Journal of Sailing Technology 2021, volume 6, issue 1, pp. 21 – 43.
The Society of Naval Architects and Marine Engineers.

Multi-Fidelity Surrogate Models for VPP Aerodynamic Input Data

Tanya Peart
University of Auckland and Doyle Sails, New Zealand, tanya.peart@doylesails.com.

Nicolas Aubin
Doyle Sails, New Zealand.

Stefano Nava
Doyle Sails, New Zealand.

John Cater
University of Auckland, New Zealand.

Stuart Norris
University of Auckland, New Zealand.

Manuscript received October 5, 2020; revision received December 29, 2020; accepted February 2, 2021.

Abstract. Predicting the performance of a sail design is important for optimising the performance of a yacht, and Velocity Prediction Programs (VPPs) are commonly used for this purpose. The aerodynamic force data for a VPP is often calculated using Computational Fluid Dynamics (CFD) models, but these can be computationally expensive. A full VPP analysis for sail design is therefore usually restricted to high-budget design projects or research activities and is not practical for many industry projects.

This work presents a method to reduce the computational cost of creating lift and drag force coefficient curves for input into a VPP using both multi-fidelity kriging surrogate modelling and data from existing sail designs. This method is shown to reduce the number of CFD simulations required for a desired accuracy when compared to a single-fidelity model. A maximum reduction in the required computational effort of 57% was achieved for model-scale symmetric spinnaker sails. For the same number of simulations, the accuracy of the model predictions was improved by up to 72% for scale-symmetric spinnaker sails, and 90% for asymmetric spinnakers.

Keywords: sail aerodynamics; surrogate model; multi-fidelity; spinnakers; existing data.

NOMENCLATURE

k	Number of dimensions of the surrogate model
L	Likelihood
n	Number of observed values
t	Fidelity level
s	Total number of fidelity levels
$\hat{s}^2$	Estimate of mean square error of model predictions
x_i	Infill point
$\boldsymbol{x}^*$	Prediction point
$\boldsymbol{X}$	Set of sample data
$\boldsymbol{X}_t$	High-fidelity sample data
$\boldsymbol{X}_{t-1}$	Low-fidelity sample data
$\boldsymbol{y}$	Set of observed values
$\boldsymbol{y}_t$	High-fidelity observed values
$\boldsymbol{y}_{t-1}$	Low-fidelity observed values
$\hat{y}$	Maximum likelihood estimate of model prediction
$\boldsymbol{Y}$	Set of random vectors used to represent the function value
Z	Kriging model
Z_t	Kriging high-fidelity model
Z_{t-1}	Kriging low-fidelity model
δ	Model of the difference between the low- and high-fidelity points
$\boldsymbol{\theta}$	Kriging hyperparameter
$\boldsymbol{\theta_{t-1}}$	Kriging hyperparameter for the lower fidelity level model
$\boldsymbol{\theta_d}$	Kriging hyperparameter for the difference model
μ	Mean
$\hat{\mu}$	Maximum likelihood estimate of the mean
ρ_{t-1}	Scaling factor between fidelity levels
σ^2	Variance
$\hat{\sigma}^2$	Maximum likelihood estimate of the variance
$\boldsymbol{\psi}$	Correlation vector between observed data and prediction point
$\boldsymbol{\psi_t}$	Correlation vector between observed data and prediction point at fidelity level t
$\boldsymbol{\Psi}$	Correlation matrix between observed data points
$\boldsymbol{\Psi_t}$	Correlation matrix between observed data points at fidelity level t
ABL	Atmospheric Boundary Layer
ANN	Artificial Neural Network
AWA	Apparent Wind Angle
CFD	Computational Fluid Dynamics
FSI	Fluid Structure Interaction
MFK	Multi-Fidelity Kriging
MLE	Maximum Likelihood Estimate
MSE	Mean Square Error
RANS	Reynolds-Averaged Navier-Stokes
RMS	Root Mean Square
SFK	Single-Fidelity Kriging
SYRF	Sailing Yacht Research Foundation
TP52	Transpac 52
VPP	Velocity Prediction Program

1. INTRODUCTION

Small improvements in the performance of a racing yacht can have a large impact on the outcome of a race, making it important to be able to predict the performance at the design stage (Lasher, 2011). One of the most common methods for this is a Velocity Prediction Program (VPP) (Fossati, 2009). By applying the principles of equilibrium, a VPP allows performance measures such as the speed and heel angle of a yacht to be predicted by balancing the forces and moments for up to three different axes.

In a VPP, equilibrium is usually calculated at points corresponding to a certain set of wind conditions, specified by the true wind angle and true wind speed (Fossati, 2009). To calculate the equilibrium, the aerodynamic forces at this wind input condition need to be found, along with the hydrodynamic forces resulting from the yacht's speed in these conditions (Alza, 2012). Depending on the number of forces and moments that are balanced in the VPP, both the aero- and hydrodynamic forces can also include other variables such as the heel angle (Korpus, 2007). The accuracy of a VPP is strongly dependent on the accuracy of the input forces (Lasher *et al.*, 2003).

There are several methods to obtain the aero- and hydrodynamic forces. The forces can be directly determined through experimental tests or numerical simulations for each required input combination (Roux *et al.*, 2008; Böhm and Graf, 2010). However, this can require many tests or simulations due to the iterations which may be required for convergence of the equilibrium. When flow separation or other viscous effects are important, such as for the flow around downwind sails, Reynolds-Averaged Navier-Stokes (RANS) Computational Fluid Dynamics (CFD) modelling needs to be used (Lasher and Richards, 2007). This can have a high computational cost (Doyle *et al.*, 2016). Analytical formulations have also been proposed, such as the IMS model for sail aerodynamics (Fossati *et al.*, 2006). While models of this type are relatively quick to compute, the predicted lift and drag coefficients calculated only consider the sail area, not the design shape of the specific sail. This limits the accuracy of these methods, particularly when performance of a certain sail design is the goal (Korpus, 2007).

Surrogate models (also known as metamodels or response surface models) are a third method for determining the input forces for a VPP, which can balance computational cost and accuracy. The models can reduce the computational time required to evaluate complex functions by using statistical techniques to fit a curve or surface through pre-computed data points (Forrester *et al.*, 2008). Using interpolation, the previously unknown value of a function at a certain point can be predicted by querying the surrogate model. Therefore, surrogate models allow a complex function which may take hours or days to be solved by a simulation program to be replaced with a function that can be evaluated in a few seconds (Harries and Abt, 2019). This technique can be used in a VPP to interpolate previously obtained experimental or numerical data for the required aerodynamic and hydrodynamic input force curves (Ruiz de Elvira, 2015). The determination of the aerodynamic input forces is the focus of this study, using pre-computed CFD data points.

Surrogate modelling techniques have previously been presented to approximate the input force curves for a VPP. Cubic spline polynomial methods have been used (Roux *et al.*, 2002; Korpus, 2007; Doyle *et al.*, 2016), as well as the more complex methods of radial basis functions (RBF) (Prince and Claughton, 2013) and kriging (Ploé *et al.*, 2017). Kriging has also been used in other yacht design applications such as the optimisation of sail trim (Sacher *et al.*, 2017). Artificial neural networks (ANNs) have been trained for both yacht aero- and hydrodynamic data (Mason, 2010; Courtel, 2020). However, all the surrogate models previously used in VPPs have been restricted to a single fidelity. Single-fidelity surrogate models interpolate the data points and therefore reduce the number of computations which are required when compared to direct computation techniques. A large number of data points, and therefore many simulations, can be required to construct an accurate force coefficient curve or surface (Korpus, 2007). The use of RANS for these simulations leads to a significant computational cost, which is prohibitive in the sail design industry without access to high performance computing facilities, that are usually only available for high-budget projects (Doyle *et al.*, 2016; Courtel, 2020).

This computational cost can be further reduced through the use of multi-fidelity methods. These methods combine data points from two or more fidelity (accuracy) levels to construct a single surrogate model (Fernández-Godino *et al*., 2019). The changes in fidelity presented in literature have been derived through simplifying the physics of the model, using a different simulation solver (Pellegrini *et al.,* 2016), changing the mesh resolution (Serani *et al*., 2019), or using a combination of experimental and numerical methods (Kuya *et al.,* 2011).

Using multiple fidelities in a surrogate model can improve the accuracy for the same computational cost (de Baar and Roberts, 2015). The high-fidelity solutions increase the accuracy of the prediction value, but when the number of these solutions is limited, points in between the observed values can be inaccurate. The inclusion of low-fidelity solutions increases the coverage of the data points, and therefore improves the prediction of the model trends.

The use of multiple fidelities to construct a single surrogate model has been presented for non-VPP applications in the marine industry. Kriging with fidelity differences due to mesh size has been used to estimate the uncertainty of input parameters for yacht hydrodynamics (Baar *et al.,* 2015). RBF models for hydrodynamic resistance evaluation (Pellegrini *et al.,* 2016) and optimisation (Pellegrini *et al*., 2018) have been presented, both using potential flow and RANS as different fidelity solvers. The use of Artificial Neural Networks (ANNs) has also been presented for sail aerodynamics (Peña *et al.,* 2012), where the ANN corrects low-fidelity potential flow simulations with high-fidelity wind tunnel data. However, to the authors knowledge, none of the multi-fidelity approaches presented have been used for VPP input data.

Using existing data also has the potential to reduce the computational cost. The input parameter combinations for a converged VPP calculation can be used to determine the best locations for CFD solutions for a calculation of a similar design (Korpus, 2007; Böhm and Graf, 2010). This helps to concentrate the computational effort in areas of interest but does not reduce the total number of calculations required. Another study suggested the use of a known force curve for a parent design to predict the force curve of a parametrically similar design, but no implementations of this method were shown (Ruiz de Elvira, 2015). Surrogate models were not used, instead the data from the parent model was transformed to the new design by applying a series of geometrical transformations.

In this work, a novel methodology was developed to predict the lift or drag force coefficient curve of a sail design using a multi-fidelity kriging surrogate model and data from an existing similar sail design. RANS CFD simulations are used for both the high- and low-fidelity points, with the difference in fidelity due to the shape of the sail analysed. An initial investigation into the performance of this methodology is presented, applied to two pre-computed data sets consisting of symmetric and asymmetric spinnaker sails. The accuracy of the method presented is compared to a single-fidelity kriging model.

Section 2 describes the theoretical background for single- and multi-fidelity kriging. Section 3 discusses the methods for the data collection of the two sail data sets, and the construction of the single- and multi-fidelity surrogate models. Section 4 presents the performance of the method. Section 5 discusses the limitations of the model, followed by the conclusions in Section 6.

2. BACKGROUND

2.1. Kriging

The surrogate models constructed in this work are based on the original kriging method developed by Krige (1951) and applied to the approximation of computer experiments in engineering design by Sacks *et al*. (1989). The equations used to construct the single-fidelity kriging surrogate models are based on Kennedy and O'Hagan (2000), for further explanation also see Forrester *et al.* (2007), and Forrester *et al.* (2008). The training data are assumed to be without noise.

Given a set of sample data, $\boldsymbol{X} = [\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(n)}]^{T}$, with n observed values, and observed responses, $\boldsymbol{y} = [y^{(1)}, y^{(2)}, \ldots, y^{(n)}]^{T}$, we want to predict the value of the function at new points $\boldsymbol{x}^{*}$. To do this, a kriging model $Z(\boldsymbol{x})$ is constructed. We assume that the value of the function at $\boldsymbol{x}^{*}$ is the realisation of a stochastic process (even if the data is from a deterministic computer code), represented by a set of random vectors $\boldsymbol{Y} = \left[Y(\boldsymbol{x}^{(1)}), \ldots, Y(\boldsymbol{x}^{(n)})\right]^{T}$. This set of random vectors is defined by a mean $\mathbf{1}\mu$, where $\mathbf{1} = [1, \ldots, 1]^{T}$, and a covariance which is related to the correlation, $\boldsymbol{\Psi}$, by Equation 1 where σ^{2} is the variance of the function.

$$\mathrm{cov}(\mathbf{Y}, \mathbf{Y}) = \sigma^{2}\boldsymbol{\Psi} \tag{1}$$

The correlation between each of the training points is modelled using the Gaussian correlation function in Equation 2, with k the number of dimensions of the function. Stationarity of the model is assumed, where the mean is constant, and the covariance between points in the model is dependent only on the absolute distance between the sample points, not on their position within the domain. Points that have a similar location are assumed to have a similar function value. The correlation function includes the hyperparameter θ_j which forms the vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_k]$. This vector determines the extent of the influence of the training points.

$$\boldsymbol{\Psi}\left[Y\left(x^{(i)}\right), Y\left(x^{(l)}\right)\right] = \exp\left(-\sum_{j=1}^{k} \theta_j \left|x_j^{(i)} - x_j^{(l)}\right|^{2}\right) \tag{2}$$

The value of the kriging hyperparameter $\boldsymbol{\theta}$ for each model is found by maximising the likelihood of the model fitting the measured data $\mathbf{y}$ using a Maximum Likelihood Estimate (MLE) of the concentrated log-likelihood function in Equation 3. The MLE of the mean, $\hat{\mu}$ is calculated in Equation 4, and $\hat{\sigma}^{2}$ is the MLE of the variance in Equation 5. As the value of this log-likelihood function is dependent on the correlation matrix $\boldsymbol{\Psi}$ and therefore $\boldsymbol{\theta}$, the value of $\boldsymbol{\theta}$ is found which maximises Equation 3 using the SciPy (Virtanen *et al.,* 2020) differential evolution algorithm between the bounds of 1×10^{-6} and 1×10^{2}.

$$\log(L) \approx -\frac{n}{2}\log(\hat{\sigma}^{2}) - \frac{1}{2}\ln|\boldsymbol{\Psi}| \tag{3}$$

$$\hat{\mu} = \frac{\mathbf{1}^{T}\boldsymbol{\Psi}^{-1}\boldsymbol{y}}{\mathbf{1}^{T}\boldsymbol{\Psi}^{-1}\mathbf{1}} \tag{4}$$

$$\hat{\sigma}^{2} = \frac{(\boldsymbol{y} - \mathbf{1}\hat{\mu})^{T}\boldsymbol{\Psi}^{-1}(\boldsymbol{y} - \mathbf{1}\hat{\mu})}{n} \tag{5}$$

The kriging model is then used to make a prediction, which needs to be consistent with observed data. This is achieved by maximising the likelihood of the prediction given the observed sample values and the optimised kriging hyperparameters. A correlation vector $\boldsymbol{\psi}$ between the training data points and the prediction point $\boldsymbol{x}^{*}$ is constructed, as in Equation 6. This results in an MLE for the prediction $\hat{y}(\boldsymbol{x}^{*})$ in Equation 7. The mean squared error (MSE) of the model predictions can be estimated using Equation 8.

$$\boldsymbol{\psi} = \left[\exp\left(-\sum_{j=1}^{k} \theta_j \left|x_j^{*} - x_j^{(1)}\right|^{2}\right), \ldots, \exp\left(-\sum_{j=1}^{k} \theta_j \left|x_j^{*} - x_j^{(n)}\right|^{2}\right)\right]^{T} \tag{6}$$

$$\hat{y}(\boldsymbol{x}^{*}) = \hat{\mu} + \boldsymbol{\psi}^{\mathrm{T}}\,\boldsymbol{\Psi}^{-1}(\boldsymbol{y} - \mathbf{1}\hat{\mu}) \tag{7}$$

$$\hat{s}^{2}(\boldsymbol{x}^{*}) = \hat{\sigma}^{2}\left[1 - \boldsymbol{\psi}^{T}\boldsymbol{\Psi}^{-1}\boldsymbol{\psi} + \frac{1 - \mathbf{1}^{T}\boldsymbol{\Psi}^{-1}\boldsymbol{\psi}}{\mathbf{1}^{T}\boldsymbol{\Psi}^{-1}\mathbf{1}}\right] \tag{8}$$

2.2. Multi-Fidelity Kriging

Multi-fidelity kriging (MFK) is based on the kriging equations with data at two or more fidelity levels. The MFK models constructed in this work using the SMT package (Bouhlel *et al.*, 2019) are based on the recursive formulation of Le Gratiet (2013). This formulation decouples the multi-fidelity model into separate kriging models for each fidelity level, t, out of the total fidelity levels, s. The model for a particular fidelity level, $Z_t(\boldsymbol{x})$, is constructed using the training data from that level as well as all the lower fidelity levels. The models in this work are restricted to two fidelity levels.

The process starts with a set of low-fidelity training points $\boldsymbol{X}_{t-1}$ with observed values $\boldsymbol{y}_{t-1}$, and high-fidelity training points $\boldsymbol{X}_t$ with observed values $\boldsymbol{y}_t$. The low- and high-fidelity data points are assumed to be independent. Where a location has both low- and high-fidelity data points available, the high-fidelity values are assumed to be correct, with differences in the low- and high-fidelity values assumed to be due to errors or uncertainty in the low-fidelity data.

The kriging model of the high-fidelity data, $Z_t(\boldsymbol{x})$, is constructed in Equation 9 as the predictions of the low-fidelity model, $Z_{t-1}(\boldsymbol{x})$, multiplied by a scaling factor $\rho_{t-1}(x)$, and a difference function $\delta(x)$ which also uses kriging to model the difference between the low- and high-fidelity data. The scaling factor is constant for the models in this work, set as $\rho_{t-1}(\boldsymbol{x}) = [1, \dots, 1]^T$.

$$Z_t(\boldsymbol{x}) = \rho_{t-1}(\boldsymbol{x}) \times Z_{t-1}(\boldsymbol{x}) + \delta(\boldsymbol{x}), \qquad t = 2, \dots, s \tag{9}$$

The de-coupled kriging models at each fidelity are constructed in the same way as single-fidelity kriging, with Equation 3 to construct $\boldsymbol{\Psi}_t$ for the correlation between the training data points used for that model. Likewise, to make a prediction for each model, Equation 6 is used to construct $\boldsymbol{\psi}_t$. The hyperparameter vectors $\boldsymbol{\theta}_{t-1}$ for the model $Z_{t-1}$and $\boldsymbol{\theta}_d$ for the difference model δ therefore are also estimated using the MLEs in Equations 4 and 5. The prediction of the model for a new location, $\boldsymbol{x}_t^*$, at fidelity level *t* can then be calculated using Equation 10. The MSE of the prediction can be estimated with Equation 11.

$$\hat{y}_t(\boldsymbol{x}_t^*) = \hat{\mu}_t + \hat{\rho}_{t-1}\hat{y}_{t-1}(\boldsymbol{x}_t^*) + \boldsymbol{\psi}_t^T \boldsymbol{\Psi}_t^{-1}\left(y_t(\boldsymbol{x}_t) - \mathbf{1}\hat{\mu}_t - \hat{\rho}_{t-1}\hat{y}_{t-1}(\boldsymbol{x}_t)\right) \tag{10}$$

$$\hat{s}^2(\boldsymbol{x}_t^*) = \hat{\rho}_{t-1}^2 \hat{s}_{t-1}^2(\boldsymbol{x}_t^*) + \hat{\sigma}_t^2(1 - \boldsymbol{\psi}_t^T \boldsymbol{\Psi}_t^{-1} \boldsymbol{\psi}_t) \tag{11}$$

2.3. Adaptive Sampling

Adaptive sampling is a process that can be used to determine the training data points for a surrogate model (Forrester *et al*., 2008). To make the most of the available computational resources, rather than selecting all the training points *a-priori* according to a sampling plan, points are dynamically added to the training data set based on a criterion (Pellegrini *et al.*, 2018). The placement of the training points is therefore influenced by information provided by the surrogate model itself.

The focus of the surrogate models in this work is exploration, with the aim to achieve an accurate model over the full range of AWAs. The adaptive sampling strategy used in this work therefore differs from that used in the more common surrogate model application of optimisation. The goal of the strategy is to place a new infill point, x_i at locations where the kriging model predicts the maximum MSE, as in Equation 12. The MSE is calculated from Equation 8 for single-fidelity kriging or Equation 11 for multi-fidelity kriging.

$$x_i = \operatorname{argmax}(\hat{s}^2(\boldsymbol{x})) \tag{12}$$

3. METHODOLOGY

The multi-fidelity surrogate model methodology presented in this paper is evaluated using two separate data sets of downwind sails. The first consists of RANS simulations previously undertaken on a series of parametrically varying rigid symmetric spinnaker sails (Peart, 2018). These sails are illustrated in Figure 1 and are named Sail I to IV. The collection of these data points and the construction of the corresponding surrogate models is discussed in Section 3.1. The second data

set is provided by the Sailing Yacht Research Foundation (SYRF) (Braun, 2016) where Fluid Structure Interaction (FSI) simulations were performed on several asymmetric spinnaker sails, discussed in Section 3.2.

The surrogate models constructed for both data sets use the same methodology. Multi-fidelity surrogate models commonly use a difference in mesh resolution (de Baar and Roberts, 2015; Serani *et al.*, 2019), or a difference in solver (Pellegrini *et al*., 2016; Pellegrini *et al*., 2018), to define the high and low-fidelity training points. This work presents an alternate hierarchical approach, where fidelity is instead a measure of how a set of training points model the problem. High-fidelity points directly model the problem. They are obtained from CFD simulations which model the exact sail for which the force coefficient curve is to be predicted. In contrast, low-fidelity points are acquired for similar, but different, sail designs (noting that differences in sail shape result in differences in the force coefficient curves compared to the sail to be predicted). High-fidelity points for the sail being modelled are combined with more numerous low-fidelity points from one of the other sail designs to create a single multi-fidelity surrogate model curve.

For example, when the object of the surrogate model is to predict the force coefficient curves of Sail I, the CFD force coefficient data points for Sail I directly model this curve and are therefore considered to be high-fidelity training points. When data points from one of the other sail shapes (Sails II – IV, which have different force coefficients than Sail I) are used in the same MFK surrogate model, they are low-fidelity training points.

For the application of this method to real-world cases, the low-fidelity data points will be from different but similar sail designs which have previously been analysed, referred to as “existing sails”. The low-fidelity points will therefore have already been calculated and are associated with no additional computational cost. To train a new model, only new high-fidelity CFD simulations will need to be performed. When analysing the computational cost required to construct a multi-fidelity model of a certain accuracy, consequently only the number of high-fidelity simulations required is considered.

The MFK surrogate modelling methodology is implemented in this project using the open-source Surrogate Modelling Toolbox (SMT) (Bouhlel *et al*., 2019) Python package. The kriging method was chosen due to its comparative performance against other methods in previous testing (Jin *et al.*, 2001; Zhao and Xue, 2010; Alizadeh *et al.*, 2020), and its ability to predict the error of the surrogate model prediction (Forrester *et al.*, 2008). SMT’s MFK surrogate model object was used as the basis, with functionality added for adaptive sampling. The set-up using the SMT package implements ordinary kriging where the mean of the model $Z(\boldsymbol{x})$ is assumed to be constant.

After the SFK and MFK models were trained and adaptively sampled, the accuracy of the model predictions was assessed. For exploration surrogate models, where the overall accuracy of the model is important and all points of the model are weighted equally, a suitable approach for evaluating the accuracy of the model is the RMS error (Forrester *et al*., 2008). The RMS error in this work is calculated between the predictions of the surrogate models and the full CFD data set of the high-fidelity sail. When adaptive sampling is considered, this RMS error is calculated after each adaptive sampling cycle so that the changes in model accuracy with number of high-fidelity data points can be evaluated.

3.1. Symmetric Spinnaker Sails

3.1.1 Data Collection

The first data set used in this paper consists of a series of four semi-rigid symmetric spinnaker sails, Sails I – IV. The sails are 1/25th scale and all have a nominal surface area of 0.8 m^2. The sail shapes vary with parametric equations in their width and depth, details of which can be found in Moore (2004), and Lasher and Richards (2007). In Figure 1 it can be seen that Sail I has a flat shape with a wide mid-girth and shoulder, while the mid-girth distance and shoulder decreases and the depth of the sail increases progressively to Sail IV.

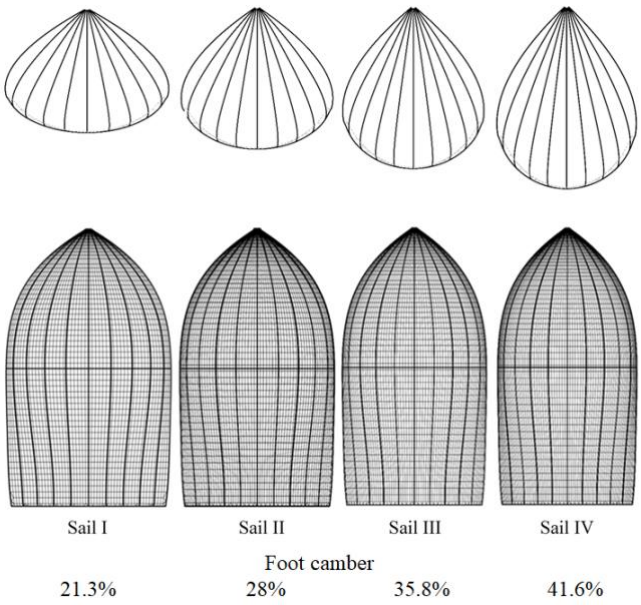
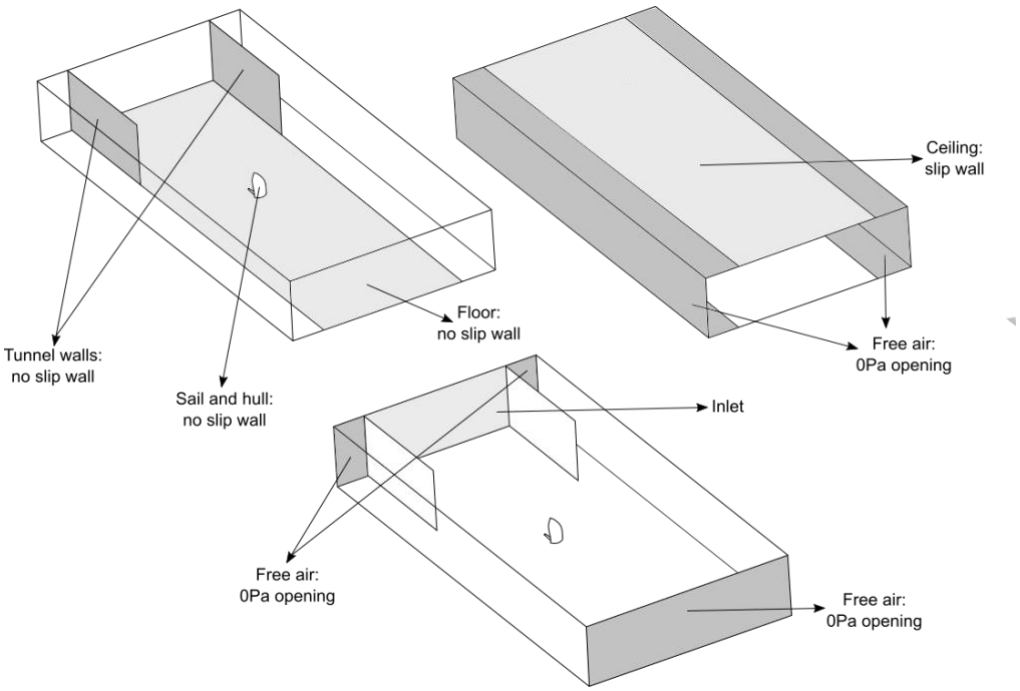


Figure 1: Parametrically varying series of symmetric spinnaker sails with the foot camber specified.

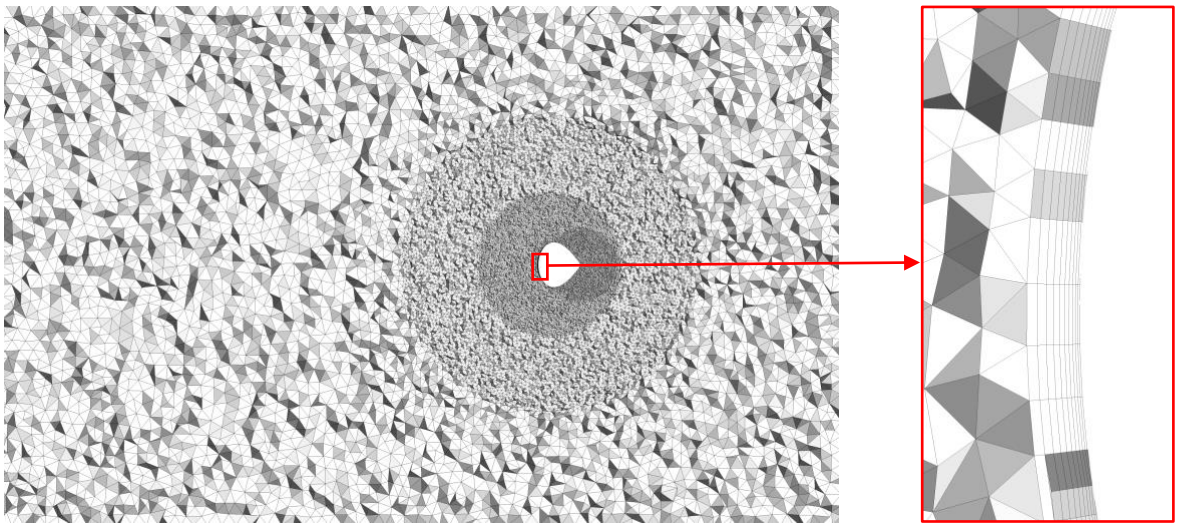
Figure 2: Experimental testing twisted flow wind tunnel set-up with the yacht model, Sail II, and a cobra probe to the left of the model (Peart, 2018).

Testing was performed for each sail over a range of AWAs with a fixed apparent wind speed. From analysis of the pressure distributions on Sail I from the CFD modelling of Moore (2004), the lowest practical AWA for the flying of the spinnaker, where the luff is just starting to curl, is likely to be between 110° and 120°. The lift and drag coefficients for each of the sails were therefore calculated for AWAs between 110° and 180° at 5° increments. As the sails alter in shape, this lower AWA limit will not be accurate for all the sails, causing variability in the accuracy of data points, particularly for AWAs lower than 120°.

RANS CFD is used, rather than more computationally expensive methods, as it is commonly applied to spinnaker sail modelling (Lasher and Richards, 2007; Viola, 2009; Ramolini, 2019; Cirello *et al.*, 2019). The RANS model was set up to match experimental testing conducted in the Twisted Flow Wind Tunnel by Moore (2004), Lasher and Richards (2007), and Peart (2018) as closely as possible. An example of the experimental arrangement from Peart (2018) is presented in Figure 2. In this set-up, semi-rigid scale models of the symmetric spinnaker sails were used, with their position aligned with the centre line of the hull. The hull was then rotated for each of the different AWAs. To improve the rigidity of the sail model, the clew and tack points were attached with poles in addition to the clew and brace lines. The trim of the sails was not considered in the CFD simulations as it was not changed during the experimental set-up. It can therefore be considered that the variation of AWA is



(a) Domain modelling of the twisted flow wind tunnel, and boundary conditions for each face.



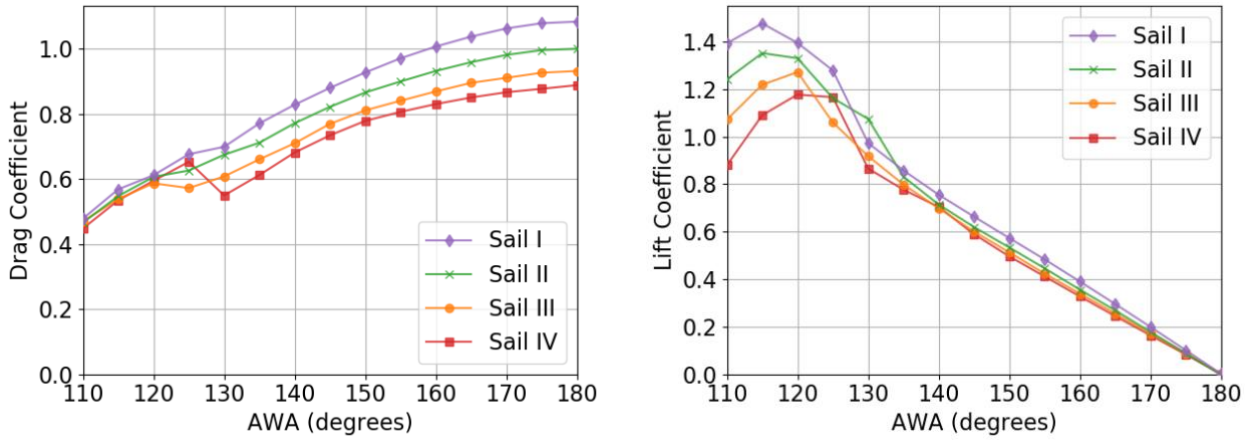
(b) Overall mesh refinement around the sail and inflation layers on the sail surface, total of 1×10^6 cells.

Figure 3: Modelling details of the ANSYS CFX simulations to replicate the twisted flow wind tunnel conditions.

a change only in the angle of attack, without any trimming effect. The spinnaker sails were tested without a mainsail present.

The CFD model for each of the sails used a uniform inlet velocity profile taken from the experimental testing where there is only a small boundary layer present. This profile had a reference velocity of 3.25 m/s at a height of 0.4 m (an equivalent full-scale height of 10 m), selected to match the previous experimental test set-up (a cobra probe measuring the velocity can be seen on the left-hand side of Figure 2). This reference velocity was used to compute the force coefficients, with the head of the sails positioned at a height of 1.36 m above the floor of the domain. The set-up of the other boundary conditions and the CFD domain is presented in Figure 3a. The full test area of the wind tunnel was included in the CFD domain to properly capture the flow conditions (Nava *et al*., 2018).

Each of the sail models used the same unstructured mesh parameters generated by the ANSYS® CFX meshing tool, with approximately 1×10^6 cells, including prismatic inflation layers on the sail surfaces as well as around the hull and on the floor of the domain. Mesh convergence was performed



(a) Drag coefficients. **(b) Lift coefficients.**

Figure 4: Force coefficient data from CFD modelling, Sails I-IV tested with a uniform inlet velocity profile.

for the 180° case, with a grid convergence index of 0.2%. The y^+ values were in the range 0.88 − 10.9. A depiction of the mesh for Sail I at an AWA of 180° is presented in Figure 3b.

The simulations were conducted using RANS modelling in ANSYS® CFX release 18.2, using 8 cores on an Intel® Xeon® CPU E5-2698 machine with 16 GB RAM and a clock speed of 2.30 GHz. Based on the findings of Norris and Richards (2012), the SST turbulence model was used. Each RANS simulation took approximately 1.5 hours to complete, with the simulations terminated when the residuals reached a convergence of 1×10^{-7}. Validation against experimental data points was performed, with a Root Mean Square (RMS) error of 0.06 for the Sail I drag coefficients, corresponding to 7.1% of the mean drag coefficient value.

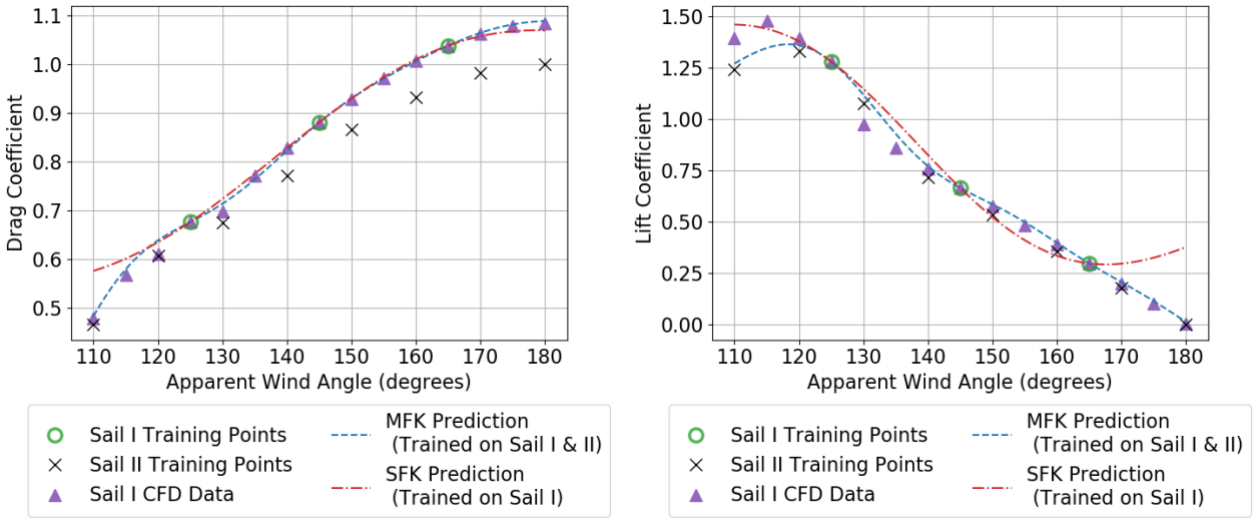
The drag coefficients for the four sails investigated in this paper are shown in Figure 4a, and the lift coefficients in Figure 4b. The drag coefficients for all the sails show similar trends for AWAs above 130°, with the different sail geometries leading to different drag performance. The trend of the drag coefficient curve for Sail IV differs from the other sails between 120° and 130°, having a much sharper change. This is caused by a drop in lift, and therefore the induced drag at this point, as seen in the lift coefficients from Figure 4b. As the lift is a maximum in this region, the induced drag is a significant component of the overall drag, resulting in the noticeable overall drag coefficient drop. The drag coefficient curve for Sail II is the closest to that of Sail I, with Sail IV being the least similar. This matches the sail geometries in Figure 1, where the geometry of Sail II is the most similar to Sail I. The lift coefficients in Figure 4b also show similar trends for AWAs above 130°, with differences in the magnitude and trend of the lift coefficients below this angle.

3.1.2 Surrogate Model Construction

To assess the performance of the presented surrogate modelling methodology, MFK surrogate models were constructed for different combinations of the symmetric spinnaker sails and compared to the pre-computed RANS data points. Single-Fidelity Kriging (SFK) models were also constructed, trained only on the data points of the high-fidelity sail. This represents an equivalent model when the low-fidelity data is not used. The comparison of these two models allows the effect of using data points from an existing sail in a MFK surrogate model to be assessed.

Both the SFK and MFK models are trained using the same initial data points from the high-fidelity sail at the AWAs of 125°, 145°, and 165°. These three points were chosen to have an approximately uniform coverage of the design space. The MFK model was additionally trained with low-fidelity points from one of the other sails, as specified in each of the results. These lower-fidelity data points were positioned at 10° increments of AWA between 110° and 180°.

For the results presented in Section 4.2, both the MFK and SFK surrogate models use an adaptive sampling method to supplement the initial training points with additional data points, known as infill



(a) Drag coefficient model. **(b) Lift coefficient model.**

Figure 5: MFK prediction of symmetric spinnaker Sail I force coefficients with training points from Sails I and II, compared to SFK prediction with training points from Sail I, and the CFD data set of Sail I.

points, to improve the accuracy of the model (Forrester *et al.*, 2008). In this work, the focus of the surrogate models is exploration, with the aim to achieve an accurate model over the full range of AWAs. The adaptive sampling strategy used in this work therefore differs from that used in the more common surrogate model application of optimisation. Equation 12 is used to find the AWA with the maximum MSE. The pre-computed CFD data point which is the closest in value to this AWA is then added as a training point for the model, and the model is retrained. This adaptive sampling process is then repeated.

3.2. Asymmetric Spinnaker Sails

The second data set used in this work consists of CFD data for asymmetric spinnaker and mainsail combinations of two different yachts, provided by the Sailing Yacht Research Foundation (SYRF) (Braun, 2016). The data points were calculated using FSI with the fluid modelled by OpenFOAM (Bouhlel *et al*., 2019), and structural modelling with MemBrain™. The FSI calculations were performed at different wind angles, wind speeds, yacht heel angles, and sail traveller angles. The trim of the sails was adjusted between each iteration of the FSI process to achieve a neutral pressure on the luff. The meshes were constructed using SnappyHexMesh with approximately 10×10^6 elements. Full details of the mesh and domain are not provided in the SYRF report and validation of the CFD results with experimental results was not undertaken. The data used for this model along with further details on the calculation methods can be found in the SYRF project report (Braun, 2016).

The sails used from this data set for this work were A3 asymmetric spinnaker sails of a modified Transpac 52 (TP52) yacht and a Swan 42 yacht, with force coefficients at a true wind speed of 12 knots. Both the A3 spinnaker and mainsail were present in the simulation domain, with the drag coefficients of only the A3 spinnaker considered at this initial stage to simplify the analysis. This is a significant simplification and a limitation of this work; the application of the MFK method to the overall sail plan will need to be analysed in future work.

The results for these sails were used to create a MFK model to predict the drag coefficient curve of the TP52's sail, with the low-fidelity training points from the Swan 42's sail. A SFK model was also constructed, trained solely on the TP52 sail's drag coefficients. All of the data points had the traveller in the centre position, with data at four AWAs (45°, 65°, 85°, 105°). As a minimum of three high-fidelity training points are required for the surrogate model construction, there is only one remaining AWA at which the prediction error of both the SFK and MFK models can be assessed against the known CFD data point. Consequently, adaptive sampling could not be undertaken. For this work,

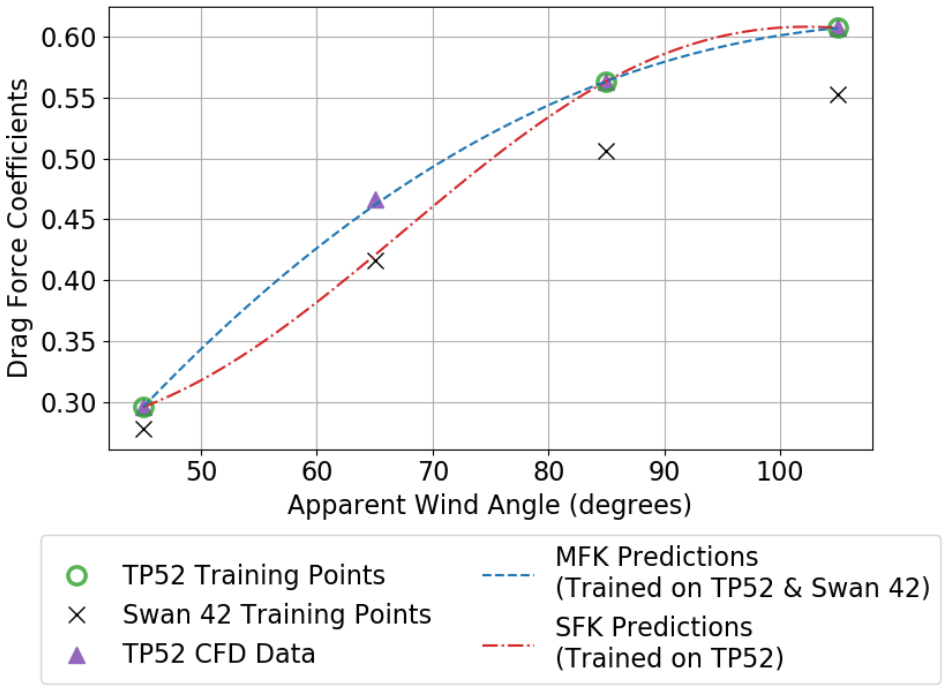


Figure 6: The SFK and MFK model predictions for the TP52 asymmetric spinnaker drag coefficient curve. The TP52 points are considered high-fidelity, and Swan 42 low-fidelity.

the case of interpolating between known data points was chosen, and the AWA of 65° is to be predicted by the models, with the data points at the remaining AWAs (45°, 85°, 105°) used as high-fidelity training points. All four available AWAs for the Swan 42 were used as low-fidelity data points.

4. RESULTS

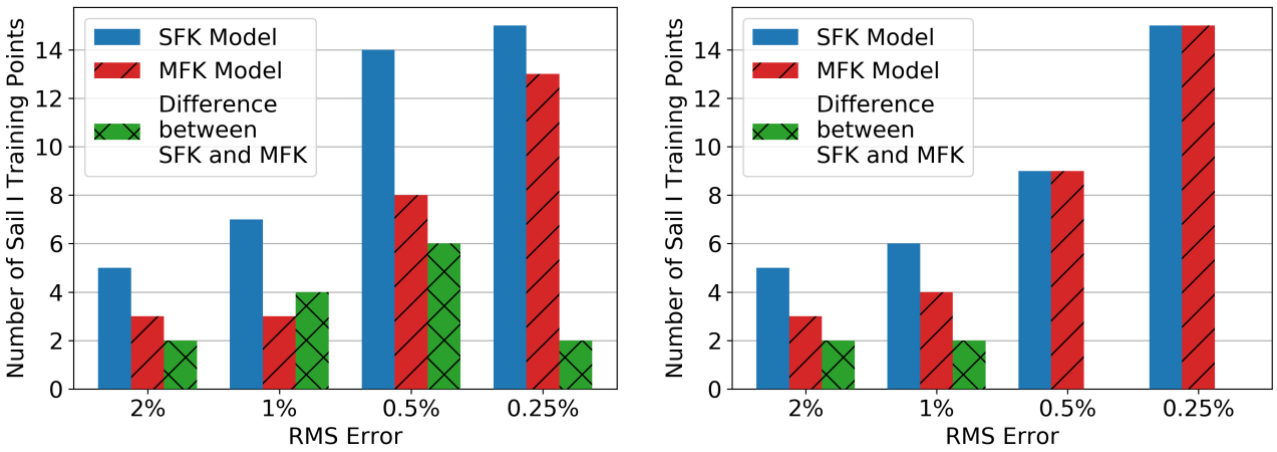
The performance of the MFK methodology was assessed for both the symmetric and asymmetric spinnaker data sets. The form of the MFK and SFK models for both data sets is analysed in Section 4.1. The reduction in computational cost which can be achieved for the symmetric spinnaker sails is discussed in Section 4.2. Both data sets are then used to assess how the error of the MFK models vary in Section 4.3, and how the errors of the MFK and SFK models compare in Section 4.4.

4.1. Multi- and Single-Fidelity Surrogate Models

To illustrate the form of the surrogate models used in this work, the MFK model prediction curve for both the lift and drag coefficients of Sail I from the symmetric spinnaker sail series is presented in Figure 5. Data points from Sail I were used as high-fidelity points, with Sail II providing low-fidelity points. No adaptive sampling was used.

The MFK predictions of the drag coefficients of Sail I, in Figure 5a, are close to the known CFD data points over the majority of the coefficient curve, with differences near the AWA of 120°. The lift predictions of the MFK model, in Figure 5b, are accurate for AWAs above 140° where the lift coefficient curve is approximately linear. There is a difference in the predicted and CFD values between 125° and 140° where the trend of the lift coefficient curve changes, and the peak lift coefficient values at low AWAs are underpredicted. The MFK model for the drag coefficients had the lowest error of the two models, with an RMS error of 1.1%, compared to 8.7% for the lift coefficients, taken as a proportion of the mean drag or lift coefficient value for Sail I. This could be due to the sharp increase in lift coefficient value for AWAs between 125° and 140°.

Figure 5 also shows the predictions of an equivalent SFK model, trained solely on the three Sail I high-fidelity points. This model has similar predictions for the drag coefficient curve, in Figure 5a, with an overprediction of the coefficient value for AWAs less than 125°. Overall, the drag coefficient SFK model had an RMS error of 3.4%. The SFK model for the lift coefficients in Figure 5b shows an



(a) Drag coefficients. **(b) Lift coefficients.**

Figure 7: MFK prediction of Sail I force coefficients with training points from Sails I and II, compared to SFK prediction with training points from Sail I, and the CFD data set of Sail I. The RMS errors are measured as a proportion of the mean drag and lift coefficient values for Sail I.

increased overall error compared to the drag coefficients, with an RMS error of 18.6%. The lift coefficients below 165° are overpredicted by the SFK model, as well as between 125° and 145°. While the SFK model is close to the CFD points for the peak lift coefficient values at low AWAs, the trend of the predictions in this region do not show the presence of a peak lift coefficient value. It therefore may be the case that the MFK model, while having a larger error in this region, is better at capturing the important characteristics of the lift coefficient curve.

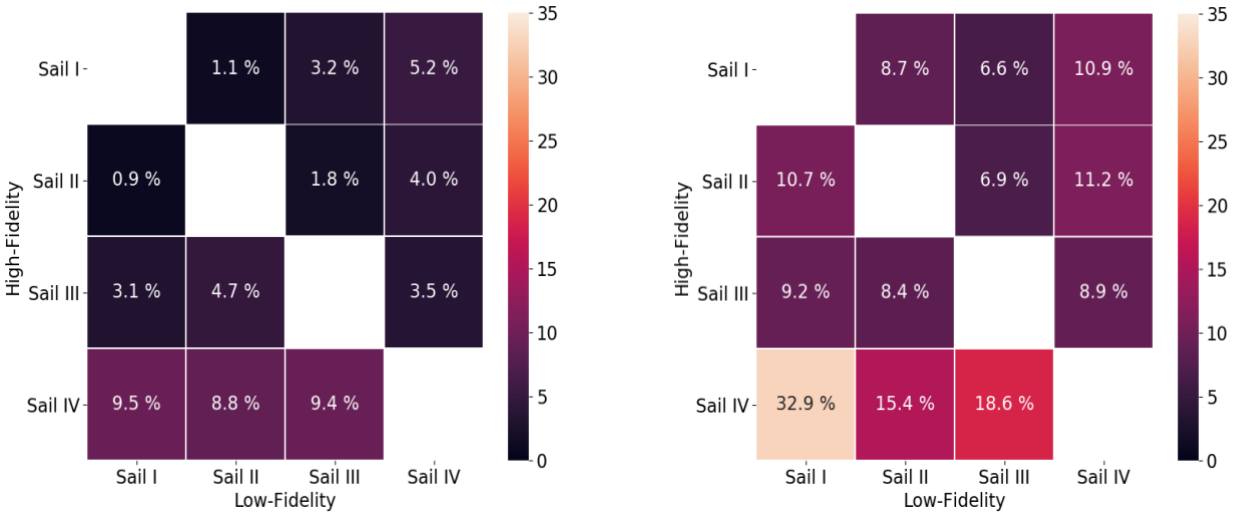
For this symmetric spinnaker sail combination, the MFK models have a lower error than the SFK models for the same computational cost, and therefore represent an improvement in the application of surrogate models. Both models display a smooth fit through the training points. However, this analysis has only considered one of the possible sail combinations for the MFK model. The application of both models to the other symmetric spinnaker sail combinations will be discussed in Section 4.2.

The MFK and SFK models for the TP52 asymmetric spinnaker drag coefficients are presented in Figure 6. The drag coefficients for the asymmetric spinnaker are considerably smoother than for the symmetric spinnaker example in Figure 5a. The MFK model therefore achieves a more accurate representation of the drag coefficient curve for the TP52 sail, with an RMS error of 0.2% as a proportion of the average drag coefficient value for this sail. The SFK model does not capture the trend of the curve between 45° and 85° as well, leading to an RMS error of 2.1%.

For the same number of initial high-fidelity data points, in this case the CFD simulations of the TP52 sail, a lower error was achieved using a MFK model. This demonstrates the use of low-fidelity training points to infer the trend of the drag coefficient curve for a different sail design. The application of the MFK method to data points calculated using FSI for non-rigid downwind sails shows promise for the method's future application to more real-world scenarios, despite the dataset being limited in size and the results not able to be experimentally validated.

4.2. Reduction in Computational Cost

One of the main reasons for using a MFK model is to reduce the computational cost required to construct a surrogate model surface with a certain error when compared to an SFK model. To assess the extent to which the methodology achieves this, the number of high-fidelity training points required to attain a force coefficient curve with a certain RMS error value was compared between the two models for the symmetric spinnaker sail data set. The force coefficient curves of Sail I were predicted by both models, with the MFK model using low-fidelity points from Sail II. Adaptive sampling was used to add high-fidelity infill points in addition to the three starting points. As the MFK models are



(a) Drag coefficients. **(b) Lift coefficients.**

Figure 8: RMS errors for symmetric spinnaker surrogate models trained with different sail combinations of high- and low-fidelity sails, as a proportion of the mean drag and lift coefficient values for the high-fidelity sail.

predicting the force coefficient curves for Sail I, Sail I represents a new sail design, and therefore the infill points are new CFD simulations which would need to be performed. A reduction in the number of Sail I training points required would consequently reduce the computational cost of constructing the force coefficient curves.

Figure 7 presents the number of Sail I training points required using the MFK and SFK models for the drag coefficients in Figure 7a and lift coefficients in Figure 7b. The reduction in computational cost was most pronounced for the drag coefficient models in Figure 7a, where the MFK model consistently reduced the number of Sail I training points required to attain a curve with a certain RMS error value. The largest relative reduction is for an RMS error of 1%, where there is a decrease of 57% in the number of Sail I training points required. This computational cost decrease is of a similar magnitude to that presented in the literature, with (Pellegrini *et al.*, 2016) finding a decrease of 50% for a one-dimensional test function when using a multi-fidelity model compared to single-fidelity, and a decrease of 23% for the modelling of the resistance of a ship hull.

The difference in training points required varies for each RMS error level, as demonstrated by the green bars in Figure 7a. This increases up to an RMS error of 0.5% where 6 fewer training points are needed. However, for lower RMS errors the difference reduces to 2 simulations. To achieve these low errors, new CFD simulations are required rather than inferring the trends from the data of other sail designs. However, a model with an error this low may not be required for all sail design applications. By creating a model with an RMS error of 1% rather than 0.5%, the number of new CFD simulations required for a MFK model is reduced by 63%. Ensuring that the surrogate model is only refined to the necessary error level can therefore be used to reduce the computational cost of predicting the force coefficient curves.

A feature seen in Figure 7a is that the number of Sail I training points for RMS errors of 2% and 1% are the same. This is because the RMS error of the MFK model is 0.009 when trained using the initial three high-fidelity points. Therefore, even if an RMS error of 2% was desired, a lower error than needed was achieved by the MFK model. However, this feature will likely not be the case for all implementations of this method as there will be variations in the model error for different sail combinations.

The reduction in computational cost for the lift coefficients in Figure 7b is lower, and the overall error of the models is higher, than for the drag coefficients. There is a reduction in computational cost for RMS errors of 2% and 1%, with two fewer high-fidelity points needed. For a lower RMS error, there is then no difference in the high-fidelity points needed for the MFK and SFK models. The benefit of

Table 1: Comparison of Prediction Errors of SFK and MFK Models for Different Prediction AWAs as a Proportion of the Mean Drag Coefficient Value

Prediction AWA	SFK Error	MFK Error
45°	11.8%	1.2%
65°	2.1%	0.2%
85°	3.9%	0.2%
105°	3.1%	0.2%

using a MFK model is therefore more limited for the lift coefficients, highlighting a potential limitation or area of improvement for the method. This difference may be due to the trend of the lift coefficient curve changing at low AWAs, an issue which may be mitigated by a different placement of high-fidelity points to better capture this region.

The MFK method therefore has the potential to reduce the computational cost required to construct a model for a certain error, with a relative reduction of up to 57% achieved. However, this reduction is dependent on the error level desired, with lower errors having a lower cost reduction advantage. The cost reduction for the MFK method is also reduced for the lift coefficients.

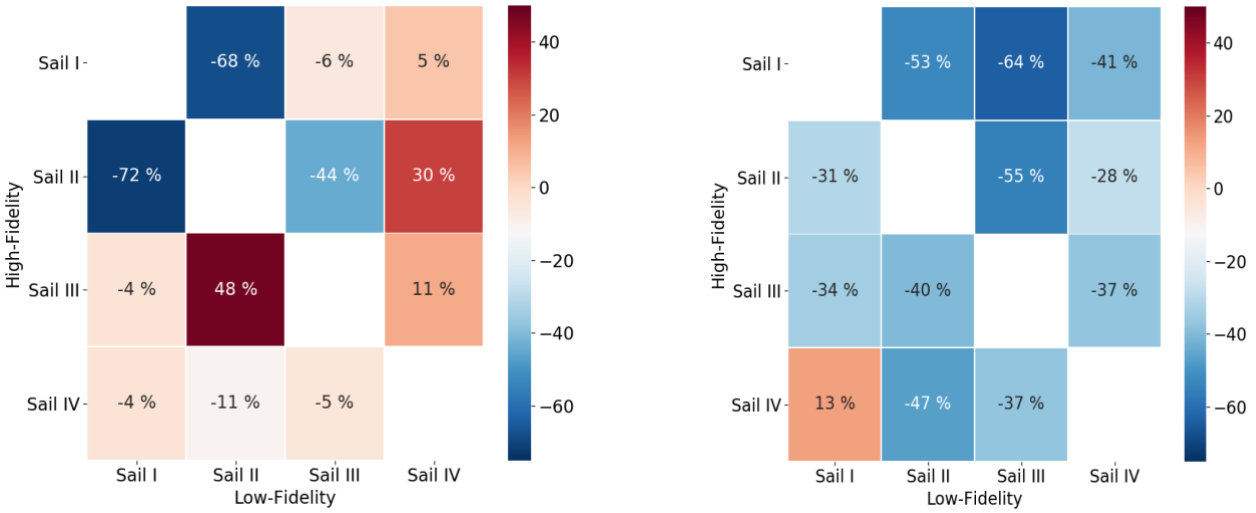
4.3. Variation in Model Error

The variation in the error of the model can be assessed for the symmetric spinnaker data set by comparing the error of different sail combinations, and for the asymmetric spinnaker data set by assessing how the error varies when different combinations of high-fidelity points are used, and the coefficient prediction is therefore made at different AWAs.

The first aspect to consider is how the error of the MFK models vary depending on the choice of existing sail for the low-fidelity data. This was assessed using the symmetric spinnaker sail data set. The RMS error of the MFK model trained on each of the possible combinations of two sails (one sail high-fidelity on the y-axis, and the other sail low-fidelity on the x-axis) is presented for the drag coefficients in Figure 8a and the lift coefficients in Figure 8b. The RMS error is measured as a proportion of the mean drag or lift coefficient value for the respective high-fidelity sail. While there is variation within each of the force coefficient types, the drag coefficient models have a consistently lower RMS error compared to the lift coefficient models for all the sail combinations.

For each high-fidelity sail, the error of the MFK model varies depending on the sail used for the low-fidelity data. For the drag coefficient models of Sail I in Figure 8a, the RMS error increases from 1.1% when Sail II is used, to 5.2% with Sail IV, a nearly 5-fold increase in estimated error. As expected, the choice of sail for the low-fidelity data therefore has a significant effect on the resulting model error and is an important consideration when constructing a model. The choice of existing sail is analysed in Section 4.4.

An asymmetry can be seen in the sail combination matrices for both the drag and lift coefficient models in Figures 8a and 8b. Combinations which are in the upper right of the matrix tend to have a lower RMS error than their equivalent on the lower left. This means that for the same combination of sails, for example Sails II and III, the order in which the sails are high- or low-fidelity has an impact on the resulting model error. This asymmetry is expected, as if the underlying force coefficient curve of the high-fidelity sail has a complex trend which is not present in the low-fidelity sail data, it is difficult for this trend to be captured with only three high-fidelity data points. Conversely, when a simpler force coefficient curve is to be predicted, the more complex low-fidelity curve trend is more easily compensated for. This asymmetry of the model error for different sail combinations may also be caused by the placement of high-fidelity points. The same high-fidelity point locations are used



(a) Drag coefficients. **(b) Lift coefficients.**

Figure 9: Percentage difference in RMS error of MFK models compared to SFK models, negative values when the MFK model has a lower RMS error compared to SFK, positive values when the MFK model has a higher RMS error.

for all the models, but this will not be the optimal placement of high-fidelity points for all the combinations.

For the asymmetric spinnaker sails, while there is only one point at which the predictions can be made, the sensitivity of the error of the MFK model can be assessed by training the model on different combinations of three high-fidelity points. The RMS errors for MFK and SFK models trained to predict the drag coefficient curve of the TP52 sail with different high-fidelity point combinations, and therefore prediction points, are shown in Table 1.

The MFK model errors in Table 1 for the prediction locations are very similar, with three of the RMS errors being 0.2% and a maximum error of 1.2% at an AWA of 45°.The RMS error of the MFK model is therefore not sensitive to which points are used to train the model, suggesting that it is robust. The maximum SFK model error is also at an AWA of 45°, suggesting that this is a more difficult point to predict due to its position at the end of the curve. At this point the difference between the low- and high-fidelity curves is also closer than would be suggested by the difference at the other AWAs, as seen in Figure 6. The MFK model performs better than SFK for all AWAs, with the error an order of magnitude lower.

4.4. Reduction in Model Prediction Error

Another potential benefit of MFK models is a reduction in error of the model for the same computational cost when compared to the SFK method (Forrester *et al.*, 2008). When a number of high-fidelity data points have been obtained, for example the minimum of three required to create a model, a choice has to be made between creating a SFK model using these high-fidelity points, or whether to supplement them with existing data points and instead create a MFK model (Toal, 2015). The decision of when to use a SFK or MFK model is therefore important.

An important aspect to investigate is how the errors of the MFK models compared to the equivalent SFK models. To assess this, the MFK models constructed for each symmetric spinnaker sail combination were compared to the equivalent SFK model. For each sail combination and for both drag and lift coefficients, the percentage difference between the RMS error of these two models was calculated. These are presented in Figure 9. Negative percentage values represent the MFK model having a reduced RMS error compared to the SFK model, with the MFK model therefore having a better performance, and positive values are where the SFK model has a lower error instead.

Table 2: r^2 Correlations of the Drag Coefficient for each Sail Combination

	Sail I	Sail II	Sail III	Sail IV
Sail I	-	0.998	0.994	0.982
Sail II	0.998	-	0.997	0.984
Sail III	0.994	0.997	-	0.986
Sail IV	0.982	0.984	0.986	-

Table 3: r^2 Correlations of the Lift Coefficient for each Sail Combination

	Sail I	Sail II	Sail III	Sail IV
Sail I	-	0.980	0.955	0.889
Sail II	0.980	-	0.987	0.937
Sail III	0.955	0.987	-	0.971
Sail IV	0.889	0.937	0.971	-

For the drag coefficients in Figure 9a, there is variability in the relative error of the two models, with the MFK model not always having a reduction in error when compared to the equivalent SFK model. This variation is present when assessing models for a particular high-fidelity sail. The prediction for Sail II has the largest relative difference in performance, with a 72% decrease in model error achieved when a MFK model is used with the low-fidelity data from Sail I, but a 30% increase when Sail IV is used instead. These two cases represent combinations where the choice between MFK and SFK models is evident; a MFK model with Sail I, and an SFK model with Sail IV. This highlights the need for a method to determine which existing sail design is the most suited for use in a MFK model as well as when the MFK method is suitable.

For the lift coefficients in Figure 9b, all but one of the MFK models has a reduced RMS error compared to the equivalent SFK model, with a maximum decrease of 64%. This demonstrates that while the MFK models had difficulty modelling the lift coefficients, with higher RMS errors in Figure 8 than for the drag coefficients, the SFK models were also not able to capture the trends well. The only combination where the MFK performance was worse than SFK was the prediction of Sail IV using Sail I. This combination has the highest difference in sail geometries as measured by the foot camber in Figure 1, and therefore different force coefficient curve trends, and is predicting the more complex curve of Sail IV using the simpler curve of Sail I. This is therefore one of the most difficult sail combinations to model, and one in which the performance of the MFK method was not expected to be optimal.

The magnitude of error reduction found for the drag and lift coefficients in Figure 9 is within the same order of magnitude as presented in literature. For example, Peña *et al.* (2012) found a decrease in the mean error for the prediction of drive and side forces of asymmetric spinnakers from 11.5% to 1.4%, using an ANN to correct low fidelity CFD data. This represents a decrease of 88%, similar to the maximum decrease of 72% found for the drag coefficient model of Sail II and I.

It is difficult to predict which low-fidelity data will improve the model accuracy *a-priori*. This is required in real-world applications as the accuracy of the model predictions are not readily available. Toal (2015) found that the correlation between the low- and high-fidelity data points relates to the accuracy of a MFK model and recommended that a criterion of $r^2 > 0.9$ be used when deciding to use a MFK model instead of SFK. The importance of the correlation between the input curves has also been discussed by Fernández-Godino *et al.* (2019), who suggest that by having a higher correlation the variance of the discrepancy function $\delta(\boldsymbol{x})$ in Equation 9 is reduced, resulting in a more accurate correction between the low- and high-fidelity points and therefore a more accurate MFK model. The r^2 correlations of the symmetric spinnaker sail combinations were therefore assessed to see if there was a relationship with the relative performance of the two modelling methods.

The r^2 correlations for each combination of the symmetric spinnaker sail input curves are presented for the drag coefficients in Table 2, and lift coefficients in Table 3. The r^2 correlation is a measure of the similarity of two curves, calculated as how much variance in one of the curves is explained by the other curve (James *et al.*, 2013). The highest correlation for the drag coefficients is between Sails I and II, and for the lift coefficients between Sails II and III. The lift coefficient r^2 correlations are also noticeably lower than for the drag coefficients. This matches the trend of the lift coefficient

MFK models having a larger error, suggesting that one reason for this is the lower similarity between the low- and high-fidelity coefficient curves. All the r^2 correlations for the drag coefficients of each sail combination are above the criterion of $r^2 = 0.9$ recommended by Toal (2015), and so are all but two of the lift coefficient combinations. This suggests that almost all the sail combinations are suitable for constructing a MFK model. However, as previously discussed, this was not found to be the case.

One possible reason for the failure of the criterion from Toal (2015) for the symmetric spinnaker data set is that the r^2 correlation is not able to account for the previously discussed asymmetry in the error of the MFK models, and therefore the relative error of the MFK and SFK models. This is illustrated by the combinations of Sails II and III. By definition, the two drag coefficient MFK models constructed for this sail combination, one to predict Sail II and the other Sail III, will have the same r^2 correlation value as the same two sails are being used. However, the two models have different RMS errors in Figure 8 and therefore a different relative performance compared to the equivalent SFK models in Figure 9. The model predicting the drag coefficient curve of Sail II has a reduction in error of 44% compared to a SFK model, while the model predicting Sail III has an error increase of 48%. A criterion for determining the suitability of a MFK model will therefore need to account for the relative fidelity of the data of the two sails being used. These asymmetric effects with an identical r^2 are also seen when analysing the lift coefficient models, in particular Sails I and IV.

Another possible reason for the failure of this criterion is that having a high r^2 correlation does not uniquely achieve a smooth difference function between the low- and high-fidelity data. Having a smooth difference function is important, as the simplicity of the difference function is related to the error of the model (Forrester *et al.*, 2007). This may explain the lower error for the MFK model of the asymmetric spinnaker data set in Figure 6. While the r^2 correlation was high ($r^2 = 0.997$), the difference function is also smooth compared to the symmetric spinnaker models. The difference between the low- and high-fidelity points changes with the AWA, however the increase is smooth. There are no significant local variations in the difference function, as are present for the symmetric spinnaker sails. For example, for the symmetric spinnaker sail in Figure 5b there is a local variation in the difference function where the low-fidelity points alternate between having a higher and low lift coefficient value than the high-fidelity points between AWAs of 120° and 140°. If the difference function is not smooth, this can cause the MFK model to predict erroneous trends and have a reduced accuracy compared to SFK (Raven and Scholcz, 2019).

Therefore, the r^2 correlation is not appropriate for determining when a MFK model is suitable when applied to the case of spinnaker sails, and other factors including those which alter the smoothness of the difference function need to be considered. The effect of the placement of high-fidelity points should also be considered, as this can help to model local variations in a difference function.

5. DISCUSSION

This work is an initial investigation into the use of MFK models for sail performance prediction. The MFK method has the potential to reduce the computational cost required to achieve a model with a certain accuracy, or to have an increased accuracy for the same computational cost. This reduction in cost or improvement in accuracy is dependent on the sail combination used and the force coefficient being analysed. While most combinations had an improved performance over SFK models, some resulted in a worse performance. It is therefore important to investigate additional criteria which would indicate *a-priori* whether the MFK method is better suited than SFK for each case.

The basis of the MFK method is that the low-fidelity data points are for existing sail designs and have previously been calculated as part of the performance assessment of those designs. Therefore, there would be no additional computational cost involved in the use of these data points. New CFD simulations would only be computed for the high-fidelity points of the new sail under consideration. The advantages of the presented MFK method are therefore reliant on having an existing database

of data points for similar sails. In cases where this information is not present, the advantages of the MFK methodology are limited.

One limitation of the modelling is the consideration of the sail flying shapes and trimming. The data points used for the symmetric spinnaker surrogate models were simulated as rigid, with no investigation into the trim or shape deformations of the sails. This differs from reality as soft-cloth spinnakers have a complex non-linear coupling between the pressure distribution and sail shape which will influence the trend of the force coefficient curves (Ramolini, 2019). These sails were also tested without the presence of a mainsail, resulting in a less realistic flow field. A more realistic data set is that of the SYRF asymmetric spinnakers, where the data points were simulated using an FSI procedure with the mainsail present, and the trim adjusted between iterations. While this data set contained more complex effects, the presented MFK methodology showed promising performance. Nevertheless, the evaluation of the presented MFK method for a wider range of sail shapes and more realistic trim and flying shape characteristics is a necessary next stage to assess the method's applicability to real-world sail design cases.

Another limitation of this work is the sole use of pre-computed CFD data for the adaptive sampling of the MFK model. Using pre-computed data restricts the AWAs at which the infill points can be placed, and consequently they may not be placed at the point of maximum variance. The ability to obtain a new CFD data point at the AWA where the maximum variance occurs would likely improve the accuracy of both the MFK and SFK models where adaptive sampling is used.

A limitation of assessing the performance of the surrogate modelling methods is the data to which the model predictions can be compared. The accuracy of the model is determined by comparing the model predictions to the available CFD points for the high-fidelity sail, and these data points are also being used as the high-fidelity training points for the MFK model. These points have errors due to approximations in the model construction, and the solving of the RANS equations. At this stage, the aim is to assess the initial applicability of the MFK method. This is measured by its ability to construct a curve with a similar accuracy to that solely from RANS CFD calculations, but with a reduced computational cost, as discussed in Section 4.2. The accuracy of the MFK model predictions will still need to be compared against full-scale data.

A less significant factor is the use of only two fidelity levels for the construction of the MFK model. The architecture of the MFK model allows data from several different levels of fidelity to be used, but further testing is required to successfully implement this into the MFK modelling methodology. Having data from more than two fidelities has the potential to improve the accuracy of the model but, as discussed, the similarity of the curves used is important and there is the potential for the error of the model to increase instead. The MFK methodology used requires the fidelities of the data to be specified in a hierarchical manner. This means that when using two (or more) low-fidelity sails, their relative fidelity level will need to be determined, which may not be known if they have a similar design. An alternative method would be to apply a multi-information source surrogate model, where the fidelity level of the data points is non-hierarchical and can change over the solution space (Ghoreishi and Allaire, 2019). However, the challenge is then how to accurately specify the relative fidelity between the different data sets. This is an area which could be investigated in future work.

This work has also considered the lift and drag coefficients to be a function of only the AWA. For practical use in a VPP, the force coefficients are likely to be modelled as a function of both the AWA and apparent wind speed, with other aspects such as the heel angle also possible. This would require the MFK method to be extended to higher dimensions. This capability is present in the SMT framework and is an area of investigation for future work. Increasing the number of dimensions of a surrogate model usually results in an increase in the number of training points required (Forrester *et al.*, 2007). An increase in the number of dimensions may therefore increase the benefits of using MFK models to reduce the computational cost of the surrogate model construction.

The effect of noise in the training data points on the accuracy of the MFK model predictions has not been considered. This is important to assess when data points with significant noise, such as those

from experimental testing, or with low mesh resolution, are used in the model. However, this is not the case for both the low- and high-fidelity data points used for the MFK models. The effect of noise is therefore likely to be reduced compared to approaches where the differences in fidelity are achieved by a difference in mesh resolution, resulting in data points from low mesh resolution simulations which likely have numerical noise present (Wackers *et al.*, 2020).

6. CONCLUSIONS

- Data from existing sail designs have been used for the first time to assist in the prediction of the force coefficient curve of a different symmetric or asymmetric spinnaker sail. While single-fidelity surrogate models have previously been used to interpolate the force coefficients of a sail for input into a VPP, more sophisticated surrogate modelling methods such as multi-fidelity kriging have not been used, and the potential of using data from existing sails has not previously been explored.
- For a target accuracy, the MFK model for one of the symmetric spinnaker sail combinations reduced the number of high-fidelity training points required, and therefore the computational cost, by a maximum of 6 training points for the drag and 2 training points for the lift coefficients.
- The accuracy of a MFK model is dependent on the sail used for the low-fidelity training points, with an increase in RMS error from 1.1% to 5.2% for the symmetric spinnaker Sail II. The estimated error from the MFK model was not found to significantly vary with the prediction point location for the asymmetric spinnaker data, with a maximum increase in RMS error from 0.2% to 1.2%.
- The use of a MFK model for the symmetric spinnaker sails reduced the error of the force coefficient predictions by a maximum of 72% for the drag coefficients and 64% for lift coefficients compared to a SFK model with the same number of training points. For the asymmetric spinnaker sail set, the RMS error was reduced by 90%.
- MFK models did not always have a lower error than SFK models. The maximum increase in error found was 48% for the drag and 13% for the lift coefficients of the symmetric spinnaker sails. The r^2 correlation between the low- and high-fidelity force coefficient curves was not found to be sufficient to determine when a MFK model is suitable.

7. ACKNOWLEDGEMENTS

This work is supported by Doyle Sails. The authors would like to thank Callaghan Innovation for their funding of this work through the R&D Fellowship Grant number DSYX1801.

8. REFERENCES

Alizadeh, R., Allen, J. and Mistree, F. (2020), 'Managing computational complexity using surrogate models: a critical review', *Research in Engineering Design* **31**, 275–298.

Alza, P. (2012), Numerical and experimental studies of sail aerodynamics, PhD thesis, Polytechnic University of Madrid, Spain.

Böhm, C. and Graf, K. (2010), Coupling of RANSE-CFD with VPP methods: From the numerical tank to virtual boat testing, *in* '2nd International Conference on Innovation in High Performance Sailing Yachts', Lorient, France.

Bouhlel, M. A., Hwang, J. T., Bartoli, N., Lafage, R., Morlier, J. and Martins, J. R. R. A. (2019), 'A python surrogate modeling framework with derivatives', *Advances in Engineering Software* .

Braun, J. (2016), 'SYRF downwind aero moments and forces'. Retrieved from http://sailyachtresearch.org/projects/downwind-aero-1. Accessed February 2020.

Cirello, A., Cucinotta, F., Ingrassia, T., Nigrelli, V. and Sfravara, F. (2019), 'Fluid-structure interaction of downwind sails: A new computational method', *Journal of Marine Science and Technology* **24**(1), 86.

Courtel, P. (2020), Preliminary sail shape optimisation using neural network aerodynamic and hydrodynamic models, Master's thesis, University of Liège, Belgium.

de Baar, J. and Roberts, S. (2015), Multi-fidelity surrogate-based parameter estimation for a sailing yacht hull, *in* '21st International Congress on Modelling and Simulation', Gold Coast, Australia.

de Baar, J., Roberts, S., Dwight, R. and Mallol, B. (2015), 'Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging', *Computers & Fluids* **123**, 185–201.

Doyle, T., Knight, B. and Swain, D. (2016), A comparison of a RANS based VPP to on the water sailing performance, *in* '22nd Chesapeake Sailing Yacht Symposium', Annapolis, USA.

Fernández-Godino, M., Park, C., Kim, N. and Haftka, R. (2019), 'Issues in deciding whether to use multifidelity surrogates', *AIAA Journal* **57**(5).

Forrester, A., Sobester, A. and Keane, A. (2007), 'Multi-fidelity optimization via surrogate modelling', *Proceedings of the Royal Society A* **463**, 3251–3269.

Forrester, A., Sobester, A. and Keane, A. (2008), *Engineering Design via Surrogate Modelling: A Practical Guide*, Wiley.

Fossati, F. (2009), *Aero-Hydrodynamics and the Performance of Sailing Yachts: The Science Behind Sailing Yachts and Their Design*, A&C Black.

Fossati, F., Muggiasca, S. and Viola, I. (2006), An investigation of aerodynamic force modelling for IMS rule using wind tunnel techniques, *in* '19th International HISWA Symposium on Yacht Design and Yacht Construction', Amsterdam, Netherlands.

Ghoreishi, S. and Allaire, D. (2019), 'Multi-information source constrained Bayesian optimization', *Structural and Multidisciplinary Optimization* **59**, 977–991.

Harries, S. and Abt, C. (2019), CAESES - the HOLISHIP platform for process integration and design optimization, *in* 'A Holistic Approach to Ship Design: volume 1: Optimisation of Ship Design and Operation for Life Cycle', Springer International Publishing, pp. 247–293.

James, G., Witten, D., Hastie, D. and Tibshirani, R. (2013), *An Introduction to Statistical Learning*, Springer.

Jin, R., Chen, W. and Simpson, T. (2001), 'Comparative studies of metamodelling techniques under multiple modelling criteria', *Structural and Multidisciplinary Optimisation* **23**, 1–13.

Kennedy, M. and O'Hagan, A. (2000), 'Predicting the output from a complex computer code when fast approximations are available', *Biometrika* **87**(1), 1–13.

Korpus, R. (2007), Performance prediction without empiricism: A RANS-based VPP and design optimization capability, *in* '18th Chesapeake Sailing Yacht Symposium', Annapolis, USA.

Krige, D. (1951), 'A statistical approach to some basic mine valuation problems on the Witwatersrand', *Journal of the Southern African Institute of Mining and Metallurgy* **52**(6), 119–139.

Kuya, Y., Takeda, K., Zhang, X. and J. Forrester, A. I. (2011), 'Multifidelity surrogate modeling of experimental and computational aerodynamic data sets', *AIAA Journal* **49**(2), 289–298.

Lasher, W. (2011), The determination of aerodynamic forces on sails - challenges and status, *in* 'Atmospheric Turbulence, Meteorological Modeling and Aerodynamics', Nova Science Publishers, Inc., pp. 487–504.

Lasher, W. and Richards, P. (2007), 'Validation of Reynolds-averaged Navier-Stokes simulations for international America's Cup class spinnaker force coefficients in an atmospheric boundary layer', *Journal of Ship Research* **51**, 22–38.

Lasher, W., Sonnenmeier, J., Forsman, D., Zhang, C. and White, K. (2003), Experimental force coefficients for a parametric series of spinnakers, *in* '16th Chesapeake Sailing Yacht Symposium', Annapolis, USA.

Le Gratiet, L. (2013), Multi-Fidelity Gaussian Process Regression for Computer Experiments, PhD thesis, Paris Diderot University, France.

Mason, A. (2010), Stochastic Optimisation of America's Cup Class Yachts, PhD thesis, Australian Maritime College, Australia.

Moore, W. (2004), Numerical modelling of a series of parametric spinnakers, Master's thesis, University of Auckland.

Nava, S., Cater, J. and Norris, S. (2018), 'Large eddy simulation of an asymmetric spinnaker', *Ocean Engineering* **196**, 99–109.

Norris, S. and Richards, P. (2012), CFD models of downwind sail aerodynamics using eddy viscosity and reynolds stress turbulence models, *in* '4th High Performance Yacht Design Conference', Auckland, New Zealand.

Peña, F. L., Casás, V. D., Gosset, A. and Duro, R. (2012), 'A surrogate method based on the enhancement of low fidelity computational fluid dynamics approximations by artificial neural networks', *Computers & Fluids* **58**, 112–119.

Peart, T. (2018), CFD modelling of downwind sails using RANS, BE (Hons) Part IV Project Report 2018-ME073, University of Auckland.

Pellegrini, R., Leotardi, C., Zaghi, S., Broglia, R., Campana, E., Iemma, U. and Diez, M. (2016), Multi-fidelity adaptive metamodel for ship hull performance via CFD, *in* '19th Numerical Towing Tank Symposium', France.

Pellegrini, R., Serani, A., Broglia, R., Diez, M. and Harries, S. (2018), Resistance and payload optimization of a sea vehicle by adaptive multi-fidelity metamodeling, *in* 'AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference', Florida, USA.

Ploé, P., Lanos, R., Visonneau, M. and Wackers, J. (2017), Bayesian strategies for simulation-based optimisation and response surface creation using a single tool, *in* '4th International Conference on Innovation in High Performance Sailing Yachts', Lorient, France.

Prince, M. and Claughton, A. (2013), Estimating a yacht's hull-sailplan balance and sailing performance using experimental results and VPP methods, *in* '3rd International Conference on Innovation in High Performance Sailing Yachts', Lorient, France.

Ramolini, A. (2019), 'Implementation of a fluid-structure interaction solver for a spinnaker sail', *Journal of Sailing Technology* **4**.

Raven, H. and Scholcz, T. (2019), An assessment of multifidelity procedures for ship hull form optimisation, *in* '8th International Conference on Computational Methods in Marine Engineering', Sweden.

Roux, Y., Durand, M., Leroyer, A., Queutey, P., Visonneau, M., Raymond, J., Finot, J., Hauville, F. and Purwanto, A. (2008), Strongly coupled VPP and CFD RANSE code for sailing yacht performance prediction, *in* '3rd High Performance Yacht Design Conference', Auckland, New Zealand.

Roux, Y., Huberson, S., Hauville, F., Boin, J., Guilbaud, M. and Ba, M. (2002), Yacht performance prediction: Towards a numerical VPP, *in* 'High Performance Yacht Design Conference', Auckland, New Zealand.

Ruiz de Elvira, M. (2015), Design of a generalized tool for the performance assessment under sail based on analytical, numerical and empirical Results, PhD thesis, The Polytechnic University of Madrid, Spain.

Sacher, M., Hauville, F., Duvigneau, R., Le Matre, O., Aubin, N. and Durand, M. (2017), 'Efficient optimization procedure in non-linear fluid-structure interaction problem: Application to mainsail trimming in upwind conditions', *Journal of Fluids and Structures* **69**, 209–231.

Sacks, J., Welch, W., Mitchell, T. and Wynn, H. (1989), 'Design and analysis of computer experiments', *Statistical Science* **4**(4), 409–423.

Serani, A., Pellegrini, R., Wackers, J., Jeanson, C.-E., Queutey, P., Visonneau, M. and Diez, M. (2019), 'Adaptive multi-fidelity sampling for CFD-based optimisation via radial basis function metamodels', *International Journal of Computational Fluid Dynamics* **33**.

Toal, D. J. (2015), 'Some considerations regarding the use of multi-fidelity kriging in the construction of surrogate models', *Structural and Multidisciplinary Optimization* **51**(6), 1223–1245.

Viola, I. (2009), 'Downwind sail aerodynamics: A CFD investigation with high grid resolution', *Ocean Engineering* **36**, 974–984.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P. and SciPy 1.0 Contributors. (2020), 'SciPy 1.0: Fundamental algorithms for scientific computing in Python', *Nature Methods* **17**, 261–272.

Wackers, J., Visonneau, M., Serani, A., Pellegrini, P., Broglia, R. and Diez, M. (2020), Multi-fidelity machine learning from adaptive- and multi-grid RANS simulations, *in* '33rd Symposium on Naval Hydrodynamics', Japan.

Weller, H. G., Tabor, G., Jasak, H. and Fureby, C. (1998), 'A tensorial approach to computational continuum mechanics using object-oriented techniques', *Computers in Physics* **12**(6), 620–631.

Zhao, D. and Xue, D. (2010), 'A comparative study of metamodeling methods considering sample quality merits', *Structural and Multidisciplinary Optimization* **42**, 923–938.