Physically plausible propeller noise prediction via recursive corrections leveraging prior knowledge and experimental data

Kalikatzarakis, Miltiadis; Coraddu, Andrea; Atlar, Mehmet; Gaggero, Stefano; Tani, Giorgio; Oneto, Luca

DOI
10.1016/j.engappai.2022.105660

Publication date
2023

Document Version
Final published version

Published in
Engineering Applications of Artificial Intelligence

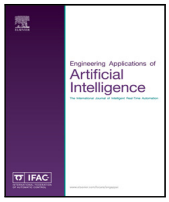
Citation (APA)
Kalikatzarakis, M., Coraddu, A., Atlar, M., Gaggero, S., Tani, G., & Oneto, L. (2023). Physically plausible propeller noise prediction via recursive corrections leveraging prior knowledge and experimental data. *Engineering Applications of Artificial Intelligence*, *118*, Article 105660. https://doi.org/10.1016/j.engappai.2022.105660

Engineering Applications of Artificial Intelligence 118 (2023) 105660

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Physically plausible propeller noise prediction via recursive corrections leveraging prior knowledge and experimental data

Miltiadis Kalikatzarakis [a], Andrea Coraddu [b,*], Mehmet Atlar [a], Stefano Gaggero [c], Giorgio Tani [c], Luca Oneto [d]

[a] Department of Naval Architecture, Ocean & Marine Engineering, University of Strathclyde, Glasgow, UK
[b] Department of Maritime & Transport Technology, Delft University of Technology, The Netherlands
[c] DITEN, Polytechnic School, University of Genoa, Genoa, Italy
[d] DIBRIS, University of Genova, Genova, Italy

ARTICLE INFO

Keywords:
Propeller cavitation noise
Hybrid models
Prior knowledge
Recursive corrections
Physical plausibility
Extrapolation
Meridian standard propeller series

ABSTRACT

For propeller-driven vessels, cavitation is the most dominant noise source producing both structure-borne and radiated noise impacting wildlife, passenger comfort, and underwater warfare. Physically plausible and accurate predictions of the underwater radiated noise at design stage, i.e., for previously untested geometries and operating conditions, are fundamental for designing silent and efficient propellers. State-of-the-art predictive models are based on physical, data-driven, and hybrid approaches. Physical models (PMs) meet the need for physically plausible predictions but are either too computationally demanding or not accurate enough at design stage. Data-driven models (DDMs) are computationally inexpensive ad accurate on average but sometimes produce physically implausible results. Hybrid models (HMs) combine PMs and DDMs trying to take advantage of their strengths while limiting their weaknesses but state-of-the-art hybridisation strategies do not actually blend them, failing to achieve the HMs full potential. In this work, for the first time, we propose a novel HM that recursively correct a state-of-the-art PM by means of a DDM which simultaneously exploits the prior physical knowledge in the definition of its feature set and the data coming from a vast experimental campaign at the Emerson Cavitation Tunnel on the Meridian standard propeller series behind different severities of the axial wake. Results in different extrapolating conditions, i.e., extrapolation with respect to propeller rotational speed, wakefield, and geometry, will support our proposal both in terms of accuracy and physical plausibility.

1. Introduction

The sustainability of anthropogenic activities is nowadays a fundamental problem requiring a multidisciplinary approach in order to be properly addressed. In a broad sense, sustainability concerns the control of the adverse effects of human activities on the environment.[1] In the maritime industry, heretofore, the impact of vessels on the environment was considered just in terms of atmospheric pollutant formation from the prime movers (Trivyza et al., 2021), release of toxic compounds from hull coatings (Torres and De-la Torre, 2021), or importation of exotic biological species through ballast water (Lakshmi et al., 2021) and biofouling (Song et al., 2020). Underwater Radiated Noise (URN) was just recently categorised as a form of pollution (Vakili et al., 2020a) due to the substantial increase of noise pollution on oceans worldwide (Sezen et al., 2021b). URN not only has severe effect on the marine ecosystem (Ferrier-Pagès et al., 2021; Di Franco et al., 2020) but also affects the crew and passenger comfort (Oldenburg et al., 2010). Several regulatory bodies have urged the industry to address the harmful effects of noise pollution (Chou et al., 2021; Vakili et al., 2020b) and take precautionary measures (European Union, 2017). For instance, the International Maritime Organisation (IMO) adopted a regulation defined by the International Convention for the Safety of Life at Sea (IMO, 2012) that requires shipowners to take measures that reduce on-board noise pollution, by setting mandatory maximum noise level limits. More recently, guidelines to reduce URN in commercial shipping were released, in the attempt to address adverse impacts on marine life (IMO, 2014). As a consequence, several classification societies have now introduced the "silent" class notation (Det Norkse Veritas, 2010). Apart from environmental concerns, URN abatement is compulsory in maritime warfare. Navies are conducting continuous research and

* Corresponding author.
E-mail addresses: miltiadis.kalikatzarakis@strath.ac.uk (M. Kalikatzarakis), a.coraddu@tudelft.nl (A. Coraddu), mehmet.atlar@strath.ac.uk (M. Atlar), stefano.gaggero@unige.it (S. Gaggero), giorgio.tani@unige.it (G. Tani), luca.oneto@unige.it (L. Oneto).

[1] https://ec.europa.eu/info/strategy/international-strategies/sustainable-development-goals

https://doi.org/10.1016/j.engappai.2022.105660
Received 22 June 2022; Received in revised form 7 November 2022; Accepted 21 November 2022
Available online xxxx

development activities in stealth technology optimisation and quiet submarines, with the aim of reducing their noise signature which determines their detectability, operability, and survivability (Tucker and Azimi-Sadjadi, 2011; Li and Wang, 2020).

The noise signature of a seagoing vessel is composed of a variety of sources including rotating machinery, hull-form, and propellers (Lurton, 2002). Among these, the propeller is the most significant noise contributor and the subject of increased interest in academia and the industry (Carlton, 2018). Non-cavitating propellers generate the highest noise levels at frequencies below 200 [Hz] while cavitating propellers can emit noise up to frequencies of 1 [MHz] dominating the audible noise spectrum (Ross, 1976; Carlton, 2018). As a consequence, a considerable amount of effort has been spent in investigating and implementing propeller URN mitigation strategies (Tong et al., 2021; Tong and Chen, 2020; Kim and Kinnas, 2021; Hu et al., 2019; Huang et al., 2019; Capone et al., 2021). Nevertheless, there is a natural tension between the need for silent yet efficient propellers that requires to find the best application specific trade-off (Gaggero et al., 2017a,b, 2016; Nouri et al., 2018; Valdenazzi et al., 2019). For this purpose, computational tools able to support the design process for finding these break points are needed (Valdenazzi et al., 2019; Li et al., 2018). In fact, at the design stage, it is not feasible to investigate the propeller's behaviour by means of experimental facilities (e.g., cavitation tunnels) since experiments are time consuming and expensive, thus they cannot be carried out for too many different designs (Miglianti et al., 2020).

State-of-the-art computational tools are based on three main approaches: Physical Models (PMs), Data Driven Models (DDMs), and Hybrid Models (HMs). Here we briefly summarise them but in Section 2 we will perform a thorough review

- PMs (see Section 2.1) rely on the knowledge of the phenomena and can be further subdivided in two main families
  - Empirical and semi-empirical models (see Section 2.1.1) utilise empirical formulas, to approximate with different levels of accuracy the physical phenomena, fine tuned by means of measurement data (Bosschers et al., 2017). These models are computationally efficient but usually not enough accurate to be exploited at design stage (Carlton, 2018).
  - Computational Fluid Dynamics (CFD) based models (see Section 2.1.2) decouple the sound propagation from its source generation, allowing to separate the flow solution from the acoustic analysis (Sezen et al., 2021a). The viscous flow field, where the sound source is generated, is solved by means of a CFD method with an appropriate turbulence model and the sound propagation is treated by an integral method based on acoustic analogy (Sezen et al., 2021b). CFD-based models can be quite accurate and reliable at the expense of large computational requirements, which prevent their use at the design stage (Bosschers, 2018a).
- DDMs (see Section 2.2) rely on Machine Learning (ML) and historical observations to build models of the phenomena with no prior physical knowledge about them (Coraddu et al., 2017). While DDMs can be quite computationally expensive during the model creation phase, they can be highly accurate and computationally inexpensive during the prediction phase, which is considered the crucial one at design stage (Miglianti et al., 2020). DDMs main limitation lies on their accuracy, which is high on average but not pointwise. Therefore, in some cases, DDMs can provide physically inconsistent predictions (Coraddu et al., 2021a).
- HMs (see Section 2.3) leverage on both PMs and DDMs. By combining them, they take advantage of their strengths while limiting their weaknesses (Miglianti et al., 2020). Specifically, HMs can achieve the same or higher accuracy with respect to DDMs (fully leveraging historical data), but they also leverage prior physical knowledge (exploiting computationally efficient outputs or partial computations behind PMs) to deliver physically plausible results (Coraddu et al., 2018).

HMs have the potential to deeply change the design stage of propellers. As a first advantage, they can rely on and fully exploit both physical knowledge of the phenomena and historical data to deliver both accurate and physically plausible results (Coraddu et al., 2021a). As a second advantage, they can be quite efficient in making predictions, allowing to include them in automatic tools for design optimisation (Coraddu et al., 2020). Finally, they reduce the need for historical observations, thanks to the exploitation of the prior knowledge about the phenomena, reducing consequently the computational effort needed to build them. Unfortunately, current hybridisation strategies do not allow to fully achieve HMs true potential, since they naively pull together PMs and DDMs without actually blending them. The limitation of this naive approach is twofold. The first one is that the current works available in the literature do not fully investigate the physical plausibility of the prediction (e.g., with comparisons between the expected behaviour of the models in circumstances where physical knowledge of the phenomena is high) (Vesting and Bensow, 2014). The second one is that current body of literature do not show the advantage of using HM in terms of the ability to better extrapolate (to be more accurate) with respect to PMs and DDM (and this limitation is also due to the unavailability of large datasets covering multiple propellers and multiple working conditions) (Miglianti et al., 2020).

For these reasons, in this work, we focus on tackling these three main gaps. In particular, we propose a novel hybridisation strategy able to truly blend PMs and DDMs into single HMs. The knowledge of the phenomena is exploited in model structure, in model building, and in model enrichment: we consider a dedicated feature engineering process to extract meaningful information from the available experimental data and several additional quantities derived from a computationally cheap Boundary Element Method (BEM) (Gaggero et al., 2017b; Gaggero and Villa, 2018). Furthermore, we exploited the computationally cheap PMs from our preliminary work (Kalikatzarakis et al., 2021) to estimate the noise generated by sheet cavitation and Tip Vortex Cavitation (TVC). Everything is empowered by a novel approach based on a state-of-the-art DDM which recursively correct a state-of-the-art PM, leveraging the information contained in the historical data. We tested the models in many different extrapolation scenarios, and we tested the physical plausibility of the results. In order to achieve this goal and support our statements, we will test our proposal, for the first time in the literature, on real data collected in a vast experimental campaign at the Emerson Cavitation Tunnel (Atlar, 2011) on the Meridian standard propeller series (Aktas, 2017) behind different severities of the axial wake.

The rest of the paper is organised as follows. Section 2 summarise the literature on the topics faced in this work. Section 3 describe the data collected during the experimental campaign at the Emerson Cavitation Tunnel on the Meridian standard propeller series. Section 4 presents our proposal. Section 5 exploits the data described in Section 3 to test the proposal detailed in Section 4. Finally, Section 6 concludes the paper. For the sake of readability, Table 1 summarises the acronyms used in the paper.

2. Related work

This section reviews the state-of-the-art computational tools for propeller URN prediction. Specifically, PMs, DDMs, and HMs will be reviewed in Sections 2.1, 2.2, and 2.3 respectively.

2.1. Physical models

PMs for propeller URN prediction can be further subdivided in two main families Empirical and semi-empirical models and CFD based model that we will review separately in Sections 2.1.1 and 2.1.2, respectively.

Table 1
Acronyms.

Acronym	Description	Acronym	Description
BEM	Boundary Elements Method	CFD	Computational Fluid Dynamics
DDM	Data Driven Model	DES	Detached Eddy Simulation
ETV	Empirical Tip Vortex Model	FI	Features Importance
FWH	Ffowcs William-Hawkings	HM	Hybrid Model
IMO	International Maritime Organisation	KRR	Kernel Ridge Regression
LOGO	Leave One Geometry Out	LORSO	Leave One Rotational Speed Out
LOWO	Leave One Wakefield Out	ML	Machine Learning
MAE	Mean Absolute Error	MAPE	Mean Absolute Percentage Error
NN	Neural Network	PM	Physical Model
PPMCC	Pearson Product-Moment Correlation Coefficient	RANS	Reynolds-Averaged Navier–Stokes
RF	Random Forests	RNL	Radiated Noise Level
SNN	Single-layered Neural Network	SPL	Sound pressure level
TPK	Test on Prior Knowledge	TVI	Tip Vortex Index
TVC	Tip Vortex Cavitation	URN	Underwater Radiated Noise
VLM	Vortex Lattice Method		

2.1.1. Empirical and semi-empirical models

Historically, empirical and semi-empirical models were the first attempts to predict propeller URN (Brown, 1976; Bosschers, 2018a, 2009; Brown, 1999; Bosschers, 2018c,b; Raestad, 1996). They utilise empirical formulas to approximate with different levels of accuracy the physical phenomena, and then they are fine tuned by means of measurement data (Bosschers et al., 2017). These models are computationally efficient but usually not enough accurate to be effectively exploited at design stage (Carlton, 2018) even if some researchers have investigated this research direction (Bosschers et al., 2017).

Brown (1976) were among the first researchers investigating empirical models for broadband noise. Exploiting measurements taken from thruster propellers, they proposed a simple model able to describe its upper limit for frequencies between 100 [Hz] to 10 [kHz]. The original model has been subsequently modified by several researchers (Okamura and Asano, 1988; Ekinci et al., 2010) improving prediction quality. Nevertheless, detailed comparisons for a reasonably large set of propellers were not reported.

Semi-empirical models focusing on sheet cavitation and its effects on propeller URN have been studied by Brown (1999), Matusiak (1992) and further extended by Brown (1999). Specifically, Brown (1999) presented a simple empirical formulation for the noise generated by thruster propellers. Their work provided a simple relation between the amplitude of noise and the area of sheet cavitation. Despite its lack of insight into the underlying physical phenomena, the method is able to accurately predict the URN of open propellers as recently demonstrated by Lafeber and Bosschers (2016).

Matusiak (1992) proposed a more theoretically grounded approach to model the collapse of free bubbles due to sheet cavitation, assuming that the number and the average size of cavitation bubbles generated by a cavity break-off to follow a beta distribution characterised by one parameter. This model produces a broadband propeller-induced pressure spectrum strongly correlated with the one measured from sheet cavitating propellers. This idea was then evolved in Kamiirisa and Goto (2005) who estimated the behaviour of sheet cavitation using Lifting Surface Method and model tests introducing the effect of compressibility and damping in bubble flow to improve the prediction accuracy.

In order to deal with noise sources different from sheet cavitation, Bosschers (2018c,b), Raestad (1996) studied tip vortex induced propeller URN. Specifically, Raestad (1996) proposed empirical formula for the amplitude of tip vortex noise. Later, Bosschers (2009) presented another empirical formula for the prediction of the characteristic frequency of tip vortex generated noise. Both studies succeeded in approximating the behaviour of vortex noise, but they both require extensive case-by-case tuning due to their high parameters sensitivity (Bosschers, 2018a).

Recently, Bosschers (2018c) combined the semi-empirical vortex model of Bosschers (2018a) with BEM to predict the hump-shaped pattern of the propeller cavitation noise spectrum for a variety of vessels. This approach was later improved by Bosschers (2018b) noting that the main weakness of the original method was its inability to model the underlying dynamics interactions that explain the inaccuracies observed in several case studies. For the sake of completeness and readability, Table 2 provides an overview of Empirical and Semi-Empirical state-of-the-art methods for propeller URN prediction reporting for each one of the considered works its scopes, case studies, exploited data, performance, and takeovers.

2.1.2. CFD-based models

CFD-based models, contrarily to Empirical and Semi Empirical models, are able to achieve very accurate results at the expense of considerable computational requirements preventing their use at design stage (Ianniello et al., 2013). Nevertheless, few large-scale studies involving CFD-based numerical models exists (Lidtke, 2017).

The most common CFD-based approach involves coupling hydrodynamic solvers with the acoustic analogy, effectively decoupling the noise source and noise propagation fields (Sezen et al., 2021a). For instance, Salvatore and Ianniello (2003) employed this approach on the David Taylor Research Center (DTRC) 4148 propeller, utilising the Ffowcs Williams–Hawkings (FWH) acoustic analogy evaluated on the blade surface (to deduce the URN radiated by the propeller) and potential flow coupled with BEM (for the hydrodynamic analysis).

Similarly, Seol et al. (2002) conducted a numerical study for the prediction of non-cavitating propeller radiated noise utilising the FWH coupled with BEM considering a wide range of operating conditions. The aim of the study was to investigate the effects of duct geometry on the overall propeller URN considering sound reflection and scattering effects. This study was further extended by Seol et al. (2005), studying propeller URN levels and noise directivity patterns of various noise sources under the presence of sheet cavitation on the DTRC-4119 propeller.

Several studies (Ianniello et al., 2013, 2014a,b) employed Reynolds-Averaged Navier–Stokes (RANS) coupled with FHW to investigate the propeller URN induced by a model- and full-scale ships and propellers in non-cavitating conditions. The authors concluded that the propeller URN is considerably affected by the contribution of nonlinear noise sources which must be taken into account regardless of propeller operational conditions. The authors also challenged the commonly accepted idea that cavitation is in most cases the dominant propeller URN source showing possible significant noise levels in non-cavitating conditions. Finally, the authors concluded that the current state-of-the-art in cavitation modelling is unlikely to be sufficient to accurately capture the associated propeller URN as the compressibility effects are usually excluded from the analysis.

The previous studies were further extended by Ianniello and De Bernardis (2015), where RANS and Detached Eddy Simulation (DES) solvers were coupled with the FWH equation, showing that RANS solvers were inadequate for hydroacoustics analyses. The

Table 2
Propeller UNR empirical and semi-empirical models review.

Ref.	Scopes	Case studies	Exploited data	Performances	Takeovers
Brown (1999)	URN prediction of thruster propellers and reduction technologies.	Discoverer Enterprise drill-ship.	Measurements from 6 thruster propellers.	≤ 30 [dB]	Acceptable qualitative agreement between the predicted and measured highest levels of cavitation noise.
Bosschers (2018c)	URN prediction of broadband hull pressure fluctuations and propeller tip vortex cavitation.	Combi freighter vessel and MS Statendam cruise vessel (Kipple, 2002).	Experiments from various vessels.	Average error less than 25 [dB] with high variance among the case studies.	Aspects influencing the shape of the hump and the high frequency slope need further investigation.
Bosschers (2018b)	Prediction of the azimuthal velocity distribution of a cavitating vortex in a two-dimensional viscous flow.	Wing tip-vortex in cavitating and non-cavitating conditions.	Experimental data from Pennings et al. (2015).	Good qualitative agreement between the predicted and measured azimuthal velocity distributions.	Model still needs to be extended with a relation for the axial and radial velocity distribution.
Raestad (1996)	Tip vortex noise and acoustic pressure prediction.	Various twin-screw passenger vessels.	Experimental studies involving 15 vessels of various types.	Average error less than 25 [dB].	Acceptable qualitative agreement between predictions and experimental data of model scale tests and full-scale trials.
Okamura and Asano (1988)	Prediction of tonal and broadband noise from marine propellers.	Full-scale tests.	Cargo liner and training ship.	Average error less than 5 [dB] for tonal and less than 20 [dB] for broadband noise.	More detailed measurements of cavitation patterns required for validation.
Ekinci et al. (2010)	URN due to blade sheet cavitation prediction.	DTMB-4148, DTMB-4119, and Seium-maru HSP propellers.	Numerical results (Fraser, 1986).	Average error less than 25 [dB].	Models produces fast and reliable predictions but validation on additional propellers required.
Matusiak (1992)	Model the generation and collapse of free bubbles from sheet cavitation.	Sydney Express coastal service vessel.	Measurements from the Pasadena oil tanker.	Average error less than 20 [dB].	Further validation on highly-skewed propellers is required together with extensions to include viscous effects.

same approach was exploited by Lidtke et al. (2019) for the E779 A propeller with the purposed to explore the crucial solution parameters for the prediction of the propeller URN. The authors concluded that larger porous data surface in the coarse grid region might risk information loss due to discretisation errors and dissipation.

Finally, it is worth mentioning an interesting comparison conducted by Testa et al. (2008) between DES and BEM solvers coupled with a porous FHW equation for the E779 A propeller under uniform and cavitating conditions. The authors concluded that the BEM solver was computationally cheap but inadequate to reflect the effects of non-linear noise sources as moving downstream of the propeller.

For the sake of completeness and readability, Table 3 provides an overview of CFD-based state-of-the-art methods for propeller URN prediction reporting for each one of the considered works its scope, employed methods, case studies, and takeovers.

2.2. Data-driven models

DDMs have proven to be valuable instruments in several maritime applications (Miglianti et al., 2020; Coraddu et al., 2020, 2019; Cipollini et al., 2018; Coraddu et al., 2017; Wang et al., 2021; Vesting and Bensow, 2014; Coraddu et al., 2021b; Sha et al., 2022; Banan et al., 2020; Shao et al., 2021; Fan et al., 2020), and have also been employed by several researchers in propeller design and analysis.

Koushan (2000) was among the first to explore the use of DDMs in propeller hydrodynamics, developing a Neural Network (NN) to predict propeller induced pressure pulses utilising 470 model-scale tests. The authors reported results of satisfactory accuracy with an average error below 20% for the predictions of the 1st and 2nd harmonic pressure pulse coefficient. They further commented that the use of DDMs can accelerate the propeller design loop and help researchers develop highly efficient unconventional designs.

NN were also exploited by Roddy et al. (2008) for the prediction of propeller forces and moments during crash-back maneuvers. Utilising 155 experiments, the authors demonstrated that NN can estimate propeller forces and moments quite accurately, with a determination coefficient ranging in 0.95÷0.99. The authors also commented on the potential of NN as a valuable tool for large-scale maneuvering simulation.

DDMs for propeller design optimisation exploiting NN was explored by Calcagni et al. (2010). Utilising a dataset of BEM simulations and experimental data, they were able to build a model for propeller thrust and torque coefficients prediction with an average error smaller than 4%. The authors primarily commented on the computational advantages of using NN in large-scale performance optimisation studies. In fact, they were able to successfully optimise two unconventional propeller geometries, used in high-speed vessels, with minimal time requirements.

Similarly to Calcagni et al. (2010), Jiang et al. (2021), employed a NN-based URN predictor coupled with a genetic optimisation algorithm to find a propeller designs with minimal low-frequency discrete spectrum thrust. Leveraging 336 thousand simulations, they were able to build a NN showing average error below 0.01% for the thrust coefficient and below 0.2% and 5.7% for the and the 1st and 2nd order discrete spectrum thrust respectively. Exploiting the NN-based predictor, the authors obtained impressive reductions in computational requirements (150 times faster) with respect to using a CFD-based predictor.

A quantitative comparison between NN- and Kriging-based predictor, to be used in propeller design optimisation, was reported by Vesting and Bensow (2014). Authors first exploited a total of 350 numerical simulations obtained by the application of the Vortex Lattice Method (VLM) of He et al. (2011), Sun et al. (2004) to train a NN- and a Kriging-based predictor. Then they used them to predict propeller dynamics and the necessary cavitation constraints needed for the design optimisation process. The NN-based models showed 1% of average error while the Kriging-based were able to achieved smaller average error ranging in 0.3÷0.9%. The authors concluded that DDMs can facilitate rapid convergence and greatly reduce computational requirements of large-scale optimisation processes.

Table 3
Propeller URN prediction CFD-based models review.

Ref.	Scopes	Methods	Case studies	Takeovers
Salvatore and Ianniello (2003)	Comparison of noise waveforms using FWH or the Bernoulli equation.	BEM and FWH or Bernoulli equation.	DTRC-4148.	Waveforms are fully comparable and both models identify the major features of cavitation induced propeller radiated noise generation. Discrepancies are present even at non-cavitating flows.
Seol et al. (2002)	Investigation of the effects of duct geometry on overall URN considering sound reflection and scattering effects.	BEM and FWH.	DTMB-4119 and KA4-70.	Noise generated by a marine propeller under non-cavitating conditions has a long fundamental wavelength and the effect of the duct is not so important at the far-field from the viewpoint of acoustic performance.
Seol et al. (2005)	Investigation of URN levels and noise directivity patterns of various noise sources under the presence of sheet cavitation.	BEM and FWH.	DTMB-4119.	In non-uniform flows noise, directivity patterns are a direct result of dipole noise under non-cavitating flows.
Ianniello et al. (2013)	Test of the versatility and effectiveness of the numerical method in predicting URN from non-cavitating propellers.	RANS and FWH.	INSEAN-E779A.	Due to the relevant acoustic role played by the vorticity and turbulence, RANS simulations become inadequate for hydroacoustic purposes, especially at measurement points where the turbulent fluctuating component of the velocity field becomes relevant.
Ianniello et al. (2014a)	URN Prediction for a scaled ship model in a steady course.	RANS and FWH.	Scaled patrol boat model.	Results demonstrate the effectiveness and robustness of the numerical method to assess URN and to identify the main generating noise mechanisms.
Ianniello et al. (2014b)	URN prediction for a large ROPAX.	RANS and FWH.	ROPAX vessel.	The numerical method can be used as a standard approach to provide hydroacoustic characterisation during the design stage.
Ianniello and De Bernardis (2015)	Demonstration of the potential of the numerical method for hydroacoustic analysis.	RANS or DES and FWH	INSEAN-E779A	Reliable hydroacoustic analysis of a marine propeller requires the computation of the nonlinear quadrupole sources and cannot avoid an accurate estimation of the three-dimensional turbulence and vorticity fields.
Lidtke et al. (2019)	Exploration of the sensitivity of the FWH for defining the porous data surfaces and key simulation parameters such as time step and grid resolution.	RANS or DES and FWH.	INSEAN-E779A.	Particular care must be adopted while defining the porous data surfaces to minimise the amount of upstream vorticity penetrating them while ensuring the effect of noise-generating flow features is aptly captured.
Testa et al. (2008)	Comparison of FWH and the Bernoulli equation for the evaluation of far-field pressure distribution from cavitating marine propellers.	BEM and FWH or Bernoulli equation.	INSEAN-E779A	Numerical results demonstrate the superiority of the FWH and support its usage for naval applications of hydroacoustics.

Gaggero et al. (2021) demonstrated the feasibility of a multi-fidelity and multi-objective design optimisation of the E779 A propeller (Salvatore et al., 2009) employing Kriging-based prediction models. The letters were developed exploiting data coming from different low-fidelity BEM and high-fidelity RANS (Siemens Digital Industries Software, 2021) simulations. Predictions were able to achieve average error ranging in 0.3÷5.8% on 160 BEM simulations and average error ranging in 0.06÷4.8% when exploiting data from 320 BEM and 40 RANS simulations.

URN prediction capabilities of NN were also studied by Wang et al. (2021), where they were exploited to predict the Sound Pressure Levels (SPLs) at the first 3 blade passing frequencies exploiting a dataset of 3098 hydroacoustic simulations. The authors experimented with various NN architectures considering a variety of interpolation and extrapolation scenarios. The best performing NN architecture was able to achieve an average error of 7.8±1.0 [dB] across all targets.

For the sake of completeness and readability, Table 4 provides an overview of state-of-the-art DDMs for propeller noise prediction, reporting for each one of the considered works, the scopes, the methods, the inputs and the outputs of the models, the data exploited to conduct the study, and the final performance. DDMs also appear as components of HMs, and are therefore included in all of the studies involving HMs. For the sake of clarity, these are further discussed in Section 2.3.

2.3. Hybrid models

HMs are able to exploit the physical knowledge of the phenomenon and available historical data, to deliver both accurate and physically plausible results, usually surpassing the performance of both PM or DDM (Coraddu et al., 2018, 2021a, 2022).

Recently, HMs have also been proposed for propeller URN prediction, and their effectiveness has been demonstrated in several studies (Cipollini et al., 2019b,a; Miglianti et al., 2019; Oneto et al., 2020; Miglianti et al., 2020). A seminal work was the one of Cipollini et al. (2019b) aiming at predicting the cavitating vortex frequency and its corresponding SPLs. The authors exploited a dataset of 164 cavitation tunnel model scale tests, a Kernel Ridge Regression (KRR) (Hainmueller and Hazlett, 2014) based DDM, and a semi-empirical (Bosschers, 2018a; Raestad, 1996) based PM. Combining the DDM and PM they demonstrated that the resulting HM has high potential in predicting vortex noise. This study was further extended by Cipollini et al. (2019a) demonstrating the superiority of HMs over DDMs and PMs in various interpolation and extrapolation scenarios. Subsequently, Miglianti et al. (2019) proposed HMs to predict an artificial simplification of URN spectra. In particular, they developed 3 different modelling approaches: A PM, combining BEM simulations (Gaggero et al., 2013, 2016) with the Empirical Tip Vortex (ETV) model (Bosschers, 2018a,c), a DDM based on KRR, and a Multi Task Learning (Baxter, 2000; Caruana, 1997) based HM. An extensive experimental analysis conducted on a dataset

Table 4
Propeller URN prediction DDMs review.

Ref.	Scopes	Methods exploited	Inputs	Outputs	Data	Performance
Koushan (2000)	Propeller-induced pressure pulse prediction.	NN.	16 wakefield quantities and 19 propeller geometry quantities.	1st and 2nd order harmonic pressure pulse coefficients.	Experimental data of 470 model-scale tests.	4% of average error for the 1st order coefficient and 20% of average error for the 2nd order coefficient.
Roddy et al. (2008)	Propeller force and moment prediction.	NN.	16 variables describing propeller geometry and submarine control.	Forces and moments in 6 degrees of freedom.	155 experiments of submarine model tests.	Determination coefficient in 0.95÷0.99.
Calcagni et al. (2010)	Hydrodynamic performance prediction.	NN.	Number of propeller blades, blade area ratio, pitch ratio, and advance coefficient.	Thrust, torque, and open water efficiency.	Wageningen-B series experiments and BEM simulations.	Average error below 4%
Jiang et al. (2021)	Discrete spectrum thrust prediction.	NN.	Coefficients of a 4th order polynomial describing the propeller skew distribution.	Thrust coefficient and 1st and 2nd order thrust coefficients.	336,000 BEM simulations.	Average error below 0.01% for the thrust coefficient and in 0.21÷5.71% for the 1st and 2nd order thrust coefficients.
Vesting and Bensow (2014)	Hydrodynamic performance prediction.	NN and Kriging.	11 variables for propeller geometry, inflow conditions. and cavitation.	4 cavity characteristics of the "key" blade.	350 VLM simulations	Average error around 1% for NN and in 0÷0.9% for Kriging.
Gaggero et al. (2021)	Optimisation of the hydrodynamic performance	Kriging.	Pitch distribution.	Thrust coefficient and 4 pressure coefficients.	40–640 BEM simulations and 4–40 RANS simulations.	Average error in 0.3÷5.8% on BEM and 0.06÷4.8% on RANS simulations.
Wang et al. (2021)	URN prediction of blade passing frequencies.	NN.	15 variables describing the propeller geometry and inflow conditions.	1st, 2nd, and 3rd blade passing frequency URN levels.	3098 BEM–FWH simulations.	Average error 7.8±1.0 [dB].

of 425 cavitation tunnel model scale tests allowed the authors to prove that HMs can remarkably improve the original DDMs or PMs.

Oneto et al. (2020) employed HMs to predict cavitating vortex frequency and its corresponding SPL, utilising Deep Learning (Goodfellow et al., 2016) based DDMs and HMs on a subset of experiments of Miglianti et al. (2019). The authors demonstrated that these new DDMs and HMs outperform the corresponding models presented by Miglianti et al. (2019). In fact, they are able to effectively take in to account several high-dimensional quantities describing the underlying phenomena (i.e., surface pressure distributions over the blades, and bound circulation). Nevertheless, authors notices how models performance need to be tested on larger datasets of geometries and for challenging real world extrapolation scenarios.

To the best of the authors' knowledge, the latest results in HMs for URN prediction are the ones of Miglianti et al. (2020), which accounts for the extensions recommended by Oneto et al. (2020), and further predicts the simplification of the URN spectra proposed by Miglianti et al. (2019), instead of the cavitating vortex frequency and its corresponding SPL. In this study, the authors utilised a subset of experiments from the dataset of Miglianti et al. (2019) for the development of the DDMs and HMs, enriched with several quantities extracted from a BEM computational model. The authors demonstrated that HMs are able to outperform both DDMs and PMs in all interpolation and extrapolation scenarios considered. Still, limitations in the cardinality and variety of geometries in the dataset are present.

For the sake of completeness and readability, Table 5 provides an overview of the state-of-the-art HMs for propeller URN prediction reporting for each one of the considered works, the scopes, the methods, the inputs and the outputs of the models, the data exploited to conduct the study, and the final performance.

3. Dataset description

In order to test our proposal, we will leverage the dataset collected by Aktas (2017), Aktas et al. (2018), which we will describe in this section for the sake of completeness. We provide all details that would allow the full reproduction of our work on any similar dataset, and note that any dataset containing experiments from a cavitation tunnel, university, research institution, or private company could have been employed.

The authors of Aktas (2017), Aktas et al. (2018) performed an extensive measurement campaign conducting a series of cavitation tunnel tests at the Emerson Cavitation Tunnel of the Newcastle University (Atlar, 2011), with 6 members of the Meridian Standard propeller series (Emerson, 1978) and 3 wakefields.

The Meridian propeller series, derived from the proprietary propeller design of the Stone Manganese Marine Ltd., is an unique standard series based solely on practical propeller designs for standardised variations in pitch-to-diameter ratio P/D, blade area ratio A_E/A_O, and number of blades Z, where P is the propeller pitch, D is the propeller diameter, A_E is the expanded area of the propeller, and A_O is the propeller disk area. Initially, the series comprised 4 parent models having a combination of $A_E/A_O \in \{0.45, 0.65, 0.85, 1.05\}$ and $P/D \in \{0.4, 0.6, 0.8, 1.0, 1.2\}$. All propellers had $D = 304.8$ [mm] and $Z = 6$ blades with a hub diameter of $D_h = 0.185D$. Currently, there are 60 propellers in the series, with $Z \in \{4, 5, 6\}$, $A_E/A_O \in [0.45, 1.05]$ and $P/D \in [0.4, 1.2]$ (Carlton, 2018). The subset of this series exploited in this work includes the propeller models King's College-D (KCD)-65, KCD-74, KCD-129, KCD-191, KCD-192, and KCD-193 represented in Fig. 1.

The wakefields were selected based on the criteria suggested by Konno et al. (2002), Angelopoulos et al. (1988), Odabasi and Fitzsimmons (1978), where it was observed that steeper velocity changes on the wakefield produce more severe pressure fluctuations, and contribute to the formation of a higher number of cavities (Aktas et al., 2018). Based on these observations, the wake non-uniformity, mean wake, half-wake width, and wake depth were controlled to generate 3 wakefields that would induce variation in the inflow velocities of varying severity. These changes would subsequently induce the formation of unsteady cavitation from the collapse and rebound of cavity volumes at the exit of the wake peak region. Fig. 2 provides a visual representation of the wakefields, which are referred to as W1, W2, and W3.

Based on these propeller geometries and wakefields, a full factorial experimental design was conducted, including 3 levels of tunnel vacuum pressures (150 [mmHg], 300 [mmHg], and 760 [mmHg]) and 8 propeller rotational speeds (600 [rpm], 800 [rpm], 1000 [rpm], 1200 [rpm], 1400 [rpm], 1500 [rpm], 1750 [rpm], and 2000 [rpm])

Table 5
Propeller Noise Prediction HMs review.

Ref.	Scopes	Methods exploited	Inputs	Outputs	Data	DDM Performance	HM Performance
Cipollini et al. (2019a)	Cavitating vortex frequency and SPL prediction.	KRR.	39 quantities for propeller geometry, inflow conditions, cavitation.	Cavitating vortex frequency and SPL.	164 cavitation tunnel tests.	Errors ranging between 10.9÷18.1% for interpolation, and 17.6÷25.1% for extrapolation.	Errors ranging between 6.4÷10.7% for interpolation, and 7.7÷12.6% for extrapolation.
Miglianti et al. (2019)	Simplified spectra URN prediction.	KRR.	39 quantities for propeller geometry, inflow conditions, cavitation.	3 2-dimensional key quantities of the URN spectrum.	425 cavitation tunnel tests.	Errors ranging between 5.3÷140.1% for interpolation, and between 15.2÷354.4% for extrapolation.	Errors ranging between 4.3÷10.1% for interpolation, and between 6.1÷12.6% for extrapolation.
Oneto et al. (2020)	Cavitating vortex frequency and SPL prediction.	NN.	39 quantities for propeller geometry, inflow conditions, cavitation.	Cavitating vortex frequency and SPL.	258 cavitation tunnel tests.	Errors approximately equal to 6% for all quantities predicted.	Errors ranging between 4.0÷4.7% for all quantities predicted.
Miglianti et al. (2020)	Simplified spectra URN prediction.	NN.	37 quantities for propeller geometry, inflow conditions, cavitation.	3 2-dimensional quantities of the URN spectrum.	258 cavitation tunnel tests.	Errors ranging between 2.9÷6.1% for interpolation, and between 5.8÷9.8% for extrapolation.	Errors ranging between 2.3÷4.3% for interpolation, and between 4.2÷8.1% for extrapolation.

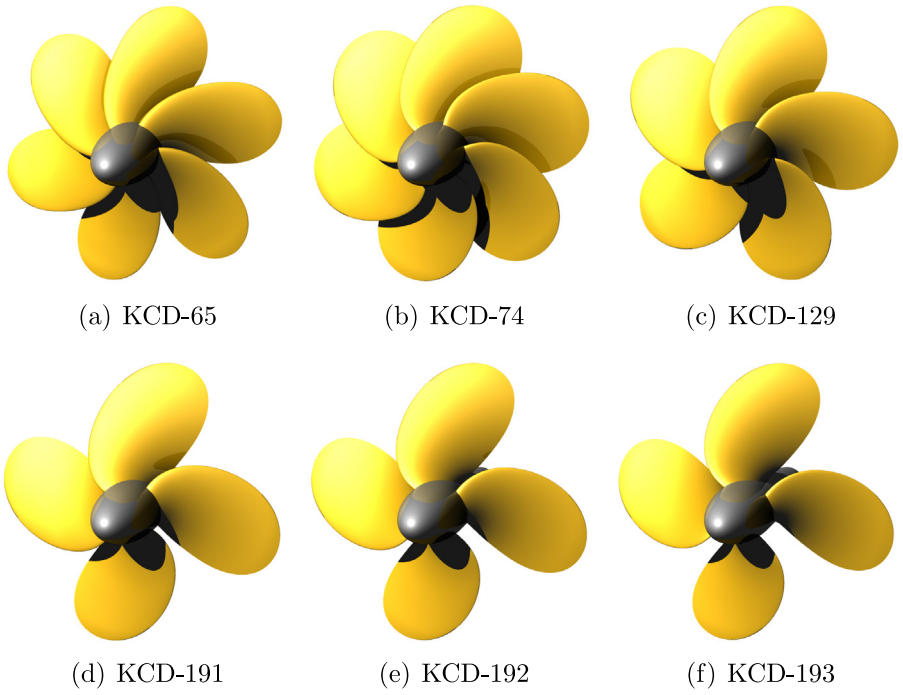


Fig. 1. Subset of the Meridian Standard propeller series exploited in this work.

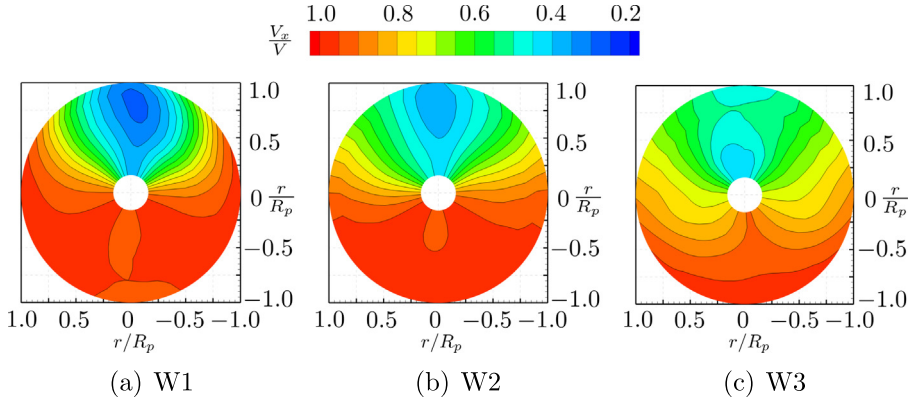
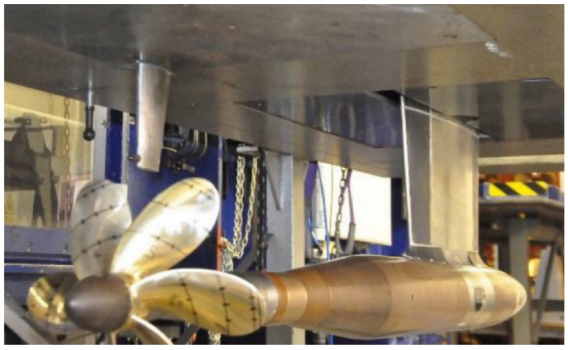


Fig. 2. Contour plots of axial velocity distributions of the dataset wakefields.

(a) Propeller behind a wake screen (Aktas, 2017).



(b) Location of the hydrophone (Aktas et al., 2018).

Fig. 3. Experimental Setup.

with a constant inflow velocity of 3 [m/s], for a total of 432 experiments (Aktas et al., 2018). The URN was measured by means of an hydrophone placed in the tunnel test section, measuring the fluctuations of the pressure over time. The pressure timeseries spectrum, computed via Fourier transformation, were subsequently converted in $^1/_3$ octave band Institute (2009), corrected for background noise, and converted to the standard measuring distance of 1 [m] according to the recommendations of International Towing Tank Conference Specialist Committee on Hydrodynamic Noise (2017). A visual representation of the experimental setup is reported in Fig. 3.

From this set of experiments we extracted a set of quantities (that we will exploit later as predictive features and targets to predict) that will compose our dataset (of 432 samples, one for each experiment).

For what concerns the targets, the final goal is to predict the Radiated Noise Levels (RNLs) in the form of $^1/_3$ octave band, namely 31 RNLs for 31 frequencies $\boldsymbol{f}$.

For what concerns the predictive features, namely, the quantity that we can extract at design stage to make the predictions, we rely on a set of quantities that provide a rich representation of the propeller geometry and the operational aspects that influence the generated noise, including cavitation phenomena, loading, and kinematic conditions. Some of these quantities can be easily and rapidly estimated via BEM.

In order to describe the propeller geometry, we rely on a combination of heterogeneous quantities. The first and most intuitive ones are D, Z, and A_E/A_O. Additionally, we took into account $\boldsymbol{P}$, chord $\boldsymbol{c}$, total rake $\boldsymbol{i}_T$, maximum thickness $\boldsymbol{t}_{\max}$, maximum camber $\boldsymbol{f}_{\max}$, and skew angle $\boldsymbol{\theta}_s$ for 8 radial sections from the root to the blade tip. Skew angle is defined as the rotation of the section seen in the transversal plane with respect to midchord measured from the vertical axis.

In order to describe the average operating conditions, we included the propeller's rotational speed n_p, its advance velocity V_a, the static pressure of the tunnel p_{rel}, the advance ratio J, thrust and torque coefficients K_t, K_q, and the propeller efficiency η_o (Carlton, 2018). The wakefields are defined as $\boldsymbol{w}_a = 1 - \frac{V_a}{V}$, with V being the undisturbed flow velocity. These quantities are usually known during the propeller design process (Miglianti et al., 2019) and are represented by two-dimensional tensors of the axial wake distribution in polar coordinates, provided for 22 radial sections and 60 angular positions. In order to indicate the presence of cavitation, we consider 4 different definitions of the cavitation index (Carlton, 2018). First σ_v and σ_n, that refer to the cavitation numbers based on V_a and n_p respectively. Then $\sigma_{\text{v,tip}}$ and $\sigma_{\text{n,tip}}$ that refer to the same values resulting at the tip and also with respect to the static pressure at the tip when the blade is in the upright position. In order to describe the hydrodynamic functioning of the blade sections, we also include the geometric angle of attack $\boldsymbol{\alpha}_G$ in the dataset, evaluated as the difference between the advance angle $\boldsymbol{\theta}_a$ of a blade section and the local pitch angle $\boldsymbol{\theta}_p$ according to International Towing Tank Conference Propulsion Committee (2008)

$$\begin{aligned} \alpha_G(r,\theta) &= \theta_p(r) - \theta_a(r,\theta), \\ \theta_a(r,\theta) &= \arctan\left[\frac{V_a\left(1-w_a(r,\theta)\right)}{2\pi n r}\right], \end{aligned} \tag{1}$$

with r and θ being the radial and angular position of the blade section.

Finally, we estimate 2 tensors with a computationally inexpensive BEM briefly recalled in Section 4.1, extensively described in Kalikatzarakis et al. (2021), and verified in many works (Gaggero and Brizzolara, 2009; Gaggero et al., 2013, 2016; Gaggero and Villa, 2017, 2018; Gaggero et al., 2010, 2014, 2019). These quantities include the distribution of the pressure coefficient $\boldsymbol{C}_{\text{pn}}$ for 44 chord-wise locations, 22 radial sections and 60 angular positions, as well as the unsteady radial circulation distribution $\boldsymbol{C}$ for the same radial sections and angular positions. $\boldsymbol{C}_{\text{pn}}$ is directly related to the occurrence of blade surface cavitation and can provide an estimation of cavitation inception (see Section 4.2.1). $\boldsymbol{C}$ is related to the forces acting on the hydrofoil, and in particular to the lift, according to the well-known Kutta–Joukowski theorem (Katz and Plotkin, 2001) and correlates with the load acting on the blades and its distribution, which in turn is strictly related to the strength of the shed vortices and the occurrence of vortex cavitation. By considering all these quantities, we can fully characterise the dynamic pressure acting on the whole surface of the blades during their rotation in any wakefield (Miglianti et al., 2020).

A summary of the quantities just described composing our dataset is reported in Table 6. Note that data have not been made publicly available due to confidentiality constraints.

4. Methodology

This section is devoted to the description of the proposed methodology with all the details needed to reproduce this study and is organised as follows. Authors first recall the PM (Section 4.1), then we exploit this PM to empower state-of-the-art DDMs (Section 4.2) to build novel HMs (Section 4.3). In particular, we will leverage the domain knowledge both for the DDMs (to design a domain-knowledge enriched feature set) and HMs (to exploit the knowledge distilled from the PM) and the pipeline of the different approaches is summarised in Section 4.4. , we describe the methods to tune and assess the performance of the methods in different interpolating and extrapolating scenarios, as reported in Section 4.5. Finally, the physical plausibility of the generated model will be tested according to what is described in Section 4.6.

Table 6
Quantities available in the dataset.

Targets to predict			
Symbol	Description	Size[a]	Units
RNLs	RNLs in 1/3 octave band	31	[dB]
Input features for the prediction			
Symbol	Description	Size[a]	Units
Propeller geometry			
D	Propeller diameter		[m]
Z	Number of blades		[–]
A_E/A_O	Blade aspect ratio		[–]
$\boldsymbol{p}$	Sectional pitch	8	[m]
$\boldsymbol{c}$	Sectional chord	8	[m]
$\boldsymbol{i}_T$	Sectional total rake	8	[deg]
$\boldsymbol{t}_{\max}$	Max. sectional thickness	8	[m]
$\boldsymbol{f}_{\max}$	Max. sectional camber	8	[m]
$\boldsymbol{\theta}_s$	Sectional skew angle	8	[deg]
Operating conditions			
n_p	Propeller rotational speed		[rpm]
V_a	Advance velocity		[m/s]
p_{rel}	Tunnel pressure		[mbar]
$\boldsymbol{w}_a$	Axial wakefield	22 × 60	[–]
J	Advance coefficient		[–]
K_t	Thrust coefficient		[–]
$10K_q$	Torque coefficient		[–]
η_o	Propeller efficiency		[–]
σ_v	Cavitation index ref. on V_a		[–]
$\sigma_{\text{v,tip}}$	Cavitation index ref. on $\sqrt{V_a^2+(\pi n_p D)^2}$ at blade tip		[–]
σ_n	Cavitation index ref. on n_pD		[–]
$\sigma_{\text{n,tip}}$	Cavitation index ref. on n_pD at blade tip		[–]
$\boldsymbol{\alpha}_G$	Geometric angle of attack	22 × 60	[deg]
Estimated quantities from BEM simulations			
$\boldsymbol{C}$	Blade circulation	22 × 60	[–]
$\boldsymbol{C}_{pn}$	Pressure coefficient	44 × 22 × 60	[–]

[a]Empty field indicates scalar quantity.

4.1. Physical models

The exploited PMs have been verified and validated by Kalikatzarakis et al. (2021), and for the sake of completeness, here we will report a summary of the method.

First, to accurately characterise the propeller's hydrodynamic field, unsteady hydrodynamic computations have been performed through an in-house developed BEM code (Gaggero et al., 2010; Gaggero and Villa, 2018). The validity of our implementation for cavitating flows has been demonstrated in the past for steady (Gaggero and Villa, 2017) and unsteady conditions (Gaggero and Villa, 2018), for strongly non-homogeneous wakefields (Brizzolara et al., 2008; Gaggero et al., 2014), for very off-design conditions (Gaggero et al., 2019), and for a variety of propeller types, including supercavitating (Gaggero and Brizzolara, 2009), ducted (Gaggero et al., 2013), and tip loaded propellers (Gaggero et al., 2016).

Then, to estimate the broadband effects of sheet cavitation, we implemented the approach suggested by Matusiak (1992), which considers as inputs the blade pressure distribution computed by the BEM code, along with 5 constants that can be estimated from available data. The model assumes that any change in the volume rate of the generated bubbles equals the rate at which the sheet cavity volume decreases and that the bubble size follows a β distribution. In combination with a bubble dynamics model (Brennen, 2014), this method can accurately estimate the high-frequency broadband spectral content of the URN due to propeller cavitation.

Finally, we utilise the ETV model proposed by Bosschers (2018a,c) to estimate the URN due to the TVC. In particular, the ETV is a semi-empirical model based on the Tip Vortex Index (TVI) method (Raestad, 1996), and relates the measured URN to the predicted size of the vortex cavity, using a computed circulation distribution on the propeller blade.

Note that this PM, as all PMs, has no hyperparameters to tune (since the model structure is given by the domain knowledge) and just parameters need to be tuned on the available data (see Table 8).

4.2. Informed data-driven models

In this section, we will present our proposal for an Informed DDMs, namely a pure shallow DDMs that leverage on the domain knowledge to design a rich and informative yet synthetic feature set.

In fact DDMs can be grouped in two main families: shallow and deep DDMs (Shalev-Shwartz and Ben-David, 2014; Goodfellow et al., 2016). Shallow DDMs usually require handcrafting, implicitly and/or explicitly, features to be able to achieve good recognition performance (Shalev-Shwartz and Ben-David, 2014; Duboue, 2020). Usually, this feature set is designed based on classical signal processing techniques (Duboue, 2020) and then enriched via ensemble (Breiman, 2001; Chen and Guestrin, 2016), kernel (Shawe-Taylor and Cristianini, 2004), random (Huang et al., 2011), or learned (Bishop et al., 1995) representations. Deep DDMs, instead, are able to automatically learn features directly from the data (Goodfellow et al., 2016) and over-perform state-of-the-art shallow models (and in some case also humans) in terms of recognition performance in many different applications (Cireşan et al., 2011; Hekler et al., 2019; Silver et al., 2017; Jumper et al., 2021; Grace et al., 2018). Unfortunately, Deep DDMs have also three main weaknesses. First, they require a huge number of samples to be trained. In our case, the dataset has less than 500 samples in the simplest scenario, namely interpolation, and much fewer samples in complex scenarios (see Section 4.5), namely extrapolation. The second problem is that Deep DDMs are very hard to interpret. It is complex to understand what they actually learned from the data, resulting in models not useful for practical applications, where insights on the problem need to be extracted (Molnar, 2020). Finally, Deep

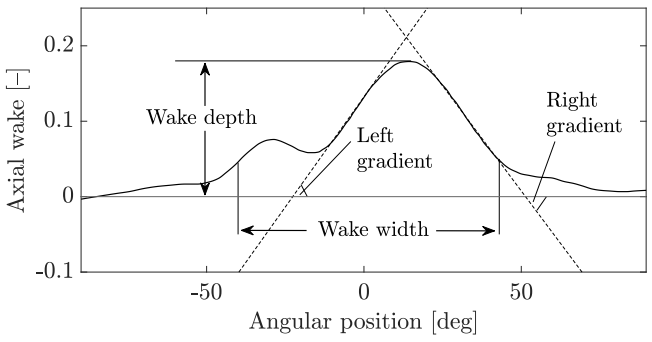


Fig. 4. Wake features defined by Odabaşi and Fitzsimmons (1978) and utilised in Miglianti et al. (2019).

DDMs are seldom able to give physically plausible prediction, see for example the well known problem of the adversarial samples (Biggio and Roli, 2018; Duan et al., 2020).

For these reasons in this work we will leverage on Informed Shallow DDMs and for this purpose we will present in Section 4.2.1 the proposed Informed Feature Engineering Phase and in Section 4.2.2 the different state-of-the-art Shallow DDMs that we will test since, a priori, it is not possible to choose a model (Adam et al., 2019) and different approaches need to be tested.

4.2.1. Informed feature engineering phase

The predictive features described in Table 6 are fundamental raw quantities to predict the URN SPLs. Nevertheless, as described above, DDMs can achieve their full potentials when these raw quantities are transformed in richer and informative, yet synthetic feature set, based on the domain knowledge.

In particular, some quantities have been kept from the original dataset (Table 6), while others ($\boldsymbol{w}_a$, $\boldsymbol{\alpha}_G$, $\boldsymbol{C}$, and $\boldsymbol{C}_{\text{pn}}$) need to be synthesised into a new set of features characterised by a more compact yet informative set of features.

Wakefield. For what concerns $\boldsymbol{w}_a$, we followed the approach of Odabaşi and Fitzsimmons (1978), as these features are known to provide a rich representation of the propeller inflow conditions (Carlton, 2018; Miglianti et al., 2019, 2020). The first quantity that we can extract from $\boldsymbol{w}_a$ is the average volumetric axial wake $\bar{w}_a$

$$\bar{w}_a = \frac{\int_{r_h}^{R} r \int_0^{2\pi} w_a(r,\theta) d\theta dr}{\pi \left(R^2 - r_h^2\right)}, \tag{2}$$

where r_h corresponds to the hub radius and R to the propeller radius. Furthermore, several quantities on two radial sections, $r/R = 0.7$, and $r/R = 0.9$, have been defined

- the left and right gradients of $\boldsymbol{w}_a$ w.r.t. the angular position $D_\theta w|_{\{07,09\}}^{\{+,-\}}$, which represent the rate of variation of blade loading during one revolution;
- the wake width $w_{\text{wd}\{07,09\}}$, which is the angular sector where the wake fraction is greater than 0.05, i.e. the sector where the axial velocity on the propeller plane is reduced by at least 5%;
- the wake depth $w_{\max\{07,09\}}$, which corresponds to the maximum value of $\boldsymbol{w}_a$ for a given radial section.

Fig. 4 provides a visual impression of these features for one radial section. In principle, all and any radial sections of a propeller can be utilised. Nevertheless, the ones described represent the best trade-off between: Richness of the features describing propeller behaviour under sheet cavitation and TVC (Carlton, 2018), and the cardinality of the feature set.

To further enrich the representation of $\boldsymbol{w}_a$, we exploit Fourier's theorem to decompose the total fluctuating component at any radial section into a finite set of sinusoidal components of various harmonic orders. We use the first 4 components of this decomposition, as they are sufficient to accurately describe $\boldsymbol{w}_a$ for the available experimental data. Using this basis, the general approximation of $\boldsymbol{w}_a$ at a particular propeller radius is given by

$$w_a(\theta)|_r = \sum_{k=0}^{4} a_{w,k|_r} \cos\left(\frac{k\theta}{2\pi}\right) + b_{w,k|_r} \sin\left(\frac{k\theta}{2\pi}\right), \tag{3}$$

with $a_{w,k|_r}, b_{w,k|_r}$ being the Fourier coefficients of order $k = \{1,\dots,4\}$ that have been utilised as additional features.

Angle of attack. The angle of attack is another important aspect with respect to cavitation occurrence, and more specifically the presence of sheet cavities on the suction sides of the blades (Carlton, 2018). The evaluation of the angle of attack requires the propeller's self-induced velocities to be known. These can be straightforwardly evaluated using a BEM or lifting surface numerical code. However, as the information it contains is implicitly included in $\boldsymbol{w}_a$ (Miglianti et al., 2019), the geometric angle of attack is used instead, for simplicity.

We extract a set of features from $\boldsymbol{\alpha}_G$ that is similar to $\boldsymbol{w}_a$. For the same radial locations of the blade, namely at the 70% and the 90% of the propeller radius, we computed

- the average value of $\boldsymbol{\alpha}_G$, denoted as $\bar{\alpha}_G$;
- its minimum value $\alpha_{G_{\min}}$ and the corresponding angular location $\theta_{\min \alpha_G}$;
- its maximum value $\alpha_{G_{\max}}$ and the corresponding angular location $\theta_{\max \alpha_G}$;
- the coefficients $(\boldsymbol{a},\boldsymbol{b})|_{\alpha_G}$ resulting from the Fourier analysis.

Propeller blade pressure distribution. The pressure distribution evaluated with BEM can approximate the presence of cavitation under the assumption that cavitation occurs when the opposite of the local pressure coefficient is higher than the cavitation index at a given operating condition, i.e., when the local pressure is lower than the vapour pressure (Gaggero and Villa, 2017). This cavitation inception criterion has been successfully applied for the design of conventional and unconventional propellers, and it allows estimating the occurrence of cavitation on the blades, its location, and to guess its extent (Gaggero and Villa, 2017; Gaggero, 2020). With these considerations in mind, in order to estimate the region where true cavitation starts from $\boldsymbol{C}_{\text{pn}}$, we evaluate the blade areas $\boldsymbol{A}_c|_{(s)}$, $\boldsymbol{A}_c|_{(p)}$, for which the pressure is lower than the vapour pressure. Subsequently, from these two vectors we further compute

- the 4th order Fourier coefficients $(\boldsymbol{a},\boldsymbol{b})|_{A_c|_{(s)}}$ and $(\boldsymbol{a},\boldsymbol{b})|_{A_c|_{(p)}}$;
- the minimum and maximum areas encountered $A_{\min c}|_{(s)}$, $A_{\min c}|_{(p)}$, $A_{\max c}|_{(s)}$, and $A_{\max c}|_{(p)}$;
- their corresponding angular positions $\theta_{\min A_c|_{(s)}}$, $\theta_{\min A_c|_{(p)}}$, $\theta_{\max A_c|_{(s)}}$, and $\theta_{\max A_c|_{(p)}}$.

Moreover, we split each side of the blade into the following 4 panels, according to Fig. 5

- Panel 1 ($P1$): From blade root to $r/R = 0.7$, and from the leading edge to 20% of the chord;
- Panel 2 ($P2$): From blade root to $r/R = 0.7$, and from 20% to 60% of the chord;
- Panel 3 ($P3$): From $r/R = 0.7$ to blade tip, and from the leading edge to 20% of the chord;
- Panel 4 ($P4$): From $r/R = 0.7$ to blade tip, and from 20% to 60% of the chord.

For panels $P1$ and $P3$ on the pressure side of the blade, and panels $P1$, $P2$, $P4$, and $P4$ on the suction side, we evaluate the minimum value of $\boldsymbol{C}_{\text{pn}}$ for each angular position of the key blade, and subsequently compute the 4th order Fourier coefficients.

Blade circulation. Finally, from $\boldsymbol{C}$, we evaluate the strength of the vortex shed in the wake $\boldsymbol{\Gamma}_{0.95}$ at $r/R = 0.95$, which is proportional to the cavitating tip vortex occurrence (Brennen, 2014), for every angular position of the key blade. From the resulting vector, we subsequently compute

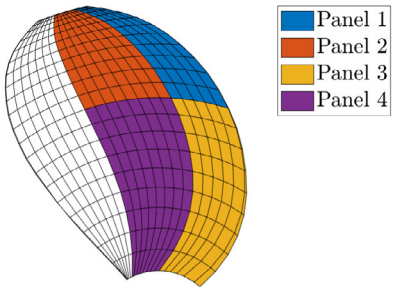


Fig. 5. Blade subdivision in panels.

- its minimum value $\Gamma_{\min 095}$ and the corresponding angular location $\theta_{\min \Gamma_{095}}$;
- its maximum value $\Gamma_{\max 095}$ and the corresponding angular location $\theta_{\max \Gamma_{095}}$;
- the coefficients $(\boldsymbol{a}, \boldsymbol{b})|_{\Gamma_{095}}$ resulting from a Fourier analysis.

Final predictive features set. The new set of predictive features coming from Table 6 according to what we described in this section is summarised in Table 7.

4.2.2. Shallow data-driven models

After the Informed Feature Engineering phase described in Section 4.2.1, we applied a series of top learning algorithms for regression. In fact, the problem of predicting the URN RNLs in 1/3 octave band based on the predictive features reported in Table 7 can be mapped to a typical multi-output regression problem in Machine Learning (Shalev-Shwartz and Ben-David, 2014; Shawe-Taylor and Cristianini, 2004).

The no-free-lunch theorem (Adam et al., 2019) ensures us that, in order to find the best algorithm for a particular application, it is necessary to test multiple algorithms. In our case, we will test 4 state-of-the-art algorithms[2] (Fernández-Delgado et al., 2014; Wainberg et al., 2016): Random Forests (RF) (Breiman, 2001), XGBoost (Chen and Guestrin, 2016), Kernel Ridge Regression (KRR) (Shawe-Taylor and Cristianini, 2004), and a Single Layered Neural Network (SNN) (Bishop et al., 1995).

In RF we need to tune the number of features to randomly sample from the whole features during each node of each tree creation n_f and the maximum number of elements in each leaf of each tree n_l. As RF performance improves increasing the number of trees n_t we set it to 1000 as a reasonably large number but yet computationally tractable.

In XGBoost, we need to tune the learning rate of the gradient l_r, the max dept of each tree n_d, the minimum loss reduction m_l, number of points to randomly sample from the whole training set for each tree creation n_b, and the number of features to randomly sample from the whole training set during the creation of each node for each tree n_f.

In KRR we chose to rely on the Gaussian kernel for the reason described in Keerthi and Lin (2003), and then the regularisation hyperparameter λ and the kernel coefficient γ need to be tuned.

In SNN, we use the sigmoid activation function in the hidden layer and the linear activation in the output layer, and we used the ADAM optimiser (Bishop et al., 1995). Then we need to tune the learning rate l_r, the number of hidden neurons h_l, the weight decay w_d, and the dropout rate in the hidden layer d_r.

The summary of these hyperparameters with the associated search space is reported in Table 8.

As we will see in Section 5, KRR will result to be the best performing algorithm for this application.

[2] Results in Kaggle www.kaggle.com, the most popular Machine Learning competition website, shows that these algorithms are the top winners.

4.3. Hybrid models

Current, all HMs rely on the same idea (see Section 2): correct/improve the results of the PMs exploiting the available data or improve the results of the DDMs encapsulating the domain knowledge into the DDMs (i.e., via feature engineering or via DDMs functional form modifications). In this paper, we consider the first approach since the second one, according to the point of view of the authors, can be inserted under the umbrella of fine tuned PMs (i.e., PMs whose parameters are tuned with available data) of informed DDMs (i.e., features engineered or enriched with domain knowledge).

In the context of HMs (see Section 2), current strategies are quite basic: either the results of the PMs are added as new features for the DDMs, or the DDMs tries to learn, simultaneously, both the data and the output of the PMs (i.e., be as much as possible close to it).

In this work, we propose a novel and more elaborated approach. Since the final goal of the HMs is to correct the PMs outputs via DDM, we will use the idea of recursion where, for a series of iterations, the DDMs input features are enriched with the difference between the output of the PMs and the one of the DDMs itself.

More formally, let us define the PM as $P_M(X)$ where X is a subset of the features of Table 6 needed to predict the URN RNLs (see Section 4.1). Then the DDM of Section 4.2 is defined as $D_M(X)$ where X is the set of features of Table 7. Finally, the proposed HM $H_M(X) = D^r_M(X)$ where

$$\begin{aligned} &D^0_M(X) = D_M(X), \\ &D^{i+1}_M(X) = D_M([X, D^i_M(X) - P_M(X)]), \quad i \in \{1, \dots, r\} \end{aligned} \tag{4}$$

where r is the number of recursion iterations, an hyperparameter to be tuned (see Table 8). In plain English, the proposed HM, is the successive application of a DDM informed by the distance from the PMs and (implicitly) the accuracy of the previous application of the DDM.

4.4. Physical, data-driven, and hybrid models pipeline

In order to improve the readability and the clarity of the manuscript, we present in Fig. 6 a simplified pipeline of the proposed methodology for the exploited and proposed PMs (Section 4.1), DDMs (Section 4.2), and HMs (Section 4.3).

4.5. Scenarios and performance tuning/assessment

In our experiment, we will study three different extrapolating scenarios based on the intrinsic hierarchy of the dataset. This will allow us to understand the extrapolation ability and the robustness of the different models described in Sections 4.1, 4.2, and 4.3:

- Leave One Rotational Speed Out (LORSO): in this scenario the models have been trained with all data except the one referring to a particular rotational speed;
- Leave One Wakefield Out (LOWO): in this scenario the models have been trained with all data except the one referring to a particular wakefield;
- Leave One Geometry Out (LOGO): in this scenario the models have been trained with all data except the one referring to a particular propeller geometry.

Note how the scenarios are the extrapolating scenario increase in complexity from LORSO to LOWO and finally to LOGO.

We report in Fig. 7 a visual representation of the intrinsic hierarchy of the dataset and these three scenarios. In particular, in the figure, we highlighted data hidden from the training phase and exploited just for testing purposes.

What remains to be addressed is how to tune the hyperparameters of the PMs, DDMs, and HMs and how to assess the final performance (Oneto, 2020).

For what concerns the last point, the answer is easy. Based on the different scenarios (LORSO, LOWO, and LOGO) we have to split

Table 7
Predictive feature set engineered from the quantities of Table 6 based on the domain knowledge.

Symbol	Description	Size[a]	Units
Quantities retained from the original dataset of Table 6.			
D	Propeller diameter		[m]
Z	Number of blades		[–]
A_E/A_O	Blade aspect ratio		[–]
$\boldsymbol{p}$	Sectional pitch ratio	1×8	[–]
$\boldsymbol{c}$	Sectional chord ratio	1×8	[–]
$\boldsymbol{i}_T$	Sectional total rake ratio	1×8	[–]
$\boldsymbol{t}_{\max}$	Max. sectional thickness ratio	1×8	[–]
$\boldsymbol{f}_{\max}$	Max. sectional camber ratio	1×8	[–]
$\boldsymbol{\theta}_s$	Sectional skew angle	1×8	[deg]
n_p	Propeller rotational speed		[rpm]
V_a	Advance velocity		[m/s]
p_{rel}	Tunnel pressure relative to ambient pressure		[mbar]
J	Advance coefficient		[–]
K_t	Thrust coefficient		[–]
$10K_q$	Torque coefficient		[–]
η_o	Propeller efficiency		[–]
σ_v	Cavitation index ref. on V_a		[–]
$\sigma_{\text{v,tip}}$	Cavitation index ref. on $\sqrt{V_a^2 + (\pi n_p D)^2}$ at blade tip		[–]
σ_n	Cavitation index ref. on $n_p D$		[–]
$\sigma_{\text{n,tip}}$	Cavitation index ref. on $n_p D$ at blade tip		[–]
Quantities substituting $\boldsymbol{w}_a$ in the original dataset of Table 6.			
$\bar{w}_a$	Average $\boldsymbol{w}_a$		[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{w_{07}}$	4th order Fourier coefficients of w at 0.7R	1×9	[–]
w_{wd07}	Wakefield width at 0.7R		[deg]
$D_\theta w\|_{07}^{+}$	Left wake gradient at 0.7R		[deg]
$D_\theta w\|_{07}^{-}$	Right wake gradient at 0.7R		[deg]
$w_{\max 07}$	Maximum w at 0.7R		[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{w_{09}}$	4th order Fourier coefficients of w at 0.9R	1×9	[–]
w_{wd09}	Wakefield width at 0.9R		[deg]
$D_\theta w\|_{09}^{+}$	Left wake gradient at 0.9R		[deg]
$D_\theta w\|_{09}^{-}$	Right wake gradient at 0.9R		[deg]
$w_{\max 09}$	Maximum w at 0.9R		[–]
Quantities substituting $\boldsymbol{\alpha}_G$ in the original dataset of Table 6.			
$\bar{\alpha}_{G_{07}}$	Average $\boldsymbol{\alpha}_G$ at 0.7R		[deg]
$\alpha_{G_{\max 07}}$	Maximum $\boldsymbol{\alpha}_G$ at 0.7R		[deg]
$\alpha_{G_{\min 07}}$	Minimum $\boldsymbol{\alpha}_G$ at 0.7R		[deg]
$\theta_{\max \alpha_{G_{07}}}$	Angular position of $\alpha_{G_{\max 07}}$		[deg]
$\theta_{\min \alpha_{G_{07}}}$	Angular position of $\alpha_{G_{\min 07}}$		[deg]
$(\boldsymbol{a}, \boldsymbol{b})\|_{\alpha_{G_{07}}}$	4th order Fourier coefficients of $\boldsymbol{\alpha}_{G_{07}}$	1×9	[deg]
$\bar{\alpha}_{G_{09}}$	Average α_G at 0.9R		[deg]
$\alpha_{\max G_{09}}$	Maximum α_G at 0.9R		[deg]
$\alpha_{\min G_{09}}$	Minimum α_G at 0.9R		[deg]
$\theta_{\max \alpha_{G_{09}}}$	Angular position of $\max \alpha_{G_{09}}$		[deg]
$\theta_{\min \alpha_{G_{09}}}$	Angular position of $\min \alpha_{G_{09}}$		[deg]
$(\boldsymbol{a}, \boldsymbol{b})\|_{\alpha_{G_{09}}}$	4th order Fourier coefficients of $\boldsymbol{\alpha}_{G_{09}}$	1×9	[deg]
Quantities substituting $\boldsymbol{C_{pn}}$ in the original dataset of Table 6.			
$A_{\max c}\|_{(s)}$	Maximum $A_c\|_{(s)}$		[m]
$A_{\min c}\|_{(s)}$	Minimum $A_c\|_{(s)}$		[m]
$\theta_{\max A_c\|_{(s)}}$	Angular position of $A_{\max c}\|_{(s)}$		[deg]
$\theta_{\min A_c\|_{(s)}}$	Angular position of $A_{\min c}\|_{(s)}$		[deg]
$(\boldsymbol{a}, \boldsymbol{b})\|_{A_c\|_{(s)}}$	4th order Fourier coefficients of $A_c\|_{(s)}$	1×9	[m]
$A_{\max c}\|_{(p)}$	Maximum $A_c\|_{(p)}$		[m]
$A_{\min c}\|_{(p)}$	Minimum $A_c\|_{(p)}$		[m]
$\theta_{\max A_c\|_{(p)}}$	Angular position of $\max A_c\|_{(p)}$		[deg]
$\theta_{\min A_c\|_{(p)}}$	Angular position of $\min A_c\|_{(p)}$		[deg]
$(\boldsymbol{a}, \boldsymbol{b})\|_{A_c\|_{(p)}}$	4th order Fourier coefficients of $A_c\|_{(p)}$	1×9	[m]
$(\boldsymbol{a}, \boldsymbol{b})\|_{C_{\text{pn}}^{\text{P1}}\|_{(p)}}$	4th order Fourier coefficients of $C_{pn}^{(\text{LE,p,r})}$	1×9	[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{C_{\text{pn}}^{\text{P2,}}\|_{(p)}}$	4th order Fourier coefficients of $C_{pn}^{(\text{LE,p,t})}$	1×9	[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{C_{\text{pn}}^{\text{P1}}\|_{(s)}}$	4th order Fourier coefficients of $C_{pn}^{(\text{LE,s,r})}$	1×9	[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{C_{\text{pn}}^{\text{P2,s}}\|_{(s)}}$	4th order Fourier coefficients of $C_{pn}^{(\text{LE,s,t})}$	1×9	[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{C_{\text{pn}}^{\text{P3,s}}\|_{(s)}}$	4th order Fourier coefficients of $C_{pn}^{(\text{m,s,r})}$	1×9	[–]
$(\boldsymbol{a}, \boldsymbol{b})\|_{C_{\text{pn}}^{\text{P4}}\|_{(s)}}$	4th order Fourier coefficients of $C_{pn}^{(\text{m,s,t})}$	1×9	[–]
Quantities substituting $\boldsymbol{C}$ in the original dataset of Table 6.			
$(\boldsymbol{a}, \boldsymbol{b})\|_{\Gamma_{095}}$	3rd order Fourier coefficients of Γ_{095}	1×7	[m^2/s]
$\Gamma_{\max 095}$	Maximum Γ_{095}		[m^2/s]
$\Gamma_{\min 095}$	Minimum Γ_{095}		[m^2/s]
$\theta_{\max \Gamma_{095}}$	Angular position of $\max \Gamma_{095}$		[deg]
$\theta_{\min \Gamma_{095}}$	Angular position of $\min \Gamma_{095}$		[deg]

[a]Empty field indicates scalar quantity.

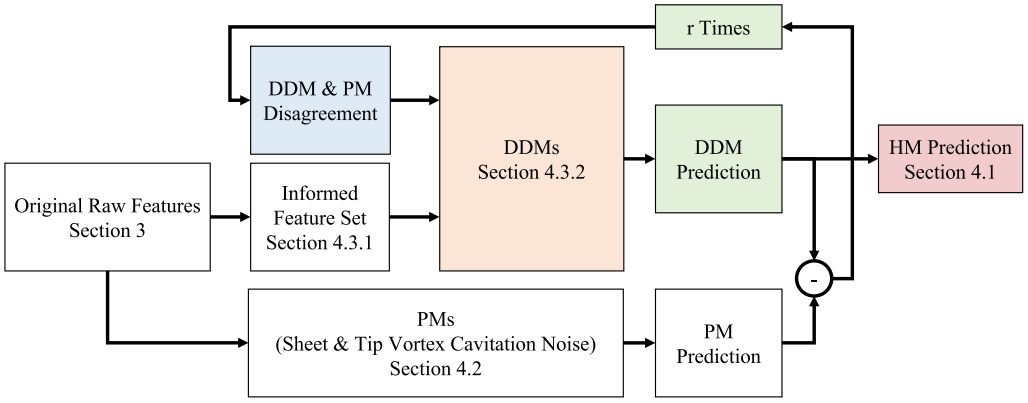
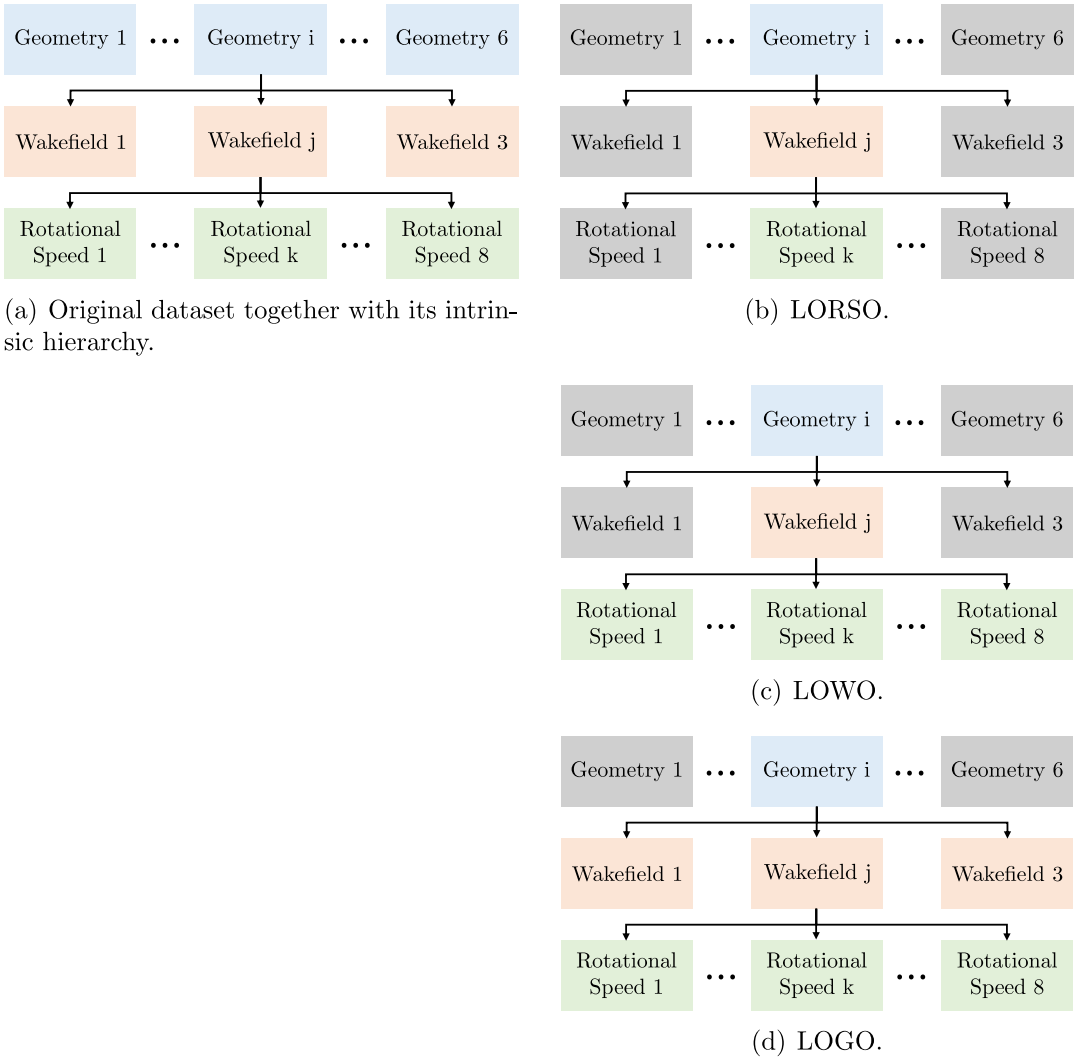


Fig. 6. Simplified pipeline of the proposed methodology for the exploited and proposed PMs (Section 4.1), DDMs (Section 4.2), and HMs (Section 4.3).



(a) Original dataset together with its intrinsic hierarchy.

(b) LORSO.

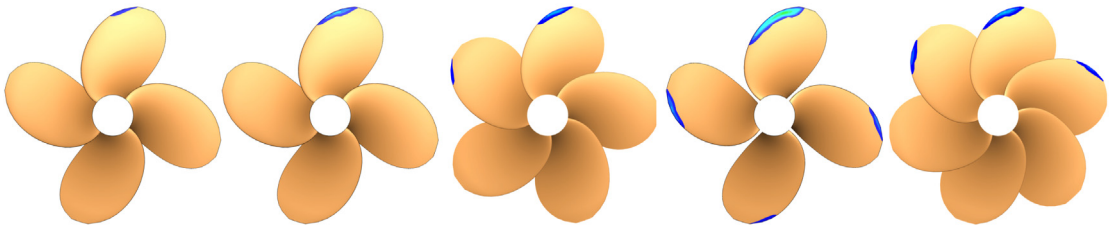
(c) LOWO.

(d) LOGO.

Fig. 7. Visual representation of the intrinsic hierarchy of the dataset and the three extrapolating scenarios. In particular we highlighted data hidden from the training phase and exploited just for testing purposes.

the data in Training $\mathcal{D}_n$ and Test $\mathcal{T}_t$ sets using the principle described above. Then we can use $\mathcal{D}_n$ to both train the model and select the best hyperparameters (architecture) of the shallow and deep models and used $\mathcal{T}_t$ to assess the performance of the final model. Repeating multiple times, this procedure will give us the average performance in the different scenarios. The performance will be measured in accordance with different metrics: three qualitative (the Mean Absolute Error – MAE – the Mean Absolute Percentage Error – MAPE – and the Pearson Product-Moment Correlation Coefficient - PPMCC) (Naser and Alavi, 2021) and one quantitative (the scatter plot actual versus predicted value) (Sainani, 2016).

Instead, for tuning the hyperparameters of the PMs, DDMs, and HMs we proceeded as follows. We took $\mathcal{D}_n$ and split it into Learning $\mathcal{L}_l$ and Validation $\mathcal{V}_v$ sets considering the LORSO, LOWO, or LOGO scenarios.



(a) Case C1. (b) Case C2. (c) Case C3. (d) Case C4. (e) Case C5.

Fig. 8. Cavitation occurrence computed by cavitating BEM for the case studies of Table 9.

Table 8
Hyperparameters and hyperparameters search space for all algorithms tested in this work, d denotes the number of features in the dataset.

Models	Algorithm	Hyperparameters
PMs	None	
DDMs	RF	$n_f : \{d^{1/3}, d^{1/2}, d^{3/4}\}$ $n_l : \{1, 3, 5, 10\}$ $n_t : \{1000\}$
	XGBoost	$l_r : \{0.01, 0.02, 0.03, 0.04, 0.05\}$ $n_d : \{3, 5, 10\}$ $m_l : \{0, 0.1, 0.2\}$ $n_b : \{0.6n, 0.8n, 1n\}$ $n_f : \{0.5d, 0.8d, 1d\}$
	KRR	$C : \{10^{-6}, 10^{-5.8}, \ldots, 10^{3}\}$ $\gamma : \{10^{-6}, 10^{-5.8}, \ldots, 10^{3}\}$
	SNN	$l_r : \{0.0001, 0.0005, 0.001, 0.005, 0.01\}$ $h_l : \{32, 64, 128, \ldots, 65536\}$ $w_d : \{10^{-6}, 10^{-5.8}, \ldots, 10^{3}\}$ $d_r : \{0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1\}$
HMs	The same of the DDMs	The same of the DDMs $r : \{1, 2, 4, 8\}$

Then we train each model with $\mathcal{L}_l$ with many different hyperparameters configurations and measure its performance on $\mathcal{V}_v$ according to the MAE. Then we repeated the experiment multiple times and selected the hyperparameters' configuration which gives the best average MAE on the validation sets. Finally, we retrained the model with the selected best configuration of the hyperparameters on the whole $\mathcal{D}_n$ which is the model that will be used for testing purposes (see the previous paragraph).

4.6. Physical plausibility

As discussed in the introduction, the purpose of this work was to truly blend PMs and DDMs to achieve the full potential of the HMs, namely: (i) overperform the top DDMs and PMs without the computational burden required by the top PMs and (ii) maintain the physical plausibility of PMs.

Regarding point (i) our proposal is depicted in the previous section. For what concerns (ii), we will leverage a twofold strategy. The first one comes from the world of the DDMs (the Feature Importance - FI explainability method Molnar, 2020) and the second one comes from the world of PMs (testing the behaviour of DDMs and HMs based on prior knowledge — TPK of the phenomena under consideration Oberkampf and Roy, 2010).

Let us start with the FI. In this setting, we decided to exploit the permutation importance (Fisher et al., 2019; Good, 2013; François et al., 2006), a statistically grounded and effective method firstly applied in RF (Breiman, 2001), using the mean decrease in MAE as a metric. This approach allows understanding of the features which mostly affect the models' performance. Failure of the learned models to properly account for the relevant features according to the domain knowledge might indicate poor quality in the measurements, or spurious correlations.

Table 9
Numerical experiments for the DDMs' physical plausibility assessment.

Case	Propeller	Wakefield	n_p [Hz]	p_{rel} [mmHg]	J [–]	σ_n [–]
C_1	KCD 191	W3	22.0	300	0.44	3.75
C_2	KCD 191	W2	25.5	100	0.38	3.42
C_3	KCD 129	W3	20.7	80	0.46	5.18
C_4	KCD 192	W3	22.5	0	0.43	4.74
C_5	KCD 74	W3	24.5	0	0.39	4.07

FI, therefore, represents an important step in the model verification process, since it examines whether the learned models generate results that are consistent with prior knowledge of the phenomena under consideration. The idea behind the permutation-based FI is quite simple: one has to estimate the increase/decrease in the model's MAE after permuting the feature. The feature is deemed important if the permutation of its values increases the model error, which implies that the model strongly relied on the feature for making accurate predictions. On the other hand, a feature is considered unimportant if the permutation of its values does not alter (or does not increase) the model MAE, as the model practically disregards (or is deceived by) that feature in its predictions.

For what concerns TPK, instead, we have conducted additional numerical experiments to test the physical plausibility of the predicted URN SPLs. The additional numerical experiments are designed to assess if the predictions of the models are in line with the knowledge of the phenomena under consideration. This assessment is especially critical for DDMs as, contrarily to HMs, they do not rely on any prior knowledge regarding the phenomena under consideration. These experiments still involve the propellers and wakefields presented in Section 3 but under different loading conditions. The loading conditions can be introduced by varying n_p and p_{rel}. Selecting values that do not appear in our data allows for investigating how the cavitating area and the vortex strength affect model predictions. The numerical experiments conducted to test these effects are provided in Table 9. The predictions of models will be compared with the general theory on cavitation noise. More specifically, using one of the cases of Table 9 as a reference, we evaluate

- the effect of the cavitation area, by examining the relative differences between:
  - the noise increment corresponding to increased or decreased cavitation area as predicted by the formula of Brown (1976),
  - the noise increment according to the DDMs, estimated on the frequency band between 4–20 [kHz],
- the effect of the vortex strength, by examining the relative differences between:
  - the URN peak level corresponding to increased or decreased cavitating tip vortex radius, as predicted by the model of Bosschers (2017),
  - the URN peak level according to the DDMs.

To better appreciate the cavitating conditions occurring in these experiments, apart from the unsteady non-cavitating BEM simulations that provide the necessary inputs for all the models, unsteady cavitating

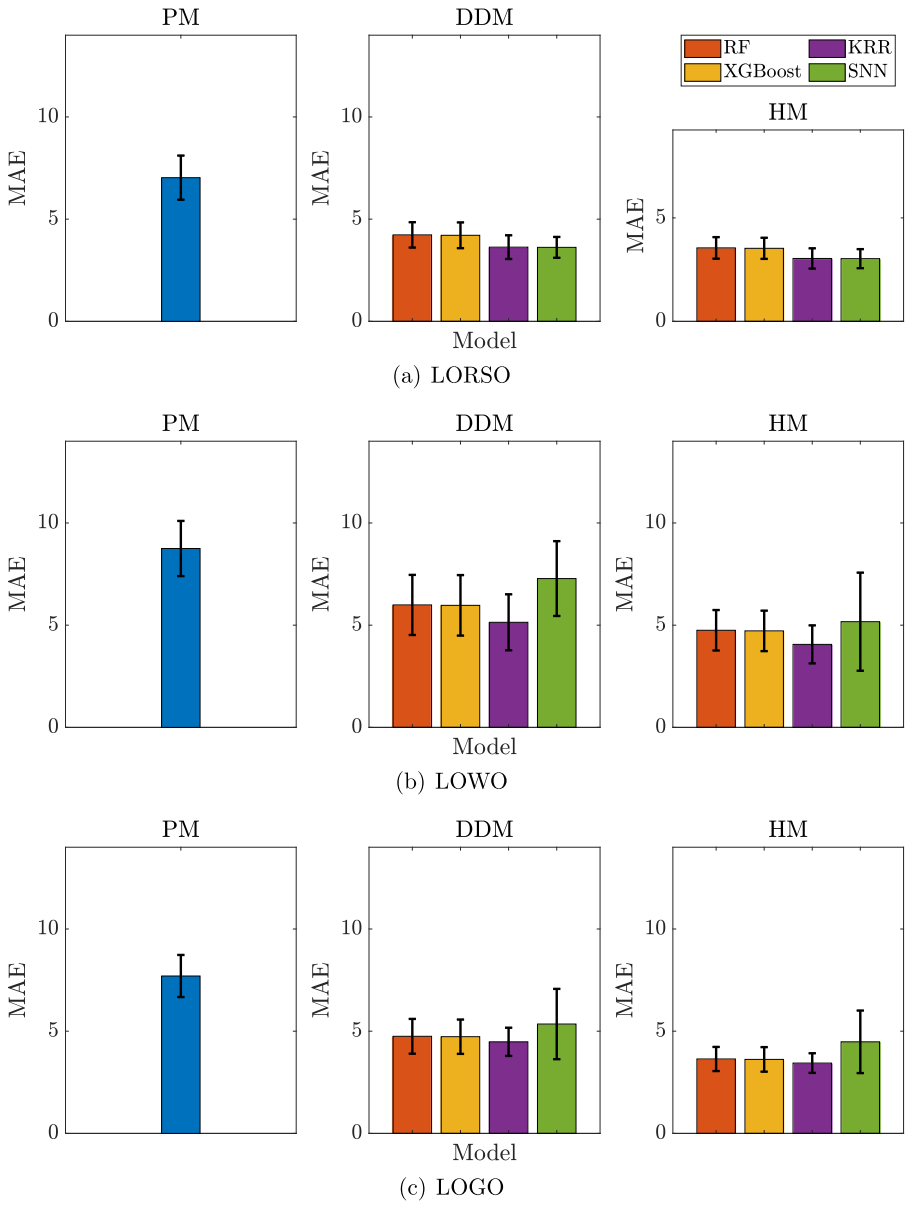


Fig. 9. Visual representation of the MAE from Table 10: overall average performance measured according to MAE for the PMs, DDMs (based on RF, XGBoost, KRR, and SNN), and the HMs counterparts of the DDMs for the different scenarios (LORSO, LOWO, and LOGO).

BEM simulations have also been performed. A visual representation of the cavities formed on the blade surfaces for all experiments of Table 9 is provided in Fig. 8, at one time-step of the simulations. In Fig. 8 the colour of the cavity bubble corresponds to the estimated thickness from the cavitating BEM conditions. The range of the thickness values is consistent among all experiments, to facilitate the visual comparison in terms of both cavitation area and cavitation thickness.

5. Results

In this section, we test the performance of the PMs, DDMs, and HMs described in Section 4 developed utilising the data described in Section 3, in terms of the performance measures discussed in Section 4.5 and analysing the physical plausibility of the DDMs and HMs according to what described in Section 4.6.

5.1. Performance of PMs, DDMs, and HMs

In this section, we will test and compare the performance of PMs, DDMs, and HMs following the approach described in Section 4.5 using the data described in Section 3. In particular, we tested three extrapolating scenarios (LORSO, LOWO, and LOGO) and we measured the performance with different metrics (three quantitative — MAE, MAPE, and PPMCC – and one qualitative – the scatter plot).

For this purpose we report a series of tables and graphs that will help the reader appreciating and understanding the quality of the different models in different scenarios.

Table 10 reports, for the different scenarios (LORSO, LOWO, and LOGO), the overall average performance measured according to MAE, MAPE, and PPMCC for the PMs, DDMs (based on RF, XGBoost, KRR, and SNN), and the HMs counterparts of the DDMs. Fig. 9 shows the MAE reported in Table 10 but with a visual representation which is easier to catch. Tables 11, 12, and 13 have been inserted for the sake of completeness and report the MAE, MAPE, and PPMCC respectively for the different scenario but detailing the accuracy on the different speed for the LORSO, the accuracy on the different wakes for LOWO, the accuracy on the different geometries for the LOGO. Fig. 10, instead, reports the scatter plots for the different scenario (LORSO, LOWO, and LOGO) for the best PM, DDM, and HM. Finally, Fig. 11 reports for a

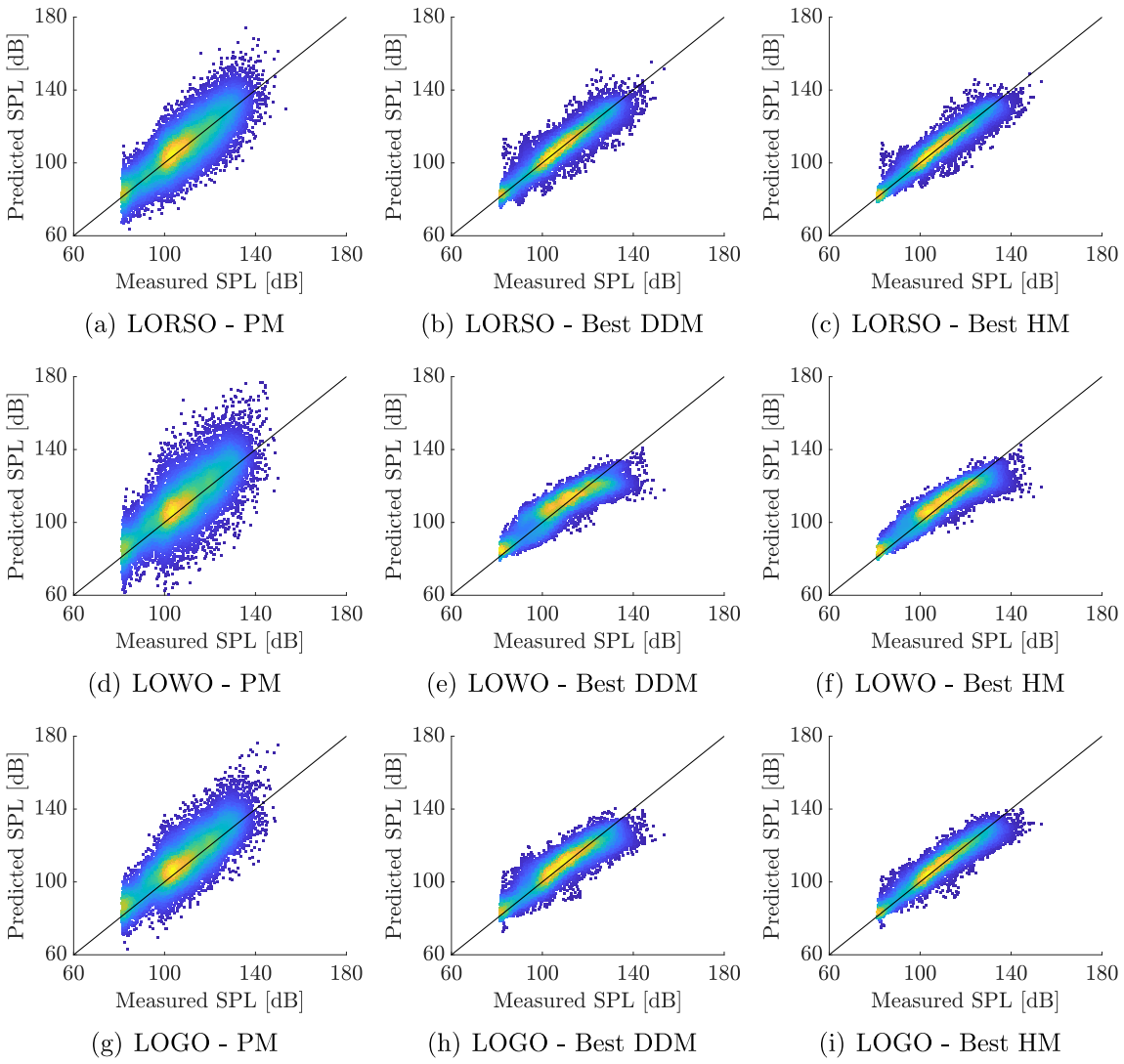
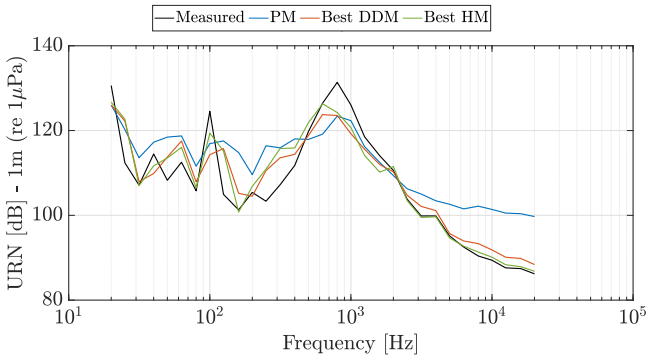


Fig. 10. Scatter plot, actual versus predicted value, for the different scenario (LORSO, LOWO, and LOGO) for the best PM, DDM, and HM.

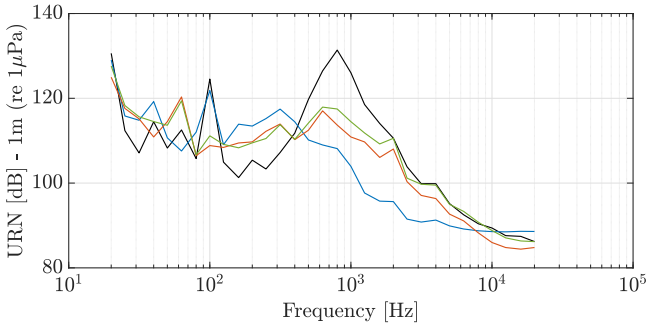
Table 10
Overall average performance measured according to MAE, MAPE, and PPMCC for the PMs, DDMs (based on RF, XGBoost, KRR, and SNN), and the HMs counterparts of the DDMs for the different scenarios (LORSO, LOWO, and LOGO).

	PM	DDM				HM			
		RF	XGBoost	KRR	SNN	RF	XGBoost	KRR	SNN
Metric	LORSO								
MAE	7.03 ± 1.08	4.23 ± 0.62	4.21 ± 0.63	3.63 ± 0.58	3.62 ± 0.51	3.55 ± 0.52	3.53 ± 0.51	3.04 ± 0.49	3.03 ± 0.46
MAPE	6.57 ± 1.03	3.79 ± 0.52	3.77 ± 0.52	3.22 ± 0.50	3.22 ± 0.43	3.24 ± 0.35	3.21 ± 0.34	2.77 ± 0.41	2.76 ± 0.38
PPMCC	0.81 ± 0.07	0.88 ± 0.04	0.87 ± 0.04	0.89 ± 0.04	0.89 ± 0.05	0.92 ± 0.04	0.92 ± 0.05	0.93 ± 0.04	0.93 ± 0.03
Metric	LOWO								
MAE	8.75 ± 1.35	5.99 ± 1.47	5.97 ± 1.48	5.14 ± 1.37	7.28 ± 1.83	4.75 ± 0.99	4.72 ± 0.99	4.06 ± 0.93	5.17 ± 2.40
MAPE	8.41 ± 1.35	5.64 ± 1.34	5.61 ± 1.35	4.79 ± 1.29	6.83 ± 1.65	4.41 ± 0.89	4.39 ± 0.89	3.77 ± 1.04	4.78 ± 2.41
PPMCC	0.79 ± 0.11	0.85 ± 0.08	0.84 ± 0.07	0.86 ± 0.07	0.81 ± 0.12	0.90 ± 0.03	0.90 ± 0.04	0.91 ± 0.03	0.89 ± 0.05
Metric	LOGO								
MAE	7.70 ± 1.03	4.75 ± 0.85	4.73 ± 0.84	4.48 ± 0.69	5.35 ± 1.72	3.64 ± 0.59	3.62 ± 0.60	3.44 ± 0.48	4.48 ± 1.53
MAPE	7.20 ± 0.97	4.57 ± 0.82	4.54 ± 0.81	4.12 ± 0.66	5.09 ± 1.68	3.45 ± 0.58	3.43 ± 0.59	3.16 ± 0.36	4.16 ± 1.25
PPMCC	0.82 ± 0.10	0.87 ± 0.06	0.87 ± 0.06	0.88 ± 0.06	0.85 ± 0.07	0.91 ± 0.03	0.91 ± 0.03	0.92 ± 0.03	0.89 ± 0.04

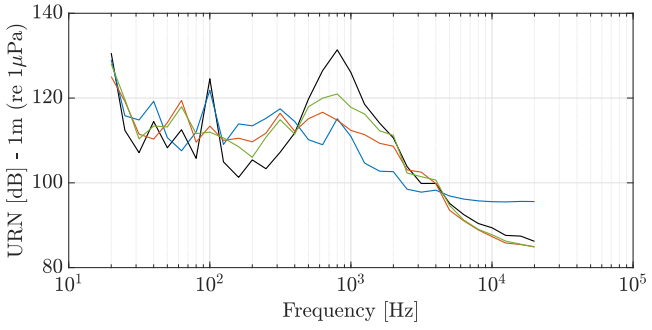
single representative spectrum the comparison between the best PM, DDM, and HM in the different scenarios (LORSO, LOWO, and LOGO) and in Fig. 12 the comparison between the different scenario (LORSO, LOWO, and LOGO) for the best PM, DDM, and HM (the dual version of Fig. 11). The representative spectra have been selected as they correspond to the experiment for which the errors committed by the PM, DDMs and HMs closely follow the average performance reported in Table 10. From these results, it is possible to observe that:



(a) LORSO - Measured, PM, Best DDM, and Best HM

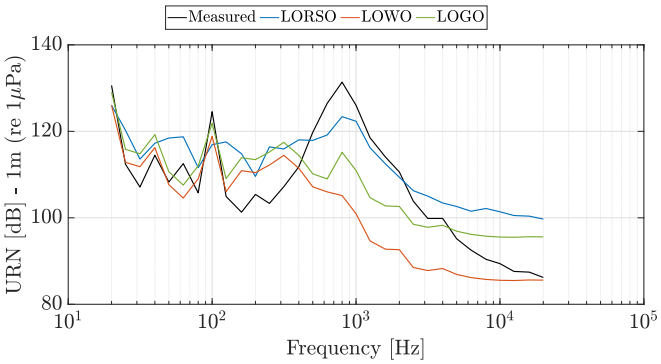


(b) LOWO - Measured, PM, Best DDM, and Best HM

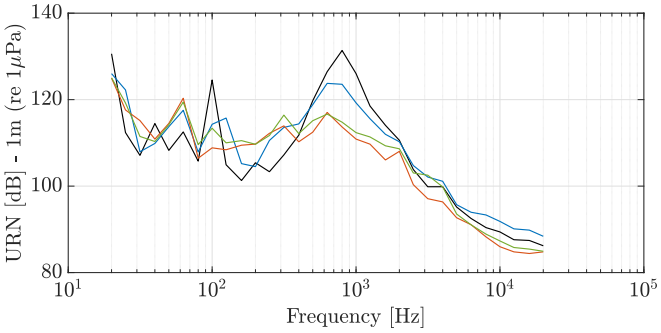


(c) LOGO - Measured, PM, Best DDM, and Best HM

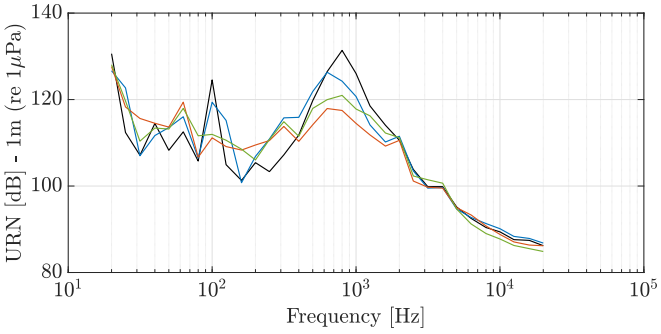
Fig. 11. Comparison between the best PM, DDM, and HM in the different scenario (LORSO, LOWO, and LOGO) for a single representative spectrum.



(a) Measured, LORSO PM, LOWO PM, and LOGO PM



(b) Measured, LORSO Best DDM, LOWO Best DDM, and LOGO Best DDM



(c) Measured, LORSO Best HM, LOWO Best HM, and LOGO Best HM

Fig. 12. Comparison between the different scenario (LORSO, LOWO, and LOGO) of the best PM, DDM, and HM in for a single representative spectrum.

- all models commit the lowest errors on the LORSO scenario, followed by the LOGO and LOWO scenarios that exhibit similar errors;
- considering the performance of each individual model, the errors committed are fairly consistent within each scenario, irrespective of the wakefield (LOWO), propeller geometry (LOGO), or rotational speed (LORSO), being present in the test set;
- the HMs are the most accurate models across all scenarios, followed by the DDMs, and the PM which is consistently the least accurate model. Specifically, the error reduction of the DDMs and HMs with respect to the PM is approximately 30% and 45%, respectively;
- the relative performance of the DDMs and HMs with respect to the PM decreases with the complexity of the scenario: for the least complex LORSO scenario, the DDMs commit approximately 50% lower errors than the PM, whereas for the HMs the same error reduction approaches 60%. On the other hand, for the LOGO scenario, the same error reductions correspond to 35% (DDMs) and 50% (HMs);
- the performance of the various DDMs is similar, except for the SNN which is consistently the least accurate DDM. RF, XGBoost, and KRR experience relatively small performance differences across all scenarios with KRR being the top performing DDM. One exception is the LORSO scenario, in which the SNN commits slightly lower errors, however the differences between the two is not statistically significant, i.e., the associated confidence intervals strongly overlap. Similar conclusions can be drawn for the HMs, with the HM based on KRR being the top performing model for all scenarios;
- the PM is able to capture the general trends of the URN spectrum, however it tends to underestimate the noise levels at frequencies higher than 500 [Hz]. The DDMs exhibit the opposite behaviour, with the highest accuracy being observed for frequencies lower than 1 [kHz]. Finally, the HMs, leveraging on the advantages of both the PM and DDMs, are able to capture the noise levels over the entire frequency range.

Table 11
Performance measured according to MAE for the PMs, DDMs (based on RF, XGBoost, KRR, and SNN), and the HMs counterparts of the DDMs for the different scenarios (LORSO, LOWO, and LOGO). MAE details for different speeds in the LORSO scenario, accuracy for different wakes in the LOWO scenario, and accuracy for different geometries in the LOGO scenario.

	PM	DDM				HM			
		RF	XGBoost	KRR	SNN	RF	XGBoost	KRR	SNN
Speed	LORSO								
600 [rpm]	7.02 ± 1.07	4.24 ± 0.58	4.02 ± 0.68	3.7 ± 0.55	3.82 ± 0.47	3.22 ± 0.56	3.26 ± 0.55	3.07 ± 0.46	2.92 ± 0.44
800 [rpm]	7.05 ± 1.02	4.08 ± 0.66	4.23 ± 0.65	3.8 ± 0.57	3.46 ± 0.46	3.67 ± 0.55	3.69 ± 0.5	3.19 ± 0.45	3.11 ± 0.47
1000 [rpm]	7.02 ± 1.12	4.43 ± 0.63	4.61 ± 0.61	3.83 ± 0.56	3.63 ± 0.56	3.74 ± 0.51	3.32 ± 0.47	3.14 ± 0.52	2.99 ± 0.47
1200 [rpm]	7.02 ± 1.08	3.92 ± 0.56	4.6 ± 0.62	3.63 ± 0.57	3.78 ± 0.56	3.58 ± 0.5	3.7 ± 0.56	3.03 ± 0.5	3.25 ± 0.46
1400 [rpm]	7.01 ± 1.09	3.92 ± 0.63	3.86 ± 0.67	3.44 ± 0.63	3.33 ± 0.56	3.86 ± 0.48	3.32 ± 0.57	3.22 ± 0.5	2.92 ± 0.49
1600 [rpm]	7.05 ± 1.05	4.42 ± 0.58	3.81 ± 0.65	3.93 ± 0.6	3.93 ± 0.46	3.47 ± 0.57	3.84 ± 0.51	2.74 ± 0.5	3.13 ± 0.46
1800 [rpm]	7.04 ± 1.01	3.82 ± 0.58	4.29 ± 0.59	3.28 ± 0.59	3.42 ± 0.48	3.66 ± 0.52	3.4 ± 0.51	2.82 ± 0.51	3.19 ± 0.43
2000 [rpm]	7.03 ± 1.14	4.35 ± 0.67	4.22 ± 0.57	3.52 ± 0.57	3.58 ± 0.54	3.32 ± 0.52	3.74 ± 0.47	3.18 ± 0.47	2.84 ± 0.42
Average	7.03 ± 1.08	4.23 ± 0.62	4.21 ± 0.63	3.63 ± 0.58	3.62 ± 0.51	3.55 ± 0.52	3.53 ± 0.51	3.04 ± 0.49	3.03 ± 0.46
Wake	LOWO								
W1	8.91 ± 1.31	6.16 ± 1.45	6.14 ± 1.44	5.91 ± 1.35	7.38 ± 1.81	4.64 ± 0.98	4.63 ± 0.99	4.45 ± 0.91	5.78 ± 2.49
W2	8.61 ± 1.30	5.78 ± 1.47	5.75 ± 1.47	4.59 ± 1.36	6.31 ± 1.75	4.68 ± 0.99	4.65 ± 0.98	3.72 ± 0.92	3.74 ± 2.20
W3	8.74 ± 1.35	6.04 ± 1.48	6.02 ± 1.49	4.92 ± 1.41	8.14 ± 1.95	4.92 ± 1.00	4.90 ± 0.99	4.01 ± 0.95	5.99 ± 2.50
Average	8.75 ± 1.35	5.99 ± 1.47	5.97 ± 1.48	5.14 ± 1.37	7.28 ± 1.83	4.75 ± 0.99	4.72 ± 0.99	4.06 ± 0.93	5.17 ± 2.40
Geometry	LOGO								
KCD-65	7.99 ± 1.10	4.87 ± 0.81	4.87 ± 0.80	4.58 ± 0.69	5.62 ± 1.70	4.30 ± 0.57	4.28 ± 0.56	4.04 ± 0.48	5.15 ± 1.67
KCD-74	7.58 ± 1.02	4.98 ± 0.92	4.95 ± 0.89	4.66 ± 0.66	4.91 ± 1.63	3.65 ± 0.60	3.63 ± 0.60	3.48 ± 0.44	4.04 ± 1.57
KCD-129	7.70 ± 1.04	5.58 ± 0.92	5.58 ± 0.92	5.16 ± 0.81	5.22 ± 1.81	4.34 ± 0.62	4.32 ± 0.61	4.01 ± 0.54	4.64 ± 1.46
KCD-191	7.67 ± 1.02	4.98 ± 0.82	4.94 ± 0.82	4.62 ± 0.68	5.17 ± 1.72	3.55 ± 0.59	3.52 ± 0.58	3.29 ± 0.48	4.98 ± 1.55
KCD-192	7.82 ± 1.04	4.12 ± 0.78	4.10 ± 0.77	3.92 ± 0.68	5.71 ± 1.83	3.02 ± 0.55	3.00 ± 0.56	2.87 ± 0.49	4.05 ± 1.49
KCD-193	7.53 ± 1.02	4.04 ± 0.87	4.02 ± 0.85	3.93 ± 0.60	5.48 ± 1.64	3.00 ± 0.65	2.99 ± 0.64	2.92 ± 0.44	4.04 ± 1.46
Average	7.70 ± 1.03	4.75 ± 0.85	4.73 ± 0.84	4.48 ± 0.69	5.35 ± 1.72	3.64 ± 0.59	3.62 ± 0.60	3.44 ± 0.48	4.48 ± 1.53

Table 12
Performance measured according to MAPE for the PMs, DDMs (based on RF, XGBoost, KRR, and SNN), and the HMs counterparts of the DDMs for the different scenarios (LORSO, LOWO, and LOGO). MAPE details for different speeds in the LORSO scenario, accuracy for different wakes in the LOWO scenario, and accuracy for different geometries in the LOGO scenario.

	PM	DDM				HM			
		RF	XGBoost	KRR	SNN	RF	XGBoost	KRR	SNN
Speed	LORSO								
600 [rpm]	6.05 ± 0.96	3.44 ± 0.51	3.73 ± 0.57	3.42 ± 0.5	3.3 ± 0.44	3.01 ± 0.32	3.2 ± 0.32	2.71 ± 0.44	3 ± 0.35
800 [rpm]	6.19 ± 1.11	3.47 ± 0.53	3.69 ± 0.52	3.48 ± 0.46	3.3 ± 0.44	2.93 ± 0.35	3.42 ± 0.32	3.01 ± 0.45	3.03 ± 0.41
1000 [rpm]	6.87 ± 0.99	3.94 ± 0.51	4.08 ± 0.54	3.44 ± 0.46	3.37 ± 0.4	3.32 ± 0.33	2.92 ± 0.35	2.53 ± 0.42	2.54 ± 0.38
1200 [rpm]	6.64 ± 1.05	3.78 ± 0.48	3.51 ± 0.53	3.08 ± 0.5	3.08 ± 0.39	3.52 ± 0.36	3.38 ± 0.36	2.62 ± 0.4	2.67 ± 0.41
1400 [rpm]	6.58 ± 1.03	3.99 ± 0.57	3.7 ± 0.48	3.52 ± 0.46	3.27 ± 0.39	3.29 ± 0.37	3.32 ± 0.33	2.97 ± 0.39	2.55 ± 0.41
1600 [rpm]	6.92 ± 1.09	4.09 ± 0.49	3.78 ± 0.5	3.02 ± 0.49	3.09 ± 0.43	3.01 ± 0.38	3.4 ± 0.31	2.71 ± 0.38	3 ± 0.39
1800 [rpm]	7.12 ± 0.98	3.82 ± 0.47	3.77 ± 0.52	3.37 ± 0.47	2.92 ± 0.42	3.29 ± 0.37	3.04 ± 0.36	2.57 ± 0.43	2.61 ± 0.39
2000 [rpm]	6.2 ± 1	3.8 ± 0.5	3.96 ± 0.49	3.26 ± 0.5	3.37 ± 0.43	3.49 ± 0.33	3.05 ± 0.36	2.97 ± 0.43	2.75 ± 0.37
Average	6.57 ± 1.03	3.79 ± 0.52	3.77 ± 0.52	3.22 ± 0.50	3.22 ± 0.43	3.24 ± 0.35	3.21 ± 0.34	2.77 ± 0.41	2.76 ± 0.38
Wake	LOWO								
W1	8.53 ± 1.41	5.86 ± 1.32	5.84 ± 1.31	5.49 ± 1.22	6.71 ± 1.59	4.31 ± 0.88	4.29 ± 0.89	4.13 ± 0.90	5.32 ± 2.40
W2	8.23 ± 1.33	5.38 ± 1.35	5.32 ± 1.35	4.28 ± 1.28	5.83 ± 1.54	4.35 ± 0.90	4.35 ± 0.89	3.46 ± 0.88	3.38 ± 2.36
W3	8.47 ± 1.38	5.69 ± 1.36	5.67 ± 1.37	4.61 ± 1.36	7.94 ± 1.81	4.57 ± 0.90	4.55 ± 0.91	3.73 ± 1.35	5.64 ± 2.48
Average	8.41 ± 1.35	5.64 ± 1.34	5.61 ± 1.35	4.79 ± 1.29	6.83 ± 1.65	4.41 ± 0.89	4.39 ± 0.89	3.77 ± 1.04	4.78 ± 2.41
Geometry	LOGO								
KCD-65	7.70 ± 1.04	4.76 ± 0.78	4.72 ± 0.77	4.34 ± 0.67	5.35 ± 1.93	4.07 ± 0.55	4.05 ± 0.56	3.83 ± 0.37	4.72 ± 1.23
KCD-74	7.10 ± 0.96	4.69 ± 0.89	4.68 ± 0.88	4.23 ± 0.62	4.78 ± 1.59	3.46 ± 0.60	3.44 ± 0.59	3.15 ± 0.31	3.86 ± 1.25
KCD-129	7.28 ± 0.97	5.38 ± 0.88	5.37 ± 0.87	4.83 ± 0.72	5.06 ± 1.83	4.11 ± 0.60	4.09 ± 0.61	3.75 ± 0.43	4.39 ± 1.21
KCD-191	7.14 ± 0.94	4.77 ± 0.78	4.73 ± 0.78	4.21 ± 0.67	5.01 ± 1.71	3.36 ± 0.56	3.35 ± 0.57	3.00 ± 0.37	4.46 ± 1.31
KCD-192	7.31 ± 0.96	3.96 ± 0.73	3.93 ± 0.73	3.59 ± 0.69	5.35 ± 1.61	2.86 ± 0.55	2.85 ± 0.55	2.63 ± 0.38	3.76 ± 1.27
KCD-193	7.00 ± 0.95	3.84 ± 0.85	3.83 ± 0.84	3.52 ± 0.57	5.01 ± 1.39	2.84 ± 0.62	2.81 ± 0.61	2.62 ± 0.32	3.78 ± 1.25
Average	7.20 ± 0.97	4.57 ± 0.82	4.54 ± 0.81	4.12 ± 0.66	5.09 ± 1.68	3.45 ± 0.58	3.43 ± 0.59	3.16 ± 0.36	4.16 ± 1.25

5.2. Physical plausibility of DDMs and HMs

In order to test the physical plausibility of the DDMs and HMs, we will leverage a twofold strategy as described in Section 4.6: the use of the FI (Section 5.2.1) and the TPK (Section 5.2.2).

5.2.1. Feature importance

Having studied the performance of the different models developed in the paper and showed that the HM based on KRR is the top performing one, we computed the FI described in Section 4.6 to verify if the model actually learned physically plausible knowledge from the

Table 13
Performance measured according to PPMCC for the PMs, DDMs (based on RF, XGBoost, KRR, and SNN), and the HMs counterparts of the DDMs for the different scenarios (LORSO, LOWO, and LOGO). PPMCC details for different speeds in the LORSO scenario, accuracy for different wakes in the LOWO scenario, and accuracy for different geometries in the LOGO scenario.

	PM	DDM				HM			
		RF	XGBoost	KRR	SNN	RF	XGBoost	KRR	SNN
Speed	LORSO								
600 [rpm]	0.776 ± 0.065	0.852 ± 0.044	0.816 ± 0.037	0.938 ± 0.041	0.803 ± 0.048	0.927 ± 0.039	1.012 ± 0.05	0.876 ± 0.039	0.886 ± 0.027
800 [rpm]	0.844 ± 0.069	0.888 ± 0.043	0.79 ± 0.039	0.878 ± 0.038	0.836 ± 0.049	0.865 ± 0.039	0.971 ± 0.055	1.011 ± 0.038	0.906 ± 0.032
1000 [rpm]	0.845 ± 0.076	0.87 ± 0.042	0.92 ± 0.043	0.959 ± 0.039	0.853 ± 0.047	0.966 ± 0.044	0.843 ± 0.051	0.94 ± 0.042	1.017 ± 0.029
1200 [rpm]	0.876 ± 0.071	0.891 ± 0.037	0.834 ± 0.038	0.892 ± 0.043	0.902 ± 0.054	0.838 ± 0.041	0.855 ± 0.046	0.985 ± 0.039	0.889 ± 0.029
1400 [rpm]	0.825 ± 0.072	0.964 ± 0.041	0.91 ± 0.04	0.812 ± 0.042	0.842 ± 0.048	0.94 ± 0.04	0.864 ± 0.048	0.838 ± 0.04	0.865 ± 0.031
1600 [rpm]	0.755 ± 0.072	0.84 ± 0.041	0.903 ± 0.044	0.872 ± 0.039	0.934 ± 0.051	0.896 ± 0.044	0.919 ± 0.049	0.971 ± 0.037	0.967 ± 0.029
1800 [rpm]	0.817 ± 0.069	0.916 ± 0.038	0.949 ± 0.04	0.977 ± 0.04	0.968 ± 0.055	0.957 ± 0.04	0.966 ± 0.048	0.885 ± 0.043	1.016 ± 0.029
2000 [rpm]	0.756 ± 0.072	0.839 ± 0.043	0.902 ± 0.041	0.803 ± 0.044	0.965 ± 0.05	0.956 ± 0.042	0.937 ± 0.054	0.91 ± 0.041	0.865 ± 0.027
Average	0.81 ± 0.07	0.88 ± 0.04	0.87 ± 0.04	0.89 ± 0.04	0.89 ± 0.05	0.92 ± 0.04	0.92 ± 0.05	0.93 ± 0.04	0.93 ± 0.03
Wake	LOWO								
W1	0.78 ± 0.12	0.83 ± 0.08	0.83 ± 0.07	0.85 ± 0.07	0.82 ± 0.12	0.88 ± 0.03	0.89 ± 0.04	0.90 ± 0.03	0.87 ± 0.05
W2	0.80 ± 0.11	0.86 ± 0.07	0.85 ± 0.07	0.87 ± 0.06	0.81 ± 0.10	0.91 ± 0.04	0.90 ± 0.04	0.92 ± 0.04	0.93 ± 0.04
W3	0.81 ± 0.14	0.85 ± 0.08	0.84 ± 0.08	0.86 ± 0.07	0.80 ± 0.13	0.90 ± 0.03	0.91 ± 0.03	0.91 ± 0.03	0.88 ± 0.05
Average	0.79 ± 0.11	0.85 ± 0.08	0.84 ± 0.07	0.86 ± 0.07	0.81 ± 0.12	0.90 ± 0.03	0.90 ± 0.04	0.91 ± 0.03	0.89 ± 0.05
Geometry	LOGO								
KCD-65	0.80 ± 0.12	0.87 ± 0.05	0.87 ± 0.04	0.87 ± 0.05	0.84 ± 0.08	0.91 ± 0.03	0.92 ± 0.03	0.91 ± 0.03	0.85 ± 0.05
KCD-74	0.85 ± 0.08	0.89 ± 0.05	0.88 ± 0.06	0.88 ± 0.05	0.87 ± 0.06	0.92 ± 0.03	0.92 ± 0.04	0.91 ± 0.03	0.92 ± 0.04
KCD-129	0.84 ± 0.08	0.85 ± 0.07	0.86 ± 0.08	0.87 ± 0.06	0.86 ± 0.08	0.89 ± 0.04	0.88 ± 0.03	0.91 ± 0.03	0.86 ± 0.03
KCD-191	0.82 ± 0.11	0.87 ± 0.06	0.87 ± 0.06	0.87 ± 0.05	0.84 ± 0.07	0.92 ± 0.02	0.91 ± 0.03	0.92 ± 0.02	0.85 ± 0.05
KCD-192	0.80 ± 0.10	0.89 ± 0.05	0.88 ± 0.06	0.88 ± 0.05	0.85 ± 0.06	0.94 ± 0.02	0.94 ± 0.02	0.93 ± 0.02	0.93 ± 0.03
KCD-193	0.83 ± 0.9	0.87 ± 0.08	0.88 ± 0.06	0.90 ± 0.07	0.86 ± 0.07	0.90 ± 0.03	0.91 ± 0.04	0.93 ± 0.03	0.94 ± 0.04
Average	0.82 ± 0.10	0.87 ± 0.06	0.87 ± 0.06	0.88 ± 0.06	0.85 ± 0.07	0.91 ± 0.03	0.91 ± 0.03	0.92 ± 0.03	0.89 ± 0.04

data. Before analysing the results of this procedure, it should be noted that the many features in the dataset are strongly correlated or even redundant and that FI is not a so precise information to retrieve. Consequently, we should look at global trends and not go into details of results that would not provide any real insight.

Table 14 reports the top 20 features computed during the FI test for the top performing model (HM based on KRR). From Table 14, it is possible to observe that:

- the outputs of the PM at various frequencies correspond to 6 out of the 20 most important features;
- the relative tunnel pressure and the cavitation index based on rotational speed are also ranked highly as they are directly related to the URN levels;
- the angular locations in which the size of the blade area on the propeller suction side having pressure lower than the vapour pressure reaches its minimum and maximum value are also highly ranked. Note that these features were included to estimate the region where true cavitation starts and ends, which correlate with the time duration of bubble growth and collapse, hence significantly influencing the characteristics of noise spectra both in terms of amplitude and frequency distribution;
- several Fourier coefficients extracted from the wakefield as well as propeller design parameters are also present, including the number of propeller blades, the total rake and skew angle at $r = 0.75R$, as well as the skew angle near the tip ($r = 0.95R$), which are known to be important design parameters with respect to URN (Da-Qing, 2006; Tong et al., 2021; Carlton, 2018);
- it is interesting to note the absence in the top positions of features assumed to be strongly related to the cavitation noise, such as the features related to the angle of attack. This can be justified by recalling that these features are directly dependent on, or derived by, the wakefield. Hence, these features might be redundant.

Summarising, the results of the FI test are consistent with general cavitation noise theory, with the top positions occupied by important parameters and PM features directly containing the theoretical information about propeller cavitation noise.

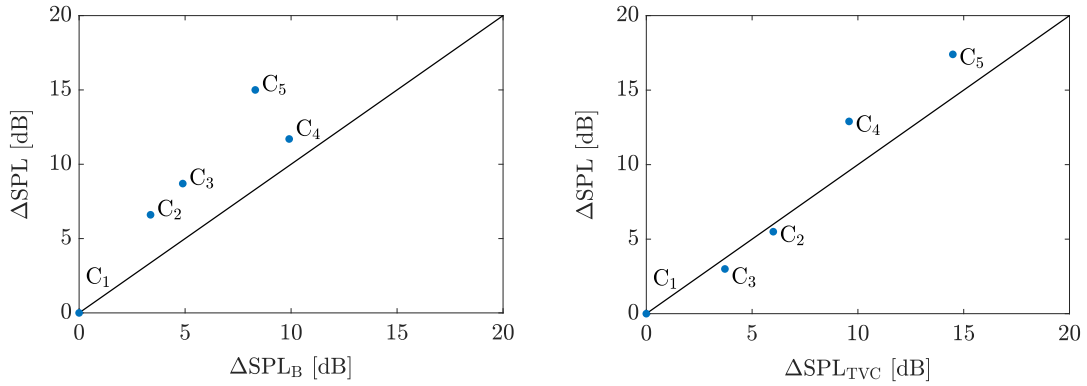
Table 14
Top 20 features computed during the FI test for the top performing model (HM based on KRR).

#	Feature	#	Feature
1	$P_M(X; f = 500\ [\text{Hz}])$	11	$P_M(X; f = 630\ [\text{Hz}])$
2	$P_M(X; f = 1000\ [\text{Hz}])$	12	Z
3	p_{rel}	13	$a^{(3)}_{w_{07}}$
4	σ_n	14	$P_M(X; f = 800\ [\text{Hz}])$
5	$\theta_{\max A_c\|_{(s)}}$	15	$a^{(1)}_{w_{07}}$
6	$\theta_{\min A_c\|_{(s)}}$	16	$a^{(2)}_{w_{07}}$
7	$P_M(X; f = 400\ [\text{Hz}])$	17	$P_M(X; f = 1250\ [\text{Hz}])$
8	$\theta_s(r/R = 0.75)$	18	$a^{(1)}_{w_{09}}$
9	$\theta_s(r/R = 0.95)$	19	$a^{(2)}_{w_{09}}$
10	$i_T\|(r/R = 0.75)$	20	$\Gamma_{\max 095}$

5.2.2. Test against prior knowledge

As discussed in Section 4.6, it is important to assess if the predictions of the best model developed – HM based on KRR (HM-KRR) – are aligned with the general theory of cavitation noise. To this end, for the cases reported in Table 9 the predictions of the HM-KRR are compared with the models of Brown (1976), Bosschers (2018c).

More in detail, the noise increments of the HM-KRR, estimated on the frequency range between 4÷20 [kHz], are compared with the noise increment predicted with the formula of Brown (1976). Results are reported in Fig. 13(a). Utilising the URN SPLs of case C_1 as a reference value, which represents the origin of the axes in Fig. 13(a), the relative noise increase ΔSPL with respect to C_1 is evaluated with the HM-KRR and is plotted along the y-axis. The same predictions ΔSPL_B from the formula of Brown (1976) are plotted along the x-axis. A perfect correlation between the HM-KRR's and Brown's predictions is expected when $\Delta SPL = \Delta SPL_B$ for all cases, which corresponds to the straight line of Fig. 13(a). Whereas this is not the case for Fig. 13(a), the predictions of the HM-KRR do correlate with Brown's formula. The largest discrepancy is observed for Case C_5, in which ΔSPL = 15 [dB] for the HM-KRR, whereas the formula of Brown (1976) estimates an increase of 7.5 [dB]. For cases C_2 and C_3, the HM-KRR estimates that ΔSPL equals 6 [dB], and 8.5 [dB], respectively. According to Brown (1976), ΔSPL should be equal to 4 [dB], and 5 [dB] respectively. In



(a) URN increment. (b) URN peak level.

Fig. 13. URN increment and peak levels predicted with the HM-KRR and the models of Brown (1976), Bosschers (2017) for the cases of Table 9.

addition, for Case C_4, the HM-KRR predict ΔSPL = 11 [dB], whereas a value of 10 [dB] is expected according to Brown (1976). Overall, the predictions of the two models are in reasonable agreement.

Fig. 13(b) presents the results of the same analysis conducted for the peak noise levels predicted by the HM-KRR and the model of Bosschers (2017). Once again, case C_1 is considered as the reference (axis origin), and the relative increase of the peak SPL with respect to C_1 (ΔSPL) is evaluated with the HM-KRR and plotted along the y-axis. The x-axis contains the same values (ΔSPL_{TVC}) as predicted by the model of Bosschers (2017). Similarly to the previous analysis, there is a strong correlation between the two models. However, the HM-KRR either over- or under-predict the peak SPL increase with respect to the predictions obtained from Bosschers (2017). The lowest discrepancy is observed for case C_3, with the HM-KRR matching almost exactly the estimations obtained from Bosschers (2017). The highest deviation is observed for cases C_4 and C_5, with the HM-KRR overestimating ΔSPL by approximately 4 [dB], with respect to the model of Bosschers (2017). Overall, the predictions of the two models are fairly close.

6. Conclusions

Silent vessels preserve marine wildlife, ensure the comfort of passengers, and are compulsory in underwater warfare. For propeller-driven vessels, cavitation is the most dominant noise source, producing both structure-borne and radiated noise. For this reason, accurate predictions of the noise signature at the design stage are fundamental for engineering simultaneously silent and efficient propellers. State-of-the-art predictive models are based on physical, data-driven, and hybrid approaches. PMs rely on the knowledge of the phenomena and can be quite accurate and reliable at the expense of huge computational requirements which prevent their use at the design stage. On the other side, DDMs rely on historical observations of the phenomena and can be computationally inexpensive and accurate but may produce physically implausible results. For this reason, HMs exploit the combination of PMs and DDMs taking advantage of their strengths while limiting their weaknesses. Currently, hybridisation does not fully achieve its true potential since it naively pulls together PMs and DDMs without actually blending them. For this reason, in this work, we propose a novel hybridisation strategy able to truly blend the knowledge of the phenomena with the information contained in the historical data producing models able to properly, i.e., physically plausible, extrapolating as PMs while being extremely accurate and computationally inexpensive as DDMs. The performance of all models in several real-world scenarios has been investigated. More specifically, three different tests have been conducted:

- extrapolation on the rotational speed, in which the models are tested with previously unseen rotational speeds;
- extrapolation on the wakefield, in which the models are tested with previously unseen wakefields,
- extrapolation on the propeller geometry, in which the models are tested with previously unseen geometries.

In order to support our statements, we tested our proposal, for the first time in the literature, on real data collected in a vast experimental campaign at the Emerson Cavitation Tunnel on the Meridian standard propeller series behind different severities of the axial wake. In all the considered scenarios, DDMs and HMs have shown increased capabilities with respect to PMs in predicting the URN spectra, with errors that are certainly acceptable during the early stage of the design process. The proposed HMs, consistently demonstrate the lowest average errors in all scenarios. Moreover, various tests showed that the top performing HMs learns and deliver predictions which are in agreement with the state-of-the-art physical knowledge of the phenomena.

In the future, we plan to further enlargement of the existing dataset with additional experiments that cover a broader set of propeller geometries and operating conditions using the facilities available at the University of Genoa and Delft University of Technology. We also plan to make a more tight collaboration with leading research institution active in the sector like the Maritime Research Institute Netherlands. This allows us not only to evaluate more thoroughly the capabilities of the proposed approaches, but to further validate and evolve them in additional scenarios encountered in practice. Furthermore, we aim at utilising the DDMs and HMs for design optimisation studies with multiple and conflicting operating goals, towards the development of unconventional propeller designs. Finally, we are trying to understand if Physically Informed surrogate models of CFD simulator could allow us to further develop a PM with a good trade-off between accuracy and computational requirements.

CRediT authorship contribution statement

Miltiadis Kalikatzarakis: Conceptualization, Methodology, Software, Validation, Data curation, Writing – original draft, Writing – review & editing. **Andrea Coraddu:** Conceptualization, Methodology, Software, Validation, Data curation, Supervision, Writing – original draft, Writing – review & editing. **Mehmet Atlar:** Dataset supply, Writing - Reviewing. **Stefano Gaggero:** Conceptualization, Methodology, Data curation, Writing – review & editing. **Giorgio Tani:** Conceptualization, Methodology, Data curation, Writing – review & editing. **Luca Oneto:** Conceptualization, Methodology, Data curation, Supervision, Writing – original draft.

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.

Data availability

The data that has been used is confidential.

Acknowledgement

The authors would like to express their gratitude to Dr. Batuhan Aktas, who generated and kindly provided the experimental data used in the present work.

References

Adam, S.P., Alexandropoulos, S.A.N., Pardalos, P.M., Vrahatis, M.N., 2019. No free lunch theorem: A review. Approx. Optim. 57–82.

Aktas, B., 2017. A systematic experimental approach to cavitation noise prediction of marine propellers (Ph.D. thesis). Newcastle University.

Aktas, B., Atlar, M., Fitzsimmons, P., Shi, W., 2018. An advanced joint time-frequency analysis procedure to study cavitation-induced noise by using standard series propeller data. Ocean Eng. 170, 329–350.

American National Standards Institute, 2009. ANSI S1.11: Specification for octave, half-octave, and third octave band filter sets. URL https://law.resource.org/pub/us/cfr/ibr/002/ansi.s1.11.2004.pdf.

Angelopoulos, A., Fitzsimmons, P.A., Odabasi, A.Y., 1988. A Semi-Empirical Method for Propeller Broad-Band Noise. Technical report, British Maritime Technology Limited.

Atlar, M., 2011. Recent upgrading of marine testing facilities at newcastle university. In: International Conference on Advanced Model Measurement Technology for the EU Maritime Industry.

Banan, A., Nasiri, A., Taheri-Garavand, A., 2020. Deep learning-based appearance features extraction for automated carp species identification. Aquac. Eng. 89, 102053.

Baxter, J., 2000. A model of inductive bias learning. J. Artificial Intelligence Res. 12, 149–198.

Biggio, B., Roli, F., 2018. Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognit. 84, 317–331.

Bishop, C.M., et al., 1995. Neural Networks for Pattern Recognition. Oxford University Press.

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

Bosschers, J., 2017. A semi-empirical method to predict broadband hull pressure fluctuations and underwater radiated noise by cavitating tip vortices. In: International Symposium on Marine Propulsors.

Bosschers, J., 2018a. Propeller Tip-Vortex Cavitation and its Broadband Noise (Ph.D. thesis). University of Twente.

Bosschers, J., 2018b. An analytical and semi-empirical model for the viscous flow around a vortex cavity. Int. J. Multiph. Flow. 105, 122–133.

Bosschers, J., 2018c. 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.

Bosschers, J., Choi, G.H., Hyundai, H.I., Farabee, K.T., Fréchou, D., Korkut, E., Sato, K., et al., 2017. Specialist committee on hydrodynamic noise. In: International Towing Tank Conference. Specialist Committee on HYdrodynamic Noise.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brennen, C.E., 2014. Cavitation and Bubble Dynamics. Cambridge University Press.

Brizzolara, S., Villa, D., Gaggero, S., 2008. A systematic comparison between RANS and panel methods for propeller analysis. In: International Conference on Hydrodynamics.

Brown, N.A., 1976. Cavitation noise problems and solutions. In: International Symposium on Shipboard Acoustics.

Brown, N.A., 1999. Thruster noise. In: Dynamic Positioning Conference.

Calcagni, D., Salvatore, F., Bernardini, G., Miozzi, M., 2010. Automated marine propeller design combining hydrodynamics models and neural networks. In: International Symposium on Fishing Vessel Energy Efficiency.

Capone, A., Di Felice, F., Pereira, F.A., 2021. On the flow field induced by two counter-rotating propellers at varying load conditions. Ocean Eng. 221, 108322.

Carlton, J., 2018. Marine Propellers and Propulsion. Butterworth-Heinemann.

Caruana, R., 1997. Multitask learning. Mach. Learn. 28 (1), 41–75.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In: ASM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Chou, E., Southall, B.L., Robards, M., Rosenbaum, H.C., 2021. International policy, recommendations, actions and mitigation efforts of anthropogenic underwater noise. Ocean Coast. Manag. 202, 105427.

Cipollini, F., Miglianti, F., Oneto, L., Tani, G., Viviani, M., 2019a. Hybrid model for cavitation noise spectra prediction. In: 2019 International Joint Conference on Neural Networks.

Cipollini, F., Miglianti, F., Oneto, L., Tani, G., Viviani, M., Anguita, D., 2019b. Cavitation noise spectra prediction with hybrid models. In: INNS Big Data and Deep Learning Conference.

Cipollini, F., Oneto, L., Coraddu, A., Murphy, A.J., Anguita, D., 2018. Condition-based maintenance of naval propulsion systems: Data analysis with minimal feedback. Reliab. Eng. Syst. Saf. 177, 12–23.

Cireşan, D., Meier, U., Masci, J., Schmidhuber, J., 2011. A committee of neural networks for traffic sign classification. In: International Joint Conference on Neural Networks.

Coraddu, A., Kalikatzarakis, M., Oneto, L., Meijn, G.J., Godjevac, M., Geertsma, R.D., 2018. Ship diesel engine performance modelling with combined physical and machine learning approach. In: International Ship Control Systems Symposium.

Coraddu, A., Kalikatzarakis, M., Theotokatos, G., Geertsma, R., Oneto, L., 2022. Physical and data-driven models hybridisation for modelling the dynamic state of a four-stroke marine diesel engine. In: Engine Modeling and Simulation.

Coraddu, A., Oneto, L., Baldi, F., Anguita, D., 2017. Vessels fuel consumption forecast and trim optimisation: a data analytics perspective. Ocean Eng. 130, 351–370.

Coraddu, A., Oneto, L., Baldi, F., Cipollini, F., Atlar, M., Savio, S., 2019. Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. Ocean Eng. 186, 106063.

Coraddu, A., Oneto, L., Cipollini, F., Kalikatzarakis, M., Meijn, G.J., Geertsma, R., 2021a. Physical, data-driven and hybrid approaches to model engine exhaust gas temperatures in operational conditions. Ships Offshore Struct. 1–22.

Coraddu, A., Oneto, L., Ilardi, D., Stoumpos, S., Theotokatos, G., 2021b. Marine dual fuel engines monitoring in the wild through weakly supervised data analytics. Eng. Appl. Artif. Intell. 100, 104179.

Coraddu, A., Oneto, L., Kalikatzarakis, M., Ilardi, D., Collu, M., 2020. Floating spar-type offshore wind turbine hydrodynamic response characterisation: A computational cost aware approach. In: Global Oceans 2020.

Da-Qing, L., 2006. Validation of RANS predictions of open water performance of a highly skewed propeller with experiments. J. Hydrodyn., Ser. B 18 (3), 520–528.

Det Norkse Veritas, 2010. Silent class notation - rules for classification of ships-newbuildings. URL https://rules.dnv.com/docs/pdf/dnvpm/rulesship/2010-01/ts624.pdf.

Di Franco, E., Pierson, P., Di Iorio, L., Calò, A., et al., 2020. Effects of marine noise pollution on mediterranean fishes and invertebrates: A review. Mar. Pollut. Bull. 159, 111450.

Duan, R., Ma, X., Wang, Y., Bailey, J., Qin, A.K., Yang, Y., 2020. Adversarial camouflage: Hiding physical-world attacks with natural styles. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition.

Duboue, P., 2020. The Art of Feature Engineering: Essentials for Machine Learning. Cambridge University Press.

Ekinci, S., Celik, F., Guner, M., 2010. A practical noise prediction method for cavitating marine propellers. Shipbuilding: Theory Pract. Shipbuilding Mar. Eng. 61 (4), 359–366.

Emerson, A., 1978. Propeller design and model experiments. Trans. North-East Coast Inst. Eng. Shipbuilders 944, 199–234.

European Union, 2017. United nations open-ended informal consulative process on oceans and the law of the sea - the effects of anthropogenic underwater noise, contribution of the European union. URL https://www.un.org/Depts/los/consultative_process/contributions_19cp/EU.pdf.

Fan, Y., Xu, K., Wu, H., Zheng, Y., Tao, B., 2020. Spatiotemporal modeling for nonlinear distributed thermal processes based on KL decomposition, MLP and LSTM network. IEEE Access 8, 25111–25121.

Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D., 2014. Do we need hundreds of classifiers to solve real world classification problems? J. Mach. Learn. Res. 15 (1), 3133–3181.

Ferrier-Pagès, C., Leal, M.C., Calado, R., Schmid, D.W., Bertucci, F., Lecchini, D., Allemand, D., 2021. Noise pollution on coral reefs?-A yet underestimated threat to coral reef communities. Mar. Pollut. Bull. 165, 112129.

Fisher, A., Rudin, C., Dominici, F., 2019. All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. J. Mach. Learn. Res. 20 (177), 1–81.

François, D., Wertz, V., Verleysen, M., 2006. The permutation test for feature selection by mutual information.. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.

Fraser, A., 1986. The Prediction and Minimisation of Propeller Induced Waterborne Noise. Technical report, British Maritime Technology Limited, Hydromechanics Division, Computational Hydromechanics Department.

Gaggero, S., 2020. Numerical design of a RIM-driven thruster using a RANS-based optimization approach. Appl. Ocean Res. 94, 101941.

Gaggero, S., Brizzolara, S., 2009. A panel method for trans-cavitating marine propellers. In: International Symposium on Cavitation.

Gaggero, S., Dubbioso, G., Villa, D., Muscari, R., Viviani, M., 2019. Propeller modeling approaches for off-design operative conditions. Ocean Eng. 178, 283–305.

Gaggero, S., Gonzalez-Adalid, J., Sobrino, M.P., 2016. Design of contracted and tip loaded propellers by using boundary element methods and optimization algorithms. Appl. Ocean Res. 55, 102–129.

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

Gaggero, S., Vernengo, G., Villa, D., 2021. A two-fidelity level approach for marine propeller design. In: Conference on Computational Methods in Marine Engineering.

Gaggero, S., Villa, D., 2017. Steady cavitating propeller performance by using OpenFOAM, StarCCM+ and a boundary element method. Proc. Inst. Mech. Eng., Part M: J. Eng. Marit. Environ. 231 (2), 411–440.

Gaggero, S., Villa, D., 2018. Cavitating propeller performance in inclined shaft conditions with OpenFOAM: PPTC 2015 test case. J. Mar. Sci. Appl. 17 (1), 1–20.

Gaggero, S., Villa, D., Brizzolara, S., 2010. RANS and PANEL method for unsteady flow propeller analysis. J. Hydrodyn., Ser. B 22 (5), 564–569.

Gaggero, S., Villa, D., Tani, G., Viviani, M., Bertetta, D., 2017b. Design of ducted propeller nozzles through a RANSE-based optimization approach. Ocean Eng. 145, 444–463.

Gaggero, S., Villa, D., Viviani, M., 2014. 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.

Gaggero, S., Viviani, M., Tani, G., Conti, F., Becchi, P., Valdenazzi, F., 2013. Comparison of different approaches for the design and analysis of ducted propellers. In: International Conference on Computational Methods in Marine Engineering.

Good, P., 2013. Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses. Springer Science & Business Media.

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

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., Evans, O., 2018. Viewpoint: When will AI exceed human performance? Evidence from AI experts. J. Artificial Intelligence Res. 62, 729–754.

Hainmueller, J., Hazlett, C., 2014. Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach. Polit. Anal. 5 (2), 143–168.

He, L., Tian, Y., Kinnas, S., 2011. MPUF-3A (version 3. 1) user's manual and documentation 11-1. URL https://www.caee.utexas.edu/prof/kinnas/home.html.

Hekler, A., Utikal, J.S., Enk, A.H., Solass, W., et al., 2019. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. Eur. J. Cancer 118, 91–96.

Hu, J., Wang, Y., Zhang, W., Chang, X., Zhao, W., 2019. Tip vortex prediction for contra-rotating propeller using large eddy simulation. Ocean Eng. 194, 106410.

Huang, Y.S., Dong, X.Q., Yang, C.J., Li, W., Noblesse, F., 2019. Design of wake-adapted contra-rotating propellers for high-speed underwater vehicles. Appl. Ocean Res. 91, 101880.

Huang, G.B., Wang, D.H., Lan, Y., 2011. Extreme learning machines: A survey. Int. J. Mach. Learn. Cybern. 2 (2), 107–122.

Ianniello, S., De Bernardis, E., 2015. Farassat's formulations in marine propeller hydroacoustics. Int. J. Aeroacoust. 14 (1–2), 87–103.

Ianniello, S., Muscari, R., Di Mascio, A., 2013. Ship underwater noise assessment by the acoustic analogy. Part I: nonlinear analysis of a marine propeller in a uniform flow. J. Mar. Sci. Technol. 18 (4), 547–570.

Ianniello, S., Muscari, R., Di Mascio, A., 2014a. Ship underwater noise assessment by the acoustic analogy Part II: hydroacoustic analysis of a ship scaled model. J. Mar. Sci. Technol. 19 (1), 52–74.

Ianniello, S., Muscari, R., Di Mascio, A., 2014b. Ship underwater noise assessment by the acoustic analogy, Part III: measurements versus numerical predictions on a full-scale ship. J. Mar. Sci. Technol. 19 (2), 125–142.

IMO, 2012. Resolution MSC. 337(91) (XII) - code on noise levels on board ships. URL https://wwwcdn.imo.org/localresources/en/KnowledgeCentre/IndexofIMOResolutions/Documents/MSC%20-%20Maritime%20Safety/337(91).pdf.

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

International Towing Tank Conference Propulsion Committee, 2008. Model manufacture, propeller models terminology and nomenclature for propeller geometry. In: Recommended Procedures and Guidelines 7. 5-01-02-01 International Towing Tank Conference.

International Towing Tank Conference Specialist Committee on Hydrodynamic Noise, 2017. Model scale cavitation test. In: Recommended Procedures and Guidelines 7. 5-01-02-05 International Towing Tank Conference.

Jiang, J.W., Yang, Y., Ren, T.W., Wang, F., Huang, W.X., 2021. Evolutionary optimisation for reduction of the low-frequency discrete-spectrum force of marine propeller based on a data-driven surrogate model. J. Mar. Sci. Eng. 9 (1), 18.

Jumper, J., Evans, R., Pritzel, A., Green, T., et al., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596 (7873), 583–589.

Kalikatzarakis, M., Coraddu, A., Atlar, M., Tani, G., Gaggero, S., Villa, D., Oneto, L., 2021. Computational prediction of propeller cavitation noise. In: Conference on Computational Methods in Marine Engineering.

Kamiirisa, H., Goto, H., 2005. Development of prediction method for ship underwater noise by bubble dynamics. Mitsui Zosen Tech. Rev. 2 (185), 38–44.

Katz, J., Plotkin, A., 2001. Low-Speed Aerodynamics. Cambridge University Press.

Keerthi, S.S., Lin, C.J., 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Comput. 15 (7), 1667–1689.

Kim, S., Kinnas, S.A., 2021. Prediction of cavitating performance of a tip loaded propeller and its induced hull pressures. Ocean Eng. 229, 108961.

Kipple, B., 2002. Southeast Alaska Cruise Ship Underwater Acoustic Noise, NSWCCD-71-TR-2002/574. Technical report, Naval Surface Warface Center.

Konno, A., Wakabayashi, K., Yamaguchi, H., Maeda, M., Ishii, N., Soejima, S., Kimura, K., 2002. On the mechanism of the bursting phenomena of propeller tip vortex cavitation. J. Mar. Sci. Technol. 6 (4), 181–192.

Koushan, K., 2000. Prediction of propeller induced pressure pulses using artificial neural networks. In: International Conference on Computer Application and Information Technology in Maritime Industries.

Lafeber, F.H., Bosschers, J., 2016. Validation of computational and experimental prediction methods for the underwater radiated noise of a small research vessel. In: Proceedings of PRADS.

Lakshmi, E., Priya, M., Achari, V.S., 2021. An overview on the treatment of ballast water in ships. Ocean Coast. Manag. 199, 105296.

Li, D., Hallander, J., Johansson, T., 2018. Predicting underwater radiated noise of a full scale ship with model testing and numerical methods. Ocean Eng. 161, 121–135.

Li, Y., Wang, L., 2020. A novel noise reduction technique for underwater acoustic signals based on complete ensemble empirical mode decomposition with adaptive noise, minimum mean square variance criterion and least mean square adaptive filter. Def. Technol. 16 (3), 543–554.

Lidtke, A.K., 2017. Predicting radiated noise of marine propellers using acoustic analogies and hybrid Eulerian-Lagrangian cavitation models (Ph.D. thesis). University of Southampton.

Lidtke, A.K., Lloyd, T., Vaz, G., 2019. Acoustic modelling of a propeller subject to non-uniform inflow. In: International Symposium on Marine Propulsors.

Lurton, X., 2002. An Introduction to Underwater Acoustics: Principles and Applications. Springer Science & Business Media.

Matusiak, J., 1992. Pressure and noise induced by a cavitating marine screw propeller (Ph.D. thesis). Aalto University.

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.

Miglianti, F., Cipollini, F., Oneto, L., Tani, G., Viviani, M., 2019. Model scale cavitation noise spectra prediction: Combining physical knowledge with data science. Ocean Eng. 178, 185–203.

Molnar, C., 2020. Interpretable Machine Learning. Lulu.com.

Naser, M.Z., Alavi, A.H., 2021. Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences. Archit. Struct. Constr. 1–19.

Nouri, N.M., Mohammadi, S., Zarezadeh, M., 2018. Optimization of a marine contra-rotating propellers set. Ocean Eng. 167, 397–404.

Oberkampf, W.L., Roy, C.J., 2010. Verification and Validation in Scientific Computing. Cambridge University Press.

Odabasi, Y.A., Fitzsimmons, P.A., 1978. Alternative methods for wake quality assessment. Int. Shipbuild. Prog. 25 (282), 34–42.

Odabaşi, Y.A., Fitzsimmons, P.A., 1978. Alternative methods for wake quality assessment. Int. Shipbuild. Prog. 25 (282), 34–42.

Okamura, N., Asano, T., 1988. Prediction of propeller cavitation noise and its comparison with full scale measurement. J. Soc. Naval Archit. Japan 1988 (164), 43–53.

Oldenburg, M., Baur, X., Schlaich, C., 2010. Occupational risks and challenges of seafaring. J. Occup. Health 1007–1036.

Oneto, L., 2020. Model Selection and Error Estimation in a Nutshell. Springer.

Oneto, L., Cipollini, F., Miglianti, L., Tani, G., Gaggero, S., Viviani, M., Coraddu, A., 2020. Deep learning for cavitating marine propeller noise prediction at design stage. In: International Joint Conference on Neural Networks.

Pennings, P.C., Westerweel, J., Van Terwisga, T.J.C., 2015. Flow field measurement around vortex cavitation. Exp. Fluids 56 (11), 1–13.

Raestad, A.E., 1996. Tip vortex index-an engineering approach to propeller noise prediction. Naval Archit..

Roddy, R., Hess, D., Falleer, W., 2008. Utilizing neural networks to predict forces and moments on a submarine propeller. In: AIAA Aerospace Sciences Meeting and Exhibit.

Ross, D., 1976. Mechanics of Underwater Noise. Peninsula Publishing.

Sainani, K.L., 2016. The value of scatter plots. PM&R 8 (12), 1213–1217.

Salvatore, F., Ianniello, S., 2003. Preliminary results on acoustic modelling of cavitating propellers. Comput. Mech. 32 (4), 291–300.

Salvatore, F., Streckwall, H., Van Terwisga, T., 2009. Propeller cavitation modelling by CFD-results from the VIRTUE 2008 rome workshop. In: International Symposium on Marine Propulsors.

Seol, H., Jung, B., Suh, J.C., Lee, S., 2002. Prediction of non-cavitating underwater propeller noise. J. Sound Vib. 257 (1), 131–156.

Seol, H., Suh, J.C., Lee, S., 2005. Development of hybrid method for the prediction of underwater propeller noise. J. Sound Vib. 288 (1–2), 345–360.

Sezen, S., Cosgun, T., Yurtseven, A., Atlar, M., 2021a. Numerical investigation of marine propeller underwater radiated noise using acoustic analogy Part 1: The influence of grid resolution. Ocean Eng. 220, 108448.

Sezen, S., Cosgun, T., Yurtseven, A., Atlar, M., 2021b. Numerical investigation of marine propeller underwater radiated noise using acoustic analogy Part II: The influence of eddy viscosity turbulence models. Ocean Eng. 220, 108353.

Sha, Y., Faber, J., Gou, S., Liu, B., Li, W., Schramm, S., Stoecker, H., Steckenreiter, T., Vnucec, D., Wetzstein, N., 2022. A multi-task learning for cavitation detection and cavitation intensity recognition of valve acoustic signals. Eng. Appl. Artif. Intell. 113, 104904.

Shalev-Shwartz, S., Ben-David, S., 2014. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press.

Shao, Q., Faizollahzadeh A., S., Mafarja, M., Turabieh, H., Zhang, Q., Band, S.S., Chau, K., Mosavi, A., 2021. Diffusion analysis with high and low concentration regions by the finite difference method, the adaptive network-based fuzzy inference system, and the bilayered neural network method. Eng. Appl. Comput. Fluid Mech. 15 (1), 1392–1399.

Shawe-Taylor, J., Cristianini, N., 2004. Kernel Methods for Pattern Analysis. Cambridge University Press.

Siemens Digital Industries Software, 2021. Simcenter STAR-CCM+ user guide, version 2021. 1. URL https://www.plm.automation.siemens.com/global/en/products/simcenter/STAR-CCM.html.

Silver, D., Schrittwieser, J., Simonyan, K., et al., 2017. Mastering the game of go without human knowledge. Nature 550 (7676), 354–359.

Song, S., Demirel, Y.K., Muscat-Fenech, C., Tezdogan, T., Atlar, M., 2020. Fouling effect on the resistance of different ship types. Ocean Eng. 216, 107736.

Sun, H., Young, J., Kinnas, S.A., 2004. HULLFPP (version 1. 3) HULL field point potential-user's manual. URL https://www.caee.utexas.edu/prof/kinnas/home.html.

Testa, C., Ianniello, S., Salvatore, F., Gennaretti, M., 2008. Numerical approaches for hydroacoustic analysis of marine propellers. J. Ship Res. 52 (01), 57–70.

Tong, X.D., Chen, Y., 2020. Random response of highly skewed propeller-shafting system induced by inflow turbulence. Ocean Eng. 195, 106750.

Tong, X.D., Chen, H.Y., Chen, Y., 2021. Low frequency broadband noise radiated by highly skewed propeller ingesting inflow turbulence. J. Sound Vib. 490, 115709.

Torres, F.G., De-la Torre, G.E., 2021. Environmental pollution with antifouling paint particles: Distribution, ecotoxicology, and sustainable alternatives. Mar. Pollut. Bull. 169, 112529.

Trivyza, N., Rentizelas, A., Theotokatos, G., Boulougouris, E., 2021. Decision support methods for sustainable ship energy systems: A state-of-the-art review. Energy 122288.

Tucker, J.D., Azimi-Sadjadi, M.R., 2011. Coherence-based underwater target detection from multiple disparate sonar platforms. IEEE J. Ocean. Eng. 36 (1), 37–51.

Vakili, S.V., Ölçer, A.I., Ballini, F., 2020a. The development of a policy framework to mitigate underwater noise pollution from commercial vessels: The role of ports. Mar. Policy 120, 104132.

Vakili, S.V., Ölçer, A.I., Ballini, F., 2020b. The development of a policy framework to mitigate underwater noise pollution from commercial vessels: The role of ports. Mar. Policy 120, 104132.

Valdenazzi, F., Conti, F., Gaggero, S., Vaccaro, C., Grassi, D., Villa, D., 2019. A practical tool for the hydro-acoustic optimization of naval propellers. In: International Conference on Computational Methods in Marine Engineering.

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

Wainberg, M., Alipanahi, B., Frey, B.J., 2016. Are random forests truly the best classifiers? J. Mach. Learn. Res. 17 (1), 3837–3841.

Wang, Y., Wang, K., Abdel-Maksoud, M., 2021. Noisenet: A neural network to predict marine propellers' underwater radiated noise. Ocean Eng. 236, 109542.