X International Conference on Computational Methods in Marine Engineering
MARINE 2023
J. García-Espinosa, L. González, J.E. Gutiérrez & B. Serván-Camas (Eds)

A Non-Deterministic Propeller Design Optimization Framework Leveraging Machine Learning Based Boundary Element Methods Surrogates

Andrea Coraddu[1*], Stefano Gaggero[2], Diego Villa[2], and Luca Oneto[2]

[1] Delft University of Technology, Delft, The Netherland e-mail: a.coraddu@tudelft.nl

[2] University of Genoa, Genoa, Italy e-mail: {Stefano.Gaggero, Diego.Villa, Luca.Oneto}@unige.it

[*] Corresponding author: First Author, a.coraddu@tudelft.nl

ABSTRACT

Marine propeller design is a cornerstone of naval architecture and marine engineering, serving as a classic optimization problem that has captivated the attention of scholars, engineers, and practitioners for generations. The challenge lies in developing efficient, reliable, and cost-effective propellers that excel in a variety of operating conditions while adhering to strict environmental regulations and addressing noise and vibration concerns. As a critical component in the propulsion systems of ships, submarines, and other marine vessels, the marine propeller's design directly impacts vessel performance, fuel consumption, and emissions. Marine propellers often operate under uncertain conditions, including inflow, rate of revolutions, and manufacturing tolerances. On the one hand, a deterministic design approach that does not consider these stochastic inputs can lead to excessive sensitivity to slight variations in operational conditions, resulting in a geometry that is sub-optimal for real-world functioning scenarios. On the other hand, the computational effort required to quantify the design's uncertainties may be prohibitive even when mid-fidelity solvers, like Boundary Element Methods (BEM), are employed in Simulation-Based Design Optimization. As in the case of deterministic design optimizations driven by mid-fidelity codes, Machine Learning methodologies represent a computational booster of the procedure. By realizing computationally cheap yet accurate surrogate models of the key performance indicators of the design, they allow for the hundreds of thousands of calculations needed by sampling methods to evaluate the uncertainty of the design and drive the process towards configurations less sensitive to inputs variations, making the non-deterministic design optimization a feasible alternative to conventional deterministic design-by-optimization methodologies. In this paper, deterministic and non-deterministic designs are carried out in the case of a conventional propeller, considering uncertainties of the nominal functioning conditions. BEMs computations are used to train accurate Machine Learning-based surrogate models. Thanks to their cost-effectiveness in the forward phase, they enable the assessment of the uncertainties necessary for a non-deterministic design optimization framework. In order to demonstrate the efficacy of the proposed methodology, the optimal geometries attained through the utilization of surrogates are re-tested utilizing medium-fidelity BEM solvers to validate the superior performance of the non-deterministic design approach.

Keywords: Robust Optimization; Boundary Element Methods; Surrogate Models; Machine Learning; Simulation-Based Design Optimization.

NOMENCLATURE

SBDO	Simulation-Based Design Optimization
CFD	Computational Fluid Dynamics
BEM	Boundary Element Methods
RANS	Reynolds-averaged Navier–Stokes
ML	Machine Learning
DDSs	Data-driven Surrogates
GA	Genetic Algorithm
SVR	Support Vector Regression
MAE	Mean Absolute Error
AT	Average Time
MS	Model Selection
EE	Error Estimation

1. INTRODUCTION

In the last few years, the design of marine propellers has significantly evolved. Basic vortex-lattice-like lifting line and lifting surface methods were outclassed by Simulation-Based Design Optimization (SBDO) approaches fed by medium- and high-fidelity Computational Fluid Dynamics (CFD) calculations [2, 10], often supported by Machine Learning (ML) based surrogate models to overcome computational resources' limitations or comply with relatively short design deadlines [13]. In most cases, only deterministic designs, i.e., for given functioning conditions, were addressed. This was a natural consequence of the lifting line/lifting surface commonly used design methods and can be considered a limitation of the design process that can be addressed using SBDO methods. However, in real functioning scenarios, propellers operate in conditions (e.g., inflow velocities and rate of revolutions) that may have a certain level of uncertainty or, from the geometrical point of view, they may be affected by tolerances that can nullify some of the advantages foreseen and expected by the optimization of geometrical details not sufficiently monitored and accurately handled during the manufacturing process. This uncertainty of the input parameters is reflected in the design performance and, consequently, in the overall efficacy of the propulsion. Therefore, without considering the stochastic nature of inputs, a deterministic design approach may result in excessive sensitivity to slight variations of the operational conditions, leading to a final geometry that is not optimal in its real and uncertain functioning scenario.

If, when using traditional vortex-based design methods, such as the lifting line methods [27, 32], this issue related to sensitivity to disturbances was addressed by a balanced choice of the nominal functioning of the propeller and a certain number of (manual) adjustments of the final geometry based on off-design functioning prediction, the adoption of SBDO approaches fostered a more quantitative and rationale process. As proposed in Bertetta et al. [2], controllable pitch propellers were designed to comply with two functioning conditions (design and reduced ship seed), addressed only with pitch variation and leading to suction and pressure side cavitation, by a design-by-optimization process. Performance indicators of the propeller were collected for both functioning, and the optimal geometry was finally selected as a balance of the contrasting objectives. Similar criteria inspired the design of the high-speed propellers of [12], where the selection of the optimal propeller was supported by the quantification of efficiency and cavitation in correspondence with some very off-design functioning conditions related to the nature of the small and very high-speed ship the propeller was designed for. The detailed investigation of selected propeller variants accounting for different phenomena not directly

assessed in the optimization process was used in [1] to choose a final geometry suitable for pressure pulses and hull interaction improvements. Recently, the propulsion of wind-assisted and wind-powered vessels [17] required a reformulation of the design objectives through a re-definition of the operational profile with the aim of determining the most important conditions for the route and arranging a design procedure capable of accounting for the very different working regime associated with the expected variations of the environmental conditions. Interactive Genetic Algorithm (GA) supported the simulation-based design optimization [18] to include the designer experience automatically in the design process for the reduction of the total energy consumption of the ship, which was assumed as a sort of integrated design objective simultaneously accounting for the different functioning during the ship route. Also, in this case, however, the propeller design was addressed within a deterministic paradigm by designing several alternatives, one for each of the relevant functioning conditions observed for the ship and by selecting the geometry which provides the best performances when tested in all the other conditions. None of the previous cases, however, were formulated as a truly non-deterministic design-by-optimization, and generally speaking, there are very few examples of robust design processes applied to marine propellers in the literature. An example is given by [33] where CFD calculations were employed for the uncertainty estimations of ducted propeller performances that drove the design towards a robust design. Perturbation of the geometrical features and the functioning point (such as the inflow velocities) of the propulsor were included in the analyses. Several techniques, like the sparse grid quadrature method [16, 31], were used to reduce the theoretically exponential growth of CFD simulations required for the characterization of uncertainties.

The cost of robust approaches consists, indeed, of the computational effort required to quantify the uncertainties of the design, which may result in a prohibitive computational burden also when mid-fidelity solvers, like BEM [14] of the current paper, are employed in the SBDO. However, robust design through optimization is often necessary for marine propellers due to the highly uncertain and complex operating conditions they are exposed to during their life [11]. Propellers are required to operate with high efficiency across a range of operating conditions, including inflow and rate of revolutions, minimizing at the same time the occurrence of cavitation. Achieving these objectives is extremely important because cavitation can cause significant damage to the propeller blades and surrounding structures, reducing the overall performance and service life of the propeller [25]. As such, propeller design must be optimized to account for the complex hydrodynamic interactions between the flow field and the propeller blades, ensuring that the risk of cavitation is always minimized while maintaining high levels of efficiency. The development of reliable and flexible tools to address this problem is a fundamental step towards overall better propulsive performances as required by international regulation bodies like IMO by means of EEOI and EEDI efficiency-related indexes.

As in the case of deterministic design optimizations driven by high-fidelity codes, ML methodologies represent a computational booster of the procedure [22]. Surrogate-based optimization techniques may replace the computationally expensive objective function with properly defined surrogate models, enabling optimization with fewer high-cost function evaluations. By realizing computationally cheap yet accurate surrogate models of the key performance indicators of the design, they allow for the hundreds of thousands of calculations needed by sampling methods to evaluate the uncertainty of the design and drive the process towards configurations less sensitive to input variations. Surrogate models have been successfully applied for handling complex and non-linear optimization problems [9] when coupled with inspired by natural selection search techniques that iteratively modifies a population of candidate solutions to find the optimal solution making the non-deterministic design optimization a feasible alternative to conventional deterministic design-by-optimization methodologies. In order to address and reduce the computational requirements induced by the optimization process, in the present work, we propose to overcome this limitation by exploiting a data-driven approach [15, 29]. Data-driven

Surrogates (DDSs) allow to approximate the mid-fidelity BEM calculations accurately, exploited in the optimization and as the primary reason for its computational requirements, without requiring to perform too strong simplifications. DDSs can automatically learn a functional representing the output of the BEM evaluations using a series of data generated by running multiple times (but less than the optimization requirements) the solver. The main advantage of this approach is that the learned functional is computationally inexpensive to apply, so overcoming the limitations of using the selected BEM solver. On the contrary, the disadvantage is that the evaluation of this function is computationally expensive to build. In fact, building a model using DDSs requires running the solver code multiple times to generate the data and then training the functional with an ML algorithm. Nevertheless, once this procedure is completed, the resulting learned functional can be reused inexpensively as often as necessary. Note that this approach has already been exploited in the past to resolve many similar problems [23].

Specifically, in this paper, deterministic and non-deterministic designs are carried out in the case of a conventional propeller, considering uncertainties of the nominal functioning conditions. Among the possible stochastic inputs to the propeller, the geometrical characteristics and functioning conditions are the most important ones. The latter is the most usual since they are strictly related to the vessel's operations. Slight modifications of the ship displacement, or the occurrence of fouling on the hull surface, can easily determine variations of the ship resistance and, then, of the propeller rate of revolution to achieve the desired speed. Interactions with environmental currents change the inflow field to the propeller. Moreover, ensuring robustness against functioning conditions intrinsically presuppose robustness against geometrical uncertainties since any variation in the rate of revolution or the inflow velocity can be interpreted such as a modification of the angle of attack and, consequently, of the pitch/camber of the blade.

Mid-fidelity BEM calculations are used to train accurate DDSs, which, in turn, support the evaluation of the uncertainties needed for non-deterministic optimization. They are also used to verify the final geometries coming from the robust design process driven by the DDSs and assess the stabler behaviour of these optimal geometries compared to that of the initial propeller assumed as the reference and, in particular, to those obtained from a conventional, deterministic, design-by-optimization process. This latter SBDO is carried out entirely using BEM calculations to follow the well-consolidated design practice.

The rest of the paper is organised as follows. Section 2 will describe the baseline reference four-blade propeller used to evaluate the improvements of the proposed non-deterministic design framework and the data generated by running the BEM multiple times to create a dataset for the DDSs. Section 3 will present our DDSs to approximate the BEM using the generated data of Section 2. Section 4 will present the formalization of deterministic and non-deterministic optimization problems under consideration, how we optimized them, and where the surrogates are employed to reduce the computational burden. Section 5 will report the result of applying the methodology presented in Section 4 and 3 using the data described in Section 2. Performances of propellers from deterministic and non-deterministic designs are compared, and the effects of input variability are discussed. Section 6 will conclude the paper.

2. AVAILABLE DATA

This Section presents the dataset for the surrogate DDMs learning phase generated using the BEM. The dataset, created through BEM simulations, forms a critical aspect of our study, providing a rich source of information to train ML models for predicting propeller behaviour under different rotational speeds and ship velocities. The dataset was carefully designed, processed, and validated to ensure its

quality, enabling us to train DDSs models that generalize to real-world scenarios. We first introduce in Section 2.1 the baseline reference propeller used to evaluate the improvements from the application of our methodology. Then we describe the nature and characteristics of the dataset generated through the BEM simulations in Section 2.2.

2.1 Reference Propeller

The improvements provided by a non-deterministic design have been tested starting from a custom propeller design obtained for the MARINE Ferry [8]. This vessel is a fast twin-screw passenger vessel designed by the Maritime Research Institute Netherlands as a benchmark test case representative of modern ferries. The ship has a length of 190 m and can reach a design speed of 25 kn. It is equipped with two inward over-the-top propellers and two spade rudders. The main particulars of the propeller blade were not available. To this aim, we used a traditional lifting line/lifting surface method [3] to identify the geometry of a four-blade propeller that represents the baseline reference to evaluate the possible improvements granted by the application of i) a conventional deterministic SBDO (i.e., considering only the nominal functioning in the design by optimization process), ii) the SBDO process carried out under uncertainties. This initial geometry has a diameter of 5.4 m and was designed for a ship speed of 25 kn at 200 RPM to operate in the wake (circumferentially averaged) of Figure 1 having a wake fraction $(1 - w)$ equal to 0.869. This corresponds to a nominal advance coefficient of 0.616, a thrust coefficient K_T equal to 0.0894 at a cavitation index σ_N based on the propeller rate of revolution equal to 0.85. The wake-fraction has been obtained by means of the high-fidelity RANS solver (OpenFOAM) calculations performed following the ITTC Recommended Procedure and Guideline. An in-house developed procedure has been considered to set up the solver and the mesh grids. A detailed description is omitted here for brevity, however, more details can be found in [34].

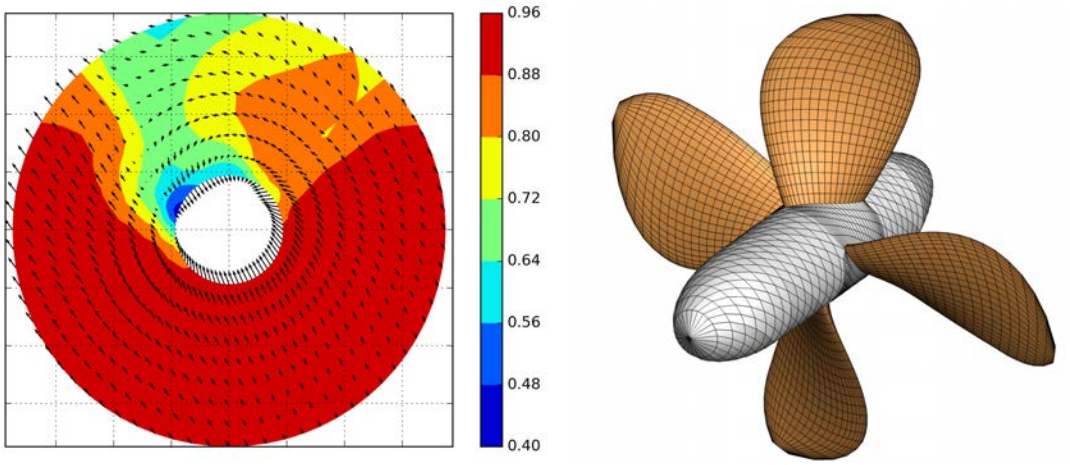

Figure 1: Nominal wake of the MARINE Ferry for the initial propeller design (left) and reference propeller with computational grid for BEM calculations (right).

In correspondence with this design functioning condition, the propeller has an open water efficiency of 0.7255 and a cavitation extension on the back side at the tip equal to 0.055 m^2 on each blade. Those results have been achieved employing the BEM calculations using a surface mesh of 1000 panels per blade, as shown in Figure 1. Those settings have been successively adopted for any propeller geometry tested in this work when using the BEM code. Pressure side cavitation at the design functioning does not occur. This shape has been considered the baseline performance at the nominal design functioning to be improved by the optimization processes.

To address the non-deterministic design, we perturbed the nominal functioning both in terms of vessel speed and propeller rate of revolution, assuming for these two quantities probability normal distributions centered in the nominal functioning condition (i.e., $\mu_{v_s} = 25$ kn and $\mu_n = 200$ RPM). Variances of the distributions were assumed equal to 0.04 and 16 for the ship speed and the propeller rate of revolution respectively, leading to a range of variations ($\pm 2\sigma$ which represent the 95% of values) of the advance coefficient J between 0.58 and 0.67 and of the cavitation index between 0.78 and 0.91. The predicted cavitation extension in correspondence of maximum and minimum values of these parameters and at the design functioning are summarized in Figure 2. Pressure side cavitation is always avoided, also in off-design functioning. In contrast, suction side cavitation always appears as leading-edge sheet cavitation with significant tip vortex in correspondence with the most loaded conditions.

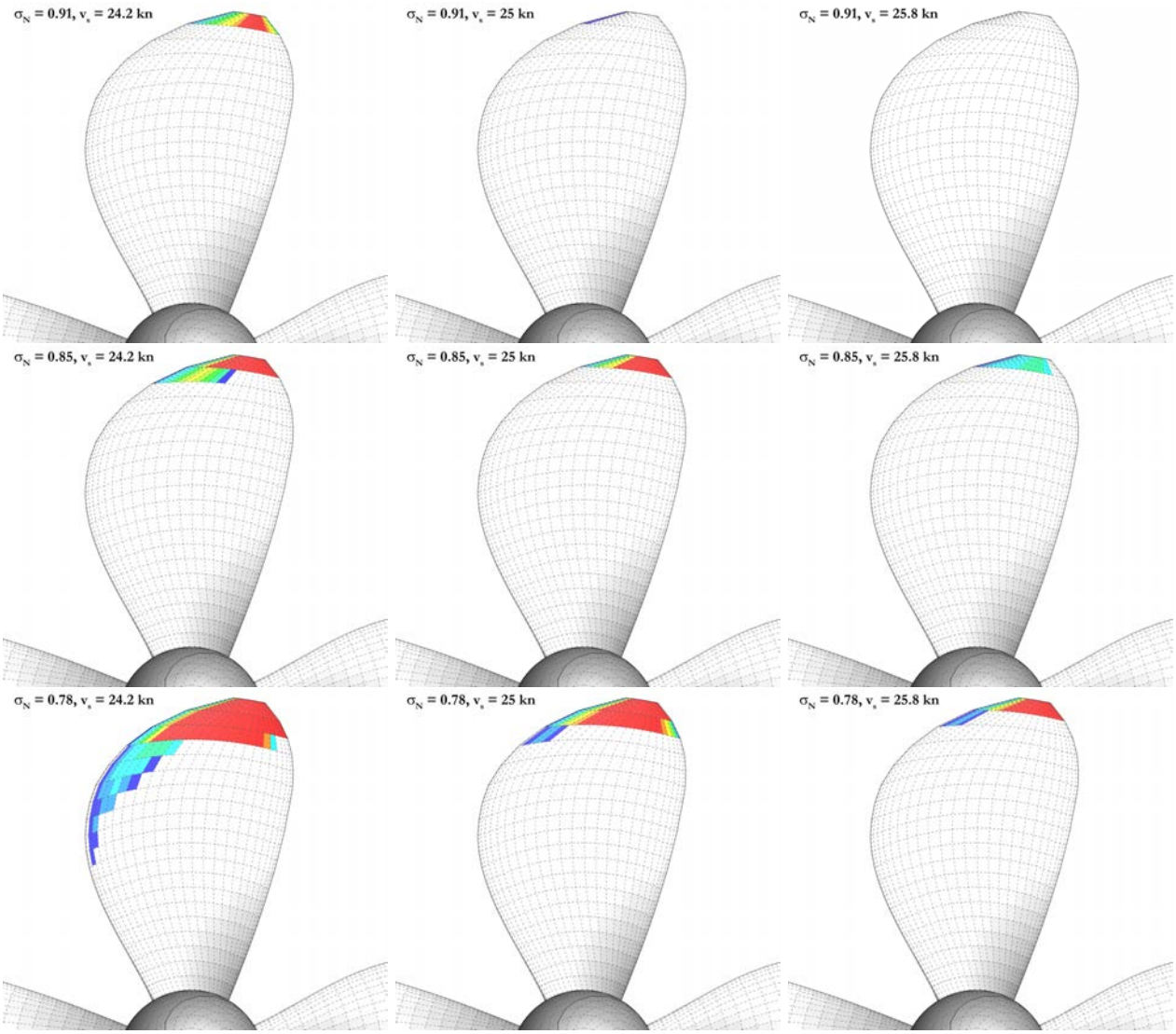

Figure 2: Reference propeller estimated cavitation extension on the blade suction side using BEM calculations. From top to bottom: $\sigma_N = 0.91, 0.85, 0.78$. From left to right: Ship speed of 24.2, 25 and 25.8 kn.

To train the DDSs, i.e. to populate a design space with thousands of geometrically different configu-

rations, each characterized by their performances and then to run the SBDO, we applied our standard parametrization of the propeller blade [2]. The blade parametrization is described using B-spline curves representing the radial and sectional most relevant characteristics of the blade geometry (i.e. pitch, chord, camber distributions and so on), as already successfully done in several other applications [10]. These curves are controlled by means of as few control points as possible able to generate all the desired shapes, ensuring at the same time the smoothness of the final shape. Control points of the curves turn into the design parameters and realize the design space for both the training phase of the DDSs and the optimization. For this particular case, we assume a blade geometry described using 38 design parameters controlling the chord, the pitch, the skew and the maximum sectional camber distributions along the blade span, in addition to the sectional hydrofoil shape through the camber line and thickness along the chord. To simplify the description, the non-dimensional hydrofoil shape is considered constant at any radial position. Blade thickness is recalculated for any analyzed geometry depending on the design loading condition and chord, pitch and skew choice to account for structural reasons, while the blade rake is always kept constant as the reference propeller. An example of the parametric description is given in Figure 3, while the complete list of parameters utilized as input space for the learning phase of the DDSS and design variables for the optimization problems of this work and the relative range of variations are summarized in Table 1.

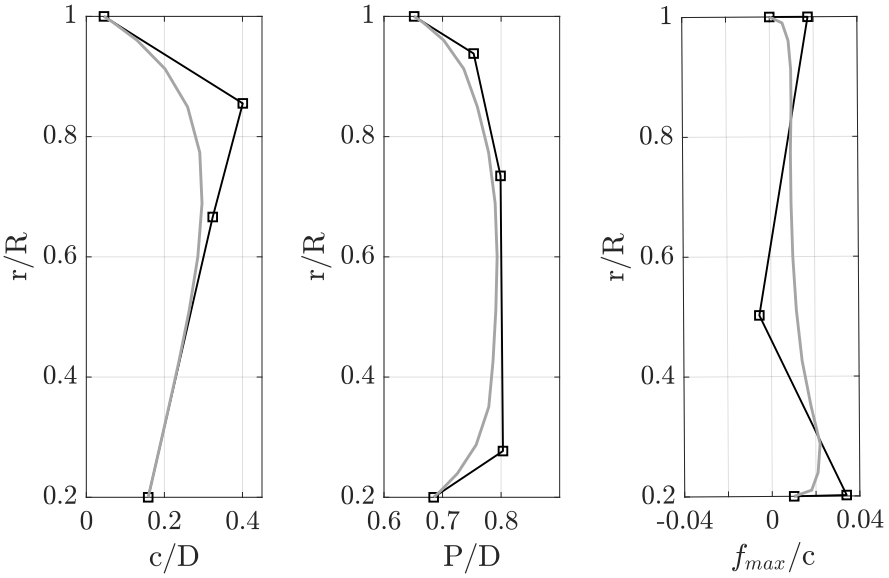

Figure 3: Parametric description using B-spline polygons of the reference propeller geometry.

2.2 Dataset Generation

In order to reduce the computational requirements needed to predict the propeller efficiency η_o, the trust coefficient K_T, torque coefficient K_Q, pressure-side cavitation area $A_{cav,front}$, and suction-side cavitation area $A_{cav,back}$), we propose to substitute the hydrodynamics analysis performed with the BEM, with surrogate DDMs. In particular, the summary of the inputs feeding the surrogate models and the related outputs is reported in Table 1, together with the ranges for building the dataset.

The database contains $50,000$ geometries generated randomly in the ranges reported in Table 1. For each of the $50,000$ geometry, the BEM has been run to compute the output space quantities described in Table 1. Each simulation took an average of 20 seconds on a machine equipped with two Intel

Table 1: List of inputs and outputs of the surrogate models.

	Name	Description	Range
Input Space	v_s	Ship Speed [m/s]	12.2÷13.5
	n	Propeller rate of revolution [rpm]	190÷210
	$c_{2,x}$	Radial position of chord control point n.2 [%]	0.55÷0.90
	$c_{3,x}$	Radial position of chord control point n.3	0.65÷0.85
	$c_{1,y}$	Chord of control point n.1	0.10÷0.20
	$c_{2,y}$	Chord of control point n.2 [%]	0.00÷0.65
	$c_{3,y}$	Chord of control point n.3	0.30÷0.45
	$s_{2,x}$	Radial position of skew control point n.2	0.50÷0.70
	$s_{3,x}$	Radial position of skew control point n.3	0.80÷0.90
	$s_{2,y}$	Skew of control point n.2	-6.0÷ -2.0
	$s_{3,y}$	Skew of control point n.3	-4.0÷ -1.0
	$s_{4,y}$	Skew of control point n.4	11.0÷15.0
	$pd_{2,x}$	Radial position of pitch control point n.2	0.25÷0.35
	$pd_{3,x}$	Radial position of pitch control point n.3	0.635÷0.78
	$pd_{4,x}$	Radial position of pitch control point n.4	0.85÷0.92
	$pd_{1,y}$	Pitch of control point n.1 [%]	0.20÷0.80
	$pd_{2,y}$	Pitch of control point n.2	0.75÷0.90
	$pd_{3,y}$	Pitch of control point n.3	0.75÷0.90
	$pd_{4,y}$	Pitch of control point n.4	0.60÷0.75
	$pd_{5,y}$	Pitch of control point n.5 [%]	0.40÷0.80
	$f_{2,x}$	Radial position of maximum sectional camber control point n.2	0.21÷0.35
	$f_{3,x}$	Radial position of maximum sectional camber control point n.3	0.45÷0.65
	$f_{4,x}$	Radial position of maximum sectional camber control point n.4	0.82÷0.90
	$f_{1,y}$	Maximum sectional camber of control point n.1	0.00÷0.004
	$f_{2,y}$	Maximum sectional camber of control point n.2	0.00÷0.004
	$f_{3,y}$	Maximum sectional camber of control point n.3	0.00÷0.004
	$f_{4,y}$	Maximum sectional camber of control point n.4	0.00÷0.004
	$f_{5,y}$	Maximum sectional camber of control point n.5	0.00÷0.015
	$x_{t,3}$	Chordwise position of hydrofoil thickness control point n.3 [%]	0.30÷0.60
	$x_{t,4}$	Chordwise position of hydrofoil thickness control point n.4	0.40÷0.60
	$x_{t,5}$	Chordwise position of hydrofoil thickness control point n.5 [%]	0.20÷0.70
	$y_{t,2}$	Hydrofoil thickness of control point n.2	0.15÷0.20
	$y_{t,3}$	Hydrofoil thickness of control point n.3 [%]	0.40÷0.80
	$y_{t,4}$	Hydrofoil thickness of control point n.4	0.40÷0.70
	$y_{t,5}$	Hydrofoil thickness of control point n.5 [%]	0.30÷0.70
	$x_{f,2}$	Chordwise position of hydrofoil camber control point n.2 [%]	0.15÷0.35
	$x_{f,3}$	Chordwise position of hydrofoil camber control point n.3	0.30÷0.40
	$x_{f,4}$	Chordwise position of hydrofoil camber control point n.4 [%]	0.45÷0.70
	$y_{f,2}$	Hydrofoil camber of control point n.2 [%]	0.20÷0.65
	$y_{f,3}$	Hydrofoil camber of control point n.3	0.80÷1.15
	$y_{f,4}$	Hydrofoil camber of control point n.4 [%]	0.40÷0.75
Output Space	$A_{cav,back}$	Suction-side Cavitation Area	–
	$A_{cav,face}$	Face-side Cavitation Area	–
	η_0	Efficiency	–
	K_Q	Torque Coefficient	–
	K_T	Thrust Coefficient	–

Xeon Silver 4216, 128 GB of RAM, and 512 GB SSD running Windows Server 2019 for a total of 280 hours for creating the entire dataset. Significant man-hours are also required for the modelling phase before each simulation. Based on these computational requirements, it becomes evident that even though the BEM solver is generally considered with a low computational impact, its use is inadequate to optimize the propeller design for the robust optimization problem presented in this work. This is also due to the fact that for each design, several calculations are needed to generate the distributions of the output quantities. To address these computational barriers, we leverage surrogate DDMs.

3. SURROGATE MODELS

The problem that we want to face in this work is to predict the result of the BEM (i.e., $A_{cav,face}$, $A_{cav,back}$, η_o, K_Q, K_T) based on the input parameters reported in Table 1 using the data produced by a $50,000$ BEM simulations (see Section 2). This problem can be easily mapped into a now-classical supervised ML problem, particularly an ML regression problem [29]. In regression, we have an input space $\mathcal{X} \subseteq \mathbb{R}^d$ composed of d features (in our case the 41 input features), an output space $\mathcal{Y} \subseteq \mathbb{R}$ (in our case the 5 target features), and a series of n examples, a dataset, input/output relation $\mathcal{D}_n = \{(\boldsymbol{x}_1, y_1), \cdots, (\boldsymbol{x}_n, y_n)\}$ where $\boldsymbol{x}_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$ $\forall i \in \{1, \cdots, n\}$. The scope is to learn the input/output relation $\mu : \mathcal{X} \to \mathcal{Y}$ based just on $\mathcal{D}_n$. Generally, μ is a probabilistic relation, but in our case, this relation is induced by the BEM, so it is deterministic. An ML regression algorithm $\mathscr{A}$, characterized by its hyperparameters $\mathcal{H}$, selects a model f inside a set of possible ones $\mathcal{F}$ based on the available data $\mathscr{A}_{\mathcal{H}} : \mathcal{D}_n \times \mathcal{F} \to f$. In this work, we also request f to require as little computational requirement as possible so the prediction (i.e., computation) of $f(X)$ will be less time-consuming. $\mathcal{F}$ is generally unknown and depends on the choice of $\mathscr{A}$ and $\mathcal{H}$. Many different ML algorithms exist in the literature [7, 15, 29, 35] but, even if the no-free-lunch theorem states [36] there is no way to determine a-priory the best ML algorithms to use for a specific application, in this work we detail just the model that performed best for space constraints. The error and the computational requirements of f in approximating μ are measured by a prescribed metric $M : f \to \mathbb{R}$. We will use the Mean Absolute Error (MAE) for the error, and for the computational requirements, we will use the Average Time (AT) in making a prediction. For the sake of selecting the best ML algorithms and the related optimal hyperparameters and to estimate the performance of the final model according to the desired metrics, a statistically consistent Model Selection (MS) and Error Estimation (EE) phase needs to be performed [26].

In this work, we leveraged the Kernel Methods, a family of techniques that exploits the "Kernel trick" for distances to extend linear techniques to the solution of non-linear problems [28]. Kernel methods select the model which minimizes the trade-off between the performance, measured with a defined metric, over the data and the complexity of the solution, measured with different measures of complexities [29, 30]. represents the most known and effective Kernel methods techniques. The hyperparameters of the SVR are: the kernel, which is usually fixed to be Gaussian because of the reasons described in [19], the kernel hyperparameter γ, which regulates the non-linearity of the solution, the γ regularization hyperparameter C, which trades-off accuracy and complexity of the solution, and ϵ which regulates the sparsity of the solution and then its computational requirements. C, γ, and ϵ need to be tuned during the MS phase.

MS and EE deal with the problem of tuning and assessing the performance of an ML algorithm [26]. Researchers and practitioners commonly use resampling techniques since they work well in most situations, which is why we will exploit them in this work. Other alternatives exist based on the Statistical Learning Theory but tend to underperform resampling techniques in practice. Resampling techniques are based on a simple idea: the original dataset $\mathcal{D}_n$ is resampled once or many (n_r) times,

with or without replacement, to build three independent datasets called learning, validation and test sets, respectively $\mathcal{L}_l^r$, $\mathcal{V}_v^r$, and $\mathcal{T}_t^r$, with $r \in \{1, \cdots, n_r\}$ such that $\mathcal{L}_l^r \cap \mathcal{V}_v^r = \oslash$, $\mathcal{L}_l^r \cap \mathcal{T}_t^r = \oslash$, $\mathcal{V}_v^r \cap \mathcal{T}_t^r = \oslash$, and $\mathcal{L}_l^r \cup \mathcal{V}_v^r \cup \mathcal{T}_t^r = \mathcal{D}_n$. Subsequently, to select the best hyperparameters' combination $\mathcal{H}$ in a set of possible ones $\mathfrak{H} = \{\mathcal{H}_1, \mathcal{H}_2, \cdots\}$ for the algorithm $\mathscr{A}_{\mathcal{H}}$ or, in other words, to perform the MS phase, the following procedure has to be applied:

$$\mathcal{H}^* : \quad \arg\min_{\mathcal{H} \in \mathfrak{H}} \sum_{r=1}^{n_r} M(\mathscr{A}_{\mathcal{H}}(\mathcal{L}_l^r), \mathcal{V}_v^r), \tag{1}$$

where $\mathscr{A}_{\mathcal{H}}(\mathcal{L}_l^r)$ is a model built with the algorithm $\mathscr{A}$ with its set of hyperparameters $\mathcal{H}$ and with the data $\mathcal{L}_l^r$, and where $M(f, \mathcal{V}_v^r)$ is a desired metric. Since the data in $\mathcal{L}_l^r$ are independent of the data in $\mathcal{V}_v^r$, $\mathcal{H}^*$ should be the set of hyperparameters which allows achieving a small error on a data set that is independent of the training set.

Then, to evaluate the performance of the optimal model, which is $f_{\mathscr{A}}^* = \mathscr{A}_{\mathcal{H}^*}(\mathcal{D}_n)$ or, in other words, to perform the EE phase, the following procedure has to be applied:

$$M(f_{\mathscr{A}}^*) = \frac{1}{n_r} \sum_{r=1}^{n_r} M(\mathscr{A}_{\mathcal{H}^*}(\mathcal{L}_l^r \cup \mathcal{V}_v^r), \mathcal{T}_t^r). \tag{2}$$

Since the data in $\mathcal{L}_l^r \cup \mathcal{V}_v^r$ are independent of the ones in $\mathcal{T}_t^r$, $M(f_{\mathscr{A}}^*)$ is an unbiased estimator of the true performance, measured with the metric M, of the final model [26].

In this work we will rely on Complete k-fold cross validation which means setting $n_r \leq \binom{n}{k}\binom{n-\frac{n}{k}}{k}$, $l = (k-2)\frac{n}{k}$, $v = \frac{n}{k}$, and $t = \frac{n}{k}$ and the resampling must be done without replacement [26].

4. PROPELLER DESIGN OPTIMIZATION FRAMEWORK

In this section, we introduce the formulation of optimization problems pertaining to both deterministic and non-deterministic methodologies for marine propeller design. We untangle the conventional deterministic marine propeller optimization problem (Section 4.1), as well as a non-deterministic optimization problem (Section 4.2) that incorporates uncertainties associated with the operational environment and design parameters.

4.1 Classic Optimization Problem - the Deterministic Approach

The optimization problem for marine propeller design can be formulated as a multi-objective optimization problem, where the aim is to find the best trade-off between competing objectives, subject to various constraints. The objectives and constraints can be tailored to the specific requirements of the propeller design project. Due to the objectives' complex and conflicting nature, multi-objective optimization techniques [4, 5] can be employed to efficiently search the design space and generate a set of optimal solutions, known as the Pareto front. The final propeller design can be selected from this front based on the designer's preferences and priorities. A general formulation of the problem can be given as follows:

$$\max_{\mathbf{X}\in\mathbb{R}} \quad \eta_0(\mathbf{X}) - (1-\lambda_1)A_{cav,face}(\mathbf{X}) - (1-\lambda_2)A_{cav,back}(\mathbf{X}) \tag{3}$$
$$\text{subject to} \quad \begin{cases} K_T(\mathbf{X}) \leq K_{T,min} \\ K_T(\mathbf{X}) \geq K_{T,max} \\ x_1 = v_s \\ x_2 = n \end{cases},$$

where $\mathbf{X} = \{x_1, x_2, \cdots, x_n\}$ represents the vector of deterministic design variables (Input Space reported in Table 1), obtained keeping the design variables $x_1 = v_s$ and $x_2 = n$ equal to their design values of 25 kn and 200 RPM, respectively. λ_1 and λ_2 in $[0, 1]$ define the importance of the different objectives, i.e., for $\lambda \to 1$ we care more about the weight than the pressure-side cavitation area and vice-versa for $\lambda_1 \to 0$. The same concept applies to λ_2, for $\lambda \to 1$ we care more about the weight than the suction-side cavitation area and vice-versa for $\lambda_2 \to 0$. Solving Equation (3) for different values of λ_1 and λ_2 allows for creating the so-called Pareto frontier in a computationally efficient way [6].

The optimization problem of Equation 3 has a non-linear and non-convex objective and a series of non-linear constraints. In order to solve this problem, different approaches can be exploited [20]. A series of no-free-lunch theorems [37] ensure us that there is no way to choose apriori the best optimization algorithms for a particular problem, and the only option is to empirically test multiple approaches verifying which is actually the best one. Nonetheless, in this case, we decided to apply the GA [21], supported by the other scholars' findings and results as reported in []. Moreover, to the author's best knowledge, and based on the recent literature [20], this optimization algorithm reasonably covers the most important approaches to solving the optimization problem of Equation (3). Since the convergence of all these algorithms is influenced by the starting point, we employed a multi-start strategy [24]. In particular, as starting point, we used: (i) the initial geometry described in Table 1 and (ii) 100 random points uniformly distributed in the domain induced by the linear constraints of the optimization problem of Equation (3). The optimization methods have been implemented using the Matlab 2022a[1] environment. Table 2 summarizes the parameter setting of selected algorithms.

Table 2: Parameters setting for the optimization algorithm.

Algorithm	**Matlab Function**	**Parameter**	**Value(s)**
GA	ga	Population size	5000
		Elite count	250
		Crossover Fraction	0.8
		Pareto Fraction	0.35
		Mutation function	Uniform mutation
		Crossover function	Scattered crossover

4.2 Non-Deterministic Optimization Problem

This Section presents a methodological approach to address marine propellers' non-deterministic optimization design problem. The objective is to optimize propeller performance while considering uncertainties in design variables. The methodology combines stochastic optimization techniques, surrogate

[1] `https://www.mathworks.com/`

modelling, and robust design methods to find an optimal propeller design that is efficient, minimizing cavitation occurrence, and resilient to uncertainties. In particular, we are interested in maximising the cost function $J(\mathbf{X})$, which is a measure of propeller performance (e.g., efficiency and cavitation) that is to be optimized, requiring cavitation and efficiency variances to be minimal

$$\begin{aligned} \max_{\mathbf{X}\in\mathbb{R}} \quad & \int_{v_s}\int_{n}[\eta_0(\mathbf{X}) - (1-\lambda_1)A_{cav,face}(\mathbf{X}) - (1-\lambda_2)A_{cav,back}(\mathbf{X})]P(v_s)P(n)dv_s dn \\ \text{subject to} \quad & \begin{cases} K_T(\mathbf{X}) \geq K_{T,min} \\ K_T(\mathbf{X}) \leq K_{T,max} \end{cases}, \end{aligned} \tag{4}$$

where $\mathbf{X}$ represents the vector of deterministic design variables (i.e., the features reported in Table 1). $\mathbf{\Theta} = \{v_s, n\}$ represent the uncertain parameters considered in our non-deterministic optimization problem, following known probability distributions, $P(\mathbf{\Theta})$, which have been determined based on the author's experience. In particular for this problem, $P(v_s) = (\mu_{v_s}, \sigma_{v_s})$ and $P(n) = (\mu_n, \sigma_n)$, where $\mu_{v_s} = 25$ kn, $\mu_n = 200$ RPM, $\sigma_{v_s} = 0.02$ kn, and $\sigma_{v_s} = 4$ RPM.

The non-deterministic optimization problem presented in Equation 4 demands computationally efficient approaches, as using high-fidelity methods such as Reynolds-averaged Navier-Stokes equations, large-eddy simulations, or direct numerical simulations would be prohibitively expensive and time-consuming. Furthermore, even employing mid-fidelity techniques, like the BEM method, is still computationally demanding due to the presence of uncertainties in the design process. To overcome these challenges, we utilize surrogate DDSs developed in Section 3. These surrogate models offer rapid and accurate approximations of the objective function, significantly enhancing the efficiency of the optimization process. They reduce the computational load associated with evaluating the various components (i.e., η_o, $A_{cav,face}$, $A_{cav,back}$) of $J(\mathbf{X})$ while accounting for uncertainties.

Incorporating the uncertainties into the DDSs is achieved by constructing a probabilistic representation of the objective function. This approach enables a robust optimization process that takes into account the possible variations in design parameters and their potential effects on propeller performance. Additionally, the DDSs can approximate multiple objective functions simultaneously, facilitating efficient design space exploration and identification of Pareto-optimal solutions.

The optimization problem in Equation 4 is characterized by a non-linear and non-convex objective function along with a set of non-linear constraints. Consequently, we apply the same solution strategy discussed in Section 4.1 to address this problem effectively.

5. EXPERIMENTAL RESULTS

In this section, we will report the results of applying the methodology described in Sections 3 and 4 to solve the problem described in this work using the data described in Section 2. Specifically, in Section 5.1, we tested the quality of the DDSs in predicting the 5 target features reported in Table 1, in Section 5.2 we focused on classic optimization problem results, testing the DDS model, and in Section 5.3, we report the results of the non-deterministic optimization framework. For both optimisation problems, a sub-set of the best candidates has been re-tested with the BEM code to asses the functioning point within the functioning space (a regular distribution is considered made of 5x5 conditions) as for the reference propeller. This can give an equal description of the overall quality of the new designs in a real scenario.

5.1 Surrogate Accuracy

In this section, we will evaluate the performance of the DDSs utilizing the validation techniques outlined in Section 3. Specifically, we will examine the accuracy of the SVR algorithm used for constructing the DDSs by employing the MAE for error quantification and the AT for assessing computational requirements. We present the performance results using both quantitative and qualitative measures. Table 3 provides the metrics employed for evaluating the algorithm's performance across various target features in constructing the surrogate model. Additionally, scatter plots illustrating these relationships are depicted in Figures 4 and 5. By examining Table 3 and Figures 4 and 5, we observe that the DDSs exhibit excellent performance in terms of both accuracies (with an error lower than 4% for all the features) and time requirement (approximately $2 \cdot 10^{-2}$ [s] for all target features). This performance makes these surrogates highly suitable for integration within an optimization framework, as discussed in Section 5.3. Notably, to achieve this level of accuracy, a mid-fidelity BEM simulation is typically required. However, such simulations demand significantly higher computational time (approximately 20 seconds for each evaluation).

Table 3: Surrogate Models Validation: metrics employed to evaluate the performance (MAE and AT) for the algorithm employed to build the DDSs

Target Feature	**MAE**	**Average Time** 10^{-2} **[s]**
$A_{cav,back}$	0.1156 ± 0.0365	2.05
$A_{cav,front}$	0.01355 ± 0.0113	2.12
η_o	0.0049 ± 0.0163	1.93
K_Q	0.0021 ± 0.0112	1.84
K_T	0.0013 ± 0.0125	1.88

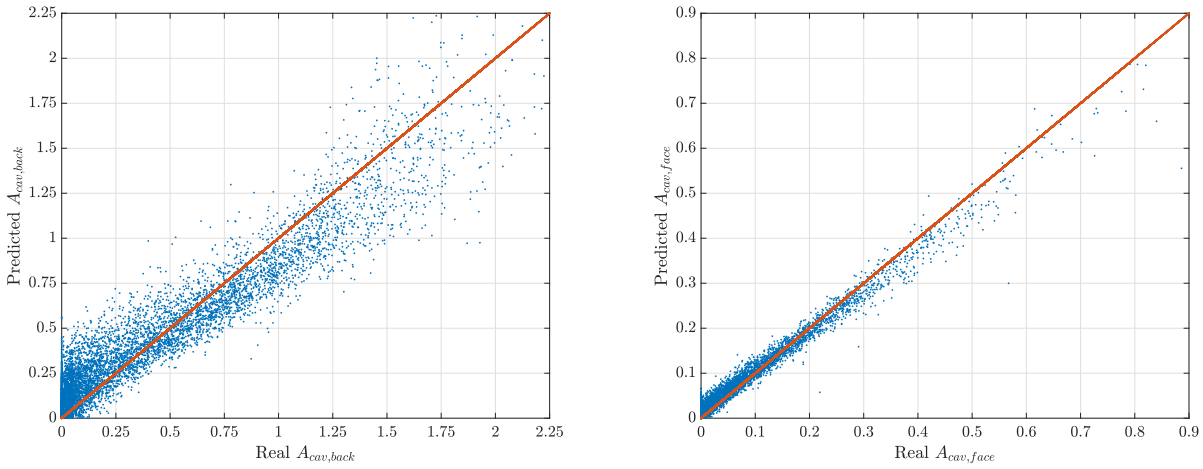

Figure 4: DDSs validation: suction-side cavitation area $A_{cav,back}$ (left) and face-side cavitation area $A_{cav,face}$ (right) scatter plots.

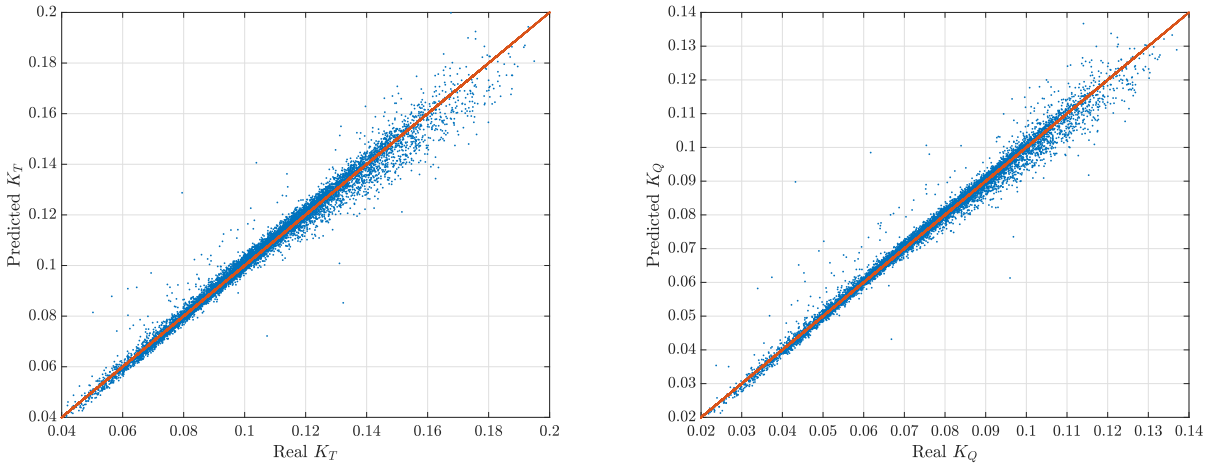

Figure 5: DDSs validation: torque coefficient K_T (left) and thrust coefficient K_Q (right) scatter plots.

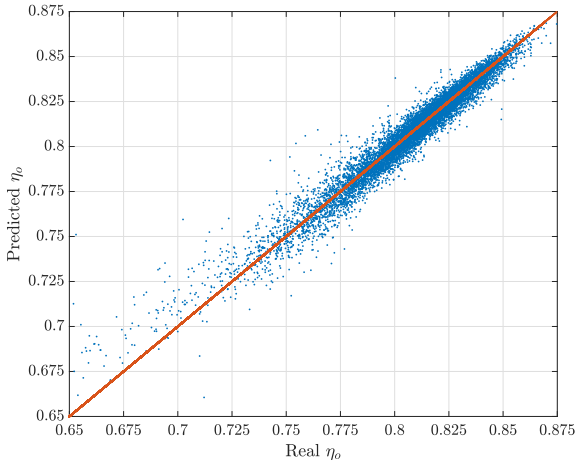

Figure 6: DDSs validation: efficiency η_o scatter plot.

5.2 Classic Optimization Problem Results

The computational equivalence between the BEM-based and surrogate-based optimization solvers is primarily attributable to the training phase needed for the DDSs. Nonetheless, in the context of robust optimisation, a BEM-based approach becomes infeasible. This infeasibility arises because each design necessitates multiple evaluations under slightly varying operating conditions to determine the objective function distribution. Given these considerations, the conventional design is derived using the BEM-based procedure, while, conversely, the robust design employs a surrogate-based approach. Figure 7 illustrates the optimization results obtained by means of the conventional approach, displaying efficiency versus the extension area of suction-side cavitation evaluated at the design functioning only. The data points are colour-coded based on the extension area of pressure-side cavitation, with the red dot representing the reference propeller. The optimization process identifies propeller configurations capable of simultaneous improvements of the two contrasting objectives of the design, which is a very interesting result given the performances of the reference propeller that exhibits an already significantly reduced extension of the cavitation at the design point (limited only to the outermost

strip of panels at the tip of the blade) at an interesting level of efficiency. At constant suction side cavitation area extensions, there are geometries ensuring an efficiency increase of up to 2.5%. Keeping constant the delivered efficiency, cavitation extension can be completely nullified. In this range of improvements, two optimal candidates emerge, C-29655 and C-61557. The former exhibits exceptionally high efficiency, close to 74%, with a very limited suction-side cavitation area, making it an attractive choice if higher efficiencies are preferred at the cost of a small amount of cavitation. Considering the baseline propeller geometry, in correspondence with this design functioning condition, has an open water efficiency of 0.7255, the improvement in efficiency is about 1.7% with a cavitation extension more than halved. It is worth noting from the predicted cavitation extension at the design point of Figure 14 that the cavitating behaviour of the selected propeller is completely different compared to the reference geometry. The reduction of cavitation is achieved by the unloading of the tip (not permitting further increases of the efficiency), which turns into the complete avoidance of tip cavitation at the cost of an almost negligible leading edge sheet cavitation at midspan due to the redistribution of the blade load to satisfy the delivered thrust constraint. Conversely, the latter candidate maintains the efficiency of the reference propeller while entirely avoiding cavitation on the blade at the design point. In both cases, pressure-side cavitation is absent. Only a few designs located in the lower portion of the graph exhibit, indeed, substantial pressure-side cavitation extension. This phenomenon typically occurs when the blade section experiences a negative angle of attack. However, it seldom arises due to the lightly loaded design condition and moderate wake disturbances. Since pressure-side cavitation also diminishes efficiency, which is one of the design objectives, the optimization process infrequently identifies such geometries that, in any case, are discarded during the selection process of the optimal candidates. The choice of these two candidates has twofold reasons. While it demonstrates the capabilities of the SBDO in a usual design scenario, it will also point out the limitations, in a real functioning environment, of propeller geometries obtained by a pure deterministic design approach also when they are selected to favour the complete avoidance of undesired phenomena like cavitation at the cost of no improvements in terms of efficiency.

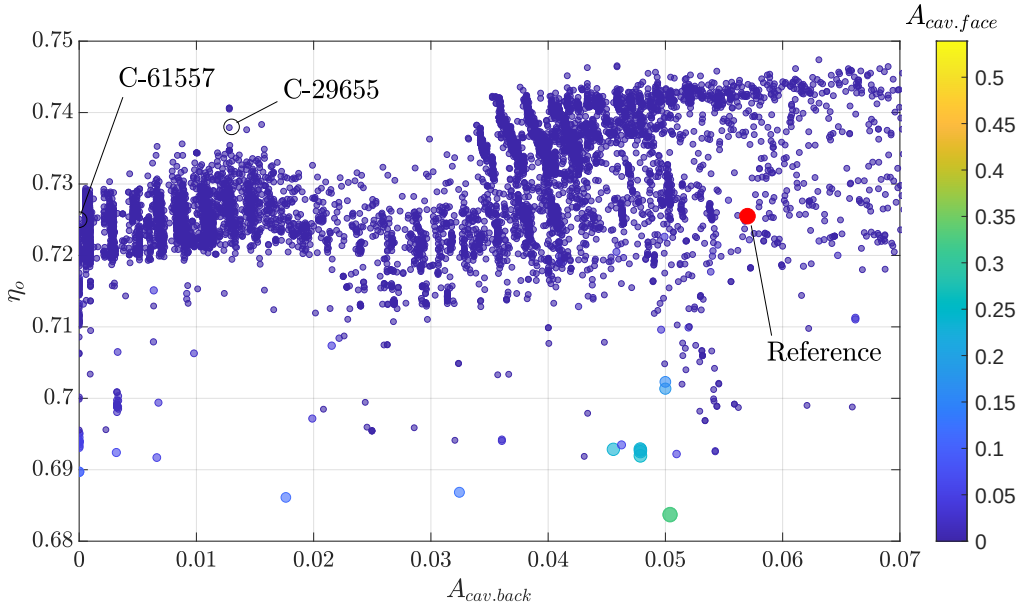

Figure 7: Conventional optimization history using BEM.

5.3 Non-deterministic Optimization Problem Results

Figure 8 illustrates the robust optimization results obtained by using the DDS in place of truly BEM calculations throughout the design process. The size of each data point represents the standard deviation of propeller efficiency, while the colours indicate the standard deviation of suction-side cavitation. Pressure side cavitation, not represented in this 4-dimensions diagram, sometimes occurs at the design point, especially in off-design conditions. The optimal geometry is always chosen by discarding all the propellers affected by this side effect.

Compared to the classic optimization problem, this formulation, which also takes into account variations in input parameters (i.e., propeller working conditions), tends to yield designs with lower efficiency. While designs with efficiencies exceeding 2% at the design point were identified in the previous strategy, none surpass the 74% threshold in the robust optimization, not even when cavitation extensions larger than the reference propeller are accepted. Nevertheless, many designs exhibit efficiencies higher than the reference propeller. As expected, the best design improves efficiency compared to the reference while simultaneously reducing cavitation extension, but based solely on the findings presented in Figure 8, it appears that the robust optimization problem yields geometries that are less overall effective compared to those from the conventional (truly BEM based) optimization problem. This is undoubtedly true when looking at the propeller performances calculated with the DDS at the design condition since the influence (and the performances) of off-design functioning are not immediately appreciable from this diagram. Alongside the reference design, marked with a red dot, one of the best candidates, R-28709, is highlighted. This choice is based on values of sensitivity to modifications of functioning conditions, measured through the variance of delivered efficiency and cavitation area, well reduced with respect to the reference propeller. We accepted a non-excessively reduced value of suction side cavitation considering the slightly higher uncertainty of DDS associated with this performance parameter, especially for very low predicted values (Figure 4) and further verified on a series of preliminary calculations.

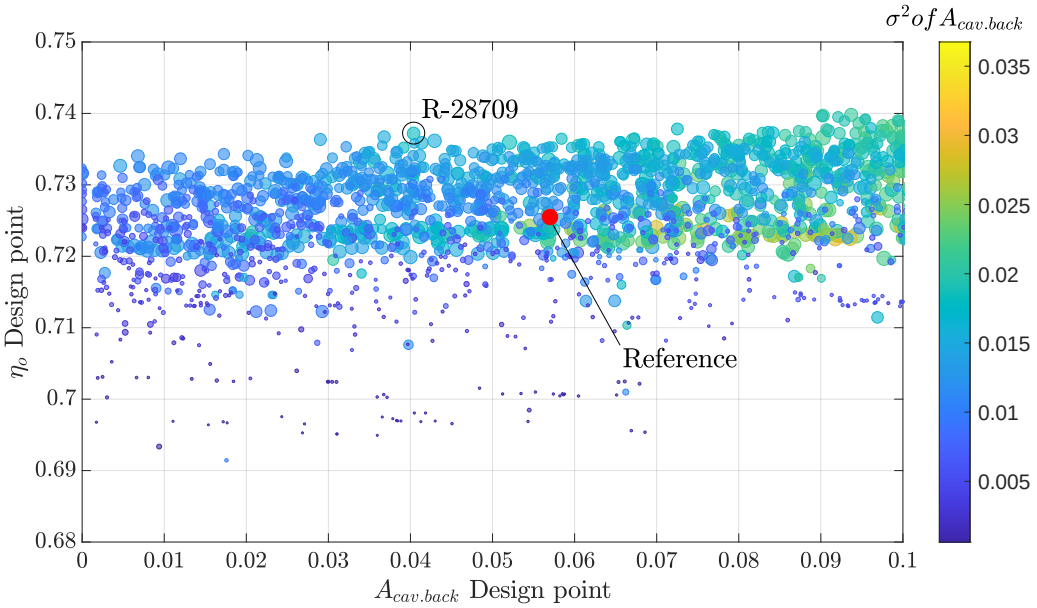

Figure 8: Robust optimization using the DDS. Bubble color identifies the standard deviation of the suction side cavitation area. Bubble size identifies the standard deviation of the propeller efficiency.

The improvements provided by the non-deterministic optimization are discussed in comparison to the reference propeller that has been specifically analyzed in the same functioning scenario assumed for the robust process. Also, the two optimal configurations from the conventional process are included to highlight the risk of designs for nominal conditions only specifically. Quantitative comparisons with C-61557 are given in Figures 9 and 11. Figures 10 and 12 account, instead, for geometry C-29655 from the conventional design process. Efficiency and predicted cavitation area, recomputed with the BEM (then not those from the DDS), in both cases, are compared at the ship speeds encompassing the uncertainty of working conditions as a function of the cavitation index (i.e., of the propeller rate of revolution). Truly BEM calculations at the design point immediately highlight some differences between the calculated (BEM) and the predicted (using DDS) value of efficiency of propeller R-28709, which results slightly lower (from 0.736 using DDS to 0.727 using BEM) but higher than the reference and in line with the uncertainties associated to the DDSs model. They also disclose the limitations of a deterministic design, especially looking at the efficiency of C-29655, which is a clear example of very high sensitivity to perturbations of design parameters. At nominal functioning, this propeller exhibits the highest efficiency among the selected geometries. The rate of change of efficiency with respect to the functioning condition (a measure of the sensitivity to perturbation that robust optimization tries to minimize) is however the fastest and performances easily worsen in off-design even compared to those of the reference propeller. C-61557, which has a design efficiency slightly lower than the reference propeller as the cost of a complete cavitation avoidance at the same functioning point, evidence very similar limitations of a conventional design process when tested in off-design conditions. Compared to the reference propeller, the degradation of efficiency is faster, especially at lower cavitation indexes (i.e., lower propeller advance coefficient). On the contrary, the performances of the propeller obtained by the robust design are much more stable, as expected from non-deterministic methods. Although the computed increase in efficiency may appear small, it is consistently maintained across a wide range of functioning conditions. As a result, this geometry is clearly "overall" superior to conventional designs and is either slightly better or at least equivalent to the reference geometry.

The convenience of the robust optimization framework is crystal clear when cavitation is considered. Compared to the reference propeller, R-28709 shows a cavitation extension on the suction side always lower regardless of the functioning condition. At the design point, the improvement is not outstanding, as already observed in the Pareto diagram of Figure 8, and the propeller still shows a cavitating blade tip only slightly reduced (Figure 15) if compared to that of the reference propeller of Figure 2. This is the cost of dealing with uncertainties and of a balanced action of minimization of suction side cavitation at the design functioning and minimization of the variance of the predicted cavitation extension in off-design. Monitoring off-design cavitation and minimizing its variance lead to geometries capable of substantial reduction of the occurrence of cavitation at the very low cavitation indexes included in the analyses. In correspondence with the highly loaded conditions (ship speed lower than design, cavitation index lower than 0.85), the advantages of the robust optimization are extraordinary, and propeller R-28709 behaves far better than the reference and any conventionally optimized geometry. These latter, in particular, disclose the risks of a deterministic design. Negligible variations of the functioning point, easily ascribable to uncertainties of the ship working conditions, lead to disruptive consequences in terms of cavitation occurrence. Even C-61557, which is the propeller with the best performances in terms of cavitation at the design point (that indeed is completely nullified) and which would have been a reasonable choice for cavitation avoidance based only on the results and the Pareto frontier from the deterministic optimization, is entirely unsuitable to address with success off-design functioning since even a 3% reduction of the cavitation index determines cavitation extensions larger than the reference geometry without any appreciable advantage in terms of propulsive efficiency.

From a qualitative point of view, the calculated extensions of the sheet cavity bubble of Figures 13,

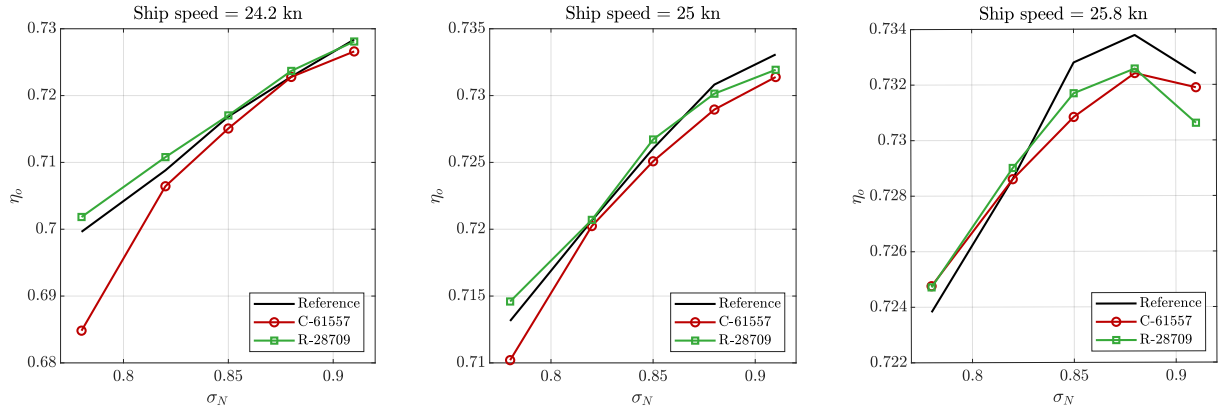

Figure 9: Comparison of propeller efficiency as a function of the cavitation index for three ship speeds. Robust propeller R-28709 compared to reference propeller and optimized propeller C-61557 from conventional optimization.

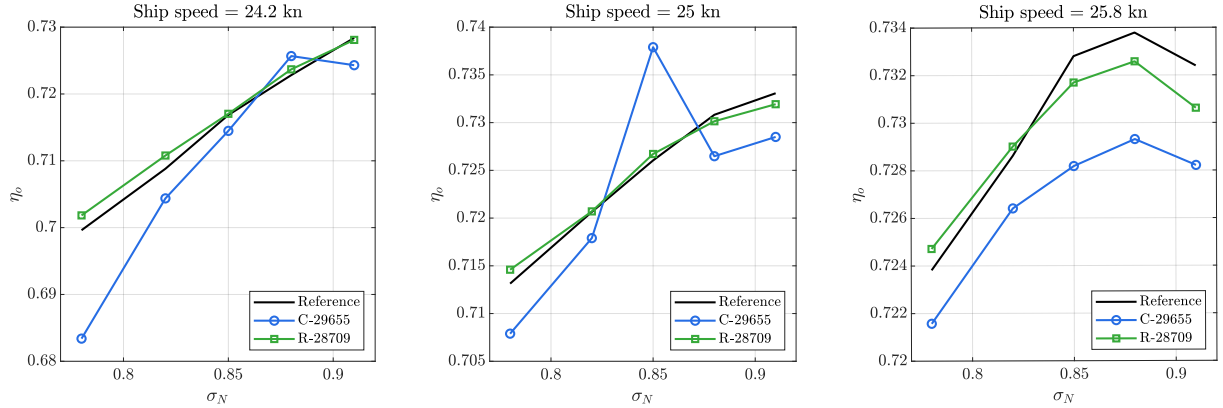

Figure 10: Comparison of propeller efficiency as a function of the cavitation index for three ship speeds. Robust propeller R-28709 compared to reference propeller and optimized propeller C-29655 from conventional optimization.

14 and 15 further highlight the convenience of a robust design process. The occurrence of cavitation, and the way this occurrence happens at different cavitation indexes, moreover, shed light on the blade shapes of the different optimization methods favoured. The geometry of the reference propeller is intrinsically robust against mid-chord bubble cavitation. The original pitch and chord distribution, together with the sectional hydrofoil shape, realize over the entire range of advance coefficients under investigation cavitation phenomena that always have the characteristics of tip vortex/sheet cavitation from the leading edge. This means a shock-free geometry with a certain tendency to suction peaks on the back side to ensure a certain margin against mid-chord bubble cavitation. Both propellers C-61557 and C-29655 have, instead, combined pitch/camber distributions designed to exploit the maximum efficiency at the design point. In terms of pressure distributions, this means almost constant, and extremely close to the vapour pressure, values of suction on the back of the propeller, that easily turn into cavitation over a larger portion of the blade when the cavitation index is even slightly reduced. At the lowest analyzed cavitation indexes, this explains the abrupt occurrence of cavitation observed for C-61557 and C-29655 that, based on the simplified assumption of the BEM employed for calculations, can be assumed of bubble type. As expected, R-28709 is much less sensitive to inflow conditions. Including the sensitivity to the variations of the functioning conditions as additional objectives of the

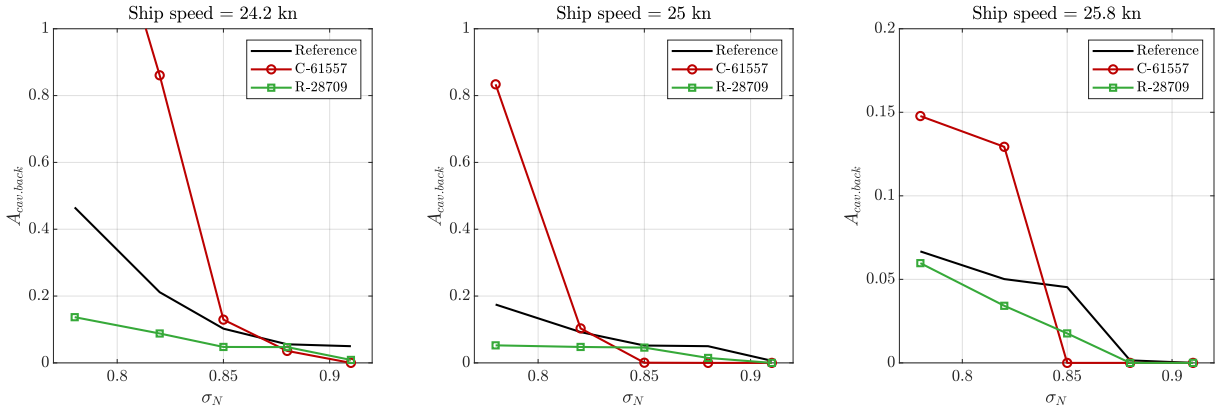

Figure 11: Comparison of suctions-side cavitation extension as a function of the cavitation index for three ship speeds. Robust propeller R-28709 compared to reference propeller and optimized propeller C-61557 from conventional optimization.

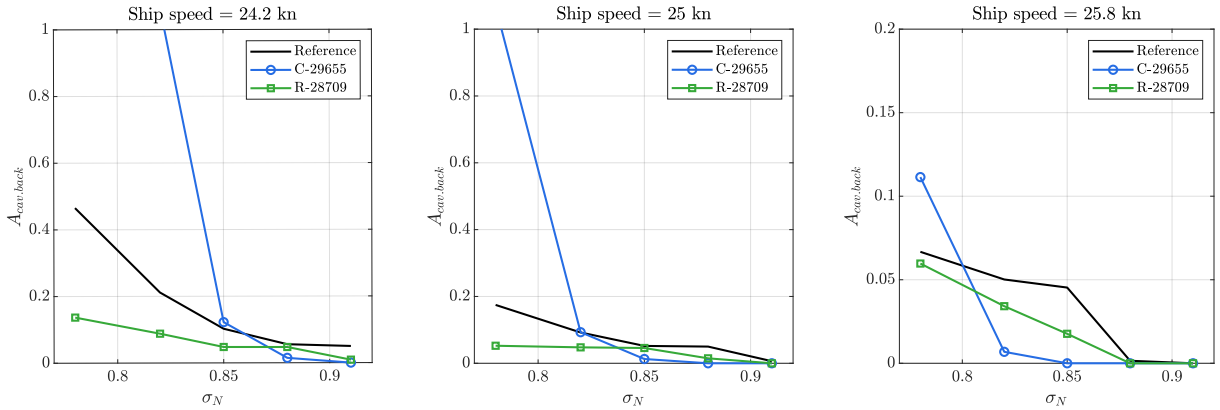

Figure 12: Comparison of suctions-side cavitation extension as a function of the cavitation index for three ship speeds. Robust propeller R-28709 compared to reference propeller and optimized propeller C-29655 from conventional optimization.

design process fostered the selection of geometries less prone to sudden variations of the performances. This results in a propeller having a slightly higher expanded area ratio needed for a more uniform and balanced pressure distribution on the blades at the cost of the negligible reduction in efficiency and in a pitch distribution more loaded at the tip (Figure 16). Tip cavitation is observable earlier for this geometry, but thanks to the blade shape, it is the only (and stabler) type of cavitation occurring for this geometry.

6. CONCLUSIONS

This paper presents a non-deterministic design optimization framework for the design of marine propellers under uncertain functioning conditions. Traditional deterministic design approaches can lead to blade geometries' excessive sensitivity to even slight variations in operational conditions, easily related to uncertainties about the ship displacement or to the influence of sea currents, that in the end, can nullify the effort of shape optimization. Accounting for the uncertainties of the design, however, would require unaffordable computational costs, even if using mid-fidelity codes like the Boundary Element Methods usually employed for propeller optimization, since hundreds of thousands of calcu-

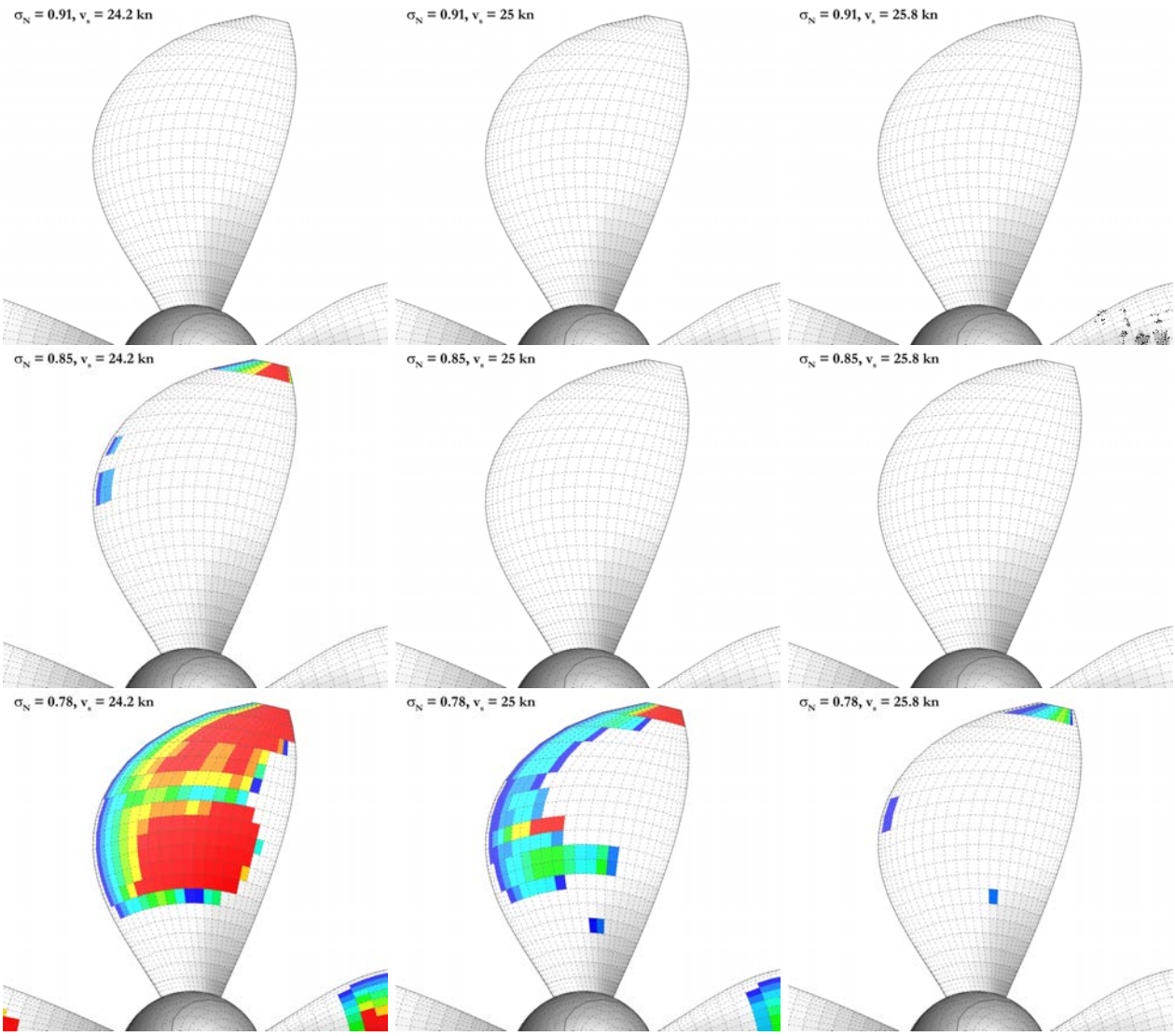

Figure 13: Cavitation extension on the blade-suction side. Optimized propeller $C-61557$ from the conventional optimization process. From top to bottom: $\sigma_N = 0.91, 0.85, 0.78$. From left to right: Ship speed of 24.2, 25 and 25.8 kn.

lations would be needed to evaluate the sensitivity to geometric uncertainty or, as in the present case, to variations of the functioning conditions. To overcome this "curse of dimensionality", we leveraged BEM-based Data-driven Surrogates models. They can be easily incorporated, and accurately substitute Boundary Element Methods calculations, into an optimization approach, to develop an efficient Simulation-Based Design Optimization framework for optimising marine propellers adhering to strict performance requirements and, simultaneously, considering the uncertainties in design variables. To this aim, we generated a dataset by running the Boundary Element Method multiple times on an extensive set of randomly sampled geometries in the given design space to feed the learning phase of Data-driven Surrogates and realize highly accurate surrogate models of the target performance indicators to be efficiently used for the reduction of the computational burden of the optimization process. The effectiveness of this approach strictly depends on the quality and quantity of data available for training the surrogate models. Results of Section 5.1 confirm this achievement. Then, we employed these Data-driven Surrogates in a non-deterministic optimization process. Their computa-

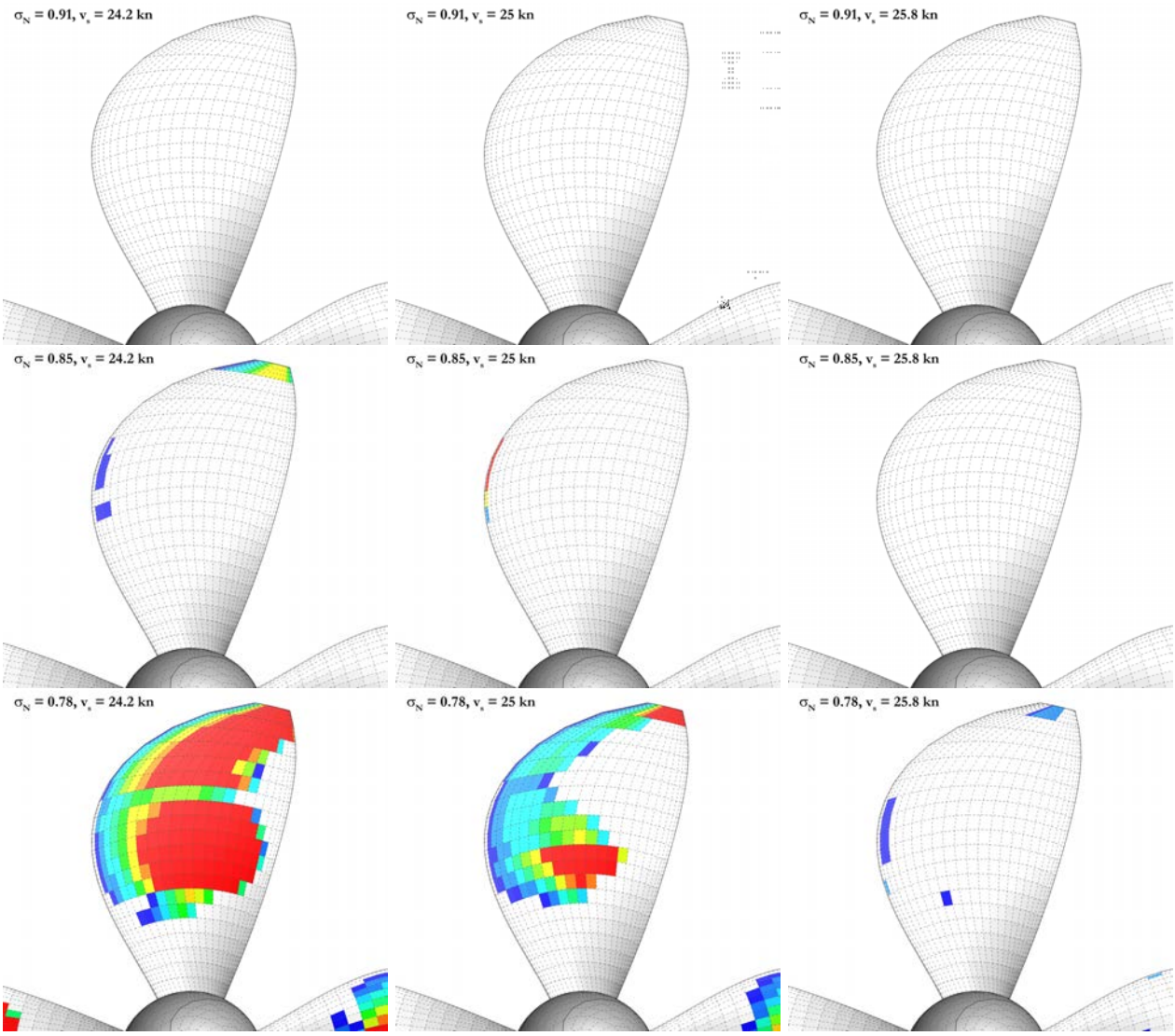

Figure 14: Cavitation extension on the blade suction side. Optimized propeller $C - 29655$ from the conventional optimization process. From top to bottom: $\sigma_N = 0.91, 0.85, 0.78$. From left to right: Ship speed of 24.2, 25 and 25.8 kn.

tional efficiency permitted the evaluation of the uncertainties of performances and, consequently, their inclusion, as additional objectives and constraints, into the iterative design process. Specifically, the variance of efficiency and suction side cavitation, as a consequence of given Gaussian distributions of inflow speed and rate of revolution centered in the nominal functioning, were minimized to reduce the rate of variations of these target features when the propellers operate in off-design conditions. We also solved a conventional, deterministic, design-by-optimization problem starting from the same reference propeller and considering the nominal functioning only. The comparison of the performances of the optimal geometry from the robust design process with those of the initial propeller and those of the geometries from the deterministic optimization framework unveils the convenience of the non-deterministic approach. Including performance uncertainties makes propeller geometries significantly less sensitive to functioning conditions. At the design functioning, the performances of the propeller from the robust optimization are, as expected, not as good as those of propellers optimized for nominal functioning only. The optimal robust design has a hint of tip vortex cavitation, and its efficiency is

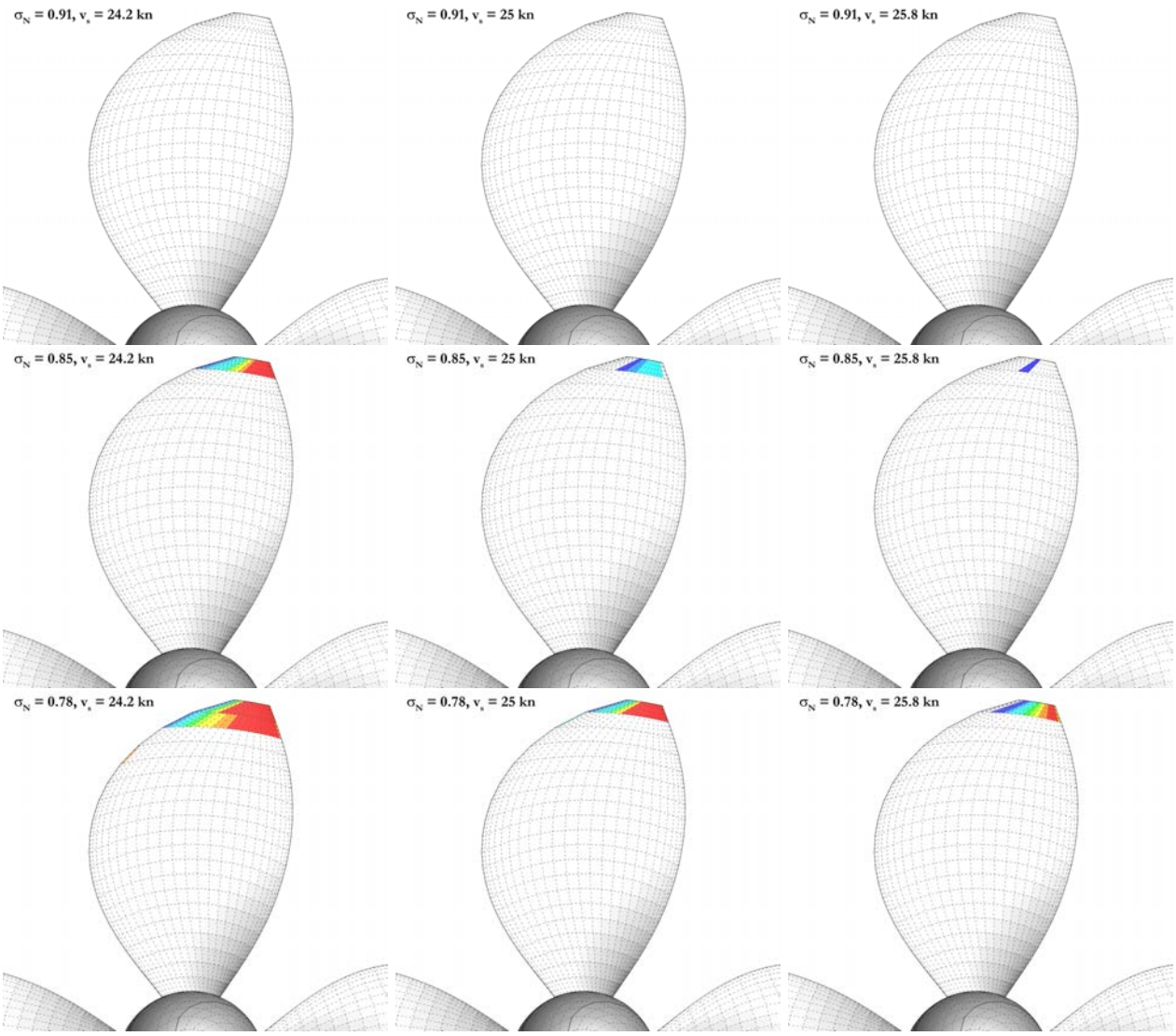

Figure 15: Cavitation extension on the blade suction side. Optimized propeller R-28709 from the robust optimization process. From top to bottom: $\sigma_N = 0.91, 0.85, 0.78$. From left to right: Ship speed of 24.2, 25 and 25.8 kn.

only slightly higher than the reference geometry, while the propellers from the deterministic optimization completely avoid cavitation, or ensure an efficiency gain close to 2%. When the entire range of functioning conditions is considered, the propellers designed under the robust assumption outperform those from the conventional optimization framework, showing the benefits of a design under uncertainties. Regardless of the functioning condition, the predicted cavitation extension on the suction side is always substantially lower than the reference propeller and shows a more stable behaviour (especially compared to conventionally optimized propellers) against variation of the advance coefficient and/or the cavitation index. Efficiency, too, is less affected by variations of the functioning point and remains higher than the reference on a relatively large set of functioning conditions. Overall, the non-deterministic optimization ensured better-balanced propellers, substantially less sensitive to input variations at the cost of marginal reductions of nominal performances compared to propellers optimized for nominal functioning only. This was possible thanks to Machine Learning methods that accurately evaluated the hundreds of thousands of samples needed for the uncertainty assessment.

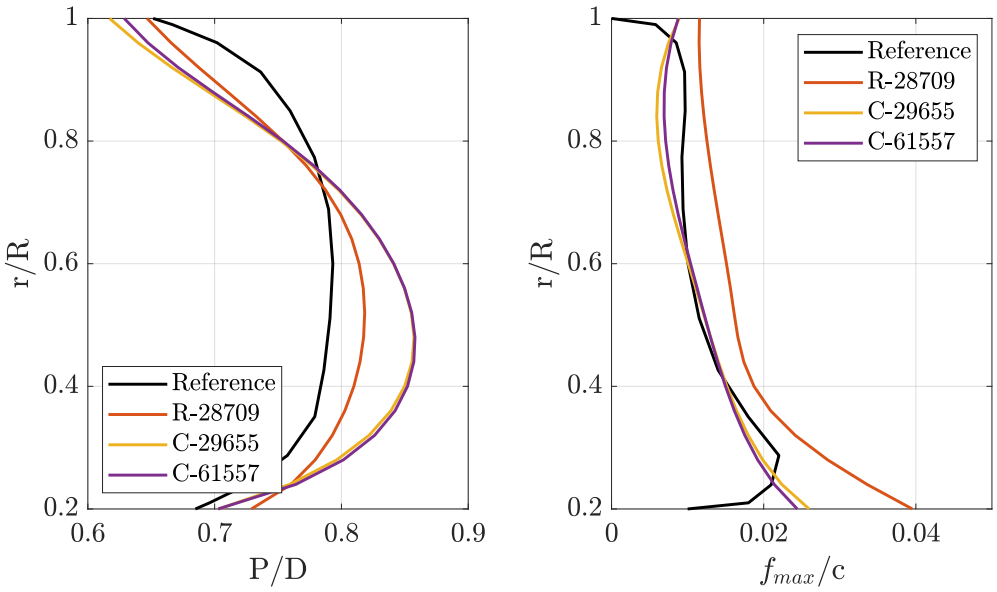

Figure 16: Pitch and maximum sectional camber radial distributions of the propeller under investigation.

ACKNOWLEDGEMENTS

This work has been partially supported by Spoke 3 "Waterways" and Spoke 10 "Logistics and Freight" within the Italian PNRR National Centre for Sustainable Mobility (MOST), CUP I53C22000720001, funded by the European Union - NextGenerationEU. Opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

[1] S. Berger, M. Druckenbrod, M. Pergande, and M. Abdel-Maksoud. A two-stage optimisation method for full-scale marine propellers working behind a ship. *Ship Technology Research*, 61(2):64 – 79, 2014.

[2] D. Bertetta, S. Brizzolara, S. Gaggero, M. Viviani, and L. Savio. Cpp propeller cavitation and noise optimization at different pitches with panel code and validation by cavitation tunnel measurements. *Ocean engineering*, 53:177–195, 2012.

[3] J. Carlton. *Marine propellers and propulsion*. Butterworth-Heinemann, 2018.

[4] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga. Handling multiple objectives with particle swarm optimization. *IEEE Transactions on evolutionary computation*, 8(3):256–279, 2004.

[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

[6] M. Emmerich and A. H. Deutz. A tutorial on multiobjective optimization: fundamentals and evolutionary methods. *Natural computing*, 17(3):585–609, 2018.

[7] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim. Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research*, 15(1):3133–3181, 2014.

[8] V. Ferrari, R. Tonelli, A.S. Kisjes, and R. Hallmann. Manoeuvring experiments, mathematical model and sensitivity analysis for test-case ferry. *Trends in Maritime Technology and Engineering Volume 1*, pages 327 – 335, 2022.

[9] S. Forrest. Genetic algorithms: principles of natural selection applied to computation. *Science*, 261(5123):872–878, 1993.

[10] S. Gaggero and M Martinelli. Design and analysis of pumpjet propulsors using cfd-based optimization. *Ocean Engineering*, 277:114304, 2023.

[11] S. Gaggero, G. Tani, D. Villa, M. Viviani, P. Ausonio, P. Travi, G. Bizzarri, and F. Serra. Application of multi-objective optimization based design to high-speed craft propellers. In *Proceedings of Fifth International Symposium on Marine Propulsors. Espoo, Finland*, volume 1, pages 278–290, 2017.

[12] S. Gaggero, G. Tani, D. Villa, M. Viviani, P. Ausonio, P. Travi, G. Bizzarri, and F. Serra. Efficient and multi-objective cavitating propeller optimization: An application to a high-speed craft. *Applied Ocean Research*, 64:31–57, 2017.

[13] S. Gaggero, G. Vernengo, and D. Villa. A marine propeller design method based on two-fidelity data levels. *Applied Ocean Research*, 123:103156, 2022.

[14] S. Gaggero, D. Villa, and M. Viviani. An investigation on the discrepancies between ranse and bem approaches for the prediction of marine propeller unsteady performances in strongly non-homogeneous wakes. In *International Conference on Offshore Mechanics and Arctic Engineering*, volume 45400. American Society of Mechanical Engineers, 2014.

[15] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep learning*. MIT press Cambridge, 2016.

[16] M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Mathematics of computation*, 78(268):2223–2257, 2009.

[17] I. Gypa, M. Jansson, R. Gustafsson, S. Werner, and R. Bensow. Controllable-pitch propeller design process for a wind-powered car-carrier optimising for total energy consumption. *Ocean Engineering*, 269:113426, 2023.

[18] I. Gypa, M. Jansson, K. Wolff, and R. Bensow. Propeller optimization by interactive genetic algorithms and machine learning. *Ship Technology Research*, 70(1):56–71, 2023.

[19] S. S. Keerthi and C. J. Lin. Asymptotic behaviors of support vector machines with gaussian kernel. *Neural computation*, 15(7):1667–1689, 2003.

[20] M. J. Kochenderfer and T. A. Wheeler. *Algorithms for decision making*. Mit Press, 2022. `https://algorithmsbook.com/`.

[21] O. Kramer. *Genetic algorithm essentials*. Springer, 2017.

[22] S. Li, A. Coraddu, and L. Oneto. Computationally aware estimation of ultimate strength reduction of stiffened panels caused by welding residual stress: From finite element to data-driven methods. *Engineering Structures*, 264:114423, 2022.

[23] S. Li, A. Coraddu, and L. Oneto. Computationally aware estimation of ultimate strength reduction of stiffened panels caused by welding residual stress: from finite element to data-driven methods. *Engineering Structures*, 264:114423, 2022.

[24] R. Martí. Multi-start methods. In *Handbook of Metaheuristics*, 2003.

[25] T. Melissaris, N. Bulten, and T. van Terwisga. On cavitation aggressiveness and cavitation erosion on marine propellers using a urans method. In *CAV2018: 10th International Symposium on Cavitation*, pages 838–843. ASME, 2018.

[26] L. Oneto. *Model Selection and Error Estimation in a Nutshell*. Springer, 2020.

[27] A. Rahman, R. Ullah, and M. Karim. Marine propeller design method based on lifting line theory and lifting surface correction factors. *Procedia engineering*, 194:174–181, 2017.

[28] B. Scholkopf. The kernel trick for distances. In *Advances in neural information processing systems*, pages 301–307, 2001.

[29] S. Shalev-Shwartz and S. Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[30] J. S. Shawe-Taylor and N. Cristianini. *Kernel methods for pattern analysis*. Cambridge university press, 2004.

[31] S. A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. In *Doklady Akademii Nauk*, volume 148, pages 1042–1045. Russian Academy of Sciences, 1963.

[32] J. M. Stubblefield. *Numerically-based ducted propeller design using vortex lattice lifting line theory*. PhD thesis, Massachusetts Institute of Technology, 2008.

[33] K. Vidal, L. Poppelier, B. Mallol, and C. Hirsch. The importance of a non-deterministic design optimization for predicting real life propeller performances. *Propellers & Impellers–Research, Design, Construction & Applications*, 2019.

[34] D. Villa, S. Gaggero, T. Gaggero, G. Tani, G. Vernengo, and M. Viviani. An efficient and robust approach to predict ship self-propulsion coefficients. *Applied Ocean Research*, 92:101862, 2019.

[35] M. Wainberg, B. Alipanahi, and B. J. Frey. Are random forests truly the best classifiers? *The Journal of Machine Learning Research*, 17(1):3837–3841, 2016.

[36] D. H. Wolpert. The supervised learning no-free-lunch theorems. *Soft computing and industry*, pages 25–42, 2002.

[37] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, 1(1):67–82, 1997.