DATA DRIVEN UNCERTAINTY QUANTIFICATION FOR COMPUTATIONAL FLUID DYNAMICS BASED SHIP DESIGN

T.P. Scholcz

Maritime Research Institute Netherlands (MARIN)
Haagsteeg 2, 6708 PM Wageningen, The Netherlands
e-mail: t.p.scholcz@marin.nl, web page: http://www.marin.nl/

Key words: Robust ship design, data-driven, uncertainty quantification, computational fluid dynamics

Abstract. Maritime transport is responsible for an annual emission of around 1000 million tonnes of CO2, which is around 2.5% of the global greenhouse gas emissions. Nowadays, ships are designed using simplified operational profiles representing the expected operational profile during the lifetime of the ship. However, there is a discrepancy between these simplified profiles used for design and the actual full operational profile of a ship during its lifetime. This discrepancy leads to inefficient hydrodynamic ship design resulting in a waste of fuel and an increase of greenhouse gas emissions.

The amount of available data on actual operational conditions of ships is rapidly increasing. The Automatic Identification System (AIS) and onboard monitoring systems produce a huge amount of historical data on ship operations. These developments call for efficient data-driven methods that account for this data. Knowledge of operational conditions can be used for Computational Fluids Dynamics (CFD) -based probabilistic uncertainty quantification leading to robust design: A hull shape that is optimal with respect to uncertain operational conditions. Robust design is a promising approach since it makes ships energy efficient for the real usage situation.

Three UQ-methods are discussed: The perturbation method, the Polynomial Chaos Expansion (PCE) method and the multi-fidelity PCE method. The methods are applied to a simple one-dimensional test case to compute the stochastic moments of the effective power. The multi-fidelity Polynomial Chaos Method is found to be the most efficient UQ method. Moreover, the multi-fidelity PCE can be used as a surrogate for efficient Monte Carlo integration. This makes the method suitable for an Optimisation Under Uncertainty (OUU) algorithm leading to robust design.

1 INTRODUCTION

Greenhouse gas emissions from the combustion of oil-based fuels are directly proportional to fuel consumption. Some options to reduce fuel consumption are, amongst others, hull optimisation, weather routing, propeller polishing, slow steaming and trim optimisation, see [1]. Here, we will focus on improving hull optimisation techniques to reduce fuel consumption. By including realistic operational conditions in the hull form optimisation process it is expected that ships can be made more robust with respect to these conditions, leading to an improved energy efficiency.

Robust design optimisation is a promising technique that enables the ship designer to optimise the hull for the actual usage situation [2, 3]. When operational data is available of comparable ships and missions, *new* ships can be optimised to account for exactly these conditions. The resulting ship designs will be more energy efficient than designs resulting from optimisation using only a few representative design conditions.

In order for robust design optimisation to be useful it is critical to have reliable operational data relevant to the ship and mission to be designed. Sources of data are for example the Automatic Identification System (AIS) [2] and onboard monitoring systems [1]. Using AIS data, speed is recorded during the voyage and displacement at the beginning of the voyage. AIS signals are detected by satellites which provide a capability for monitoring all vessels with AIS equipment. When operational data is collected onboard, the ship has a data logging system which processes the data from on board sensors.

Robust design optimisation requires efficient Uncertainty Quantification (UQ) methods to propagate the uncertain operational conditions to the uncertain design objectives. This contribution will focus on UQ using Polynomial Chaos Expansions (PCE). The methods are ranked on computational efficiency and the capability to predict uncertainty for large input data. A simple one-dimensional test case is defined in Section 1.1 involving a container ship sailing at random speed. Two deterministic solvers are used: a cheap to evaluate low fidelity solver and an expensive high fidelity solver.

1.1 Numerical test case: 22000 TEU container vessel

The particulars of the containership are summerized in Table 1. The speed is assumed to be a random

Table 1: Main particulars

Parameter	Symbol	Value	Units
Ship	Lpp	383	m
Beam	B	61.2	m
Draft	T	14.5	m
Froude number	Fn	$N(0.1847, 008397)$	[-]

variable which can be decomposed in m standard independent random variables:

$$V = V(\xi_i) \quad \xi_i = i, ..., m \tag{1}$$

For simplicity, here the speed is assumed to be a function of a single standard Gaussian random variable

$$V = V_0 + \sigma_V \cdot \xi, \tag{2}$$

where the speed V_0 and σ_V are chosen to represent a 'realistic' operational profile. Here, we choose $V_0 = 22$ knots and $\sigma_V = 1$ knot which leads to a random Froude number $Fn = N(0.1847, 0.008397)$. The deterministic design objective is the effective power of the ship

$$P_{\mathrm{E}} = R \cdot V, \tag{3}$$

where R is the total resistance of the ship. The total resistance can be predicted using the low fidelity potential flow solver RAPID in order to compute a quick and rough estimate and using the viscous high fidelity solver ReFRESCO to compute a reliable value.

1.1.1 Low fidelity solver: free-surface potential code RAPID

The fully nonlinear free-surface potential flow solver Raised-Panel Iterative Dawson (RAPID) [4] is a panel method neglecting viscous effects. Using RAPID, the total (viscous) resistance is estimated using

$$R = (1+k)R_{f_0} + R_w, \tag{4}$$

where R_{f_0} is the frictional resistance of a flat plate with a surface equal to the wetted surface area S moving at speed V, the factor k is an estimated form factor and R_w is the wave resistance computed from integration of the pressure resistance along the hull. Expression (4) yields a Low Fidelity (LF) prediction of the total resistance which is computed in a few minutes on a desktop PC.

1.1.2 High fidelity solver: free surface RANS code ReFRESO

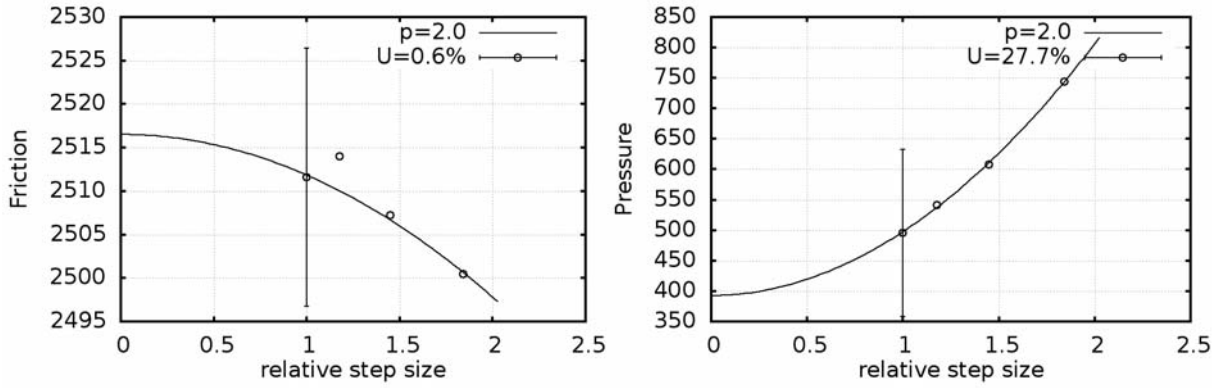
The free surface RANS code ReFRESCO solves incompressible viscous flows using the Navier-Stokes equations, see [5]. ReFRESCO is used for High Fidelity (HF) predictions of the resistance. ReFRESCO is an unstructured finite volume code based on the volume-of-fluid formulation. To assess the prediction accuracy, the numerical uncertainty needs to be estimated using a grid refinement study. Richardson extrapolation is used in the least square sense since scatter in the solution may deteriorate the observed order of convergence p. The goal is to obtain the discretisation uncertainty U of prediction $\hat{R}$ such that with 95% confidence: $\hat{R} - U \leq R^* \leq \hat{R} + U$, where R^* denotes the exact solution. Equivalently, this can be written as

$$P\left[|\varepsilon_R| \geq \frac{U}{R^*}\right] \leq 1 - C \quad \text{with} \quad \varepsilon_R = \frac{\hat{R} - R^*}{R^*}, \tag{5}$$

where ε_R is the discretisation error and $C = 0.95$. The uncertainty U is obtained using the method described in [6]. Four solutions on geometrically similar grids with sizes $N = 2.7, 4.6, 7.4$ and 10.8 million cells are used to fit $\delta_{RE} = R_i - R^* = \alpha h_i^p$. When the observed order p satisfies $1.0 \leq p \leq 2.1$ and the standard deviation of the fit $\hat{\sigma}$ is smaller than the mean change in the data, the value for U is estimated using

$$U = F_s \delta_{RE} + \hat{\sigma} + |\hat{R} - R_{\text{fit}}|, \tag{6}$$

with a safety factor $F_s = 1.25$. A distinction is made between the friction resistance and the pressure resistance, see Figure 1.



(a) Friction resistance [kN]

(b) Pressure resistance [kN]

Figure 1: Grid convergence of the friction and pressure resistance.

The results for the finest grid with $N = 10.8$ million cells are summarized in Table 2.

Table 2: Errors and uncertainties of the finest grid (N=10.8 million cells)

	Friction	Pressure	Total
Approximation ($\hat{R}$)	2512 kN	496 kN	3007 kN
Uncertainty ($U/\hat{R}$)	5.90×10^{-3}	2.77×10^{-1}	4.60×10^{-2}

The uncertainty of the total resistance is obtained using $U_T = \sqrt{U_f^2 + U_p^2}$, which assumes independent convergence of pressure and friction. Note that the numerical uncertainty of the pressure is much higher than the uncertainty of the friction and that the total uncertainty is almost 5% of the estimated resistance. However, grid convergence is observed such that this uncertainty can always be reduced when more grid cells are used for the computation. To limit computational costs, we use $N = 10.8$ million cells in the remainder of the paper.

1.1.3 Quantities of Interest

The Quantities of Interest (QoI) are the stochastic moments of the effective power with respect to the uncertain operational conditions. The effective power depends on the random variable as

$$P_{\mathrm{E}}(\xi) = R\left(V(\xi)\right) \cdot V(\xi). \tag{7}$$

The statistical moments of the effective power are then defined by

$$\begin{aligned} \mu_{P_{\mathrm{E}}} &= \int P_{\mathrm{E}}(\xi) p_\xi d\xi \\ \sigma^2_{P_{\mathrm{E}}} &= \int \left(P_{\mathrm{E}}(\xi) - \mu_{P_{\mathrm{E}}}\right)^2 p_\xi d\xi, \end{aligned} \tag{8}$$

where p_ξ denotes the probability density function of random variable ξ. Computing the integrals in (8) efficiently is required in order to be able to optimise the hull shape for minimal expected power. Efficient UQ methods are the subject of Section 2.

2 Uncertainty Quantification Methods

The Monte Carlo method is the most general and reliable method to propagate uncertainty to the objectives. Unfortunately, it is also the most expensive method which makes it not suitable for robust optimisation. More efficient UQ methods exploit the smoothness of the objectives with respect to the uncertain variables. Three UQ methods are applied: The perturbation method, the Polynomial Chaos Expansion method and the multi-fidelity Polynomial Chaos Expansion method. To this end, Sandia National Laboratories open-source DAKOTA tool is used, see [7].

2.1 Perturbation method

If the objective function is smooth, the effective power can be expanded in a multi-dimensional second-order Taylor expansion around a point α (see [8])

$$\begin{aligned} P_\mathrm{E}(\xi) &= P_\mathrm{E}\Big|_\alpha + \nabla P_\mathrm{E}\Big|_\alpha (\xi-\alpha) + \frac{1}{2}(\xi-\alpha)^T \nabla^{(2)} P_\mathrm{E}\Big|_\alpha (\xi-\alpha) \\ &= a + \mathbf{b}^T\xi + \frac{1}{2}\xi^T\mathbf{C}\xi, \end{aligned} \tag{9}$$

where a, $\mathbf{b}$ and $\mathbf{C}$ are given by

$$\begin{aligned} a &= P_\mathrm{E}\Big|_\alpha - \nabla P_\mathrm{E}\Big|_\alpha \alpha \\ \mathbf{b}^T &= \nabla P_\mathrm{E}\Big|_\alpha - \alpha^T \nabla^{(2)} P_\mathrm{E}\Big|_\alpha \\ \mathbf{C} &= \nabla^{(2)} P_\mathrm{E}\Big|_\alpha . \end{aligned} \tag{10}$$

Here, ∇P_E denotes the gradient of the effective power with respect to the uncertain variables and $\nabla^{(2)} P_E$ the Hessian of the effective power. From equation (7), the gradient and Hesssian are computed as

$$\frac{dP_\mathrm{E}}{d\xi_i} = \left(\frac{\partial R}{\partial V}\cdot V + R\right)\cdot\frac{\partial V}{\partial \xi_i} \quad \text{for} \quad i = 1,...,m \tag{11}$$

and

$$\frac{d^2P_\mathrm{E}}{d\xi_i d\xi_j} = \left(\frac{\partial^2 R}{\partial V^2}\frac{\partial V}{\partial \xi_j}V + 2\frac{\partial R}{\partial V}\frac{\partial V}{\partial \xi_j}\right)\cdot\frac{\partial V}{\partial \xi_i} + \left(\frac{\partial R}{\partial V}\cdot V + R\right)\cdot\frac{\partial^2 V}{\partial \xi_i \partial \xi_j} \quad \text{for} \quad i,j = 1,...,m. \tag{12}$$

If the random variables are independent and Gaussian distributed, the statistical moments follow from

$$\begin{aligned} \mu_{P_\mathrm{E}} &= a + \frac{1}{2}\mathrm{Tr}[\mathbf{C}] \\ \sigma^2_{P_\mathrm{E}} &= \mathbf{b}^T\mathbf{b} + \frac{1}{2}\mathrm{Tr}[\mathbf{C}^2], \end{aligned} \tag{13}$$

where Tr[] is the trace operator. The mean-centred perturbation method results when the center of expansion α equals the mean of the input random variables: $\alpha = 0$ in case of Gaussian distributed variables.

2.2 Polynomial Chaos Expansion Method

A (truncated) polynomial chaos expansion is an expansion in a finite number of basis functions. The basis functions ψ_i with $i = 0...P$ are chosen such that

$$\langle \psi_i(\xi)\psi_j(\xi)\rangle = \int \psi_i(\xi)\psi_j(\xi)p_\xi(\xi)d\xi = \delta_{ij}, \tag{14}$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The effective power is expanded in the basis functions

$$\begin{aligned} P_{\mathrm{E}}(\xi) &= a_0 B_0 + \sum_{i_1=1}^{\infty} a_{i_1} B_1(\xi_{i_1}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} a_{i_1 i_2} B_2(\xi_{i1},\xi_{i2}) + ... \\ &\approx \sum_{i=0}^{P} \alpha_i \psi_i(\xi), \end{aligned} \tag{15}$$

where P is the expansion order. Standard Hermite functions are used for the (Askey) polynomial basis which satisfy (14) in case p_ξ is normally distributed. Two options are available to compute the expansion coefficients: the spectral method and the regression method ([9]). Once the coefficients are computed, the statistical moments of the effective power are determined using

$$\begin{aligned} \mu_{P_{\mathrm{E}}} &= \alpha_0 \\ \sigma^2_{P_{\mathrm{E}}} &= \sum_{i=1}^{P} \alpha_i^2 \langle \Psi_i^2 \rangle. \end{aligned} \tag{16}$$

Using the spectral method, Equation (15) is multiplied with basisfunctions ψ_i and integrated over the random variable space to obtain

$$\alpha_i = \frac{\langle P_{\mathrm{E}}(\xi), \psi_i(\xi)\rangle}{\langle \psi_i^2(\xi)\rangle} = \frac{1}{\langle \psi_i^2(\xi)\rangle} \int P_{\mathrm{E}}(\xi)\psi_i(\xi)\rho(\xi)d\xi, \tag{17}$$

where use has been made of the orthogonality of the basis functions with respect to the inner product $\langle\cdot\rangle$. To evaluate the integrals in (17) sparse grid integration can be used. So called Genz-Keister ([10]) quadrature results in a number of quadrature points and corresponding expansion orders that depend on the dimension of the problem and the level of the grid. For example, in one dimension, the number of required quadrature points is given in Table 3. For higher dimensions, tables can be found in [10, 11].

Table 3: Number of quadrature points (samples) for a 1-dimensional Genz-Keister quadrature

Level	N_t	P	Number of samples
0	1	0	1
1	3	2	3
2	7	6	9
3	15	14	19

As can be seen in Figure 2 the quadrature rule is nested which means there is an overlap between the quadrature points on different grid levels. The advantage of nested quadrature rules is the possibility of

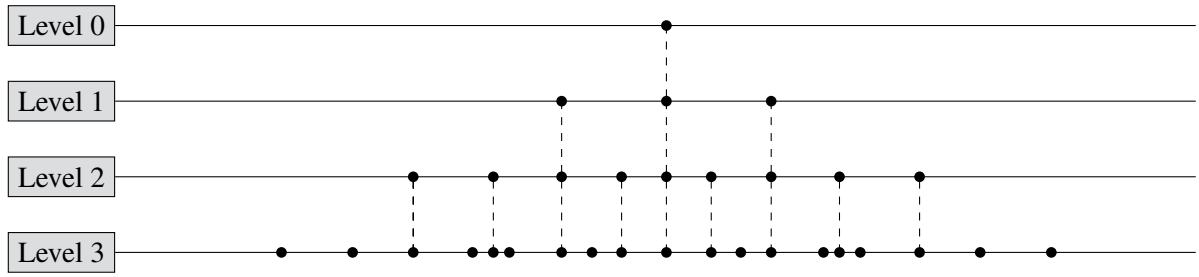


Figure 2: Genz Keister quadrature points in 1 dimension. The dashed lines indicate the overlap/nesting.

reusing results from one grid to the other and the possibility of combining high and low fidelity results on different grid levels in a multi-fidelity method. Using Genz-Keister quadrature, the PCE coefficients can be computed from (17) and substituted in (16) to compute the stochastic moments.

2.2.1 Multi-fidelity Polynomial Chaos Expansion

In case of an additive multifidelity correction model, the model discrepancy is defined as

$$\Delta(\xi) = P_{\mathrm{E}}^{\mathrm{HF}}(\xi) - P_{\mathrm{E}}^{\mathrm{LF}}(\xi), \tag{18}$$

which can be expanded in the polynomial basis as

$$\Delta(\xi) = \sum_{i=0}^{P_{\mathrm{HF}}} \delta_i \psi_i(\xi). \tag{19}$$

If the HF-solver is used at a certain grid level and the LF-solver at a higher grid level, the model discrepancy can only be computed on the overlapping points, see Figure 2. The discrepancy expansion is therefore constructed on the lower grid level. Once the model discrepancy expansion is constructed it can be evaluated at the higher grid level to correct the LF prediction:

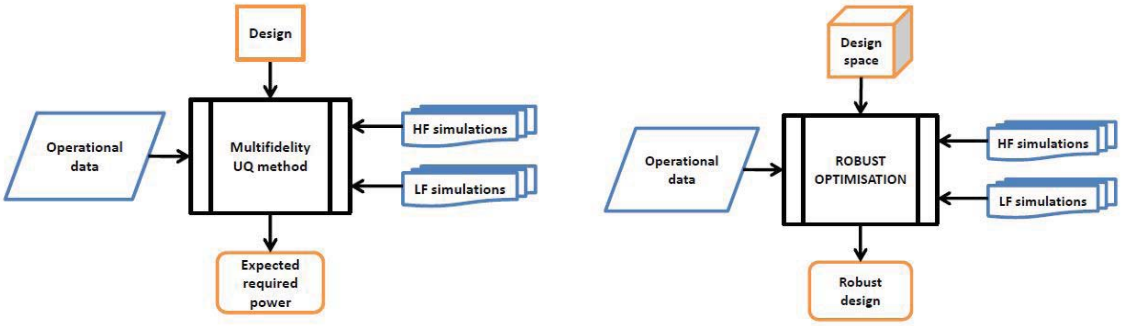
$$\begin{aligned} P_{\mathrm{E}} &\approx \sum_{i=0}^{P_{\mathrm{HF}}} \delta_i \psi_i(\xi) + \sum_{i=0}^{P_{\mathrm{LF}}} \alpha_i^{\mathrm{LF}} \psi_i(\xi) && (20) \\ &= \sum_{i \in \mathrm{CB}} (\alpha_i^{\mathrm{LF}} + \delta_i)\psi_i(\xi) + \sum_{i \in \mathrm{B_{LF}} \setminus \mathrm{CB}} \alpha_i^{\mathrm{LF}} \psi_i(\xi). && (21) \end{aligned}$$

Here, $\mathrm{B_{rapid}}$ is the polynomial basis for the LF expansion and CB the Common polynomial Basis of the HF and LF expansion, see [12, 13]. The stochastic moments μ_{P_E} and $\sigma^2_{P_E}$ are then computed in the same way as in Equation (16)

$$\begin{aligned} \mu_{P_{\mathrm{E}}} &= \alpha_0^{\mathrm{LF}} + \delta_0 \\ \sigma^2_{P_{\mathrm{E}}} &= \sum_{i \in \mathrm{CB} \setminus 0} (\alpha_i^{\mathrm{LF}} + \delta_i)^2 \langle \Psi_i^2 \rangle + \sum_{i \in \mathrm{B_{LF}} \setminus \mathrm{CB}} (\alpha_i^{\mathrm{LF}})^2 \langle \Psi_i^2 \rangle. \end{aligned} \tag{22}$$

3 Robust optimisation using data driven UQ

An Optimisation Under Uncertainty (OUU) procedure allows to account for the uncertainties in the operational conditions during the hull form optimisation. The challenge is to efficiently and reliably include the uncertainties in the analysis such that the result will be a hull design which is robust with respect to given representative operational conditions.



(a) Multifidelity Uncertainty Quantification of the expexted required power for reprentative operational data

(b) Robust Optimisation of the expected required power. Optimisation Under Uncertainty (OUU), see [14].

Figure 3: Data driven design methodology: (a) Prediction (b) Optimisation.

Figure 3a shows the procedure to predict the expected required power using a multifidelity UQ method. When for example a PCE surrogate is constructed from multifidelity simulations it is possible to predict the expected required power in a very cheap and efficient way since the surrogate can be evaluated at a negligible cost. Any operational data can be used as long as the surrogate represents the high fidelity objective in the range of conditions.

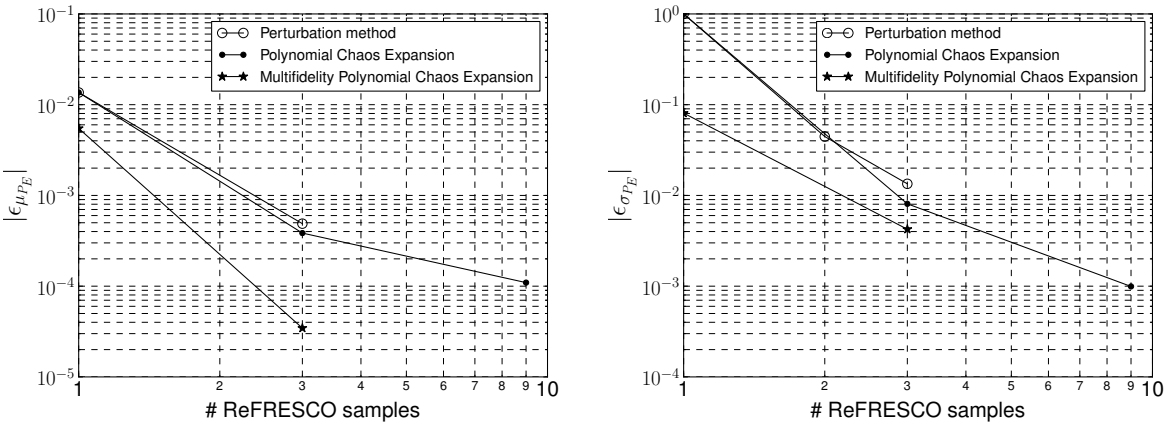
Figure 3b shows the procedure to optimise a hull form for expected required power given representative operational data. Here, the operational data drives the design which means that representative operational conditons will determine the optimal hull shape from the design space. Different formulations exist to perform the optimisation: Nested, Layered/Nested, Nested/Layered and Layered/Nested/Layered, see [14]. The latter constructs a surrogate at the UQ level *and* at the optimisation level. In other words, each point (hull shape) of the Design of Experiment at the optimisation level contains an entire surrogate based UQ analysis to predict the expected required power. This is the most attractive option from a computational point of view but only works if the quality of the surrogates are verified.

4 Results and discussion

In this section we examine the quality of the UQ methods from section 2. Assuming converged moments $\mu^*_{P_E}, \sigma^*_{P_E}$ at grid level 3, the errors in the predicted moments $\hat{\mu}_{P_E}, \hat{\sigma}_{P_E}$ are defined as

$$\varepsilon_{\mu_{P_E}} = \frac{\hat{\mu}_{P_E} - \mu^*_{P_E}}{\mu^*_{P_E}} \quad \text{and} \quad \varepsilon_{\sigma_{P_E}} = \frac{\hat{\sigma}_{P_E} - \sigma^*_{P_E}}{\sigma^*_{P_E}} \tag{23}$$

As can be seen from Table 3 and Figure 2, 19 high fidelity calculations (samples) were required to compute $\mu^*_{P_E}$ and $\sigma^*_{P_E}$. Figure 4a and 4b show the convergence of the mean and standard deviation of the effective power with respect to the number of required samples for each method from section 2.



(a) Convergence of the mean μ_{P_E}

(b) Convergence of the standard deviation σ_{P_E}

Figure 4: Convergence of statistical moments with the number of required ReFRESCO samples

The predictions are collected in Table 4 and ranked in performance for predicting the expected effective power μ_{P_E}. It can be seen that the UQ methods all converge rapidly to $\mu^*_{P_E}$ and $\sigma^*_{P_E}$ and that the multifidelity PCE method is the most efficient UQ technique. Multi-fidelity PCE at levels 1,2 is even more accurate in predicting the expected effective power than PCE at level 2. However, note that this is not true for the standard deviations of the effective power. The accuracy of multi-fidelity PCE at levels 0,1 falls exactly between the accuracy of the first and second order perturbation method: a constant correction to the LF data performs better than linearisation but worse than a curved approximation . Due to the uncertain speed ($\sigma_V = 1$ knot), the expected power is about 1.4% higher than the power at $V_0 = 22$ knots.

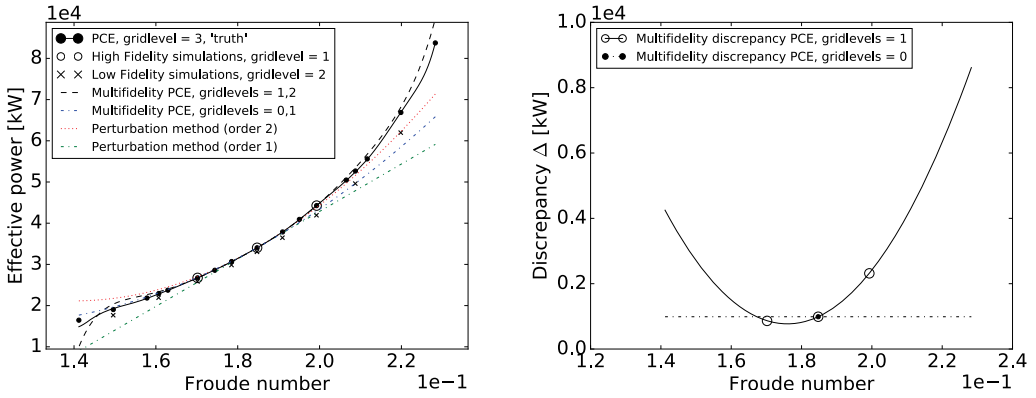
Method	Resistance		Effective power	
	μ_R (ε_μ) [kN]	σ_R (ε_σ) [kN]	μ_{P_E} (ε_μ) [kW]	σ_{P_E} (ε_σ) [kW]
PCE, level = 3, 'truth'	**3035.0 (0.00%)**	**306.1 (0.00%)**	**34505.6 (0.00%)**	**5069.3 (0.00%)**
Multi-fi PCE, levels = 1,2	3035.5 (0.01%)	307.5 (0.47%)	34506.7 (0.00%)	5090.8 (0.42%)
PCE, level = 2	3035.4 (0.01%)	305.5 (-0.18%)	34509.3 (0.01%)	5064.3 (-0.10%)
PCE, level = 1	3036.4 (0.05%)	309.5 (1.11%)	34518.8 (0.04%)	5110.3 (0.81%)
Second order perturbation	3034.1 (-0.03%)	299.8 (-2.04%)	34488.7 (-0.05%)	5001.2 (-1.34%)
Multi-fi PCE, levels = 0,1	3020.1 (-0.49%)	276.4 (-9.71%)	34316.5 (-0.55%)	4660.1 (-8.07%)
First order perturbation	3007.2 (-0.91%)	290.6 (-5.06%)	34035.4 (-1.36%)	4843.5 (-4.46%)
PCE, level = 0	3007.2 (-0.91%)	-	34035.4 (-1.36%)	-
Zeroth order perturbation	3007.2 (-0.91%)	-	34035.4 (-1.36%)	-

Table 4: Comparison of UQ methods

PCE can also be used as a surrogate for the high fidelity solver in order to estimate the effective power. Using Monte Carlo integration, stochastic moments can be computed for arbitrary input data at almost negligible cost. This is the subject of section 4.1.

4.1 Comparison of surrogates and Monte Carlo integration

The surrogates of each UQ method from section 2 are shown in Figure 5a together with the high fidelity evaluations at grid level 1 and the low fidelity evaluations at grid level 2. The conclusions of the previous section are now visualised: Multi-fidelity PCE at grid levels 1,2 performs best and and the multi-fidelity PCE at grid levels 0,1 falls exactly between the first and second order perturbation curves. Figure 5b shows the model discrepancy expansions of the multi-fidelity PCEs. The discrepancies are small such that the multi-fidelity approach works well in this case.

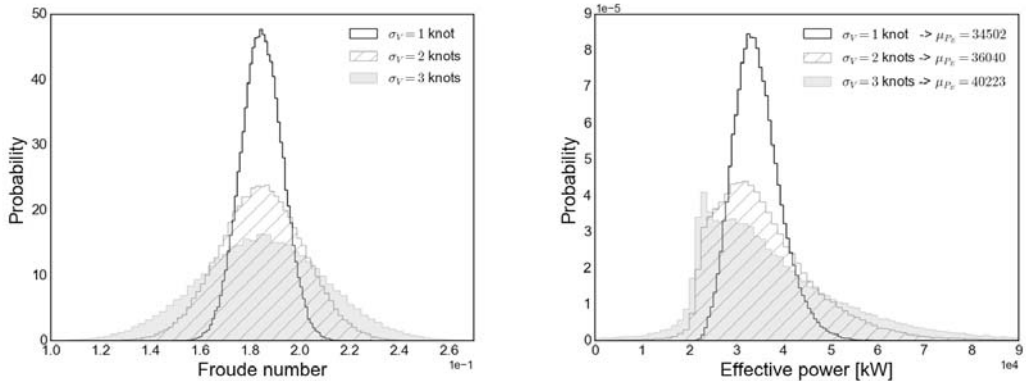


(a) Surrogate predictions vs speed

(b) Multifidelity discrepancy predictions vs speed

Figure 5: UQ Surrogate comparison

Monte Carlo integration is now applied to the multi-fidelity PCE surrogate at grid levels 1,2 using 100,000 samples drawn from three Gaussian distributions with $\sigma_V = 1,2,3$ knots, see Figure 6. A rapid increase in expected effective power is observed due to the stretched tail which is a consequence of increased effective power at high speeds. The entire computation takes only a few seconds.



(a) Input uncertainty: speed

(b) Output uncertainty: effective power

Figure 6: Monte Carlo on the multi-fidelity PCE surrogate constructed at grid levels 1,2. $N_{mc} = 100,000$.

5 Conclusions

Greenhouse gas emissions are directly proportional to fuel consumption. One option to mitigate fuel consumption is to optimise the ship hull with respect to realistic operational conditions leading to a so called 'robust design'.

Robust design optimisation requires efficient Uncertainty Quantification in order to propagate the uncertainty to the objectives. In addition to efficiency, the UQ method should also be able to deal with large (measured) input data. Three UQ methods are discussed: the perturbation method, the Polynomial Chaos Expansion method and the multi-fidelity Polynomial Chaos Expansion method. The methods are applied to a simple one-dimensional test case: a container ship sailing at random speed.

The multi-fidelity Polynomial Chaos Expansion is found to be the most efficient UQ method for this test case. Moreover, the PCE method can be used as a surrogate for efficient Monte Carlo integration in case large input data of arbitrary distribution is available. Since surrogate evaluations are computationally cheap, they are well suited for an Optimization Under Uncertainty (OUU) algorithm leading to robust design. Such a data-driven hull optimisation approach will be the subject of future work.

6 Acknowledgements

This research was funded from the TKI-allowance of the Dutch Ministry of Economic Affairs. The support is gratefully acknowledged.

REFERENCES

[1] Coraddu, A., Oneto, L. and Anguita D. Vessels Fuel Consumption: A Data Analytics Perspective to Sustainability. in: Cruz Corona C. (eds) *Soft Computing for Sustainability Science. Studies in*

Fuzziness and Soft Computing, vol 358. Springer, Cham (2018).

[2] Peri, D. 'Robust Design Optimization for the refit of a cargo ship using real seagoing data' , *Ocean Engineering*, Vol. 123, pp. 103-115, (2016).

[3] Wei, H., Diez, M, Campana, E., Stern, F. and Zao-jian, Z. 'A one-dimensional polynomial chaos method in CFD-Based uncertainty quantification for ship hydrodynamics performance' *Journal of Hydrodynamics*, Vol. 25(5), pp. 655–662, (2013).

[4] Raven, H. C., "A solution Method for the Nonlinear Ship Wave Resistance Problem" PhD Thesis Delft University of Technology (1996).

[5] Vaz, G., Jaouen, F., Hoekstra, M. "Free-surface viscous flow computations. Validation of URANS code FRESCO", *Proceedings of OMAE2009*, Hawaii, USA, (2009).

[6] Eça, L., Hoekstra, M., "Evaluation of numerical error estimation based on grid refinement studies with the method of the manufactured solutions", *Computer & Fluids*, Vol. 38, pp. 1580–1591, (2009).

[7] Adams, B.M., Ebeida, M.S., Eldred, M.S., Geraci, G., Jakeman, J.D., Maupin, K.A., Monschke, J. A., Stephens, J.A., Swiler, L.P., Vigil, D.M., Wildey, T.M., "Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and sensitivity Analysis: Version 6.9 User's Manual" *Sandia Technical Report SAND2014-4633*.

[8] Verhoosel, C.V., Scholcz, T.P., Hulshof S.J. and Gutierrez, M.A. 'Uncertainty and reliability analysis of fluid-structure stability boundaries', *AIAA Journal*, Vol. 47(1), pp. 91-104, (2009).

[9] Duz, B., Ypma, E. "Uncertainty quantification in numerical simulations of parametric roll", *Proceedings of the ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering OMAE2018, 17-22 June 2018, Madrid, Spain. OMAE2018-77801*

[10] Genz, A. and Keister, B.D. 'Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight', *Journal of Computational and Applied Mathematics*, Vol. 71, pp. 299-309, (1996).

[11] Chen, P. 'Sparse quadrature for high dimensional integration with Gaussian measure', *ESAIM Journal*, Vol. 52(2), pp. 631-657, (2018).

[12] Ng, L. W. T. and Eldred, M. S. 'Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation', *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, 23-26 April 2012, Honolulu, Hawaii. AIAA 2012-1852

[13] Palar, P.S., Zuhal, L.R., Shimoyama, K., Tsuchiya, T. 'Global sensitivity analysis via multi-fidelity polynomial chaos expansion', *Reliability Engineering & System Safety*, Vol. 170, pp.175–190, (2018).

[14] Eldred, M. S., Giunta, A. A., Wojtkiewicz, S. F. and Trucano, T. G. "Formulations for surrogate-based optimization under uncertainty", *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA-2002-5585 (2002).