Delft University of Technology

A machine learning approach for propeller design and optimization
Part II

Doijode, Pranav Sumanth; Hickel, Stefan; van Terwisga, Tom; Visser, Klaas

Important note
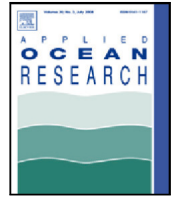To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A machine learning approach for propeller design and optimization: Part II

Pranav Sumanth Doijode [a],[*], Stefan Hickel [b], Tom van Terwisga [a],[c], Klaas Visser [a]

[a] *Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, The Netherlands*
[b] *Faculty of Aerospace Engineering, Delft University of Technology, The Netherlands*
[c] *Maritime Research Institute Netherlands (MARIN), Wageningen, The Netherlands*

ABSTRACT

We propose and analyse an optimization method that uses a machine learning approach to solve multi-objective, constrained propeller optimization problems. The method uses an online learning strategy where explainable supervised classifiers learn the location of the Pareto front and advise search strategies. The classifiers are trained with orthogonal features that capture geometric variation in radial distribution of pitch, skew, camber and chordlength. Based on orthogonal features, the classifiers predict whether or not a design lies on the Pareto front. If the design is predicted to lie on the Pareto front, the method verifies this with an evaluation. If the design is predicted to not lie on the Pareto front with a high confidence level, then the design is ignored. This skipped evaluation reduces the computational effort of optimization. The method is demonstrated on a cavitating, unsteady flow case of the Wageningen B-4 70 propeller with P/D = 1.0 operating in the Seiun-Maru wake. Compared to the classical Non-dominated Sorting Genetic Algorithm — III (NSGA-III) the optimization method is able to reduce 30% of evaluations per generation while reproducing a comparable Pareto front. Trade-offs between suction side, pressure side, tip-vortex cavitation and efficiency are identified from the Pareto front. The non-elitist NSGA-III search algorithm in conjunction with the explainable supervised classifiers also find very diverse solutions. Among the solutions, a design with no pressure side cavitation, low suction side cavitation and reasonable tip-vortex cavitation is found.

1. Introduction

Global trade is expected to increase the traffic density of shipping lanes on seas, oceans, and also rivers. Consequently, the Green House Gas (GHG), NOx, SOx, and noise emissions from shipping will increase making it more difficult to meet stricter emission regulations (International Maritime Organization, 2021) and noise reduction guidelines (International Maritime Organization, 2014). The International Maritime Organisation (IMO) aims to achieve a 50% reduction in GHG emissions by 2050 compared to 2008 levels (IMO, 2019) and ports have also started to incentivize silent ships with harbour due rate discounts (Port of Vancouver, 2017). Against this background, there is strong emphasis on making ship propellers more quiet and more efficient.

For vessels there are various sources of airborne and underwater noise (Abrahamsen, 2012). Compared to air, noise travels faster and farther in water. Underwater noise is generated by water flow on vessels, auxiliary machinery and equipment, diesel generators, prime movers, electric motors and propellers; however, cavitating propellers are expected to be the dominant source of underwater noise. Averson et al. (Arveson and Vendittis, 2000) report the underwater radiated noise for M/V Overseas Harriette at different ship speeds (propeller rpm). It was found that after cavitation inception, the noise hump between 50–100 Hz gets more pronounced with increasing speed. This noise hump is typically associated with the broad band spectrum of tip vortex noise (Pennings et al., 2016) which is often the dominant source of underwater radiated noise. Thus, one opportunity to make propellers quieter is by reducing noise radiation of tip-vortex cavities.

To improve efficiency and reduce tip-vortex cavitation noise, an optimal blade loading is required. Blade loading plays an important role in propeller design as blade sections are continuously exposed to different angles of attack in the non-uniform wake-field they operate in. For a rigid propeller, the blade loading distribution is optimized by modifying the geometry. Traditionally, in addition to efficiency objectives, indirect noise constraints as a limit on cavitation volume or area have been implemented (Brizzolara and Gaggero, 2009; Garg et al., 2017). Acoustic objectives have been considered but the focus has been limited to radiated sound power at blade passage frequencies (Mulcahy et al., 2014). The hydroacoustic design and optimization problem of propellers is of immense strategic importance and is most likely well researched in defence. However, there is limited knowledge

---

Nomenclature

Symbol	Description
$\bar{f}_*$	Posterior mean prediction for input x_*
ϵ	independent and identically distributed noise
η	Open-water efficiency, $J.K_t/2\pi K_q$
γ	Skew angle in radians
λ	Rake angle in radians
$\mathbb{V}$	Posterior variance prediction
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution
$\mathfrak{R}_{cart}$	A Cartesian frame of reference
$\mathfrak{R}_{cyl}$	A cylindrical frame of reference
$\mathfrak{R}_{orth}$	A reference frame defined by orthogonal vectors
$\mathbf{f}$	Latent function values, $\mathbf{f} = (f(x_1), \dots, f(x_N))^T$
$\phi(\vec{x})$	Feature map of input $\vec{x}$
ρ_∞	Free-stream density
Σ	Diagonal matrix of singular values
σ_n^2	Noise Variance
E[X]	Expectation of any matrix X
K	$n \times n$ Gram matrix
k_*	The vector $K(X, x_*)$
R[X, X]	Correlation coefficient matrix of any matrix X
θ	Pitch angle in radians
$\vec{\mu}, \vec{\mu}^*$	The mean vector for rows of M and M^*
$\vec{p}$	A point on a B-Spline surface
$\vec{r}_i$	Vector from origin to the mid-chord location
$\vec{s}$	A design vector of classical design variables
$\vec{x}$	The design vector of orthogonal parameters
C_f	Skin Friction Coefficient
D	Propeller diameter
f	Real valued function $f(x) : \mathbb{R}^N \to \mathbb{R}$
f_{obj}	Real valued objective function $f_{obj}(x) : \mathbb{R}^n \to \mathbb{R}$
g	A real valued constraint function $g(x) : \mathbb{R}^n \to \mathbb{R}$
J	Advance ratio, $v_\infty/\omega D$
k'	Number of orthonormal parameters
$k(x, x_*)$	Kernel function evaluation at inputs x and x_*
K_q	Torque Coefficient, $M_x/\rho_\infty v_{ref}^2 L_{ref}^3$
K_t	Thrust Coefficient, $T/\rho_\infty v_{ref}^2 L_{ref}^2$
M	Data Matrix of all propeller geometries in $\mathfrak{R}_{cyl}^{N' \times k}$
M^*	Parameter matrix for B-Spline discretization
M_x	Torque
N	Number of propellers in universal set
$N_{i,p}(t)$	A 3rd order pth degree basis function
p_{atm}	Absolute atmospheric pressure
S	The sample matrix of classical design variables
$s(\vec{x})$	The Silhouette Score of a vector $\vec{x}$
s_i	The ith classical design variable in $\vec{s}$
T	Thrust
t	Knots of the B-Spline surface
T_0	Target Thrust
U	Left singular vector matrix
u, v	B-Spline surface parameters
V	Right singular vector matrix
v_∞	Open-Water Speed
v_{cav}	Cavitation inception speed
v_{ship}	Ship speed
X	The parameter matrix
x_*	A new query vector
x_i	The ith orthogonal parameter in $\vec{x}$
y	A real valued range
y_c	Performance label for designs
k	Dimensionality

optimization of propellers being tractable to very few academic research groups. In North America, Yin Lu Young's group at Michigan University focuses on flexible composite propellers (Young et al., 2016) as a way towards designing quieter and more efficient propellers. One of the focus areas at the group is to design shape adaptive propellers, which have a bend twist coupling to optimize propeller loading (Plucinski et al., 2007; Motley et al., 2009; Liu and Young, 2009) and improve efficiency. Given the challenging structural modelling for composites (Maljaars and Kaminski, 2015), cavitation and noise constraints are secondary considerations. In Europe, MARIN has focused on developing and demonstrating tools for hydroacoustic design and optimization of propellers (Foeth, 2015; Huisman and Foeth, 2017). Rickard Bensow's group at Chalmers University has investigated algorithms, strategies and methods for hydrodynamic optimization of propellers with cavitation considerations (FlorianVesting, 2011; Vesting and Bensow, 2014; Vesting, 2015). In Asia, Nakashima propellers, for example, have focused on achieving better cavitation behaviour and improved efficiency for large vessels (Tendo et al., 2019). Better cavitation behaviour is expected to result in lower cavity volume fluctuations and thus lower pressure fluctuations and noise.

Up to approximately 2015, the focus in the aforementioned studies was predominantly on optimizing for efficiency and controlling sheet cavitation. In parallel, efforts were underway to better predict tip-vortex cavitation behaviour with simulations (Oprea, 2013; Gaggero et al., 2014; Liebrand et al., 2021). Thus, while propellers with better sheet cavitation behaviour were designed and introduced, the dominant source of noise was tending to be the cavitating tip-vortex. However, predicting tip-vortex cavitation on propellers continues to be too resource intensive for optimization. Recently, Bosschers .J (Bosschers, 2018a) proposed an Empirical Tip Vortex (ETV) model which offers one approach to hydroacoustic optimization of propellers with tip-vortex considerations. In addition to noise and efficiency design objectives, propeller design must also account for constraints on hull excitation (Bosschers, 2009), cavitation erosion (van Terwisga et al., 2007), ship speed and thrust. State of the art optimization methods are able to tackle this multi-objective constrained problem; however, there are three challenges which make them computationally expensive (Doijode et al., 2022):

The *first* challenge is that most optimization methods tend to approach the optimization problem with classical design variables. However, they could lead to multicollinearity (Allen, 1997) which makes it difficult to isolate the impact of changing individual design variables on objectives and constraints. By solving multicollinearity, there is an opportunity to reduce computational cost of optimization by reducing the number of objective/constraint evaluations to isolate the sensitivities of individual classical design variables.

and reporting in the public domain. Access to specialized simulation software, models based on propriety data and intellectual property restrictions (due to sensitivity or embargo) result in hydroacoustic

Consider a function

$$\phi = f(\vec{M}(\vec{s})), \tag{1}$$

where $\vec{M}$ represents the mesh and $\vec{s}$ is the design vector. By chain rule,

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial \vec{M}} \cdot \frac{\partial \vec{M}}{\partial \vec{s}} = J_1 \cdot J_2, \tag{2}$$

where J_1 is the Jacobian of the function with respect to the mesh and J_2 is the Jacobian of the mesh with respect to the design variables. The optimization problem is free of multicollinearity if both J_1 and J_2 are orthonormal matrices. The orthonormality of J_1 is determined by the flow physics and J_2 by the chosen design variables. Any design variable which affects multiple nodes on the mesh will lead to J_2 which is not orthonormal. This is because the blade surface is often defined with splines and changes in mesh nodes are strongly correlated. Different classical design variables also tend to change the same mesh nodes. Thus, it is difficult to isolate the impact of individual design variables on objectives and constraints.

The *second* challenge is that in the preliminary design phase, Boundary Element Methods (BEM) and Empirical Tip Vortex (ETV) model tend to have uncertainties (Vaz et al., 2015). These uncertainties results from numerical methods, domain discretization and physical models. In specific regions in the design space, these uncertainties could lead to the similar efficiency or cavitation performance being predicted by BEM and ETV for very different designs. Consequently, optimization methods could converge to solutions whose operational performance is significantly different from predicted performance.

The *third* challenge is that the computational cost of optimization for realistic cases is relatively high even in the preliminary design phase. Most commonly Artificial Neural Networks, Krigging, iKrigging, Cascading Neural Networks (Vesting and Bensow, 2014) have been used to reduce the cost of optimization. More recently Deep Learning (Goodfellow et al., 2016) has also gained traction in propeller design and optimization (Miglianti et al., 2020). In addition, explainable machine learning approaches (Abedjan et al., 2019) presents very interesting opportunities to provide or extrapolate performance guarantees, which is not possible with other approaches for propeller manufacturers. These developments are relatively new in the maritime industry. However, the models used in explainable machine learning such as Gaussian mixture models and Naive-Bayes are fundamental methods which are well known. Explainable machine learning can be used to predict where lucrative designs may be found in the design space *and* also provide insight into *why* the regions are lucrative.

To address the *first* challenge of multicollinearity, we propose an orthogonal feature set which has been demonstrated to capture geometric variations more accurately than classical design variables (Doijode et al., 2022). To address the *second* challenge and to account for uncertainties in preliminary design phases, performance predictions are assumed to have a mean and variance. Solutions dominate, i.e., lie on the Pareto front only when the mean is high and there is no overlap in the 95% confidence intervals of performance with compared solutions. This reduces the risk of false positives on whether a solution dominates or not. Consequently, Design and Optimization (D&O) strategies are expected to yield a range of solutions whose predicted performance does not differ significantly from operational performance. To address the *third* challenge of computational cost, soft explainable supervised classifiers with online training are used. The classifiers are taught the location of the Pareto front and, when sufficiently trained, they could potentially exclude designs which lie meaningfully away from the Pareto front, making the optimization method efficient.

The *first* challenge is that most optimization methods tend to approach the optimization problem with classical design variables. However, they could lead to multicollinearity (Allen, 1997) which makes it difficult to isolate the impact of changing individual design variables on objectives and constraints. By solving multicollinearity, there is an opportunity to reduce computational cost of optimization by reducing the number of objective/constraint evaluations to isolate the sensitivities of individual classical design variables. Consider a function

$$\phi = f(\vec{M}(\vec{s})), \tag{3}$$

where $\vec{M}$ represents the mesh and $\vec{s}$ is the design vector. By chain rule,

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial \vec{M}} \cdot \frac{\partial \vec{M}}{\partial \vec{s}} = J_1 \cdot J_2, \tag{4}$$

where J_1 is the Jacobian of the function with respect to the mesh and J_2 is the Jacobian of the mesh with respect to the design variables. The optimization problem is free of multicollinearity if both J_1 and J_2 are orthonormal matrices. The orthonormality of J_1 is determined by the flow physics and J_2 by the chosen design variables. Any design variable which affects multiple nodes on the mesh will lead to J_2 which is not orthonormal. This is because the blade surface is often defined with splines and changes in mesh nodes are strongly correlated. Different classical design variables also tend to change the same mesh nodes. Thus, it is difficult to isolate the impact of individual design variables on objectives and constraints.

The *second* challenge is that in the preliminary design phase, BEM and ETV model tend to have uncertainties (Vaz et al., 2015). These uncertainties results from numerical methods, domain discretization and physical models. In specific regions in the design space, these uncertainties could lead to the similar efficiency or cavitation performance being predicted by BEM and ETV for very different designs. Consequently, optimization methods could converge to solutions whose operational performance is significantly different from predicted performance.

The *third* challenge is that the computational cost of optimization for realistic cases is relatively high even in the preliminary design phase. Most commonly Artificial Neural Networks, Krigging, iKrigging, Cascading Neural Networks (Vesting and Bensow, 2014) have been used to reduce the cost of optimization. More recently Deep Learning (Goodfellow et al., 2016) has also gained traction in propeller design and optimization (Miglianti et al., 2020). In addition, explainable machine learning approaches (Abedjan et al., 2019) presents very interesting opportunities to provide or extrapolate performance guarantees, which is not possible with other approaches for propeller manufacturers. These developments are relatively new in the maritime industry. However, the models used in explainable machine learning such as Gaussian mixture models and Naive-Bayes are fundamental methods which are well known. Explainable machine learning can be used to predict where lucrative designs may be found in the design space *and* also provide insight into *why* the regions are lucrative.

To address the *first* challenge of multicollinearity, we propose an orthogonal feature set which has been demonstrated to capture geometric variations more accurately than classical design variables (Doijode et al., 2022). To address the *second* challenge and to account for uncertainties in preliminary design phases, performance predictions are assumed to have a mean and variance. Solutions dominate, i.e., lie on the Pareto front only when the mean is high and there is no overlap in the 95% confidence intervals of performance with compared solutions. This reduces the risk of false positives on whether a solution dominates or not. Consequently, D&O strategies are expected to yield a range of solutions whose predicted performance does not differ significantly from operational performance. To address the *third* challenge of computational cost, soft explainable supervised classifiers with online training are used. The classifiers are taught the location of the Pareto front and, when sufficiently trained, they could potentially exclude designs which lie meaningfully away from the Pareto front, making the optimization method efficient.

In this paper, we propose an optimization method to solve multi-objective constrained optimization problems based on the above machine learning strategies. The method referred to as DO is detailed in Section 2. It is demonstrated on a case with the starting point as a Wageningen B-4 70 propeller with P/D = 1.0 operating in the Seiun-Maru wakefield. Sections 2.1–2.6 detail the modules of DO. Section 3 compares the performance of classical NSGA-III optimization with DO. the results are summarized in Section 4.
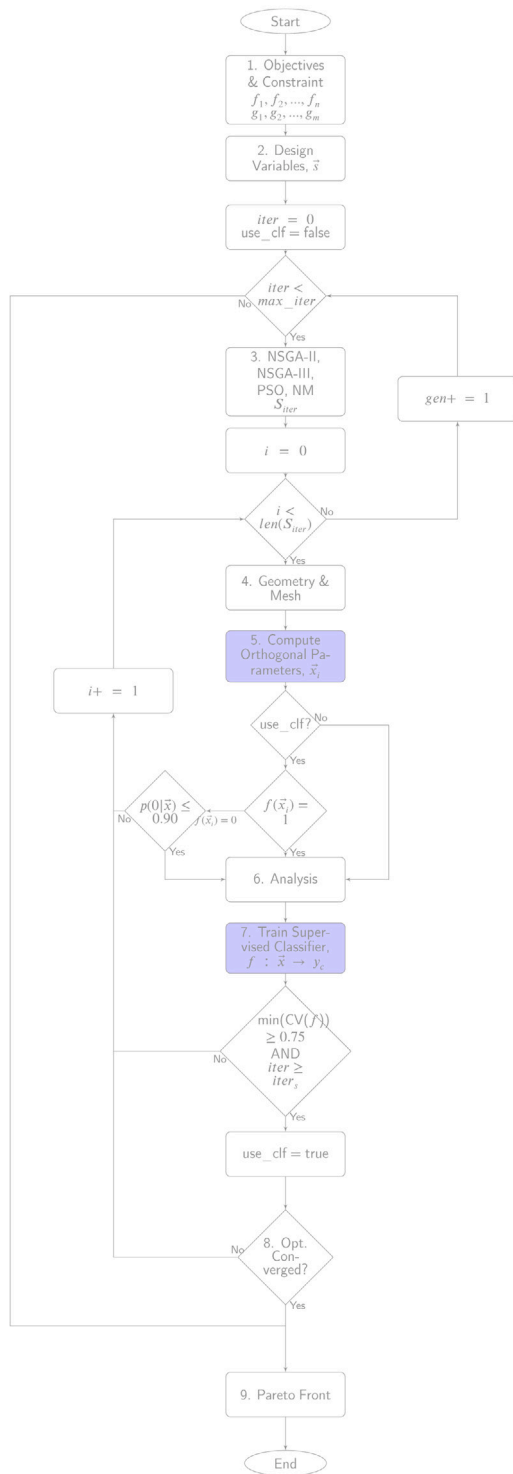
Fig. 1. Dynamic optimization flowchart.

Algorithm 1 Dynamic Optimization for multi objective applications

> **Input:** $f_1, f_2, ..., f_n,\ g_1, g_2, ..., g_m,$ max_iter, thresh_iter
> **Output:** S_{pareto}

1: **procedure** DYNAMIC OPTIMIZATION
2:     $S_{pareto}, S, F \leftarrow \emptyset$
3:     $i \leftarrow 0$
4:     use_clf $\leftarrow$ **false**
5:     **while** $i \leq$ max_iter **do**
6:         $S_i \leftarrow$ SEARCH(*params)
7:         $S \leftarrow S \cup S_i$
8:         **for** $j \leftarrow (0 \rightarrow len(S_i))$ **do**
9:             $\vec{s} \leftarrow S_{i,j}$
10:             Generate geometry and mesh $(\vec{M}^*)$ for $\vec{s}$
11:             $\vec{x}^* \leftarrow$ PROJECTION$(\Re_{ortho}, k', \vec{M}^*)$ ▷ See Algorithm 3
12:             **if** use_clf **then**
13:                 **if** $f(\vec{x}^*) = 1$ **AND** $p(\vec{x}^*|\mu_0, \sigma_0^2) \leq 0.90$ **then**   ▷ The classifier $f : \vec{\Re}_{ortho}^{k'} \rightarrow \Re$
14:                     Analyse design $\vec{s}$
15:                     $F \leftarrow F \cup \left[\begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}\right]$
16:                 **else**
17:                     continue   ▷ Skipped an evaluation, thus reduced cost of optimization
18:                 **end if**
19:             **else**
20:                 Analyse design $\vec{s}$
21:                 $F \leftarrow F \cup \left[\begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}\right]$
22:             **end if**
23:             Train Supervised Classifier, $f : \vec{\Re}_{ortho}^{k'} \rightarrow \Re$
24:             **if** min(10-k Cross Validation) ≤ 0.75 **AND** $i \geq$ thresh_iter **then**
25:                 use_clf $\leftarrow$ **true**
26:             **else**
27:                 use_clf $\leftarrow$ **false**
28:             **end if**
29:             **if** Converged **then**
30:                 break
31:             **end if**
32:         **end for**
33:     **end while**
34:     $\vec{y}_c \leftarrow LABELS(F, \epsilon)$ ▷ See Algorithm 4
35:     $S_{pareto} \leftarrow S_{pareto} \cup \{s_i | S[\vec{y}_c == 1]\}$
36:     **return** S_{pareto}
37: **end procedure**

Sorting Genetic Algorithm — II (NSGA-II), NSGA-III (Deb and Jain, 2013), Particle Swarm optimization (PSO), Nelder–Mead (NM) or others generate samples S_i. For each sample $\vec{s} \in S_i$, the geometries and meshes are generated. A set of orthogonal features $\vec{x}^* \in \Re^{k'}$ which are projections of the i-th mesh in a k' dimensional hyperspace is computed. Until there are sufficient samples to train and validate, the classifier is not used. When the classifier is not used, designs $\vec{s} \in S_i$ are analysed with Computational Fluid Dynamics (CFD) or BEM. After each design is analysed they are assigned a label. Interesting designs are labelled as $y_c = 1$ and clearly inferior designs are labelled $y_c = 0$. If the design lies on the Pareto front and has an efficiency greater than $a \times \eta_{max}$, where $a = 0.85$ is an input parameter. With a, it is possible to control the explore/exploit bias. A higher value of a makes the search more exploitative. For details on assigning labels see Algorithm 4.

After the first two samples are evaluated, the supervised classifier is trained with orthogonal parameters ($\vec{x}$) as features and target labels

2. Dynamic optimization

Dynamic Optimization is a method for multiobjective, constrained propeller optimization problems where search agents are advised by supervised classifiers. The method is detailed in Algorithm 1 and the flowchart in Fig. 1. As inputs, the method takes the objectives, constraints and iteration limits. When the optimization begins, for the ith iteration, an objective driven search strategy such as Non-dominated

Table 1
Objectives and Constraints.

Obj./Cons.	Description
max $f_1(\vec{s}) = \eta$	Maximize efficiency
min $f_2(\vec{s}) = V_{cav,b}$	Minimize suction surface cavity
min $f_2(\vec{s}) = V_{cav,f}$	Minimize pressure surface cavity
min $f_4(\vec{s}) = \frac{\partial^2 V_{cav}}{\partial^2 t}$	Minimize cavity volume acceleration
min $f_5(\vec{s}) = \max\,(r_c)$	Minimize tip vortex cavity radius
min $f_6(\vec{s}) = \max\,(|C_p|) = \max\left(\left\|\frac{p-p_{r,\infty}}{\frac{1}{2}\rho_\infty\left(v_{ship}^2+2\pi\omega r^2\right)}\right\|\right)$	Minimize peak pressures. ω is the propeller rotation rate in revolutions per second. $p_\infty = 1.01325e5$ Pa, $\rho_\infty = 1025$ kgm^{-3}.
$g_1(\vec{s}) = T - T_0 = 0$	Thrust constraint. Note that this is not changed during optimization.
$g_2(\vec{s}) = v_{ship} - v_{ship,0} = 0$	Ship speed constraint. Note that this is not changed during optimization.
$g_3(\vec{s}) = \max\left(\left\|\frac{s_{\theta,i}-s_{\theta,i+1}}{s_{\theta,i+1}}\right\|\right) - 0.3 \leq 0$	Fairing constraint for pitch
$g_4(\vec{s}) = \max\left(\left\|\frac{s_{\gamma,i}-s_{\gamma,i+1}}{s_{\gamma,i+1}}\right\|\right) - 0.3 \leq 0$	Fairing constraint for skew
$g_5(\vec{s}) = \max\left(\left\|\frac{s_{chord,i}-s_{chord,i+1}}{s_{chord,i+1}}\right\|\right) - 0.3 \leq 0$	Fairing constraint for chord-length
$g_6(\vec{s}) = \max\left(\left\|\frac{s_{cam,i}-s_{cam,i+1}}{s_{cam,i+1}}\right\|\right) - 0.3 \leq 0$	Fairing constraint for camber

($y_c \in [0,1]$). After at least 10 samples are evaluated, it is possible to validate the classifier with 10-fold cross-validation studies. In such a case, if the classifier has a mean 10-fold cross-validation score of 75% or higher the classifier is used. Weighted average precision scores are used for cross-validation. In iterations where the classifier is used, design labelled as $y_c = 1$ are analysed to confirm the prediction. Else, the design is not analysed when $p(\vec{x}|\mu_0,\sigma_0^2) \geq p_{th}(\vec{x}|\mu_0,\sigma_0^2)$, where $p_{th}(\vec{x}|\mu_0,\sigma_0^2)$ is the threshold for the probability that $y_c = 0$. A high threshold indicates a high confidence that the design is much further away from the Pareto front. Thus, by skipping analysis of designs which lie away from the Pareto front, computational cost of optimization is reduced. In our demonstration $p_{th}(\vec{x}|\mu_0,\sigma_0^2) = 0.9$. The method catches most false positives. However, there is a risk that false negatives are missed. This risk can be mitigated by increasing $p_{th}(\vec{x}|\mu_0,\sigma_0^2)$. If after an iteration or generation the convergence criteria are met, the optimizer reports the Pareto front and exits. Else, it continues to the next iteration or generation.

2.1. Objective and constraints

For demonstration of DO, we optimize the Wageningen B-4 70 propeller with P/D = 1.0 as the starting point. The blade is operating in the effective wake of the Seiun-Maru vessel with $v_{ship,0} = 10.80$ m/s and $T_0 = 577.4kN$ (Maljaars, 2019). The full scale measurements for flow characteristics around the hull and resulting propeller inflow and turbulence are reported by ITTC (Faresi, 1990). Particulars of the Seiun-Maru vessel are also provided in Kato and Kodama (2003).

The optimization problem has multiple objectives and constraints. Table 1 lists and describes the objectives and constraints. The goal is to maximize efficiency and minimize (a) cavity volume, (b) cavity volume acceleration, (c) vortex cavity radius and (d) peak pressure. The efficiency and cavitation behaviour are predicted with PROCAL (Vaz and Bosschers, 2006). The radius of the vortex cavity is predicted with the ETV-2 model (Bosschers, 2018b). In addition to thrust and ship-speed constraints, fairing constraints are specified to guide search towards solutions without high geometric variations along the radius. In order to match the thrust requirement, the propeller rotation rate is modified while the ship-speed is maintained a constant.

It is recognized that the skewed Seiun-Maru propeller is a good starting point for the demonstration. In the scope of this research, to demonstrate DO in an explore biased optimization, we prefer to start from a design much further away from a known local optimum. In

Table 2
Bounds of design variables.

Variable	Lower bound	Upper bound
$\Delta\theta$ (rad)	−0.36	036
$\Delta\gamma$ (rad)	−0.063	0.063
s_{camber} (rad)	−0.045	0.045
s_{chord} (−)	0.76	1.24
n (rps)	0	inf

the current demonstration, the baseline propeller is the Wageningen B-4 70 propeller with P/D = 1.0. The baseline performance is detailed in Section 3.1. This starting point implies a more challenging task for the optimization algorithm. At the same time, it increases the chances of finding new design solutions and search directions, compared to starting from an already more optimal design. The starting propeller geometry, effective wake and operating conditions are illustrated in Fig. 2.

2.2. Design variables

The blade is parametrized with four equispaced hydrofoil cross-sections at locations $[0, \frac{span}{3}, \frac{2span}{3}, span]$. Four parameter $[\theta, \gamma, s_{chord}, s_{camber}]$ are used as design variables to modify the cross-sections. They are intended to change pitch, skew, chordlength and camber respectively. While θ, γ are modified for all four cross-sections, s_{camber}, s_{chord} are modified for the last three cross-sections. The modifications are in the sequence θ, γ, s_{chord} and s_{camber}. The fifth design variable is the propeller rotation rate $n(rps)$.

The baseline geometry and the location of the hydrofoils is illustrated in Fig. 3. The bounds of the 14 design variables are listed in Table 2. The bounds result from the applicability range for PROCAL. The found solution lies away from the bounds (see Table 4b) thus, it is demonstrated that the applicability bounds of PROCAL are not restrictive.

The schemes for geometry modification are detailed below.

Parameter θ: Fig. 4(b) illustrates the change in geometry when θ is varied. The target cross-section is rotated about a vector parallel to z-axis passing through the midchord of the section. An increase in θ results in the blade experiencing a higher inflow angle thus also higher
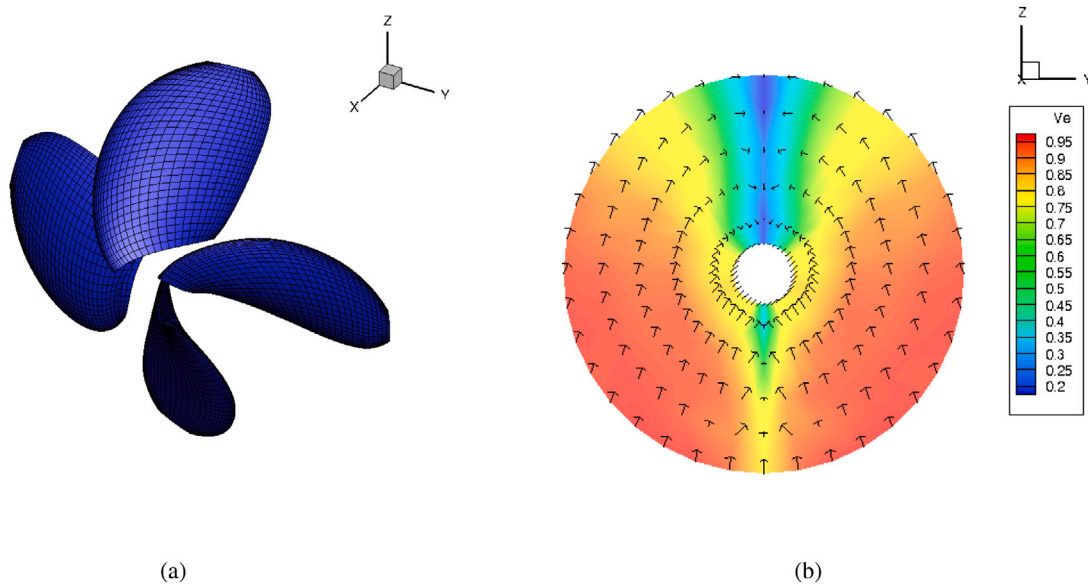
(a)

(b)

Fig. 2. (a) Wageningen B4-70 propeller base geometry and (b) Effective wakefield of Seiun-Maru vessel, $V_e = \sqrt{V_x^2 + V_y^2 + V_z^2}$. *The wakefield was provided by MARIN for this research and is available on request.*
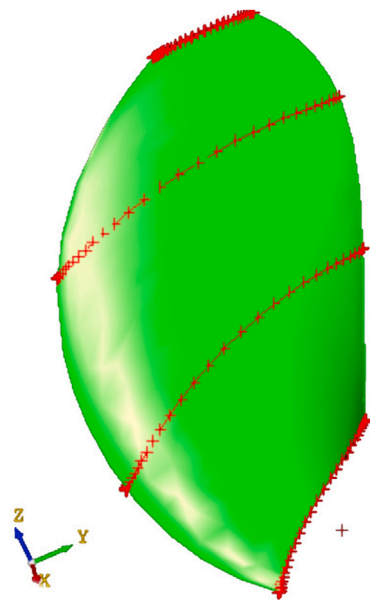

Fig. 3. Blade hydrofoils.

loading. Pitch is expected to be very sensitive to θ. It is also possible that θ influences the camber of the hydrofoil.

Parameter γ: Fig. 5 illustrates the change in geometry when γ is varied. The target cross-section is rotated at the mid-chord vector about the x-axis. Skew is expected to be very sensitive to γ. The proposed parameter helps control the rate at which loading on the blade increases in the wake-field. By doing so, it helps regulate cavity volume acceleration.

Parameter λ: Fig. 6 illustrates the change in geometry when λ is varied. The target cross-section is modified by rotating Mid-chord vector ($\vec{r}_i$) about the y-axis. Rake is expected to be very sensitive to λ.

Camber: The schematic for modifying hydrofoil camber is illustrated in Fig. 7. Camber is modified by rotating each of the coordinates defining

the hydrofoil about the x-axis. The angle of rotation is

$$\phi = \frac{\phi'^2 \sin(\phi') s_{camber,i}}{c},$$ (5)

where $c = 3.9453$ is a smoothing constant that can be defined by the designer and $\phi' = \frac{x - x_{te}}{x_{le} - x_{te}} \cdot \pi$. Note that the camber modification does not influence the leading edge or trailing edge thus, pitch is not expected to be impacted by changes in s_{camber}. The coordinates of the cross-section are transformed as

$$\vec{p}' = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & -sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix}.$$ (6)

Increasing camber for the cross-section is expected to increase loading while also shifting the centre of pressure for the blade. For flexible propellers, increased loading could introduce bending forces on the blade. Furthermore, the shift in centre of pressure could lead to a twisting moment about the inertial axis.

Chord-length: The schematic for modifying hydrofoil chord-length is illustrated in Fig. 8. Chord-length is modified in two steps. The first step, illustrated in Fig. 8(a), involves scaling the hydrofoil along the mid-chord vector by s_{chord}. The coordinates of the cross-section are transformed as

$$\vec{p}' = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} s_{chord} & 0 & 0 \\ 0 & s_{chord} & 0 \\ 0 & 0 & s_{chord} \end{bmatrix}.$$ (7)

The second step, illustrated in Fig. 8(b), involves translating it back to the initial radial location as

$$\vec{p}' = \begin{bmatrix} 0 & 0 & 0 & \delta x \\ 0 & 0 & 0 & \delta y \\ 0 & 0 & 0 & \delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix},$$ (8)

where $[\delta x, \delta y, \delta z] = \vec{r} - \vec{r}_{scaled}$.

One limitation of the above proposed parameters is that the pitch, skew and rake are indirectly influenced. Furthermore, the sequence in which the parameters are varied could determine whether skew and rake are preserved or not. It is also possible that modifying θ could impact the camber. However, it is to be noted that even with very
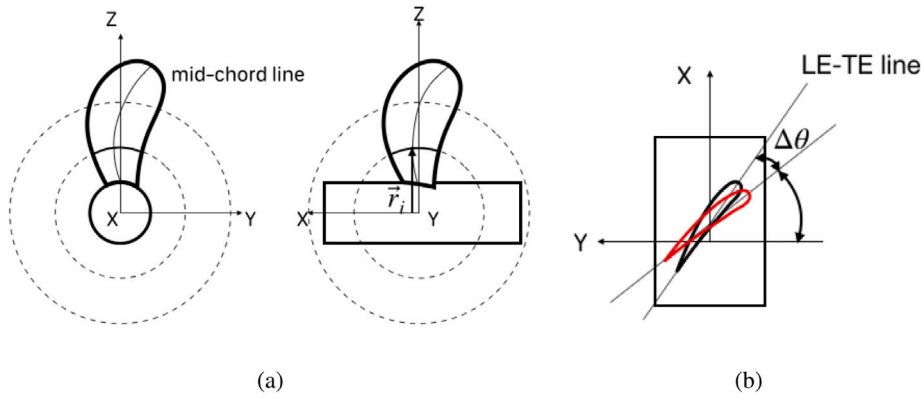
Fig. 4. (a) Mid-chord vector ($\vec{r}_i$), (b) change in geometry resulting from $\Delta\theta$.
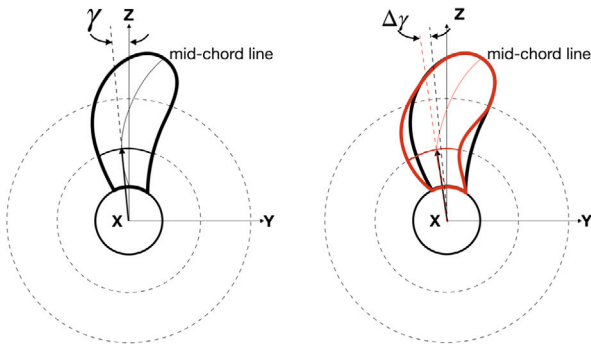

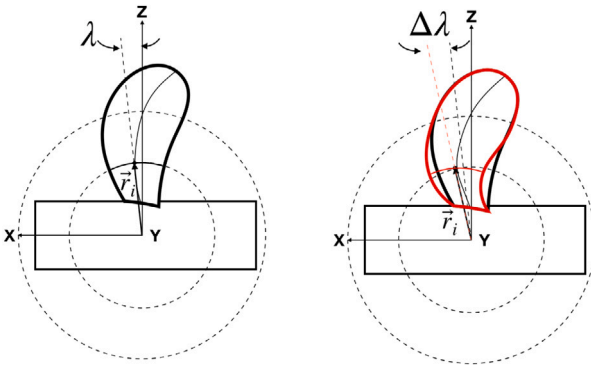**Fig. 5.** Change in geometry resulting from $\Delta\gamma$.

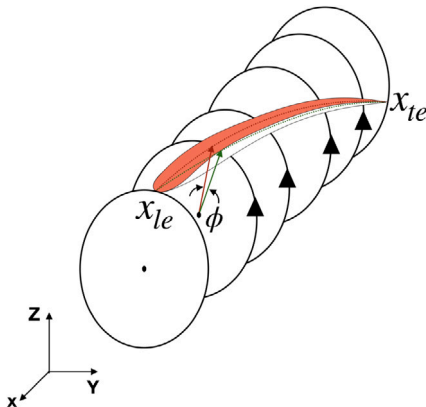**Fig. 6.** Change in geometry resulting from $\Delta\lambda$.

Fig. 7. Schematic illustrating camber modification.

carefully considered and more robust alternatives, J_2 in Eq. (4) will not be orthonormal if varying one parameter influences multiple mesh nodes on the discretized blade geometry. Consequently the parameterization is not orthogonally independent. This problem will be tackled in Section 2.5, where we derive a set of orthogonal parameters.

2.3. NSGA-III

For the current multi-objective constrained optimization problem we choose NSGA-III, as it is an established method on 3 to 15 objective optimization problems with convex, concave, disjointed and differently scaled Pareto fronts (Deb and Jain, 2013). The framework of NSGA-III is similar to that of NSGA-II with modifications to mating selection and survival (Blank et al., 2019).

Algorithm 2 NSGA-III survival selection

> **Input:** $R^t, N, Z, \hat{z}^*, \hat{z}^{nadir}$
> **Output:** $P^{(t+1)}$
> **procedure** SURVIVORS
> $(F_1, F_2, \ldots) \leftarrow non\_dominated\_sort(R^t)$
> $S = \emptyset, i = 1$
>   **while** $|S| + |F_i| \leq N$ **do**
> $S \leftarrow S \cup F_i$
> $i = i + 1$
>   **end while**
> $F_L \leftarrow F_i$
>   **if** $|S| + |F_L| > N$ **then**
> $\bar{S}, \bar{F}_L, \hat{z}^*, \hat{z}^{nadir} \leftarrow normalize(S, F_L, \hat{z}^*, \hat{z}^{nadir})$
> $\rho, \pi, d \leftarrow 0$
>   **for** $k \leftarrow 1 \; to \; |S|$ **do**
> $\pi_k, d_k \leftarrow associate(\bar{S}_k, Z)$
> $\rho_{\pi_k} \leftarrow \rho_{\pi_k} + 1$
>   **end for**
> $S \leftarrow S \cup niching(\bar{F}_L, n - |S|, \rho, \pi, d)$
>   **end if**
> $P^{(t+1)} \leftarrow S$
>   **return** $P^{(t+1)}$
> **end procedure**

Algorithm 2 details the selection of the survivors set (S) for NSGA-III. At the t-th generation, the selection procedure takes as inputs (a) $R^t = P^t \cup Q^t$, where P^t is the parent population, Q^t is the offspring population, (b) the population size N, (c) the reference directions Z and, (d) $\hat{z}^*$ and $\hat{z}^{nadir}$ which are bounds to normalize individuals in the selection set S. The output is $P^{(t+1)} = S$. The first step involves non-dominated sorting of merged population R^t into hierarchical fronts $(F_1, F_2, \ldots)$. The fronts are sequentially appended to the initially null

(a) Step 1 - Scaling hydrofoil along the mid-chord vector

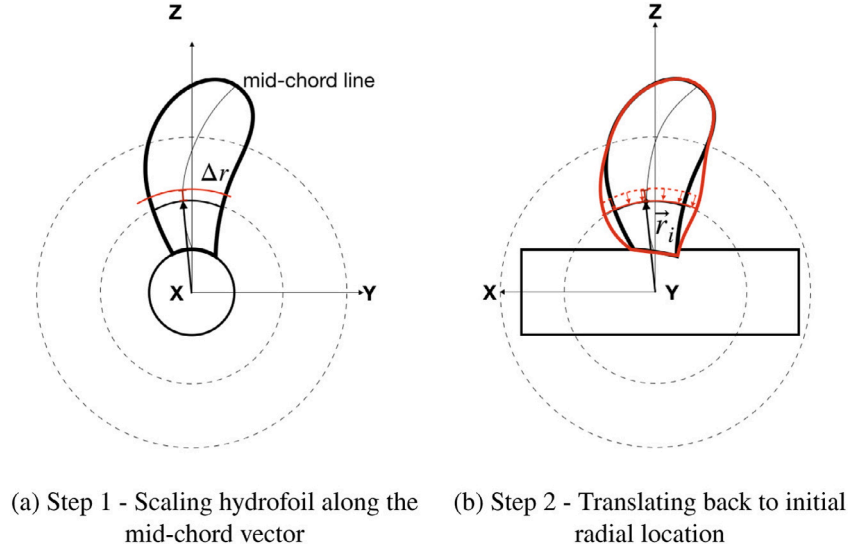(b) Step 2 - Translating back to initial radial location

Fig. 8. Schematic illustrating chordlength modification.

Table 3
Bounds for variance study.

Variable	Lower bound	Upper bound
$\Delta\theta$ (rad)	−0.6	0.6
$\Delta\gamma$ (rad)	−0.2	0.2
s_{camber} (rad)	−0.10	0.10
s_{chord} (−)	0.60	1.40
$\Delta\lambda$ (rad)	−0.10	0.10

set S until $|S| + |F_i| \leq N$. When $|S| + |F_L| > N$, the splitting front F_L is identified. Individuals in S and F_L are then normalized by using $\hat{z}^*$ and the nadir point estimation $\hat{z}^{nadir}$ as lower and upper bounds respectively. The individuals in $\bar{S}, \bar{F}_L$ are assigned to a reference direction π_k with the shortest perpendicular distance d_k. We generate the reference directions based on the Riesz s-energy concept to obtain a well-spaced distribution over the optimal Pareto front (Blank et al., 2021). The niching method selects $N - |S|$ individuals based on niche counts (ρ), π_k and d_k and appends them to S.

2.4. Geometry & mesh

The blade geometry is defined by a single B-Spline surface defined as

$$\vec{p}(u, v) = \sum_{i=0}^{m} \sum_{j=0}^{n} N_{i,d_1}(u) N_{j,d_2}(v) \vec{P}_{i,j}, \tag{9}$$

where $u, v \in [0.0, 1.0]$ are knots in the u and v directions; $\vec{P}_{i,j}$ is the control point of the ith row and jth column; $m+1$ and $n+1$ are the number of rows and columns the control points are organized in; d_1, d_2 are the degree of the B-Splines in u and v direction respectively and; N_{i,d_1} and N_{j,d_2} are the basis functions in the u and v direction respectively. The surface is fitted to hydrofoil cross-sections in Cartesian coordinates at different radial locations. In our case $d_1 = 3$ as the spline is closed in u direction and $d_2 = 8$ to find an optimal fit for a hydrofoil with camber and thickness distribution with fewer control points. The basis functions for each directions are defined as

$$N_{i,0}(t) = \begin{cases} 1; t_i \leq t < t_{i+1} \\ 0; otherwise, \end{cases} \tag{10}$$

$$N_{i,d} = \frac{t - t_i}{t_{i+d} - t_i} N_{i,d-1}(t) + \frac{t_{i+d+1} - t}{t_{i+d+1} - t_{i+1}} N_{i+1,d-1}(t). \tag{11}$$

where t is the knot value and d is the degree of the spline. The surface mesh is generated by uniformly discretizing the B-Spline surface in the B-Spline parametric space. The discrete nodes in the B-Spline parametric space are

$$M^*_{m \times n} = \begin{bmatrix} (0.0, 0.0) & \cdots & \left(0.0, \frac{(j-1)}{(n-1)}\right) & \cdots & (0.0, 1.0) \\ \vdots & \ddots & & & \vdots \\ \left(\frac{(i-1)}{(m-1)}, 0.0\right) & & \left(\frac{(i-1)}{(m-1)}, \frac{(j-1)}{(n-1)}\right) & & \vdots \\ \vdots & & & \ddots & \vdots \\ (1.0, 0.0) & \cdots & \cdots & \cdots & (1.0, 1.0) \end{bmatrix}, \tag{12}$$

where m is the number of nodes on the hydrofoil cross-section and n is the number of cross-sections along the radius. The nodes correspond to the mesh for the ith propeller $\vec{M}_i \in \mathfrak{R}^{3 \times m \times n}_{cart}$

$$\vec{M}_i = \{p(m^*_{ij}) | m^*_{ij} \in M^*_{m \times n}\}. \tag{13}$$

2.5. Orthogonal parameters

The primary objective of the proposed parametric model is to alleviate the limitations of classical design variables. Thus, the parametric model must (a) accurately quantify geometric variations, (b) yield orthogonally independent parameters and, (c) be generalized for all possible geometries. To satisfy the above requirements, the parameterization is based on the coordinates of the blades' surface mesh such that all possible variations of all possible families are represented and the relation between geometry and design objectives is not obscured. The parameters of the proposed model are derived by projecting the surface mesh of propellers in an orthonormal hyperspace $\mathfrak{R}_{orth}$ to ensure that the parameters are orthogonally independent. We define an orthonormal hyperspace $\mathfrak{R}_{orth}$ with a geometry variance study where pitch, skew, rake, chord-length and camber are individually varied within the bounds listed in Table 3. Chord-length and camber are varied for the last three hydrofoils. 100 equispaced samples are taken within the bounds for each variable.

Our data matrix M is stacked with meshes of all propellers in the variance study and then transformed into cylindrical coordinates:

$$\tilde{M} = \begin{bmatrix} \vec{M}_1 \\ \vdots \\ \vec{M}_N \end{bmatrix}, \tag{14}$$
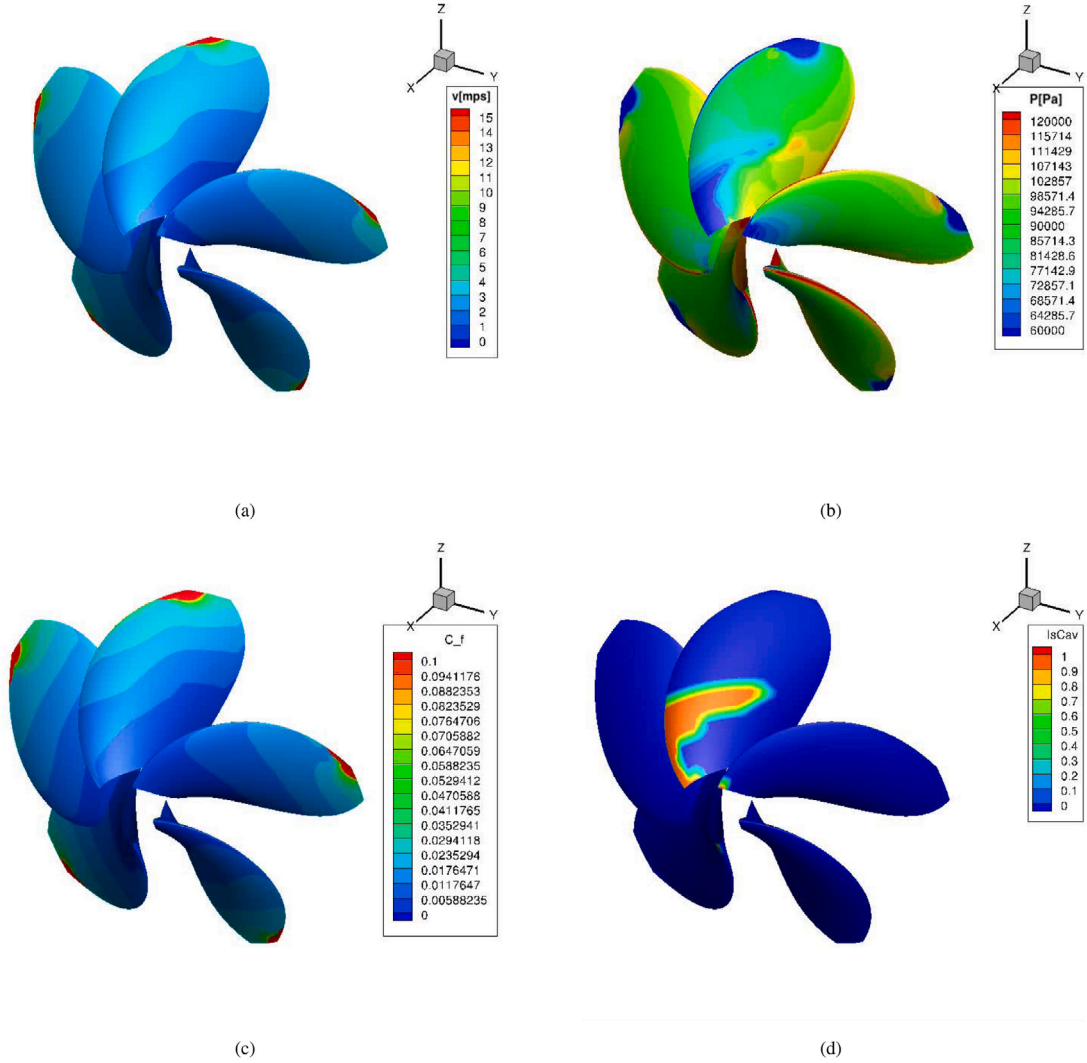
(a)

(b)

(c)

(d)

Fig. 9. B4-70 propeller (a) inviscid velocity distribution, (b) pressure field on blade, (c) skin friction coefficient and, (d) fraction of cavitating vertices on mesh panel. $|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ and IsCav $\in [0.00, 0.25, 0.50, 0.75, 1.00]$.

Algorithm 3 Subspace Projection

> **Input:** $\Re_{ortho}, k', \vec{M}^*$
> **Output:** $\vec{x}$
> **procedure** Projection
> $\quad \Re \leftarrow \Re_{ortho}$
> $\quad \vec{x}^* \leftarrow 0_{k'+1}$
> $\quad \vec{x}^*_{k'+1} \leftarrow |\vec{M}^*|$
> $\quad$ **for** $j = 1$ to k' **do**
> $\qquad \vec{x}^*_j \leftarrow cos^{-1}\left(\dfrac{\vec{M}^* \cdot \Re_{:,j}}{|\vec{M}^*||\Re_{:,j}|}\right)$
> $\quad$ **end for**
> $\quad$ **return** $\vec{x}^*$
> **end procedure**

$$M = T(\tilde{M}), T : \Re_{cart}^{N \times k} \rightarrow \Re_{cyl}^{N \times k}. \tag{15}$$

The origin of $\Re_{orth}$ is chosen to be the mean of all propeller meshes,

$$\vec{\mu} = \frac{\sum_{i=1}^{N} \vec{M}_i}{N}, \tag{16}$$

thus, the translated data matrix is

$$M' = \begin{bmatrix} \vec{M}_1 - \vec{\mu} \\ \vdots \\ \vec{M}_N - \vec{\mu} \end{bmatrix}. \tag{17}$$

We factorize M' using thin Singular Value Decomposition (SVD)

$$M'_{k \times N} = U_{k \times N} \Sigma_{N \times N} V'_{N \times N}. \tag{18}$$

In this worked example, the orthonormal reference frame is defined as the product of $k' = 10$ right and left singular vectors.

$$\Re_{orth} = U_{k \times k'} \cdot V'_{k' \times k'}. \tag{19}$$

Algorithm 3 details the procedure to project $\vec{M}^*$ sampled by NSGA-III in $\Re_{ortho}$ yielding the orthogonal parameters $\vec{x}^*$. For the orthogonal parameters, $\frac{\partial \tilde{M}}{\partial \vec{x}}$ is an orthonormal matrix. This is expected to solve the problem of multicollinearity.

2.6. Naive-Bayes classifier

We use an explainable supervised classifier to guide search towards more lucrative regions in the design space. Naive-Bayes classifier was
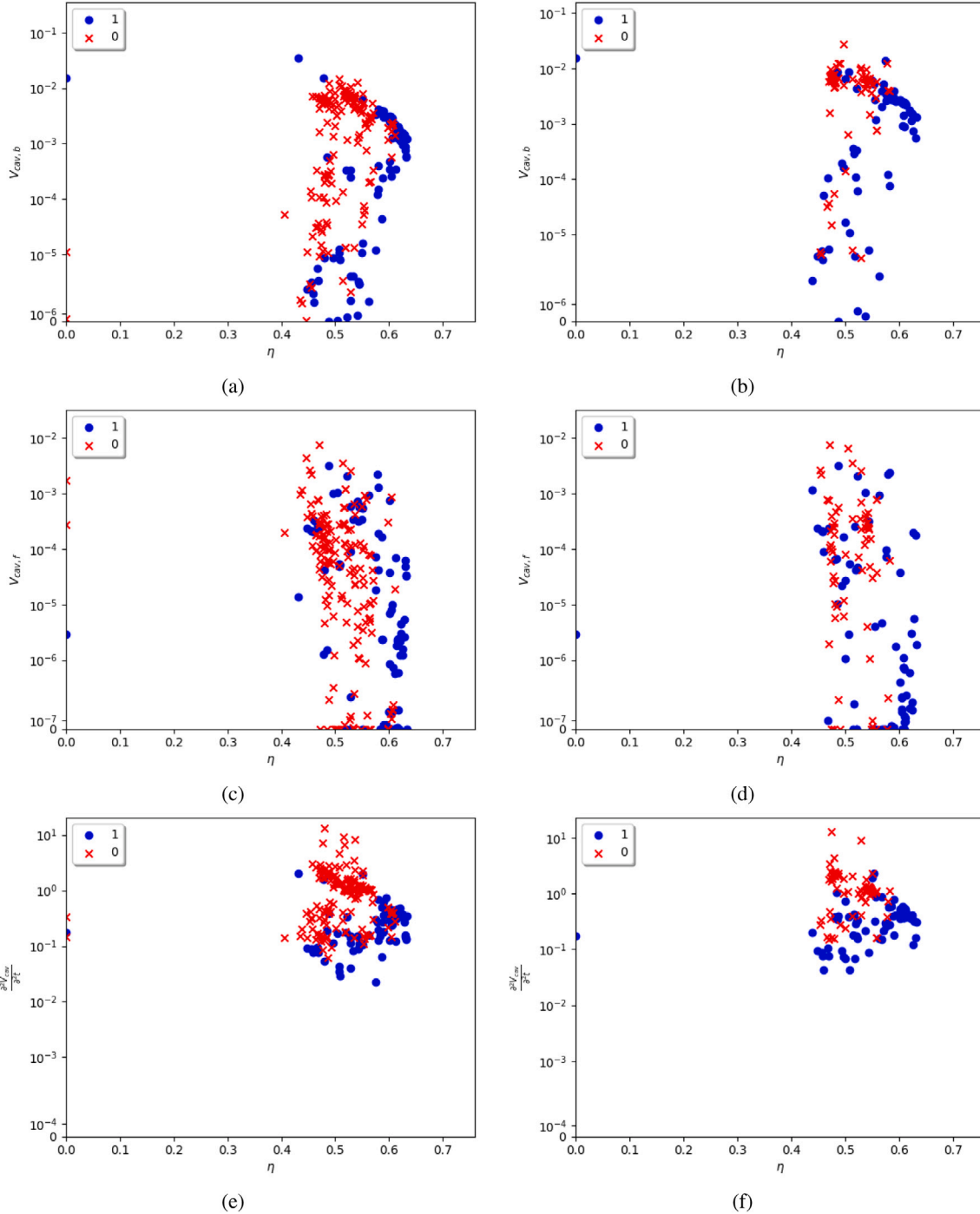
Fig. 10. Comparison of Pareto fronts of classical NSGA-III optimization (left) with DO (right) after 7 generations. The high dimensional Pareto front is illustrated as a combinations of efficiency vs. $V_{cav,f}$, $V_{cav,b}$ and $\frac{\partial^2 V_{cav}}{\partial^2 t}$.

demonstrated to work reliably (Doijode et al., 2022) and is also considered in this demonstration. The Naive-Bayes classifier predicts labels as

$$y_c = \text{argmax}_j \ \pi_j p(\vec{x}_i | \vec{\mu}_j, \vec{\sigma}^2_j), \qquad (20)$$

where j maximizes the probability of $\vec{x}_i$ occurring. The probability of $\vec{x}_i$ occurring given the mean and variance $\vec{\mu}_j, \vec{\sigma}^2_j$ for the jth cluster is

$$p(\vec{x}_i | \mu_j, \sigma^2_j) = \prod_{d=1}^{k'} \frac{1}{\sqrt{2\pi\sigma^2_{j,d}}} . \exp\left(-\frac{|x_{i,d} - \mu_{j,d}|^2}{2\sigma^2_{j,d}}\right). \qquad (21)$$

Where k' is the number of orthogonal parameters, $x_{i,d}$ is the dth parameter in $\vec{x}_i$, $\mu_{j,d}$ and $\sigma^2_{j,d}$ are the mean and variance of $X_{:,d}$. Implementations in SciKit-Learn (Pedregosa et al., 2011) are used.

For the worked optimization problem in this paper, it is assumed that the clusters are fundamentally of two types (a) interesting designs and (b) un-interesting designs. The former lies on the Pareto front and the latter either lies meaningfully away and/or cannot be evaluated due to limitations of the numerical method. The expectation for propeller design is that the current hypothesis of the two fundamental cluster types holds for most cases. It is, however, recognized that propeller
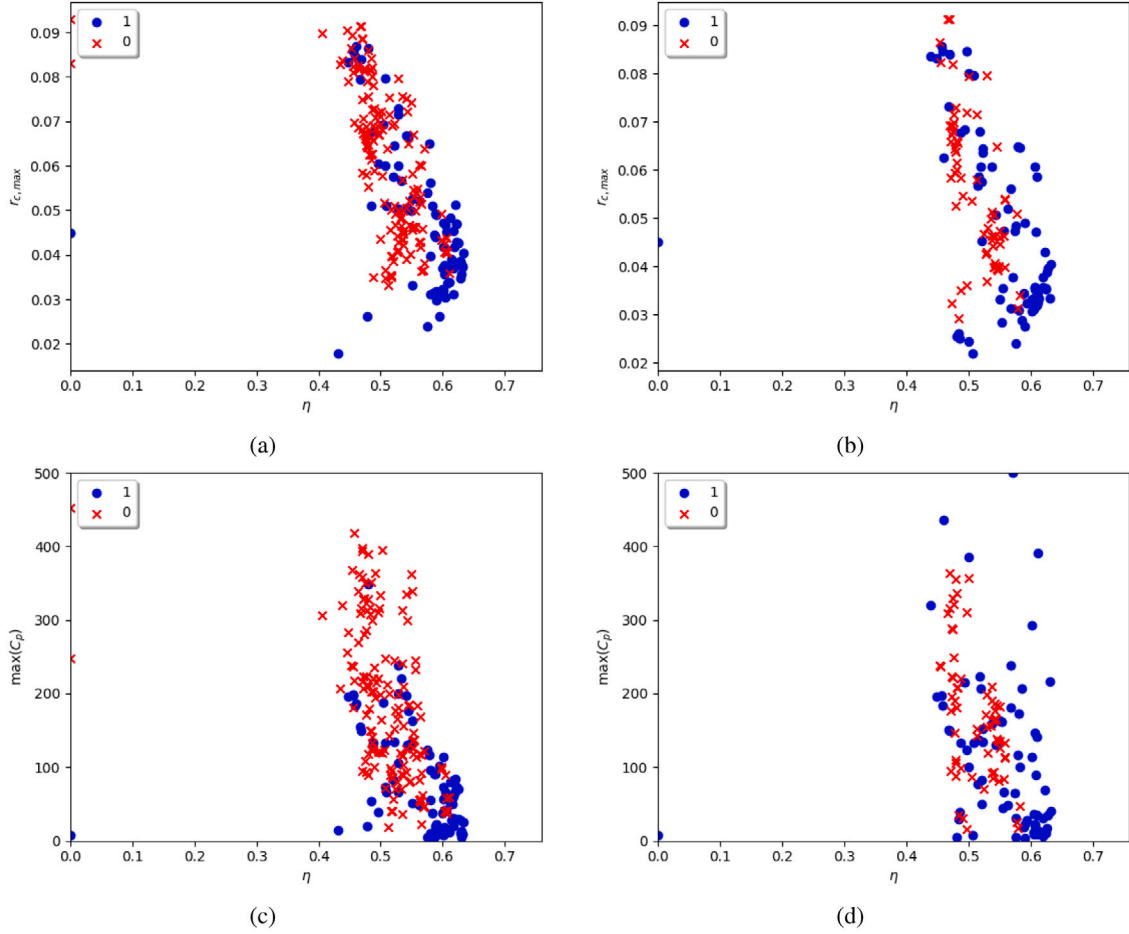
Fig. 11. Comparison of Pareto fronts of classical NSGA-III optimization (left) with DO (right) after 7 generations. The high dimensional Pareto front is illustrated as a combinations of efficiency vs. $r_{c,max}$ and $max(C_p)$.

Algorithm 4 Assign Labels

> **Input:** F, ϵ
> **Output:** $\vec{y}_c$
> **procedure** LABELS
> > $N \leftarrow len(F)$
> > $\vec{y}_c \leftarrow [0, ..., N-1]$
> > $next\_point \leftarrow 0$
> > $\sigma \leftarrow \epsilon F$
> > $F_{low} \leftarrow F - 1.96 \frac{\sigma}{\sqrt{N}}$ ▷ Lower bound of 95% confidence interval
> > $F_{high} \leftarrow F + 1.96 \frac{\sigma}{\sqrt{N}}$ ▷ Upper bound of 95% confidence interval
> > **while** $next\_point < N$ **do**
> > > $mask \leftarrow \{y | F_{high}[i] < F_{low}[next\_point] \ \forall i \in [0, N-1]\}$
> > > $mask[next\_point] \leftarrow 1$
> > > $\vec{y}_c \leftarrow \vec{y}_c[mask]$
> > > $F \leftarrow F[mask]$
> > > $F_{low} \leftarrow F_{low}[mask]$
> > > $F_{high} \leftarrow F_{high}[mask]$
> > > $next\_point \leftarrow \sum_{i=0}^{next\_point-1} mask[i] + 1$
> > **end while**
> > $\vec{y}_c \leftarrow \{y | \vec{y}_c[i] = 1 \ \textbf{AND} \ F_i \geq 0.85 \eta_{max} \ \forall i \in [0, N-1]\}$
> > **return** $\vec{y}_c$
> **end procedure**

---

design is a complex problem. It is quite possible that the set of interesting designs will be spread out over multiple clusters with additional

constraints/requirements which define an acceptable design. In such cases, One-vs-Rest (OvR) multiclass strategies (Scikit Learn, 2021) must be employed to have reliable classifier performance.

The labels $y_c \in [0, 1]$ are determined by whether the designs lie on the Pareto front and whether they have a high efficiency. To account for uncertainties, it is assumed that designs have a mean and variance in performance. A design dominates when the mean of the predicted performance is higher and there is no overlap in the 95% confidence interval with the compared solution. The 95% confidence interval is based on the assumption that there is a 15% uncertainty in predicted performance (Vaz et al., 2015). Algorithm 4 detail the procedure to assign labels. As inputs it takes objective function values and assumed uncertainty (ϵ). It returns the labels for all sampled points.

3. Results

The Pareto fronts resulting from classical NSGA-III and DO optimization are compared. The trade-offs in objectives as observed from the Pareto front are detailed. A solution with minimal suction surface cavitation, no pressure surface cavitation and moderate tip-vortex cavitation is identified and described.

3.1. Baseline performance

The performance of the baseline Wageningen B-4 70 P/D = 1.0 propeller operating in the wake of Seiun-Maru is estimated by PROCAL. For this worked example, the baseline performance prediction has local discontinuities in the pressure and velocity fields at the blade tip as
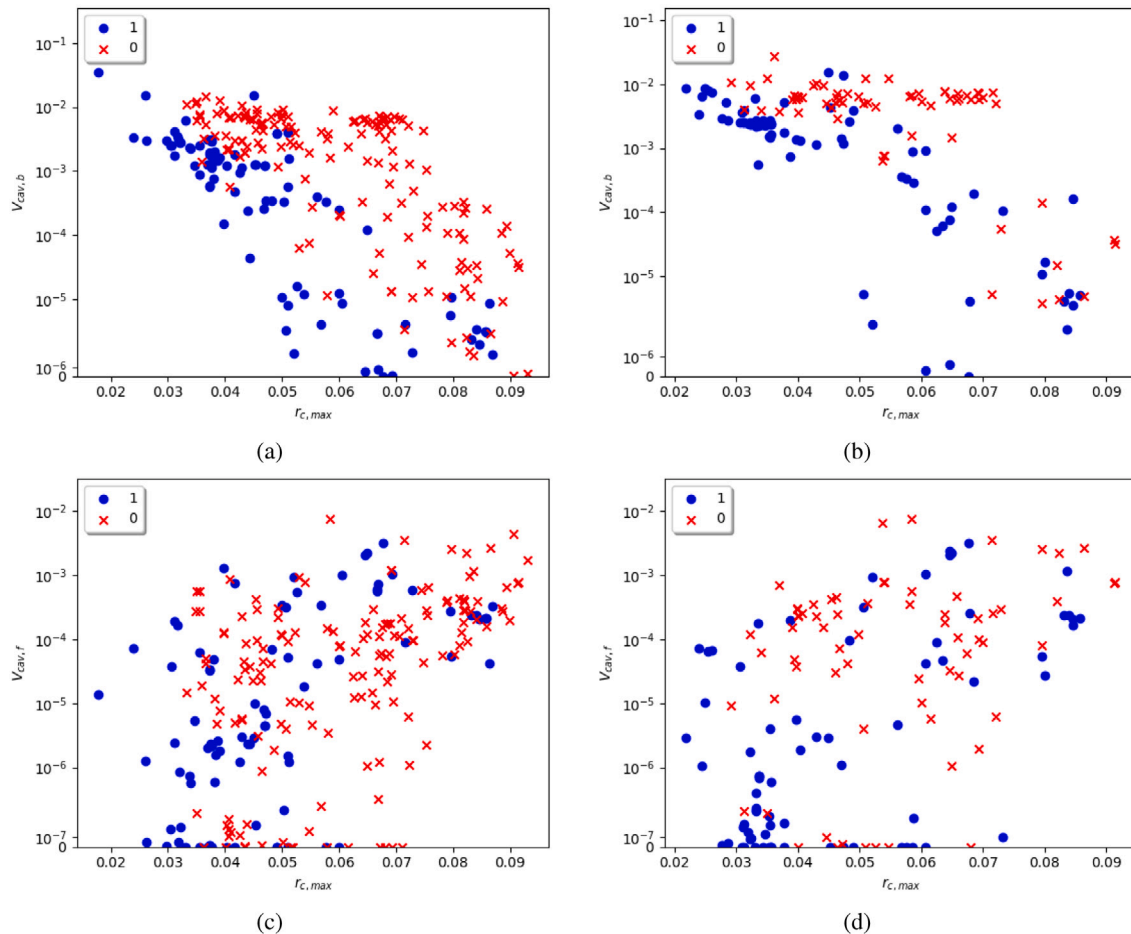
Fig. 12. Trade-off between V_{cav} and $r_{c,max}$ as found by NSGA-III optimization (left) and DO (right). The Pareto front visualized with (a, b) $V_{cav,b}$ as abscissa and, (c, d) $V_{cav,f}$ as abscissa.

illustrated in Fig. 9. Thus, it is difficult to interpret the results to identify sources of energy losses and cavitation behaviour. These discontinuities are expected to result from panels being relatively more skewed at the blade tip. Consequently, DO is also tasked with identifying designs which can be analysed reliably by PROCAL.

3.2. Pareto front

Both classical NSGA-III optimization and DO are iterated through 7 generation with 72 individuals in the population. For DO the classifier is used to guide search from the second generation onwards. Figs. 10–11 compares the Pareto front after the 7-th generation for classical NSGA-III and DO. The high dimensional Pareto front is plotted with η as the abscissa and $V_{cav,b}$, $V_{cav,f}$, $\frac{\partial^2 V_{cav}}{\partial^2 t}$, $r_{c,max}$, $max(C_p)$ as ordinates to illustrate the trade-off with efficiency. Blue dots represent designs which lie on the Pareto front ($y_c = 1$) and the red crosses represent dominated solutions ($y_c = 0$). It is observed that the Pareto fronts are comparable for both NSGA-III and DO.

Figs. 10(a)–10(d) illustrate the trade-off between $V_{cav,b}$, $V_{cav,f}$ and η. It is observed that improving efficiency over 52% is always accompanied by a cavitation penalty. This is because up to $\eta = 52\%$, efficiency can be increased without significantly loading the mid section of the propeller however, the tip is heavily loaded. This leads to larger tip-vortices as shown in Fig. 11(a). In this region, where $\eta \in [40\%, 52\%]$, lower mid-section loading also suggests lower supervelocities at the

mid-section compared to the tip. Consequently, higher rotational losses are expected. Furthermore, up to $\eta = 52\%$, viscous losses are expected to be reduced since, the gain in efficiency does not lead to an exponential cavitation penalty, . Thus, for region $\eta > 52\%$, gain in efficiency is expected to result from reduced rotational losses.

From Figs. 10(e)–10(f), it is observed that the cavity volume acceleration is better contained despite a growth in cavity volume. This is likely because skew is also optimized. From Figs. 11(c)–11(d) it can be observed that reducing $max(C_p)$ results in more efficient propellers. This is because higher peak pressures lead to higher induced losses which can be modelled in BEM. In addition to induced losses, higher peak pressures may also lead to flow separation which requires scale resolving simulations, such as Large Eddy Simulations (LES) or Detached Eddy Simulations (DES).

Figs. 11(c)–11(d) also show that very high values of $max(C_p)$ are observed. The bounds of $max(C_p)$ are determined by the stagnation pressure at the leading edge and the vapour pressure due to cavitation. Excessively high values of $max(C_p)$ may indicate numerical artefacts, which are often observed at nodes on the leading edge when BEM is used. By specifying the minimization of $max(C_p)$ as an objective, we encourage search algorithms to minimize these unrealistic spikes. However, since we use the same panel distribution for each propeller, we do not expect an effect on the relative ranking of the designs.

From Figs. 11(a)–11(b), it is observed that $r_{c,max}$ can be reduced down to 4 cm and efficiency can be increased suggesting a reduction in rotational losses. However, further reducing $r_{c,max}$ results in an
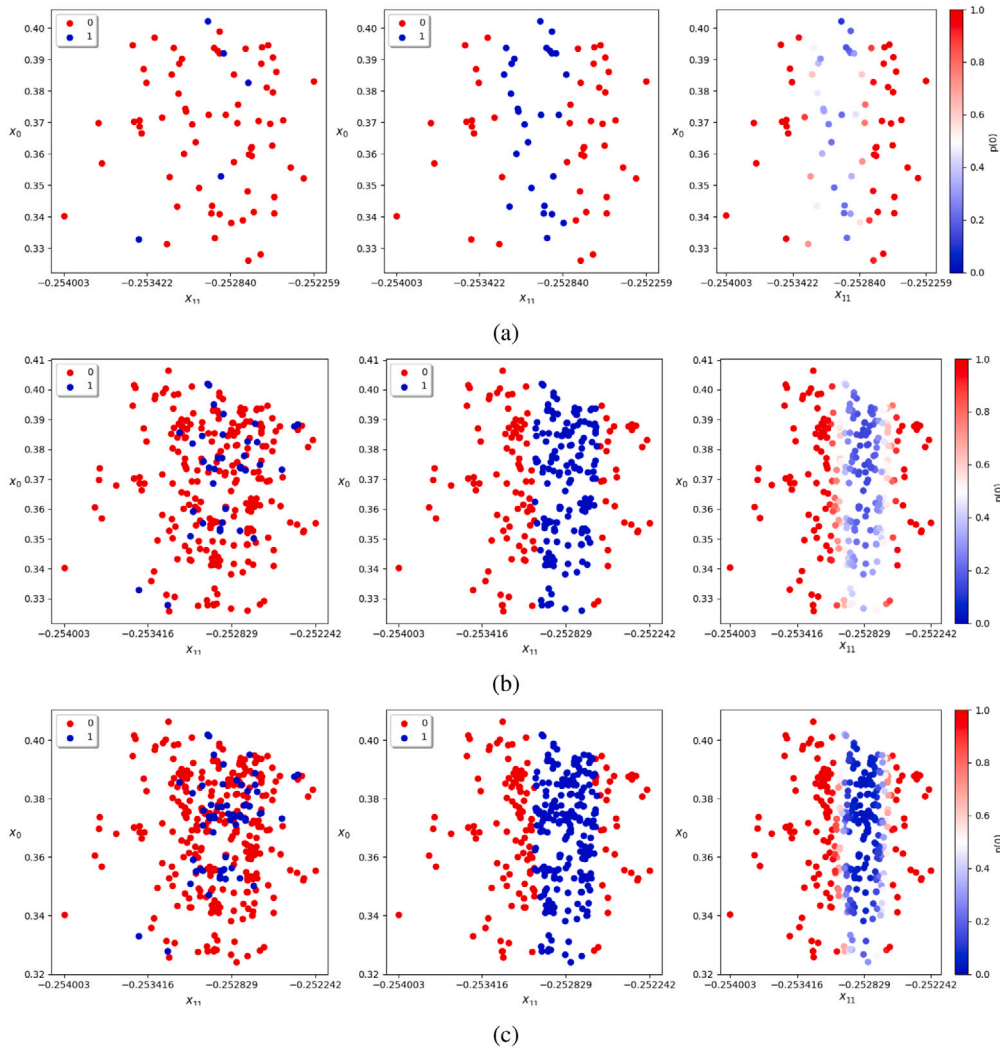
Fig. 13. Evolution of Naive-Bayes label predictions over generations, with blue dots ($y_c = 1$) indicating solutions that lie on the Pareto front. In all plots, the *x*-axis is x_0 and *y*-axis is x_{11}. The plots are for (a) Generation I, (b) Generation IV and (c) Generation VII. Left: Ground reality, Center: The label predicted by Naive-Bayes classifier, Right: The cluster probabilities.
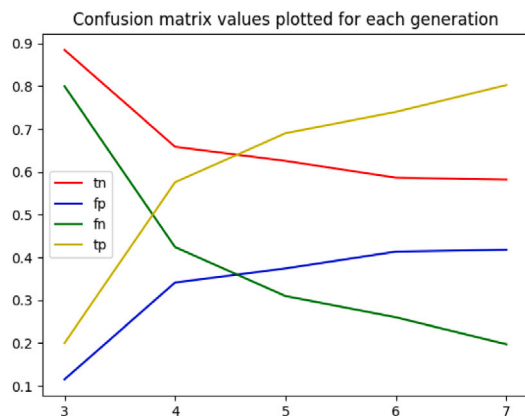
Fig. 14. Values of the confusion matrix, tp — True Positive, tn — True Negative, fp — False Positive, fn — False Negative.

efficiency penalty. The penalty is expected to result from growth in cavity volume both on the pressure and suction surface of the propeller. Fig. 12 illustrates the trade-off between V_{cav} and $r_{c,max}$. It is observed from Figs. 12(a)–12(b) that reducing $r_{c,max}$ results in a higher $V_{cav,b}$. This suggests that a trade-off exists between higher tip v/s higher mid-section loading.

3.3. Naive-Bayes classifier

Fig. 13 illustrates the label predictions made by the trained Naive-Bayes classifier. In Figs. 13(a)–13(c), the left most subplot illustrates the ground reality. Interesting designs which lie on the Pareto front have $y_c = 1$. The centre subplot illustrates the label predicted by the trained classifier and the right most subplot illustrates the $p(\vec{x}|\mu_0, \sigma_0^2)$. It is observed that the classifier learns the location of the Pareto front by the first generation visible as a light blue band in Fig. 13(a). The learning is reinforced in further generations, i.e. the classifier gets more confident of the location of the Pareto front. This is visible as the blue band intensifies. This is also observed from the so-called confusion matrix whose values are plotted in Fig. 14. It is observed that the true positive predictions(tp) for designs that lie on the Pareto front steadily increases
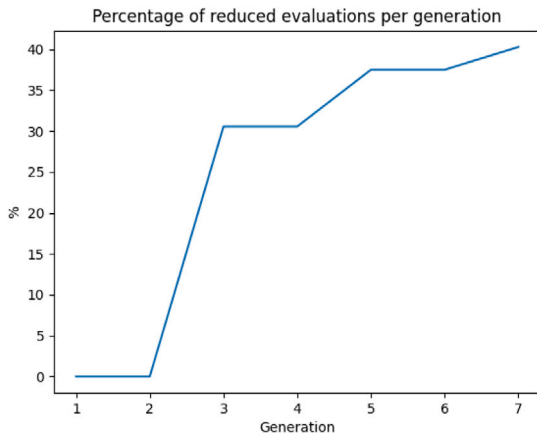
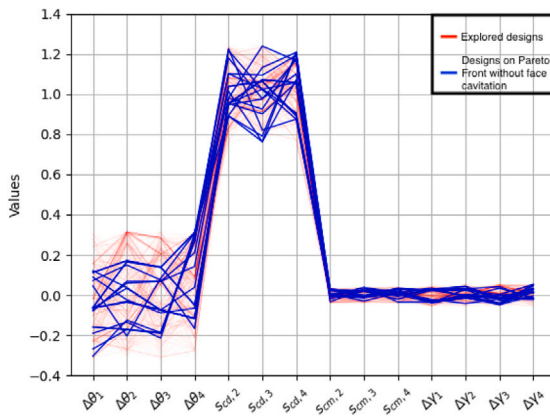**Fig. 15.** Percentage of evaluations reduced per generation.

Fig. 16. Parallel coordinate plot of explored designs and designs that lie on the Pareto front.

Table 4

Design parameters and performance for the least cavitating design.

f_i	Units	Value
η	–	5.2023e1 %
$V_{cav,b}$	m³	3.2972e−4
$V_{cav,f}$	m³	0.0000
$\frac{\partial^2 V_{cav}}{\partial^2 t}$	m³s⁻¹	1.7486e−1
$r_{c,max}$	m	5.761e−2
$max(C_p)$	–	8.2648e1

(a) Predicted objective function values.

Variable	Units	Value
$\Delta\theta_1$	rad	4.1000e−2
$\Delta\theta_2$	rad	−1.8405e−1
$\Delta\theta_3$	rad	7.5432e−2
$\Delta\theta_4$	rad	2.7701e−1
$\Delta\gamma_1$	rad	3.5201e−1
$\Delta\gamma_2$	rad	1.3164e−1
$\Delta\gamma_3$	rad	3.6064e−1
$\Delta\gamma_4$	rad	−2.2871e−1
$s_{chord,2}$	–	1.1440
$s_{chord,3}$	–	9.8151e−1
$s_{chord,4}$	–	1.0800
$s_{camber,2}$	–	1.9978e−2
$s_{camber,3}$	–	−3.7130e−4
$s_{camber,4}$	–	2.0544e−3

(b) Design parameters

generation over generation. As the model becomes more accurate in predicting the location of the Pareto front, it is possible to reduce evaluations of designs which are expected to lie meaningfully away from the Pareto front reducing the computational cost of optimization.

Fig. 15 shows the percentage of reduced evaluations per generation as the optimization progresses. From the third generation onwards the classifier reduces at least 30% of evaluations. With online learning, the classifier is better able to identify the Pareto front as the optimization progresses and the percentage of reduced evaluations increases steadily to 40%.

3.4. Pareto solutions

Fig. 16 shows the so-called Parallel Coordinate plot for explored designs (-) and the designs of interest that lie on the Pareto front with $V_{cav,f} = 0.0m^3$ (-). There are 19 diverse designs of interest. Among the designs, increasing efficiency comes with a trade-off with either increasing suction side cavitation, higher cavity volume acceleration or tip-vortex radius as detailed in Section 3.2. In the Parallel Coordinate plot, if $\Delta\theta, \Delta\gamma, s_{camber} < 0$, the pitch, skew and camber for the hydrofoil cross-section have reduced. Otherwise, they have remained the same or increased compared to the baseline geometry. If $s_{chord} < 1.0$, the hydrofoil chord-length has reduced. Otherwise, the chord-length has either remained the same or increased.

For the current operating point, a cavitation free solution was not found. Although the maximum efficiency is $max(\eta) = 63.49\%$, the corresponding solution has a large suction surface cavity. Solutions without pressure side cavitation and much lower suction surface cavitation are found on the Pareto front. Among solutions without pressure side cavitation, the design with the minimum suction side cavitation is illustrated in Fig. 17. Fig. 17(c) compares the pitch distribution of this optimized design and baseline. Compared to the baseline design, the pitch at the root section is slightly decreased while the pitch at the blade tip has been increased substantially. Fig. 17(d) compares the skew distribution of the optimized design and baseline. The skew at the mid-section and at the tip has been increased. Fig. 17(e) compares the chordlength distribution of this optimized design and baseline. The chordlength is measured for the expanded hydrofoil. It is noted that the chordlength of the mid-section has increased substantially. The chordlength of the tip-section has reduced compared to the baseline. From Fig. 17(f) it is observed that the rake of the blade has also changed from the baseline while λ has not been varied. As mentioned in Section 2.2, the change most likely results from the sequence in which $\theta, \gamma, s_{chord}, s_{camber}$ are varied. Furthermore, it could also result from residuals in fitting of B-Spline curve to the cross-section and lofting operation of a B-Spline surface through the fit curves. Table 4b lists the design parameters. In the table, the design parameters provide insight into how the geometry has changed. The chordlength at the root and tip has reduced while that at the mid-section has increased. The camber of the hydrofoil has been reduced at the root and mid-section while it has increased at the tip.

Fig. 18 illustrates the pressure, velocity, skin-friction coefficient and cavitation behaviour resulting from the previously mentioned changes to the baseline. Table 4a lists the objective function values. It is observed that the blade loading is optimized to eliminate discontinuities at the blade tip in velocity and pressure field when compared with the baseline. Nonetheless, *Design 1* seems to have mitigated adverse pressure gradients. However, suction pressure at the blade root and along the leading edge indicate a risk of cavitation. It is observed that loading on the mid-section is reduced as the blade is de-pitched and camber is reduced. In order to meet the thrust constraint, the loading is shifted more towards the blade tip by increasing the tip pitch and camber. Consequently, $V_{cav,b}$ is reduced and $r_{c,max}$ increases. Due to higher tip loading the torque requirement and rotational losses increases resulting in lower hydrodynamic efficiency.
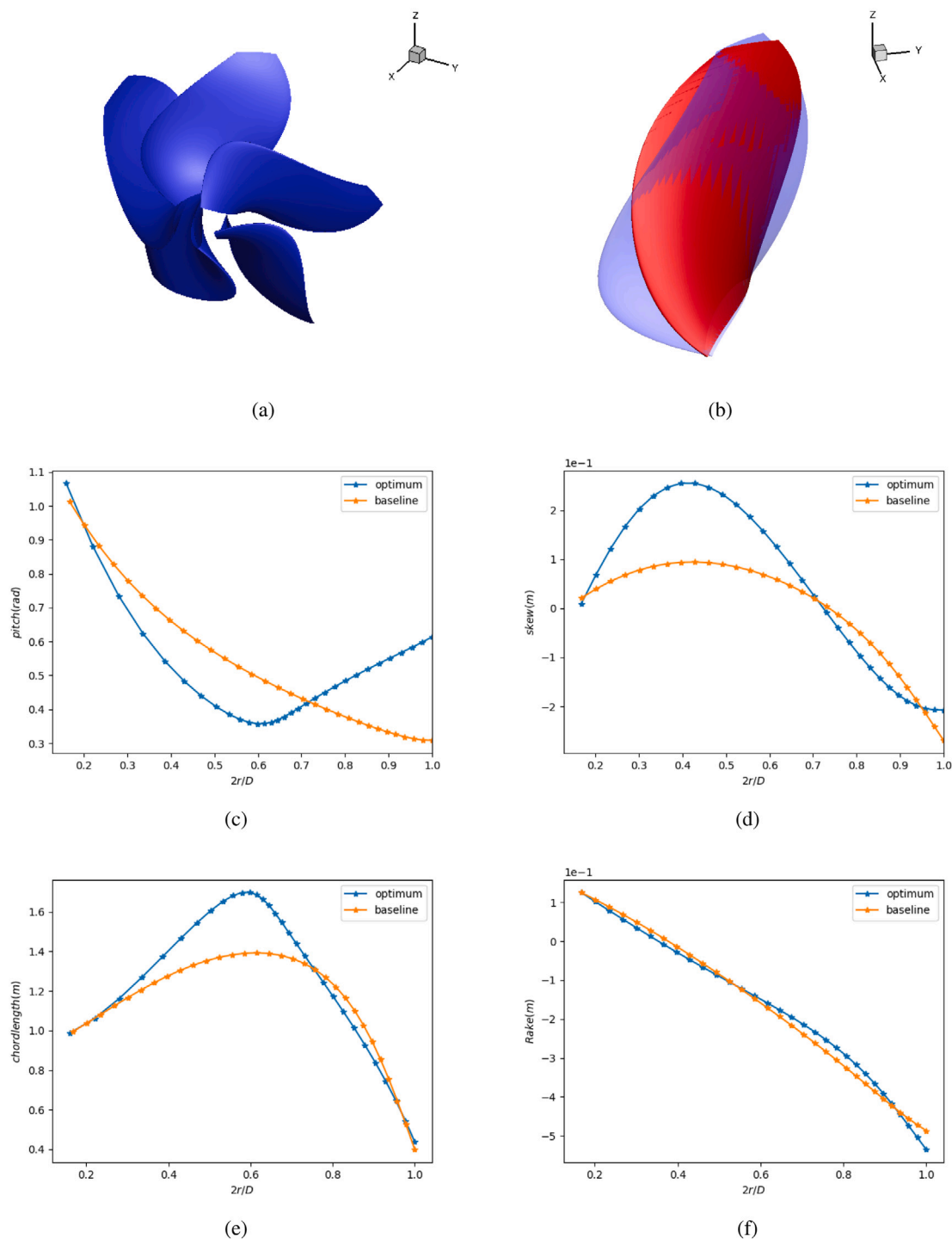
Fig. 17. Design solution with minimum suction side cavitation volume:(a) geometry, (b) comparison with baseline geometry, (c) pitch distribution of baseline and design, (d) skew distribution of baseline and design and, (e) chordlength distribution of baseline and design and, (f) rake distribution of baseline and design.

4. Conclusion

An optimization method to solve multi-objective, constrained optimization problems based on machine learning strategies is proposed. The method referred to as *Dynamic Optimization* (DO) is demonstrated on the hydrodynamic optimization of Wageningen B-4 70 propeller with P/D = 1.0 operating in the Seiun-Maru wakefield. Compared to classical NSGA-III optimization, DO is able to reduce at least 30% of evaluations per generation while reproducing a comparable Pareto front. The gain in computational efficiency of optimization is achieved with an online learning strategy using the Naive-Bayes classifier, which identifies the location of the Pareto front and designs that are located close to the front. Both classical NSGA-III and DO identify similar trade-offs between suction side, pressure side, tip-vortex cavitation and efficiency. For the current demonstration case, improving efficiency is accompanied by a cavitation penalty. It is possible to reduce suction side cavitation and eliminate pressure side cavitation fully. Reducing suction side cavitation requires a higher tip loading resulting in a larger tip vortex. Nonetheless, a non-elitist search strategy results in a Pareto front with very diverse designs.
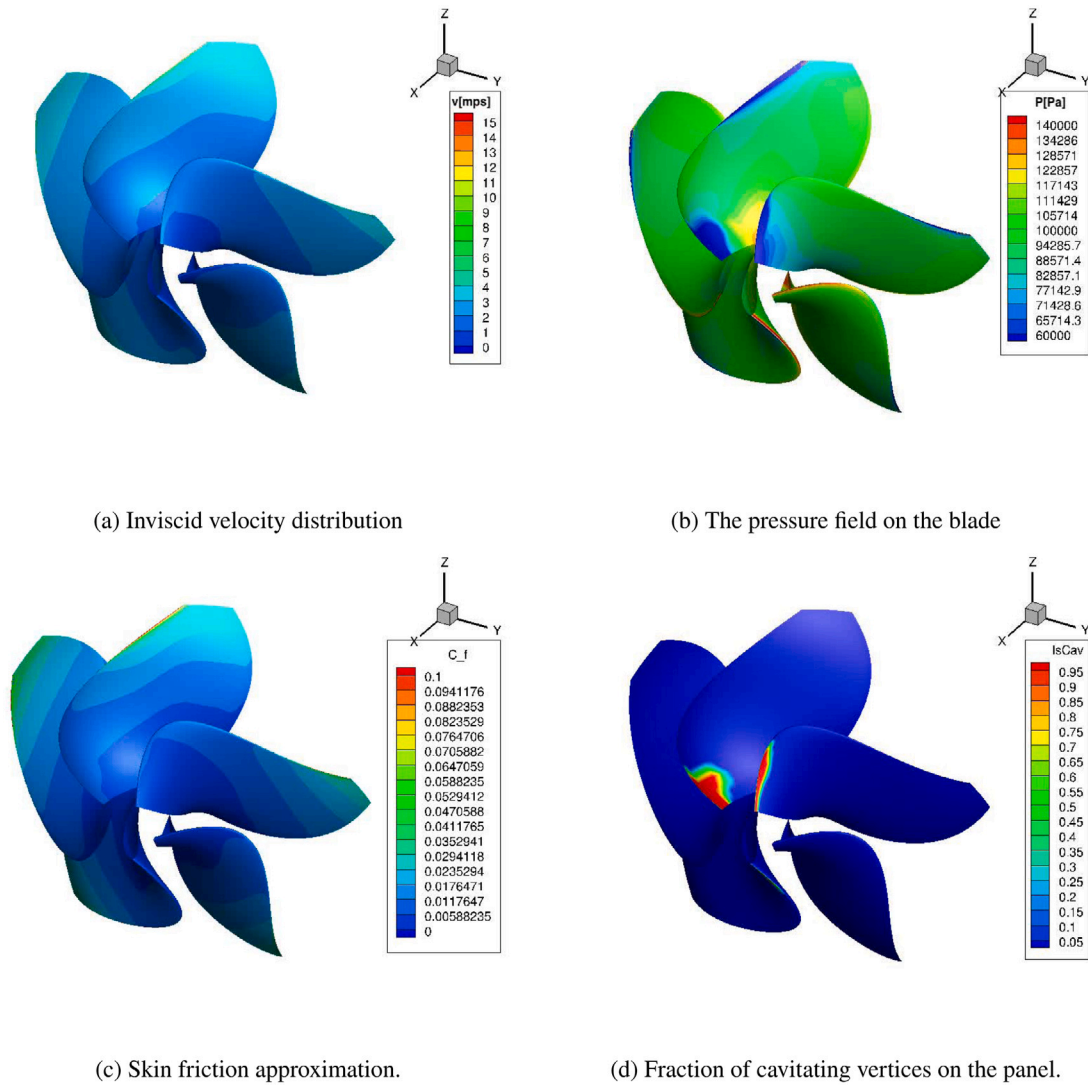
(a) Inviscid velocity distribution

(b) The pressure field on the blade

(c) Skin friction approximation.

(d) Fraction of cavitating vertices on the panel.

Fig. 18. Design 1 (a) inviscid velocity distribution, (b) pressure field on the blade, (c) skin friction coefficient and, (d) fraction of cavitating mesh panels. $|\vec{v}| = \sqrt{v_x^2 + v_y^2 + v_z^2}$ and IsCav $\in [0.00, 0.25, 0.50, 0.75, 1.00]$.

CRediT authorship contribution statement

Pranav Sumanth Doijode: Research (conceptualization, methodology), Software, Writing. **Stefan Hickel:** Guidance, Supervision and Review. **Tom van Terwisga:** Guidance, Supervision and Review. **Klaas Visser:** Guidance, Supervision and Review.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

Abedjan, Z., Boujemaa, N., Campbell, S., Casla, P., Chatterjea, S., Consoli, S., Costa-Soria, C., Czech, P., Despenic, M., Garattini, C., et al., 2019. Data science in healthcare: benefits, challenges and opportunities. In: Data Science for Healthcare. Springer, pp. 3–38.

Abrahamsen, K., 2012. The ship as an underwater noise source. In: Proceedings of Meetings on Acoustics ECUA2012, Vol. 17. Acoustical Society of America, 070058.

Allen, M.P., 1997. The Problem of Multicollinearity. Springer US, Boston, MA, pp. 176–180.

Arveson, P.T., Vendittis, D.J., 2000. Radiated noise characteristics of a modern cargo ship. J. Acoust. Soc. Am. 107 (1), 118–129.

Blank, J., Deb, K., Dhebar, Y., Bandaru, S., Seada, H., 2021. Generating well-spaced points on a unit simplex for evolutionary many-objective optimization. IEEE Trans. Evol. Comput. 1.

Blank, J., Deb, K., Roy, P.C., 2019. Investigating the normalization procedure of NSGA-III. In: International Conference on Evolutionary Multi-Criterion Optimization, Vol. 1. Springer, pp. 229–240.

Bosschers, J., 2009. Investigation of hull pressure fluctuations generated by cavitating vortices. In: Proc. First Symposium on Marine Propulsors.

Bosschers, J., 2018a. Propeller tip-vortex cavitation and its broadband noise. (Ph.D. thesis). University of Twente, Netherlands.

Bosschers, J., 2018b. A semi-empirical prediction method for broadband hull-pressure fluctuations and underwater radiated noise by propeller tip vortex cavitation. J. Mar. Sci. Eng. 6 (2), 49.

Brizzolara, S., Gaggero, S., 2009. Silent propellers with unconventional profile shapes. examples obtained with a new automatic optimization method. In: Proceedings

of the 4th Annual Conference MAST2009 on Maritime Systems and Technology. Stockholm, Sweden.

Deb, K., Jain, H., 2013. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. IEEE Trans. Evol. Comput. 18 (4), 577–601.

Doijode, P.S., Hickel, S., van Terwisga, T., Visser, K., 2022. A machine learning approach for propeller design and optimization: Part i. Applied Ocean Research Accepted for publishing.

Faresi, R., 1990. Full scale measurements. In: ITTC Group Discussions.

FlorianVesting, R.B., 2011. Propeller optimisation considering sheet cavitation and hull interaction. In: Second International Symposium on Marine Propulsors.

Foeth, E.-J., 2015. Propeller optimization using an unsteady boundary-element method. In: Fourth International Symposium on Marine Propulsors. pp. 27–32.

Gaggero, S., Tani, G., Viviani, M., Conti, F., 2014. A study on the numerical prediction of propellers cavitating tip vortex. Ocean Eng. 92, 137–161.

Garg, N., Kenway, G.K., Martins, J.R., Young, Y.L., 2017. High-fidelity multipoint hydrostructural optimization of a 3D hydrofoil. J. Fluids Struct. 71, 15–39.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep Learning, Vol. 1. MIT Press, Cambridge.

Huisman, J., Foeth, E.-J., 2017. Automated multi-objective optimization of ship propellers. In: Proceedings of the Fifth International Symposium on Marine Propulsors (SMP'17), Espoo, Finland.

IMO, 2019. Reducing greenhouse gas emissions from ships.

International Maritime Organization, 2014. Guidelines for the reduction of underwater noise from commercial shipping to address adverse impacts on marine life. https://cetsound.noaa.gov/Assets/cetsound/documents/MEPC.1-Circ%20883%20Noise%20Guidelines%20April%202014.pdf.

International Maritime Organization, 2021. Prevention of air pollution from ships. https://www.imo.org/en/OurWork/Environment/Pages/Air-Pollution.aspx.

Kato, H., Kodama, Y., 2003. Microbubbles as a skin friction reduction device–a midterm review of the research. In: Proceedings of the 4th Symposium on Smart Control of Turbulence, Tokyo, Japan.

Liebrand, R., Klapwijk, M., Lloyd, T., Vaz, G., 2021. Transition and turbulence modeling for the prediction of cavitating tip vortices. J. Fluids Eng. 143 (1), 011202.

Liu, Z., Young, Y.L., 2009. Utilization of bend–twist coupling for performance enhancement of composite marine propellers. J. Fluids Struct. 25 (6), 1102–1116.

Maljaars, P., 2019. Hydro-elastic analysis of flexible marine propellers. (Ph.D. thesis). Delft University of Technology.

Maljaars, P.J., Kaminski, M.L., 2015. Hydro-elastic analysis of flexible propellers: an overview. In: Fourth International Symposium on Marine Propulsors.

Miglianti, L., Cipollini, F., Oneto, L., Tani, G., Gaggero, S., Coraddu, A., Viviani, M., 2020. Predicting the cavitating marine propeller noise at design stage: A deep learning based approach. Ocean Eng. 209, 107481.

Motley, M., Liu, Z., Young, Y., 2009. Utilizing fluid–structure interactions to improve energy efficiency of composite marine propellers in spatially varying wake. Compos. Struct. 90 (3), 304–313.

Mulcahy, M.L., Croaker, P., McGuckin, D.G., Brandner, P.A., Kississoglou, N., 2014. Optimisation applied to composite marine propeller noise. In: INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Vol. 249. Institute of Noise Control Engineering, pp. 4012–4019.

Oprea, A.I., 2013. Prediction of tip vortex cavitation for ship propellers. (Ph.D. thesis). University of Twente.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Pennings, P., Westerweel, J., van Terwisga, T., 2016. Cavitation tunnel analysis of radiated sound from the resonance of a propeller tip vortex cavity. Int. J. Multiph. Flow. 83, 1–11.

Plucinski, M.M., Young, Y.L., Liu, Z., 2007. Optimization of a self-twisting composite marine propeller using genetic algorithms. In: 16th International Conference on Composite Materials, Kyoto, Japan. pp. 8–13.

Port of Vancouver, 2017. New incentive for cargo and cruise vessels intended to quiet waters around the port of vancouver for at-risk whales – port of vancouver. https://www.portvancouver.com/news-and-media/news/new-incentive-for-cargo-and-cruise-vessels-intended-to-quiet-waters-around-the-port-of-vancouver-for-at-risk-whales/.

Scikit Learn, 2021. SKLearn Multiclass One-vs-Rest Classifier https://scikit-learn.org/stable/modules/generated/sklearn.multiclass.OneVsRestClassifier.html.

Tendo, Y., Okada, Y., Okazaki, A., 2019. Developing the high efficiency propeller with optimization of skeg profile for twin-screw LNG carrier. In: Sixth International Symposium on Marine Propulsors (SMP'19), Rome, Italy.

van Terwisga, T., van Wijngaarden, E., Bosschers, J., Kuiper, G., 2007. Achievements and challenges in cavitation research on ship propellers. Int. Shipbuild. Prog. 54 (2–3), 165–187.

Vaz, G., Bosschers, J., 2006. Modelling three dimensional sheet cavitation on marine propellers using a boundary element method. In: Sixth International Symposium on Cavitation CAV2006, Wageningen, the Netherlands.

Vaz, G., Hally, D., Huuva, T., Bulten, N., Muller, P., Becchi, P., Herrer, J.L., Whitworth, S., Macé, R., Korsström, A., 2015. Cavitating flow calculations for the e779a propeller in open water and behind conditions: code comparison and solution validation. In: Proceedings of the 4th International Symposium on Marine Propulsors (SMP'15), Austin, TX, USA. pp. 330–345.

Vesting, F., 2015. Marine propeller optimisation-strategy and algorithm development. (Ph.D. thesis). Chalmers University of Technology.

Vesting, F., Bensow, R.E., 2014. On surrogate methods in propeller optimisation. Ocean Eng. 88, 214–227.

Young, Y.L., Motley, M.R., Barber, R., Chae, E.J., Garg, N., 2016. Adaptive composite marine propulsors and turbines: progress and challenges. Appl. Mech. Rev. 68 (6), 060803.