Efficient Optimization Procedure in Non-Linear Fluid-Structure Interaction Problem: Application to Mainsail Trimming in Upwind Conditions

Matthieu Sacher[a,*], Frédéric Hauville[a], Régis Duvigneau[b], Olivier Le Maître[c],
Nicolas Aubin[a], Mathieu Durand[d]

[a]*Naval Academy Research Institute - IRENAV CC600, 29240 BREST Cedex 9, France*
[b]*INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902 Valbonne, France*
[c]*LIMSI - CNRS, Rue John Von Neumann, 91400 Orsay, France*
[d]*K-EPSILON, 1300 Route des Crêtes, 06560 Valbonne, France*

Abstract

This paper investigates the use of Gaussian processes to solve sail trimming optimization problems. The Gaussian process, used to model the dependence of the performance with the trimming parameters, is constructed from a limited number of performance estimations at carefully selected trimming points, potentially enabling the optimization of complex sail systems with multiple trimming parameters. The proposed approach is tested on a two-parameter trimming for a scaled IMOCA mainsail in upwind sailing conditions. We focus on the robustness of the proposed approach and study especially the sensitivity of the results to noise and model error in the point estimations of the performance. In particular, we contrast the optimization performed on a real physical model set in a wind tunnel with a fully non-linear numerical fluid-structure interaction model of the same experiments. For this problem with a limited number of trimming parameters, the numerical optimization was affordable and found to require a comparable amount of performance estimation as for the experimental case. The results reveal a satisfactory agreement for the numerical and experimental optimal trimming parameters, considering the inherent sources of errors and uncertainties in both numerical and experimental approaches. Sensitivity analyses have been eventually performed in the numerical optimization problem to determine the dominant source of uncertainties and characterize the robustness of the optima.

Keywords: Sail trimming optimization, fluid-structure interaction, Gaussian process model, experimental, numerical, uncertainty quantification

1. Introduction

Research on sailing yachts has fostered the advancement of methods to predict and improve the performances of racing yachts. Yacht performance is usually assessed using so-called Velocity Prediction Programs (VPPs) [1], which by equilibrating loads on hull, appendages and sails, determine several performance indicators, such as Boat Speed (BS) and Velocity Made Good (VMG). The loads estimation in VPPs can be based on empirical formulas, experimental data and/or numerical models of various complexity level [2, 3]. Due to the complexity and multi-physics character of yacht dynamics, performance studies are often separated in hydrodynamic [4] and aerodynamic [5, 6, 7] aspects. In the present work, we focus on the aerodynamics optimization for the performance of a sail system; however, the procedure developed below can be applied to hydrodynamic optimization and even fully coupled (hydro-aero) yacht performance optimization.

The physics of sail systems involves very complex phenomena, such as nonlinear Fluid-Structure Interaction (FSI) effects and aero-elastic instabilities. Moreover, the modeling of real sailing conditions is still an

---

open challenge because of the large uncertainties in the prediction of wind and sea states. To our knowledge, the sails optimization has thus been limited so far to idealized situations. For instance, sail shape optimizations (without accounting for the full FSI problem) were reported in [8], while the numerical trimming of two-dimensional sails was considered in [9]. Regarding three-dimensional FSI problems, the authors of [10] mention an optimization of the trimming of sails, but within an inviscid flow approximation and few details are provided on the actual optimization procedure used.

The present work aims toward the development of efficient numerical optimization procedures, capable of dealing with complex sail systems with realistic physical models (*e.g.* nonlinear FSI and turbulent flows) and large number of optimization variables (*i.e.* trimming parameters). Abstractly, the optimization problem can be expressed as

$$\mathbf{x}_{\text{opt}} = \underset{\mathbf{x} \in \Omega}{\arg \min} -\mathcal{P}(\mathbf{x}), \tag{1}$$

where $\mathbf{x}$ are the optimization variables, Ω the domain of variation of the optimization variables, and $\mathcal{P} : \Omega \mapsto \mathbb{R}$ is the measurement of performance to be maximized. The main difficulty in solving problem (1), in the context of sail-trimming, is related to the cost of numerically estimating the performance $\mathcal{P}$ at tentative values $\mathbf{x}$ of the parameters.

Indeed, the estimation of $\mathcal{P}(\mathbf{x})$ involves the resolution of a complex nonlinear FSI problem, with typically several convergence iterations between the nonlinear elastic and flow solvers. Further, adjoint-based techniques are hardly amenable to non-linear FSI problems, particularly when the resolution is based on the coupling of distinct solvers. This fact precludes the use of efficient gradient-based optimization methods, and favors the use of derivative free optimization algorithms such as the simplex based [11] and evolutionary [12, 13] methods. These approaches classically require many performance evaluations of $\mathcal{P}(\mathbf{x})$, making applications to sail systems very costly as a single evaluation of $\mathcal{P}$ can routinely require several hours of CPU-time even on modern parallel computers.

From these observations, we advocate the use of meta-models to mitigate the large computational cost of optimizing the trimming parameters of sail systems. Meta-models-based optimization methods have been experiencing a growing interest for the last years, and are currently used in several other disciplines, such as aerodynamic drag reduction [14, 15], vibration minimization for rotating aircrafts [16], civil engineering for the design of water distribution network [17] and geological carbon sequestration [18], or FSI problems [19, 20]. Specifically, we rely in this work on Gaussian Process (GP) to approximate the mapping $\mathcal{P} : \Omega \mapsto \mathbb{R}$. This statistical approach uses a coarse set of performance evaluations at some selected parameters values $\mathbf{x} \in \Omega$ to infer a GP $\mathcal{G}(\mathbf{x}) \approx \mathcal{P}(\mathbf{x})$. Given the GP approximation one can apply his favorite optimization procedure substituting $\mathcal{G}$ to $\mathcal{P}$ in (1), and obtain an approximation of $\mathbf{x}_{\text{opt}}$. This surrogate-based optimization procedure is embedded in an iterative scheme, where new evaluations of the performance at carefully selected new points $\mathbf{x}$ are introduced in order to refine the GP approximation in regions of Ω of interest, that is susceptible to include the optimum. The GP approach is then expected to improve the direct optimization of $\mathcal{P}$ by a) requiring an overall lower number of performance evaluations, compared to direct gradient-free approaches, and b) enabling the use of efficient global optimization tools.

Another interest of GP-based optimization is that it naturally accommodates for errors and noise in the performance evaluation. This feature is especially attractive in the case of optimizations relying on complex numerical simulations, where both modeling and numerical errors are expected to be significant and hardly reducible. To illustrate the interest of the robustness of the GP-based optimization to inherent error, this work focuses on the optimization of a scaled IMOCA mainsail in upwind conditions. The objective is to find the optimal trimming of the sail, for a performance criterion combining the drive and side aerodynamic force coefficients. The GP-based optimization is performed first on a physical model using measurements of $\mathcal{P}(\mathbf{x})$ performed in the wind tunnel of the Yacht Research Unit (the University of Auckland), for the sequence of trimming points $\mathbf{x}$ requested by the iterative optimization procedure. In this case, the error in the observed values of the performance is due to the imperfections in the experimental apparatus and the inherent noise in the measurements. This experiment shows that the significant error in the measurements of $\mathcal{P}(\mathbf{x})$ compromises the convergence of descent methods [21] without its implicit account in the GP reconstruction.

The experimental optimization is subsequently used to assess the relevance of numerical optimizations,

where a non-linear FSI solver is used to compute the performance. To this end, a high fidelity numerical model was created from measurements of the experimental model (e.g. dimensions, and sail geometry, mechanical characteristics of mast and boom, wind tunnel inflow conditions, . . .). A state of the art FSI solver is then used to solve the numerical model at the sequence of trimming points requested by the GP-based optimizer. The FSI solver involves a nonlinear structural solver coupled to a finite volume flow solver with an Unsteady Reynolds-Average Navier-Stokes Equations (URANS) model. The comparison of the experimental and numerical optimal trimming parameters shows a significant discrepancy, as expected from both numerical modeling and experimental errors. However, it is found that the agreement between the numerical and experimental optimal performances is consistent with the error and noise levels estimated by the GP constructions. This observation has motivated further uncertainty quantification studies to determine the dominant source of model error affecting the optimal trimming parameters and the performance. Because of the highly non-linear and coupled nature of the FSI solver, the uncertainty analyses rely again on local surrogates of the performance, using a Polynomial Chaos (PC) expansion [22] constructed on sparse grids. These surrogates allow to quantify the variability in the achieved performance and in the optimal trimming parameters. They are also used to propose a robust optimum.

The paper is organized as follows. Section 2 reviews the construction of the Gaussian process approximation and the iterative optimization procedure detailing the selection of the new parameters. We then detail the experimental set-up and report the results of the experimental optimization in Section 3, contrasting the cases of GP with and without noise assumption. The numerical modeling of the experiment and the FSI solver are discussed in Section 4, which also presents the corresponding optimization results. Section 5 discusses the various possible source of disagreement between the experimental and numerical optima, and presents several uncertainty quantification analyses to assess their respective importance and impact. Finally, conclusions of this work and direction for future developments are provided in Section 6.

2. GP-Based Optimization

The use of surrogate models is a classical approach to reduce the computational burden related to the optimization of complex systems [23]. Polynomial models were firstly used as surrogates, thanks to their ease of construction. However, more sophisticated meta-models such as Gaussian processes (GP) [24] or Support Vector Machines (SVM) [25] have emerged so far. GP models have been found especially appealing for optimization purpose, because their statistical nature allows to provide both a prediction of the performance function, in terms of model mean, and an error estimate, in terms of model variance. Optimization procedures need also to be adapted to the statistical nature of the approximation. Here, we shall rely on the rigorous optimization framework based on the maximization of the Expected Improvement (EI) criterion, referred as the Efficient Global Optimization (EGO) [26] method. An additional key feature of the GP construction exploited in this work, is that it can naturally deal with observation noise. We also mention the recent extensions to multi-point optimization proposed [27, 28].

In this Section we start by introducing the definitions of a Gaussian process and briefly summarizing the construction of a Gaussian process to model a function from noisy observations. More details on GP models can be found, *e.g.*, in [29, 30]. We then describe the GP-based optimization procedure [31], detailing the selection of successive optimal candidates.

2.1. Gaussian Process

Gaussian Process modeling is a statistical method to approximate functions from a finite set of observations, possibly noisy, at arbitrary points. The method takes advantage of the full definition of a Gaussian process from its second-order characteristics, namely its mean and covariance functions. Using the observations, it explicitly updates the prior mean and covariance functions into their posterior counterparts, minimizing the mean squared-error of the estimator. For Gaussian observation noise, the estimator corresponds to the optimal Bayesian posterior. Compared to alternative regression-type approaches, the Gaussian process modeling also avoids the need to prescribe explicitly a basis for the approximation space, which can be infinite. Instead, one needs to specify the prior covariance structure, which is usually selected using objective criteria (see below) in a family of parametrized kernels.

For a formal definition of the Gaussian processes, we are considering a probability space Ω, a measurable space E and a random variable defined by the measurable application $X : \Omega \to E$. The space $E \subset \mathbb{R}^d$ and X is called a real-valued random variable when $d = 1$. Gaussian variables are a particular class of the random ones and the real-valued random variable X is a Gaussian variable if $\exists (\mu, \Sigma) \in \mathbb{R} \times \mathbb{R}^+$, $Z \sim \mathcal{N}(0, 1)$ and $X = \mu + \Sigma Z$. We write $X \sim \mathcal{N}(\mu, \Sigma^2)$.

Definition 1. *A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution [30].*

The mean function $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$, that fully parameterize a Gaussian process of a real process $f(\mathbf{x})$, are defined by

$$m(\mathbf{x}) = \mathbb{E}\{f(\mathbf{x})\}, \tag{2}$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\{f(\mathbf{x}) - m(\mathbf{x}), f(\mathbf{x}') - m(\mathbf{x}')\}, \tag{3}$$

and a Gaussian process is written as

$$f(\mathbf{x}) \sim \mathcal{G}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')). \tag{4}$$

2.2. Gaussian Process Model

We first focus on the construction of a GP to approximate a mapping $f : \mathbf{x} \in \Omega \mapsto f(\mathbf{x}) \in \mathbb{R}$, from a set $\mathbf{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^{\mathrm{T}}$ of n training inputs vectors, or observation points, $\mathbf{x}_i \in \Omega$. Each component $\mathbf{x}_i \in \mathbf{X}_n$ is associated to an observation (or measurement) $y_i \in \mathbb{R}$ which is assumed to be dependent on the latent function $f(\mathbf{x})$ through

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{5}$$

where ε_i is a random measurement error (*i.e.* the measurement noise). In this work, the ε_i are assumed *independent* and identically distributed, with (centered) Gaussian distribution:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma_\epsilon{}^2). \tag{6}$$

In the following, $\sigma_\epsilon{}^2$ is referred to as the noise variance.

In the GP model, the latent function is considered as a realization of a zero-mean multivariate Gaussian process F, with unknown covariance function C_F; that is $F \sim \mathcal{N}(0, C_F)$, with

$$C_F(\mathbf{x}, \mathbf{x}') \doteq \mathbb{E}\{F(\mathbf{x}), F(\mathbf{x}')\}, \tag{7}$$

where $\mathbb{E}\{\cdot\}$ denotes the expectation operator.

The covariance function of F must be specified. We shall consider the Matérn class [32] of stationary covariance functions, given by one-dimensional generator with form

$$M_\nu(r, l) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}r}{l}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}r}{l}\right). \tag{8}$$

Here $r \doteq |x - x'|$, ν and l are two positive parameters, and K_ν is the modified Bessel function of the second kind. We shall further restrict ourselves to covariances with $\nu \to \infty$, leading to the classical squared exponential covariance family with generator

$$M_\infty(r, l) = \exp\left(\frac{-r^2}{2l^2}\right). \tag{9}$$

Other kernel functions could be used as well. However, as shown in the benchmark from [33], the influence of the kernel choice is low in general, and only noticeable for multi-modal objective functions, which is not the case in the present study. Various other definitions of kernel are given in [30].

The multidimensional counterpart is defined as the product of the one-dimensional generator. The final expression of the covariance function for the GP approximation is

$$C_F(\mathbf{x}, \mathbf{x}'; \boldsymbol{\Theta}) \doteq \theta_1 \prod_{i=1}^{d} \exp\left(\frac{-(x_i - x_i')^2}{2l_i^2}\right) + \theta_2. \tag{10}$$

In the expression (10) of C_F, $\boldsymbol{\Theta} = \{\theta_1, \theta_2, l_1, l_2, \ldots, l_d\}$ is a vector of hyper-parameters that will have to be inferred from the observations. The first hyper-parameter, θ_1, scales the distance-dependent correlation, while θ_2 is an offset from zero. The remaining parameters l_i are the anisotropic correlation lengths associated to the d directions of Ω. We denote $\mathbf{C}(\boldsymbol{\Theta}) \in \mathbb{R}^{n \times n}$ the covariance matrix for the observation points in $\mathbf{X}_n$, having entries

$$C_{i,j}(\boldsymbol{\Theta}) \doteq C_F(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\Theta}), \quad 1 \leq i, j, \leq n. \tag{11}$$

The vector $\mathbf{Y}_n = (y_1, \ldots, y_n)^{\mathrm{T}}$ of noisy observations and the predicted observation $y(\mathbf{x})$ at a point $\mathbf{x} \in \Omega$ have for joint Gaussian distribution

$$\begin{pmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{pmatrix} \Big| \mathbf{X}_n, \boldsymbol{\Theta}, \sigma_\epsilon^2 \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{C} + \sigma_\epsilon^2 \mathbf{I} & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^{\mathrm{T}}(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\epsilon^2 \end{bmatrix}\right). \tag{12}$$

The dependence of $\mathbf{C}$ on the hyper-parameters has been removed in (12) to simplify the notation; in addition, we denoted

$$\kappa(\mathbf{x}) \doteq C_F(\mathbf{x}, \mathbf{x}; \boldsymbol{\Theta}), \quad \mathbf{k} \doteq (C_F(\mathbf{x}, \mathbf{x}_1; \boldsymbol{\Theta}), \ldots, C_F(\mathbf{x}, \mathbf{x}_n; \boldsymbol{\Theta}))^{\mathrm{T}},$$

while $\mathbf{I}$ is the identity of $\mathbb{R}^n$. The rule of conditional probabilities [34, 30] allows to write the inference of $y(\mathbf{x})$ given the noisy observations $\mathbf{Y}_n$ as

$$p(y(\mathbf{x})|\mathbf{Y}_n) = \frac{p(y(\mathbf{x}), \mathbf{Y}_n)}{p(\mathbf{Y}_n)}, \tag{13}$$

and noting $\mathbf{C}_{n+1} = \begin{bmatrix} \mathbf{C} + \sigma_\epsilon^2 \mathbf{I} & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^{\mathrm{T}}(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\epsilon^2 \end{bmatrix}$, that predictive posterior distribution is

$$p(y(\mathbf{x})|\mathbf{Y}_n) \propto \exp\left(-\frac{1}{2}\begin{bmatrix} \mathbf{Y}_n & y(\mathbf{x}) \end{bmatrix} \mathbf{C}_{n+1}^{-1} \begin{bmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{bmatrix}\right). \tag{14}$$

Using the partitioned inverse equation [35], $\mathbf{C}_{n+1}^{-1}$ is written

$$\mathbf{C}_{n+1}^{-1} = \begin{bmatrix} \mathbf{M}_n & \mathbf{m}_{n+1} \\ \mathbf{m}_{n+1}^{\mathrm{T}} & \mu \end{bmatrix}, \tag{15}$$

where

$$\mu = \left(\kappa(\mathbf{x}) + \sigma_\epsilon^2 - \mathbf{k}^{\mathrm{T}}(\mathbf{x})\left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}\right)^{-1} \mathbf{k}(\mathbf{x})\right)^{-1},$$

$$\mathbf{m}_{n+1} = -\mu\left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}\right)^{-1} \mathbf{k}(\mathbf{x}),$$

$$\mathbf{M}_n = \left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}\right)^{-1} + \frac{1}{\mu}\mathbf{m}_{n+1}\mathbf{m}_{n+1}^{\mathrm{T}}.$$

Then, we substitute in equation (14) and we find

$$p(y(\mathbf{x})|\mathbf{Y}_n) \propto \exp\left(-\frac{(y(\mathbf{x}) - \hat{y}(\mathbf{x}))^2}{2\hat{\sigma}_y^2(\mathbf{x})}\right). \tag{16}$$

To the end, we have $y(\mathbf{x})|\,\mathbf{Y}_n, \mathbf{X}_n, \boldsymbol{\Theta}, {\sigma_\epsilon}^2 \sim \mathcal{N}\left(\hat{y}(\mathbf{x}), \hat{\sigma}_y^2(\mathbf{x})\right)$, with *best* prediction (mean of $y(\mathbf{x})$) $\hat{y}(\mathbf{x})$ and prediction variance $\hat{\sigma}_y^2(\mathbf{x})$ given by

$$\hat{y}(\mathbf{x}) = \mathbf{k}^{\mathrm{T}}(\mathbf{x})\left(\mathbf{C}(\boldsymbol{\Theta}) + {\sigma_\epsilon}^2\mathbf{I}\right)^{-1}\mathbf{Y}_n, \tag{17}$$

$$\hat{\sigma}_y^2(\mathbf{x}) = \kappa(\mathbf{x}) + {\sigma_\epsilon}^2 - \mathbf{k}^{\mathrm{T}}(\mathbf{x})\left(\mathbf{C}(\boldsymbol{\Theta}) + {\sigma_\epsilon}^2\mathbf{I}\right)^{-1}\mathbf{k}(\mathbf{x}). \tag{18}$$

The hyper-parameters $\boldsymbol{\Theta}$ and noise variance ${\sigma_\epsilon}^2$ are unknown *a priori* and must to be learned from the observations. It is done by maximization of the evidence, which involves to maximize the probability density of the observations:

$$f\left(\mathbf{Y}_n|\,\mathbf{X}_n, \boldsymbol{\Theta}, {\sigma_\epsilon}^2\right) = (2\pi)^{\frac{n}{2}}\left|\mathbf{C}(\boldsymbol{\Theta}) + {\sigma_\epsilon}^2\mathbf{I}\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\mathbf{Y}_n^{\mathrm{T}}\left(\mathbf{C}(\boldsymbol{\Theta}) + {\sigma_\epsilon}^2\mathbf{I}\right)^{-1}\mathbf{Y}_n\right). \tag{19}$$

The likelihood function is given by $L(\boldsymbol{\Theta}, {\sigma_\epsilon}^2) \doteq L(\boldsymbol{\Theta}, {\sigma_\epsilon}^2; \mathbf{Y}_n) \doteq f\left(\mathbf{Y}_n|\,\mathbf{X}_n, \boldsymbol{\Theta}, {\sigma_\epsilon}^2\right) \in \mathbb{R}^+$. In the present work, $(\boldsymbol{\Theta}, {\sigma_\epsilon}^2)$ are determined by maximizing the log-marginal likelihood [30] defined by

$$\mathcal{L}(\boldsymbol{\Theta}, {\sigma_\epsilon}^2) = -\frac{n}{2}\log\left(2\pi\right) - \frac{1}{2}\log\left|\mathbf{C}(\boldsymbol{\Theta}) + {\sigma_\epsilon}^2\mathbf{I}\right| - \frac{1}{2}\mathbf{Y}_n^{\mathrm{T}}\left(\mathbf{C}(\boldsymbol{\Theta}) + {\sigma_\epsilon}^2\mathbf{I}\right)^{-1}\mathbf{Y}_n. \tag{20}$$

$\mathcal{L}(\boldsymbol{\Theta}, {\sigma_\epsilon}^2) \doteq \log\left[L(\boldsymbol{\Theta}, {\sigma_\epsilon}^2)\right] \in \mathbb{R}$ is the natural logarithm of the likelihood and it avoids to work with very small probability density values. An evolution strategy algorithm [13] is used to maximize $\mathcal{L}(\boldsymbol{\Theta}, \sigma_\epsilon)$ with respect to its arguments. The GP model can be used to predict values at new points $\mathbf{x}$ using (17) and (18). The most computationally demanding part of the GP construction is the inversion (and determinant calculation) of the (full) matrix $\left(\mathbf{C} + {\sigma_\epsilon}^2\mathbf{I}\right)$, required in the definition of the log-marginal likelihood and for new predictions. For the problems treated in this work, this is efficiently performed by LU decomposition.

The effect of the observation noise σ_ϵ on the constructed GP model is illustrated in Figure 1 for a one-dimensional function. The construction is based on 6 observations points, depicted with circles in the plots, and the covariance hyper-parameters $\boldsymbol{\Theta}$ are determined by maximizing the log-marginal likelihood. However, in Figure 1(a) a value $\sigma_\epsilon = 0$ is imposed, while in Figure 1(b) the noise level is also optimized. The two plots report the mean $\hat{y}$ of the GP models with the standard $\pm 3\hat{\sigma}$ uncertainty range ($\approx 95\%$ confidence interval). In the noise-free ($\sigma_\epsilon = 0$) case, the resulting GP model interpolates the observations with a variance of the prediction equal to zero at the observation points. However, the mean of the GP model exhibits significant oscillations such that over-fitting can be suspected. In contrast, optimizing the noise level σ_ϵ in addition to $\boldsymbol{\Theta}$ results in a mean process mostly free of oscillations but no more interpolating, as it can be appreciated from Figure 1(b). The optimal value $\sigma_\epsilon = 0.143$ is characteristic of the averaged distance between the best prediction and the observations.

2.3. *Optimization using Gaussian Process Models*

GP models constitute an attractive alternative to approximate performance functions in complex optimization problems where $\mathcal{P}(\mathbf{x})$ is costly to evaluate. More precisely, from a set of (possibly noisy) estimates of $\mathcal{P}(\mathbf{x}_i)$, at n observation points $\mathbf{x}_i \in \mathbf{X}_n$, we construct a GP model. This GP model can then be used in the optimization procedure, to estimate at a low computational cost the performance $\mathcal{P}(\mathbf{x})$ at new tentative points $\mathbf{x}$. In the following, we shall denote $\widehat{\mathcal{P}}_n$ the predictive mean of the performance, and $\hat{\mathbf{x}}_n$ the corresponding optimum

$$\hat{\mathbf{x}}_n = \arg\min_{\mathbf{x}\in\Omega} -\widehat{\mathcal{P}}_n(\mathbf{x}). \tag{21}$$

It is expected that $\hat{\mathbf{x}}_n \approx \mathbf{x}_{\mathrm{opt}}$ if the approximation error $\widehat{\mathcal{P}}_n - \mathcal{P}$ is small enough. This claim can be made more formal assuming sufficient regularity on $\mathcal{P}(\mathbf{x})$ and $\widehat{\mathcal{P}}_n$. The advantage is that $\widehat{\mathcal{P}}_n(\mathbf{x})$ is inexpensive to evaluate (compared to the evaluation of $\mathcal{P}(\mathbf{x})$) and can be easily differentiated (see (17)), enabling efficient gradient-based optimization methods. The error between $\hat{\mathbf{x}}_n$ and $\mathbf{x}_{\mathrm{opt}}$ is a concern; in the approach below, the error on the optimum is iteratively reduced by completing the set of observation points $\mathbf{X}_n$.
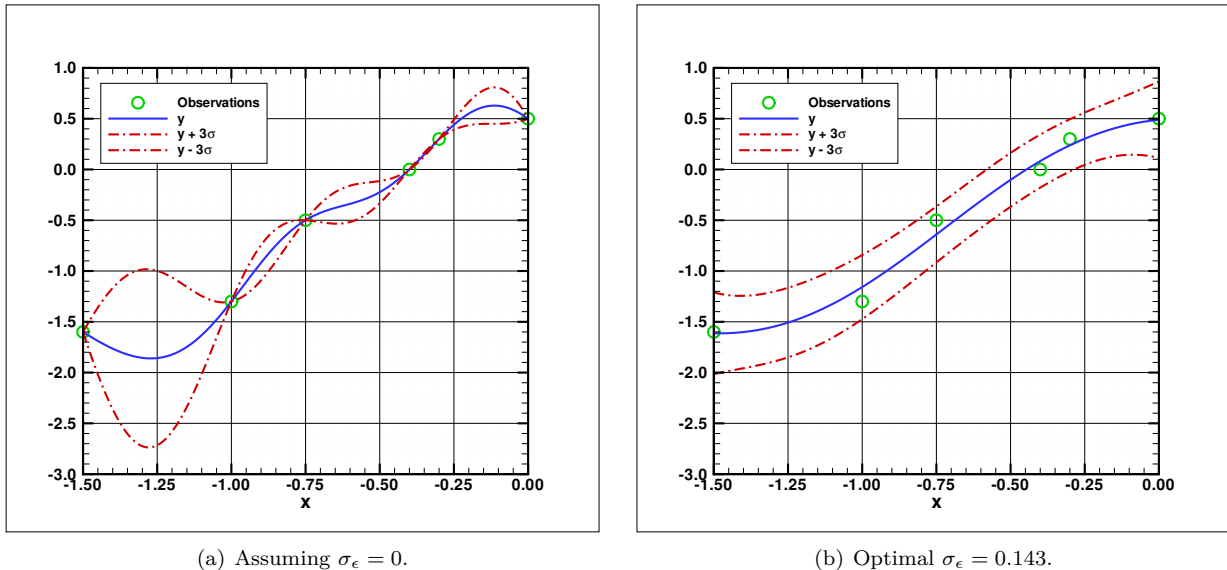
(a) Assuming $\sigma_\epsilon = 0$.

(b) Optimal $\sigma_\epsilon = 0.143$.

Figure 1: Effect of σ_ϵ on the GP model.

It is clear that the error between $\mathcal{P}$ and $\widehat{\mathcal{P}}_n$ only matters in the neighborhood of the optimum, so the completion of $\mathbf{X}_n$ with a new point $\mathbf{x}_{n+1}$ should aim at reducing the predictive error close to $\mathbf{x}_{\text{opt}}$ while ensuring that $\widehat{\mathcal{P}}_n(\mathbf{x})$ has no other spurious optimum. A deterministic optimization procedure would classically choose $\mathbf{x}_{n+1}$ as the optimal point given our current approximation, that is $\mathbf{x}_{n+1} = \hat{\mathbf{x}}_n$ with $\hat{\mathbf{x}}_n$ given in (21). However, the GP model of $\mathcal{P}(\mathbf{x})$ provides not just a predictive mean, but a whole distribution of values at any $\mathbf{x}$. This probabilistic information can be exploited to construct more robust approaches for the selection of $\mathbf{x}_{n+1}$, particularly during the initial stages of the optimization when $\widehat{\mathcal{P}}_n$ may have significant errors. Such improved strategies are based on merit functions which usually are composite functions involving the predictive mean $(\widehat{\mathcal{P}}_n)$ and the variance $(\hat{\sigma}_{\mathcal{P}}^2)$ of the GP model; the new point $\mathbf{x}_{n+1}$ is selected as the optimum of such composite merit function. Specifically, the merit functions should balance an optimization of $\widehat{\mathcal{P}}_n$ (optimality) with the selection of $\mathbf{x}_{n+1}$ in areas of large variance $\hat{\sigma}_{\mathcal{P}}^2$ to reduce the GP model error. A review of various merit functions proposed in the literature is provided in [36]. In this work, we use the Augmented Expected Improvement (AEI) merit function [37], which is an extension of the popular Expected Improvement (EI) of [26] to the case of noisy estimations. The AEI merit function $AEI(\mathbf{x})$ estimates the expected increase in the performance, taking into account the noise in the observed values and penalizing areas where the variance $\hat{\sigma}_{\mathcal{P}}^2$ is small. It writes as

$$AEI(\mathbf{x}) = EI(\mathbf{x})\left(1 - \frac{\sigma_\epsilon}{\sqrt{\hat{\sigma}_{\mathcal{P}}^2(\mathbf{x}) + \sigma_\epsilon{}^2}}\right), \tag{22}$$

where the Expected Improvement is defined by

$$EI(\mathbf{x}) = \hat{\sigma}_{\mathcal{P}}^2(\mathbf{x})\left[u(\mathbf{x})\Phi\left(u(\mathbf{x})\right) + \phi\left(u(\mathbf{x})\right)\right], \quad u(\mathbf{x}) = \frac{\widehat{\mathcal{P}}_n(\mathbf{x}_{*,n}) - \widehat{\mathcal{P}}_n(\mathbf{x})}{\hat{\sigma}_{\mathcal{P}}^2(\mathbf{x})}, \tag{23}$$

with Φ and $\phi = \Phi'$ the cumulative distributions (Erf-function) and density of the standard Gaussian distribution, and $\mathbf{x}_{*,n}$ is the effective best solution over the set of current observation points $\mathbf{X}_n$:

$$\mathbf{x}_{*,n} \doteq \underset{\mathbf{x}_i \in \mathbf{X}_n}{\arg\min}\left[\widehat{\mathcal{P}}_n(\mathbf{x}_i) + \hat{\sigma}_{\mathcal{P}}(\mathbf{x}_i)\right]. \tag{24}$$

7

Then, we set the new points $\mathbf{x}_{n+1}$ as the optimum point of the AEI, and add the new point to $\mathbf{X}_n$ setting $n \leftarrow n + 1$. The iteration is completed by evaluating the performance $\mathcal{P}(\mathbf{x}_n)$ of the new point, solving the full model. The next iteration can then proceed, starting by updating the GP model using the new observation point and its performance. Overall, each iteration requires one numerical model resolution to compute $\mathcal{P}(\mathbf{x}_n)$, and the resolution of two optimization problems, the first one to determine the covariance parameters of the GP model, and the second one to find AEI optimum. For these optimization problems, the results reported in this work were obtained using the nonlinear non-convex black-box optimization library based on the Covariance Matrix Adaptation Evolution Strategy [38, 13]. The iterations of the GP-based optimization problem are continued until a stopping criterion is satisfied or the resources allocated to the optimization procedure have been exhausted (*e.g.*, fixing a maximum number of performance evaluations). Classically, the stopping criterion compares the distance between successive iterates. The distance can be measured from the successive AEI optimal points, $\mathbf{x}_n$, successive predictive optimal points, $\hat{\mathbf{x}}_n$ given by (21), or based on the differences in successive AEI values, $AEI(\mathbf{x}_{n+1})$, or optimal predictive performance $\widehat{\mathcal{P}}_n(\hat{\mathbf{x}}_n)$.

The optimization procedure is schematically presented in Figure 2, in the case of the numerical optimization considered in Section 4. In the top block, labeled sampling loop, an initial set of observation points $\mathbf{x}_i \in \Omega$ are generated. In the present work we rely on a Latin Hypercube Sample (LHS) scheme [39] to randomly generate the initial sample set. The performance $\mathcal{P}(\mathbf{x}_i)$ is also evaluated for all the initial sample points. The lower block, label iterative loop in Figure 2, is the core of the optimization procedure consisting in the repetition of the sequence of GP model construction, AEI optimization and performance evaluation at the new point $\mathbf{x}_n$ selected, as well as the convergence check.
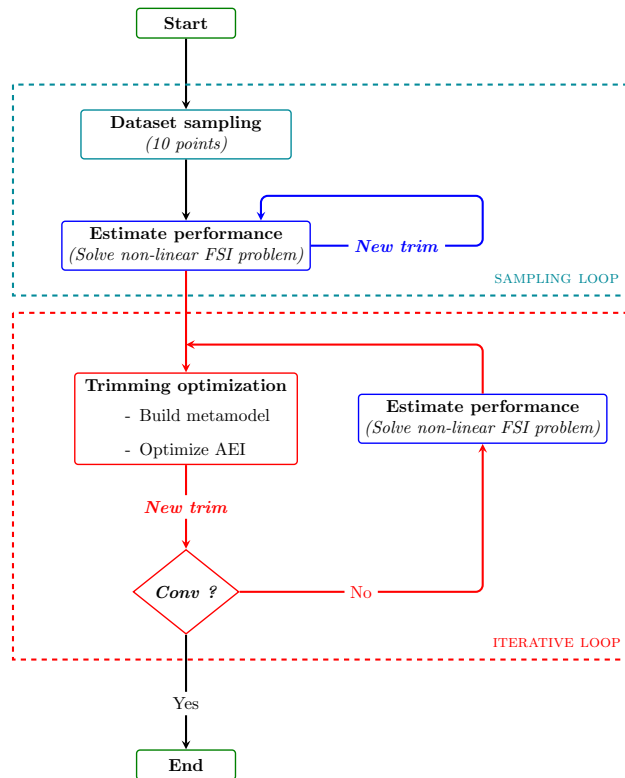

Figure 2: Schematic of the GP-model based optimization procedure.

As explained above, the performance evaluations, which represent the bulk of the computational cost, are divided between the initial sampling of observation points and the additional entries governed by the AEI criterion. As with most derivative-free approaches, the number of evaluations required to determine a

satisfactory approximation of the optimum can grow rapidly with the number of parameters. Obviously, this cost depends strongly on the characteristics of the performance function, such as its smoothness, anisotropy, separability, making the prescription of guidelines tedious. As discussed in references provided, GP models are usually restricted to few dozens of parameters. The reader interested in a more precise quantification of the computational cost and the best strategy to allocate resources, in particular for problems in larger dimension, can refer to [40].

3. Experimental Optimization

This Section concerns the sail trimming optimization performed in the wind tunnel of the Yacht Research Unit (YRU) [41].

3.1. Experimental Setup

The experimental sail is inspired by an IMOCA 60-foot design mainsail at 1:13 scale, designed and manufactured by INCIDENCE SAILS company. The surface area is 1 m^2, for a height of $h = 2$ m; it is supported by a rig made of a flexible circular section carbon mast (constant diameter 14 mm). The mast is simply clamped at its base and has no spreader, backstay or forestay. The sail and rig are set in the open jet test section of the YRU wind tunnel, see Figure 3(a). The test section is 7.2 m wide and 3.5 m high.

Three stepper motors and a control card are used to remotely control the main sheet length (L_sheet) and main car position (L_car), which are the two trimming parameters to be optimized. The remote control system, shown in Figure 3(b), allows us to enforce new trimmings with a precision of ± 2 mm, without switching off the wind tunnel flow and recalibrating the measurement devices. These devices consist in a six-component force balance, located under the floor of the wind tunnel, to measure the aerodynamic forces, a load sensor (± 0.02 N precision, range 5 daN) to measure the load in the sheet and a V-SPARS acquisition system [42] to track the position of five fluo-red stripes across the sail (see Figure 3(a)) and record the sail's flying shape. The force balance measures the thrust force in the X-direction corresponding to the model longitudinal forward direction, the side force in the Y-direction defined as the positive port-side. The Z-direction of the balance is pointing upward. After careful calibration checking, the balance precisions in the X, Y and Z directions were estimated to be ± 0.09 N, ± 0.11 N and ± 0.27 N respectively.

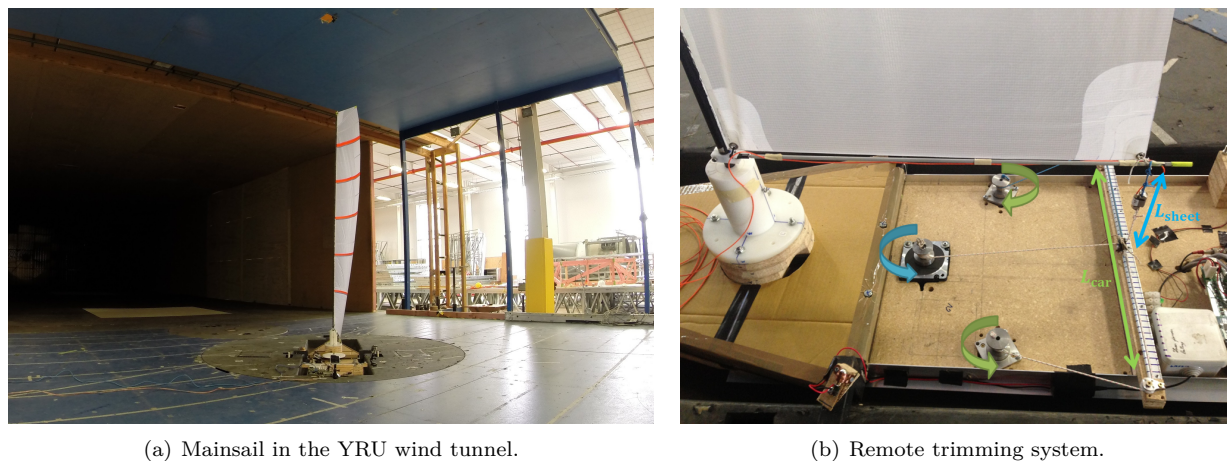(a) Mainsail in the YRU wind tunnel.

(b) Remote trimming system.

Figure 3: Experimental setup.

The wind tunnel inflow velocity was measured and found to have an apparent wind speed (AWS) of 3.5 ± 0.15 m/s for an apparent wind angle (AWA) of 40 ± 2 deg. The experimental Reynolds number, based on the sail chord length $c = S/h = 0.5$ m, is Re $= 1.2\,10^5$. A multi-hole pressure probe (Cobra Probe) was used to measure profiles of the flow velocity at several locations inside the wind tunnel, in particular in the

inflow section. These measurements were repeated with and without the sail model to observe its effects on the flow field.

The optimization procedure is used to maximize $\mathcal{P}(\mathbf{x})$, where $\mathbf{x} = (L_{\text{sheet}}, L_{\text{car}})$. The performance (objective) function is a linear combination of the thrust and side force aerodynamic coefficients, denoted respectively C_X and C_Y:

$$\mathcal{P}(\mathbf{x}) = C_X(\mathbf{x}) + 0.1 C_Y(\mathbf{x}). \tag{25}$$

The aerodynamic coefficients are the aerodynamic forces normalized by $S \times p_\infty$, where p_∞ is the measured dynamic pressure (precision ± 1 Pa) of the inflow. In addition to the trust coefficient to be maximized, the objective function in (25) penalizes the (negative) side force coefficient C_Y, with a weight coefficient $1/10$, to account for the hydrodynamic drag and leeway that are detrimental to yacht performance. The weight value is selected on the basis of previous numerical hydrodynamic studies on a Figaro sailing yacht. Changing the weight value affects the objective function and optimal trimmings such that a sensitivity analysis would be necessary if it is not well known. Alternatively, a more complex Velocity Prediction Program can be used to estimate the performance in lieu of (25). Nevertheless the proposed optimization procedure will remain the same.

The optimization follows the procedure illustrated in Figure 2, except that the performance is not computed but measured experimentally. Specifically, for each new point $\mathbf{x}_n$, the sail is remotely trimmed to the corresponding L_{car} and L_{sheet} values and, after the flow has stabilized, the measured aerodynamic forces are averaged over a 30 s period to filter-out noisy fluctuations. The averaged forces are subsequently normalized to estimate the experimental performance denoted $\mathcal{P}^{\text{exp}}(\mathbf{x}_n)$. In these experiments, the optimization procedure is considered converged when two successive new trimming points $\mathbf{x}_n$ are within a distance less than 1 mm, a tolerance corresponding to the precision on the trimming parameters that can be applied experimentally.

In the rest of the section, we apply the optimization procedure in two different settings in order to demonstrate the effectiveness of the GP-based optimization on noisy experimental performance evaluations. To this end, we first assume in section 3.2 a perfect performance measure, enforcing a measurement noise $\sigma_\epsilon = 0$ (noise-free situation) in the GP construct, so the GP model exactly interpolates the experimental points. This noise-free assumption is subsequently compared in section 3.3 to the case where the GP construction uses a finite measurement noise estimated from the experimental observations.

3.2. Experimental Optimization with $\sigma_\epsilon = 0$

In a first series of experiments, the optimization is carried out using GPs constructed fixing $\sigma_\epsilon = 0$. This situation corresponds to perfect performance measure assumption for the points in the initial sample set and subsequent optima of the AEI merit functions. As the result of setting $\sigma_\epsilon = 0$, the posterior mean of the GPs interpolates the experimental values. This is illustrated in Figure 4 which depicts the GP model of $\mathcal{P}^{\text{exp}}(\mathbf{x})$ after iteration 36 of the optimization procedure. Figure 4(a) shows the color contours of the posterior mean as a function of the trimming parameters L_{car} and L_{sheet} over the optimization domain, and Figure 4(b) shows the corresponding standard deviation. Also shown using black dots are the 36 sample points of the initial set and successive AEI optima, that are used to construct the GP model. A large dispersion of the sample points is reported, the points being scattered over all the domain Ω. This dispersion denotes that the successive AEI merit function optima are not converging. Instead, the algorithm continues to explore Ω without discovering a particular sub-domain more likely to contain the true global optimum. This is confirmed by inspecting the posterior mean of the GP model, in Figure 4(a), which presents at least 2 local minima, but without significantly attracting sample points. This behavior can be explained from the standard deviation field reported in Figure 4(b); it has a zero value at the sample points (consistently with enforcing $\sigma_\epsilon = 0$) but quickly increases when moving away from these points: this structure dominates the AEI merit function and leads the optimization procedure to propose new points $\mathbf{x}_{n+1}$ in less explored areas because of the high variance of the GP model. In addition to distracting the iterates from promising areas to favor the exploration of highly uncertain regions, the inappropriately high confidence (low standard deviation) of the posterior mean in the immediate neighborhood of the sample points induces spurious local

minima, due to the interpolatory nature of the mean. This behavior highlights the fact that the GP model fits the experimental noise.

(a) GP model mean.

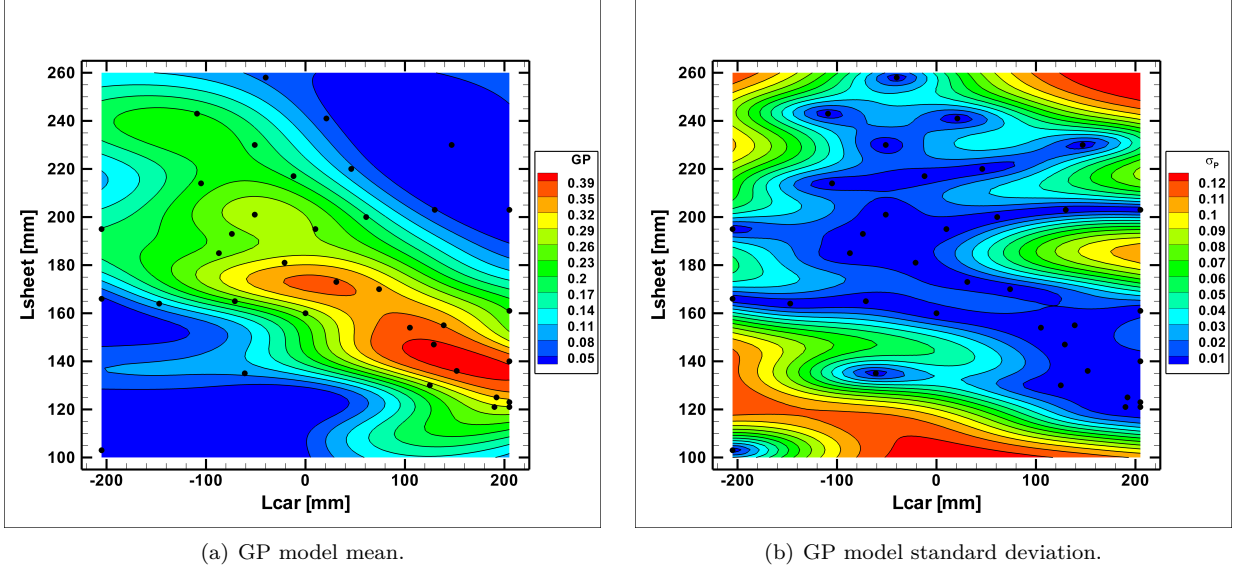(b) GP model standard deviation.

Figure 4: GP model of $\mathcal{P}^{\text{exp}}(\mathbf{x})$ at iteration 36 and assuming $\sigma_\epsilon = 0$.

Fitting the experimental noise is highly prejudicial in the present example, because the error in the performance is significant. The effect can be appreciated in Figure 5 which shows the sequence of experimental performance measurements $\mathcal{P}^{\text{exp}}(\mathbf{x}_n)$ as a function of the iteration index (the first 10 iterations correspond to the initial sampling and are not actual iterations of the optimization algorithm). The plot shows that the sequence is not converging but sustains large fluctuations from an iterate to another. The magnitude of the fluctuations have remained essentially the same as for the initial random sampling, denoting the absence of improvement of the proposed optimum performance with the iterations. Clearly, the over-fitting of the noisy measurement $\mathcal{P}^{\text{exp}}$ is responsible for the failure of the present approach, which keeps exploring the domain without converging to the maximum underlying performance.

3.3. Experimental Optimization with finite noise

We next repeat the previous experiment, starting from the same initial sample set, but using a fixed finite measurement noise in the construction of the GPs. In this experiment, the value of σ_ϵ is not determined as part of the GP construction optimizing the log-marginal likelihood in (20), but is fixed to $\sigma_\epsilon = 0.027$. This value for σ_ϵ was experimentally estimated from repeated measurements at different trimmings points. Other than using a finite fixed non-zero value for σ_ϵ, the optimization procedure is the same as previously. However, the initial GP constructed from the initial sample set being different, the sequence of optima proposed is entirely different. In particular, the optimization reaches the stopping criterion at iteration 33.

Figure 6 depicts the GP at convergence and should be contrasted with the previous case shown in Figure 4. In particular, we observe that the posterior mean is much smoother than previously, with a single well-defined global optimum in Ω. In addition, it is seen that the sample points cluster in the neighborhood of the optimum, in sharp contrast to the previous case with $\sigma_\epsilon = 0$. The optimal trimming at convergence, $\hat{\mathbf{x}}_n$, is found for $L_{\text{sheet}} = 133$ mm and $L_{\text{car}} = 138$ mm corresponding to a predicted (GP) performance $\widehat{\mathcal{P}}_n^{\text{exp}}(\hat{\mathbf{x}}_n) = 0.397$. The standard deviation of the GP model, shown in Figure 6(b), is minimal in the neighborhood of the optimum, but remains finite and assumes values $\gtrsim \sigma_\epsilon$. Other regions of Ω far from the optimum are not explored by the optimization process, although the variance can be large.

The convergence of the optimization procedure can also be appreciated from the plot of Figure 7, to be contrasted with Figure 5. The improvement in the measured performance, after the first 10 initial random
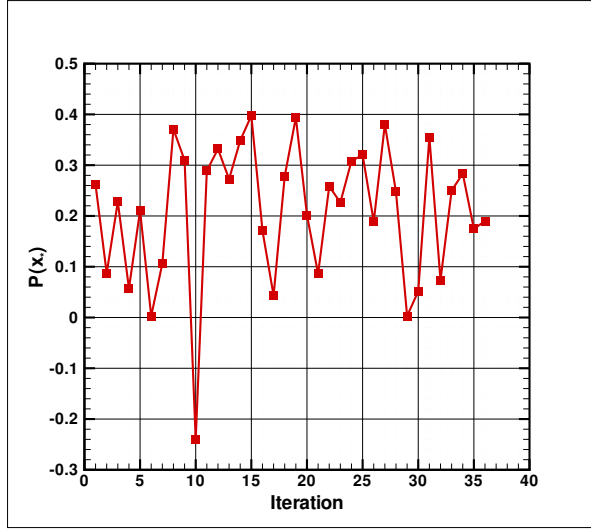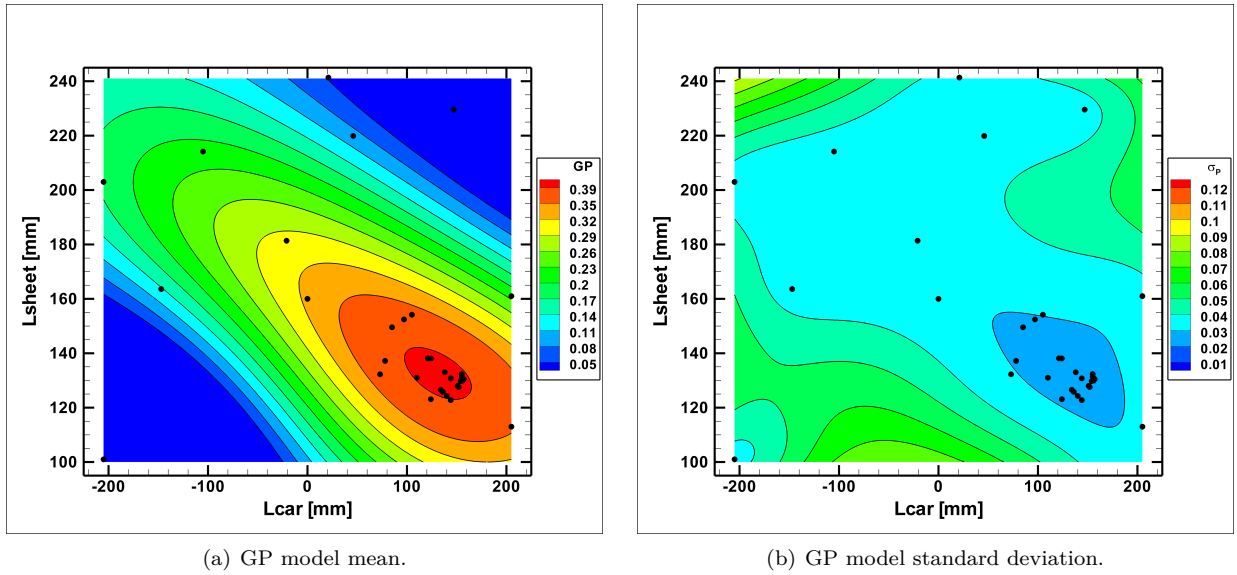
11

Figure 5: Sequence of measured $\mathcal{P}^{\mathrm{exp}}(\mathbf{x}_n)$ as a function of the iteration index. Case of $\sigma_\epsilon = 0$.

(a) GP model mean.

(b) GP model standard deviation.

Figure 6: GP model of $\mathcal{P}^{\mathrm{exp}}(\mathbf{x})$ at iteration 33. Case of $\sigma_\epsilon = 0.027$.

sample points, is now clear, and the magnitude of the fluctuations quickly decreases and saturates at a low level with small amplitude $\approx \sigma_\epsilon$. Experimental imperfections can be reasonably deemed responsible for the remaining fluctuations, rather than an ineffective selection strategy of the successive optima.

The two experiments reported in this section have underlined the importance of properly treating the measurement noise. For the settings of the experimental test case, not accounting for the noise results in a complete failure of optimization procedure. On the contrary, when the experimental noise is accounted for, the GP-based optimization is robust, owing to a suitable AEI merit function able to disregard non-interesting areas of Ω, even when they support large prediction variance. Further, the GP model with correct noise level is able to smooth out the measurement noise, allowing to find the optimum in few iterations only.
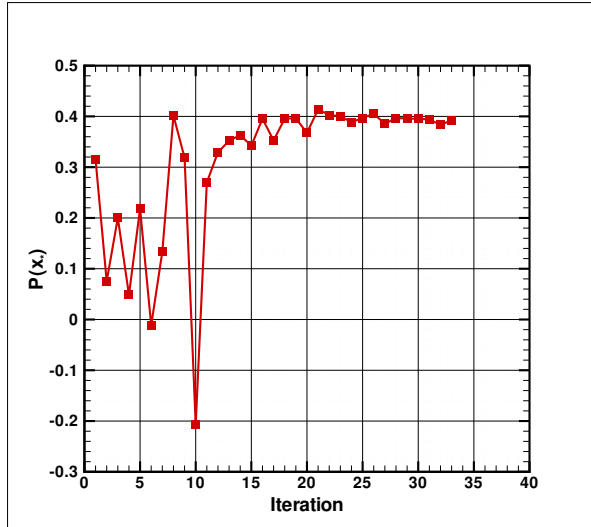
12

Figure 7: Sequence of measured $\mathcal{P}^{\mathrm{exp}}(\mathbf{x}_n)$ as a function of the iteration index. Case of $\sigma_\epsilon = 0.027$.

4. Numerical Optimization

The results reported in the previous section demonstrated the effectiveness of the GP-Based optimization procedure in determining the optimal trimming parameters in just a limited number of iterations, provided the experimental noise is appropriately accounted for. In this section, the same optimization procedure is used, but with the performance at trimming points estimated numerically. Our objective is to assess the capabilities of the optimization method applied to a coupled FSI solver, and compare the computed optimum with the experimental one. To this end, a numerical model of the wind tunnel experiment has been set up.

4.1. FSI Solver

We assume that the FSI problem has a steady solution for all values of the trimming parameters in the optimization range, such that a quasi-steady approach can be used to couple the structural and fluid solvers. These solvers are now briefly described.

4.1.1. Structural solver

We rely on the ARA software developed by K-EPSILON for the structural model of the sail. The sail model involves Timoshenko beams (mast, boom, battens), cables (sheet) and Constant Strain Triangles (CST) membrane elements (cloth) of various types in large displacement formulation [43]. A wrinkle model is added to capture the local sail deformations with a reasonable number of membrane elements [44]. See Durand [45] or Nakashino and Natori [46] for further details on CST elements and the wrinkle model. The nonlinear solution is obtained by means of a Newton method with Aitken relaxation. The geometry of the model is imported from the SAILPACK software developed by BSG Developments, to best fit the experimental geometry. The mechanical characteristics of the mast, boom, and sail fabrics were measured during the experimental campaign and used in the numerical model. The number of membrane elements is a key parameter to control the convergence of the sail geometry and of the whole FSI solution. Figure 8 shows 3 meshes of the sail with different number of membrane elements. The first mesh with ≈ 100 elements in Figure 8(a) is quite coarse and is compared with an intermediate and fine meshes having $\approx 2,700$ and $\approx 9,000$ elements respectively (shown in Figure 8(b) and Figure 8(c)). Solving the complete FSI problem with these 3 meshes, we found that the coarse mesh yield a solution far from being converged. On the contrary the two most refined meshes yield fluid loads in excellent agreement with less than 0.5% discrepancies. To minimize the computational cost, and given the other sources of error, the structural mesh with $\approx 2,700$ membrane elements in Figure 8(b) is used in the following.

13

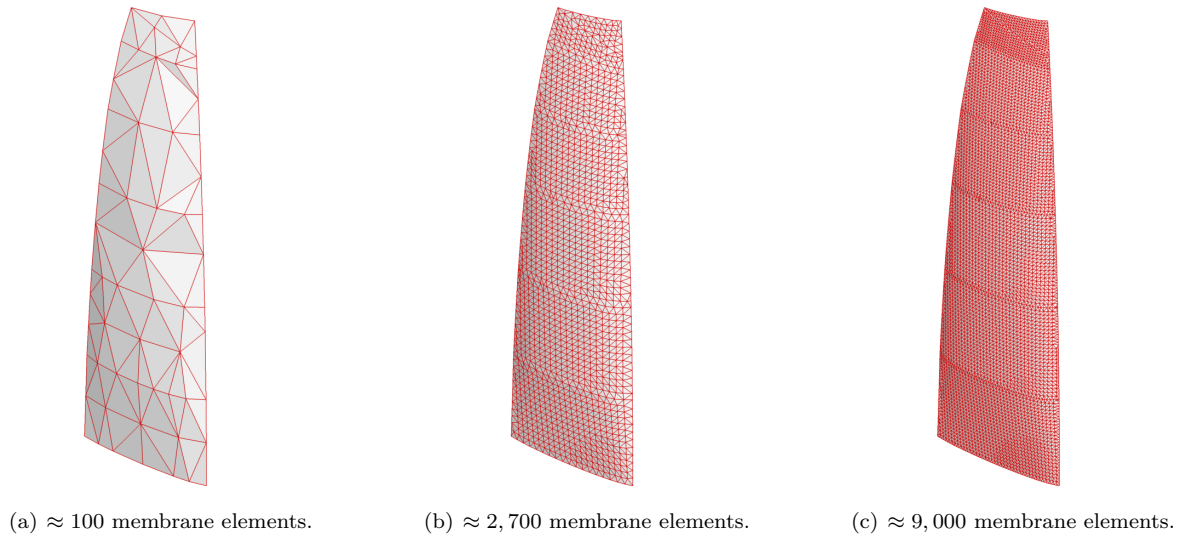(a) ≈ 100 membrane elements.   (b) ≈ 2,700 membrane elements.   (c) ≈ 9,000 membrane elements.

Figure 8: Numerical meshes with different number of membrane elements.

The resulting structural model is illustrated in Figure 9, where plotted are the membrane elements in Figure 9(a) and the fabric stiffness contours in Figure 9(b). The higher fabric stiffness at the clew, head and tack of the sail is clearly visible.
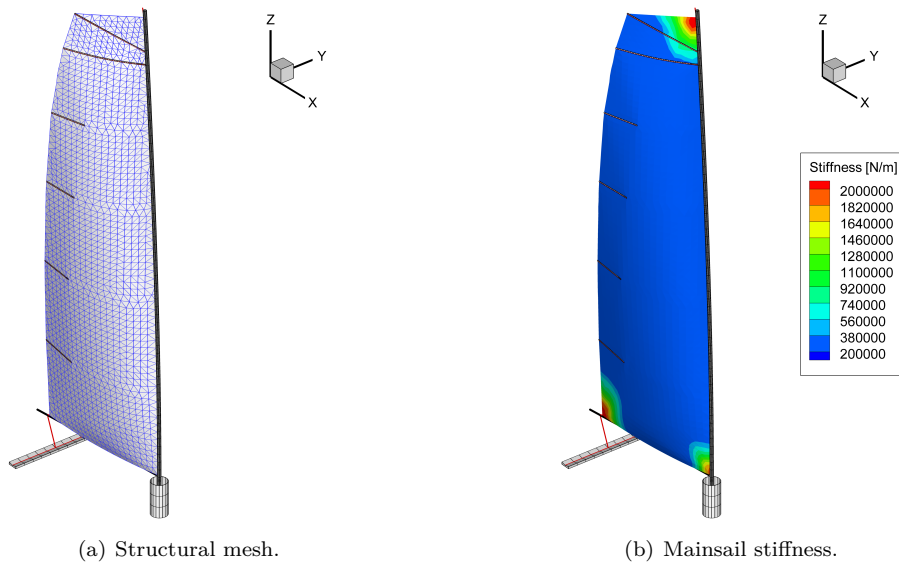

(a) Structural mesh.   (b) Mainsail stiffness.

Figure 9: Numerical model of the sail.

4.1.2. Flow solvers

The structural solver has been coupled with two different flow solvers.

First, it is coupled with the steady and unsteady inviscid flow solver AVANTI [47, 48]. The inviscid solver is based on the Vortex Lattice Method (VLM) [44] where the sail surface is discretized in rectangular panels supporting doublets distributions. The Kutta condition is enforced at the trailing edge to generate vorticity

particles that model the wake and its nonlinear dynamics. This inviscid model assumes slip conditions over all the sail surface and therefore cannot account for detached flow situations. Its domain of validity is then limited to low-camber sails trimmings and a moderate angle of attacks. The steady inviscid solution is identified with the limit as $t \to \infty$ of the unsteady solutions.

The structural solver is also coupled with the parallel (MPI) incompressible viscous flow solver ISIS-CFD (of FINE$^{\text{TM}}$/Marine). This software solves the Unsteady Reynolds-Average Navier-Stokes Equations (URANS) with finite volume discretizations, accommodating both structured and unstructured meshes. We considered a parallelepiped computational domain enclosing the sail, with extension $7.5 \times 12 \times 1.8h^3$ (in the X, Y and Z directions), following the numerical tests in [49]. The boundary conditions applied on the different faces of the computational domain are schematically illustrated in Figure 10(a). The velocity profile at the inlet boundary is set to the profile measured experimentally in the wind-tunnel, which is shown in Figure 10(b). The fluid domain is meshed to fit the sail geometries using HEXPRESS$^{\text{TM}}$, a semi-automated mesh generator. Note that the mast is not represented in the fluid model. Regarding the turbulence modeling, we used the two equations shear stress transport (SST) $k - \omega$ model [50] with wall function boundary conditions [51].
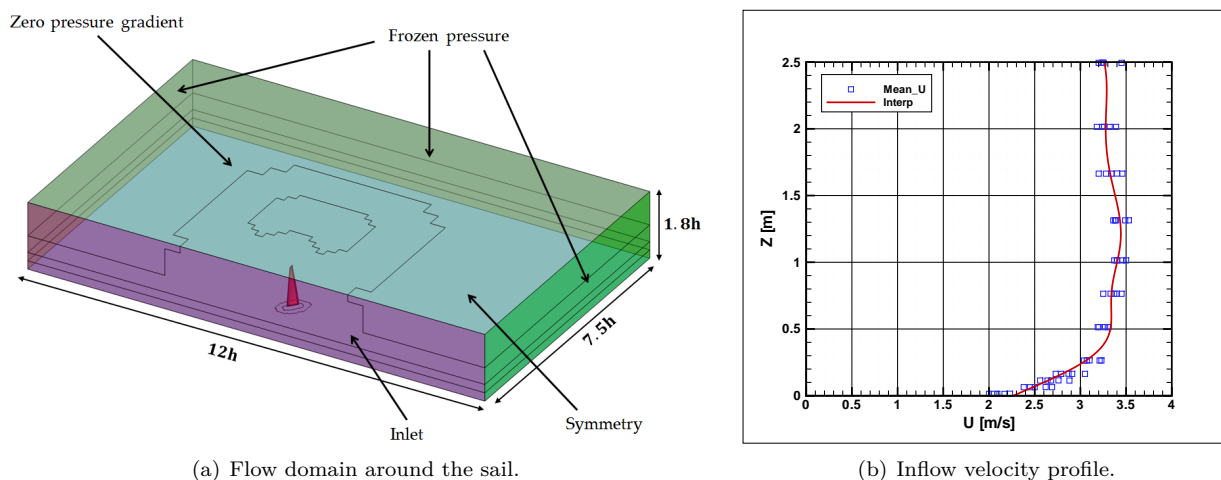

(a) Flow domain around the sail.

(b) Inflow velocity profile.

Figure 10: Computational fluid domain and boundary conditions.

4.1.3. Coupled problem

The FSI problems are solved coupling the structural and flow solvers through a quasi-monolithic algorithm [45], which is an implicit coupling procedure adapted to partitioned solvers. Briefly, the resolution of the structural problem is nested inside the iterations on the nonlinear steady flow solution. This approach preserves the convergence and stability properties of the monolithic approach. As the sail sustains large deformations, the flow solver must account for the change of the boundary geometry. This is easily achieved in the inviscid solver where only the sail surface is discretized. This is less trivial and computationally more demanding in the case of the viscous solver, which requires a volume mesh. In this work we use the mesh deformation propagation (MDP) method [52], proceeding from the sail boundary toward the inside of the fluid domain. More details on the solvers and the coupling algorithms can be found in [45, 48, 47].

4.2. Numerical Validation

The inviscid and viscous FSI solvers have been extensively validated at the Naval Academy Research Institute with many experimental comparisons, see for instance Augier et al. [5]. A specific convergence analysis was conducted on the present problem to determine the discretization parameters of the fluid and structural problems needed to correctly capture the physics of the FSI models, while maintaining

an affordable computational cost. Achieving minimal computational time is indeed crucial to enable the optimization of the trimming parameters. A comparison with experimental measurements is also provided to assess the predictive capability of the (viscous flow) FSI solver.

4.2.1. Influence of mesh deformations

In these numerical tests, we analyze the impact of the fluid mesh generation on the computed FSI solutions. More precisely, we compare two solutions for trimming parameters $(L_{\text{sheet}}, L_{\text{car}}) = (123, 124)$ mm. The first one uses the MDP method [52], starting from the solution mesh at trimming parameters $(L_{\text{sheet}}, L_{\text{car}}) = (160, 0)$ mm, called S_0. This first solution is called S_{Def}; it requires a large deformation of the initial geometry of S_0 to adapt the flying shape at the new trimming parameters. The difference between the two flying shapes can be appreciated from Figure 11(a), and the mesh deformation can be assessed comparing Figures 11(b) and 11(c). Subsequently, we build a new mesh around the geometry of S_{Def} with ≈ 1.5 million finite volumes, and restart the FSI solver to accommodate changes in the fluid loads induced by the new mesh. This leads to a second solution over a much less deformed mesh, as can be seen in Figure 11(d); we call this solution S_{Remesh}. We only report the differences between S_{Def} and S_{Remesh} in terms of fluid loads as the two flying shapes are almost indistinguishable. Values of C_X, C_Y and the sheet tensions T_{sheet} are reported in Table 1 for the two solutions. It is seen that they agree very well with less than 1% differences in the fluid loads, demonstrating the robustness of the MDP method that remains accurate even for these large changes in the sail geometry and mesh deformation. We shall exploit this robustness of the MDP method and avoid systematically iterating on the initial mesh at every new optimal point proposed when optimizing for the trimming parameters. Instead we shall perform a remeshing of the fluid domain only in the case of extreme mesh deformations.

	C_X	C_Y	T_{sheet} $[N]$
S_{Def}	0.476	-0.841	12.6
S_{Remesh}	0.477	-0.847	12.5
Differences	0.3%	0.7%	0.8%

Table 1: Comparison of the computed fluid loads for the Mesh Deformation Propagation method (S_{Def}) and the corresponding remeshed solution (S_{Remesh}).

4.2.2. Experimental comparison

The predictive capability of the viscous numerical FSI solver is also assessed by comparison with experimental measurements. In Table 2 we compare to the experimental values the computed fluid loads for 3 different meshes of the fluid domain having increasing resolution. The computations correspond to the experimental optimal trimming points $(L_{\text{sheet}}, L_{\text{car}}) = (133, 138)$ mm. The three meshes, M_0, M_1 and M_2 have ≈ 1.5, ≈ 3.4 and ≈ 4.3 million cells respectively. They were generated with the remeshing procedure discussed above. We can see that the numerical loads are close to each other, denoting again the convergence of the computation with respect to the fluid mesh size. The thrust force coefficient C_X and sheet tension T_{sheet} are also in excellent agreement with their experimental values. The magnitude of the side force coefficient C_Y, on the contrary, appears to be numerically under-estimated by roughly 15%. This discrepancy level is not surprising and can be explained by both the experimental uncertainties and the numerical modeling errors (see below). Consequently, the fluid mesh with around 1.5 million cells is selected in the following as it appears to provide sufficiently converged numerical fluid loads, given the discrepancies with the experimental loads. Specifically, our numerical tests show that the discretization error can not explain the discrepancy between numerical and experimental estimates. We shall discuss and investigate further the causes of the remaining discrepancies, namely the model errors, later in in Section 5.

The comparison with the experiment measurements can also be made for the geometries (flying shapes). For instance, Figure 12(a) and Figure 12(b) report the evolutions of the sail's twist angle and camber as functions of the height Z above the bottom boundary (the bottom of the sail is at $Z = 0.2$ m). These two plots stress the convergence of the numerical prediction with respect to the mesh size, as no clear differences
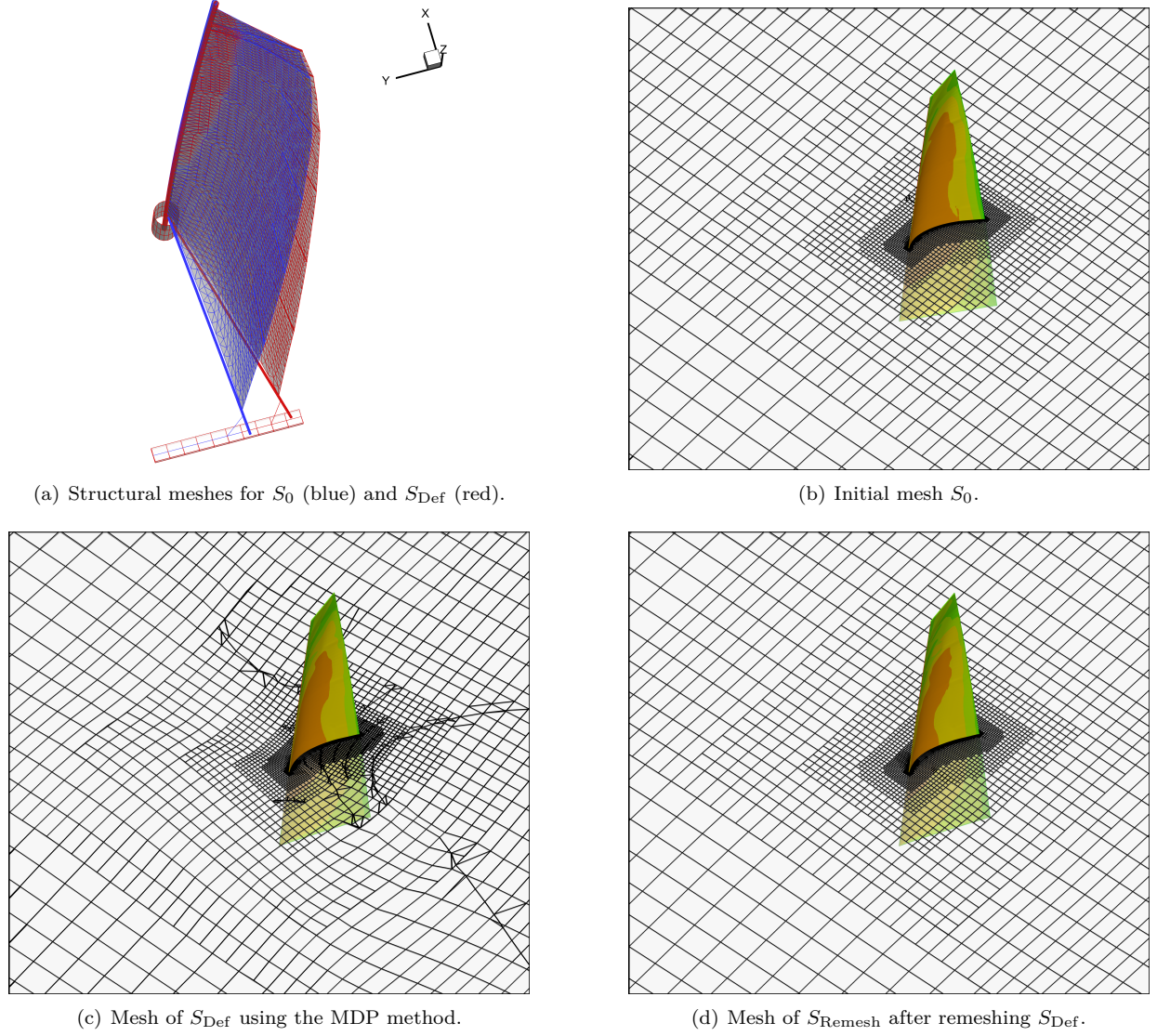
16

(a) Structural meshes for S_0 (blue) and S_Def (red).

(b) Initial mesh S_0.

(c) Mesh of S_Def using the MDP method.

(d) Mesh of S_Remesh after remeshing S_Def.

Figure 11: Structural and fluid meshes for different trimming parameters and mesh generation methods (see text for more details).

	C_X	C_Y	T_sheet $[N]$	N_cell
Experimental	0.497	-1.026	14.9	–
M_0	0.476	-0.844	14.5	1.5 M
M_1	0.467	-0.831	14.3	3.4 M
M_2	0.483	-0.849	14.6	4.3 M

Table 2: Comparison of the fluid loads computed using 3 different meshes with increasing resolution and their corresponding experimental values. Computational case corresponding to the experimental optimal trimming point.

between the 3 meshes are reported. The twist distribution is also in excellent agreement with the measured one. On the contrary, differences between numerical and experimental values are not negligible for the camber distribution, with a maximum difference as large as $\approx 4.5\%$ of the chord length around 1.5 m height. Specifically, the camber of the sail appears to be numerically overestimated in the middle part of the sail.

17

Regarding the top experimental point, we remark that it is the most sensitive to measurement errors, and the experimental camber is likely to be overestimated at this point. We however report this point for the completeness of the comparison.

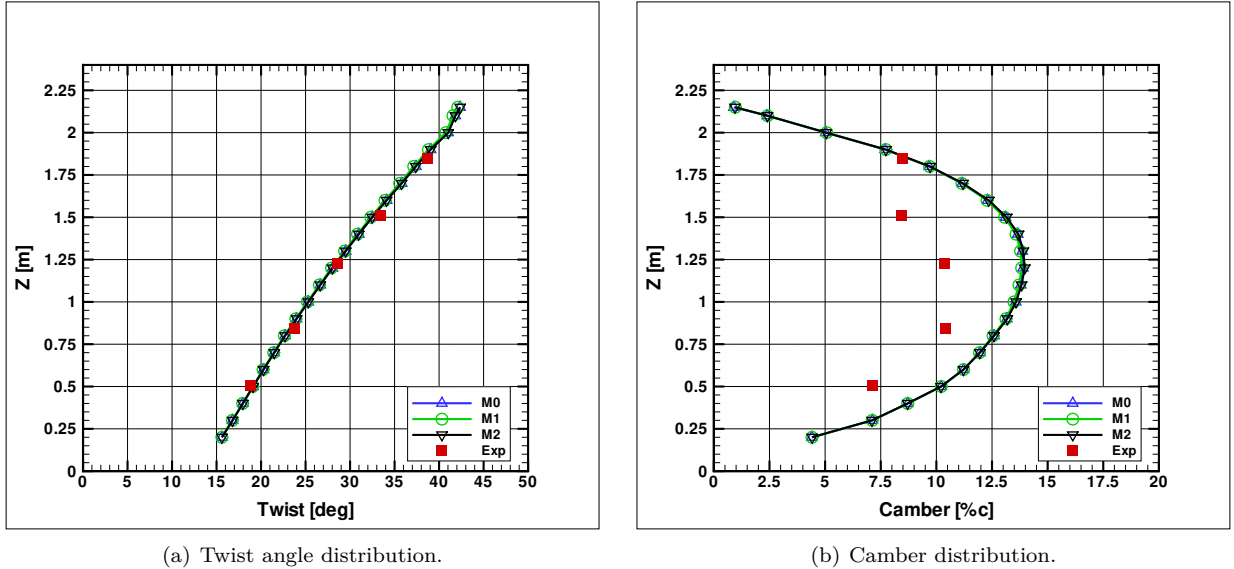
(a) Twist angle distribution.

(b) Camber distribution.

Figure 12: Comparison of the experimental and numerical twist and camber distributions at the optimal experimental trimming points. Three numerical meshes $M_{0,1,2}$ are tested with size reported in Table 2.

To better appreciate the differences between the numerical and experimental results, we provide in Figure 13 a visual comparison of the two flying shapes. In the experimental case shown in Figure 13(a), a diagonal wrinkle of the sail is clearly visible. This is in contrast to the corresponding numerical flying shape, shown in 13(b), which is smooth and presents no such wrinkle. The wrinkle can explain the differences in the camber distribution (Figure 12(b)), as wrinkles generally tend to reduce the sail camber. The absence of wrinkle in the numerical solution could be due to an incorrect prescription of the tension in the two top battens (horizontal beams) of the sail, which is difficult to measure, and the inability of the numerical model to capture these types of wrinkles. Another important possible source of discrepancies between the flying shapes is the boundary conditions and confinement effects, which are known to have a significant impact on the computed aerodynamic forces [49].

Based on this numerical experiment, we use in the following $\approx 2,700$ membrane elements for the structural mesh of the sail, and meshing parameters such that the fluid mesh size has ≈ 1.5 million finite volumes. With these discretization parameters, well converged viscous FSI solutions are obtained in a reasonable time of ≈ 5 h on a 64 CPUs cluster per computation.

4.3. Numerical Optimization

The results of the numerical sail trimming optimization are now presented. The optimization considers again the performance function defined in (25). The optimization procedure is the same as for the experimental case, except that $\mathcal{P}(\mathbf{x})$ is numerically evaluated, solving the FSI problem at the proposed trimming points $\mathbf{x}_i$. We shall denote $\mathcal{P}^{\mathrm{num}}(\mathbf{x})$ the numerical performance. In addition, the "noise" parameter σ_ϵ of the Gaussian process is not fixed to an a priori value, as in Section 3, but is inferred from the observations (see (20)).

4.3.1. Inviscid flow solver

The optimization is first applied on the structural solver coupled to the inviscid flow solver. The advantages of the inviscid model are its low computation time, compared to the viscous solver, and the simplicity

(a) Experimental flying shape.
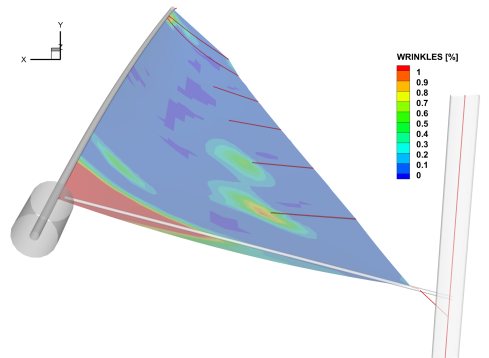
(b) Numerical flying shape. The color contours indicate the wrinkles rate.

Figure 13: Comparison of the experimental and numerical (viscous model) flying shapes at trimming parameters $L_{\text{sheet}} = 133$ and $L_{\text{car}} = 138$.

of dealing with the sail deformations (no volume mesh needs be generated).

Figure 14 shows the predictive mean $\widehat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ at convergence, which is achieved at iterations 38 (it must be remember that the first 10 iterates are selected randomly and are not actual iterations of the procedure). This predictive mean should be compared with its experimental counter-part depicted in Figure 6(a). It can be seen that the inviscid and experimental predictive mean differ qualitatively, both in their magnitudes (note the change in the color range) and in the location of their respective optimum. Specifically, though $\widehat{\mathcal{P}}_n^{\text{exp}}(\mathbf{x})$ and $\widehat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ are both smooth with a unique maximum, the inviscid optimum corresponds to a trimming with L_{car} fully trimmed to the windward side and L_{sheet} quite eased, while the car is trimmed to the leeward side and the sheet is tightened in the experimental optimum. The optimal inviscid trimming corresponds to an unrealistically cambered sail. Such trimming is favored because the inviscid model cannot account for detached flow situations, yielding non-physical solutions with large overestimation of the performance in this regime. This example illustrates the importance of using a physically realistic predictive model that remains valid over the whole parameter domain. Otherwise, the numerical optimization procedure can produce an unphysical solution, even if the model is actually fairly accurate in the neighborhood of the physical optimum. In the present case, the problem could be remedied by means of an ad hoc penalization of the performance function; for instance one could add a righting moment penalty term or artificially decrease the performance for trimmings likely to produce a detached flow [53].

4.3.2. Viscous flow solver

The optimization is now applied coupling the structural solver with the viscous solver. For reasons that will be explained soon, the parameter domain is extended to $L_{\text{car}} \in [-100, 310]$ (mm), while the range of L_{sheet} is kept as before. However, the optimization procedure is still initialized with the same LHS points as previously.

Figure 15 shows the first two moments of the GP model of $\mathcal{P}^{\text{num}}(\mathbf{x})$ at the convergence of the optimization procedure, which is achieved at iteration 34. Note that the number of iterations to converge is very close to the experimental case (with noise). The predictive mean $\widehat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ reported in Figure 15(a) is again smooth with a unique optimum. Further, for $L_{\text{car}} \in [-210, 150]$ mm $\widehat{\mathcal{P}}_n^{\text{num}}$ and $\mathcal{P}^{\text{exp}}$ are in good agreement (see Figure 6(a)). However, the hill of $\widehat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ containing the numerical optimum has a larger extension and is flatter compared to the experimental case. Consequently, the sequence of proposed numerical optima converges to a value of L_{car} larger than for the experimental case, in fact slightly outside of the range of the experimental test. This was the motivation for extending the L_{car} range: to encompass the numerical optimum. The standard deviation of the GP model of $\mathcal{P}^{\text{num}}(\mathbf{x})$, shown in Figure 15(b), has a structure comparable in shape and magnitude to the experimental one in Figure 6(b), at least in the areas that were
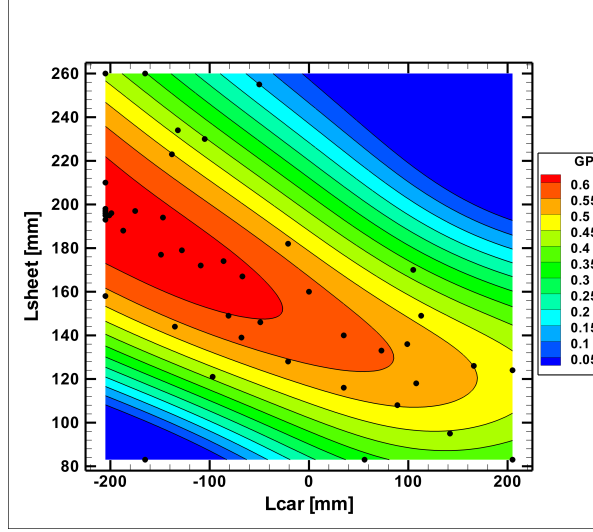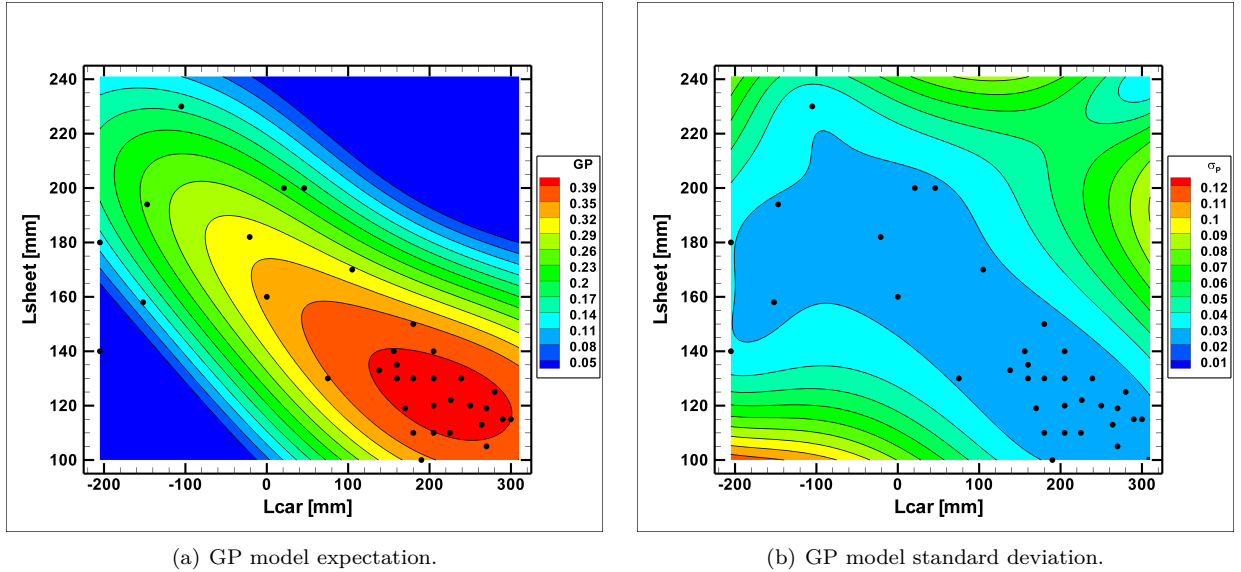
19

Figure 14: Predictive mean $\widehat{\mathcal{P}}_n^{\mathrm{num}}(\mathbf{x})$ at iteration 38 (inviscid model, $\sigma_\epsilon = 0.006$).

jointly sampled areas. Also, the inferred $\sigma_\epsilon = 0.022$ is not far from the experimental value, even though they have very different origins and account for totally different effects (experimental variability and numerical errors).

(a) GP model expectation.

(b) GP model standard deviation.

Figure 15: GP model of $\mathcal{P}^{\mathrm{num}}(\mathbf{x})$ at iteration 34 (viscous model, $\sigma_\epsilon = 0.022$).

The convergence of the viscous optimization procedure is reported in Figure 16, where the computed $\mathcal{P}^{\mathrm{num}}(\mathbf{x}_n)$ is plotted as a function of the iteration index. Again, the first ten values correspond to the initial sampling of the parameter domain. After these first points, the computed performance improves and eventually stabilizes with slowly decaying fluctuations after iteration 20. This behavior is similar to the experimental results reported in Figure 7. Figure 16 also provides the evolution with the iteration index of the inferred σ_ϵ. It is seen that σ_ϵ decreases (on average) as the optimization proceeds and asymptotically converges to a finite value $\sigma_\epsilon \approx 0.022$ (read on the right-hand axis). The fact that σ_ϵ does not go to

zero asymptotically indicates that the GP construction is informed of the numerical errors and does not attempt to fit exactly the computed performances. In other words, as for the experimental case, the GP-based optimization is able to smooth out part of the numerical errors inherent, for example, to the mesh deformations, the approximation of steady states or the convergence of nonlinear solvers.
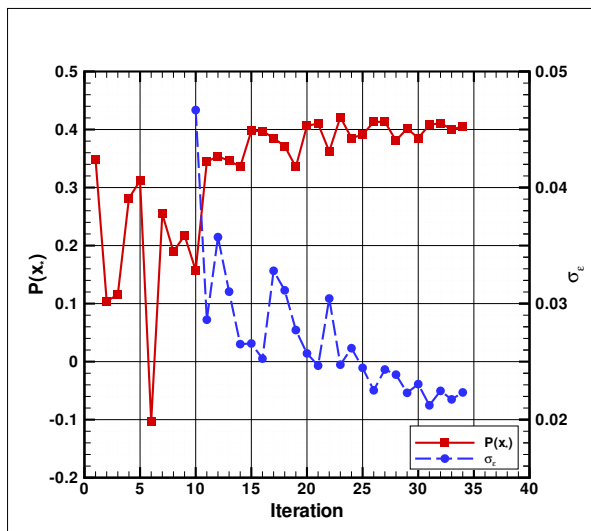

Figure 16: Sequence of computed $\mathcal{P}^{\mathrm{num}}(\mathbf{x}_n)$ as a function of the iteration index (in red, left axis), and corresponding inferred σ_ϵ (in blue, right axis). Case of the viscous model.

To complete the comparison of the experimental and numerical solutions, we report in Table 3 the optimal trimming parameters and the corresponding optima of $\widehat{\mathcal{P}}_n^{\mathrm{num}}$ and $\widehat{\mathcal{P}}_n^{\mathrm{exp}}$ at convergence. We observe that the experimental and numerical solution are significantly different in terms of optimum locations, particularly for L_{car}: 138 mm in the experimental case and 226 mm in the numerical case. In contrast, the predicted performances in the fourth column of Table 3 are in much better agreement. The flatness of the performance around the experimental and numerical optima can explain the resulting large distance in terms of parameters: changing L_{car} around its optimal value weakens the numerical performance and, conversely, small numerical and modeling errors can drastically affect the optimal value of L_{car}. This claim is further confirmed by inspecting the fluid loads reported in the last column of Table 3. The loads reveal that the numerical and experimental thrust coefficients are in exact agreement, such that the side force coefficient explains all the difference in the performance. As mentioned previously, this is not surprising as the side force is known to be more sensitive to modeling errors than the thrust force. Finally, the last column of Table 3 compares the tensions in the sheet for the two optima, which significantly differ from one would have expected from the differences in the car trimming. In the next section we further refine the analysis of the different source of numerical and modeling errors and their impact on the predicted optimal trimmings.

	L_{sheet} [mm]	L_{car} [mm]	$\widehat{\mathcal{P}}_n(\mathbf{x}_{\mathrm{opt}})$	C_X	$-C_Y$	T_{sheet} [N]
Exp.	133	138	0.397 ± 0.027	0.497 ± 0.012	1.026 ± 0.015	14.9 ± 0.02
Num.	122	226	0.413 ± 0.024	0.497	0.803	20.1

Table 3: Comparison of the experimental and numerical optima: location, predicted optimal performance (see (21)), thrust and side force coefficients, and sheet tension.

5. Optimum Uncertainty Quantification

The comparison of the experimental and numerical optimizations has shown differences in the optimal trimming parameters that are consistent with the discrepancies between the measured and computed fluid

forces. In addition, the numerical discretization parameters of the FSI solvers were selected as to ensure a sufficiently small numerical error, such that it can only be responsible for a limited contribution to the discrepancies between numerical and experimental optima. In fact, model errors and experimental imperfections are needed to explain most of the differences reported above. In a effort toward investigating model errors, we follow an uncertainty quantification approach in this Section, we propagate some uncertainties in the inflow conditions through the numerical optimization problem. Specifically, we are concerned with the impact of the model uncertainties caused by an incomplete knowledge of the actual wind-tunnel experiment, on the outcome of the optimization procedure. Two points of view will be considered. First, in Section 5.1, we analyze the impact of some model uncertainties on the predicted performance at the numerical optimal trimming determined in the previous section. Second, in Section 5.2, we directly estimate the uncertainty on the numerical optimal trimming due to the model uncertainties.

To perform these uncertainty analyses, we need to select the most relevant experimental uncertainties and define their variability range. To this end, we rely on expertise (from the staff of the Auckland wind-tunnel as well as our own appreciation) to review the different sources of experimental uncertainties likely to have a noticeable impact on the solution of the numerical model and on the optimization problem. The main sources identified were the measured apparent wind angle (AWA), apparent wind speed (AWS) and mechanical properties of the sail and spar elements. Regarding the mechanical properties of the mast and boom, as well as the stiffnesses of the sail fabrics, measured before and after the experimental tests were found to be quite accurate and also to have a small impact on the numerical solution. Assuming an overly large 2% range of variation for these parameters, no significant effect was observed in the predicted performance with consequently a negligible impact on the optimal solution. The tensions in the 2 top battens of the sail were identified as another potential source of uncertainties, that could significantly affect the numerical prediction, in particular through the direct modification of the sail camber and the creation-resorbing of wrinkles. Unfortunately, the batten tensions were not systematically measured during the experimental campaign, such that determining a realistic variability range is not possible and we have preferred to disregard these model uncertainties in the present work. We are then left with the two main characteristics of the wind-tunnel inflow: the AWA and AWS. Based on experimental measurements and expertise, we have set the uncertainty ranges at 40 ± 2 degrees for the AWA and at 3.5 ± 0.175 m/s for the AWS. We further assume uniform distributions for the AWA and AWS which are considered independent, that is the least informative distribution given the variability range.

5.1. Uncertainty in the Optimal Performance

In this section we quantify the uncertainty, due to the uncertain AWA and AWS, in the numerical performance prediction at the optimal trimming point approximated by $\hat{\mathbf{x}}_n$ given by (21), that is $(L_{\text{sheet}}, L_{\text{car}}) = (122, 226)$ mm. To this end, $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$ is seen as a function of AWA and AWS which approximated using a Polynomial Chaos expansion [54, 22]. The expansion coefficients are obtained using the non-intrusive Sparse Pseudo Spectral Projection (SPSP) method [55, 56, 57] which requires the resolution of the FSI model for selected AWA and AWS values located on a sparse grid. An isotropic PC expansion with maximal degree 4 was found necessary to accurately approximate the dependences, corresponding to a sparse grid having a total number of 17 grid points.

Figure 17(a) shows the PC approximation of $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$ as a function of AWA and AWS in the uncertainty range. The sparse grid points involved in the PC construction are also shown using black squares. The performance is seen to have a low dependence on the wind speed AWS over the considered uncertainty range, especially at the lowest wind angle. The surface plot also indicated that the performance monotonically depends on the AWA, and this dependence is stronger at the highest wind speed.

The influence of AWA and AWS can be quantified using global sensitivity indices, also known as the Sobol indices [58, 59], which can be easily computed from the PC approximation [60, 61]. First and total sensitivity indices of the predicted performance are reported in Table 4. The first order indices measure the fractions of the variance due to the respective factor (AWA or AWS) without any interaction with other factors. On the contrary, the total indices account for the effect of the factor and all its interactions with other factor. The table confirms the low influence of AWS compared to AWA, with a total sensitivity index barely exceeding 2%. Conversely, the AWA variability is responsible for 98% of the optimal prediction

variability, and the two factors are weakly interacting. The low influence of the wind speed can be explained by the normalized form of the performance and the relative stiffness of the sail system that does not undergo significant elastic deformations over the AWS uncertainty range.

	AWA	AWS
First order	0.9776	0.016
Total order	0.9838	0.022

Table 4: First and total order sensitivity indices of $\mathcal{P}^{\mathrm{num}}(\hat{\mathbf{x}}_n)$, associated to AWA and AWS.

The PC approximation can be averaged over the uncertainty domain to give $\overline{\mathcal{P}}^{\mathrm{num}}(\hat{\mathbf{x}}_n) = 0.413$; This averaged performance agrees up to the third significant digit with the nominal predicted performance $\widehat{\mathcal{P}}_n^{\mathrm{num}}(\hat{\mathbf{x}}_n)$ (that is for AWA= 40 degrees and AWS= 3.5 m/s). Then similarly, the standard deviation over the uncertainty range can be estimated from the PC expansion; we find $\sigma_{\mathcal{P}} = 0.014$, corresponding to a low coefficient of variation (ratio of mean and standard deviation) less than 3.5%. Having characterized the uncertainty in the performance at the nominal optimum point $\hat{\mathbf{x}}_n$, we would like to deduce an uncertainty in the optimum itself. In Section 5.2 we shall consider the resolution of all the optimization problems with uncertain AWA and AWS, to assess the uncertainty in $\hat{\mathbf{x}}_n$. Here, we adopt a simpler characterization based on the distance to the nominal optimal performance $\mathcal{P}^{\mathrm{num}}(\hat{\mathbf{x}}_n)$; we define the family of uncertainty sets $X(\delta) \subset \Omega$, parametrized by $\delta \geq 0$, as follows:

$$X(\delta) = \{\mathbf{x} \in \Omega; \ \mathcal{P}^{\mathrm{num}}(\hat{\mathbf{x}}_n) - \mathcal{P}^{\mathrm{num}}(\mathbf{x}) < \delta\}.$$

In words, $X(\delta)$ is the set of trimmings points whose nominal performance is in a distance less than δ to the optimal nominal performance. In particular, $X(\sigma_{\mathcal{P}})$ is the set of trimmings points that are within a distance less than one AWA and AWS-induced standard deviation. To compute this set, we substitute the numerical performance $\mathcal{P}^{\mathrm{num}}(\mathbf{x})$ with $\widehat{\mathcal{P}}_n^{\mathrm{num}}(\mathbf{x})$. The boundary of the set $X(\sigma_{\mathcal{P}})$ is plotted in Figure 17(b) using a dotted line. It is superposed to the color contours of the experimental performance $\widehat{\mathcal{P}}_n^{\mathrm{exp}}(\mathbf{x})$. We see that the experimental optimum is not far from being included in $X(\sigma_{\mathcal{P}})$. Using $\delta = k\sigma_{\mathcal{P}}$ with $k > 1$ in the definition of $X(\delta)$ defines larger and larger sets of parameters. However, k can be made arbitrarily large, in the present case, because of the finite variability of the performance caused by the AWA and AWS. Instead we consider for largest set $X(\delta_{\max})$ the set defined by the upperbound distance $\delta_{\max} = \overline{\mathcal{P}}^{\mathrm{num}}(\hat{\mathbf{x}}_n) - \mathcal{P}_{\min}^{\mathrm{num}}$ where $\mathcal{P}_{\min}^{\mathrm{num}}$ is the minimal performance over the AWA and AWS uncertainty range. The boundary of $X(\delta_{\max}) \supset X(\sigma_{\mathcal{P}})$ is also shown in Figure 17(b) using a solid line. It is seen that the set $X(\delta_{\max})$ contains the experimental optimum. This result suggests that the discrepancy between the experimental and numerical optima can be partly explained by the experimental and model uncertainties, though other sources of modeling error not considered in this work could also have a significant impact.

5.2. Uncertainty on Optimal Trimming

The analyses presented in Section 5.1 demonstrate the effect of uncertainties in the AWA and AWS on the performances. In this section, we directly examine the uncertainty in the optimal trimming and its robustness. Following the previous findings on the dominant influence of the AWA relative to the AWS, we restrict ourselves to this unique uncertainty source in the following, keeping the same uncertainty range as before. Our objective is to characterize the dependence with the AWA of the numerical optimum, that is to determine $\hat{\mathbf{x}}_n(\mathrm{AWA})$ and subsequently propose a robust optimum. To this end, we start by generating a PC expansion to approximate the dependences of the numerical performance with respect to the trimming parameters and the AWA. We again rely on the SPSP method assuming a uniform distribution of L_{car}, L_{sheet} and AWA in the domain $[100, 150] \times [100, 300] \times [38, 42]$. Note that the trimming domain has been restricted to a small portion of Ω around the nominal numerical optimum. The isotropic sparse grid used to construct the PC expansion of the performance is shown in Figure 18(a); it has 31 points, requiring as many resolutions of the FSI problem, and leading to a polynomial approximation with maximum partial degree of 3. The resulting PC approximation is illustrated in Figure 18(b), where shown are the colour contours of the
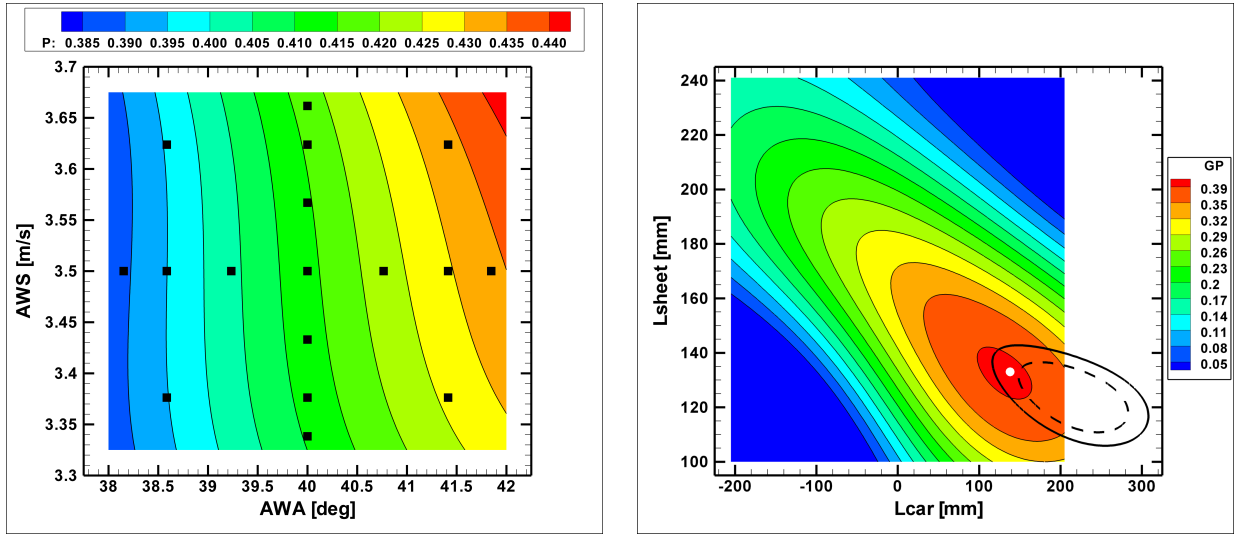
(a) Color contours of the PC approximation of the dependences of $\mathcal{P}^{\mathrm{num}}(\hat{\mathbf{x}}_n)$ with the AWA and AWS. The dark points indicate the sparse grid points supporting the PC approximation.

(b) Color contours of $\widehat{\mathcal{P}}_n^{\mathrm{exp}}(\mathbf{x})$ superposed with the boundaries of the uncertainty sets $X(\sigma_{\mathcal{P}})$ (dashed line) and $X(\delta_{\max})$ (solid line) and experimental optimum (white point).

Figure 17: Effects of uncertainties in the AWA and AWS.

numerical performance in the $(L_{\mathrm{car}}, L_{\mathrm{sheet}})$ domain for three values of AWA $= 38.5, 40.0$ and 41.5 deg. The contours for AWA $= 40$ deg should be compared with the predicted mean $\widehat{\mathcal{P}}_n^{\mathrm{opt}}$ of the GP model shown in Figure 15(a). Differences between the PC and GP models are visible, with in particular the presence of a second local maximum along the domain boundary. However, increasing accuracy of the PC expansion would require significantly more FSI solves, and the PC approximation is believed accurate enough for the uncertainty characterization of the optimum.

For the characterization, we rely on the determination of the optimal trimming solution $\hat{\mathbf{x}}_n(\mathrm{AWA})$, for AWA $\in [38, 42]$. For a given value of the AWA, the optimum is determined by a simple Newton optimization method, exploiting the polynomial form of the PC approximation. The function $\hat{\mathbf{x}}_n(\mathrm{AWA})$ is shown as a red line in Figure 18(b). We can see that the optimal trimming corresponds to a car moved leeward and a sheet progressively eased as the AWA increases.

The uncertainty in the optimum $\hat{\mathbf{x}}_n$ can be appreciated from the values collected in Table 5, which reports the (AWA) averages and standard deviations of the trimming parameters and optimal performance. The AWA uncertainty is seen to have a greater influence on the optimal car trimming (13 mm standard deviation) than on the sheet trimming (1 mm standard deviation). Note that the averaged trimming point is close to the nominal numerical solution but remains quite different from the experimental one. In addition, the coefficient of variation of the optimal performance is seen to be slightly less than 6%, a significantly high value given the AWA range. This behavior is consistent with the findings of the previous of Section 5.1.

	L_{sheet} [mm]	L_{car} [mm]	$\mathcal{P}^{\mathrm{num}}(\hat{\mathbf{x}}_n)$ (PC approx.)
Mean	120	238	0.427
Standard deviation	1	13	0.025

Table 5: AWA averages and standard deviations of the numerical optimum trimmings and optimal performance. Quantities are based of the PC expansion of the numerical performance.

Finally, Table 6 presents the optimal trimming for the AWA-averaged numerical performance. It corresponds to the best (most robust) fixed trimming solution that one would select knowing that the AWA is uncertain, to yield the best performance on average. The standard deviation with respect to the AWA of
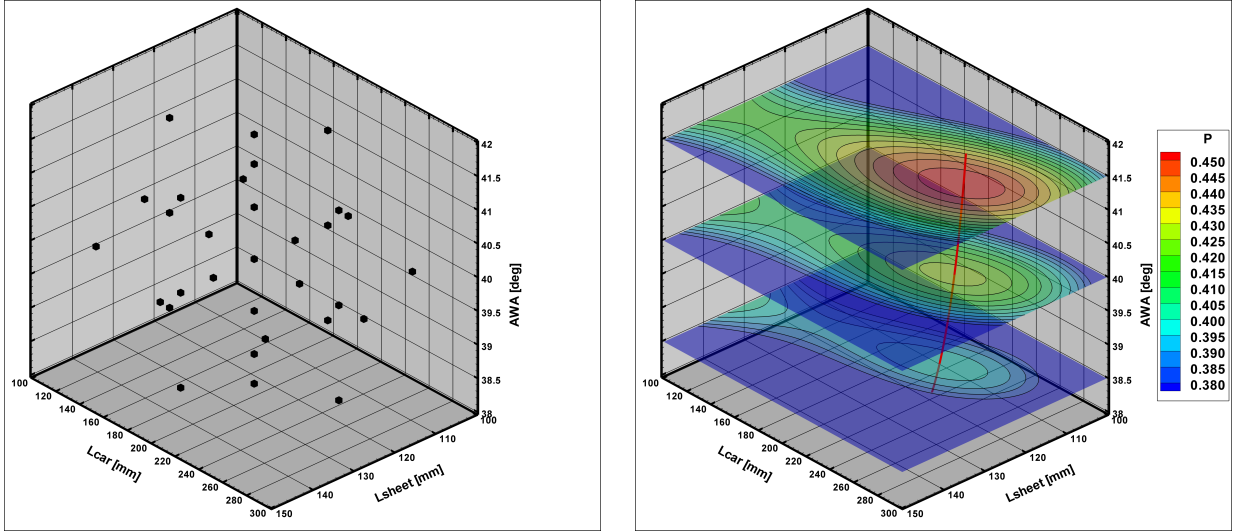
(a) Sparse grid for the PC approximation of the numerical performance in the L_sheet, L_car and AWA domain.

(b) Color contours of the PC approximation of the performance for 3 values of the AWA: 38.5, 40 and 41.5 degrees. Also shown is the optimal trimming as a function of the AWA (red line).

Figure 18: Effect of the AWA on the performance and optimal trimming.

this robust optimal solution is also provided. It is remarkable that the robust trimming parameters are close to the averaged trimming solution reported in Table 5 and to the optimal solution for the nominal AWA value. However, the robust and nominal trimmings yield different averaged performances, with an increase from 0.413, using the optimum for the nominal AWA, to 0.426 for the robust optimum solution.

L_sheet [mm]	L_car [mm]	Pred. $\mathcal{P}$
120	240	0.426 ± 0.026

Table 6: Robust optimal trimming based on (AWA)-averaged numerical performance. Also reported is the mean performance with $\pm$ standard deviation. Quantities are based on the PC expansion of the numerical.

6. Conclusion

We have proposed to use a Gaussian process model based optimization procedure to enable the optimization of trimming parameters in complex nonlinear sail systems. The approach was demonstrated on experimental optimization of a sail trimming in the Yacht Research Unit wind tunnel. These experiments have demonstrated the validity of the approach, which converges quickly on this problem, without exploring uninteresting trimming values. It was also shown that the GP model approach is very robust against experimental noise in the performance measurements.

To assess the validity and relevance of the experimental optimization, a comparison with a purely numerical optimization has been proposed. To this end, detailed numerical models of the wind tunnel experiment and sail have been established, considering the resolution of the nonlinear Fluid-Structure Interactions problems. An inviscid and a viscous flow solvers have been compared to highlight the importance of using a model physically correct over the whole optimization domain. The most complex, and physically correct, numerical model is based on a turbulent flow solver coupled with a nonlinear elastic solver, with a mesh deformation method to handle the changes in the sail geometries. This FSI solver can be considered representative of the state of the art for this kind of simulations. Discretization parameters were carefully selected

to ensure sufficient convergence while maintaining a computational cost low enough to enable the numerical optimization, and the FSI solution was validated using several experimental measurements.

Performing the optimization on the numerical model confirmed the effectiveness of the GP-based approach. The convergence on the numerical optimum was achieved in a number of iterations comparable to the experimental case. In fact, we have shown that, using a limited set of computations, the GP model is able to accurately reconstruct the computed performance. Comparing the experimental and numerical optima, we found that while the corresponding performances and fluid forces are in excellent agreement, the optimal trimming parameters differ significantly, in particular the trimming of the car which is set more leeward in the numerical case than in the experimental one. Several modeling imperfections can explain these differences and it can be reasonably claimed that they are consistent with the current predictive capabilities of state of the art FSI solvers. In any cases, the GP-based optimization is not responsible for these differences, since it does not introduce noticeable errors, and only a better numerical modeling of the experimental set-up would help reducing the observed discrepancies. Possible avenues in this direction have been discussed such as improved batten models, better boundary conditions for the flow, accounting for mast/flow interaction or more advanced turbulence model. We also focused on the influence of experimental uncertainties on the optimal performance and trimming, focusing principally on the apparent wind angle. We have shown that the uncertainty in the AWA induces an uncertainty in the performance which is consistent with the distance between the experimental and numerical optimal trimmings, and the important effect of the AWA on the trimming of the car. We finally were interested in considering uncertainties in the definition of the optimum and introducing robust optimization.

Though the results reported in the paper have demonstrated the efficiency of the procedure, in terms of number of performance estimations, the method remains to be tested on problems involving larger sets of trimming parameters, for instance an optimization of a realistic full rig system. This is currently investigated numerically. In terms of future improvements of the optimization procedure, current developments concern the treatment of constrained problems and the incorporation of a prior belief on the optimal trimming parameters. We are also planning to introduce multi-fidelity concepts [62] to further reduce the computational cost of optimizing complex nonlinear fluid-structure interaction problems.

Acknowledgments

References

[1] P. V. Oossanen, Predicting the Speed of Sailing Yachts, SNAME 101 (1993) 337–397.

[2] H. Hansen, P. S. Jackson, K. Hochkirch, Real-time velocity prediction program for wind tunnel testing of sailing yachts, in: Proc. The Modern Yacht, Southampton, UK, 2003.

[3] R. Korpus, Performance Prediction without Empiricism: A RANS-Based VPP and Design Optimization Capability, in: The 18th Chesapeake Sailing Yacht Symposium, SNAME, 2007.

[4] L. Huetz, P. E. Guillerm, Database building and statistical methods to predict sailing yacht hydrodynamics, Ocean Engineering 90 (2014) 21–33.

[5] B. Augier, P. Bot, F. Hauville, M. Durand, Experimental validation of unsteady models for fluid structure interaction: Application to yacht sails and rigs, Journal of Wind Engineering and Industrial Aerodynamics 101 (2012) 53–66.

[6] D. Trimarchi, Analysis of downwind sail structures using non-linear shell finite elements, PhD Thesis, University of Southampton, 2012.

[7] W. Menotti, M. Durand, D. Gross, Y. Roux, D. Glehen, L. Dorez, An unsteady FSI investigation into the cause of the dismating of the Volvo 70 Groupama 4, in: INNOVSail, Innovation in high performance sailing Yacht, Lorient, 197, 2013.

[8] N. Rousselon, Optimization for Sail Design, in: modeFRONTIER Conference, 2008.

[9] V. G. Chapin, R. Neyhousser, G. Dulliand, P. Chassaing, Design optimization of interacting sails through viscous CFD, in: INNOVSail, Innovation in high performance sailing Yacht, Lorient, 2008.

[10] R. Ranzenbach, D. Armitage, A. Carrau, Mainsail Planform Optimization for IRC 52 Using Fluid Structure Interaction, in: The 21st Chesapeake Sailing Yacht Symposium, SNAME, 2013.

[11] J. Nelder, R. Mead, A Simplex Method for Function Minimization, Computer Journal 7 (4) (1965) 208–313.

[12] T. Bäck, H. P. Schwefel, An Overview of Evolutionary Algorithms for Parameter Optimization, Evolutionary computation 1 (1) (1993) 1–23.

[13] N. Hansen, The CMA Evolution Strategy: A Comparing Review, in: Towards a new evolutionary computation, Springer, 75–102, 2006.

[14] J. Laurenceau, P. Sagaut, Efficient response surfaces of aerodynamic functions with Kriging and Cokriging, AIAA Journal 46 (2) (2008) 498–507.

[15] S. Jeong, M. Murayama, K. Yamamoto, Efficient Optimization Design Method Using Kriging Model, Journal of aircraft 42 (2) (2005) 413–420.

[16] B. Glaz, P. P. Friedmann, L. Liu, Helicopter Vibration Reduction throughout the Entire Flight Envelope Using Surrogate-Based Optimization, Journal of the American Helicopter Society 54 (1) (2009) 12007–12007.

[17] F. di Pierro, S.-T. Khu, D. Savić, L. Berardi, Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms, Environmental Modelling & Software 24 (2) (2009) 202 – 213, ISSN 1364-8152, doi:http://dx.doi.org/10.1016/j.envsoft.2008.06.008, URL http://www.sciencedirect.com/science/article/pii/S1364815208001084.

[18] A. J. Espinet, C. A. Shoemaker, Comparison of optimization algorithms for parameter estimation of multi-phase flow models with application to geological carbon sequestration, Advances in Water Resources 54 (2013) 133 – 148, ISSN 0309-1708, doi:http://dx.doi.org/10.1016/j.advwatres.2013.01.003, URL http://www.sciencedirect.com/science/article/pii/S0309170813000122.

[19] N. Aghajari, M. Schäfer, Efficient shape optimization for fluid–structure interaction problems, Journal of Fluids and Structures 57 (2015) 298 – 313.

[20] J. Degroote, I. Couckuyt, J. Vierendeels, P. Segers, T. Dhaene, Inverse modelling of an aneurysm's stiffness using surrogate-based optimization and fluid-structure interaction simulations, Structural and Multidisciplinary Optimization 46 (3) (2012) 457–469.

[21] V. K. Saul'ev, I. I. Samoilova, Approximation methods for the unconstrained optimization of functions of several variables, Journal of Soviet Mathematics 4 (6) (1975) 681–705.

[22] O. P. Le Maître, O. M. Knio, Spectral Methods for Uncertainty Quantification, Scientific Computation, Springer, New York, NY, 2010.

[23] T. Simpson, J. Poplinski, N. P. Koch, J. Allen, Metamodels for Computer-based Engineering Design: Survey and recommendations, Engineering with Computers 17 (2) (2001) 129–150, ISSN 1435-5663, doi:10.1007/PL00007198, URL http://dx.doi.org/10.1007/PL00007198.

[24] J. P. Kleijnen, Kriging metamodeling in simulation: A review, European Journal of Operational Research 192 (3) (2009) 707 – 716, ISSN 0377-2217, doi:http://dx.doi.org/10.1016/j.ejor.2007.10.013, URL http://www.sciencedirect.com/science/article/pii/S0377221707010090.

[25] V. Vapnik, The nature of statistical learning theory, Springer-Verlag New York, Inc., New York, NY, USA, ISBN 0-387-94559-8, 1995.

[26] D. R. Jones, M. Schonlau, W. J. Welch, Efficient Global Optimization of Expensive Black-Box Functions, Journal of Global optimization 13 (4) (1998) 455–492.

[27] D. Ginsbourger, R. Le Riche, L. Carraro, Computational Intelligence in Expensive Optimization Problems, chap. Kriging Is Well-Suited to Parallelize Optimization, Springer Berlin Heidelberg, Berlin, Heidelberg, 131–162, 2010.

[28] J. Liu, W.-P. Song, Z.-H. Han, Y. Zhang, Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models, Structural and Multidisciplinary Optimization (2016) 1–19ISSN 1615-1488, doi: 10.1007/s00158-016-1546-7, URL http://dx.doi.org/10.1007/s00158-016-1546-7.

[29] M. N. Gibbs, Bayesian Gaussian Processes for Classification and Regression, Ph.D. thesis, University of Cambridge, 1997.

[30] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[31] R. Duvigneau, P. Chandrashekar, Kriging-based optimization applied to flow control, International Journal for Numerical Methods in Fluids 69 (11) (2012) 1701–1714.

[32] M. L. Stein, Interpolation of Spatial Data - Some Theory for Kriging, Springer Science & Business Media, 2012.

[33] V. Picheny, T. Wagner, D. Ginsbourger, A benchmark of kriging-based infill criteria for noisy optimization, Structural and Multidisciplinary Optimization 48 (3) (2013) 607–626, ISSN 1615-1488, doi:10.1007/s00158-013-0919-4, URL http://dx.doi.org/10.1007/s00158-013-0919-4.

[34] R. Von Mises, Mathematical Theory of Probability and Statistics, Academic press, 1964.

[35] S. Barnett, S. Barnett, Matrix methods for engineers and scientists, McGraw-Hill, 1979.

[36] D. R. Jones, A Taxonomy of Global Optimization Methods Based on Response Surfaces, Journal of global optimization 21 (4) (2001) 345–383.

[37] D. Huang, T. T. Allen, W. I. Notz, N. Zeng, Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models, Journal of global optimization 34 (3) (2006) 441–466.

[38] N. Hansen, A. Ostermeier, Completely Derandomized Self-Adaptation in Evolution. Strategies, Evolutionary Computation

9 (2) (2001) 159–195.

[39] M. D. McKay, R. J. Beckman, W. J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics 42 (1) (2000) 55–61.

[40] P. Chandrashekar, R. Duvigneau, Study of some strategies for global optimization using Gaussian process models with application to aerodynamic design, Tech. Rep. RR-6964, INRIA, 2009.

[41] R. G. J. Flay, A twisted flow wind tunnel for testing yacht sails, Journal of Wind Engineering and Industrial Aerodynamics 63 (1–3) (1996) 171 – 182, ISSN 0167-6105, special issue on sail aerodynamics.

[42] D. J. Le Pelley, O. Modral, V-Spars: A Combined Sail and Rig Shape Recognition System Using Imaging Techniques, in: Proc. 3rd High Performance Yacht Design Conference Auckland, New Zealand, Dec, 2–4, 2008.

[43] B. Augier, Etudes expérimentales de l'interaction fluide-structure sur surface souple: application aux voiles de bateaux., PhD Thesis, Université de Bretagne Occidentale, 2012.

[44] J. Katz, A. Plotkin, Low-Speed Aerodynamics, Cambridge University Press, second edn., 2001.

[45] M. Durand, Interaction fluide-structure souple et légère, application aux voiliers, PhD Thesis, Ecole Centrale de Nantes, 2012.

[46] K. Nakashino, M. C. Natori, Efficient modification scheme of stress-strain tensor for wrinkled membranes, AIAA journal 43 (1) (2005) 206–215.

[47] Y. Roux, S. Huberson, F. Hauville, J. P. Boin, M. Guilbaud, B. Malick, Yacht performance prediction : Towards a numerical VPP, in: 1st High Performance Yacht Design Conference, 4-6 December, Auckland, 2002.

[48] Y. Roux, M. Durand, A. Leroyer, P. Queutey, M. Visonneau, J. Raymond, J. M. Finot, F. Hauville, A. Purwanto, Strongly coupled VPP and CFD RANSE code for sailing yacht performance prediction, in: 3rd High Performance Yacht Design Conference, 2-4 December, Auckland, 215–225, 2008.

[49] I. M. Viola, P. Bot, M. Riotte, Upwind sail aerodynamics: A RANS numerical investigation validated with wind tunnel pressure measurements, International Journal of Heat and Fluid Flow 39 (2013) 90–101.

[50] F. R. Menter, M. Kuntz, R. Langtry, Ten Years of Industrial Experience with the SST Turbulence Model, Turbulence, heat and mass transfer 4 (2003) 625–632.

[51] G. Kalitzin, G. Medic, G. Iaccarino, P. Durbin, Near-wall behavior of RANS turbulence models and implications for wall functions, Journal of Computational Physics 204 (1) (2005) 265–291.

[52] M. Durand, A. Leroyer, C. Lothodé, F. Hauville, M. Visonneau, R. Floch, L. Guillaume, FSI investigation on stability of downwind sails with an automatic dynamic trimming, Ocean Engineering 90 (2014) 129–139.

[53] N. Aubin, B. Augier, P. Bot, F. Hauville, R. Floch, Inviscid approach for upwind sails aerodynamics. How far can we go?, Journal of Wind Engineering and Industrial Aerodynamics 155 (2016) 208 – 215, ISSN 0167-6105, doi:http://dx.doi.org/10.1016/j.jweia.2016.06.005, URL http://www.sciencedirect.com/science/article/pii/S016761051530177X.

[54] R. G. Ghanem, P. D. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer Verlag, 1991.

[55] P. G. Constantine, M. S. Eldred, E. T. Phipps, Sparse pseudospectral approximation method, Computer Methods in Applied Mechanics and Engineering 229–232 (2012) 1–12.

[56] P. R. Conrad, Y. M. Marzouk, Adaptive Smolyak Pseudospectral Approximations, SIAM Journal on Scientific Computing 35 (6) (2013) A2643–A2670.

[57] J. Winokur, D. Kim, F. Bisetti, O. P. Le Maître, O. M. Knio, Sparse Pseudo Spectral Projection Methods with Directional Adaptation for Uncertainty Quantification, Journal of Scientific Computing (2016) 1–28In press.

[58] I. M. Sobol', Sensitivity Estimates for Nonlinear Mathematical Models, Math. Modeling & Comput. Exp. 1 (1993) 407–414.

[59] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models, John Wiley & Sons, 2004.

[60] T. Crestaux, O. Le Maître, J. M. Martinez, Polynomial chaos expansion for sensitivity analysis, Reliability Engineering & System Safety 94 (7) (2009) 1161–1172.

[61] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering & System Safety 93 (7) (2008) 964–979.

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