; this independence is one of the main advantages of this sampling scheme. Another advantage is that random samples can be taken one at a time, remembering which samples have been taken so far. In summary, the  $s$ -factor (and the  $p$ -factor analogically, if needed) is approximated according to the following procedure.

- The subdivision topology (number of bulkheads, tanks and openings) is fixed. This makes sure that the number of damage cases is the same for each run. This also implies that the subdivision topology must be optimized in advance, thus for instance, the optimal

number of bulkheads can be found as described in Refs. [13, 14].

- Generate subdivision design alternatives via LHS, prescribing ranges of variation for each bulkhead. The ranges should be selected in such a way that they do not overlap. The recommended sample size is  $3N$  [15] with  $N$  being the number of design variables (e.g. bulkhead positions) to be optimized.

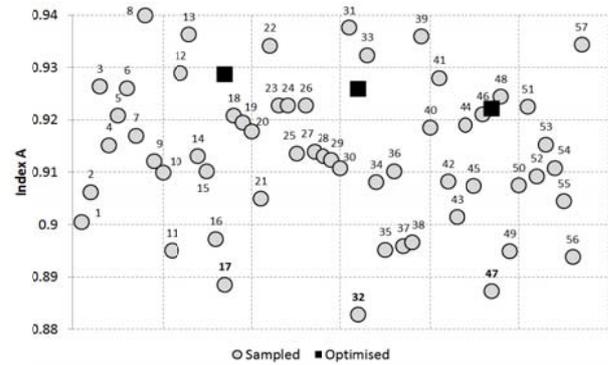
- Run SOLAS 2009 calculations for each subdivision alternative and record their output, which contains all damage cases with s-factor values, damage description (zones damaged etc.) etc. per each damage case.

- Relate each damage case, and hence the s-factor, to design variable through damage descriptions. Thus for example, four zones damage “Z3 / Z4 / Z5 / Z6” would be related to transversal bulkheads TBH2 and TBH6 and a corresponding longitudinal bulkhead between them.

- Apply logistic regression for each damage case, to estimate corresponding regression coefficients  $\beta_i$  in Eq. (6).

In the initial design of the subdivision, we positioned the transversal bulkhead equidistantly, so that  $(x_{i+1} - x_i)/(x_{i+2} - x_{i+1}) = 1/2$ . The first (from the aft) and the last bulkheads were fixed and the total number of transversal bulkheads to optimize was 20. Offsets for longitudinal bulkheads were chosen arbitrary. We produced 57 sample points via LHS and evaluated them by naval architecture software NAPA ([www.napa.fi](http://www.napa.fi)). We used software platform spiral™ [16] for automated sampling, controlled design evaluation and logistic regression. Fig. 6 shows the sample points covering a quite wide range of index “A” values within [0.883, 0.94] interval.

Further, we selected three worst design variations with index “A” values of 0.888, 0.883 and 0.887 (points 17, 32 and 47 in Fig. 6, respectively) and optimized them using the probabilistic method called Simulated Annealing (SA) [17]. In principle, there was little difference in which optimization method was used, for the optimization time was not limited. We allowed SA



**Fig. 6 Index “A” (SOLAS 2009) values for design variations of subdivision model in Fig. 2. The three black circles are the variations optimized using the surrogate function. A label next to a sample point indicates the sample index.**

to run 10,000 function evaluations (which took ca. three minutes), thus making sure that the design space was explored extensively. Fig. 6 also shows the optimization results in the form of three black circles corresponding 17, 32 and 47 sample points. It is important to note that the found design improvements (0.929, 0.926, and 0.922) are significant, but they are clearly inferior to some solutions in the sample set, meaning that the surrogate function did not sufficiently well incorporate those superior designs in its structure. Hence, even though the optimization of the constructed surrogate Eq. (9) converged well, it was not accurate enough to find the relative global optimum, i.e. best solution within the given region.

The apparent reason for this deficiency in the surrogate model lays in the regression error (see Fig. 3), which notably grows with the number of compartments involved in a damage case. Therefore, further analysis is needed to improve the regression model. However, this is beyond the scope of this paper and the interested reader is referred to a related work by Pusa and Vassalos [18].

## 5. Conclusions

It is generally accepted that a proper definition of a problem is much more difficult than its solution afterwards. In absence of a correct problem definition or its understanding, its solution is flawed or at least

inefficient. This paper has shed light on the nature of the subdivision optimization problem (as implied by SOLAS 2009) with the aim to reducing uncertainty while selecting the most relevant optimization strategy.

In particular, we have performed topology analysis of the index “A” sub-functions and concluded that they are generally non-convex, multimodal (the presence of multiple maxima) and can be highly irregular with flat regions where the function derivatives vanish. The existence of large plateaux with  $s$ -factor values 0/1 also means that a Taylor series expansion, which is often employed by some gradient-based optimization algorithms, cannot be applied to the index “A” function, as there is no convergent infinite power series.

On this basis we have suggested that such irregularity of the index “A” function is reduced as the index increases. In other words, the higher the initial flooding survivability (i.e. the subdivision index “A”), the easier is to raise it further by optimizing positions of watertight bulkheads. This constitutes useful knowledge for practitioners who deal with the subdivision design problem on the regular basis.

Further, due to the highlighted irregularity of the index “A” function, optimization algorithms that require computation of derivatives are obviously irrelevant. Therefore, probabilistic methods such as genetic algorithms (GA) (or evolutionary algorithms in general), simulated annealing (SA) and other heuristics should be used instead. It is important to note that particularly because of the presence of the plateau and multiple function maxima, a selected probabilistic method must have mechanisms to handle these search impairing difficulties. Such mechanisms are usually related to diversity preserving strategies. Thus for example, the mutation probability in GA can be made adaptive, automatically increasing when the search stagnates and loses diversity [19]. In SA the cooling schedule can be made less steep/fast or/and the search can be restarted (resetting the initial temperature) when it starts stagnating.

We have also attempted to derive a surrogate

function for the index “A”, aiming to reduce the optimization time. The surrogate function is based on the logistic regression used to approximate the  $s$ -factor. The main reason to use the logistic regression for  $s$ -factor approximation was due to similarity between the topological shapes of the logistic function and the  $s$ -factor function. Specifically, the both functions follow the shape of a sigmoidal function. The logistic regression has proven to be a good way of approximating the  $s$ -factor, although the approximation error is present and it increases with the number of compartments involved in a damage case. We have tested the surrogate index “A” function in optimizing the subdivision of a cruise ship and found that its inaccuracy is detrimental, although the optimization results were feasible and represented significant improvement of initial designs. We hence conclude that the proposed surrogate model is generally unsuitable. It is worth reminding the reader that the approximation of the subdivision index “A” has not been an objective but rather a natural consequence of the analysis presented in this paper.

## References

- [1] MSC. 216 (82), Annex 2: Adoption of amendments to the international convention for the safety of life at sea, 1974, as amended, IMO, London, 2006.
- [2] Z. Karaszewski, M. Pawłowski, A general framework of new subdivision regulations, in: Proceedings of the 8th International Marine Design Conference (IMDC), Athens, 2003, pp. 254-265.
- [3] D. Cabaj, Floodable length surfaces as a novel approach to design of ship subdivision, Ph.D. Thesis, Department of Naval Architecture and Marine Engineering, University of Strathclyde, Glasgow, 2009.
- [4] M. Pawłowski, Probability of flooding a compartment (the  $\pi$  factor)—A critique and a proposal, *Engineering for the Maritime Environment* 219 (2005) 185-201.
- [5] D.W. Hosmer, S. Lemeshow, *Applied Logistic Regression*, John Wiley and Sons, 1989.
- [6] M. Gen, R. Cheng, *Genetic Algorithms and Engineering Optimization (Engineering Design and Automation)*, John Wiley and Sons, 2000.
- [7] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1st ed., Addison-Wesley Professional, 1989.

- [8] E.K. Boulougouris, Optimization of arrangements of Ro-Ro passenger ships with genetic algorithms, *Ship Technology Research* 51 (2004) 99-105.
- [9] A. Papanikolaou, Holistic ship design optimization, *Computer-Aided Design* 42 (2010) 1028-1044.
- [10] A.İ. Ölçer, A hybrid approach for multi-objective combinatorial optimisation problems in ship design and shipping, *Computers and Operations Research* 35 (2008) 2760-2775.
- [11] D.H. Wolpert, W.G. Macready, *No Free Lunch Theorems for Search*, Santa Fe Institute, 1995.
- [12] R.L. Iman, An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment, *Journal of Quality Technology* 13 (1981) 174-183.
- [13] N. Tsakalakis, R. Puisa, D. Vassalos, Goal-based subdivision and layout, in: *10th International Conference on Stability of Ships and Ocean Vehicles*, St Petersburg, Russia, 2009.
- [14] G. Simopoulos, Sensitivity analysis of the probabilistic damage stability regulations for RoPax vessels, *Maritime Science and Technology* 13 (2008) 164-177.
- [15] R. Jin, Comparative studies of metamodeling techniques under multiple modeling criteria, *Structural and Multidisciplinary Optimization* 23 (2000) 1-13.
- [16] R. Puisa, K. Mohamed, Prudent platform for multidisciplinary ship design exploration, analysis and optimisation, in: *ICCAS 2011, International Conference on Computer Applications in Shipbuilding*, Trieste, Italy, 2011.
- [17] S. Kirkpatrick, Optimization by simulated annealing, *Science* 220 (1983) 671-680.
- [18] R. Puisa, D. Vassalos, Surrogate optimisation of probabilistic subdivision index, in: *Design and Operation of Passenger Ships Conference*, London, 2011.
- [19] R. Puisa, The adaptive stochastic algorithms for structure optimization of mechanical parts, Ph.D. Thesis, Vilnius Gediminas Technical University, Vilnius, 2005.