Struct Multidisc Optim (2015) 51:347–368
DOI 10.1007/s00158-014-1128-5

RESEARCH PAPER

Development and validation of a dynamic metamodel based on stochastic radial basis functions and uncertainty quantification

Silvia Volpi · Matteo Diez · Nicholas J. Gaul · Hyeongjin Song · Umberto Iemma · K. K. Choi · Emilio F. Campana · Frederick Stern

Received: 26 August 2013 / Revised: 16 May 2014 / Accepted: 3 June 2014 / Published online: 7 August 2014

Abstract A dynamic radial basis function (DRBF) metamodel is derived and validated, based on stochastic RBF and uncertainty quantification (UQ). A metric for assessing metamodel efficiency is developed and used. The validation includes comparisons with a dynamic implementation of Kriging (DKG) and static metamodels for both deterministic test functions (with dimensionality ranging from two to six) and industrial UQ problems with analytical and numerical benchmarks, respectively. DRBF extends standard RBF using stochastic kernel functions defined by an uncertain tuning parameter whose distribution is arbitrary and whose effects on the prediction are determined using UQ methods. Auto-tuning based on curvature, adaptive sampling based on prediction uncertainty, parallel infill, and multiple response criteria are used. Industrial problems are two UQ applications in ship hydrodynamics using high-fidelity computational fluid dynamics for the high-speed Delft catamaran with stochastic operating and environmental conditions: (1) calm water resistance, sinkage and trim with variable Froude number; and (2) mean value and root mean square of resistance and heave and pitch motions with variable regular head wave. The number of high-fidelity evaluations required to achieve prescribed error levels is considered as the efficiency metric, focusing on fitting accuracy and UQ variables. DKG is found more efficient for fitting low-dimensional test functions and one-dimensional UQ, whereas DRBF has a greater efficiency for fitting higher-dimensional test functions and two-dimensional UQ.

Keywords Simulation-based design · Dynamic metamodels · Uncertainty quantification · Radial basis function networks · Kriging

S. Volpi · M. Diez · N. J. Gaul · H. Song · K. K. Choi · F. Stern (✉)
The University of Iowa, Iowa City, IA, USA
e-mail: frederick-stern@uiowa.edu

S. Volpi · U. Iemma
Department of Engineering, Roma Tre University, Rome, Italy

M. Diez · E. F. Campana
CNR-INSEAN, National Research Council-Marine Technology Research Institute, Rome, Italy

1 Introduction

Simulation-based design (SBD) of complex engineering systems requires high-fidelity solvers to guarantee the accuracy of the solution. Real-world problems are affected by different sources of uncertainty (environmental, operational, geometrical) and therefore need uncertainty quantification (UQ) methods. Combining design optimization and UQ into stochastic SBD, such as robust and reliability-based design optimization, requires a high number of function evaluations and large computational resources. This represents a significant challenge from the algorithmic and technological viewpoints, requiring efficient computational methods and high-performance computer systems.

The application of surrogate models, i.e. metamodels, alleviates the computational cost by reducing the number of high-fidelity evaluations needed. Metamodels have been widely used in several engineering contexts, such as structural optimization (Jansson et al. 2003), aeronautics and aerospace (Sobieszczanski-Sobieski and Haftka 1997), and ground vehicles (Yang et al. 2005), including stochastic applications and UQ (Giunta et al. 2006; Kennedy et al. 2006).

The choice of the metamodelling technique is based on accuracy and efficiency. The latter is based on the number of high-fidelity evaluations required, since the computational cost of the metamodelling algorithm is deemed negligible in comparison. Radial basis functions (RBF, Hardy 1971) have been demonstrated accurate and efficient in several applications such as analytical test problems (Jin et al. 2001), stochastic search in optimization (Regis and Shoemaker 2007; Regis 2011) and UQ (Loeven et al. 2007, He et al. 2013). Accuracy and efficiency of Kriging (Matheron 1963) has been demonstrated for several applications including optimization subject to uncertainty (Jin et al. 2003); moreover, extensions of Kriging to stochastic approaches including uncertain basis/correlation functions and tuning parameters has been addressed in Bayesian Kriging, using UQ (Pilz and Spock 2008; Gramacy and Lee 2008). Although RBF and Kriging are found adequate for most problems, it is difficult to predict their effectiveness for new applications. Metamodels performance is problem dependent and determined by different factors such as the degree of non-linearity, the problem dimensionality, the noisy or smooth behavior of the function and the approach used for training (Jin et al. 2001).

In order to develop accurate and efficient methods for metamodel-based analysis and optimization, research has recently moved from standard (or static) metamodelling techniques to function-adaptive approaches, also referred to as dynamic metamodels. A dynamic metamodel is able to improve its fitting capability by exploiting the information that becomes available during the analysis process. Two main characteristics identify metamodels as dynamic: an auto-tuning of the metamodel itself and an adaptive sampling technique.

In auto-tuning, the metamodel itself is not defined *a priori*. Auto-tuning can be applied considering several degrees of freedom, from tuning parameters to the choice of the metamodel itself. Auto-tuning approaches to RBF have been applied to auto-configure RBF networks: Mullur and Messac (2005) propose an extended RBF approach, where more than one basis function per data point is used resulting in an under-determined system of equations; Acar and Rais-Rohani (2009) and Zhou et al. (2011) present weighted sum of multiple metamodels, updating the weights at each iteration for improving the accuracy; Billings and Zheng (1995) propose a global optimization process using genetic algorithms; Sarimveis et al. (2004) make use of a genetic algorithm to minimize the prediction error and auto-configure dynamically RBF neural networks; Meng et al. (2009) present self-adaptive RBF neural networks using differential evolution. Auto-tuning applications to Kriging have been shown in Peri (2009), Zhao et al. (2011) and Song et al. (2013). Auto-tuning involves an optimization procedure, for which the choice of the degrees of freedom and the minimization algorithm represents a critical issue.

Adaptive sampling supports the design of experiments (DoE) used for training, which is not defined anymore *a priori* but dynamically updated using available information. The purpose of performing an adaptive DoE is to add training points where it is most useful and to use the minimum number of high-fidelity evaluations to represent the function. Li et al. (2010) shows a classification of sampling techniques, setting apart non-adaptive DoE from adaptive DoE. The latter can follow different approaches, depending on application and aim of the analysis. Relevant issues for efficient adaptive DoE are the possibility to add more than one training point per iteration, in order to take advantage of parallel computing systems (referred to as parallel infill, Forrester and Keane 2009) and the capability of managing more than one function at a time (when the relevant outputs are multiple).

Several metrics are used to evaluate metamodels accuracy. R-Square, relative average absolute error and relative maximum absolute error (Jin et al. 2001), root mean square error (Jones et al. 1998) and relative root mean square error (Zhao et al. 2011) are some of the most widely used. These are based on the L^p norm (suitably normalized) of the difference between observations and predictions. Metrics based on absolute ($p = 1$) and square ($p = 2$) errors give an assessment of the global fitting, whereas maximum errors ($p = \infty$) focus on local differences; their application to metamodels comparison gives similar trends, especially for smooth functions and predictions. Metrics and methods for validation of metamodel-based UQ have been presented by Mousaviraad et al. (2013), providing errors for function fitting, expected value (EV), standard deviation (SD) and cumulative distribution function (CDF) versus numerical benchmark obtained by computational fluid dynamics (CFD). Metrics for accuracy do not directly provide the quantification of metamodels efficiency, which is of primarily importance when high-fidelity solvers are used.

The objective of the present work is the development and validation of an efficient dynamic RBF (DRBF) metamodel based on stochastic kernel functions. A metric for assessing metamodels efficiency is developed and used. The validation includes comparison with an existing dynamic Kriging (DKG) method (Zhao et al. 2011; Song et al. 2013), available through collaboration among co-authors, and static metamodels for both deterministic test functions (with dimensionality ranging from two to six) and industrial UQ problems (Diez et al. 2014, He et al. 2013) with analytical and numerical benchmarks, respectively. Development and validation of DKG are beyond the scope of the current work, which focuses on DRBF and uses a well-established implementation of DKG for comparison.

DRBF extends standard RBF using stochastic kernel functions defined by an uncertain tuning parameter whose distribution is arbitrary and whose effects on the prediction are determined using UQ methods, similarly to Bayesian Kriging. Auto-tuning based on curvature, adaptive sampling based on prediction uncertainty, parallel infill, and multiple response criteria are used. Since DRBF is aimed at UQ in ship hydrodynamics problems with low dimensionality, test functions with dimensionality ranging from two to six are used. Industrial problems are two UQ applications in ship hydrodynamics making use of high fidelity CFD for the high-speed Delft catamaran with stochastic operating and environmental conditions: (1) calm water resistance, sinkage and trim for variable Froude number (Fr) (Diez et al. 2014); and (2) mean value and root mean square (RMS) deviation from mean of resistance and heave and pitch motions for variable regular head wave He et al. (2013). The number of evaluations required to achieve prescribed error levels is considered as the efficiency metric, focusing on fitting capability and UQ variables. An appendix provides the UQ equations for EV, SD, CDF and stochastic uncertainty U_f, where f indicates the output function; they are used both for determining prediction uncertainty stemming from stochastic tuning parameter in DRBF and for industrial UQ problems.

The paper is organized as follows. Section 2 introduces the proposed approach for DRBF and the validation metric used. Section 3 presents the static metamodels and the dynamic implementation of Kriging used for comparison. Deterministic and stochastic analytical and numerical benchmarks for validation of current methodology are presented in Section 4, whereas the associated numerical results are discussed in Section 5. Final remarks and future work are given in Section 6. Finally, Appendices A and B provide the UQ methods used and the equations for the analytical test functions, respectively.

2 Stochastic dynamic radial basis functions and evaluation metrics

Given a training set $\mathcal{T}$ of M points $\{\mathbf{x}_i\}_{i=1}^{M}$ with associated function (f) evaluations $y_i = f(\mathbf{x}_i)$, standard RBF (with centers coincident with $\mathbf{x}_i$) provides the prediction ($\hat{f}$) at $\mathbf{x}$ as per

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{M} w_i \varphi(||\mathbf{x} - \mathbf{x}_i||) \tag{1}$$

where φ is the kernel function and the w_i are the coefficients of the combination. These are solution of the linear system, which provides exact predictions at the training points:

$$\mathbf{A}\mathbf{w} = \mathbf{y} \tag{2}$$

where the elements of $\mathbf{A}$ are $a_{ij} = \varphi(||\mathbf{x}_i - \mathbf{x}_j||)$, with $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{T}$, $\mathbf{w} = \{w_j\}$ and $\mathbf{y} = \{y_i\}$. The ε power of the Euclidean distance is used as the kernel function, with ε treated as a tuning parameter:

$$\varphi||\mathbf{x} - \mathbf{x}_i|| = ||\mathbf{x} - \mathbf{x}_i||^{\varepsilon} = \left[\sqrt{\sum_{k=1}^{n}(x_k - x_{k,i})^2}\right]^{\varepsilon} \tag{3}$$

where n is the number of independent variables, $x_k, k = 1, ..., n$. The methodology proposed consists in considering a stochastic sample of RBF predictions $\mathcal{S}$, defined assuming the tuning parameter ε as a stochastic exponent, following a uniform distribution, as per

$$\mathcal{S} = \left\{\hat{f}(\mathbf{x}, \varepsilon); \mathbf{x} \in \mathcal{D}, \varepsilon \sim \text{unif}[\varepsilon_{\min}; \varepsilon_{\max}]\right\} \tag{4}$$

RBF has been widely applied using linear and cubic kernels, corresponding to $\varepsilon = 1$ (polyharmonic spline of first order) and $\varepsilon = 3$ (polyharmonic spline of third order), respectively (Gutmann 2001; Forrester and Keane 2009). This suggests the range of ε to be defined within $\varepsilon_{\min} = 1$ and $\varepsilon_{\max} = 3$. Note that the choice of the distribution for ε is arbitrary and, from a Bayesian viewpoint, this represents the degree of belief in the definition of the tuning parameter.

The prediction provided by the metamodel is given at each $\mathbf{x}$ by the EV of $\hat{f}$ over ε:

$$\hat{f}(\mathbf{x}) = \text{EV}[\hat{f}(\mathbf{x}, \varepsilon)] \tag{5}$$

which is solved by UQ, using (39) with $\xi = \varepsilon$. The metamodel stochastic uncertainty $U_{\hat{f}}(\mathbf{x})$, is quantified at each $\mathbf{x}$ by the 95%-confidence band of $\hat{f}(\mathbf{x})$, using UQ as per (38) and (41). Equations (39) and (41) are solved by Monte Carlo (MC) method, with random sample $\{\varepsilon_i\}_{i=1}^{N_\varepsilon} \sim \text{unif}[1; 3]$.

2.1 Auto-tuning

RBF capability in functions approximation is known to be sensitive to curvature and non-linearities. Hence, the following approaches are proposed to scale the variables domain:

1. *Non-adaptive*. Each independent variable is scaled to the interval [0; 1].
2. *Adaptive*. Each independent variable is scaled according to the curvature of the function.

The latter is performed with the introduction of a scaling factor resulting from the evaluation of the curvature. The main idea is that of having a maximum second derivative having the same value for all variables. Accordingly, adaptive scaling is performed only when $n \geq 2$. An analytical expression of the second derivative is available from (1); however, the function is strongly influenced by the local behavior of the metamodel and the quality of the approxi-

mation deteriorates with higher derivatives. Hence, a finite differences method is applied hereafter. Accordingly, the RBF kernel is defined as

$$\varphi||\mathbf{x}-\mathbf{x}_i|| = \left[\sqrt{\sum_{k=1}^{n} c_k^2 (x_k - x_{k,i})^2}\right]^{\varepsilon} \tag{6}$$

with c_k given by

$$c_k = \left(\sqrt{\max\left|\frac{\partial^2 \hat{f}}{\partial x_k^2}\right|} - 1\right) r_f + 1 \tag{7}$$

where r_f is a relaxation factor. Herein, the identification of the maximum value of the second derivative in (7) is performed using a deterministic particle swarm optimization algorithm (Campana et al. 2009) over $\mathbf{x}$.

Note that if a non-adaptive normalization of the independent variables is used, c_k becomes

$$c_k = \frac{1}{\max\{x_k\} - \min\{x_k\}} \tag{8}$$

2.2 Adaptive sampling

An initial training set $\mathcal{T}$ is built by evaluating the function at $M_0 = 2n + 1$ points: one training point is set at the center of the domain and the other training points are set at the center of each boundary hyper-face. Predictions are made available as EV of $\hat{f}$ (5) with related metamodel stochastic uncertainty $U_{\hat{f}}$ (38). Additional training points are placed where the metamodel stochastic uncertainty is largest:

$$\mathbf{x}_{M+1} = \text{argmax}[U_{\hat{f}}(\mathbf{x})] \tag{9}$$

Equation 9 is used to increase the size of $\mathcal{T}$ and update the metamodel iteratively, as shown in Fig. 1. For the solution of (9), the same deterministic particle swarm optimization used for (7) is applied.

It may be noted that the approach proposed for adaptive sampling is similar to considering the maximum mean square error (MMSE) in Kriging (Sacks et al. 1989). This can be extended taking into account a global metric, by integration of $U_{\hat{f}}$ over the domain, similarly to Kriging's integrated mean square error (IMSE, Sacks et al. 1989). Herein, the maximum value of $U_{\hat{f}}$ through (9) is preferred, since relatively easy to use and implement, leaving alternative sampling criteria for future work.

2.3 Parallel infill

The parallel infill is performed by applying sequentially (9) using a group of I dummy predictions $\hat{f}$ as follows.

Step 1. For $i = 1 : I$, do
identify $\mathbf{x}_{M+i}$ as per (9);
predict $\hat{f}(\mathbf{x}_{M+i})$ using the metamodel;
add $[\mathbf{x}_{M+i}; \hat{f}(\mathbf{x}_{M+i})]$ to the training set $\mathcal{T}$.
End
Step 2. For $i = 1 : I$, do (in parallel)
evaluate the function $f(\mathbf{x}_{M+i})$;
add $[\mathbf{x}_{M+i}; f(\mathbf{x}_{M+i})]$ to the training set $\mathcal{T}$.
End

(10)

This method makes an effective use of parallel computing resources; however, its accuracy is affected by the number of training points per group, I, and may not be as high as using a purely sequential scheme, $I = 1$ (Forrester and Keane 2009).

2.4 Multiple response criteria

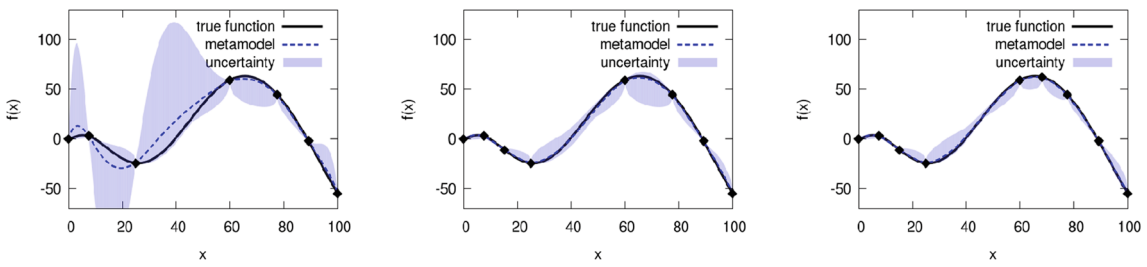
In order to perform adaptive sampling when high-fidelity simulations provide multiple responses, two criteria are formulated and applied:

U_{ave}: new training points are placed based on the maximum average value of the uncertainty among the functions:

$$\mathbf{x}_{M+1} = \text{argmax}[\bar{U}_{\hat{f}}(\mathbf{x})] \tag{11}$$

where

$$\bar{U}_{\hat{f}}(\mathbf{x}) = \frac{1}{m}\sum_{j=1}^{m} U_{\hat{f}_j}(\mathbf{x}) R_j^{-1} \tag{12}$$



(a) $M = 7$ (b) $M = 8$ (c) $M = 9$

Fig. 1 DRBF model: uncertainty-based adaptive sampling

where $U_{\hat{f}_j}$ indicates the stochastic uncertainty of the j^{th} function, $R_j = \max\{f_j\} - \min\{f_j\}$, and m is the number of responses.

U_{max}: new training points are defined, based on maximum absolute uncertainty among the functions:

$$\mathbf{x}_j = \text{argmax}[U_{\hat{f}_j}(\mathbf{x})] \tag{13}$$

$$\mathbf{x}_{M+1} = \text{argmax}[U_{\hat{f}_j}(\mathbf{x}_j) R_j^{-1}] \tag{14}$$

Note that in the case of multiple responses, (2) may be rewritten in the compact form:

$$\mathbf{AW} = \mathbf{Y} \tag{15}$$

where $\mathbf{W} = [\mathbf{w}_1 | \ldots | \mathbf{w}_m]$ and $\mathbf{Y} = [\mathbf{y}_1 | \ldots | \mathbf{y}_m]$, with subscripts indicating different output responses. The above system of equations may be solved at once, taking advantage of a single factorization of the matrix $\mathbf{A}$.

2.5 Evaluation metrics

In order to investigate the effectiveness of metamodels, function predictions and metamodel-based results are systematically validated. For each training set, the fitting error is computed as the normalized error between predictions and benchmark values:

$$E(\mathbf{x}) = \frac{[f(\mathbf{x}) - \hat{f}(\mathbf{x})]}{\max\{f\} - \min\{f\}} \tag{16}$$

For a given validation set $\mathcal{V} = \{\mathbf{x}_i\}_{i=1}^{P}$, the normalized root mean square error is given by

$$E_{\text{RMS}} = \sqrt{\frac{1}{P}\sum_{i=1}^{P} E(\mathbf{x}_i)^2} \tag{17}$$

For current UQ applications, metamodel-based estimators are evaluated following Mousaviraad et al. (2013) and the error for EV, SD and CDF are defined by:

$$E_{\text{EV}} = \frac{\text{EV} - \text{EV}^V}{\text{EV}^V} \tag{18}$$

$$E_{\text{SD}} = \frac{\text{SD} - \text{SD}^V}{\text{SD}^V} \tag{19}$$

$$E_{\text{CDF}} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left[\text{CDF}^V(y_k) - \text{CDF}(y_k)\right]^2} \tag{20}$$

where EV, SD and CDF are evaluated using (39-41) (Appendix A) substituting ξ with $\mathbf{x}$. Superscript V indicates validation values, obtained using f instead of metamodel predictions $\hat{f}$.

Finally, the average UQ error is defined as

$$E_{\text{UQ}} = \frac{|E_{\text{EV}}| + |E_{\text{SD}}| + E_{\text{CDF}}}{3} \tag{21}$$

When multiple output functions are assessed, $f_j, j = 1, \ldots, m$, all errors in (17-21) are studied by their average among multiple responses:

$$\bar{E}_X = \frac{1}{m}\sum_{j=1}^{m} E_{X,j} \tag{22}$$

where $E_{X,j}$ is respectively E_{RMS}, E_{EV}, E_{SD}, E_{CDF}, and finally E_{UQ} for output function f_j. $\bar{E}_{\text{RMS}}$ and $\bar{E}_{\text{UQ}}$ are used to provide an overall assessment of accuracy with focus on function fitting and UQ, respectively; the convergence of such parameters versus the training set size M provides an insight of the metamodel efficiency. Thus, the number M required to achieve specified errors is introduced as an important metric, providing the computational cost. 5, 2.5 and 1.25% error levels are considered, since comparable to the typical uncertainty of CFD outputs due to iterative grid and time step convergence:

$$\begin{array}{ll} M_{5\%} & \text{the minimum number of training points required to achieve } \bar{E}_X < 5\% \\ M_{2.5\%} & \text{the minimum number of training points required to achieve } \bar{E}_X < 2.5\% \\ M_{1.25\%} & \text{the minimum number of training points required to achieve } \bar{E}_X < 1.25\% \\ M_{\text{ave}} & \text{average of the above, } M_{\text{ave}} = (M_{5\%} + M_{2.5\%} + M_{1.25\%})/3 \end{array} \tag{23}$$

where M_{ave} is used as an overall index for metamodels efficiency.

3 Static metamodels and dynamic Kriging used for comparison

3.1 Static metamodels

The following static metamodels are used for comparison: k^{th} order inverse distance weighting, IDW (Shepard 1968); radial basis function network with multiquadric and inverse multiquadric kernels, RBF MQ/IMQ (Buhmann 2003); k^{th} order polyharmonic spline, PHS (Wahba 1990); least-square support vector machine with multiquadric and inverse multiquadric kernels, LS-SVM MQ/IMQ (Suykens et al. 2002); ordinary Kriging with linear and exponential correlation functions, OKG lin./exp. (Peri 2009); an implementation of stochastic RBF without adaptive scaling and sampling, implementing power-law, multiquadric and inverse multiquadric kernels, SRBF P/MQ/IMQ; and DKG without adaptive sampling. A summary of static techniques used is given in Table 1.

Table 1 Summary of the static metamodels used for comparison

Technique	Acronym	Kernel	Tuning parameter	
			Stochastic-speed	Stochastic-wave
Inverse distance weighting	IDW	$1/r^k$	k = 2, 4, 6	$k = 2, 4, 6$
Radial basis function network	RBF	$\sqrt{1+(\alpha\, r)^2}$ (MQ)	$\alpha = 10; 100; 1,000$	$\alpha = 1; 10; 100$
		$[1+(\alpha\, r)^2]^{-0.5}$ (IMQ)	$\alpha = 5; 10; 15$	$\alpha = 0.75; 1.0; 1.25$
Polyharmonic Spline	PHS	r^k, odd k; $r^k \log(r)$, even k	$k = 1, 2, 3$	$k = 1, 2, 3$
Least-square support vector machine	LS-SVM	$\sqrt{1+(\alpha\, r)^2}$ (MQ)	$\alpha = 10; 100; 1,000$	$\alpha = 1; 10; 100$
		$[1+(\alpha\, r)^2]^{-0.5}$ (IMQ)	$\alpha = 5; 10; 15$	$\alpha = 0.75; 1.0; 1.25$
Ordinary Kriging	OKG	$1-(\alpha\, r)$ (lin.)	not applied	$\alpha = 0.25; 0.5; 1$
		$\exp(-\alpha\, r)$ (exp-)	not applied	$\alpha = 0.25; 0.5; 1$
Stochastic radial basis functions	SRBF	r^ε (P)	not applied	$\varepsilon \in 1;3$
		$\sqrt{1+(\alpha\, r)^2}$ (MQ)	not applied	$\alpha \in [1;100]$
		$[1+(\alpha\, r)^2]^{-0.5}$ (IMQ)	not applied	$\alpha \in [0.75;1.25]$
Dynamic Kriging (without adaptive sampling)	DKG		auto-selected	auto-selected

3.2 Dynamic Kriging

The response $\hat{f}$ is modeled in two parts, namely the mean structure and the stochastic process:

$$\hat{f} = \boldsymbol{\phi}^T \boldsymbol{\beta} + \mathbf{Z} \tag{24}$$

$\mathbf{Z}$ is a Gaussian random process with zero-mean and covariance given by

$$C(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \rho(\boldsymbol{\theta}, \mathbf{x}_1, \mathbf{x}_2) \tag{25}$$

where σ^2 is the process variance, ρ is the spatial correlation function and $\boldsymbol{\theta}$ collects the correlation function parameters. The prediction provided by the Kriging model is

$$\hat{f}(\mathbf{x}) = \boldsymbol{\phi}^T \boldsymbol{\beta} + \mathbf{r}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta}) \tag{26}$$

where $\mathbf{y}$ are the observations at the training points, $\mathbf{y} = \{f(\mathbf{x}_i)\}$, $\mathbf{R} = \{\rho(\boldsymbol{\theta}, \mathbf{x}_i, \mathbf{x}_j)\}$, $\mathbf{r}(\mathbf{x}) = \{\rho(\boldsymbol{\theta}, \mathbf{x}, \mathbf{x}_i\}$, $\mathbf{F} = [\boldsymbol{\phi}(\boldsymbol{\gamma}, \mathbf{x}_1), \boldsymbol{\phi}(\boldsymbol{\gamma}, \mathbf{x}_2), ..., \boldsymbol{\phi}(\boldsymbol{\gamma}, \mathbf{x}_M)]^T$, and $\boldsymbol{\phi}$ are the basis functions with tuning parameters collected in $\boldsymbol{\gamma}$. Using a least-squares estimation, $\boldsymbol{\beta}$ is approximated by $\hat{\boldsymbol{\beta}}$ as

$$\hat{\boldsymbol{\beta}}(\mathbf{x}) = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{y} \tag{27}$$

3.2.1 Auto-tuning

The dynamic Kriging presents an automatic selection of the basis-functions, of the correlation function and of the correlation function parameter. Note that if auto-tuning is not performed the expected value of the prediction and the intrinsic uncertainty stemming from the choice of the basis and correlation functions including $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ may be evaluated by UQ as in Bayesian Kriging. Herein, the mean structure $\mathbf{F}$ is selected minimizing the cross-validation error among ordinary (OKG) and first/second-order universal Kriging (UKG) methods; the best $\rho(\boldsymbol{\theta}, \mathbf{x}_1, \mathbf{x}_2)$ and $\boldsymbol{\theta}$ are identified using the maximum likelihood estimation (MLE, Martin and Simpson 2005). Assuming a Gaussian distribution, the log-likelihood of the model parameters is defined as

$$l = -\frac{M}{2} ln(2\pi\sigma^2) - \frac{1}{2} ln(|\mathbf{R}|) - \frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{F}\boldsymbol{\beta})^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{F}\boldsymbol{\beta}) \tag{28}$$

Computing the derivative with respect to $\boldsymbol{\beta}$ and σ^2 and using the approximation in (27), leads to an estimation of the process variance

$$\hat{\sigma}^2 = \frac{1}{M}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}}) \tag{29}$$

and then of l. The goal is to find the optimal $\boldsymbol{\theta}$ that maximizes the likelihood function. To solve the optimization problem a generalized pattern search (GPS, Torczon 1995) is applied; GPS is sensitive to the initial guess, therefore, a genetic algorithm (GA) is used to get a best-candidate initial guess. The advantage of using derivative-free approaches such as GPS and GA is that gradient information of the log-likelihood is not required and the method has good performances also in the presence of noisy functions. When the training set size is small, MLE could be inaccurate, therefore a penalized-MLE (PMLE) is introduced (Li and Sudjianto 2005). The log-

likelihood function is modified adding a penalty function, such as

$$Q = -\frac{M}{2}ln(2\pi\sigma^2) - \frac{1}{2}ln(|\mathbf{R}|) - \frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{F}\boldsymbol{\beta})^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{F}\boldsymbol{\beta}) - M\sum_{i=1}^{N}\lambda|\theta_i| \tag{30}$$

λ is the penalty function parameter and is identified using a GPS with the cross-validation error.

3.2.2 Adaptive sampling

M_0 initial training points are distributed in the design space using a Latin centroidal Voronoi tassellation (LCVT, Saka et al. 2007) sampling, then the Kriging model is built. Once the metamodel is constructed, it is used to predict the function and to approximately bound the errors in such predictions. The prediction mean square error (MSE) from the Kriging model (Sacks et al. 1989),

$$\mathrm{MSE}[\hat{f}(\mathbf{x})] = \sigma^2\left\{1 - \left[\boldsymbol{\phi}^T(\mathbf{x})\,\mathbf{r}^T(\mathbf{x})\right]\begin{pmatrix}\mathbf{0} & \mathbf{F}^T\\ \mathbf{F} & \mathbf{R}\end{pmatrix}^{-1}\begin{bmatrix}\boldsymbol{\phi}(\mathbf{x})\\ \mathbf{r}(\mathbf{x})\end{bmatrix}\right\} \tag{31}$$

is taken as the metric of accuracy. Larger values of prediction MSE are associated with larger uncertainty in prediction (Booker et al. 1999); since the current DKG has been developed for general purpose, namely to minimize the prediction variance over the domain, the MMSE is used to determine the choice of new training points:

$$\mathbf{x}_{M+1} = \mathrm{argmax}\{\mathrm{MSE}[\hat{f}(\mathbf{x})]\} \tag{32}$$

As alternative metric to MMSE, one could use the integrated value of MSE over the domain, IMSE. This provides a global MSE-based metric and has been used for efficient DoE in several Kriging applications (see, e.g., Buslig et al. 2014). Herein, MMSE is preferred for its straightforward implementation and ease of use, in analogy with the choice made for adaptive sampling with DRBF.

3.2.3 Parallel infill

In the previous section, training points are selected one-at-a-time following a sequential sampling approach. If parallel computing is available, multiple training points by parallel infill may be more convenient than one-at-a-time sampling. However, if multiple training points are selected only by prediction MSE, many training points could be clustered in some small region. This is because test points near the test point with the largest prediction MSE also have large prediction MSE. Therefore, multiple training points need to be distributed considering the distances to existing and new training points. Thus, I new training points are selected based on the following steps:

Step 1. Prediction MSE (PMSE) is calculated using current DoE training points.

Step 2. The test point with the largest PMSE is inserted. Set $i = 1$.

Step 3. If $i = I$, go to Step 5. Otherwise, go to Step 4.

Step 4. The nearest distance $D(\mathbf{x})$ from existing training points is calculated from the test points. The test point with the largest $\mathrm{PMSE}(\mathbf{x}) \cdot D(\mathbf{x})$ is inserted. Set $i = i + 1$. Go to Step 3.

Step 5. At all selected test points, responses are evaluated in parallel.

(33)

It may be noted that a common procedure to implement the parallel infill with Kriging is that of using I dummy predictions, similarly to the approach shown for DRBF in Section 2.3. This approach requires a number of $I-1$ sequential evaluations of the Kriging model and could be computationally expensive, depending on the problem dimension and the value of I. For this reason, herein the use of PMSE together with the nearest distance D from existing training points is preferred, allowing for a good level of accuracy at a reasonable computational cost. Issues connected with the parallel infill of Kriging models are beyond the scope of the current work, and not further addressed.

3.2.4 Multiple response criterion

If multiple responses are assessed, the maximum normalized PMSE among multiple responses is used. Thus, new training points are inserted for more highly non-linear responses, similarly to (13) and (14) for DRBF.

4 Analytical and numerical benchmark for deterministic and stochastic validation problems

4.1 Test functions fitting

Test functions (Lucidi and Piccioni 1989; Ali et al. 2005; Taddy et al. 2009) are used as analytical benchmark and presented in Appendix B. These have dimensionality ranging from two (low-dimensionality) to six (medium-dimensionality), with different degree of non-linearities. Among 2D functions, three are polynomial (two of fourth order and one of sixth order), one is a combination of a fourth order polynomial and a trigonometric function, one is trigonometric. Among 3D functions, one is a fourth order polynomial and one is exponential. 4D functions include

two fourth order polynomial, whereas the 6D function is exponential. The number of independent variables and their bounds are summarized in Table 2.

4.2 Uncertainty quantification problems for high-speed catamaran

The two ship hydrodynamics problems represent extensions of the basic resistance, sinkage and trim and seakeeping (resistance, heave and pitch) deterministic problems to UQ problems. In ship design, resistance and seakeeping are usually evaluated using towing tank tests; however, recently CFD is replacing the build and test approach with SBD, which offers more detailed analysis and innovative optimized designs. Formulation of resistance and seakeeping as UQ problems is relatively new and preparatory to development of stochastic optimization approaches. The present research builds on previous Delft catamaran studies of deterministic single and multiple objective optimizations for resistance (Kandasamy et al. 2013) and resistance and seakeeping (Tahara et al. 2012) and UQ for calm water (Diez et al. 2014) and seakeeping (He et al. 2013) including comparison of several static metamodels.

4.2.1 Uncertainty quantification for high-speed catamaran in calm water with stochastic speed

The first ship hydrodynamics problem is the one-dimensional UQ of the Delft catamaran performances in calm water, presented in (Diez et al. 2014). Main particulars and conditions are shown in Table 3, whereas the hull geometry is presented in Fig. 2. The towing tank model scale (Kandasamy et al. 2013) is assumed for the current problem. The model has two degrees of freedom (it is free to sink and trim). The speed U [m/s] is taken as an operational uncertainty. Accordingly, the Froude number (used as non-dimensional speed, Fr $= U/\sqrt{gL_{pp}}$, with g acceleration of gravity [m/s^2] and L_{pp} length between perpendiculars [m]) is assumed to have a Normal distribution with expected value EV(Fr) = 0.5 and a standard deviation SD(Fr) = 0.05. The distribution is truncated to its 95% confidence interval, which have lower and upper bounds approximately equal to EV(Fr)$\pm$2SD(Fr) respectively.

Numerical benchmark consists in f_i(Fr), EV(f_i), SD(f_i), and CDF(f_i) with $i = 11, ..., 13$. Specifically, $f_{11} = C_T = R_T/0.5\rho U^2 S$ (where R_T is the total resistance [N]; ρ is the water density [kg/m^3]; S is the static wetted area) is the non-dimensional total resistance, $f_{12} = \sigma = z/L_{pp}$ (where z is the stationary sinkage at the center of gravity G [m]) is the non-dimensional sinkage and $f_{13} = \tau$ is the stationary trim angle [rad]. Output functions and input variable bounds are summarized in Table 4. Numerical benchmark values are given by converged MC with Latin hypercube sampling, using $N = 257$ URANS computations. Benchmark functions are shown in Fig. 3.

4.2.2 Uncertainty quantification for high-speed catamaran in stochastic regular wave

The second ship hydrodynamics problem is the two-dimensional UQ of the Delft catamaran performance in regular head waves as presented in He et al. (2013). A full-scale ship is considered and main particulars and conditions are included in Table 3. The ship (Fig. 2) is free to heave and pitch. A design speed corresponding to Fr = 0.5 is assumed and stochastic wave conditions pertain to sea state 6, described by the Bretschneider spectrum. Specifically, the wave period and height, T and H, follow a joint probability density function which depends on the spectrum parameters. The focus of the analysis is on the output variables: x-force coefficient $C_x = -F_x/0.5\rho U^2 S$ (where F_x is the force in x direction), heave $\zeta = z/L_{pp}$

Table 2 Test functions

Function	No. of variables	Variable bounds	Max. training set size, $M_{\max}$	Validation set size, P
f_1 Branin-Hoo	2	$-5 < x_1 < 10; 0 < x_2 < 15$	100	10,000
f_2 Six-hump camel back	2	$-2 < x_1 < 2; -1 < x_2 < 1$	100	10,000
f_3 Rosenbrock	2	$-2 < x_1 < 2; -1.5 < x_2 < 2$	100	10,000
f_4 Quartic	2	$-2 < x_1 < 2; -1 < x_2 < 1$	100	10,000
f_5 Shubert	2	$-4 < x_k < 2, \forall k$	250	10,000
f_6 Rosenbrock	3	$-2 < x_k < 2, \forall k$	150	15,625
f_7 Hartman	3	$0 < x_k < 1, \forall k$	150	15,625
f_8 Rosenbrock	4	$-2 < x_k < 2, \forall k$	300	160,000
f_9 Styblinski-Tang	4	$-5 < x_k < 5, \forall k$	300	160,000
f_{10} Hartman	6	$0 < x_k < 1, \forall k$	500	531,441

Table 3 Delft catamaran main particulars and simulation conditions

Main Particular/condition	Symbol	Value, stochastic-speed	Value, stochastic-wave
Length overall [m]	L_{oa}	3.822	105.4
Length between perpendiculars [m]	L_{pp}	3.627	100.0
Breadth overall	B/L_{pp}	0.313	0.313
Breadth demi-hull	b/L_{pp}	0.080	0.080
Draught at mid-ship	T/L_{pp}	0.050	0.050
Distance between center of demi-hulls	s/L_{pp}	0.234	0.234
Longitudinal center of gravity	L_{cg}/L_{pp}	0.527	0.527
Vertical center of gravity	K_g/L_{pp}	0.074	0.113
Pitch radius of gyration	ρ_y/L_{pp}	not used	0.261
Froude number	Fr	[0.402; 0.598]	0.5
Reynolds number	Re	$1.019\,10^7$	$7.144\,10^6$
Wave period [s]	T	not used	[2.2;17.7]
Wave height [m]	H	not used	[0.5;6.4]

and pitch θ motions. The following global output parameters are assessed: time mean and root mean square deviation from mean (RMS) of output time-histories $\eta(t)$, evaluated as

$$\bar{\eta} = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \eta(t)dt \tag{34}$$

$$\eta_{\text{RMS}} = \left\{\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} [\eta(t) - \bar{\eta}]^2 dt\right\}^{0.5}$$

where $t_2 - t_1 = T$.

Numerical benchmark is defined using fully non-linear irregular wave URANS statistically converged computation, and used to validate UQ methods based on regular wave models. Numerical benchmark to validate metamodel-based UQ is defined using converged Markov-chain MC with $N = 129$ regular wave URANS and consists in $f_i(T, H)$, EV(f_i), SD(f_i), and CDF(f_i) with $i = 14, ..., 19$. Specifically, $f_{14} = \bar{C}_x$, $f_{15} = C_{x,\text{RMS}}$, $f_{16} = \bar{\zeta}$, $f_{17} = \zeta_{\text{RMS}}$, $f_{18} = \bar{\theta}$ and $f_{19} = \theta_{\text{RMS}}$, as summarized in Table 4. Benchmark functions are shown in Fig. 4. These are made available for adaptive sampling by a thin plate spline model.

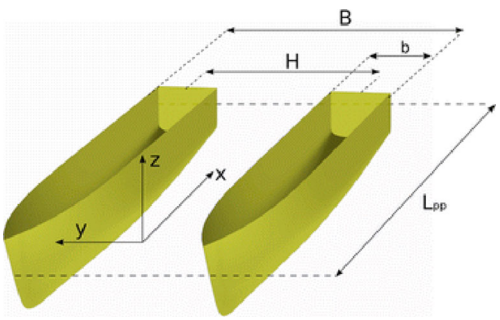


Fig. 2 Delft catamaran geometry and dimensions

5 Numerical results

Test functions results are presented showing: convergence of DRBF MC method for EV and $U_{\hat{f}}$ for f_1; average fitting error $\bar{E}_{\text{RMS}}$, using different relaxation factors ($r_f = 0, 0.25, 0.5, 0.75, 1.0$ as per (7)) and non-adaptive scaling as per (8); breakdown of results among test functions, comparing DRBF and DKG, and effects of sequential sampling (no parallel infill, $I = 1$) and parallel infill using $I = 5$ and $I = 10$ (10). Table 2 shows the maximum training set size $M_{\max}$ and the size of the validation set $\{\mathbf{x}_i\}_{i=1}^{P}$, the latter defined by a regularly distributed Cartesian grid; values of $M_{\max}$ and P are chosen according to the problem dimensionality and the degree of non-linearity.

Catamaran problems are assessed presenting: adaptive sampling and fitting error E; average $\bar{E}_{\text{RMS}}$ and $\bar{E}_{\text{UQ}}$, using different criteria for multiple response and different relaxation factor; breakdown of results among functions, comparing DRBF and DKG; effects of sequential sampling ($I = 1$) and parallel infill ($I = 5$ and 10) and convergence of $\bar{E}_{\text{RMS}}$ and $\bar{E}_{\text{UQ}}$ versus M, with comparison to static metamodels from earlier work. No scaling approach is applied for the stochastic speed case since the problem is one-dimensional, whereas different relaxation factors ($r_f = 0, 0.25, 0.5, 0.75, 1.0$) and non-adaptive scaling are used for the stochastic wave problem. The maximum training set size is set to $M_{\max} = 33$ and 65 (Table 4), according to Diez et al. (2014) and He et al. (2013) for the two catamaran problems, respectively.

5.1 Test functions fitting

Figure 5 shows the convergence of the MC method used to build the DRBF. f_1 is shown as an example, however, the

Table 4 Delft catamaran UQ problems

Problem	Function	No. of variables	Variable bounds	Max. training set size, $M_{\max}$	Validation set size, P
Stochastic-speed	f_{11} resistance, C_T f_{12} sinkage, σ f_{13} trim, τ	1	$0.402 < \text{Fr} < 0.598$	33	128
Stochastic-wave	f_{14} x-force mean, $\bar{C}_x$ f_{15} x-force RMS, $C_{x,\text{RMS}}$ f_{16} heave mean, $\bar{\zeta}$ f_{17} heave RMS, ζ_{RMS} f_{18} pitch mean, $\bar{\theta}$ f_{19} pitch RMS, θ_{RMS}	2	$2.2 < T < 17.7$ $0.5 < H < 6.4$	65	64

plot in Fig. 5 displays the typical behavior of all the problems considered. EV and $U_{\hat{f}}$ are shown for $\mathbf{x} = \mathbf{x}_{M_0+1}$, given by (9), using a number of training points N_ε up to 1,000. Variations of EV are within $\pm 1\%$ of final value for all $N_\varepsilon > 200$ whereas $U_{\hat{f}}$ is within $\pm 1\%$ of its final value for $N_\varepsilon > 800$. Accordingly, hereafter a number of $N_\varepsilon = 1,000$ is selected. The cost of the resulting MC procedure is reasonable for test functions and negligible for UQ, if compared to the cost of CFD simulations.

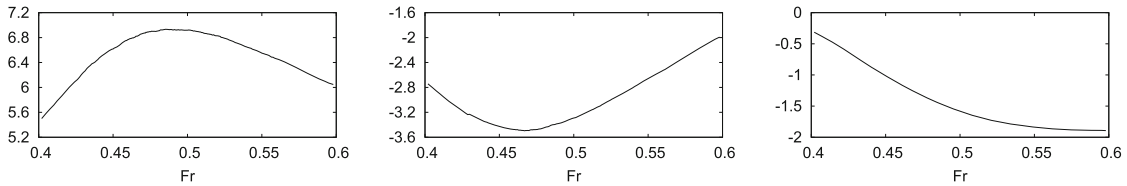
Table 5 shows a comparison among scaling approaches as performed by DRBF, using the average fitting error $\bar{E}_{\text{RMS}}$. Although performance differences are not large, using $r_f = 0.25$ is found the best approach and is used for comparison with DKG. The associated M_{ave} is found 4.7% smaller than the average over all scaling approaches. Non-adaptive domain normalization (8) provides the slowest convergence.

$M_{5\%}$, $M_{2.5\%}$, $M_{1.25\%}$ and M_{ave} for each test function are presented in Fig. 6. A grey bar is used whenever the required training set size exceeds the evaluations budget as per Table 2, and therefore indicates that the specific error level was not achieved. Both metamodels achieve $E_{\text{RMS}} < 1.25\%$ for test functions f_1 (Branin-Hoo, 2D), f_3 (Rosenbrock, 2D), f_4 (Quartic, 2D), f_6 (Rosenbrock, 3D), f_8 (Rosenbrock, 4D) and f_9 (Styblinsky-Tang, 4D). Both DRBF and DKG do not achieve errors $< 2.5\%$ for f_5 (Shubert, 2D). DRBF is not able to achieve errors $< 1.25\%$ for f_2 (Six-hump camelback, 2D) and errors $< 2.5\%$ for f_7 (Hartman, 3D) whereas DKG does not achieve errors $< 5\%$ for f_{10} (Hartman, 6D). Overall, DRBF shows a better performance for f_5 and f_{10}, likely due to the trigonometric and medium-dimensional transcendental nature of the functions, whereas DKG performs better for test functions f_{1-4} and f_{6-9}, providing mostly a sudden convergence. Table 6 shows the average performance of DRBF and DKG using both sequential sampling and parallel infill. For $I = 1$ the metamodels provide similar performance in achieving $\bar{E}_{\text{RMS}} < 5\%$. DKG requires a smaller training set size for achieving $\bar{E}_{\text{RMS}} < 2.5\%$ and 1.25%, and is found the best method on average, as per M_{ave} values. Using $I = 1$ and parallel infill with $I = 5$ and 10 presents similar trends and close results, revealing the high scalability of the approach used. The relative increase in the number of evaluations required with $I = 5$ and 10 compared to $I = 1$ is fairly small.

5.2 Delft catamaran with stochastic speed

5.2.1 Fitting

Figure 7 shows the normalized fitting error E versus Fr. Black diamonds indicate the training set $\{x_i\}_{i=1}^{M}$, with $M = 33$, using DRBF by U_{ave} and $U_{\max}$ and DKG by $U_{\max}$. The error presents similar trends and is within $\pm 1\%$ for all functions and methods. The error peaks are given by DKG for f_{11} (close to 1%) and f_{12} (-1%). Generally, errors



(a) f_{11}: calm water resistance, $C_T 10^3$ (b) f_{12}: sinkage, $\sigma\, 10^3$ (c) f_{13}: trim, τ [deg]

Fig. 3 Stochastic-speed problem: output functions provided by URANS (Diez et al. 2014)

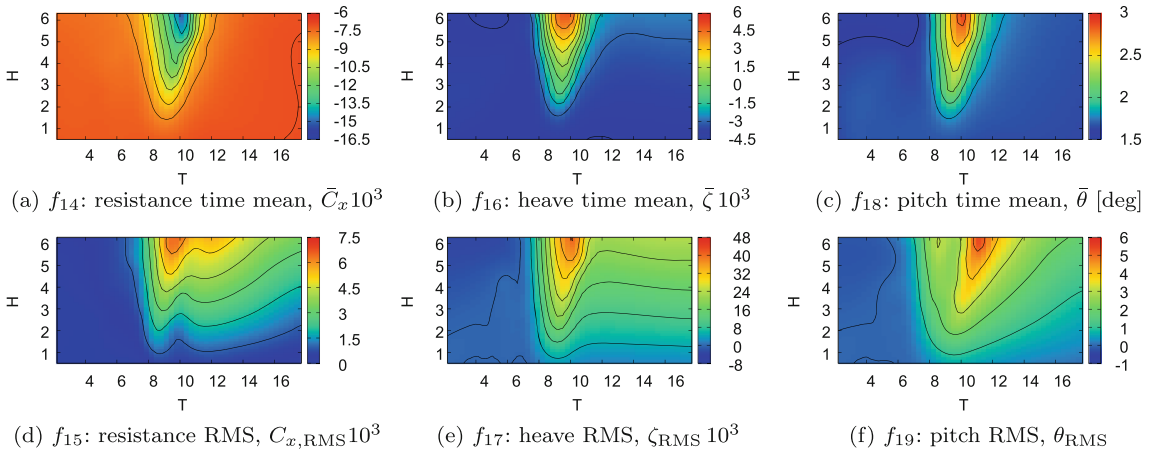


Fig. 4 Stochastic-wave problem: output functions provided by URANS (He et al. 2013)

are larger for f_{11} and f_{12}, and very small for f_{13}. Comparing Figure 7 with Figure 3 shows that DRBF training points are mostly located where the average curvature is greater; training points are found more evenly distributed with DKG.

Figure 8 presents the fitting performance for each function, comparing DRBF and DKG. DRBF sampling by U_{ave} and U_{max} gives the same results in terms of $M_{5\%}$, $M_{2.5\%}$, $M_{1.25\%}$ and M_{ave}. Both metamodels achieve $E_{\mathrm{RMS}} < 1.25\%$ for each function. DRBF shows equal or better performance than DKG for $M_{5\%}$ in f_{11}, f_{12} and f_{13}, and for $M_{2.5\%}$ in f_{11} and f_{12}. Overall, performances do not differ by more than 2 function evaluations. Table 6 shows the average fitting error $\bar{E}_{\mathrm{RMS}}$ over all functions. Using sequential sampling ($I = 1$), metamodels provide very close performances; $M_{\mathrm{ave}} \approx 6$ for both DRBF and DKG with a maximum difference of one function evaluation. Since the number of function evaluations needed to achieve the convergence is less than 7, using the parallel infill with $I = 5$ and $I = 10$ is not needed.

Table 7 shows a summary of the results with comparison to earlier studies based on static metamodels, with $M = 3$, 5, 9, 17, 33 (Diez et al. 2014). Figure 9 (a) shows the convergence of $\bar{E}_{\mathrm{RMS}}$ versus the training set size M. Average error and range of static approaches are shown by diamonds and error bars. DRBF and DKG present similar trends. Table 8 shows the percentile of the average error provided by DRBF and DKG compared to other static metamodels, for $M = 3$, 5, 9, 17 and 33. The error difference with the best metamodel is also shown. Specifically, DRBF is found the best metamodel for $M = 33$, whereas DKG is found the best metamodel for $M = 9$; for $M = 17$ the metamodels provide the same results. For $M \geq 5$ DRBF and DKG presents an error difference with the best metamodel always less than 1.65%. The best static metamodel is, on average, LS-SVM with IMQ kernel and $\alpha = 5$.

5.2.2 Uncertainty quantification

Figure 10 presents the UQ performance for each function comparing DRBF and DKG. DRBF sampling by U_{ave} and U_{max} gives the same results in terms of $M_{5\%}$, $M_{2.5\%}$, $M_{1.25\%}$ and M_{ave}. DRBF is the most effective in achieving $E_{\mathrm{UQ}} < 5\%$ for f_{11}, providing $M_{5\%}$, $M_{2.5\%}$, $M_{1.25\%}$ equal to 4, 6 and 7, whereas DKG shows a sudden convergence with $M_{5\%}$, $M_{2.5\%}$, $M_{1.25\%}$ equal to 5. Both DRBF and DKG have a

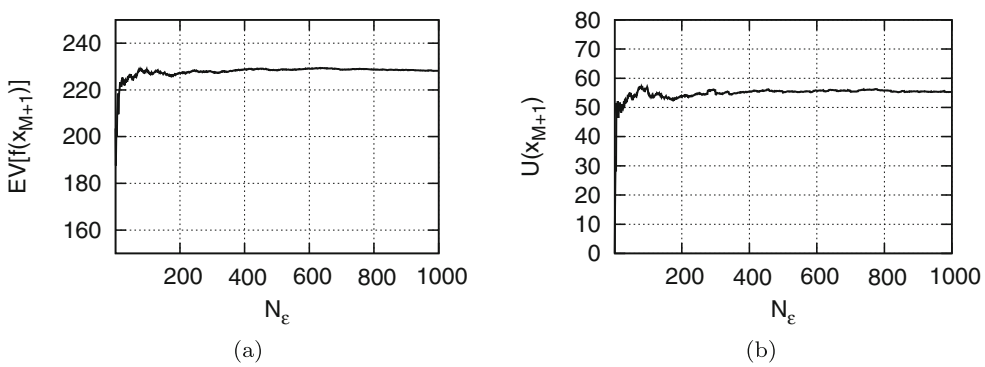


Fig. 5 Convergence of the Monte Carlo method in (39) (**a**) and (41) (**b**) for test function f_1

Table 5 DRBF relaxation factor studies for adaptive scaling

Problem	Error assessed	Multiple-resp. approach	Metric	r_f 0	0.25	0.50	0.75	1	Non-adapt. scaling
Test functions	$\bar{E}_{\text{RMS}}$	n.a.	$M_{5\%}$	106.3	90.9	90.1	89.4	106.1	107.1
			$M_{2.5\%}$	156.6	147.4	153.1	159.2	151.7	162.1
			$M_{1.25\%}$	185.8	178.3	180.1	184	184.3	191.0
			M_{ave}	149.6	138.9	141.1	144.2	147.4	153.4
Stochastic-wave	$\bar{E}_{\text{RMS}}$	U_{ave}	$M_{5\%}$	27	33	23	27	28	59
			$M_{2.5\%}$	46	48	44	44	-	-
			$M_{1.25\%}$	-	-	-	-	-	-
			M_{ave}	46.0	48.7	44.0	45.3	52.7	63.0
		U_{max}	$M_{5\%}$	34	40	38	36	29	54
			$M_{2.5\%}$	41	49	44	53	56	-
			$M_{1.25\%}$	-	-	-	-	-	-
			M_{ave}	46.7	51.3	49.0	51.3	50.0	61.3
	$\bar{E}_{\text{UQ}}$	U_{ave}	$M_{5\%}$	23	19	23	24	28	22
			$M_{2.5\%}$	43	35	43	27	44	51
			$M_{1.25\%}$	60	42	50	-	-	-
			M_{ave}	42.0	32.0	38.7	38.7	45.7	46.0
		U_{max}	$M_{5\%}$	24	21	17	27	28	24
			$M_{2.5\%}$	34	40	42	49	46	38
			$M_{1.25\%}$	50	49	58	62	-	-
			M_{ave}	36.0	36.7	39.0	46.0	46.3	42.3

sudden convergence for f_{12} and f_{13}, with DRBF performing better for f_{12} and DKG for f_{13}. Table 6 shows the average UQ error, $\bar{E}_{\text{UQ}}$, over all functions. Using sequential sampling ($I = 1$) DRBF needs 4, 5 and 7 evaluations to achieve average errors $< 5\%$, 2.5% and 1.25%, respectively. DKG reveals a sudden convergence, showing $M_{5\%}$, $M_{2.5\%}$, $M_{1.25\%}$ equal to 5. Overall performance by M_{ave} are very close for DRBF and DKG. Since the number of function evaluations needed to achieve the convergence is less than 7, using the parallel infill with $I = 5$ and $I = 10$ is not needed.

Finally, Table 7 shows a summary of the results with comparison to static metamodels from Diez et al. (2014). The convergence of the $\bar{E}_{\text{UQ}}$ versus M is depicted in Fig. 9 (b). Average error and range of static approaches are shown by diamonds and error bars. DRBF and DKG present similar trends. Generally, errors are small and few evaluations (< 7) are required to achieve errors $< 1.25\%$. Table 8 shows error percentiles and differences with the best metamodel. Specifically, DRBF is found the best metamodel for $M = 9$ and 17, whereas DKG is found the best metamodel for $M = 33$. For $M \geq 9$ DRBF presents an error difference with the best metamodel always less or equal to 0.04%. For $M \geq 5$ the error difference of DKG with the best metamodel is always less or equal to 0.04%. The best static metamodel is, on average, LS-SVM with MQ kernel and $\alpha = 5$.

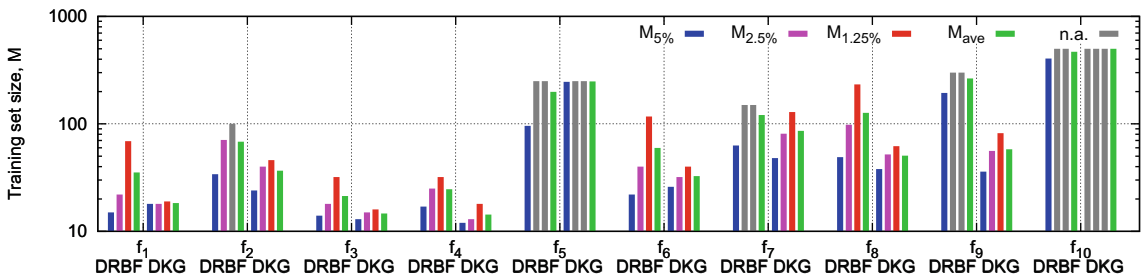


Fig. 6 Test functions: performance of DRBF and DKG for each function

Table 6 Overall performance summary

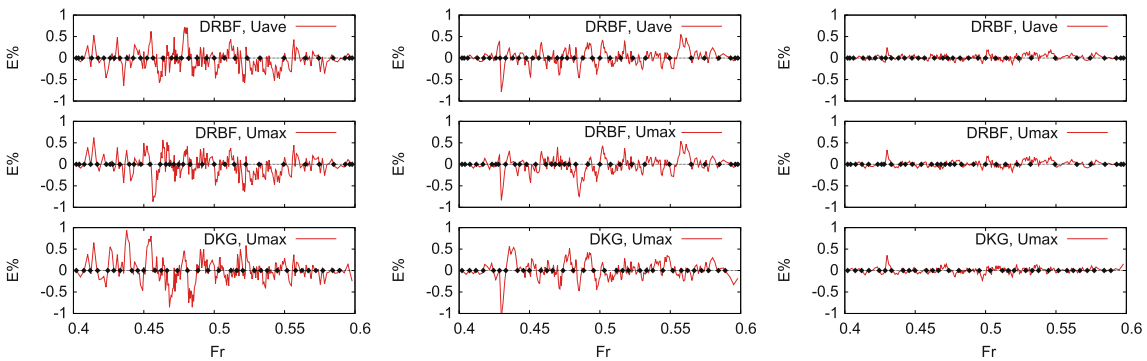
Problem	Error assessed	Metric	r_f	$I=1$		$I=5$		$I=10$	
				DRBF	DKG	DRBF	DKG	DRBF	DKG
Test functions	$\bar{E}_{\text{RMS}}$	$M_{5\%}$	0.25	91	96	100	95	112	99
		$M_{2.5\%}$		147	105	152	109	156	115
		$M_{1.25\%}$		178	116	181	119	176	123
		M_{ave}		138.6	105.6	144.6	107.5	147.8	112.5
Stochastic-speed	$\bar{E}_{\text{RMS}}$	$M_{5\%}$	n.a.	5	6	8	8	13	13
		$M_{2.5\%}$		5	6	8	8	13	13
		$M_{1.25\%}$		7	6	13	8	13	13
		M_{ave}		5.7	6.0	9.7	8.0	13.	13.
	$\bar{E}_{\text{UQ}}$	$M_{5\%}$		4	5	8	8	13	13
		$M_{2.5\%}$		5	5	8	8	13	13
		$M_{1.25\%}$		7	5	13	8	13	13
		M_{ave}		5.3	5.0	9.7	8.0	13.	13.
Stochastic-wave	$\bar{E}_{\text{RMS}}$	$M_{5\%}$	0.5 (U_{ave})	23	28	35	35	35	35
		$M_{2.5\%}$		44	54	45	-	55	-
		$M_{1.25\%}$		-	-	-	-	-	-
		M_{ave}		44.	49.	48.	55.	52.	55.
	$\bar{E}_{\text{UQ}}$	$M_{5\%}$	0.25 (U_{ave})	19	28	20	45	25	25
		$M_{2.5\%}$		35	51	45	45	45	-
		$M_{1.25\%}$		42	61	50	-	55	-
		M_{ave}		32.	47.	38.	52.	42.	52.

5.3 Delft catamaran with stochastic regular wave

5.3.1 Fitting

Figures 11, 12 and 13 show the normalized fitting error E, as a function of T and H for each function, using DRBF by U_{ave} and U_{max} and DKG by U_{max}. Black diamonds indicate the training points used, with $M = 65$. Normalized errors are found greater than in the stochastic speed case and range from -15 to 15%. Errors are generally found larger for DKG than DRBF, especially for f_{14} and f_{19}. Comparing Fig. 4 with Figs. 11, 12 and 13 shows that training points for DRBF are located in those regions where the curvature is higher, whereas they are more uniformly distributed using DKG. DRBF training points by U_{ave} and U_{max} present similar trends.

Table 5 shows a comparison among scaling approaches as performed by DRBF, using the average fitting error



(a) f_{11}, calm water resistance (b) f_{12}, sinkage (c) f_{13}, trim

Fig. 7 Stochastic-speed problem: DRBF and DKG normalized errors, $E\%$

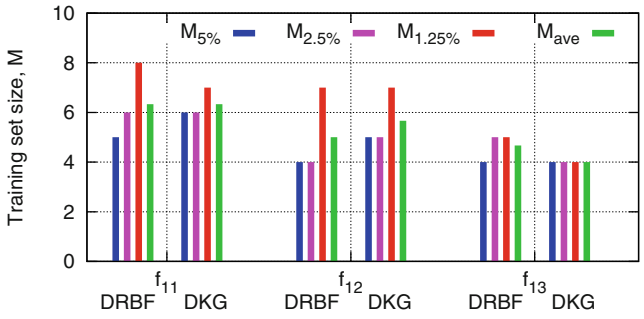


Fig. 8 Stochastic-speed problem: fitting performance of DRBF and DKG for each function

$\bar{E}_{\mathrm{RMS}}$. Both U_{ave} and U_{max} approaches for multiple-response sampling are presented. Sampling by U_{ave} with $r_f = 0.5$ is identified as the best overall approach and is used for comparison with DKG. Associated M_{ave} is found 13.3% smaller than the overall average.

Figure 14 presents E_{RMS} for each function, comparing DRBF and DKG. DRBF is found the most effective metamodel for f_{14}, f_{15}, f_{18} and f_{19}, whereas DRBF and DKG are comparable for f_{16} and f_{17}. Table 6 shows the average fitting performance over all functions, in terms of $\bar{E}_{\mathrm{RMS}}$. Using sequential sampling ($I = 1$) both DRBF and DKG do not achieve an error lower than 1.25%; overall, DRBF is found performing better than DKG requiring on average five evaluations less than DKG. Using sequential and parallel infill with $I = 5$ and 10 presents similar trends, with an increase in the number of evaluations required by the parallel infill with $I = 5$ and $I = 10$, compared to $I = 1$.

Table 9 shows a summary of present results with comparison to static metamodels used in earlier research, with $M = 9, 17, 33, 65$ (He et al. 2013). Fig. 15 (a) shows the convergence of $\bar{E}_{\mathrm{RMS}}$ versus M for DRBF and DKG, with comparison to static metamodels. Average error and range of static approaches are shown by diamonds and error bars. DRBF and DKG outperform static metamodels for $M > 33$. Table 8 shows the percentile of the average error provided by DRBF and DKG compared to other

Table 7 Summary of results for stochastic-speed (average errors are given in %)

Techniques	E_{RMS} M						E_{UQ} M					
	3	5	9	17	33	Average	3	5	9	17	33	Average
Static												
IDW (k =2)	13.3	5.73	3.95	2.92	1.91	5.56	10.4	5.9	3.96	4.02	3.08	5.46
IDW (k =4)	15.5	7.84	4.31	2.39	1.25	6.26	16.5	6.22	2.00	0.66	0.34	5.15
IDW (k =5)	17.1	9.09	5.02	2.78	1.41	7.09	16.8	7.56	2.44	0.81	0.29	5.57
RBF (MQ, α =10)	7.47	0.64	1.23	1.74	1.18	2.45	7.59	0.73	0.73	2.41	0.97	2.48
RBF (MQ, α =100)	12.8	4.74	2.24	1.01	0.47	4.25	9.55	6.22	3.52	2.18	0.34	4.36
RBF (MQ, α =1000)	11.8	3.45	1.17	0.45	0.24	3.42	7.12	3.57	2.09	0.29	0.15	2.64
RBF (IMQ, α =5)	6.77	0.61	0.87	4.78	4.86	3.58	6.56	0.50	0.66	4.16	3.04	2.98
RBF (IMQ, α =10)	7.51	0.94	1.27	5.84	2.69	3.65	6.96	0.64	0.85	4.64	1.98	3.01
RBF (IMQ, α =15)	8.87	1.59	0.97	18.1	1.36	6.18	7.26	1.04	0.72	18.0	1.13	5.62
PHS (k =1)	11.6	3.33	1.07	0.41	0.21	3.33	6.98	3.44	0.83	0.20	0.15	2.32
PHS (k =2)	11.4	3.76	1.46	0.49	0.23	3.46	9.65	5.32	2.58	0.42	0.14	3.62
PHS (k =3)	76.1	17.7	4.26	1.30	0.38	19.9	43.9	12.3	7.35	2.48	0.35	13.3
LS-SVM (MQ, α =10)	6.84	9.23	0.64	0.44	0.56	3.54	6.44	7.41	0.30	0.30	0.37	2.96
LS-SVM (MQ, α =100)	10.2	2.91	1.10	0.44	0.50	3.04	6.95	3.09	2.05	0.30	0.22	2.52
LS-SVM (MQ, α =1000)	11.5	3.24	1.04	0.39	0.20	3.26	6.88	3.36	0.78	0.19	0.10	2.26
LS-SVM (IMQ, α =5)	6.95	1.48	0.88	0.63	0.52	2.09	6.62	1.95	0.67	0.52	0.51	2.05
LS-SVM (IMQ, α =10)	8.49	1.56	0.67	0.39	0.32	2.29	8.32	2.62	0.64	0.20	0.21	2.39
LS-SVM (IMQ, α =15)	9.52	2.58	1.01	0.38	0.25	2.75	8.76	3.62	1.00	0.27	0.16	2.76
Dynamic												
DRBF	13.6	2.03	0.36	0.27	0.17	3.28	11.1	2.43	0.23	0.17	0.13	2.80
DKG	13.0	2.24	0.35	0.27	0.18	3.21	17.5	0.60	0.27	0.21	0.09	3.74

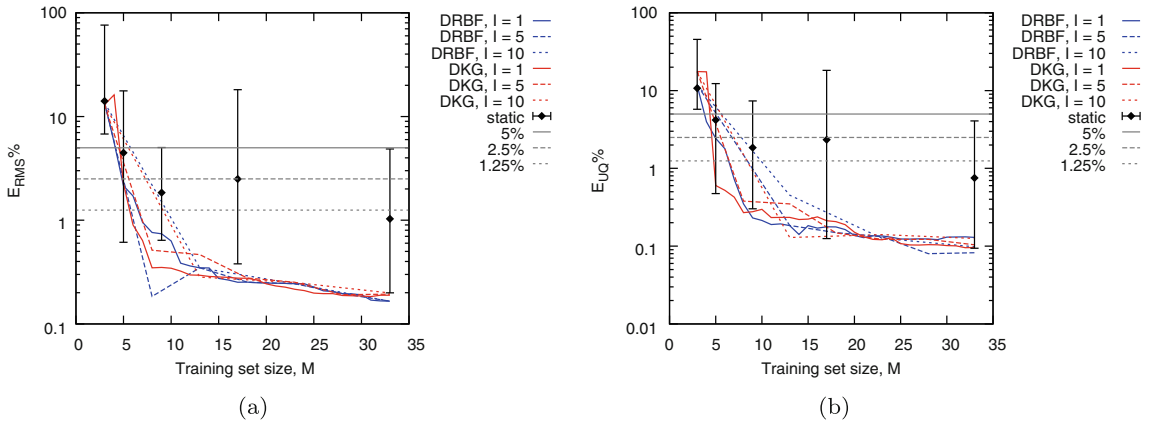


Fig. 9 Stochastic-speed problem: convergence of average fitting (**a**) and UQ (**b**) errors comparing dynamic and static metamodels

Table 8 Comparison of DRBF and DKG performance with static metamodels ($P\%$ refers to the percentile)

	M	DRBF				DKG			
		E_{RMS}		E_{UQ}		E_{RMS}		E_{UQ}	
		$P(\%)$	$E - E_{best}$	$P(\%)$	$E - E_{best}$	$P(\%)$	$E - E_{best}$	$P(\%)$	$E - E_{best}$
Stochastic-speed	3	15.8	6.83	21.1	4.62	21.1	6.25	5.26	11.1
	5	68.4	1.42	68.4	1.93	63.2	1.63	94.7	0.04
	9	94.7	0.01	100	0.00	100	0.00	78.9	0.04
	17	100	0.00	100	0.00	100	0.00	94.7	0.04
	33	100	0.00	89.5	0.04	94.7	0.01	100	0.00
Stochastic-wave	9	10.3	4.17	24.1	4.85	3.44	7.47	20.7	5.19
	17	31.0	1.22	20.7	3.92	0.00	13.0	0.00	15.8
	33	79.3	0.75	82.8	0.33	31.0	1.03	72.4	1.05
	65	100	0.00	100	0.00	96.6	0.62	96.6	0.31

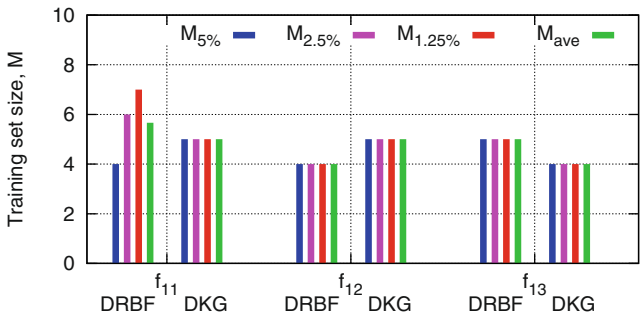


Fig. 10 Stochastic-speed problem: performance of DRBF and DKG for each function

static metamodels, for $M = 9$, 17, 33 and 65, also providing error difference with the best metamodel. Specifically, DRBF is found the best metamodel for $M = 65$. For $M \geq 33$ dynamic metamodels present an error difference with the best metamodel always less than 0.75%. The best static metamodel is, on average, LS-SVM with MQ kernel and $\alpha = 1$.

5.3.2 Uncertainty quantification

Table 5 shows a comparison among scaling approaches as performed by DRBF, using the average UQ error $\bar{E}_{\text{UQ}}$. Both U_{ave} and $U_{\max}$ approaches for multiple-response sampling are presented. Sampling by U_{ave} with $r_f = 0.25$ is identified as the overall best approach and is used for com-

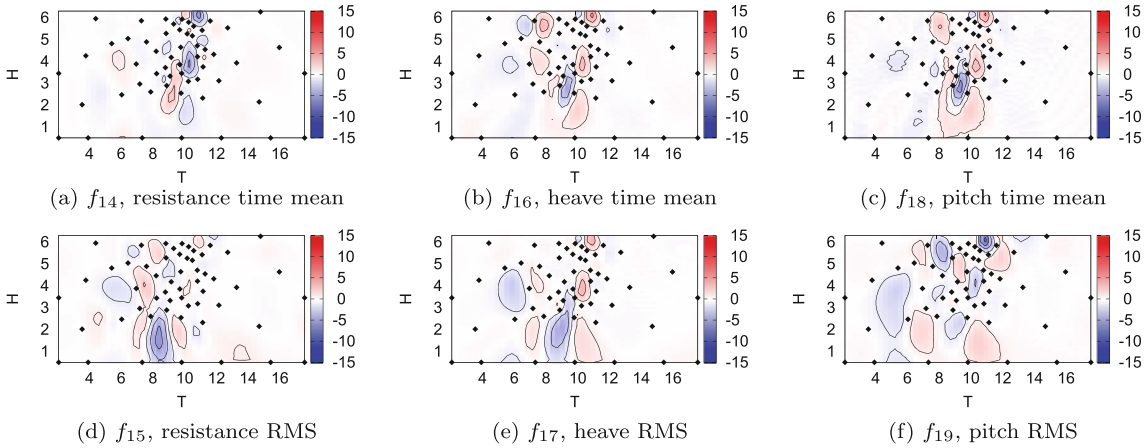


Fig. 11 Stochastic-wave problem: DRBF (U_{ave}) normalized error, $E\%$

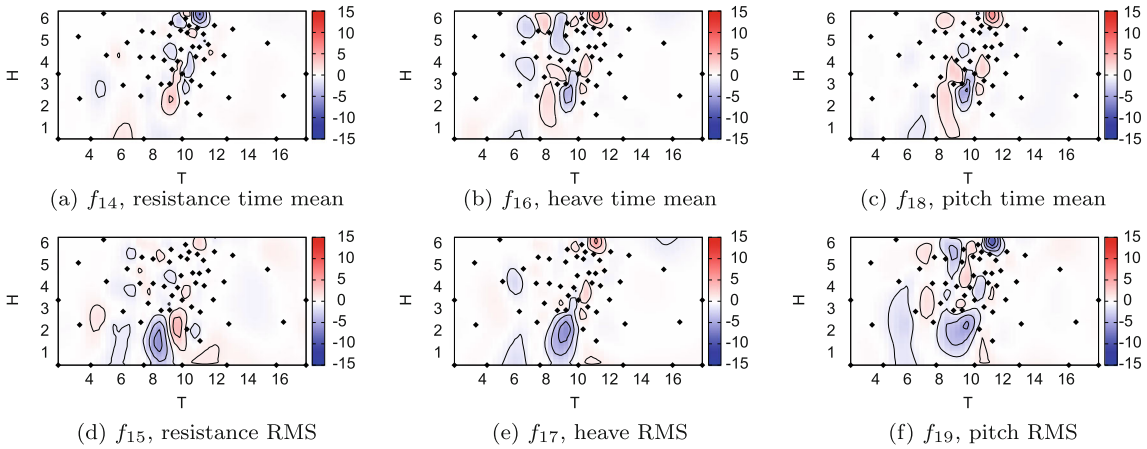


Fig. 12 Stochastic-wave problem: DRBF (U_{max}) normalized error, $E\%$

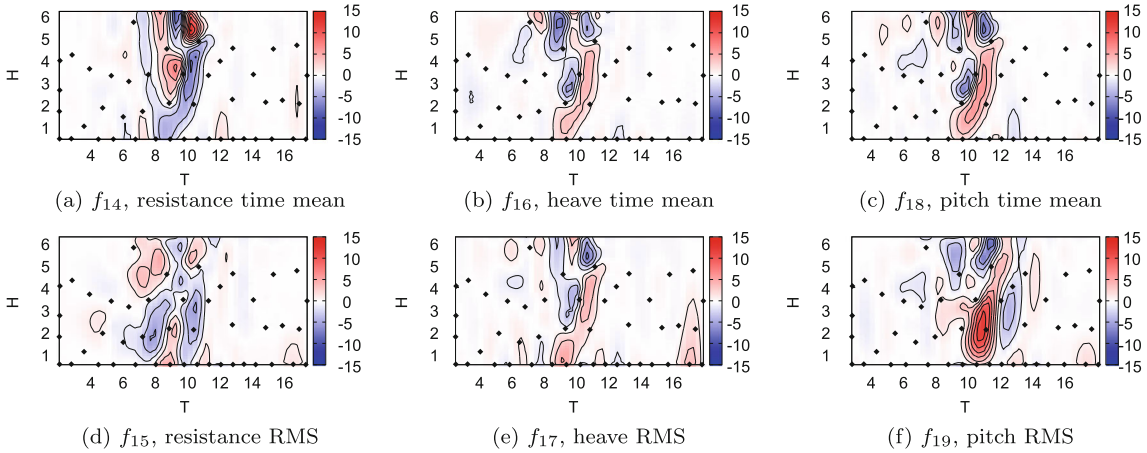


Fig. 13 Stochastic-wave problem: DKG (U_{max}) normalized error, $E\%$

parison with DKG. The associated M_{ave} is found 21.5% smaller than the overall average over different scaling approaches.

Figure 16 presents results for each function. The most effective performance with DRBF is found for f_{14} and f_{15}; DKG provides the best performance for f_{17}.

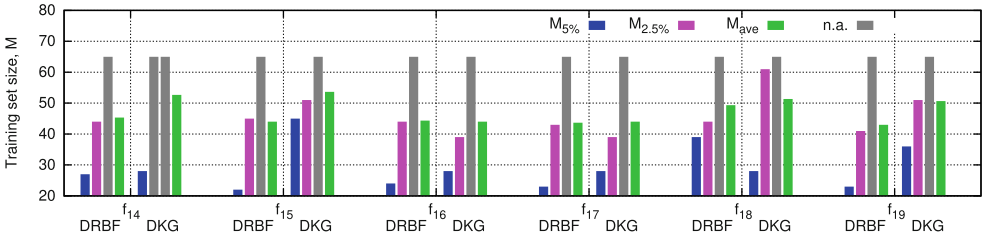


Fig. 14 Stochastic-wave problem: fitting performance of DRBF and DKG for each function

Table 6 shows the average UQ error, $\bar{E}_{\mathrm{UQ}}$, over all functions. Using sequential sampling ($I = 1$) DRBF and DKG present similar trends achieving $\bar{E}_{\mathrm{UQ}} < 1.25\%$; overall, DRBF is found more efficient than DKG requiring on average 14 functions evaluations less than DKG. Using sequential and parallel infill with $I = 5$ and 10, DRBF provides similar trends. DKG is found quite affected by parallel infill, showing different perfor-

Table 9 Summary of results for stochastic-wave (average errors are given in %)

Techniques	E_{RMS} M					E_{UQ} M				
	9	17	33	65	Average	9	17	33	65	Average
Static										
IDW (k =2)	14.0	12.1	11.3	8.96	11.6	17.9	16.4	13.7	7.37	13.8
IDW (k =4)	13.4	10.1	8.43	5.77	9.42	11.7	6.89	4.92	2.12	6.40
IDW (k =6)	14.1	10.5	8.91	6.44	9.97	9.88	6.22	3.87	1.78	5.44
RBF (MQ, α =1)	11.3	7.28	3.99	2.78	6.33	10.3	5.95	2.23	1.54	5.01
RBF (MQ, α =10)	11.5	7.39	4.51	2.99	6.58	12.1	5.51	4.05	1.20	5.71
RBF (MQ, α =100)	11.5	7.49	4.67	3.12	6.70	12.5	5.79	4.29	1.37	5.98
RBF (IMQ, α =0.75)	11.4	8.14	4.43	2.73	7.67	14.0	8.37	4.15	1.34	6.96
RBF (IMQ, α =1.0)	17.7	9.84	5.33	2.86	8.94	16.7	10.5	5.21	2.30	8.66
RBF (IMQ, α =1.25)	20.1	11.6	6.42	3.26	10.4	19.1	12.3	6.19	2.83	10.1
PHS (k =1)	11.6	7.50	4.69	3.14	6.72	12.5	5.82	4.32	1.37	6.01
PHS (k =2)	14.5	7.46	4.13	2.89	7.24	13.2	6.20	2.33	1.28	5.74
PHS (k =3)	18.4	8.43	4.45	2.95	8.55	16.0	6.80	3.34	1.29	6.85
LS-SVM (MQ, α =1)	11.3	7.06	3.96	2.82	6.28	10.1	5.73	2.25	1.43	4.88
LS-SVM (MQ, α =10)	11.3	7.10	4.40	2.93	6.42	12.2	5.41	3.04	1.19	5.46
LS-SVM (MQ, α =100)	11.3	7.19	4.56	3.06	6.54	12.5	5.73	4.19	1.30	5.94
LS-SVM (IMQ, α =0.75)	12.4	7.46	4.19	2.57	6.66	13.1	6.71	3.25	1.31	6.09
LS-SVM (IMQ, α =1.0)	13.1	7.99	4.67	2.62	7.09	14.5	8.48	4.05	1.17	7.06
LS-SVM (IMQ, α =1.25)	13.7	8.64	5.24	2.82	7.60	15.9	10.5	6.08	1.30	8.46
OKG (lin., α =0.25)	11.3	7.20	4.58	3.08	6.55	12. 6	5.75	4.23	1.32	5.96
OKG (lin., α =0.5)	11.3	7.20	4.58	3.08	6.55	12. 6	5.75	4.23	1.32	5.96
OKG (lin., α =1.0)	11.3	7.20	4.58	3.08	6.55	12. 6	5.75	4.23	1.32	5.96
OKG (exp., α =0.25)	11.4	7.20	4.60	3.08	6.56	12.7	5.82	4.26	1.30	6.02
OKG (exp., α =0.5)	11.4	7.20	4.61	3.07	6.57	12.8	5.84	4.27	1.29	6.05
OKG (exp., α =1.0)	11.5	7.20	4.65	3.07	6.60	13.0	6.00	4.37	1.30	6.16
SRBF (P)	11.4	7.30	4.18	2.90	6.45	10.4	5.10	2.43	1.14	4.75
SRBF (MQ)	11.5	7.45	4.60	3.07	6.65	12.3	5.69	4.21	1.29	5.88
SRBF (IMQ)	17.7	9.80	5.33	2.86	8.93	16.6	10.4	5.18	2.31	8.63
DKG	12.9	6.95	3.63	2.29	6.44	10.9	5.90	2.51	1.26	5.15
Dynamic										
DRBF	15.4	8.17	4.38	1.45	7.35	14.7	9.02	2.56	0.78	6.77
DKG	18.7	19.9	4.66	2.07	11.4	15.1	20.9	3.28	1.09	10.1

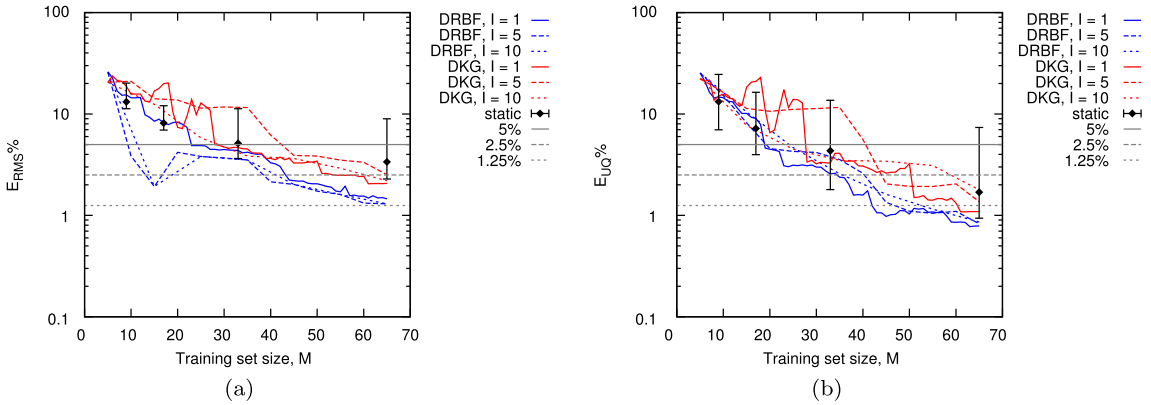


Fig. 15 Stochastic-wave problem: convergence of average fitting (**a**) and UQ (**b**) errors comparing dynamic and static metamodels

mance trends, varying I. Overall efficiency decreases as I increases.

Table 9 shows a summary of the present results with comparison to static metamodels (He et al. 2013). Fig. 15 (b) shows the convergence of $\bar{E}_{\mathrm{UQ}}$ versus M for DRBF and DKG, with comparison to static metamodels. Average errors using static approaches are shown using a diamond, whereas an error bar indicates their range. Table 8 shows the percentile of average error provided by DRBF and DKG compared to other static metamodels and the error difference with the best metamodel. DRBF is found the best metamodel for $M = 33$. For $M \geq 33$ DRBF presents an error difference with the best metmodel always less or equal to 0.3%. For $M \geq 33$ the error difference of DKG with the best metamodel is always less or equal to 1%. The best static metamodel is, on average, SRBF with P kernel.

6 Conclusions and future work

A dynamic metamodel based on stochastic RBF has been derived and validated by comparison with an existing DKG method and static metamodels used in earlier research. A metric for the evaluation of the efficiency of the metamodels has been introduced and applied to both deterministic test functions (with dimensionality ranging from two to six) and ship hydrodynamics UQ problems with analytical and numerical benchmarks, respectively.

Assessing test functions fitting, DRBF is found the most effective for trigonometric and medium dimensional functions, whereas DKG has the best fitting capability when applied to polynomial and low dimensional functions. Overall, average number of training points required, M_{ave}, equals 139 for DRBF and 106 for DKG (Table 6).

Assessment of Delft catamaran performance (total resistance, sinkage and trim) in calm water with stochastic speed reveals that multiple response criterion has no significant effect on DRBF. Relatively few training points are needed by DRBF and DKG for getting small fitting errors. Specifically, M_{ave} for average fitting error $\bar{E}_{\mathrm{RMS}}$ equals 5.66 using DRBF and 6 using DKG (Table 6). Comparing to static metamodels used in earlier research, dynamic approaches are found the most accurate for $M \geq 9$ (Tables 7 and 8). Best static approach has been found on average LS-SVM with IMQ kernel. Also for UQ analyses, DRBF and DKG need few training points to achieve fairly good accuracy. Specifically, M_{ave} for average uncertainty-quantification error $\bar{E}_{\mathrm{UQ}}$ equals 5.33 using DRBF and 5 using DKG (Table 6). Comparing to static

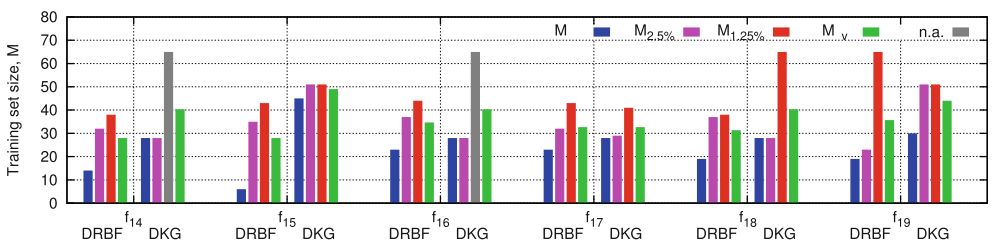


Fig. 16 Stochastic-wave problem: performance of DRBF and DKG for each function

metamodels, dynamic approaches are found the most accurate for $M \geq 9$ (Tables 7 and 8). LS-SVM with IMQ kernel has been identified as the best static approach on average.

Assessment of Delft catamaran performances (time mean and RMS of x-force, heave and pitch motions) in stochastic regular wave shows that DRBF has the most effective performances using multiple response criterion U_{ave}. Both DRBF and DKG have fairly good fitting performance, compared to static metamodels. Specifically, DRBF shows $M_{\text{ave}} = 44$ for average fitting error $\bar{E}_{\text{RMS}}$ whereas $M_{\text{ave}} = 49$ using DKG (Table 6.) Comparing to static metamodels, the error difference between dynamic and best metamodels is always less or equal than 0.75% for $M \geq 33$ (Tables 8 and 9). DRBF is found the best metamodel for $M = 65$. Best static approach has been found on average LS-SVM with MQ kernel. DRBF is found more efficient than DKG for UQ analyses. Both metamodels perform well, compared to static metamodels. Specifically, DRBF presents $M_{\text{ave}} = 32$ for average uncertainty-quantification error $\bar{E}_{\text{UQ}}$, whereas $M_{\text{ave}} = 46.7$ using DKG (Table 6). Comparing to static metamodels, the error difference between dynamic and best metamodels is always less or equal than 1.05% for $M \geq 33$ (Tables 8 and 9). DRBF is found the most accurate overall for $M = 65$. Best static approach has been found on average SRBF with P kernel.

Overall, training points for DRBF are located in high-curvature regions, whereas they are more uniformly distributed using DKG. Since DRBF locates training points where the uncertainty is larger, current results indicate that the metamodel stochastic uncertainty is larger in high-curvature regions. Generally, accuracy of metamodel prediction in high-curvature regions is difficult to achieve and many training points are required. Thus, the metamodel stochastic uncertainty is an effective metric for adaptive sampling. DKG fills more uniformly the domain, which is reasonable since prediction MSE by (31) depends on training-points distribution.

In conclusion, the introduction of a metric based on the number of evaluations required has allowed for a straightforward assessment of metamodels efficiency, which is of primarily interest when high-fidelity solvers are used. DRBF has been found with a greater efficiency for fitting trigonometric and medium-dimensional functions and two-dimensional UQ. The use of an adaptive scaling with r_f ranging from 0 to 0.5 has been found efficient compared to a non-adaptive scaling approach; however, the optimal value of r_f is problem dependent. DKG is found more effective for fitting polynomial and low-dimensional functions and one-dimensional UQ. In general, as the training set size increases, dynamic approaches are found more efficient than static metamodels used in earlier research. In addition, comparison of fitting with uncertainty-quantification errors reveals similar trends. Fitting errors have been found generally larger than those found in uncertainty quantification. Errors for stochastic wave problem (2D) have been found nearly twice than those for stochastic speed problem (1D). Finally, the use of parallel infill with groups of 5 and 10 training points is found affordable, showing an acceptable loss of efficiency compared to the gain in wall-clock time.

Future research includes the use of multiple kernels and their automatic selection in order to auto-configure the DRBF network, and the application of DRBF and DKG to design optimization problems in ship hydrodynamics, including deterministic and stochastic applications with global and local optimization algorithms. The possibility to extend current adaptive sampling approaches, based on maximum uncertainty or MSE, to integral metrics will be addressed in future studies. Applicability of present methodology to large problems (with a number of independent variables greater than 10) will also be investigated in future work. Comparison with Bayesian Kriging methods is of interest and will be evaluated in the future. In addition, model tests campaign to collect experimental benchmark values for UQ for Delft catamaran in wave is in progress.

Acknowledgments The present research is supported by the US Navy Office of Naval Research, Grant N00014-11-1-0237 and Office of Naval Research Global, NICOP Grant N62909-11-1-7011, under the administration of Dr. Ki-Han Kim and Dr. Woei-Min Lin, and by the Italian Flagship Project RITMARE, coordinated by the Italian National Research Council and funded by the Italian Ministry of Education, within the National Research Program 2011-2013.

Appendix A: Uncertainty quantification

UQ studies assess the effects of uncertain parameters ξ with probability density function $\phi(\xi)$ on the relevant outputs f, quantifying EV, SD, CDF and 95%-confidence band of CDF, herein called output stochastic uncertainty U_f, as:

$$\text{EV}(f) = \int (\xi)\phi(\xi)d\xi \tag{35}$$

$$\text{SD}(f) = \sqrt{\int [f(\xi) - EV(f)]^2\phi(\xi)d\xi} \tag{36}$$

$$\text{CDF}(y) = \int H[y - f(\xi)]\phi(\xi)d\xi \tag{37}$$

$$U_f = \text{CDF}^{-1}(0.975) - \text{CDF}^{-1}(0.025) \tag{38}$$

$H(\cdot)$ is the Heaviside step function. Using the Monte Carlo (MC) method with $\{\xi_i\}_{i=1}^{N} \sim \phi$, (35-37) are solved respectively by

$$\mathrm{EV}(f) = \frac{1}{N}\sum_{i=1}^{N} f(\xi_i) \tag{39}$$

$$\mathrm{SD}(f) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left[f(\xi_i) - \mathrm{EV}\right]^2} \tag{40}$$

$$\mathrm{CDF}(y) = \frac{1}{N}\sum_{i=1}^{N} H\left[y - f(\xi_i)\right] \tag{41}$$

Appendix B: Analytical formulation of test functions

This appendix provides the analytical formulation used for the test functions.

Branin-Hoo function (2D)

$$f_1(\mathbf{x}) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos x_1 + 10 \tag{42}$$

Six-hump camelback function (2D)

$$f_2(\mathbf{x}) = \left(4 - 2.1x_1^2 + \frac{1}{3}x_1^4\right)x_1^2 + x_1x_2 + \left(4x_2^2 - 4\right)x_2^2 \tag{43}$$

Rosenbrock function (2D)

$$f_3(\mathbf{x}) = (1 - x_1)^2 + 100\left(x_2 - x_1^2\right)^2 \tag{44}$$

Quartic function (2D)

$$f_4(\mathbf{x}) = \frac{x_1^4}{4} - \frac{x_1^2}{2} + \frac{x_1}{10} + \frac{x_2^2}{2} \tag{45}$$

Shubert function (2D)

$$f_5(\mathbf{x}) = \left\{\sum_{i=1}^{5} i\cos\left[(i+1)x_1 + i\right]\right\}\left\{\sum_{i=1}^{5} i\cos\left[(i+1)x_2 + i\right]\right\} \tag{46}$$

Extended Rosenbrock function (3D)

$$f_6(\mathbf{x}) = \sum_{i=1}^{2}\left[(1 - x_i)^2 + 100(x_{i+1} - x_i^2)^2\right] \tag{47}$$

Hartman function (3D)

$$f_7(\mathbf{x}) = -\sum_{i=1}^{4} a_i \exp\left\{-\sum_{j=1}^{3} b_{ij}(x_j - d_{ij})^2\right\} \tag{48}$$

with

$$\mathbf{a} = \begin{Bmatrix} 1.0 \\ 1.2 \\ 3.0 \\ 3.2 \end{Bmatrix} \quad \mathbf{b} = \begin{Bmatrix} 3.0 & 10.0 & 30.0 \\ 0.1 & 10.0 & 35.0 \\ 3.0 & 10.0 & 30.0 \\ 0.1 & 10.0 & 35.0 \end{Bmatrix} \quad \mathbf{d} = \begin{Bmatrix} 0.3689 & 0.1170 & 0.2673 \\ 0.4699 & 0.4387 & 0.7470 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.03815 & 0.5743 & 0.8828 \end{Bmatrix} \tag{49}$$

Extended Rosenbrock function (4D)

$$f_8(\mathbf{x}) = \sum_{i=1}^{3}\left[(1 - x_i)^2 + 100(x_{i+1} - x_i^2)^2\right] \tag{50}$$

Styblinski-Tang function (4D)

$$f_9(\mathbf{x}) = \frac{\sum_{i=1}^{4} x_i^4 - 16x_i^2 + 5x_i}{2} \tag{51}$$

Hartman function (6D)

$$f_{10}(\mathbf{x}) = -\sum_{i=1}^{4} a_i \exp\left\{-\sum_{j=1}^{6} b_{ij}(x_j - d_{ij})^2\right\} \tag{52}$$

with

$$\mathbf{a} = \begin{Bmatrix} 1.0 \\ 1.2 \\ 3.0 \\ 3.2 \end{Bmatrix} \quad \mathbf{b} = \begin{Bmatrix} 10.0 & 3.0 & 17.0 & 3.5 & 1.7 & 8.0 \\ 0.05 & 10.0 & 17.0 & 0.1 & 8.0 & 14.0 \\ 3.0 & 3.5 & 1.7 & 10.0 & 17.0 & 8.0 \\ 17.0 & 8.0 & 0.05 & 10.0 & 0.1 & 14.0 \end{Bmatrix} \tag{53}$$

$$\mathbf{d} = \begin{Bmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{Bmatrix}$$

References

Acar E, Rais-Rohani M (2009) Ensemble of metamodels with optimized weight factors. Struct Multidisc Optim 37(3):279–294. doi:10.1007/s00158-008-0230-y

Ali MM, Khompatraporn C, Zabinsky Z (2005) A numerical evaluation of several stochastic algorithms on selected continuous global optimization test problems. J Glob Optim 31(4):635–672. doi: 10.1007/s10898-004-9972-2

Billings SA, Zheng GL (1995) Radial basis function network configuration using genetic algorithms. Neural Netw 8(6):877–890. doi:10.1016/0893-6080(95)00029-Y

Booker AJ, Dennis JE, Frank PD, Serafini DB, Torczon V, Trosset MW (1999) A rigorous framework for optimization of expensive functions by surrogates. Struct Optim 17(1):1–13. doi:10.1007/BF01197708

Buhmann MD (2003) Radial basis functions: theory and implementations, vol 12. Cambridge university press

Buslig L, Baccou J, Picheny V (2014) Construction and efficient implementation of adaptive objective-based designs of experiments. Math Geosci 46(3):285–313. doi:10.1007/s11004-013-9481-2

Campana EF, Liuzzi G, Lucidi S, Peri D, Piccialli V, Pinto A (2009) New global optimization methods for ship design problems. Optim Eng 10(4):533–555. doi:10.1007/s11081-009-9085-3

Diez M, He W, Campana EF, Stern F (2014) Uncertainty quantification of delft catamaran resistance, sinkage and trim for variable froude number and geometry using metamodels, quadrature and karhunen–loève expansion. Journal of Marine Science and Technology 19(2):143–169. doi:10.1007/s00773-013-0235-0

Forrester A, Keane A (2009) Recent advances in surrogate-based optimization. Prog Aerosp Sci 45(1-3):50–79. doi:10.1016/j.paerosci.2008.11.001

Giunta AA, McFarland JM, Swiler LP, Eldred MS (2006) The promise and peril of uncertainty quantification using response surface approximations. Struct Infrastruct Eng 2(3-4):175–189. doi:10.1080/15732470600590507

Gramacy RB, Lee HKH (2008) Bayesian treed gaussian process models with an application to computer modeling. J Am Stat Assoc 103(483):1119–1130. doi:10.1198/016214508000000689

Gutmann HM (2001) A radial basis function method for global optimization. J Glob Optim 19(3):201–227. doi:10.1023/A:1011255519438

Hardy RL (1971) Multiquadric equations of topography and other irregular surfaces. J Geophys Res 76(8):1905–1915. doi:10.1029/JB076i008p01905

He W, Diez M, Zou Z, Campana EF, Stern F (2013) Urans study of delft catamaran total/added resistance, motions and slamming loads in head sea including irregular wave and uncertainty quantification for variable regular wave and geometry. Ocean Engineering 74:189–217. doi:10.1016/j.oceaneng.2013.06.020

Jansson T, Nilsson L, Redhe M (2003) Using surrogate models and response surfaces in structural optimization with application to crashworthiness design and sheet metal forming. Struct Multidisc Optim 25(2):129–140. doi:10.1007/s00158-002-0279-y

Jin R, Chen W, Simpson TW (2001) Comparative studies of metamodeling techniques under multiple modelling criteria. Struct Multidisc Optim 23(1):1–13. doi:10.1007/s00158-001-0160-4

Jin R, Du X, Chen W (2003) The use of metamodeling techniques for optimization under uncertainty. Struct Multidiscip Optim 25(2):99–116. doi:10.1007/s00158-002-0277-0

Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. J Glob Optim 13(4):455–492. doi:10.1023/A:1008306431147

Kandasamy M, Peri D, Tahara Y, Wilson W, Miozzi M, Georgiev S, Milanov E, Campana EF, Stern F (2013) Simulation based design optimization of waterjet propelled delft catamaran. Int Shipbuild Prog 60(1):277–308. doi:10.3233/ISP-130098

Kennedy MC, Anderson CW, Conti S, O'Hagan A (2006) Case studies in gaussian process modelling of computer codes. Reliab Eng Syst Saf 91(10–11):1301–1309. doi:10.1016/j.ress.2005.11.028

Li G, Aute V, Azarm S (2010) An accumulative error based adaptive design of experiments for offline metamodeling. Struct Multidisc Optim 40(1–6):137–155. doi:10.1007/s00158-009-0395-z

Li R, Sudjianto A (2005) Analysis of computer experiments using penalized likelihood in gaussian kriging models. Technometrics 47(2):111–120. doi:10.1198/004017004000000671

Loeven GJA, Witteveen JAS, Bijl H (2007) A probabilistic radial basis function approach for uncertainty quantification. Proceedings of the NATO RTO-MP-AVT-147 Computational Uncertainty in Military Vehicle design symposium

Lucidi S, Piccioni M (1989) Random tunneling by means of acceptance-rejection sampling for global optimization. J Optim Theory Appl 62(2):255–277. doi:10.1007/BF00941057

Martin JD, Simpson TW (2005) Use of kriging models to approximate deterministic computer models. AIAA J 43(4):853–863. doi:10.2514/1.8650

Matheron G (1963) Principles of geostatistics. Econ Geol 58(8):1246–1266. doi:10.2113/gsecongeo.58.8.1246

Meng K, Dong ZY, Wong KP (2009) Self-adaptive radial basis function neural network for short-term electricity price forecasting. Generation, Transmission & Distribution. IET 3(4):325–335. doi:10.1049/iet-gtd.2008.0328

Mousaviraad SM, He W, Diez M, Stern F (2013) Framework for convergence and validation of stochastic uncertainty quantification and relationship to deterministic verification and validation. Int J Uncertain Quantif 3(5):371–395. doi:10.1615/Int.J.UncertaintyQuantification.2012003594

Mullur AA, Messac A (2005) Extended radial basis functions: more flexible and effective metamodeling. AIAA J 43(6):1306–1315. doi:10.2514/1.11292

Peri D (2009) Self-learning metamodels for optimization. Ship Technol Res 56:94–108

Pilz J, Spock G (2008) Why do we need and how should we implement bayesian kriging methods. Stoch Environ Res Risk Assess 22(5):621–632. doi:10.1007/s00477-007-0165-7

Regis RG (2011) Stochastic radial basis function algorithms for large-scale optimization involving expensive black-box objective and constraint functions. Comput & Oper Res 38(5):837–853. doi:10.1016/j.cor.2010.09.013

Regis RG, Shoemaker CA (2007) A stochastic radial basis function method for the global optimization of expensive functions. INFORMS J Comput 19(4):497–509. doi:10.1287/ijoc.1060.0182

Sacks J, Welch W, Mitchell T, Wynn H (1989) Design and analysis of computer experiment. Stat Sci 4(4):409–435

Saka Y, Gunzburger M, Burkardt J (2007) Latinized, improved LHS, and CVT point sets in hypercubes. Int J Numer Anal Model 4(3-4):729–743

Sarimveis H, Alexandridis A, Mazarakis S, Bafas G (2004) A new algorithm for developing dynamic radial basis function neural network models based on genetic algorithms. Comput & chem eng 28(1):209–217. doi:10.1016/S0098-1354(03)00169-8

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM national conference, ACM:517–524. doi:10.1145/800186.810616

Sobieszczanski-Sobieski J, Haftka RT (1997) Multidisciplinary aerospace design optimization: survey of recent developments. Struct Optim 14(1):1–23. doi:10.1007/BF01197554

Song H, Choi KK, Lamb D (2013) A study on improving the accuracy of kriging models by using correlation model/mean structure selection and penalized log-likelihood function. 10th World Congress on Structural and Multidisciplinary Optimization. Florida, Orlando

Suykens J, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J (2002). Least squares support vector machines. World Scientific

Taddy MA, Lee HKH, Gray GA, Griffin JD (2009) Bayesian guided pattern search for robust local optimization. Technometrics 51(4):389-401. doi:10.1198/TECH.2009.08007

Tahara Y, Kobayashi H, Kandasamy M, He W, Peri D, Diez M, Campana EF, Stern F (2012) CFD-based multiobjective stochastic optimization of a waterjet propelled high speed ship. 29th Symposium on Naval Hydrodynamics, Gothenburg, Sweden

Torczon V (1995) Pattern search methods for nonlinear optimization, vol 6. SIAG/OPT Views and News, pp 7–11

Wahba G (1990) Spline models for observational data. 59, Siam

Yang RJ, Wang N, Tho CH, Bobineau JP, Wang BP (2005) Metamodeling development for vehicle frontal impact simulation. J Mech Des 127(5):1014–1020. doi:10.1115/1.1906264

Zhao L, Choi KK, Lee I (2011) Metamodeling method using dynamic kriging for design optimization. AIAA J 49(9):2034–2046. doi:10.2514/1.J051017

Zhou XJ, Ma YZ, Li XF (2011) Ensemble of surrogates with recursive arithmetic average. Struct Multidisc Optim 44(5):651–671. doi:10.1007/s00158-011-0655-6