Journal of Computational Physics 394 (2019) 280–297

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

# Leveraging Bayesian analysis to improve accuracy of approximate models

Balasubramanya Nadiga [a,*], Chiyu Jiang [b], Daniel Livescu [a]

[a] Los Alamos National Laboratory, Los Alamos, NM, USA
[b] University of California, Berkeley, CA, USA

ARTICLE INFO

*Article history:*
Received 12 November 2018
Received in revised form 9 May 2019
Accepted 9 May 2019
Available online 27 May 2019

*Keywords:*
Bayesian analysis
Reduced order modeling
Turbulence modeling
Reynolds-averaged Navier Stokes
Neural network
Surrogate modeling

ABSTRACT

We focus on improving the accuracy of an approximate model of a multiscale dynamical system that uses a set of parameter-dependent terms to account for the effects of unresolved or neglected dynamics on resolved scales. We start by considering various methods of calibrating and analyzing such a model given *a few* well-resolved simulations. After presenting results for various point estimates and discussing some of their shortcomings, we demonstrate (a) the potential of hierarchical Bayesian analysis to uncover previously unanticipated physical dependencies in the approximate model, and (b) how such insights can then be used to improve the model. In effect parametric dependencies found from the Bayesian analysis are used to improve structural aspects of the model. While we choose to illustrate the procedure in the context of a closure model for buoyancy-driven, variable-density turbulence, the statistical nature of the approach makes it more generally applicable. Towards addressing issues of increased computational cost associated with the procedure, we demonstrate the use of a neural network based surrogate in accelerating the posterior sampling process and point to recent developments in variational inference as an alternative methodology for greatly mitigating such costs. We conclude by suggesting that modern validation and uncertainty quantification techniques such as the ones we consider have a valuable role to play in the development and improvement of approximate models.

Published by Elsevier Inc.

## 1. Introduction

Natural and engineered systems that exhibit multi-scale behavior due to coupling of subsystems with different spatial and temporal scales are commonplace. For example, interactions between wave modes, interactions between vortical modes, and coupled interactions between wave and vortical modes gives rise to complex phenomena in an array of fluid dynamic problems ranging from homogeneous isotropic turbulence, to common and engineering instances of turbulent fluid flow, to rotating-stratified turbulence relevant to flows on global and astrophysical scales. In a different setting, interactions between fast vibrations and slow conformational changes lead to complex behavior in molecular dynamics relevant to protein folding, and so on.

* Corresponding author.
*E-mail address:* balu@lanl.gov (B. Nadiga).

https://doi.org/10.1016/j.jcp.2019.05.015
0021-9991/Published by Elsevier Inc.

If we represent the comprehensive mathematical model of such a complex, multiscale, dynamical system, (the full Navier-Stokes equations and the full force-field based molecular dynamics respectively in the two examples above) symbolically as

$$\frac{d}{dt}\mathbf{u}^{WR} = \mathbf{f}^{WR}(\mathbf{u}^{WR}; \boldsymbol{p}), \tag{1.1}$$

where $\mathbf{u}$ represents the state vector, $\mathbf{f}$ the tendency of the state vector, $\mathbf{p}$ is a set of physical parameters, and the superscript $WR$ stands for well-resolved, then it is almost invariably the case that the computational complexity of such a model prevents it from being used routinely in design and analysis of the system of interest. Therefore, a pragmatic consideration in seeking an approximate model is to reduce the computational cost of the original comprehensive model by limiting the associated number of degrees of freedom, while keeping the error in relevant quantities of interest (QoI) small. In the context of (1.1), an approximate model may be symbolically represented as

$$\frac{d}{dt}\mathbf{u}^{RO} = \mathbf{f}^{RO}(\mathbf{u}^{RO}; \boldsymbol{p}; \boldsymbol{\phi}), \tag{1.2}$$

where superscript $RO$ stands for reduced-order, dimension of $\mathbf{u}^{RO} \ll$ dimension of $\mathbf{u}^{WR}$, and $\boldsymbol{\phi}$ is a set of modeling parameters.

We note that a Reduced Order Model (ROM) is one kind of an approximate model that is popular in the engineering fields and industry wherein the original governing equations are projected on to a reduced-order basis, but where the basis vectors are determined from snapshot data, e.g., using Principal Component Analysis (PCA) [e.g., see [22] and references therein], or Dynamic Mode Decomposition (DMD) [e.g., see [35] and others], etc., and where the snapshot data is obtained by solving the original system. In the ROM context, the model parameters $\boldsymbol{\phi}$ may be thought of as associated with terms such as closures [as in, e.g., [41] and others], etc. We also note that an approximate model may be based entirely on a data-driven approach. For example, in the field of molecular dynamics (MD), it is common to use an approximate approach such as a Markov State Model (MSM) to permit longer simulation times. In this approach the conformation space is first partitioned, based on MD simulation data, into discrete subspaces, by using techniques such as clustering or PCA. And then, in such a reduced subspace, molecular kinetics itself is reduced to transitions between the identified discrete subspaces that are governed by transition probabilities that may be parameterized and learnt, again, from MD simulation data [e.g., see [11,31,36,19] and references therein].

### 1.1. Approximate models of turbulent flows

While further combinations of such data-driven and equation-based techniques are clearly possible [e.g., see [12]], we shift attention to the specific context of a turbulent flow. A common objective in the study of such a flow, is to be able to accurately compute QoIs of practical relevance. However, in a turbulent flow not only are the relevant fields (velocity, density, and others) three-dimensional, time-dependent and random, but there is a large range of time and length scales. Whereas the domain size ($L$) and geometry directly affect the flow, the Kolmogorov length—a measure of the smallest scales of turbulence—scales as $\mathrm{Re}^{-3/4}$ and the Kolmogorov time scales as $\mathrm{Re}^{-1/2}$ [33]. Here, Re is the Reynolds number, a non-dimensional number that characterizes the relative effects of nonlinearity and viscosity, and it is large in turbulent flows.

Thus, if all the scales were to be resolved, as in the Direct Numerical Simulation (DNS) procedure, the computational cost would scale as $\mathrm{Re}^3$. Here, computational cost is estimated as scaling with the number of operations $N_{ops} = N_{cells} \times N_{timesteps}$, with $N_{cells} \sim (L/\Delta x)^3 \sim \left(Re^{3/4}\right)^3$, where the cube accounts for three-dimensionality of the flow and $N_{timesteps} \sim Re^{3/4}$ rather than $Re^{1/2}$ since DNS practitioners commonly use explicit time-stepping with Courant-Frederick-Levy number of O(1). In this case, the *computational* timestep goes down as $Re^{-3/4}$. Such steep scaling of computational cost with Reynolds number for a fully resolved simulation makes it feasible only for low to moderate Reynolds numbers, meaning to say prohibitively expensive for realistic flows of interest.

Computational intractability of being able to fully resolve a (high Reynolds number) turbulent flow, leads naturally to the question of modeling only a limited range of scales that are of direct interest. In this context, the procedure of averaging the governing equations over a scale of the order of the smallest scale of interest leads to the closure problem. Here by *closure* we refer to how the neglected degrees of freedom affect the evolution of the resolved degrees of freedom.

In the Large Eddy Simulation (LES) approach [e.g., see [33]], equations are solved for filtered fields (with a filter scale $\Delta$) that are representative of the larger-scale turbulent motions and these equations include a (closure) model for how the unresolved scales affect the resolved scales. While the computational cost of this scale-restricted model is seen straightforwardly to scale as $(L/\Delta)^4$, the requirement of $\Delta$ to lie within the inertial range of turbulence (a range of scales within which dynamics is self-similar) renders its computational cost still too high for realistic flows given current computational resources.

For these reasons, in order to obtain a model that attains acceptable levels of accuracy while being computationally affordable, an ensemble-average (Reynolds-average) of the original Navier-Stokes equations (Reynolds-Averaged Navier Stokes

or RANS equations) is considered. Since the RANS equations solve only for the ensemble-averaged fields, a solution procedure using these equations is not required to resolve either the inertial range (cf. LES) or the dissipative range (cf. DNS). The RANS equations are therefore amenable to being solved using much coarser resolutions. The flip side of solving the RANS equations at coarse resolutions is that in order to attain an acceptable level of accuracy, the effects of a large set of unrepresented scales (on resolved scales) have to be explicitly modeled. Such closure issues notwithstanding, the RANS approach has emerged as the method of choice [e.g., see [33]] in both understanding realistic turbulent flows and to address practical issues of design, optimization, and operations in the context of such flows. The RANS approach may be written symbolically as

$$\frac{d}{dt}\mathbf{u}^{RO} = \mathbf{f}^{RO}(\mathbf{u}^{RO}; \boldsymbol{p}; \boldsymbol{\phi}) = \mathbf{f}^{WR}(\mathbf{u}^{RO}; \boldsymbol{p}) + \mathbf{m}(\mathbf{u}^{RO}; \boldsymbol{\theta}), \tag{1.3}$$

where $\mathbf{u}^{RO}$ is the ensemble (or Reynolds) average of $\mathbf{u}^{WR}$ ($=\overline{\mathbf{u}^{WR}}$), so that if $\mathbf{u}' = \mathbf{u}^{WR} - \mathbf{u}^{RO}$, $\overline{\mathbf{u}'} = 0$, $\mathbf{m}$ represent the closure terms, and for convenience the vector of modeling parameters in the RANS approach is represented by $\boldsymbol{\theta}$.

The modeling of closures in the LES and RANS approaches to modeling turbulent flows is central to the accuracy and efficiency of these methods. Such modeling is based largely on phenomenological understanding of prototypical turbulent flows by subject matter experts and is mostly deterministic. However, stochastic methods that typically use stochastic models of turbulent fluid flow to evolve probability density functions of turbulent fields such as velocity have also been used to address the closure problem [e.g., see [33]], and data-driven stochastic closures are also being considered [e.g., see [13,27,32] and others]. In either case, however, improvement of a model, in terms of reducing model-form related error has almost exclusively been the domain of subject matter experts and theoreticians. Thus, whereas a computational approach for such model improvement currently does not exist, we present one.

### 1.2. Structural and parametric uncertainty in the approximate model

When additional explicit model terms, such as $\mathbf{m}$ in (1.3), are used to achieve a simplified approximation of the full system, the form of such a closure is a choice that is made and is, therefore, not unique. It also follows that the QoIs in the approximate model depend on the choice of the form of the closure. We refer to this as structural or model-form uncertainty/error. After the choice of the form of the closure has been made, the values of the parameters $\boldsymbol{\theta}$ influence the accuracy of the approximate model, leading to parameter uncertainty/error. Even when additional explicit model terms, such as $\mathbf{m}$ in (1.3), are not used, it is easy to imagine hypothetically introducing such tendencies as a means of reducing the discrepancy between the average of $\mathbf{u}^{WR}$ and $\mathbf{u}^{RO}$. Indeed such terms are routine when neural networks are used to improve the model, but with the crucial difference that the dimension of $\boldsymbol{\phi} \gg$ the dimension of $\boldsymbol{\theta}$ in such a case (and consequently structural uncertainty is less important). Once the choice of the form of the closure has been made, we next consider the process of its validation and calibration given parametric uncertainty.

### 1.3. Model calibration and validation

The calibration and validation of a model, e.g., as in the RANS approach to modeling turbulent flows is based either on experimental data or observations or data from numerical studies that resolve dynamics and physics of the full range of relevant scales. In the case of some idealized turbulent flows, such calibration data may come from DNS of the flow. Here, by calibration, we refer to the assignment or adjustment of values of model parameters, $\boldsymbol{\theta}$ in (1.3), so as to bring model prediction of QoIs, $\mathbf{q}^{RO} \equiv \mathbf{q}^{RO}(\mathbf{u}^{RO})$, in certain scenarios into agreement with known values of QoIs, $\mathbf{q}^{WR}$, for those scenarios. And, validation refers to the process of determining the degree to which the model, (1.3), is an accurate representation of the real world, here (1.1), from the perspective of the intended uses of the model [34,2,17].

Different approaches to calibration and validation exist. However, calibration and validation of RANS turbulence models has traditionally relied on point estimates—that is one optimal set of parameters are sought that best fits the calibration data [e.g., see [33,37]]:

$$\boldsymbol{\theta}_{opt} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} ||\mathbf{q}^{WR}(\mathbf{u}^{WR}) - \mathbf{q}^{RO}(\mathbf{u}^{RO}; \boldsymbol{\theta})||. \tag{1.4}$$

However, it is not guaranteed that such calibration is optimal. For example, there may be multiple, possibly very different multi-way balances of phenomena that lead to similar evolutions of QoIs. In the turbulent flow context, different balances of dissipation, transport, and decay processes could be consistent with the evolution of turbulent moments as specified by given calibration data.

In contrast to approaches that aim to find a single optimal set of parameters, a Bayesian framework models the model parameters as random variables and seeks to approximate the joint distribution of such variables. In other words, it estimates the posterior probability of $\boldsymbol{\theta}$ given $\mathbf{q}^{WR}$: $P(\boldsymbol{\theta}|\mathbf{q}^{WR})$. In so doing, the methodology gives consideration to the possibility that different balances of key modeled processes can explain the calibrating data. That is, the Bayesian framework integrally permits consideration of parameter uncertainty in the calibration and validation of a model. Furthermore, the Bayesian approach allows for better integration of both prior knowledge and probabilistic structure into the calibration and validation

process. We are, therefore, interested in examining the utility of a Bayesian approach to calibration and validation, and what advantages it may hold over the traditional point estimate based approaches in developing insights into improving the model under consideration.

A few recent studies have used Bayesian estimation techniques to calibrate RANS models. For example, Oliver and Moser [30], and Cheung et al. [7] use Bayesian uncertainty analysis to calibrate and inter-compare four well-known RANS models: the Baldwin-Lomax model [3], the SA model [39], the Chien $k$-$\epsilon$ model [8], and the $v^2 - f$ model [14]; Edeling et al. [16,15] used Bayesian estimates of parameter variability in the Launder-Sharma $k - \epsilon$ model [23] as a means to estimate errors in RANS simulations; [17] present an adaptive modeling algorithm for selection and validation of models, however, in the domain of atomistic systems. In other work that uses a Bayesian framework in the context of RANS models, Edeling et al. [16,15], note that the distribution of model parameters provides information on error associated with the model: when the joint probability density function (PDF) of the model parameters is propagated through the model, the distribution of the QoIs can be used to provide confidence bounds for the QoIs.

In addition to such uses of the Bayesian methodology, we think that the Bayesian methodology has a useful role to play in the development and improvement of the approximate or reduced-order descriptions—a role that should be thought of as complementary to that of subject matter expertise that remains central to developing approximate models. Indeed, we show how Bayesian analysis of a RANS model that uses DNS (or equivalently any other calibration data) can uncover unanticipated dependencies which in turn point to structural deficiencies in the model and specific ways in which the model can be improved.

Organization of the rest of the article is as follows. In Sec. 2 we discuss the specifics of the problem we consider, the fully resolved model, the specific approximate model we consider, and the QoIs. In Sec. 3 we discuss the details about the methods we use before presenting respective results. This includes details about point estimates and Bayesian estimation, how we go about improving the model and details about the neural network-surrogate strategy that we propose for situations wherein the approximate model itself may still be expensive enough to prohibit its direct use in Bayesian analysis. This is followed by a brief section that summarizes the work and concludes.

## 2. The problem setting, the fully resolved model, and the approximate model

Buoyancy-driven turbulence is commonplace in a wide variety of naturally-occurring flows (e.g., ocean-atmosphere dynamics, astrophysics, mantle convection etc.) and engineering flows (e.g., ranging from smoke-stacks to combustion to inertial-confinement fusion), and encompassing both, a wide range of instabilities and rich phenomenology. Homogeneous Rayleigh-Taylor (hRT) turbulence is a particular idealization of buoyancy-driven turbulence. The initial condition for the onset of hRT turbulence consists of isolated regions in the domain of interest being occupied by two miscible fluids at rest that have different densities, and the flow evolves in accordance with the resultant buoyancy force.

### 2.1. The fully resolved model

If the flow is incompressible (low-speed) and the densities of the two fluids are $\rho_1$ and $\rho_2$, the full-order governing equations (symbolically represented by Eqn. (1.1)) for this problem are the Navier-Stokes equations along with the species mass fraction transport equations.

*Species mass fraction transport equations* Given mass fractions $Y_1 = Y_1(\boldsymbol{x}, t)$ and $Y_2 = Y_2(\boldsymbol{x}, t)$ for the two fluids which sum to unity (i.e., $Y_1 + Y_2 = 1$), the density of the mixture can be written as:

$$\rho = \frac{1}{Y_1/\rho_1 + Y_2/\rho_2} \tag{2.1}$$

The species mass fraction transport equations assuming Fickian diffusion (and using the comma notation for derivatives) are:

$$(\rho Y_\alpha)_{,t} + (\rho Y_\alpha u_j)_{,j} = (\rho D_0 Y_{\alpha,j})_{,j}. \tag{2.2}$$

The continuity equations is obtained by summing over $\alpha = 1, 2$. Non-zero divergence of velocity results from mixing due to the change in specific volume, $1/\rho$:

$$u_{j,j} = -D_0 \ln \rho_{,jj}. \tag{2.3}$$

Here $D_0$ is the diffusion coefficient and it is assumed to be constant.

*Navier-Stokes equations* After non-dimensionalization, using an arithmetic mean of the two densities for the reference density $\rho_0 = \frac{1}{2}(\rho_1 + \rho_2)$, a reference velocity $U_0$, and a reference length $L_0$, the well-resolved model equations are:

$$\rho_{,t} + (\rho u_j)_{,j} = 0, \tag{2.4}$$

$$(\rho u_i)_{,t} + (\rho u_i u_j)_{,j} = -p_{,i} + \tau_{ij,j} + \frac{1}{Fr^2}\rho g_i, \tag{2.5}$$

$$u_{j,j} = -\frac{1}{Re_0 Sc}\ln\rho_{,jj}, \tag{2.6}$$

and where $\tau_{ij} = (1/Re_0)(u_{i,j} + u_{j,i} - (2/3)u_{k,k}\delta_{ij})$. In the above equations, the non-dimensional state variables are the density $\rho$, $x_i$-direction velocity $u_i$, and pressure $p$. The non-dimensional parameters in the above equations are the Reynolds number $Re_0$, Schmidt number $Sc$, and Froude number $Fr$, defined by:

$$Re_0 = \frac{\rho_0 L_0 U_0}{\mu_0}, \quad Sc = \frac{\mu_0}{\rho_0 D_0}, \quad Fr^2 = \frac{U_0^2}{gL_0} \tag{2.7}$$

and where $g$ is gravitational acceleration, and $\mu$ is for dynamic viscosity (assumed constant and equal for both fluids). In the above equations, we note that (a) because we consider large density differences, the flow is not amenable to the commonly used Boussinesq approximation, leading to the full density being used consistently in all the terms of the momentum equations, and (b) even though we consider low-speed dynamically-incompressible flows, the velocity field is not divergence-free because molecular mixing leads to changes in the specific volume.

Equations (2.4)–(2.6) are solved using the Direct Numerical Simulation procedure at four different values of Atwood number: 0.05, 0.25, 0.50, and 0.75, where the Atwood number is the normalized density ratio. The reader is referred to [24] for details. These values of the Atwood number characterize the initial conditions, when the fluids are completely segregated. As the fluids molecularly mix, the evolving Atwood number can be calculated as the largest value of the (normalized) density difference between pairs of points in the flow. Such an evolving Atwood number decreases with time.

A brief phenomenological description of the flow evolution is as follows, and the reader is referred to [24] for further details. The unstable nature of the initial distribution of density endows it with high potential energy, and the ensuing buoyancy-driven instability leads to a conversion of potential energy into kinetic energy. The flow subsequently transitions to turbulence. Once turbulent, the flow can be described in terms of the evolution of single-point second-order turbulent correlations. The relevant turbulent correlations in this setting are the Favre average Reynolds stresses $\tilde{R}_{ij}$ $(= \overline{\rho u_i'' u_j''}/\bar{\rho}$, where $u_i'' = u_i - \tilde{u}_i$, with $\tilde{u}_i$ being the Favre or density-weighted averaged velocity $\overline{\rho u_i}/\bar{\rho})$ and the turbulent kinetic energy $(K \equiv \tilde{R}_{ii}/2)$, the turbulent mass flux $a_i$ which is the correlation between density and velocity fluctuations $(a_i \equiv \bar{u}_i'' = \overline{\rho' u_i'}/\bar{\rho})$, and the density-specific-volume covariance $b$. The density-specific volume covariance is largest at initial times and begins to decay as the turbulence grows and the fluids mix (see solid lines in Fig. 2; online version of figures is in color, or contact corresponding author). In particular, at these initial times, the mean pressure gradient that develops in response to gravity couples to $b$, to produce turbulent mass flux $a_1$, which then couples back with the pressure gradient to generate turbulent kinetic energy. At early to intermediate times, the Reynolds stresses (and turbulent kinetic energy), the turbulent mass flux, and turbulent dissipation all grow. Eventually, as the fluids mix, they reach peak values at different times and eventually decay asymptotically to zero.

In this problem setting, the QoIs are the various second order turbulent correlations described previously, and a length scale as discussed further in the next section. Temporal evolution of some of the QoIs, as given by the fully resolved model, is shown in solid lines in Fig. 2.

### 2.2. The approximate model

In this context, and in continuation of the long history of turbulence model development, Schwarzkopf et al. [38] showed that the single point turbulence equations developed by Besnard et al. [5] (BHR model) could be applied to a range of self-similar turbulent mixing flows generated from different instabilities such as Rayleigh-Taylor, Kelvin-Helmholtz, and Richtmyer-Meshkov without changing the model coefficients. Key to the wide applicability of the model developed by Besnard et al. [5] was their consideration of a transport equation for density specific volume covariance $b$. This, in conjunction with full consideration of the Reynolds stress transport, allowed the proper description of anisotropy that is fundamental to buoyancy-driven turbulence. Nevertheless, a shortcoming of this model was its inability to properly represent turbulence growth rates encountered in settings that involve transient evolution of buoyancy-driven turbulence, as commonly encountered in flows dominated by the Raleigh-Taylor instability. This shortcoming is better understood by considering the popular $k$-$\epsilon$ class of turbulence models wherein the transport term is slaved to the decay length scale (by coefficients $C_\mu$ and $\sigma_k$). Consequently, like in the $k$-$\epsilon$ model, profiles of dissipation and transport of Reynolds stress (and density-specific volume covariance) are scaled versions of each other (in the BHR model). This is in contrast to the different nature of the scaling of transport and dissipation that is exhibited in the self-similar regime of Rayleigh-Taylor turbulence in [25]. To remedy this shortcoming, Schwarzkopf et al. [37] introduced a second turbulent length scale.

The RANS turbulence model discussed above contains several coefficients. These coefficients need to be calibrated so that a common set of model coefficients will allow reasonable comparisons of statistics over a wide range of turbulent flows, ranging from incompressible flows with single fluids and mixtures of different density fluids (variable density flows) to flows over shock waves. Schwarzkopf et al. [37] follow a recipe for calibrating the coefficients that considers a sequence

of simple flow configurations. The simple flow configurations considered were such that most of the terms in the equations could be neglected and the remaining non-zero terms were (mostly) different for each of the configurations considered. The specific sequence of canonical flows considered in the calibration process consisted of homogeneous isotropic decaying turbulence, homogeneous buoyancy-driven decaying turbulence, homogeneous shear (including Rapid Distortion Theory at high shear rates), wall bounded flow, Rayleigh-Taylor (RT) driven turbulence, shear driven turbulence, and shocked isotropic turbulence.

A RANS turbulence model simultaneously represents a number of physical processes (e.g., production dissipation, return to isotropy and rapid distortion of second order turbulent velocity correlations and others). Therefore, it seems important that the calibration process should account for the possibility that different combinations of such processes can lead to QoIs, either as observed in experiments or as computed in DNS. This amounts to requiring a comprehensive consideration of uncertainty in the coefficients (parametric uncertainty). The traditional calibration process adopted in [37], by producing one optimal combination of coefficient values, fails to account for such parametric uncertainty. In this sense, it is clear that a calibration process that accounts for such uncertainty will be better than the calibration process adopted in [37]. While we are currently working on a comprehensive Bayesian calibration of the RANS turbulence model that includes a full complement of test cases (e.g., as considered in [37]) that will not be the focus of the current article. The focus of this study, instead, is on demonstrating how the diagnostics resulting from the Bayesian calibration and analysis can be leveraged to improve turbulence modeling. For this, it suffices to consider a suite of homogeneous buoyancy-driven or Rayleigh-Taylor (hRT) turbulence test cases.

The RANS turbulence model equations for hRT turbulence is a set of ordinary differential equations:

$$\frac{d\tilde{R}_{11}}{dt} = \Big(2 - \frac{4}{3}C_{r1}\Big)a_1\frac{\bar{P}_{,1}}{\bar{\rho}} - C_{r3}\frac{\sqrt{K}}{S_D}\Big(\tilde{R}_{11} - \frac{2}{3}K\Big) - \frac{2}{3}\frac{K^{3/2}}{S_D} \tag{2.8}$$

$$\frac{d\tilde{R}_{22}}{dt} = \frac{2}{3}C_{r1}a_1\frac{\bar{P}_{,1}}{\bar{\rho}} - C_{r3}\frac{\sqrt{K}}{S_D}\Big(\tilde{R}_{22} - \frac{2}{3}K\Big) - \frac{2}{3}\frac{K^{3/2}}{S_D} \tag{2.9}$$

$$\frac{da_1}{dt} = (1 - C_{ap})\frac{b}{\bar{\rho}}\bar{P}_{,1} - C_{a1}\frac{\sqrt{K}}{S_D}a_1 \tag{2.10}$$

$$\frac{db}{dt} = -C_{b1}\frac{\sqrt{K}}{S_D}b \tag{2.11}$$

$$\frac{dS_D}{dt} = \frac{S_D}{\bar{\rho}K}\Big(\frac{3}{2} - C_{4v}\Big)a_1\bar{P}_{,1} - \Big(\frac{3}{2} - C_{2v}\Big)\sqrt{K}. \tag{2.12}$$

Here $t$ is time, $\bar{\rho}$ is the mean material density, $\bar{P_{,1}}$ is the mean pressure gradient in the direction of gravity, $\tilde{R}_{ij}$, $K$, $a_i$, and $b$ are the Favre average Reynolds stress, turbulent kinetic energy, mean density-weighted velocity fluctuation, and density-specific-volume, respectively, as defined previously. $S_D$ is the turbulent decay length scale. The remaining terms denoted by subscripted $C$s are model parameters and consist of $C_{r1}$, related to rapid return to isotropy; $C_{r3}$ related to slow return to isotropy; $C_{ap}$, related to rapid decay of mass flux; $C_{a1}$, related to slow decay of mass flux; $C_{4v}$, related to rapid growth/decay for hRT; and $C_{b1}$, related to decay of density-specific volume. $\tilde{R}_{11}, \tilde{R}_{22}, a_1, b$ and turbulent kinetic energy $K$ are the QoIs resulting from this model. We refer readers to Schwarzkopf et al. [37] for further details about the RANS turbulence model. For convenience, this set of equations may be written as

$$\frac{d}{dt}\mathbf{q}^{RO} = \mathbf{F}(\mathbf{q}^{RO}, \mathbf{u}^{RO}; \boldsymbol{\theta}), \tag{2.13}$$

where $\mathbf{q}$ is the vector of turbulent second-order moments of interest of the primitive variables $\mathbf{u}$, and $\boldsymbol{\theta}$ is the set of parameters. In this setting, the closure term $\mathbf{m}(\mathbf{u}^{RO}; \boldsymbol{\theta})$ in (1.3) is given by the turbulence model:

$$\mathbf{m}(\mathbf{u}^{RO}; \boldsymbol{\theta}) = \mathbf{m}(\mathbf{q}^{RO}(\mathbf{u}^{RO}; \boldsymbol{\theta})) \tag{2.14}$$

Here it is also understood that the RANS turbulence model is an approximation of

$$\frac{d}{dt}\mathbf{q}^{WR} = \mathbf{F}^{WR}(\mathbf{q}^{WR}, \mathbf{u}^{WR}), \tag{2.15}$$

but where $\mathbf{F}^{WR}$ is not explicitly known because of the dependence of $\mathbf{F}^{WR}$ on yet higher order moments and so on. As such, the closures introduced in (1.3) to approximate (1.1), equivalently approximations introduced in (2.13) to approximate (2.15), invariably lead to model error that may be attributed to structural inadequacy/error and parametric errors as discussed in the introduction. While we are interested in improving the accuracy of (1.3) by improving $\mathbf{m}$, since $\mathbf{m}$ is determined from $\mathbf{q}^{RO}$ as in (2.14), in the rest of this article, unless otherwise stated, we refer to the RANS turbulence model given by the set of equations (2.8) to (2.12) or equivalently (2.13) as the approximate model we consider and wish to improve.

## 3. Methods and results

There are multiple ways in which the set of parameters in (2.8)–(2.12) can be estimated given a set of data such as data from a few DNSs. For convenience, we broadly categorize the estimation methods based on two considerations: first, based on whether the estimation procedure pools all of the sample data together or not (see below), and, secondly, based on whether the method produces a point-estimate or a probabilistic or interval estimate. Since the distinction between methods based on the second consideration is evident, we do not devote a separate section to point them out; rather we discuss relevant details in the respective sections that present the results.

### 3.1. Pooled and unpooled analysis

To motivate the distinction between pooled and unpooled analysis, we note that we are interested in analyzing the RANS turbulence model given DNS simulations that have been previously performed at four different Atwood numbers: $\mathrm{At} \in \{0.05, 0.25, 0.50, 0.75\}$. The aim in performing such an analysis is to then examine the resulting point-estimates or posterior distribution of parameters in the turbulence model as a means to gain insights into the turbulence model itself. For example, if an *a priori* unexpected dependency is found in the study, it would lead us to further investigate the origin of the dependency in terms of turbulence phenomenology. If this investigation leads us to conclude that the dependency is a shortcoming of the model, further follow-up would consist of finding appropriate (possibly structural) modifications of the turbulence model that will remove the dependency while not introducing yet other "unexpected/unwanted" dependencies.

To this end, we note that the sample data is at different Atwood numbers, a parameter that characterizes the strength of the drive that is forcing the turbulent flow. At the same time, the RANS turbulence model in (2.8)–(2.12) does not explicitly depend on Atwood number. That is because of the local nature of the differential equation form of the RANS turbulence model; the RANS turbulence model subsumes such dependencies through its dependence on local turbulent correlations. That is, the RANS approach aims to model the flow evolution based purely on local behavior of density and related quantities. As such, assuming that the given turbulence model is correct leads us to consider the sample data as coming from the same population, and to estimate a single set of parameters for the different Atwood numbers. We call this the pooled analysis. A comparison of the QoIs with this point estimate would then allow us to estimate the effectiveness of the turbulence model.

If the RANS estimates of the QoIs do not compare well to the DNS estimates, two possibilities exist: either the sample data did not come from the same population (still implying a deficiency in the turbulence model) or the turbulence model itself is (more seriously) deficient. In order to examine the possibility that the sample data do not come from the same population, we need to perform an "unpooled" analysis. That is, we need to assume that the data at different Atwood numbers are coming from different populations, and redo the analyses to obtain a separate estimate for the set of turbulence parameters at each Atwood number. We call this the unpooled analysis. We note that while indeed partial pooling is also possible and useful in certain circumstances, it suffices to consider the two end members, viz. fully pooled or pooled and unpooled analysis, for our present purpose.

If the QoIs compare well with DNS with the unpooled analysis, then the dependence of the parameters on Atwood number would have to be investigated further to be able to remove it from the turbulence model. Needless to mention, a more serious shortcoming of the turbulence model would be implicated in the case of a poor comparison between the RANS and DNS estimates of QoIs on performing unpooled analysis.

**Algorithm 1:** Gradient-based point estimation algorithm. The gradient $\nabla_{\boldsymbol{\theta}}$ can be acquired using either forward sensitivities from finite difference, or adjoint sensitivities. For pooled estimates, the residuals are summed (in the L2 sense) over the Atwood numbers as well, whereas for the unpooled estimates, the procedure is conducted separately for each Atwood number.

**Input:** initial parameters $\boldsymbol{\theta}$, RANS estimate of QoIs as a function of $\boldsymbol{\theta}$, $\boldsymbol{q}_i^{RANS}(\boldsymbol{\theta})$, DNS QoI $\boldsymbol{q}_i^{DNS}$, step size $\lambda$ accuracy tolerance $\tau$;
**Output:** Point estimation of optimal parameter $\tilde{\boldsymbol{\theta}}$;
**for** $i \leftarrow 0$ **to** $N_{step}$ **do**
  $\epsilon = ||\boldsymbol{q}_i^{RANS} - \boldsymbol{q}_i^{DNS}||_2$;
  $\Delta\boldsymbol{\theta} = -\lambda \nabla_{\boldsymbol{\theta}} \epsilon$;
  **if** $\Delta\boldsymbol{\theta} < \tau\boldsymbol{\theta}$ **then**
    **break for**;
  **end**
  $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \Delta\boldsymbol{\theta}$;
**end**
$\tilde{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}$

### 3.2. Point estimates and results

In this category of estimation methods, the RANS model is used in conjunction with calibration data to calculate a single set of values for the turbulence parameters—a set that serves as the best estimate for those parameters given the data. Fig. 1

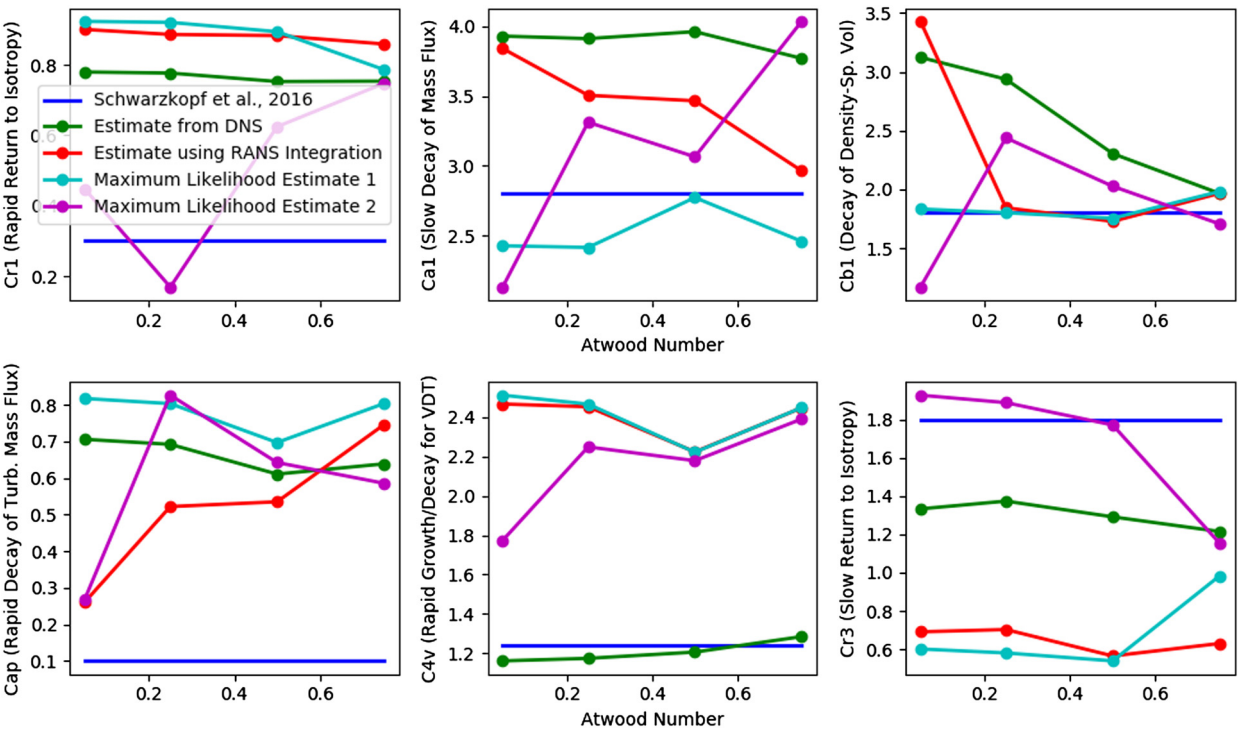



**Fig. 1.** Results from calibration of the RANS model using point estimates. The residual between the DNS and RANS estimates of QoIs (objective function) is minimized. Not only are the estimates of the various parameters different, but their variation with Atwood number is also different. Online version of figures is in color. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

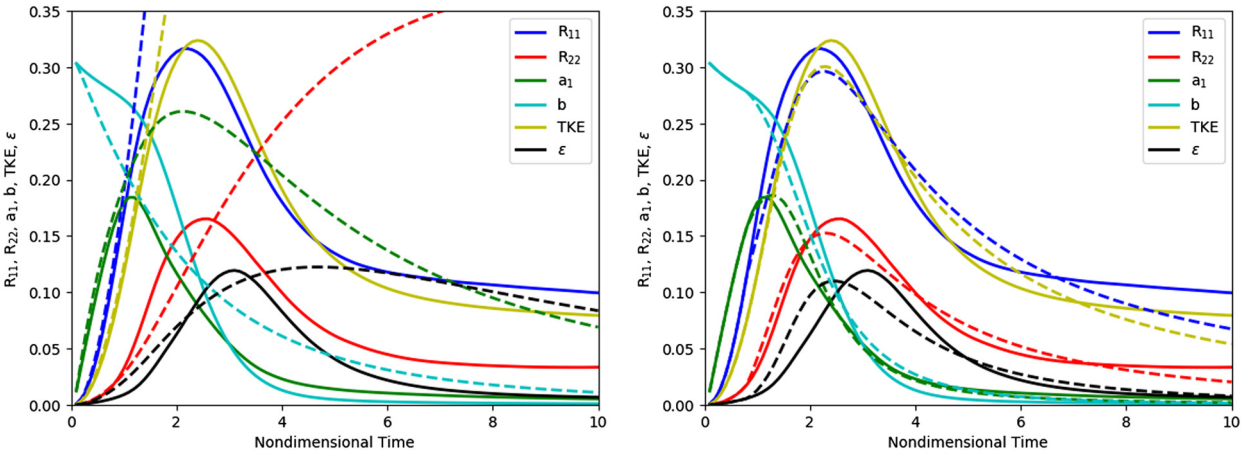



**Fig. 2.** QoIs using two of the point estimate-based calibrations. Solid lines are from DNS. Dashed lines are RANS estimates of the QoIs. The panel on the left uses the parameter settings of [37]. The panel on the right uses nonlinear least-squares to find the optimal set of parameters for each of the Atwood number cases separately (unpooled). In Table 1, the left panel corresponds to row 1, and the poor comparison of QoIs is indicated by larger residuals. The right panel corresponds to row 3, and the better comparison of QoIs is indicated by the reduced residuals. Atwood number is 0.50.

shows five different point estimates. The first estimate, in red, corresponds to that of Schwarzkopf et al., 2016. This may be considered as an example of a pooled point-estimate. Other pooled point-estimates are not shown to avoid clutter.

In the second approach all quantities other than the parameters in (2.8)–(2.12) are evaluated from the DNSs to obtain a set of algebraic equations in the unknown parameters. These equations are then solved in the least-squares sense to obtain the second, now unpooled, point-estimate of the parameter set (at each Atwood number for which the DNSs were available). This is shown in blue.

For the third estimate (green), an optimal set of parameters is sought at each Atwood number while integrating the RANS model. We note that we implemented estimation procedures using both forward sensitivities and adjoint sensitivities

**Table 1**
The (uniformly-weighted) combined residual at different Atwood numbers for the calibrations that are presented. Note that for the maximum likelihood estimates and all Bayesian estimates, the likelihood function (3.4) is different from the uniformly-weighted combined residual presented here.

| | Experiment | At=0.05 | At=0.25 | At=0.50 | At=0.75 |
|---|---|---|---|---|---|
| 1 | Schwarzkopf et al. (pooled) | 0.408 | 1.726 | 0.92 | 2.258 |
| 2 | Unpooled DNS Only | 0.144 | 0.645 | 1.088 | 1.792 |
| 3 | Unpooled RANS Integration | 0.023 | 0.143 | 0.293 | 0.384 |
| 4 | Unpooled Max. Likelihood 1 | 0.029 | 0.156 | 0.302 | 0.392 |
| 5 | Unpooled Max. Likelihood 2 | 0.069 | 0.166 | 0.385 | 0.476 |
| 6 | Pooled Bayesian | 0.060 | 0.224 | 0.678 | 1.596 |
| 7 | Unpooled Bayesian | 0.029 | 0.160 | 0.327 | 0.487 |
| 8 | Pooled Modified Bayesian | 0.028 | 0.169 | 0.500 | 0.439 |
| 9 | Unpooled Modified Bayesian | 0.029 | 0.160 | 0.331 | 0.498 |
| 10 | DNN-Surrogate Bayesian | 0.028 | 0.159 | 0.326 | 0.515 |

of the QoIs with respect to the parameters $\boldsymbol{\theta}$. Differences in these estimates were not significant, as expected, while noting that the adjoint sensitivities become more attractive when the number of parameters is larger than in the present case.

Next, we consider unpooled "maximum likelihood"estimates (MLE). The likelihood function is given by (3.4) and is described further in the section on Bayesian analysis. Two different estimates, (in cyan and magenta) are considered in this case to highlight another issue with point-estimates—their dependence on the initial value provided to the search algorithm.

The corresponding behavior of the QoIs is obtained on (re)integrating the RANS model using the estimated parameters. The uniformly-weighted combined residuals for the five QoIs at different Atwood numbers for these five estimates are presented in the first five rows of Table 1. Here the residuals are the differences between the DNS and RANS estimates of the QoIs.

We briefly discuss the nature of the point estimates from the point of view of the QoIs considered. First, the estimate from [37] is seen to result in poor comparisons of the RANS and DNS estimates of the QoIs. To a first order, the poor performance of the first estimate could be attributed to the fact that the calibration procedure of [37] considers a number of test-flows and produces one set of parameters to be used across them. Next, the second (now unpooled) estimate produces similarly poor comparisons in the QoIs (see Table 1). The poor performance of the second estimate can be understood as due to not fully considering the dynamical nature of the RANS model when estimating the parameter values [e.g., see [28]]. With the third estimate, the correspondence between RANS and DNS estimates of the QoIs is seen to be good. This is also true of the fourth and fifth (both unpooled) estimates. The reader is referred to the residuals in the first five rows of Table 1 and they are seen to be consistent with the above characterization of the parameter estimates. QoIs of only two of the five parameter estimates, estimate one (pooled) and estimate three (unpooled), are shown in Fig. 2 for reference.

The good correspondence between RANS and DNS estimates of QoIs seen in the right panel of Fig. 2 and in the third through fifth rows of Table 1 with the third, fourth, and fifth estimates establishes that the turbulence model being considered has the phenomenology required to represent the class of flows being considered. Furthermore, along the lines of reasoning for the poor performance of the first (pooled) estimate, the better performance of the third through fifth estimates may again be understood as due to the fact that the calibration was specific to each of the different Atwood number cases (unpooled).

Next, as discussed in Sec. 3.1, the favorable outcome with some of the unpooled estimates leads us to focus attention on the variation of the parameter estimates with Atwood number. Large variations with Atwood number are seen in most of the six parameters when the third through fifth estimates are considered. This would suggest that multiple aspects of the turbulence model have to be modified. However, as good as the match is between DNS and RANS at these estimates, we note that this estimation procedure does not account for parametric uncertainty discussed previously. We therefore turn our attention next to probabilistic estimates.

### 3.3. Bayesian analysis methodology

In the introduction and in the preceding parts of this section, we briefly considered the role of calibration in the overall process of validation and discussed how uncertainty has an important role to play. Here, we outline the Bayesian approach to calibration wherein uncertainty is quantified in a probabilistic sense. In the Bayesian framework, calibration of a model is formulated as an inference problem. Thus, the probability of (the vector of) the model parameters $\boldsymbol{\theta}$, given experimental or DNS data or QoIs $\boldsymbol{q}$, is determined using the Bayes rule as

$$P(\boldsymbol{\theta}|\boldsymbol{q}) = \frac{P(\boldsymbol{q}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{\int P(\boldsymbol{q}|\boldsymbol{\theta}')P(\boldsymbol{\theta}')d\boldsymbol{\theta}'} \tag{3.1}$$

$$\propto L(\boldsymbol{q}|\boldsymbol{\theta})P(\boldsymbol{\theta}) \tag{3.2}$$

In the above equation, it is common to call

- $P(\boldsymbol{\theta})$ the prior distribution of the model parameters.

- $L(\boldsymbol{q}|\boldsymbol{\theta})$ the likelihood function, and
- $P(\boldsymbol{\theta}|\boldsymbol{q})$, the posterior distribution of the model parameters

An analytical approach to Bayesian estimation requires the integration of the normalization term (denominator in (3.1)). This is difficult to evaluate when the probability model structure is complex. However, algorithms based on the Monte-Carlo Markov Chain (MCMC) allow for discretely sampling points from the posterior distribution, bypassing the need for the integration of the normalization factor. The only requirement for performing MCMC based sampling is that a function proportional to the original probability distribution be specified. This is shown in (3.2). To assist in the illustration of the Bayesian estimation process, we further make precise the QoIs vector $\mathbf{q}$ as follows:

- $\boldsymbol{q}_i^{DNS} \in \mathbb{R}^{N_t}$ is the $i$-th QoI vector from the DNS simulation data. $i \in \mathbb{Z}, i \in [1, N_q]$ where $N_q$ is the number of QoIs in the problem. $N_t$ is the total number of time-steps. Denote this as the truth vector.
- $\boldsymbol{q}_i^{RANS}(\boldsymbol{\theta}) \in \mathbb{R}^{N_t}$, results from RANS model, given model parameters $\boldsymbol{\theta} \in \mathbb{R}^{N_p}$ where $N_p$ is number of RANS model parameters. Denote this as the model vector.

*Likelihood function and discrepancy model* We parameterize the likelihood function $L(\boldsymbol{q}|\boldsymbol{\theta})$ as a multivariate Gaussian density function centered around the truth vector. For the $i$-th QoI,

$$L_i(\boldsymbol{q}|\boldsymbol{\theta}) = f(\boldsymbol{q}_i^{RANS}(\boldsymbol{\theta}); \boldsymbol{q}_i^{DNS}, \boldsymbol{\Sigma}_i) \tag{3.3}$$

$$= \frac{1}{(2\pi)^{N_t/2}|\boldsymbol{\Sigma}_i|^{1/2}} \exp\Big( -\frac{1}{2}(\boldsymbol{q}_i^{RANS} - \boldsymbol{q}_i^{DNS})^T \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{q}_i^{RANS} - \boldsymbol{q}_i^{DNS})\Big) \tag{3.4}$$

Here $\boldsymbol{\Sigma}_i$ is the covariance matrix (or equivalently $\boldsymbol{\Sigma}_i^{-1}$ is the precision matrix) for the $i$-th QoI. That is, when comparing $\boldsymbol{q}_i^{RANS}$ to $\boldsymbol{q}_i^{DNS}$, we model the discrepancy between them that arises due to various reasons including model error as $\boldsymbol{q}_i^{DNS} = \boldsymbol{q}_i^{RANS} + \epsilon_{it}$, with $\langle \epsilon_{it}\epsilon_{is}\rangle = \boldsymbol{\Sigma}_i$ and where $\boldsymbol{\Sigma}_i(t,s)$ represents the assumed covariance structure of the discrepancy. We present results here for a simple parameterization of $\boldsymbol{\Sigma}_i(t,s)$ as $\sigma_i\delta(t-s) \equiv \sigma_i \boldsymbol{I}$, and where $\sigma_i$ is a new hyper-parameter and $\boldsymbol{I}$ is the identity matrix. Clearly, while more complicated parameterizations such as Gaussian processes [e.g., see [30,29]] and other forms can be used for the parameterization of the covariance, the simple form for the discrepancy that we use here, including its independence of $\boldsymbol{\theta}$ follows from our *a priori* lack of knowledge about such dependencies including *a priori* lack of knowledge of model error. Finally, a limited amount of experimentation with the more complicated forms of the covariance showed robustness of the results we find with respect to the form of covariance.

Finally, we define the likelihood function of all QoIs as a sum of the individual likelihood functions, i.e.,

$$L(\boldsymbol{q}|\boldsymbol{\theta}) = \sum_{i=1}^{N_q} L_i(\boldsymbol{q}|\boldsymbol{\theta}) \tag{3.5}$$

where $\boldsymbol{\theta}$ is now augmented to include $\boldsymbol{\sigma}$, the $N_q$ dimensional vector of hyper-parameters $\sigma_i$: $\boldsymbol{\theta} = (\boldsymbol{\theta}^p, \boldsymbol{\sigma}^q)$. Finally, given the parameter vector $\boldsymbol{\theta}$, $\mathbf{q}^{RO}$ is obtained by accurate numerical integration of (1.3).

*Prior distributions* To ensure robustness of the analysis and results presented, all computations were performed with two different forms for the prior distribution for parameters $\boldsymbol{\theta}^p$, a normal distribution and a uniform distribution. The priors are always centered at the previous estimates of Schwarzkopf et al. [37]. As mentioned in the introduction, if the processes associated with the parameters play an important role in the class of flows we consider, we use weak, uninformative and diffuse priors; for other parameters, the width of the priors are broadly determined by physical bounds such as trying to limit regions of negative turbulent kinetic energy. Broad normal distributions are used for $\boldsymbol{\sigma}^q$.

*MCMC using delayed rejection adaptive metropolis (DRAM)* We use the DRAM algorithm [21], which is an improved MCMC algorithm, for drawing samples from the posterior distribution. DRAM combines two ideas that improve on the Metropolis-Hasting type MCMC algorithm, Delayed Rejection [40] and Adaptive Metropolis [20], whose efficiency in many scenarios outperform the original methods. The pseudo-code for the algorithm is given in Algorithm 2, and we limit $N_{stage}$ to two.

### 3.4. Results from pooled Bayesian inference

As discussed in Sec. 2, a single-point RANS turbulence model models all turbulent phenomena in a local fashion. That is, each of the turbulent processes is parameterized in terms of the local values of relevant variables and their gradients. Thus, given such locality assumptions, it is possible to imagine that the effects of Atwood number would enter implicitly through the gradients of density and gradients of other variables that are dependent on density. It may, therefore, be anticipated that the coefficients themselves should not depend on the Atwood number. Consequently, we first perform such a calibration.

The marginals of the posterior distribution of parameters resulting from such a calibration is shown in the dashed magenta line in the panels of Fig. 3. It is seen in this figure that the DNS data have moved the prior distributions—centered

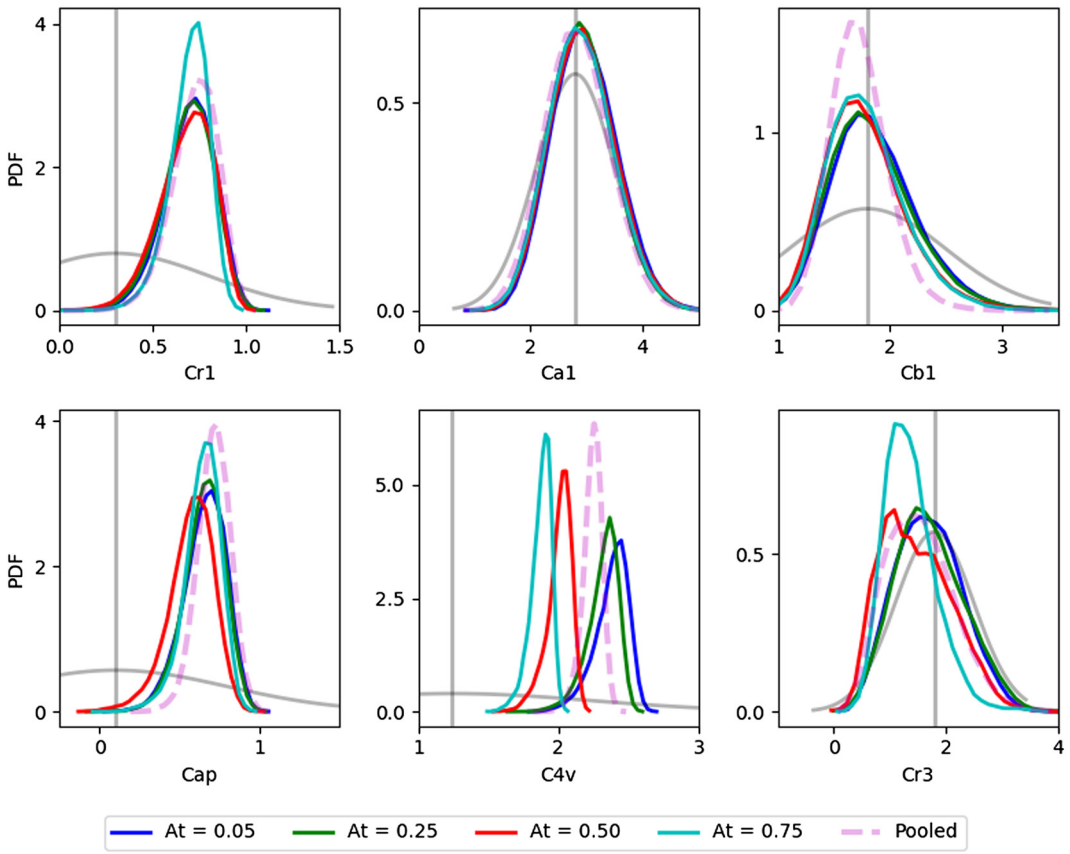
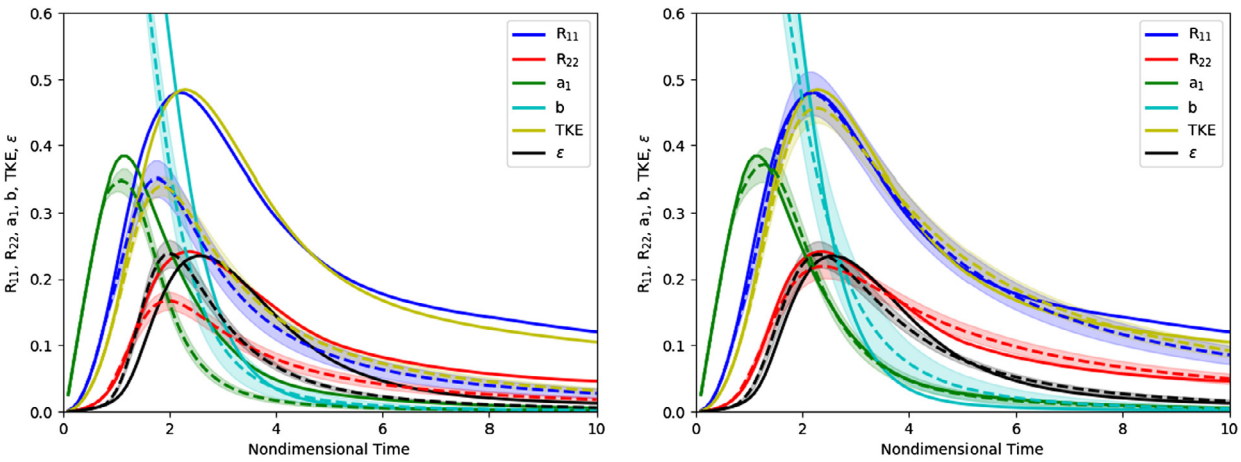



**Fig. 3.** Posterior distribution of the parameter values for the RANS model inferred in a Bayesian framework when calibrating the RANS model against DNS runs in pooled and unpooled scenarios. The colored curves are labeled in the legend. The prior distributions are shown in gray curves. Vertical gray lines are drawn at the values from [37]. Only marginal distributions are shown for convenience.




(a) Pooled analysis (b) Unpooled analysis

**Fig. 4.** Comparison of QoIs (between DNS and RANS at one Atwood (0.75) between pooled and unpooled analyses. As in Fig. 2, DNS results are shown in solid lines. Here the posterior distributions of parameter values are propagated through the RANS model to obtain the posterior distributions of the QoIs. The dashed lines now represent the mean over the posterior distribution of RANS estimates of the QoIs and the shading corresponds to $\pm$ one standard deviation. As expected the QoIs in the RANS simulations are seen to fit DNS data better in the unpooled analysis as compared to the fits in the pooled analysis.

**Algorithm 2:** Bayesian estimation of parameters using the Delayed Rejection Adaptive Metropolis (DRAM) variant of the random-walk Metropolis-Hastings algorithm used in this study. $N_{stage}$ is limited to 2, and only the second half of the chain is used. As with Algorithm 1, for pooled estimates, the likelihood at different Atwood numbers are combined by multiplying them together, whereas for unpooled estimates the procedure is repeated individually at each Atwood number.

**Input:** Initial point $\boldsymbol{\theta}^{(0)}$, initial covariance $\boldsymbol{\Sigma}^{(0)}$,
function to compute posterior probability $P(\boldsymbol{\theta}|\mathbf{q}^{WR})$;
**Output:** Bayesian estimates for parameters $\{\boldsymbol{\theta}^{(i)}|i=1,2,\cdots,N_{samp}\}$;
**for** $i \leftarrow 1$ **to** $N_{samp}$ **do**
  **for** $t \leftarrow 1$ **to** $N_{stage}$ **do**
    $\boldsymbol{\Theta}^{(t)} = \boldsymbol{\theta}^{(i-1)} + \boldsymbol{\xi}^{(t)}, \quad \boldsymbol{\xi}^{(t)} \sim \boldsymbol{N}(\mathbf{0}, \frac{1}{3^{(t-1)}}\boldsymbol{\Sigma}^{(i-1)})$;
    $\alpha^{(t)} \leftarrow \alpha(\boldsymbol{\Theta}^{(k)}, P(\boldsymbol{\Theta}^{(k)}|\boldsymbol{q}), P(\boldsymbol{\theta}^{(i-1)}|\boldsymbol{q}),\ k=1\ \textbf{to}\ N_{stage})$ ;
    // Compute acceptance probability at stage t. e.g., $\alpha^{(1)} \leftarrow P(\boldsymbol{\Theta}^{(1)}|\boldsymbol{q})/P(\boldsymbol{\theta}^{(i-1)}|\boldsymbol{q})$ ;
    **if** $u \sim U(0,1) \leq \alpha^{(t)}$ **then**
      $\boldsymbol{\theta}^{(i)} \leftarrow \boldsymbol{\Theta}^{(t)}$ // accept;
      **break for**;
    **end**
    **if** $i = N_{samp}$ **then**
      $\boldsymbol{\theta}^{(i)} \leftarrow \boldsymbol{\theta}^{(i-1)}$
    **end**
  **end**
  **if** $i < N_{samp}/2$ **then**
    $\boldsymbol{\Sigma}^{(i)} \leftarrow \eta\left(cov(\boldsymbol{\theta}^{(0)}, \cdots, \boldsymbol{\theta}^{(i)}) + \delta\boldsymbol{I}\right), \quad \delta = 10^{-5}$ // update proposal covariance periodically using entire chain up to half of total samples;
    // Adapt scaling factor $\eta$ towards 23.4% acceptance after starting at $2.4^2$/dimension($\boldsymbol{\theta}$);
  **else**
    $\boldsymbol{\Sigma}^{(i)} \leftarrow \boldsymbol{\Sigma}^{(i-1)}$
  **end**
**end**

at the point-estimate of Schwarzkopf et al. [37]—significantly away from that point estimate and greatly narrowed the distributions from their prior values for three of the six parameters. In order to ensure robustness of these inferences, the pooled analysis was repeated with flat (uniform distributions with large bounds) priors; the results did not change significantly and the results are not shown to avoid redundancy.

The comparison between RANS and DNS estimates of the QoIs are shown for only one of the Atwood numbers (0.75) in the left panel of Fig. 4, again in order to limit the number of figures; large differences are seen and this is true at other Atwood numbers as well. The reader is referred to the lower five rows of Table 1 for the actual size of the residuals at different Atwood numbers for the different Bayesian calibrations considered. Residuals for the pooled Bayesian calibration are shown in the sixth row.

As discussed earlier, the large differences seen in the QoIs leads to two possibilities: either that the model has serious deficiencies or that the model has possibly a less serious deficiency in being unable to simultaneously properly represent flows at different Atwood numbers.

### 3.5. Results from unpooled Bayesian inference

In order to find out the origin of the model deficiency and its significance, following the discussion in Sec. 3.1, we next consider unpooled or independent calibration of the RANS model at each of the Atwood numbers. The posterior distribution of parameters resulting from this unpooled calibration is shown in solid lines in Fig. 3. As in the pooled analysis and consistent with that inference, a significant shift away from the prior distribution, both in terms of mean and variance, is seen in these cases as well. Again, computations with a uniform prior distribution with large bounds did not produce any significant difference ensuring robustness of the inferences.

Comparisons of the RANS model fits to the DNS QoIs are shown in the right panel of Fig. 4 (for the same Atwood number as in the left panel). The fits are seen to be much better than with pooled inference (left panel).

### 3.6. Variation of the posterior distribution of parameters with Atwood number and its quantification

Both the unpooled Bayesian inference and the unpooled point estimate presented earlier succeed in fitting the DNS estimates of the QoIs well. However, a significant difference between the unpooled Bayesian and unpooled point-estimate is that, while with the unpooled point estimate four of the six parameters displayed large variations with Atwood number, it is seen that with the unpooled Bayesian inference only one of the parameters displays a systematic variation with Atwood number (lower middle panel of Fig. 3 vs. red, cyan, and turquoise lines in Fig. 2).

While it is visually clear that distributions of $C_{4v}$ under different Atwood numbers are distinctly different, we quantify the distributional shift of the posterior PDF of model parameters resulting from the change in Atwood number using the

**Table 2**
Janson-Shannon divergence for marginal distributions of RANS model at different Atwood numbers. $C_{4v}$, as compared to other parameters, have a strong distributional shift across Atwood numbers, suggesting a particular shortcoming of the RANS model: modeling the evolution of the turbulent length scale using (2.12).

| Parameter | $C_{r1}$ | $C_{a1}$ | $C_{b1}$ | $C_{ap}$ | $C_{4v}$ | $C_{r3}$ |
|---|---|---|---|---|---|---|
| JS-Divergence | 0.051 | 0.035 | 0.037 | 0.085 | 1.125 | 0.100 |

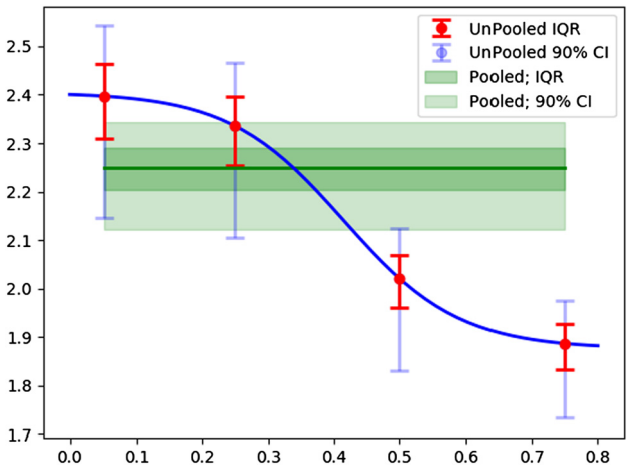



**Fig. 5.** Variation with Atwood number of RANS parameter $C_{4v}$ as inferred by Bayesian calibration.

general Jensen-Shannon divergence. The general Jenson-Shannon divergence is a distance metric that measures the similarity between $n$ probability distributions as:

$$\mathrm{JSD}_{\pi_1,\cdots,\pi_n}(P_1,\cdots,P_n) = H\Big(\sum_{i=1}^{n}\pi_i P_i\Big)P_i - \sum_{i=1}^{n}\pi_i H(P_i), \tag{3.6}$$

where $\pi_1,\cdots\pi_n$ are the weights for each probability distribution $P_1,\cdots,P_n$. Here, $n=4$ for the four different Atwood numbers, and we choose uniform weighting: $\pi_1=\pi_2=\pi_3=\pi_4=1/4$. $H(P)$ is the Shannon entropy for distribution $P$. The Shannon entropy is defined as:

$$H(P) = -\sum_{i=1}^{n} P_i \log P_i \tag{3.7}$$

The values are tabulated in Table 2.

The dependence of parameter $C_{4v}$ on Atwood number seen in the lower middle panel of Fig. 3 is better visualized in Fig. 5. A similar plot for the computations with a uniform form for the priors shows little difference.

Next, we note that, after performing the above Bayesian analysis, in order to further investigate the reason for the different dependencies we found with the least-squares point-estimate and the Bayesian inference, we considered the maximum likelihood estimate (MLE) which uses the same likelihood function as in the Bayesian analysis. Those estimates are shown in Fig. 1 and were discussed previously.

Finally, on a related note, in the probabilistic estimation context, we show the marginal posterior distributions for ease of visualization. However, the joint distribution can have a more complicated structure than in one that may be imagined by composing the marginal distributions alone. This is the case when, as is typical, the turbulence modeling parameters are not all independent. This is consistent with, and is borne out by, the fact that none of the point estimates (see Fig. 1) happen to capture the mode of the joint distribution that emerges from the Bayesian analysis. Indeed, after the Bayesian analyses were conducted, a limited amount of experimentation with "basin-hopping" extensions to gradient-based point estimation methods did not help such an "hybrid" approach to capture the mode of the Bayesian analysis either. Such hybrid approaches, however, need to be further investigated.

### *3.7. Improving the turbulence model based on results of Bayesian analysis*

As discussed above, pooled and unpooled Bayesian analyses of the RANS model given a few DNSs at different Atwood numbers helped reveal a dependence of Atwood number. That is, unlike with the point estimates, large and systematic variations were confined to just one of the parameters when Bayesian calibration was used. While is clearly beyond the

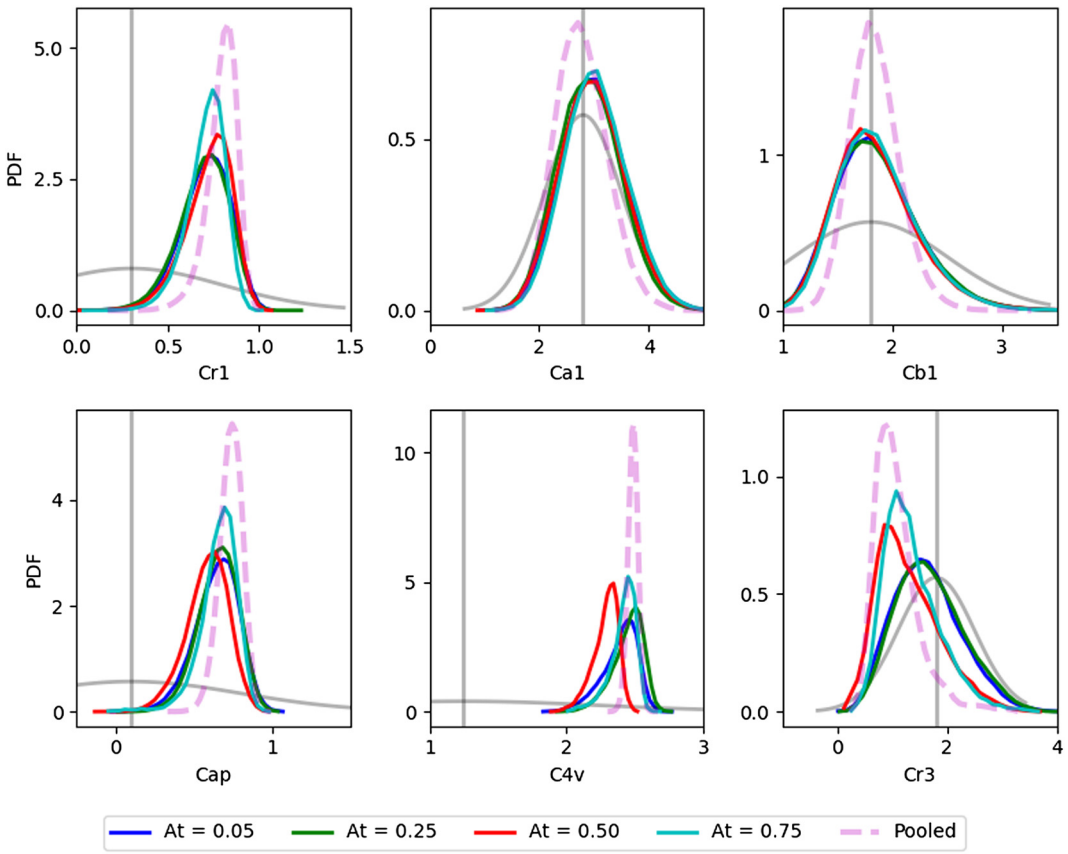



**Fig. 6.** Results from pooled and unpooled Bayesian analysis of the modified RANS model. For the first time, the modified model permits fitting of all Atwood number cases with a single set of parameter values (pooled). See Table 1. Results from earlier Bayesian analyses of the original RANS model is used to obtain the modification to the RANS model.

scope of this article to dwell on various optimization strategies and their respective advantages and disadvantages, we note that in contrast to point estimates that typically rely on a local search algorithm, a global sampling strategy that underlies the Bayesian methodology is one of the reasons why the latter framework is more robust. The flip side of this is a large added computational cost and which we will address shortly. We now turn our attention to how the results of the Bayesian analyses can be leveraged to improve the turbulence model.

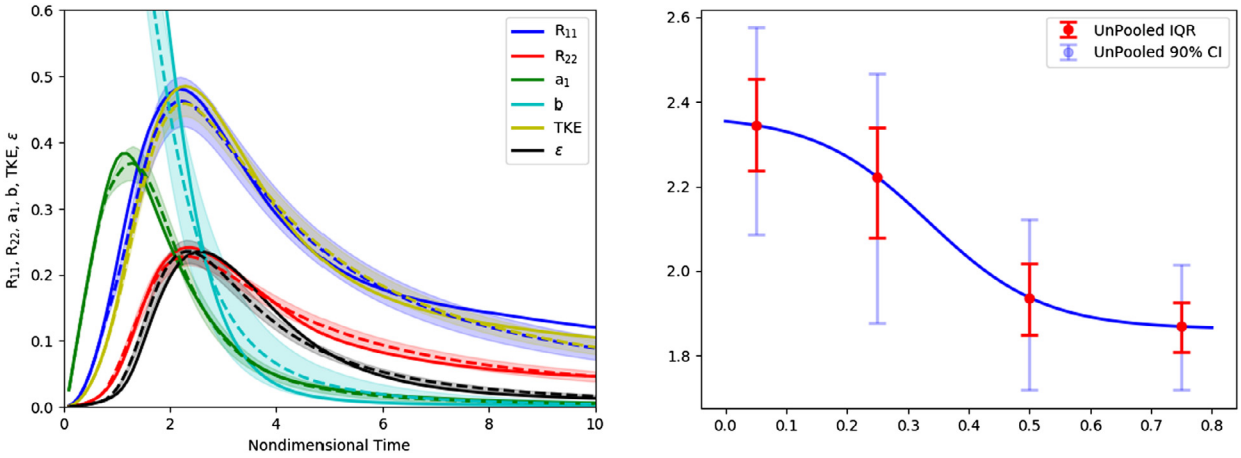
A straightforward way to do this would be to include or build in the discovered variation of $C_{4v}$ with Atwood number. However, as previously discussed, Atwood number itself is either a feature of the initial condition or, at later times, a two-point or global feature of the flow. As such, encoding the Atwood number dependence of a parameter would be inappropriate in the one-point turbulence closure model that is being considered. Therefore, we look into being able to modify the turbulence model based on a local variable/correlation instead.

We discussed earlier that in this buoyancy-driven flow, the Atwood number is a measure of the strength of the forcing. Next, towards identifying local-variable proxies for Atwood number, consider the temporal variation of various quantities in Figs. 2 and 4. It is sufficient to consider the DNSs for this purpose (and confine attention to any one panel in each of the two figures). In these figures, it is seen (a) that the only non-zero variable/correlation in the initial condition is $b$, (b) that the magnitude of $b$ increases with At (Fig. 2 is at At = 0.50 and Fig. 4 is at At=0.75). These observations suggest that $b$ may be a useful local proxy for Atwood number. Consequently, we examine the variation of the maximum of $b$ (equivalently its initial value since it decays monotonically with time) with Atwood number and find a quadratic dependence ($b_{max} \propto At^2$).

The identification of a local proxy for Atwood number paves the way for making the transition to a structural modification of the turbulence model starting from the discovered parametric dependence: we now consider a new term in (2.12) of the form

$$\frac{2}{3}\frac{S_D}{\bar{\rho}K}\sqrt{b}a_1\bar{P}_{,1}. \tag{3.8}$$

We repeat the pooled and unpooled Bayesian analysis of the modified turbulence model. The posterior probability distributions of the parameters are shown in Fig. 6. First, the systematic dependence of parameter $C_{4v}$ (lower-middle panel) on Atwood number that is seen in the corresponding panel of Fig. 3 is now absent. Second, no new dependence is seen




(a) Temporal trajectory of QoIs.

(b) Marginal distribution of $C_{4v}$ parameter vs. Atwood number

**Fig. 7.** Calibration with the RANS surrogate. The plots above are acquired by MCMC sampling on the response surface from the surrogate by the neural network. The temporal trajectories of QoIs closely resemble those acquired from the RANS model, and similar variations of the marginal distribution of the $C_{4v}$ parameter can be observed from sampling the surrogate.

to be introduced in the unpooled analysis. Furthermore, the RANS fits for the QoIs at each Atwood number are good for both pooled and unpooled calibrations. This can be seen in rows 8 and 9 of Table 1. While the good fits of the QoIs for the unpooled analysis verify that aspect of the original model is retained, the modified turbulence model now allows, *for the first time*, the *pooled* calibration to fit the QoIs at each of the four Atwood numbers well (using a single set of parameters). This is a significant improvement of the turbulence model.

### *3.8. Surrogate RANS model using neural networks*

As mentioned earlier, a disadvantage of the Bayesian approach is the added computational cost. The particular nature of the turbulent flow that we chose to analyze in this article was such that the RANS model could be integrated cheaply. As such, the added computational cost involved in the Bayesian analysis was not much of an issue. However, other turbulent flow settings may not be as forgiving in that a single integration/realization of the RANS model could be computationally expensive enough that the added computational cost of the Bayesian analysis may be prohibitive. New computationally-efficient sampling techniques may help alleviate this problem. However, we consider a possible alternate strategy—that of a surrogate RANS model.

Artificial Neural Networks (ANN) have been studied extensively for their predictive capabilities in the context of machine learning, and a feed-forward, multi-layered neural network has been shown to be capable of approximating a continuous function arbitrarily well [9]. Given our current problem, the use of an ANN for response surface exploration is of particular interest to us; this subject has been well-explored [e.g., see [1]].

In this study, we use an ANN to predict the temporal trajectories of turbulent correlation (see the RANS governing equations (2.8)–(2.12)) given the parameters $\boldsymbol{\theta}$. That is, we use the ANN to establish a mapping from the RANS parameters to the trajectories of turbulent correlations. Below, we briefly discuss some of the main aspects of developing the surrogate model including data compression, architecture of the neural network, and validation of training quality using a held-out test set.

*Compression of the output vector* While we are interested in predicting the full temporal trajectories of the QoIs (e.g., the trajectories shown in Fig. 4), the temporal trajectories contain a large number of time steps resultingly in a high-dimensional output vector. At the same time, the smoothness of these trajectories suggests temporal over-sampling and leads us to consider principle components as a means of compression. Therefore, we reduce the machine learning problem to predicting the PCA coefficients, and recover/reconstruct the full trajectories using the PCA bases used for the decomposition. Since four PCA components are found to explain in excess of 99% of the variance, we restrict ourselves to learning four components for each of five QoIs.

*Architecture of the neural network* We use a deep fully-connected neural network (DNN, a.k.a., multi-layer perceptron) with 10 hidden layers and with each layer consisting of 30 neuron units. The size of the input layer corresponds to that of the parameter vector and given the compression discussed above, the size of the output layer is 20.

*Validation on test set* We evaluate the performance of our neural network by using a held-out test set. The size of the test set is held fixed at 10% the size of the training set and both sets are sampled from the same distribution.

*Results* After training the ANN, we conducted unpooled Bayesian calibration of the RANS model, but now using the DNN-RANS surrogate instead of integrations of the actual RANS model. Representative sample results are shown in Fig. 7. With the DNN-surrogate, it is seen that not only are the fits to the DNS estimates of the QoIs good, but that the variation of the $C_{4v}$ parameter with Atwood number is also well approximated.

In order to avoid a long digression, we skip details, but note that convergence of a Markov chain to the target posterior distribution is typically slow. For this reason, chains of length a million or more samples are typically used to ensure robustness of inferences. For the training/testing phase of the DNN-surrogate, we used a set of 500 integrations of the RANS model. After the DNN-surrogate has been trained it is fairly cheap to evaluate it. Thus in a situation where individual RANS integrations are costly this strategy can lead to significant computational savings. We do not dwell on actual differences in computational costs in this proof-of-principle demonstration for the reason that the RANS model itself is presently computationally cheap. We expect the very low end of cost advantages of the surrogate approach in cases where individual RANS integrations are expensive to be ten or more; more realistic estimates require actually working through such a case.

## 4. Discussion and conclusion

In this study, we considered various methods of analyzing and calibrating an approximate model of a complex multiscale system given a few well-resolved simulations. We first considered various point-estimate approaches. These ranged from using previously published point estimates, to new ones that used ordinary least squares and maximum likelihood estimates. While it was found that some of the unpooled point estimates were able to fit the QoIs well, all pooled point estimates and some unpooled point estimates produced poor estimates of the QoIs. We rationalized the behavior of the point estimates by appealing to the effects of pooling versus unpooling of the different Atwood number cases and the effects of considering or not considering the dynamical nature of the approximate model.

Next we considered the behavior of the parameters themselves in the cases that produced good estimates of the QoIs. Not only were the parameter estimates themselves different, depending on the method used, but their variation with Atwood number was also very different. For these reasons, the main insight into the approximate model provided by the various point estimates was limited to indicating that the model under consideration contained enough of the relevant phenomenology to properly represent the particular flow that we consider, but that they had to be individually calibrated. While valuable, it did not point in further specific ways of how the model could be improved.

We next conducted Bayesian analysis of the approximate model using the same data as was used for the point estimates. Not surprisingly, pooled and unpooled Bayesian analyses first verified and confirmed the insight provided by the point estimate approaches. That is that while the approximate model could produce QoIs that compare well with DNS estimates at each Atwood number, they are unable to do so using a single set of, now probabilistic or interval estimates. Next, however, this approach provided further specific pointers towards how this shortcoming could be overcome. Compared to the point estimates where different variations with Atwood number were produced with different variants of the estimators, the Bayesian analyses produced similar posterior distributions at all Atwood numbers for all but one of the model parameters.

The minimization of parameter variability with Atwood number may be thought of as the key contribution of the Bayesian analyses as far as insights into the workings of the approximate model is concerned. This is because, in contrast to the wide range of parameter variations seen in point estimates, the systematic variation of the posterior distribution of a minimal set of parameters (in this case, a single one) suggests that the former is likely an artifact of the point-estimate approach to calibration. Furthermore, a small set of systematic parameter variations allows for an immediate improvement of the model (as we demonstrate), with initial improvements being based purely on statistical grounds. However, such improvements could then foster better physics- or dynamics-based improvements. Given the non-uniqueness of closures, if multiple such improvements should arise, then it is possible that one can use the principle of Occam's razor in conjunction with plausibility to choose from among them [17].

Next, we consider two aspects of computational cost associated with the procedure that we have considered and find useful: first the relative difference in cost between point estimates and probabilistic estimates and next, the cost of the approximate model itself. Not surprisingly, probabilistic estimates when performed using sampling are costlier than point estimates. However, this difference in cost can be greatly reduced by using variational methods of probabilistic inference, an area that has seen major improvements in the last couple of years. We present a few details below.

When using random-walk (Metropolis-Hastings) based algorithms for generating members of the Markov chain, the stepsize of MCMC scales inversely as the dimension of $\boldsymbol{\theta}$ (number of parameters being inferred): if h is the stepsize, $h \sim 2.4^2/\text{dimension}(\boldsymbol{\theta})$ [18]. Consequently, when the dimension of $\boldsymbol{\theta}$ is large, a larger number of MCMC steps have to be taken so that cost scales as $dim(\boldsymbol{\theta}) \times C_{step}$ where $C_{step}$ is the computational cost of a step. Thus, for a reasonable chain length, the computational cost of the MCMC-based scheme scales as, say, $10^3 \times dim(\boldsymbol{\theta}) \times C_{step}$. However, by similar straightforward scaling arguments, the cost of a point estimate scales only as, say, $10^2 \times C_{step}$, leading to the probabilistic estimate being costlier by a factor of about $10 \times dim(\boldsymbol{\theta})$. So for $dim(\boldsymbol{\theta}) \sim 10$, as in the present study, the cost difference is about 100 and the study was computationally feasible. However, this becomes a problem when $dim(\boldsymbol{\theta})$ is large.

Recent improvements in computational inference methods, however, can be brought to bear on this issue. For example, computational cost of Hamiltonian Monte Carlo only scales as, say, $10^3 \times C_{step}$ [e.g., see [4]]. Even better, variational inference methods have now been developed that bring the cost of probabilistic inference to levels comparable to that of point estimates [e.g., see [6]]. However, it has to be noted that with variational inference, an additional error is introduced since optimal parameters for the source distributions that best match the actual posterior distribution will be found, rather than the actual posterior distribution itself as we currently compute. Nevertheless, we think that the ideas presented in this article can be scaled up to larger problems by the use of probabilistic inference (but with caveats such as the one mentioned above).

Next, we consider the issue of computational cost of one step in the inference procedure, viz. $C_{step}$ above. It is dominated by the cost of evaluating the likelihood function which in turn requires an integration of the approximate model. If an individual integration of the approximate model itself is computationally intensive (while still being orders of magnitude cheaper than well-resolved simulations) then conducting either point estimate based analysis or probabilistic estimate based analysis can be computationally expensive. Towards reducing such costs, we explored the use of an artificial neural network as a surrogate model for the RANS solver and we demonstrated the feasibility of using such a ANN-based surrogate in conducting the Bayesian analysis: the Bayesian analysis using the ANN-surrogate recovered similar parametric dependencies as the Bayesian analysis that was performed using RANS integrations. Thus in a situation where individual integrations of the approximate model are themselves costly we anticipate that this strategy can lead to significant enough computational savings that the Bayesian analysis can be performed.

In other considerations, both the well-resolved model and the approximate model were deterministic in the case we chose to illustrate the process of improving the approximate model using probabilistic or interval inference. However, since the methods that we use are agnostic about the deterministic or stochastic nature of either the well-resolved model or the approximate model, application of this methodology to situations wherein either the well-resolved model or the approximate model is stochastic is easily achieved on taking into consideration prior information available about the nature of the stochasticity involved. That is, for example, we anticipate that the procedure for model improvement should work in the setting of the Markov State Model or other such models in the context of data coming from either a deterministic or stochastic, but well-resolved model. We also note that we have leveraged certain aspects of Bayesian estimation to serve the purpose of model improvement. It remains to be seen if other aspects resulting from Bayesian estimation, such as ensemble characteristics, can also be exploited for similar purposes.

Bayesian analysis can be used in many ways to help with the task of reduced-order modeling but, the fundamental aspect of Bayesian analysis that permits its use in such a fashion is related to its ability to create knowledge [e.g., see [2,16,15,17]]. A few recent studies have highlighted such uses and include using Bayesian approaches to estimating and characterizing errors and uncertainty in RANS models and comparing and selecting from among different models. We add to this nascent body of literature by showing how Bayesian analysis can be leveraged to improve models themselves: the pooled and unpooled Bayesian analysis we conducted revealed unanticipated parametric dependencies, and then we closed the analysis-improvement loop by using the discovered dependency to effect a structural modification of the model that removed the dependency while not introducing others and in a fashion that is consistent with the modeling approach [see also [26,10]]. Finally, we anticipate that this methodology will also be useful in improving reduced-order models when such models include other parameter-dependent model terms that serve to either stabilize or improve the fidelity of the model.

## Acknowledgements

We thank the referees for their extensive comments and suggestions. The presentation of this article has benefited greatly from them. This research was funded by the US Department of Energy, in part under the Laboratory Directed Research and Development program at the Los Alamos National Laboratory (LANL), and in part by the Mix and Burn research initiative under the Physics and Engineering Models (PEM) component of the Advanced Simulation and Computing Program (ASC) at LANL.

## References

[1] M. Farooq Anjum, Imran Tasadduq, Khaled Al-Sultan, Response surface methodology: a neural network approach, Eur. J. Oper. Res. 101 (1) (1997) 65–73.
[2] Ivo Babuska, J. Tinsley Oden, Verification and validation in computational engineering and science: basic concepts, Comput. Methods Appl. Mech. Eng. 193 (36) (2004) 4057–4066.
[3] Barrett Baldwin, Harvard Lomax, Thin-layer approximation and algebraic model for separated turbulent flows, in: 16th Aerospace Sciences Meeting, 1978, p. 257.
[4] Alexandros Beskos, Natesh Pillai, Gareth Roberts, Jesus-Maria Sanz-Serna, Andrew Stuart, et al., Optimal tuning of the hybrid Monte Carlo algorithm, Bernoulli 19 (5A) (2013) 1501–1534.
[5] Didier Besnard, F.H. Harlow, R.M. Rauenzahn, C. Zemach, Turbulence Transport Equations for Variable-Density Turbulence and Their Relationship to Two-Field Models, Technical report, Los Alamos National Lab., NM, United States, 1992.
[6] David M. Blei, Alp Kucukelbir, Jon D. McAuliffe, Variational inference: a review for statisticians, J. Am. Stat. Assoc. 112 (518) (2017) 859–877.
[7] Sai Hung Cheung, Todd A. Oliver, Ernesto E. Prudencio, Serge Prudhomme, Robert D. Moser, Bayesian uncertainty analysis with applications to turbulence modeling, Reliab. Eng. Syst. Saf. 96 (9) (2011) 1137–1149.
[8] Kuei-Yuan Chien, Predictions of channel and boundary-layer flows with a low-Reynolds-number turbulence model, AIAA J. 20 (1) (1982) 33–38.
[9] George Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Syst. 2 (4) (1989) 303–314.

[10] Anthony M. DeGennaro, Nathan M. Urban, Balasubramanya T. Nadiga, Terry Haut, Model structural inference using local dynamic operators, Int. J. Uncertain. Quantificat. 9 (1) (2019) 59–83.
[11] Peter Deuflhard, Wilhelm Huisinga, Alexander Fischer, Ch. Schütte, Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains, Linear Algebra Appl. 315 (1–3) (2000) 39–59.
[12] Jesse Dorrestijn, Daan T. Crommelin, A. Pier Siebesma, Harm J.J. Jonker, Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data, Theor. Comput. Fluid Dyn. 27 (1–2) (2013) 133–148.
[13] Jinqiao Duan, Balasubramanya Nadiga, Stochastic parameterization for large eddy simulation of geophysical flows, Proc. Am. Math. Soc. 135 (4) (2007) 1187–1196.
[14] Paul A. Durbin, Near-wall turbulence closure modeling without "damping functions", Theor. Comput. Fluid Dyn. 3 (1) (1991) 1–13.
[15] W.N. Edeling, Pasquale Cinnella, Richard P. Dwight, Predictive rans simulations via Bayesian model-scenario averaging, J. Comput. Phys. 275 (2014) 65–91.
[16] W.N. Edeling, Pasquale Cinnella, Richard P. Dwight, Hester Bijl, Bayesian estimates of parameter variability in the k–$\varepsilon$ turbulence model, J. Comput. Phys. 258 (2014) 73–94.
[17] Kathryn Farrell, J. Tinsley Oden, Danial Faghihi, A Bayesian framework for adaptive selection, calibration, and validation of coarse-grained models of atomistic systems, J. Comput. Phys. 295 (2015) 189–208.
[18] Andrew Gelman, Gareth O. Roberts Walter, R. Gilks, et al., Efficient metropolis jumping rules, Bayesian Stat. 5 (599–608) (1996) 42.
[19] Susanne Gerber, Illia Horenko, Toward a direct and scalable identification of reduced models for categorical processes, Proc. Natl. Acad. Sci. (2017) 201612619.
[20] Heikki Haario, Eero Saksman, Johanna Tamminen, et al., An adaptive metropolis algorithm, Bernoulli 7 (2) (2001) 223–242.
[21] Heikki Haario, Marko Laine, Antonietta Mira, Eero Saksman, Dram: efficient adaptive mcmc, Stat. Comput. 16 (4) (2006) 339–354.
[22] Ian Jolliffe, Principal component analysis, Springer, 2011.
[23] B.E. Launder, B.I. Sharma, Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc, Lett. Heat Mass Transf. 1 (2) (1974) 131–137.
[24] Daniel Livescu, J.R. Ristorcelli, Buoyancy-driven variable-density turbulence, J. Fluid Mech. 591 (2007) 43–71.
[25] Daniel Livescu, J.R. Ristorcelli, R.A. Gore, S.H. Dean, W.H. Cabot, A.W. Cook, High-Reynolds number Rayleigh–Taylor turbulence, J. Turbul. 10 (13) (2009).
[26] Balasubramanya T. Nadiga, Nathan M. Urban, Improved representation of ocean heat content in energy balance models, Clim. Change (2019) 1–14.
[27] B.T. Nadiga, Orientation of eddy fluxes in geostrophic turbulence, Philos. Trans. R. Soc. A, Math. Phys. Eng. Sci. 366 (1875) (2008) 2489–2508.
[28] B.T. Nadiga, D. Livescu, Instability of the perfect subgrid model in implicit-filtering large eddy simulation of geostrophic turbulence, Phys. Rev. E 75 (4) (2007) 046303.
[29] B.T. Nadiga, N.M. Urban, Dependence of inferred climate sensitivity on the discrepancy model, arXiv:1809.04068, 2018.
[30] Todd A. Oliver, Robert D. Moser, Bayesian uncertainty quantification applied to rans turbulence models, J. Phys. Conf. Ser. 318 (2011) 042032.
[31] Vijay S. Pande, Kyle Beauchamp, Gregory R. Bowman, Everything you wanted to know about Markov state models but were afraid to ask, Methods 52 (1) (2010) 99–105.
[32] R.S. Plant, George C. Craig, A stochastic parameterization for deep convection based on equilibrium statistics, J. Atmos. Sci. 65 (1) (2008) 87–105.
[33] Stephen B. Pope, Turbulent Flows, 2001.
[34] Patrick J. Roache, Verification and Validation in Computational Science and Engineering, Hermosa, 1998.
[35] Peter J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. 656 (2010) 5–28.
[36] Christof Schütte, Frank Noé, Jianfeng Lu, Marco Sarich, Eric Vanden-Eijnden, Markov state models based on milestoning, J. Chem. Phys. 134 (20) (2011) 05B609.
[37] J.D. Schwarzkopf, D. Livescu, J.R. Baltzer, R.A. Gore, J.R. Ristorcelli, A two-length scale turbulence model for single-phase multi-fluid mixing, Flow Turbul. Combust. 96 (1) (2016) 1–43.
[38] John D. Schwarzkopf, Daniel Livescu, Robert A. Gore, Rick M. Rauenzahn, J. Raymond Ristorcelli, Application of a second-moment closure model to mixing processes involving multicomponent miscible fluids, J. Turbul. 12 (49) (2011).
[39] P.Ra.A. Spalart, S. Allmaras, A one-equation turbulence model for aerodynamic flows, in: 30th Aerospace Sciences Meeting and Exhibit, 1992, p. 439.
[40] Luke Tierney, Antonietta Mira, Some adaptive Monte Carlo methods for Bayesian inference, Stat. Med. 18 (1718) (1999) 2507–2515.
[41] Zhu Wang, Imran Akhtar, Jeff Borggaard, Traian Iliescu, Proper orthogonal decomposition closure models for turbulent flows: a numerical comparison, Comput. Methods Appl. Mech. Eng. 237 (2012) 10–26.