

Titre: Adaptive surrogate modeling for response surface approximations
Title: with application to bayesian inference

Auteurs: Serge Prudhomme, & Corey M. Bryant
Authors:

Date: 2015

Type: Article de revue / Article

Référence: Prudhomme, S., & Bryant, C. M. (2015). Adaptive surrogate modeling for response surface approximations with application to bayesian inference. Advanced Modeling and Simulation in Engineering Sciences, 2(1), 22.
Citation: <https://doi.org/10.1186/s40323-015-0045-5>

Document en libre accès dans PolyPublie

Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/5097/>
PolyPublie URL:

Version: Version officielle de l'éditeur / Published version
Révisé par les pairs / Refereed

Conditions d'utilisation: CC BY
Terms of Use:

Document publié chez l'éditeur officiel

Document issued by the official publisher

Titre de la revue: Advanced Modeling and Simulation in Engineering Sciences (vol. 2, no. 1)
Journal Title:

Maison d'édition: Springer
Publisher:

URL officiel: <https://doi.org/10.1186/s40323-015-0045-5>
Official URL:

Mention légale:
Legal notice:

RESEARCH ARTICLE

Open Access



Adaptive surrogate modeling for response surface approximations with application to bayesian inference

Serge Prudhomme^{1*} and Corey M. Bryant²

*Correspondence:
serge.prudhomme@polymtl.ca
¹ Département de
Mathématiques et de
Génie Industriel, Ecole
Polytechnique de Montréal,
Montréal, QC, Canada
Full list of author information
is available at the end of the
article

Abstract

Parameter estimation for complex models using Bayesian inference is usually a very costly process as it requires a large number of solves of the forward problem. We show here how the construction of adaptive surrogate models using a posteriori error estimates for quantities of interest can significantly reduce the computational cost in problems of statistical inference. As surrogate models provide only approximations of the true solutions of the forward problem, it is nevertheless necessary to control these errors in order to construct an accurate reduced model with respect to the observables utilized in the identification of the model parameters. Effectiveness of the proposed approach is demonstrated on a numerical example dealing with the Spalart–Allmaras model for the simulation of turbulent channel flows. In particular, we illustrate how Bayesian model selection using the adapted surrogate model in place of solving the coupled nonlinear equations leads to the same quality of results while requiring fewer nonlinear PDE solves.

Keywords: Goal-oriented error estimation, Adjoint problem, Turbulence modeling, Spalart–Allmaras model, Parameter identification

Background

A general issue in parameter estimation using Bayesian inference is that one has to sample the forward model a very large number of times in order to obtain accurate posterior distributions of the parameters. In the case of complex models involving many parameters, the process often results in computational costs that far exceed computer resources currently available. Kennedy and O'Hagan [1] suggested to consider emulators of response surfaces in order to reduce the number of forward simulations and showed that it could lead to considerable computational savings. A variety of reduced models to estimate the likelihood function have since then been proposed in the literature [2–5]. However, reduced models only provide approximations of the true solution and the accuracy of the response needs to be verified in order to obtain meaningful results.

We propose in this work to develop a goal-oriented error estimation procedure to adaptively construct reduced models, also referred to as surrogate models, of the Reynolds averaged Navier–Stokes (RANS) models for the simulation of turbulent flows. Goal-oriented error estimation was initially designed as an adjoint-based method to estimate

discretization errors in finite element solutions of boundary-value problems with respect to quantities of interest. The method was later extended to the estimation and control of modeling error when a fine scale model is replaced by a coarse-scale model [6, 7]. The goal-oriented error estimation framework will be used here not only to assess the accuracy of the surrogate models but also to guide the adaptive process for improving the representation of the true response provided by the reduced model.

We illustrate the methodology on examples dealing with the simulation of turbulent channel flows. Turbulence will be modeled here by the Reynolds averaged Navier–Stokes (RANS) equations, supplemented by the Spalart–Allmaras model for the description of the Reynolds stress. This closure model involves several parameters that need to be calibrated in order to be useful. Values of the parameters have been proposed in [8–10]. As the turbulence community is well aware that these parameters may include some level of uncertainty, Bayesian inference has recently been used to quantify uncertainties in simulations of turbulence (see e.g. [10, 11]). In the present study, our objective will be to reproduce some of the numerical examples described in [10] in order to demonstrate that one can confidently use reduced models rather than the full models to estimate the parameters of the Spalart–Allmaras model.

The content of the paper is as follows: we describe in "Methods" the model problem, namely the Reynolds averaged Navier–Stokes equations and the Spalart–Allmaras model for the Reynolds stress, and derive the weak formulation of the deterministic problem. We also recall, briefly, some concepts of probability theory and present the parameterized reduced model. We then review the basic principles for goal-oriented error estimation and adaptivity methodology for the construction of the reduced model as presented in [12, 13]. Lastly, we describe the concepts of Bayesian inference and model selection. In "Results and discussion", we provide numerical results of a Bayesian model selection study, where the full model is replaced with the adapted surrogate model, before providing some concluding remarks.

Methods

Model problem

Let $\Omega \subset \mathbb{R}^3$ be the domain occupying the channel with boundary $\partial\Omega$. The RANS equations are derived from the Navier–Stokes equations using the Reynolds decomposition of the velocity, $\mathbf{u} = \mathbf{U} + \mathbf{u}'$, where $\mathbf{U} = \bar{\mathbf{u}}$ is the time-averaged velocity over a time interval $(0, T)$ and \mathbf{u}' is the fluctuation about the mean. Substituting this decomposition into the Navier–Stokes equations and taking the average over $(0, T)$, we obtain the so-called RANS equations,

$$\begin{aligned} U_j \frac{\partial U_i}{\partial x_j} &= -\frac{1}{\rho} \frac{\partial P}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial U_i}{\partial x_j} - \overline{u'_i u'_j} \right), \quad \text{in } \Omega \\ \frac{\partial U_i}{\partial x_i} &= 0, \quad \text{in } \Omega \end{aligned} \tag{1}$$

where P , ρ , and ν , denote the mean pressure, the density, and the kinematic viscosity, respectively. In order to be able to solve the equations, one usually considers a closure model for the Reynolds stress tensor $r_{ij} := \overline{u'_i u'_j}$ (scaled here by the constant density ρ)

based on the eddy viscosity assumption. That is, the Reynolds stress is expressed as the viscosity term,

$$r_{ij} = -\nu_T \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right), \quad (2)$$

where the “eddy viscosity” ν_T can be written in terms of a turbulent length and time scale. The practice of modeling the turbulence effects as a viscosity is motivated by the fact that turbulence transports momentum in a similar manner to viscosity [14–16].

Many closure models have been proposed based on the eddy viscosity assumption, see e.g. [8, 9, 14, 17–19], some of which have been examined in the study by Oliver and Moser [10]. Here we will focus on one of the most commonly used models, the eddy viscosity transport model of Spalart and Allmaras [9]. The form considered in this work, as well as in [10], has been modified to avoid negative values of turbulent production and to ignore the transition to turbulence from a laminar initial condition; see [8, 20] for more details on the modified form of the model.

Starting from the eddy viscosity assumption (2), the Spalart–Allmaras model introduces a working variable $\tilde{\nu}$ such that

$$\nu_T = \tilde{\nu} f_{\nu 1}, \quad (3)$$

where

$$f_{\nu 1} = \frac{\chi^3}{\chi^3 + c_{\nu 1}^3}, \quad \chi = \frac{\tilde{\nu}}{\nu}. \quad (4)$$

The working variable is taken to be governed by the transport equation

$$\frac{D\tilde{\nu}}{Dt} = c_{b1} \tilde{S} \tilde{\nu} - c_{w1} f_w \left(\frac{\tilde{\nu}}{d} \right)^2 + \frac{1}{\sigma_{SA}} \left[\frac{\partial}{\partial x_j} \left((\nu + \tilde{\nu}) \frac{\partial \tilde{\nu}}{\partial x_j} \right) + c_{b2} \frac{\partial \tilde{\nu}}{\partial x_j} \frac{\partial \tilde{\nu}}{\partial x_j} \right], \quad (5)$$

where d is the distance to the nearest wall and parameter c_{w1} is defined as:

$$c_{w1} = \frac{c_{b1}}{\kappa^2} + \frac{1 + c_{b2}}{\sigma_{SA}}. \quad (6)$$

The remaining undefined terms are given by the following relationships,

$$\tilde{S} = \begin{cases} S + \bar{S}, & \bar{S} \geq -c_{\nu 2} S, \\ S + \frac{S(c_{\nu 2}^2 S + c_{\nu 3} \bar{S})}{(c_{\nu 3} - 2c_{\nu 2})S - \bar{S}}, & \bar{S} < -c_{\nu 2} S, \end{cases} \quad (7)$$

where S is the magnitude of the vorticity, and

$$\bar{S} = \frac{\tilde{\nu}}{\kappa^2 d^2} f_{\nu 2}, \quad f_{\nu 2} = 1 - \frac{\chi}{1 + \chi f_{\nu 1}}, \quad (8)$$

$$f_w = g \left(\frac{1 + c_{w3}^6}{g^6 + c_{w3}^6} \right)^{1/6}, \quad g = r + c_{w2}(r^6 - r), \quad r = \frac{\tilde{\nu}}{\tilde{S} \kappa^2 d^2}. \quad (9)$$

The values of the parameters c_{b1} , σ_{SA} , c_{b2} , κ , c_{w2} , c_{w3} , c_{v1} , c_{v2} , and c_{v3} , suggested by Spalart and Allmaras, are provided in Table 1.

We suppose that our primary goal is the prediction of the centerline velocity in a fully-developed incompressible channel flow at $Re_\tau = 5000$. The turbulence is assumed non-homogeneous in the y -direction (wall normal direction) but homogeneous in the x -direction, reducing the complexity of the RANS equations significantly. Except for the mean pressure gradient in the x -direction, derivatives of statistical variables are all assumed to vanish with respect to x and t ; thus, $U_2 = 0$, and $U_1 = U(y)$ is only a function of y , and the gradient of P can be shown to be constant [14]. To simplify the presentation, we set $1/\rho \partial_x P = 1$ and control the dynamics of the flow purely through the Reynolds number.

Let $D = (0, H)$, where H represents the half height of the channel. Combining the simplified form of the RANS equations with the Spalart–Allmaras turbulence model, the strong form of the equations now reads:

$$1 = \frac{\partial}{\partial y} \left((v + \nu_T) \frac{\partial U}{\partial y} \right), \quad y \in D, \quad (10)$$

$$0 = c_{b1} \tilde{S} \tilde{\nu} - c_{w1} f_w \left(\frac{\tilde{\nu}}{y} \right)^2 + \frac{1}{\sigma_{SA}} \frac{\partial}{\partial y} \left[\left((v + \tilde{\nu}) \frac{\partial \tilde{\nu}}{\partial y} \right) + c_{b2} \left(\frac{\partial \tilde{\nu}}{\partial y} \right)^2 \right], \quad y \in D. \quad (11)$$

Equation (10) represents the RANS momentum equation, which governs the behavior of the flow variable U ; Eq. (11) is the transport equation for the Spalart–Allmaras working variable $\tilde{\nu}$. The equations are supplemented with the boundary conditions:

$$\begin{aligned} U(0) &= 0, & \partial_y U(H) &= 0, \\ \tilde{\nu}(0) &= 0, & \partial_y \tilde{\nu}(H) &= 0, \end{aligned} \quad (12)$$

which amount to symmetry boundary conditions at the center of the channel ($\partial_y U(H) = 0$ and $\partial_y \tilde{\nu}(H) = 0$) and no slip conditions at the walls. We indeed assume that the eddy viscosity is symmetric across the channel and vanishes at the wall.

Table 1 Standard parameter values for the Spalart–Allmaras turbulence model [10]

Parameter	Nominal value
c_{b1}	0.1355
σ_{SA}	2/3
c_{b2}	0.622
κ	0.41
c_{w2}	0.3
c_{w3}	2
c_{v1}	7.1
c_{v2}	0.7
c_{v3}	0.9

The weak formulation of the problem is derived in the typical manner of multiplying Eqs. (10) and (11) by suitable test functions and integrating by parts. Let $\mathcal{V} = V \times V$ where $V = \{v \in H^1(D) \mid v(0) = 0\}$. Then the problem becomes:

$$\begin{aligned} &\text{Find } (U, \tilde{v}) \in \mathcal{V} \text{ such that} \\ &\mathcal{B}((U, \tilde{v}); (v_U, v_{\tilde{v}})) = \mathcal{F}((v_U, v_{\tilde{v}})), \quad \forall (v_U, v_{\tilde{v}}) \in \mathcal{V}, \end{aligned} \quad (13)$$

where

$$\mathcal{F}((v_U, v_{\tilde{v}})) = \int_0^H v_U dy \quad (14)$$

$$\mathcal{B}((U, \tilde{v}); (v_U, v_{\tilde{v}})) = \mathcal{B}_m((U, \tilde{v}); (v_U, v_{\tilde{v}})) + \mathcal{B}_t((U, \tilde{v}); (v_U, v_{\tilde{v}})), \quad (15)$$

with

$$\mathcal{B}_m((U, \tilde{v}); (v_U, v_{\tilde{v}})) = - \int_0^H \left((v + v_T) \frac{\partial U}{\partial y} \right) \frac{\partial v_U}{\partial y} dy, \quad (16)$$

$$\begin{aligned} \mathcal{B}_t((U, \tilde{v}); (v_U, v_{\tilde{v}})) &= \int_0^H \left[c_{b1} \tilde{S} \tilde{v} - c_{w1} f_w \left(\frac{\tilde{v}}{y} \right)^2 \right] v_{\tilde{v}} \\ &\quad - \frac{1}{\sigma_{SA}} \left[\left((v + \tilde{v}) \frac{\partial \tilde{v}}{\partial y} \right) + c_{b2} \left(\frac{\partial \tilde{v}}{\partial y} \right)^2 \right] \frac{\partial v_{\tilde{v}}}{\partial y} dy. \end{aligned} \quad (17)$$

The above equations will be solved using a standard continuous finite element discretization on D . Let $\mathcal{V}^h \subset \mathcal{V}$ be the finite element subspace consisting of piecewise linear functions, on a suitable partition of D with maximal element diameter h . The finite element approximation of (13) is given by,

$$\begin{aligned} &\text{Find } (U^h, \tilde{v}^h) \in \mathcal{V}^h \text{ such that} \\ &\mathcal{B}((U^h, \tilde{v}^h); (v_U, v_{\tilde{v}})) = \mathcal{F}((v_U, v_{\tilde{v}})), \quad \forall (v_U, v_{\tilde{v}}) \in \mathcal{V}^h. \end{aligned} \quad (18)$$

A computable system of equations can then be obtained using Newton's method, linearizing about the approximate state (U^h, \tilde{v}^h) .

Problem (18) represents the model problem we shall consider throughout this work. However, since the RANS turbulence model parameters are uncertain, these equations can be parameterized by random variables; we discuss the characterizations of the uncertain parameters in the following section.

Uncertainty characterization and reduced model

As previously discussed, the parameters of the Spalart–Allmaras turbulence model (3) and (5) are usually assumed constant with the values provided in Table 1. In this work, we suppose that a subset of these parameters are in fact unknown, or random. Therefore, the boundary-value problem can be viewed as parameterized by parameters of the Spalart–Allmaras turbulence model.

We briefly review some relevant concepts of probability theory and the use of polynomial expansions, commonly referred to as generalized polynomial chaos, to represent the effects of uncertainty in the model response.

Let $\{\Theta, \Sigma, P\}$ be a probability space, where Θ is the sample space of random events, Σ is a σ -algebra, and P is the probability measure on Σ , meaning $P(\Theta) = 1$. A random variable on the probability space is defined as a P -measurable function of Θ . Let $\Xi \subset \mathbb{R}^n$. We use the notation $\xi : \Theta \rightarrow \Xi$ to denote a random variable and denote by p_ξ the associated probability density function.

Introducing the notion of random variables into the boundary-value problem implies that the solution U^h is a random process. Many authors have proposed the use of generalized polynomial chaos to construct a representation of U^h as an expansion in terms of polynomials ξ [21–27]. Of course, to ensure equality these expansions may require an infinite number of terms; to make the solution computationally feasible the expansion is truncated to produce a reduced model for U^h . To make the presentation as general as possible, we assume that the truncated expansion is given by

$$U^{h,N}(\mathbf{x}, \xi) = \sum_{i=0}^N U_i^h(\mathbf{x}) \Phi_i(\xi), \quad (19)$$

where $U_i^h(\mathbf{x}) \in V^h \subset V, \forall i = 1, \dots, N$ and each $\Phi_i(\xi)$ represents a basis function for a subset of $L^2(\Xi)$.

A number of methods for computing the coefficients U_i^h of the expansion have been developed in the uncertainty quantification literature and generally fall into two categories: intrusive and non-intrusive. Non-intrusive approaches attempt to estimate the coefficients of a generalized polynomial chaos expansion by solving the deterministic problem at a set of realizations of ξ . As a result, existing simulation codes can usually be used directly. In contrast, intrusive approaches, such as those based on Galerkin methods, solve a system of equations for the entire set of expansion coefficients; typically this requires the use of specially designed solvers. We shall focus on non-intrusive approaches here since they allow for the existing turbulence simulation codes to be used with minimal modification. Moreover, the set of independent parameter values can usually be run in parallel, making non-intrusive approaches more efficient.

As discussed previously, non-intrusive approaches aim to compute the coefficients of the expansion based on independent realizations of the deterministic solution $U^h(y, \xi)$. One can use sampling-based methods, such as Monte Carlo or Latin hypercube sampling, to compute projections onto the polynomial bases Φ . In high-dimensional parameter spaces sampling may be preferred, since convergence is based on the number of samples and not the dimension of the space. In our case, the dimension will remain relatively low so that we can rely on direct numerical integration using quadrature techniques. Even in a high number of dimensions authors have proposed the use of sparse representations to make the process more efficient [27–33].

Here we utilize the pseudo-spectral projection method [34]. Consider the differential equations to be parameterized by ξ . Let \mathcal{V}^h be as defined above; then, Problem (18) can be rewritten as,

$$\begin{aligned} &\text{Find } (U^h(\cdot, \xi), \tilde{v}^h(\cdot, \xi)) \in \mathcal{V}^h \text{ such that} \\ &\mathcal{B}_\xi \left((U^h, \tilde{v}^h); (v_u, v_{\tilde{v}}) \right) = \mathcal{F}_\xi((v_u, v_{\tilde{v}})), \quad \forall (v_u, v_{\tilde{v}}) \in \mathcal{V}^h, \end{aligned} \quad (20)$$

where \mathcal{B} and \mathcal{F} are as defined in (15) and (14), respectively; the subscript ξ is used to indicate their dependence on the parameter. Solutions of (20) at sampled ξ are independent and can be used to construct a discrete reduced model $(U^{h,N}, \tilde{v}^{h,N})$. The accuracy of the model is controlled by h , the level of discretization of the physical domain, as well as N , the maximum order of polynomials used in the expansion.

In an effort to produce more accurate surrogate models goal-oriented error estimation techniques, common to the finite element community, have recently been extended to address problems with uncertainty [27, 35, 36]. The following section outlines the adaptive approach developed in [12, 13].

Goal-oriented adaptive surrogate modeling

In this section, we provide details of the error estimation and adaptive procedure for boundary-value problems parameterized by uncertainty that was developed in [12, 13]. We review the extension of goal-oriented error estimation to the case of uncertainty as proposed in previous works [12, 13, 34–38]. In addition we suggest an adaptive procedure based on the contribution of higher-order expansion terms to the error in the quantity of interest.

Goal-oriented error estimation

Since its introduction in the 1990s [39–42], goal-oriented error estimation has grown in popularity in the finite element community. Here we only provide a brief outline of the approach and refer the interested reader to more extensive descriptions and reviews of the methodology [42–45].

Equation (20) is in fact a deterministic problem and the standard goal-oriented error estimation framework can be applied for a fixed ξ ; the effect of variability in ξ can then be handled using the expansion techniques of the previous section as we will show below.

In order to derive the error estimation and adaptive strategy, we first require the definition of a linear functional of the solution representing a quantity of interest. We will use the average of the mean flow velocity U over the channel cross-section,

$$\mathcal{Q}_\xi((U, \tilde{v})) = \int_0^H U \, dy, \quad (21)$$

where the subscript ξ is again used to indicate that the solution U depends on the value of ξ and thus so does \mathcal{Q} . Since the velocity profile is expected to reach the maximum at the center of the channel, this quantity of interest is expected to be more sensitive to the centerline velocity.

The core ingredient of the goal-oriented framework is the so-called adjoint problem, which seeks a generalized Green's function associated with the quantity of interest. In the case of a nonlinear operator, such as (20), the linearized operator is used to define

the adjoint equation [41, 43, 45–47]. For our model problem that means that the adjoint equation is given by,

$$\begin{aligned} &\text{Find } (z_U(\cdot, \xi), z_{\tilde{v}}(\cdot, \xi)) \in \mathcal{V} \text{ such that} \\ &\mathcal{B}'_{\xi}((U, \tilde{v}); (z_U, z_{\tilde{v}}), (v_U, v_{\tilde{v}})) = \mathcal{Q}_{\xi}((v_U, v_{\tilde{v}})), \quad \forall (v_U, v_{\tilde{v}}) \in \mathcal{V}, \end{aligned} \quad (22)$$

where the operator \mathcal{B}'_{ξ} is used to compute updates in Newton's method for solving the nonlinear primal problem (18). The definition of the adjoint equation allows one to establish a computable estimate for the error in the quantity of interest.

In the traditional deterministic setting one would be concerned with the error $\mathcal{Q}((U, \tilde{v})) - \mathcal{Q}((U^h, \tilde{v}^h))$, here our approximate solution is represented by the reduced model $(U^{h,N}, \tilde{v}^{h,N})$. Since $(U^{h,N}(\cdot, \xi), \tilde{v}^{h,N}(\cdot, \xi)) \in \mathcal{V}^h$ for sample ξ , the adjoint equation (22) holds and we can proceed in much the same way as is done for finite element approximations [41, 48]. Using the definition of the residual

$$\mathcal{R}_{\xi}((U^{h,N}, \tilde{v}^{h,N}); (z_U, z_{\tilde{v}})) := \mathcal{F}_{\xi}((z_U, z_{\tilde{v}})) - \mathcal{B}_{\xi}((U^{h,N}, \tilde{v}^{h,N}); (z_U, z_{\tilde{v}})), \quad (23)$$

we can write, introducing the errors in the solution as $e_U = U - U^{h,N}$ and $e_{\tilde{v}} = \tilde{v} - \tilde{v}^{h,N}$,

$$\mathcal{E}_{\mathcal{Q}} = \mathcal{Q}_{\xi}((U, \tilde{v})) - \mathcal{Q}_{\xi}((U^{h,N}, \tilde{v}^{h,N})) = \mathcal{Q}_{\xi}((e_U, e_{\tilde{v}})) \quad (24)$$

$$= \mathcal{B}'_{\xi}((U, \tilde{v}); (z_U, z_{\tilde{v}}), (e_U, e_{\tilde{v}})) \quad (25)$$

$$= \mathcal{R}_{\xi}((U^{h,N}, \tilde{v}^{h,N}); (z_U, z_{\tilde{v}})) - \Delta_{\mathcal{B}}((U, \tilde{v}), (e_U, e_{\tilde{v}}), (z_U, z_{\tilde{v}})) \quad (26)$$

$$= \mathcal{R}_{\xi}((U^{h,N}, \tilde{v}^{h,N}); (z_U - \phi_U^h, z_{\tilde{v}} - \phi_{\tilde{v}}^h)) - \Delta_{\mathcal{B}}((U, \tilde{v}), (e_U, e_{\tilde{v}}), (z_U, z_{\tilde{v}})), \quad (27)$$

for all $(\phi_U^h, \phi_{\tilde{v}}^h) \in V^h$. Note that Eq. (24) holds because of the linearity of \mathcal{Q} , that Eq. (25) follows from using the adjoint problem, that Eq. (26) is a consequence of the definition of the linearized operator \mathcal{B}' and of the residual (23), and that Eq. (27) is a result of Galerkin orthogonality. Due to the complexity of estimating $\Delta_{\mathcal{B}}$, and the fact that it is deemed to be higher-order, the term is often neglected [41, 49, 50]; we proceed assuming that the contribution due to linearization is negligible.

Furthermore, the representation (27) cannot be used directly as it involves the exact adjoint solution $(z_U, z_{\tilde{v}})$. To obtain a computable estimate of the error in the quantity of interest, we instead introduce an approximate adjoint solution from an enriched finite element space \mathcal{V}^+ where $\mathcal{V}^h \subset \mathcal{V}^+ \subset \mathcal{V}$. That is,

$$\begin{aligned} &\text{Find } (z_U^+, z_{\tilde{v}}^+) \in \mathcal{V}^+ \text{ such that} \\ &\mathcal{B}'_{\xi}((U^{h,N}, \tilde{v}^{h,N}); (z_U^+, z_{\tilde{v}}^+), (v_U, v_{\tilde{v}})) = \mathcal{Q}_{\xi}((v_U, v_{\tilde{v}})), \quad \forall (v_U, v_{\tilde{v}}) \in \mathcal{V}^+, \end{aligned} \quad (28)$$

where the enriched space will be taken as the space of piecewise quadratic functions on the same partition of D as V^h . Thus, a computable error estimate for the error in the quantity of interest at a specified value of ξ is provided by

$$\begin{aligned} \mathcal{E}_Q &= \mathcal{Q}_\xi((U, \tilde{v})) - \mathcal{Q}_\xi(U^{h,N}, \tilde{v}^{h,N}) \\ &\approx \mathcal{R}_\xi(U^{h,N}, \tilde{v}^{h,N}; (z_U^+ - \phi_U^h, z_v^+ - \phi_v^h)) := \eta_\xi(U^{h,N}, \tilde{v}^{h,N}, (z_U^+, z_v^+)). \end{aligned} \quad (29)$$

In addition, η can be broken into elementwise contributions to define refinement indicators for mesh adaptation.

To measure the error in the response of the quantity of interest over the range of parameters ξ , we use the L_2 norm of the error,

$$\left\| \mathcal{Q}_\xi((U, \tilde{v})) - \mathcal{Q}_\xi(U^{h,N}, \tilde{v}^{h,N}) \right\|_{L_2(\Xi)}. \quad (30)$$

This metric is a natural choice as it implies the control of the mean and standard deviation of the error [46].

The key to obtaining a tractable estimator for (30) is the evaluation of the norm itself. Since evaluating the norm involves integration in a possibly high-dimensional parameter space we will likely need to evaluate the error representation in a manner that scales exponentially with the parameter dimension. At each point in parameter space, in order to evaluate the error estimate (29), one requires the solution of the adjoint equation (28), a process that would quickly become prohibitively expensive. Instead, we propose to construct a reduced model of the semi-discrete adjoint solution, which can then be evaluated with minimal cost at any point in parameter space. Let

$$\hat{z}^{+,N}_\alpha(y, \xi) = \sum_{\alpha=0}^N \hat{z}^+_\alpha(y) \Psi_\alpha(\xi), \quad (31)$$

where $z^+_\alpha(y) \in V^+ \subset V$, $\forall \alpha = 1, \dots, N$ can be computed in the same way as the forward solution. With this additional modification, we then have a computable estimate for the error (30). Substituting the error estimator (29) into the norm for (30) we obtain a computable error estimator for the L_2 norm on Ξ ,

$$\eta = \left\| \mathcal{R}_\xi(U^{h,N}, \tilde{v}^{h,N}; (\hat{z}^{+,N}_U, \hat{z}^{+,N}_v)) \right\|_{L^2(\Xi)}. \quad (32)$$

The accuracy of the estimator can be shown to depend quadratically on the errors e_U and e_v [12, 13, 46].

Theorem 2.1 *Assume B_ξ to be continuously differentiable, for any $\xi \in \Xi$, in a subset of \mathcal{V} that contains (U, \tilde{v}) , (U^h, \tilde{v}^h) , $(U^{h,N}, \tilde{v}^{h,N})$. Let $(\hat{z}^{+,N}_U, \hat{z}^{+,N}_v)$ be an approximation of the adjoint solution according to (31). The error estimate (32) for (30) satisfies the bound,*

$$\begin{aligned}
& \left| \left\| \mathcal{Q}((U, \tilde{v})) - \mathcal{Q}(U^{h,N}, \tilde{v}^{h,N}) \right\|_{L^2(\Xi)} - \eta \right| \\
& \leq \mathcal{O} \left(\left\| (e_U^{h,N}, e_v^{h,N}) \right\|_V \left\| (\hat{z}_U - \hat{z}_U^{+,N}, \hat{z}_v - \hat{z}_v^{+,N}) \right\|_V \right)_{L^2(\Xi)} \\
& \quad + \mathcal{O} \left(\left\| (e_U^{h,N}, e_v^{h,N}) \right\|_V^2 \right)_{L^2(\Xi)}. \tag{33}
\end{aligned}$$

The extension of the goal-oriented error estimation framework to boundary-value problems with uncertain data simply means that uncertainty can be considered in complex physical problems using a reduced-order surrogate model since its accuracy can be assessed in terms of the quantity of interest. Even more valuable, much as in the case of deterministic goal-oriented error estimation, estimates can be used to drive adaptivity to improve the approximations further, as we discuss in the next section.

Adapting the surrogate model

One major disadvantage of using quadrature-based sampling strategies for surrogate construction is that the number of points required for a fixed expansion order grows exponentially with parameter dimension; this is commonly referred to as the *curse of dimensionality*.

A number of different approaches have been proposed to minimize the effect of the curse of dimensionality. For example, one may choose to alter the quadrature formula used in calculating the expansion coefficients; in the case a tensor product quadrature formula is employed, authors have suggested sparse quadrature grids that can reduce the number of evaluations necessary for the same level of accuracy [3, 27, 31–33, 51]. Alternatively, or in a combined manner, higher-order information can be used to identify the more influential parameters. Instead of increasing the expansion order uniformly (isotropic refinement), one may then choose to improve the surrogate model by adding basis functions that only correspond to the most influential parameters (anisotropic refinement).

While our proposed error estimates can be used with a sparse quadrature formula, we will restrict our discussion to full tensor product quadrature for simplicity. Also the advantage of using an anisotropic refinement strategy is more easily observed in this setting. In contrast to existing techniques for using higher-order information to drive anisotropic surrogate refinement, which are often based on heuristic measures, our approach identifies the components associated with the error in the quantity of interest.

In order to identify the most influential modes in the expansion, we use the residual in the error estimate (32). To do so, we must be able to identify the parameters to which the residual is most sensitive. We use a surrogate model for the residual itself

$$\mathcal{E}(\xi) := \sum_{\alpha=0}^M \mathcal{R}_\alpha \Psi_\alpha(\xi) \approx \mathcal{R}_\xi \left((U^{h,N}, \tilde{v}^{h,N}); (\hat{z}_U^{+,N}, \hat{z}_v^{+,N}) \right), \tag{34}$$

with a higher order $M > N$. The coefficients in (34) are calculated based on the same type of quadrature used for the primal and adjoint surrogates,

$$\mathcal{R}_\alpha = \sum_{k=1}^m \mathcal{R}_{\xi^k} \left((U^{h,N}, \bar{v}^{h,N}); (\hat{z}_U^{+,N}, \hat{z}_v^{+,N}) \right) \Psi_\alpha(\xi^k) w^k. \quad (35)$$

Although the forward and adjoint solutions are of order N , the residual is not restricted to be of order N . The higher-order coefficients in (34) capture additional information about the error in the quantity of interest. The relative magnitudes of the coefficients in \mathcal{E} provide a weighting of the most important parameter directions or, more specifically, which basis functions should be added to the solution expansions. To precisely define the refinement strategy, we need a generalization of the set of multi-indices; let

$$\mathcal{I}_\mathcal{N} = \{\alpha \in \mathbb{N}^n : \alpha_j \leq N_j, j = 1, \dots, n\}, \quad (36)$$

where $\mathcal{N} = (N_1, \dots, N_n)$ represents the maximum polynomial degree in each direction. Higher-order expansions of the error will now be obtained using $\mathcal{M} = \mathcal{N} + \mathbf{1} = \{N_1 + 1, \dots, N_n + 1\}$. Thus, we seek the coefficients of $\mathcal{E}(\xi)$ in (34) in the set $\mathcal{I}_\mathcal{M} \setminus \mathcal{I}_\mathcal{N}$ with the largest magnitude, which will be added to the index set $\mathcal{I}_\mathcal{N}$ for the subsequent adaptive step. Algorithm 1 describes the detailed refinement strategy.

Algorithm 1: Anisotropic p refinement in Ξ .

```

1 Construct  $\mathcal{E}(\xi) = \sum_{\alpha \in \mathcal{I}_\mathcal{M}} \mathcal{R}_\xi(u_\alpha^{h,N}, \hat{z}_\alpha^{+,N}) \Psi_\alpha(\xi)$  ;
2 Set  $\beta^* = \sum_{\alpha \in \mathcal{I}_\mathcal{M} \setminus \mathcal{I}_\mathcal{N}} |\mathcal{R}_\xi(u_\alpha^{h,N}, \hat{z}_\alpha^{+,N})|$  ;
3 For  $\alpha \in \mathcal{I}_\mathcal{M} \setminus \mathcal{I}_\mathcal{N}$ , sort  $|\mathcal{R}_\xi(u_\alpha^{h,N}, \hat{z}_\alpha^{+,N})|$  in descending order giving index  $\mathcal{I}_\alpha$  ;
4 Given  $0 \leq \eta \leq 1$ , set  $\beta = 0$  and  $i = 0$  ;
5 while  $\beta < \eta \beta^*$  do
6   Set  $i = i + 1$  and  $\alpha^* = \mathcal{I}_\alpha(i)$  ;
7   Set  $\beta = \beta + |\mathcal{R}_\xi(u_{\alpha^*}^{h,N}, \hat{z}_{\alpha^*}^{+,N})|$  ;
8   for  $j = 1 \rightarrow n$  do
9     if  $\alpha_j^* > N_j$  then
10      Increase polynomial order of approximation in component  $j$ ,  $N_j \leftarrow N_j + 1$  ;
11    end
12  end
13 end
```

Bayesian inference and model selection

To this point we have restricted the discussion to the construction of a reduced model. Often the full model is yet still in need of validation against experimental observations. A reduced model can certainly make validation studies more computationally feasible, the more pressing question is whether using the surrogate model in place of the full model will lead to the same conclusions on the model's validity.

Bayesian inference requires a large number of model simulations; replacing the full simulation with an accurate surrogate model may lead to considerable computational savings. We first provide an overview of the general Bayesian methodology and then discuss how it can be used in a model selection procedure.

Bayes' theorem is a fundamental result of probability. Relatively recently, it has been adapted toward parameter identification for complex mathematical models. The advantage of Bayesian inference for model calibration is that it provides for a distribution of probable parameter values instead of the one best fitting parameter value obtained from traditional optimization procedures. Bayesian parameter identification can be interpreted as an update of the degree of belief in the parameters.

The solution of the Bayesian calibration procedure is the posterior pdf, or the conditional distribution of the model parameters given the observed data. Let $\mathbf{q} \in \mathbb{R}^n$ represent the vector of calibration data, or observations, and let $\xi \in \Xi = \mathbb{R}^m$ be the random variable representing the model parameters we wish to calibrate. The prior distribution of the parameters is denoted by $p(\xi)$ and encapsulates the prior knowledge one has about the parameters independent of the calibration data. Bayes' theorem then states that the posterior distribution, $p(\xi|\mathbf{q})$ is proportional to the prior times the likelihood $L(\xi, \mathbf{q})$ of observing the data [52, 53],

$$p(\xi|\mathbf{q}) = \frac{L(\xi|\mathbf{q}) p(\xi)}{p(\mathbf{q})}. \quad (37)$$

More specifically, the likelihood is defined by the conditional distribution of the data as a function of the parameters $L(\xi|\mathbf{q}) = p(\mathbf{q}|\xi)$, but to emphasize the dependence on the value of the parameters it is often written in the former notation. The denominator in Bayes' theorem acts as a normalization constant and using the law of total probability can be expressed as,

$$p(\mathbf{q}) = \int_{\Xi} p(\mathbf{q}|\xi) p(\xi) d\xi. \quad (38)$$

Perhaps the most critical component of the Bayesian framework is the likelihood function. Ideally the likelihood is determined by the measurement process, or any other process contributing to uncertainty in the calibration data. For example, if the measurement error is additive, meaning the observations take the form,

$$\mathbf{q} = M(\xi) + \epsilon, \quad (39)$$

where $M(\xi)$ is the predicted value of \mathbf{q} using the model with parameters ξ and $\epsilon \in \mathbb{R}^n$ is the error model with distribution p_ϵ , then the likelihood is given by

$$L(\xi|\mathbf{q}) = p_\epsilon(\mathbf{q} - M(\xi)). \quad (40)$$

In practice, while one might have a decent estimate of the uncertainty in measurements, it is often difficult to fully characterize the distribution of experimental uncertainty. For this reason, it can be beneficial to use a model selection procedure to determine the best choice of uncertainty model.

We rely here on a Bayesian model selection procedure that compares a set of models to decide which uncertainty description most likely matches the data. We will only consider the methodology to select the most probable model among a class of uncertainty models, but one can just as well employ the technique to decide between models governing the physical response of a system, such as different RANS closure models, or a combination of physical and uncertainty models [10].

In the Bayesian model comparison procedure, the respective models are evaluated on the basis of the model posterior plausibility. Given a set of models $\mathcal{M} = \{M_1, M_2, \dots, M_n\}$, model plausibility quantifies the relative probability with which model M_i actually generates the observed data [54]. Again, we let \mathbf{q} represent the calibration data and employ Bayes' theorem to define the plausibility for model M_i by

$$p(M_i|\mathbf{q}, \mathcal{M}) = C p(\mathbf{q}|M_i, \mathcal{M}) p(M_i|\mathcal{M}), \quad (41)$$

where C is a normalization constant similar to that in (37), $p(\mathbf{q}|M_i, \mathcal{M})$ acts like a likelihood for model M_i , and $p(M_i|\mathcal{M})$ is the prior model plausibility. The likelihood of the model is given by the evidence, which is simply the normalization constant discussed in the previous section, conditioned on model M_i ,

$$E(M_i|\mathbf{q}, \mathcal{M}) := p(\mathbf{q}|M_i, \mathcal{M}) = \int_{\Xi} p(\mathbf{q}|\xi, M_i, \mathcal{M}) p(\xi|M_i, \mathcal{M}) d\xi, \quad (42)$$

and it measures the probability of observing data \mathbf{q} given the model M_i . The evidence is used to compare models relative to one another and to identify the model that is most likely capable of reproducing the data. In regards to the prior plausibility, one often chooses a uniform plausibility across the collection of models; if all models are equally likely candidates, then the natural choice for the prior is simply $p(M_i|\mathcal{M}) = 1/n$ for each model.

We have explicitly left \mathcal{M} in the conditional distributions above since the whole process is strictly conditional on the original set of models. In other words, only the models included in the set \mathcal{M} are evaluated, thus any conclusions or observations resulting from the quantitative analysis are limited by the quality of the models under consideration; the process can not be used to identify a *truth* model unless it is present in \mathcal{M} .

Results and discussion

In this section we present two sets of results. First, we illustrate the adaptive algorithm for the construction of a reduced model for the mean velocity of a turbulent flow in a channel. Secondly, we demonstrate that, as a result of the error estimation and adaptation, the surrogate model can be used in place of the full model for a Bayesian model selection procedure, providing a more computationally efficient method for performing uncertainty quantification in complex simulations.

Adapted surrogate model for turbulence model problem

We apply here the adaptive surrogate refinement procedure to the model problem of RANS turbulence modeling for incompressible flows. We will consider the physical discretization as fixed and focus on the adaptive construction of a surrogate model for the Spalart–Allmaras turbulence model with six uncertain parameters: κ , c_{b1} , σ_{SA} , c_{b2} , c_{v1} , c_{w2} . To evaluate the surrogate model, we will use both the error estimates reviewed in "Goal-oriented error estimation" as well as simulations of the quantity of interest provided by the full model to further establish the accuracy of our error estimates.

The prior distribution for all uncertain parameters will be taken to be uniform. The exact descriptions will be based on the nominal values presented in Table 1 with a range from 50 to 150 corresponding value; for example $\kappa \sim \mathcal{U}(0.205, 0.615)$.

Starting with a constant, or $N = 0$, surrogate model, we performed 17 adaptive steps adding additional polynomials to the expansion according to Algorithm 1. Figure 1 shows the convergence of the error estimate in terms of the number of nonlinear PDE solves. Compared to uniform, or isotropic p -refinement, the anisotropic refinement of the surrogate model leads to significant improvement of the error for an equal number of forward model evaluations, roughly two orders of magnitude reduction. The progression of the expansion order is shown in Table 2. We observe that the initial refinements

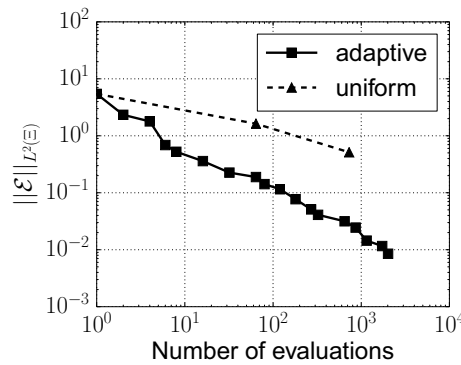


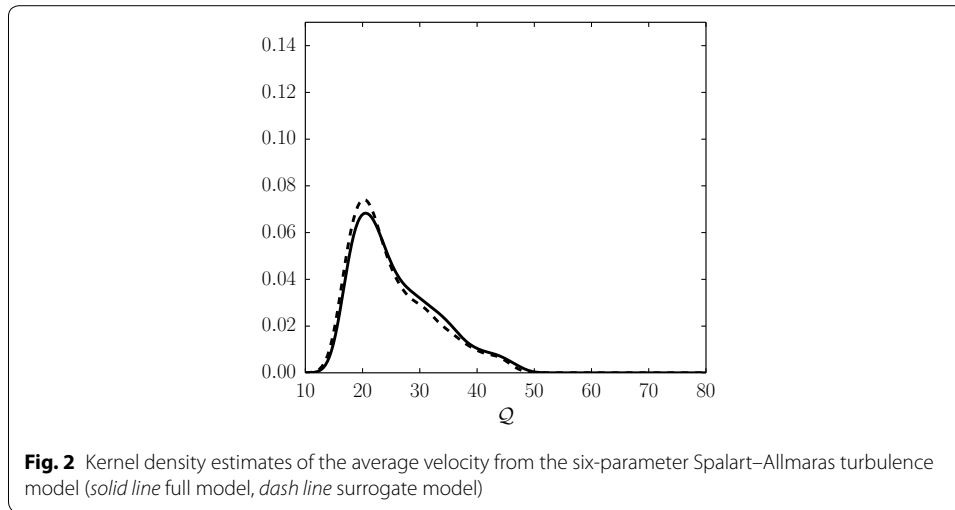
Fig. 1 Convergence of error estimate for adaptive surrogate of Spalart–Allmaras turbulence model with six uncertain parameters

Table 2 Expansion orders for parameters in adaptive surrogate of Spalart–Allmaras turbulence model

Iteration	κ	c_{b1}	σ_{SA}	c_{b2}	c_{v1}	c_{w2}
1	0	0	0	0	0	0
2	1	0	0	0	0	0
3	1	0	0	0	1	0
4	2	0	0	0	1	0
5	3	0	0	0	1	0
6	3	0	1	0	1	0
7	3	1	1	0	1	0
8	3	1	1	0	1	1
9	4	1	1	0	1	1
10	4	2	1	0	1	1
11	4	2	2	0	1	1
12	4	2	2	0	2	1
13	5	2	2	0	2	1
14	5	2	2	1	2	1
15	5	2	3	1	2	1
16	5	3	3	1	2	1
17	5	3	3	1	2	2
18	6	3	3	1	2	2

are associated with the κ and c_{v1} parameters, demonstrating that their values have the greatest influence on the quantity of interest. Following initial refinement of κ and c_{v1} , we also see a continued increase in the expansion order for κ , which we would expect to see since it has a significant impact on the flow velocity away from the wall where the velocity is higher and thus contributes more significantly to the quantity of interest. Refinements are suggested for all model parameters, though c_{b2} is only modeled linearly, suggesting that the gradient of the working variable \bar{v} does not have a notable impact on the average velocity.

From Fig. 2, one can clearly see that the adaptive surrogate model is able to capture the general response of the quantity of interest over the range of uncertain parameters. To quantify the agreement between the two distributions, we computed the



Kullback–Leibler divergence and found $D_{KL} = 0.0048$. Moreover, the mean and variance of the response were evaluated at 25.06 and 49.12 for the full model, and 25.88 and 51.52 for the surrogate model, respectively. Furthermore, the convergence of the error estimate suggests that the reduced model can be used in further studies without a significant loss in accuracy. One attractive use case is that of model validation since many techniques, particularly Bayesian methods, require repeated evaluations of the forward model for the quantity of interest. While the surrogate presented in this section is capable of producing accurate predictions of the quantity of interest, in order to be useful in a validation setting, the reduced model must also exhibit the same sensitivities as the full model. We investigate this issue in the next section based on a simplified version of the Bayesian uncertainty quantification study performed in [10].

Bayesian model selection

As a basis for comparison we use the work of Oliver and Moser [10], where Bayesian methods were used to evaluate the validity of a number of turbulence models and discrepancies. To simplify the presentation here, we restrict our investigation to the Spalart–Allmaras model for eddy viscosity discussed previously and consider four different uncertainty models: independent homogeneous, correlated homogeneous, correlated inhomogeneous, and an additive Reynolds stress error model.

Specifically, we use the same calibration data as Oliver and Moser [10], which was obtained from direct numerical simulations by Jiménez et al. [55, 56]. Mean velocity measurements were taken at $Re_\tau = 944$ and $Re_\tau = 2003$. The uncertainty in the observations from the direct simulation is the result of calculating the sample mean rather than the true mean. The authors of [57] provide an estimate of the variance in the error, however the covariance between data points in the profile is not provided. To minimize the impact of the correlation of sampling errors between measurement points, Oliver and Moser [10] downsampled the data and considered points that are farther apart; we will do the same and assume the sampling errors to be independent. Since the simulation of the channel flow is dependent on the Reynolds number Re_τ , we will construct two surrogate

models, one for each flow scenario represented in the calibration data. For all error models we will use the same surrogate construction of the approximate forward models.

As we did in the examination of the forward model, we consider six uncertain parameters. This set of parameters is naturally augmented with the calibration parameters for the uncertainty models considered. For both Reynolds numbers we will use the final expansion order in Table 2, $\mathcal{N} = (6, 3, 3, 1, 2, 2)$, which yields error estimates $\eta_{944} = 1.388788 \times 10^{-2}$ and $\eta_{2003} = 1.878746 \times 10^{-2}$.

We begin with the general description of the model error by supposing that the error is multiplicative in terms of the velocity. Thus, the observed data is taken to be governed by the equation,

$$\langle u \rangle^+(z; \xi) = (1 + \epsilon(z; \xi))U^+(z; \xi), \quad (43)$$

where $z = y/H$ is the non-dimensionalized wall-normal coordinate, $U^+ = U/u_*$ is the non-dimensionalized velocity, and $\langle u \rangle^+$ is the prediction of the true non-dimensionalized velocity. We assume a zero-mean Gaussian field for the error term $\epsilon = \epsilon(z) = \epsilon(z; \xi)$. The previous studies [10, 46] investigated three different definitions of the covariance of ϵ : independent homogeneous, correlated homogeneous, and correlated inhomogeneous.

Independent homogeneous covariance

First, we will adopt the belief that the data points provided in the DNS calibration data are independent. The covariance of ϵ is thus,

$$\langle \epsilon(z)\epsilon(z') \rangle = \sigma^2 \delta(z - z'), \quad (44)$$

where the standard deviation σ will be treated as an unknown parameter in the calibration process in addition to the turbulence model parameters.

Correlated homogeneous covariance

A straightforward extension of the independent multiplicative model is to incorporate spatial correlation into the definition of the covariance for ϵ . If we assume a homogeneous correlation length, we can write the covariance of ϵ as

$$\langle \epsilon(z)\epsilon(z') \rangle = \sigma^2 \exp\left(-\frac{(z - z')^2}{2l^2}\right), \quad (45)$$

where now both σ and the correlation length l are additional calibration parameters. While undoubtedly more reasonable than independent errors, a homogeneous correlation length still seems improbable; generally the accuracy of the turbulence models near the wall differs considerably from that in the region far from the wall.

Correlated inhomogeneous covariance

Since length scales in turbulent flows are set differently based on the region of the flow, it makes sense to incorporate that structure in the uncertainty model. To mimic the change in length scales, we use a covariance function with a variable length scale [10, 58],

$$\langle \epsilon(z)\epsilon(z') \rangle = \sigma^2 \left(\frac{2l(z)l(z')}{l^2(z) + l^2(z')} \right)^{1/2} \exp\left(-\frac{(z - z')^2}{l^2(z) + l^2(z')}\right), \quad (46)$$

where σ is a calibration parameter and the length scale function $l(z)$ is given by,

$$l(z) = \begin{cases} l_{\text{in}} & \text{for } z < z_{\text{in}} \\ l_{\text{in}} + \frac{l_{\text{out}} - l_{\text{in}}}{z_{\text{out}} - z_{\text{in}}}(z - z_{\text{in}}) & \text{for } z_{\text{in}} \leq z \leq z_{\text{out}} \\ l_{\text{out}} & \text{for } z > z_{\text{out}}. \end{cases} \quad (47)$$

Here $l_{\text{in}} = l_{\text{in}}^+/Re_\tau$, $z_{\text{in}} = z_{\text{in}}^+/Re_\tau$, and l_{in}^+ , z_{in}^+ , l_{out} , and z_{out} are additional calibration parameters.

Reynolds stress uncertainty model

Finally, we introduce an uncertainty model based on the Reynolds stress. While more complex, and thus more difficult to implement in practice, a Reynolds stress uncertainty model is appealing since it directly targets the source of error, the approximation of the Reynolds stress tensor.

The uncertainty model for Reynolds stress proposed by Oliver and Moser [10], takes the form,

$$\langle \mathbf{u}'_i \mathbf{u}'_j \rangle^+(z; \xi) = T^+(z; \xi) - \epsilon(z; \xi) \quad (48)$$

where T^+ is the Reynolds shear stress computed by the approximate turbulence model, ϵ represents the error field as before, and $\langle \mathbf{u}'_i \mathbf{u}'_j \rangle^+$ is the predicted Reynolds stress.

Note that (48) does not lead to a prediction of the mean velocity directly; in order to predict the mean flow one must first compute the Reynolds stress T^+ governed by the turbulence model. Then the predicted Reynolds stress $\langle \mathbf{u}'_i \mathbf{u}'_j \rangle^+$ obtained from (48) is used to complete the momentum equations,

$$-\frac{d}{dz} \left(\frac{1}{Re_\tau} \frac{d\langle u \rangle^+}{dz} - \langle \mathbf{u}'_i \mathbf{u}'_j \rangle \right) = 1, \quad (49)$$

which must then be solved for the mean velocity $\langle u \rangle^+$. As a result, realizations of the solution do not necessarily satisfy the turbulence momentum equations (10) [10].

We must still choose a description for the error field ϵ . Sticking with [10], we assume a zero-mean Gaussian with covariance,

$$\langle \epsilon(z) \epsilon(z') \rangle = k_{\text{in}}(z, z') + k_{\text{out}}(z, z'), \quad (50)$$

where k_{in} models the error near the wall and k_{out} represents the error far from the wall. The choices made in [10] are,

$$k_{\text{in}}(z, z') = \sigma_{\text{in}}^2 \left(1 - \frac{(z - z')^2}{l_{\text{in}}^2} - \frac{(z - z')^2}{\Delta^2} - \frac{l_{\text{in}}^2 z z'}{\Delta^2 \Delta^2} \right) \times \exp \left(-\frac{1}{2} \frac{(z - z')^2}{l_{\text{in}}^2} - \frac{1}{2} \frac{(z^2 + z'^2)^2}{\Delta^2} \right), \quad (51)$$

$$k_{\text{out}}(z, z') = \sigma_{\text{out}}^2 \left(1 - \frac{(z - z')^2}{l_{\text{out}}^2} \right) \exp \left(-\frac{1}{2} \frac{(z - z')^2}{l_{\text{out}}^2} \right), \quad (52)$$

where $l_{\text{in}} = l_{\text{in}}^+ / Re_\tau$, $\Delta = C_d l_{\text{in}}$ and $\sigma_{\text{out}} = C_s / Re_\tau$. The additional calibration parameters for this uncertainty model are σ_{in} , l_{in}^+ , C_d , C_s , l_{out} .

Numerical results

Some of the above models are obviously deficient while others may be overly complex for the present analysis; all models are now evaluated based on their agreement with the calibration data. With the model set \mathcal{M} comprised of the four models proposed in this section, we are prepared to move forward with the application of the Bayesian model comparison procedure with the adaptive surrogate model for the turbulent channel flow problem.

The same uniform distributions used to define the parameter ranges in "[Adapted surrogate model for turbulence model problem](#)" are carried over here as the prior distributions for each parameter. The results of this section were obtained using the QUESO library of algorithms for statistical inverse problems [59]; more specifically, we used the multi-level sampling functionality.

Table 3 reports the evidence computed for each of the four models and reproduces the relevant portion of Table 2 from [10]. While the numerical values themselves differ, the results are qualitatively similar. As expected, the two multiplicative error models with homogeneous covariance structures have very small evidences; clearly the error in the turbulence model is not homogeneous across the channel. Interestingly, the multiplicative model with inhomogeneous covariance appears to perform on par with the much more complex Reynolds stress model. If we actually compute the plausibility using the prior, we see that both are nearly equally plausible with $P(M_3|\mathcal{M}) = 0.53$ for the inhomogeneous covariance model and $P(M_4|\mathcal{M}) = 0.46$ for the Reynolds stress model. At this stage either model is equally probable, however the multiplicative model may be preferred for its relative simplicity.

Our conclusions are in line with those found in [10] regarding the inability of homogeneous uncertainty models to capture the turbulence modeling error. Their results favor the Reynolds stress model over the correlated inhomogeneous model, but again the two are relatively close in plausibility. In contrast to the use of the full turbulence model, our surrogate based model selection procedure required less computationally effort to complete. Table 4 displays the number of full model simulations required for each of the models using the full turbulence model, where as the use of the surrogate requires only polynomial evaluations. The more complex the full model the greater advantage the use of a reduced model will be. Of course this does not take into account the number of solutions needed to actually construct the adaptive surrogate models. A total of 7515

Table 3 Evidences computed for the four uncertainty models ($\log(E)$ is reported in the table)

	Surrogate	Full model
Independent homogeneous	-1.457	8.862
Correlated homogeneous	1.963	8.045
Correlated inhomogeneous	164.9	164.0
Reynolds stress	164.8	169.0

Table 4 Number of nonlinear PDE solves computed in the MCMC process for each uncertainty model

Independent homogeneous	7776
Correlated homogeneous	8453
Correlated inhomogeneous	8084
Reynolds stress	7595
Total	31,908

samples (nonlinear solves) were used to complete all 18 iterations of the adaptive procedure to construct the anisotropic surrogate model. Even factoring in these additional solves, it is obvious that using the reduced model in the model selection study leads to significant computational savings. The advantage of performing the analysis with a surrogate model is that the same surrogate can be used for all uncertainty models, thus the cost can be amortized over the exploration of many different uncertainty models. In other words, once a surrogate model has been constructed, a new model selection study can be performed rather efficiently whenever we wish to access a newly proposed uncertainty model, making the surrogate approach a valuable resource for modelers.

Our results suggest that for uncertainty quantification studies of this nature, requiring many simulations of a relatively complex forward model, the reduced model described here leads to the same conclusion as using the full model to perform the same analysis. In fact, some work has been done to show that, in limited cases with mostly Gaussian assumptions, the error in the surrogate model can be used to prove a bound on the error in the posterior distributions obtained through Bayesian inference [2–5]. However, further effort is needed to extend these results to more general cases.

Conclusion

We have examined the application of goal-oriented error estimation to the adaptivity of surrogate models for boundary-value problems with uncertainty. In contrast to existing anisotropic refinement strategies, a new refinement algorithm was proposed that uses higher-order information from the goal-oriented error estimate to identify the most influential parameters and adapt the surrogate model accordingly.

Based on our approach, an accurate surrogate model need to be constructed for the Spalart–Allmaras turbulence model and the solution of the RANS equations in a fully-developed channel. The reduced model was then used in a Bayesian model calibration study in place of the full simulation. Posterior distributions for the parameters showed excellent agreement with those obtained using the original forward model. The results demonstrate that the newly developed methodology can be a valuable resource to computational scientists in assessing complex physical systems using Bayesian techniques where a large number of model simulations are required. In our case, the quantity of interest and calibration observables were both defined in terms of velocities; an interesting question is how to alter the approach taken here if the observable measurements are rather different from the quantity of interest for which we would like to use the model. One would thus be required to construct surrogate models with respect to both the quantity of interest and the observable calibration data. This will be the subject of future work.

Authors' contributions

CB carried out the derivation of the error estimator and adaptive method for the construction of surrogate models, the implementation of the algorithms and numerical studies, and the drafting of the manuscript. SP participated in the development of the theoretical results on the error estimator and adaptive method, in the analysis of the numerical results, and contributed to the drafting of the manuscript. Both authors read and approved the final manuscript.

Author details

¹ Département de Mathématiques et de Génie Industriel, Ecole Polytechnique de Montréal, Montréal, QC, Canada. ² Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin, Austin, TX, USA.

Acknowledgements

SP is grateful for the support by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada. He is also a participant of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering. CB acknowledges the support by the Department of Energy [National Nuclear Security Administration] under Award Number [DE-FC52-08NA28615]. The authors are also grateful to Todd Oliver from the Institute for Computational Engineering and Sciences at The University of Texas at Austin for useful discussions on the Spalart–Allmaras model and the calibration of the model parameters using Bayesian inference.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests.

Received: 8 February 2015 Accepted: 26 August 2015

Published online: 17 September 2015

References

- Kennedy MC, O'Hagan A. Bayesian calibration of computer models. *J Royal Statist Soc Series B*. 2001;63(3):425–64.
- Li J, Marzouk Y. Adaptive construction of surrogates for the Bayesian solution of inverse problems. *SIAM J Sci Comput*. 2014;36:1164–86.
- Ma X, Zabaras N. An efficient Bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method. *Inverse Probl*. 2009;25(3):035013.
- Marzouk Y, Najm H, Rahn L. Stochastic spectral methods for efficient Bayesian solution of inverse problems. *J Comput Phys*. 2007;224:560–86.
- Marzouk Y, Xiu D. A stochastic collocation approach to bayesian inference in inverse problems. *Commun Comput Phys*. 2009;6(4):826–47.
- Oden JT, Prudhomme S. Estimation of modeling error in computational mechanics. *J Comput Phys*. 2002;182:496–515.
- Oden JT, Prudhomme S, Romkes A, Bauman PT. Multiscale modeling of physical phenomena: adaptive control of models. *SIAM J Sci Comput*. 2006;28(6):2359.
- Allmaras SR, Johnson FT, Spalart PR. Modifications and clarifications for the implementation of the Spalart–Allmaras turbulence model. In: *Seventh International Conference on Computational Fluid Dynamics (ICCF7)*, Big Island, Hawaii, USA, 9–13 July 2012.
- Spalart PR, Allmaras SR. A one-equation turbulence model for aerodynamic flows. *Recherche Aérospatiale*. 1994;1:5–21.
- Oliver TA, Moser RD. Bayesian uncertainty quantification applied to RANS turbulence models. *J Phys Conf Ser*. 2011;318(4):042032.
- Cheung SH, Oliver TA, Prudencio EE, Prudhomme S, Moser RD. Bayesian uncertainty analysis with applications to turbulence modeling. *Reliab Eng Syst Saf*. 2011;96(9):1137–49.
- Bryant CM, Prudhomme S, Wildey T. A posteriori error control for partial differential equations with random data. *ICES Report 13-08*, 2013.
- Bryant CM, Prudhomme S, Wildey T. Error decomposition and adaptivity for response surface approximations from PDEs with parametric uncertainty. *SIAM/ASA J. Uncert Quant*. 2015. **(accepted)**.
- Durbin PA, Petterson Reif BA. *Statistical theory and modeling for turbulent flows*. New York: Wiley; 2001.
- Moser RD. *Turbulence (lecture notes)*. 2013.
- Pope SB. *Turbulent flows*. New York: Cambridge University Press; 2000.
- Baldwin B, Lomax H. Thin-layer approximation and algebraic model for separated turbulent flows. *AIAA Paper 257*, 1978.
- Chien KY. Predictions of channel and boundary-layer flows with a low-Reynolds-number two-equation model of turbulence. *AIAA J*. 1982;20(1):33–8.
- Durbin PA. Separated flow computations with the $k - \epsilon - v^2$ model. *AIAA J*. 1995;33:659–64.
- Oliver TA, Darmofal DL. Impact of turbulence model irregularity on high-order discretizations. *AIAA Paper 953*, 2009.
- Ghanem R, Red-Horse J. Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach. *Phys D Nonlinear Phenom*. 1999;133:137–44.
- Ghanem R, Spanos P. *Stochastic finite elements: a spectral approach*. New York: Springer; 2002.
- Xiu D, Karniadakis G. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J Sci Comput*. 2002;24:619–44.
- Wan X, Karniadakis GE. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J Comput Phys*. 2005;209(2):617–42.

25. Wan X, Karniadakis G. Solving elliptic problems with non-Gaussian spatially-dependent random coefficients. *Comput Methods Appl Mech Eng.* 2009;198(21–26):1985–95.
26. Wan X, Karniadakis G. Error control in multi-element generalized polynomial chaos method for elliptic problems with random coefficients. *Commun Computat Phys.* 2009;5(2–4):793–820.
27. Constantine PG, Eldred MS, Phipps ET. Sparse pseudospectral approximation method. *Comput Methods Appl Mech Eng.* 2012;229–232:1–12.
28. Babuška I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J Numer Anal.* 2007;45(3):1005–34.
29. Ganis B, Klie H, Wheeler MF, Wildey T, Yotov I, Zhang D. Stochastic collocation and mixed finite elements for flow in porous media. *Comput Methods Appl Mech Eng.* 2008;197(43–44):3547–59.
30. Gerstner T, Griebel M. Dimension-adaptive tensor-product quadrature. *Computing.* 2003;71(1):65–87.
31. Nobile F, Tempone R, Webster CG. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J Numer Anal.* 2008;46(5):2411–42.
32. Nobile F, Tempone R, Webster CG. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J Numer Anal.* 2008;46(5):2309–45.
33. Todor RA, Schwab C. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J Numer Anal.* 2006;27(2):232–61.
34. Le Maître O, Knio O. Spectral methods for uncertainty quantification: with applications to computational fluid dynamics. New York: Springer; 2010.
35. Butler T, Dawson C, Wildey T. A posteriori error analysis of stochastic spectral methods. *SIAM J Sci Comput.* 2011;33:1267–91.
36. Butler T, Constantine P, Wildey T. A posteriori error analysis of parameterized linear systems using spectral methods. *SIAM J Matrix Anal Appl.* 2012;33:195–209.
37. Almeida RC, Oden JT. Solution verification, goal-oriented adaptive methods for stochastic advection-diffusion problems. *Comput Methods Appl Mech Eng.* 2010;199(37–40):2472–86.
38. Mathelin L, Le Maître O. Dual-based a posteriori error estimate for stochastic finite element methods. *Commun Appl Math Comput Sci.* 2007;2(1):83–115.
39. Eriksson K, Estep D, Hansbo P, Johnson C. Introduction to adaptive methods for differential equations. *Acta Numer.* 1995;4:105–158.
40. Eriksson K, Estep D, Hansbo P, Johnson C. Computational differential equations. New York: Cambridge University Press; 1996.
41. Becker R, Rannacher R. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.* 2001;10:1–102.
42. Oden JT, Prudhomme S. Goal-oriented error estimation and adaptivity for the finite element method. *Comput Math Appl.* 2001;41(5):735–56.
43. Ainsworth M, Oden JT. A posteriori error estimation in finite element analysis. New York: Wiley; 2000.
44. Babuška I, Strouboulis T. The finite element method and its reliability. Oxford: Oxford University Press; 2001.
45. Verfürth R. A posteriori error estimation techniques for finite element methods. Numerical mathematics and scientific computation. New York: Oxford University Press; 2013.
46. Bryant CM. On goal-oriented error estimation and adaptivity for nonlinear systems with uncertain data and application to flow problems. PhD thesis, The University of Texas at Austin, Austin, TX, 2014.
47. van der Zee KG. Goal-adaptive discretization of fluid-structure interaction. PhD thesis, Delft University of Technology, The Netherlands, 2009.
48. Prudhomme S, Oden JT. Computable error estimators and adaptive techniques for fluid flow problems. In: Barth T, Deconinck H, editors. Estimation and adaptive discretization methods in computational fluid dynamics, lecture notes in computational science and engineering, vol. 25. Heidelberg: Springer; 2003. p. 207–68.
49. Bangerth W, Rannacher R. Adaptive finite element methods for differential equations. Boston: Birkhauser; 2003.
50. Estep D, Larson MG, Williams RD. Estimating the error of numerical solutions of systems of reaction-diffusion equations. Providence: American Mathematical Soc.; 2000.
51. Ma X, Zabaras N. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *J Comput Phys.* 2009;228(8):3084–113.
52. Calvetti D, Somersalo E. Introduction to Bayesian scientific computing. New York: Springer; 2007.
53. Kaipio J, Somersalo E. Statistical and computational inverse problems. New York: Springer; 2005.
54. Jaynes ET. Probability theory: the logic of science. New York: Cambridge University Press; 2003.
55. Hoyas S, Jiménez J. Scaling of the velocity fluctuations in turbulent channels up to $Re_\tau = 2003$. *Phys Fluids.* 2006;18(1):011702.
56. Del Alamo JC, Jiménez J, Zandonade P, Moser RD. Scaling of the energy spectra of turbulent channels. *J Fluid Mech.* 2004;500:135–44.
57. Hoyas S, Jiménez J. Reynolds number effects on the Reynolds-stress budgets in turbulent channels. *Phys Fluids.* 2008;20(10):101511.
58. Rasmussen CE, Williams KI. Gaussian processes for machine learning. Cambridge: MIT Press; 2006.
59. Prudencio EE, Schulz KW. The parallel C++ statistical library QUESO: quantification of uncertainty for estimation, simulation and optimization. Euro-Par 2011: Parallel Processing Workshops, vol. 7155., Lecture Notes in Computer Science. Heidelberg: Springer; 2012. p. 398–407.