SANDIA REPORT

SAND2015-7938
Unlimited Release
Printed September, 2015

Calibration and Forward Uncertainty Propagation for Large-eddy Simulations of Engineering Flows

J.A. Templeton, M. Blaylock, S. Domino, J. Hewson, P.R. Kumar, J. Ling, H. Najm, A. Ruiz,
C. Safta, K. Sargsyan, A. Stewart, G. Wagner

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94-AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2015-7938
Unlimited Release
Printed September, 2015

Calibration and Forward Uncertainty Propagation for Large-eddy Simulations of Engineering Flows

Jeremy A. Templeton, Myra Blaylock, Stefan Domino, John Hewson, P. Raj Kumar, Julia Lin
Habib Najm, Anthony Ruiz, Cosmin Safta, Khachik Sargsyan, Alessia Stewart, Greg J. Wagn
Sandia National Laboratories, P.O. Box 969
Livermore, CA 94550-0969

Abstract

Acknowledgment

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of EnergyÕs National Nuclear Security Administration under contract DE-AC04-94AL85000. This work was funded under LDRD Project Number 151294 and Title “Multiscale Modeling for Fluid Transport in Nanosystems". The authors gratefully acknowledge helpful comments from Drs. Jackie Chen, Helmanth Kolla, Guilhelm Lacaze, and Joe Oefelein.

Contents

List of Figures

List of Tables

This page intentionally left blank.

Chapter 1

Introduction

Turbulent flows control the behavior of many applications relevant to Sandia's mission areas, including combustion and reacting flows, wind energy, atmospheric re-entry, and aerodynamics. Given turbulence's additional ubiquity in flows of interest to science and engineering generally, much work has been focused on methods to predictively simulate it. Three main simulation categories exist. The first is direct numerical simulation (DNS) which resolves all of the energy containing motions down to the scales at which viscosity completely dissipates the turbulent fluctuations. DNS is highly accurate and can be verified using mesh resolution studies, but its accuracy comes at high computational cost. As such, DNS is typically only applied to simple flows in canonical geometries as part of scientific investigations to better characterize turbulence.

At the opposite end of the spectrum from DNS lies the Reynolds Averaged Navier Stokes (RANS) method which only computes the mean velocity field and models the effects of fluctuations. While often amenable to solution verification via mesh refinement, RANS models also introduce significant but uncharacterized errors. Despite the uncertainty in the solutions, RANS simulations are the workhorse of engineering practice due to their low cost. Large-eddy simulation (LES) fills the gap between DNS and RANS by resolving the large-scale motions while modeling the nominally universal small-scales of turbulence. While the cost is higher than RANS, LES can theoretically be applied at any desired scale such that the cost of a simulation can be determined based on the required scales to be resolved.

Despite the theoretical construction of LES, practically speaking it has been much more widely used as a tool for scientific investigation rather than engineering analysis. When highly resolved, LES models need account for little energy, meaning most of the important flow structures are resolved. Indeed, even mesh refinement tends to work for these cases even though increasing amounts of turbulent energy are present. The cost of the simulations, however, scales almost as strongly with the Reynolds number of DNS. As meshes become coarser and more is required of LES models, in many cases the models as presently constructed do not sufficiently represent the missing turbulence to be useful for engineering work. The introduction of unstructured meshes and dissipative numerical methods needed for technological configurations introduce further uncertainty in the results, and established approaches to solution verification founder.

The objective of this work is to investigate the efficacy of using calibration strategies from Uncertainty Quantification (UQ) to determine model coefficients for LES. As the target methods are for engineering LES, uncertainty from numerical aspects of the model must also be quantified.

The ultimate goal of this research thread is to generate a cost versus accuracy curve for LES such that the cost could be minimized given an accuracy prescribed by an engineering need. Realization of this goal would enable LES to serve as a predictive simulation tool within the engineering design process.

This report describes the research efforts taken to determine how and if UQ calibration can lead to engineering LES. Chapter 2 describes how the theoretical foundations of LES can given rise to a means to calibrate LES models from DNS data. It also demonstrates that while the physical constants can be adjusted, numerical errors account for the majority of the uncertainty when using coarse meshes. Turning the problem around, Chapter 3 provides a calibration strategy based on engineering objectives. While the LES models can be calibrated, trade-offs in fidelity of multiple quantities of interest mean that all important flow measures cannot be equally well evaluated due to numerical and model form errors. To better understand numerical errors, Chapter 4 outlines a solution verification study of the numerical operators used in LES to show how variation of the discretization method can lead to very different flow predictions. Concluding the UQ calibration piece of the work, a mesh dependent model is calibrated in Chapter 5, which also investigates model form error.

While understanding all these various aspects of LES is important, to impact engineering work tools are needed with advanced physics models to simulate meaningful systems. Chapter 6 describes one such innovation in which understanding the behavior of a combustion model can impact how it predicts pollutant formation. As software tools are also needed, Chapter 7 overviews several technological improvements made to the code Nalu which make it vehicle for collaboration between Sandia, academia, and industry. While flow solvers are part of the technology for leveraging these new methods, management of calibration and forward UQ studies is error-prone and time-consuming. To mitigate these issues, this work developed a novel UQ management software package called Guru to facilitate organization and execution of coordinated LES solves. The combination of these new methodologies with open source software packages better positions Sandia to help transition LES from a resource for obtaining scientific knowledge to a tool suitable for application to engineering problems.

Chapter 2

Uncertainty Quantification in LES of Channel Flow

Abstract

In this chapter we present a Bayesian framework for estimating joint densities for Large-Eddy Simulation (LES) sub-grid scale model parameters based on canonical forced isotropic turbulence Direct Numerical Simulation (DNS) data. The framework accounts for noise in the independent variables and we present alternative formulations for accounting for discrepancies between model and data. To generate probability densities for flow characteristics, posterior densities for sub-grid scale model parameters are propagated forward through LES of channel flow and compared to DNS data. Synthesis of the calibration and prediction results demonstrates that model parameters have an explicit filter width dependence and are highly correlated, and this work suggests what type of training data can lead to the most successful predictions.

2.1 Introduction

Large-eddy simulation (LES) is a method to compute the solutions to turbulent flows by resolving the larger flow features while modeling the smaller (sub-grid) scale fluctuations [66]. Mathematically, these two scales are separated by use of a low-pass filter, but most LES formulations do not account for the filtering operation explicitly. However, it is not completely understood yet how to perform such filtering in practical LES calculations, particularly because numerical errors can pollute the solution [33, 42].

Several efforts have addressed this topic. Gullbrand and Chow [22] considered a dynamic reconstruction model to estimate the sub-grid stress (SGS) by approximately inverting an explicit filtering operator to compute part of it from the resolved turbulent flow field. As an alternative, Bose and Moin [5] considered directly modifying the SGS model filter width in order to obtain "mesh-independent" LES results. Moser and co-workers [45, 78] took a different strategy in which LES SGS model form and parameters were determined by applying linear stochastic estimation to channel flow in a technique known as optimal LES. Other approaches which explicitly bring information regarding the filter into the LES formulation involve boundary conditions. Previous studies have considered the effect of explicit filtering on boundary conditions for the Burgers' equation [76, 4] and the Navier-Stokes equations [71, 3], with the work of Bhattacharya *et al* [3] being in the context of optimal LES.

There have been several prior efforts which used Bayesian [7] and other machine learning techniques to estimate uncertainties in turbulence simulations with particular emphasis on turbulence models. To the best of our knowledge, the first attempt at deriving LES closures from resolved flow fields using machine learning strategies was the optimal LES formulation [78] which used linear stochastic estimation [1]. Decompositions of the Reynolds-Averaged Navier-Stokes (RANS) turbulent stress tensor have been considered as uncertain parameters and their effect on resulting RANS solutions has been quantified [49, 50, 16]. For example, the principal stress axes were computed and perturbed to estimate the effect of structural uncertainty of turbulence models on flow realizations. The effects of uncertain boundary conditions have been examined using the Burgers' equation [54, 55] and the Navier-Stokes equations [9]. Bayesian estimation of the effects of the combustion models in LES has also been carried out [59], while in cold flow isotropic turbulence, Lucor *et al* [38] assumed a distribution for the constant in the Smagorinsky model following a Polynomial Chaos Expansion (PCE) which was sampled to compute the PCE of different scales of turbulent motion. Optimization-based calibration methods have also been applied to RANS models [36]. Other recent efforts have focused on using adjoint methods to assess model sensitivities [15, 79]. A review article by Najm [47] provides a discussion on uncertainty quantification techniques, with emphasis on PCEs, in computational fluid dynamics.

The purpose of this paper is twofold: 1) we propose a framework for the estimation of uncertainties in the parameters for sub-grid scale turbulence models followed by a forward Uncertainty Quantification (UQ) study to examine the predictive capabilities of the LES approach given the calibrated set of parameters, and 2) we assess the feasibility of using directly filtered turbulence information to contribute to existing LES SGS models. In this context, we plan to investigate how calibrated LES models fare in simulations with different physical parameters, numerical methods, and meshes.

Figure 2.1 shows a schematic of this framework, consisting of two connected workflows for *Parameter Estimation* and *Forward UQ*. The joint probability density for input parameters is estimated in a Bayesian framework. The Bayesian context provides a flexible framework to handle heterogeneous information and allows for sequential updates of posterior distributions as the prior information is revised. Both the model error ε_m and the measurement error ε_d can be embedded in the parameter estimation. The computational model, f, is based on Sandia's FUEGO LES software suite [44]. In the context of this work, we explore the predictive capabilities of f in a channel flow configuration, when employing the calibrated parameters for the k^{sgs} sub-grid scale model. We note that Bayesian estimation has been successfully used to infer model parameters within multiscale settings in other applications, such as molecular dynamics [62, 67], porous media flows [8], and Carbon cycle models [18, 61].

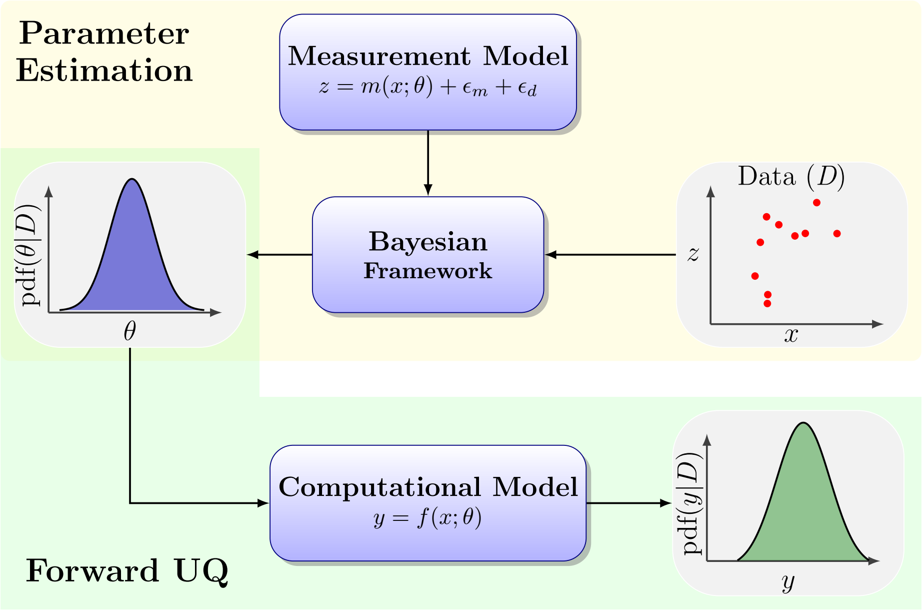


Figure 2.1: Schematic of parameter estimation and forward UQ workflows.

This paper is organized as follows. Section 2.2 provides a description of the LES formulation. Section 2.3 presents the Bayesian framework for model calibration. Section 2.4 describes propagation of uncertainties from calibrated model parameters to relevant quantities of interest in LES of channel flow. We end with conclusions in Section 2.5. The methods employed in this paper are part of Sandia's software package UQTk v3.0.[1]

---

[1] http://www.sandia.gov/UQToolkit

2.2 Large Eddy Simulation

The LES field is formally defined as a low-pass filtered turbulent flow field, i.e.

$$\tilde{u}_i(\mathbf{x}) \equiv \int_{\Omega} u_i(\mathbf{y}) G(\mathbf{x}-\mathbf{y};\Delta)\, dv, \tag{2.1}$$

where $\mathbf{x}$ and $\mathbf{y}$ are coordinate vectors in $\mathbb{R}^3$ with $dv = dy_1\, dy_2\, dy_3$, u_i is the i-th velocity component, $\tilde{u}_i$ is the filtered or LES ith velocity component, G is a convolution kernel which preferentially removes small-scale energy from u_i. Here, "small-scale" is defined relative to the parametric dependence of G on a length scale Δ, denoting the filter width. If $\int_{\Omega} G\, dv = 1$ and it satisfies assumptions of isotropy, homogeneity, and continuous differentiability, the order of convolution and differentiation can be exchanged such that the equations governing the LES field as determined from the incompressible Navier-Stokes equations are

$$\frac{\partial \rho \tilde{u}_i}{\partial t} + \frac{\partial \rho \tilde{u}_i \tilde{u}_j}{\partial x_j} = -\frac{\partial \tilde{P}}{\partial x} + \nu \frac{\partial^2 \tilde{u}_i}{\partial x_j \partial x_j} + \frac{\partial \tau_{ij}}{\partial x_j}, \tag{2.2}$$

$$\frac{\partial \tilde{u}_j}{\partial x_j} = 0. \tag{2.3}$$

where ν is the kinematic viscosity and $\tilde{P}$ is the filtered pressure. While similar to the Navier-Stokes equations, the LES equations differ through the inclusion of the SGS tensor

$$\tau_{ij} = \rho \tilde{u}_i \tilde{u}_j - \rho \widetilde{u_i u_j}. \tag{2.4}$$

Determining closure models for this term, particularly in complex engineering flows, remains an active area of research. We also note that when the assumptions on the filter are violated, as in wall-bounded flows, commutation error is incurred because differentiation and filtering no longer exactly commute.

Of relevance to our current effort, the SGS model not only expicitly depends on the convolution kernel G, such that given the kernel and a turbulent flow field, it can be exactly determined. This mathematical relationship between the "true" flow field and the LES SGS tensor enables our goals of calibration, uncertainty propagation, and error assessment. Calibration can be achieved by filtering a direct numerical simulation (DNS) to exactly construct the terms present in the SGS model, from which model constants can be inferred. Given these constants, with quantified uncertainty, the error associated with the ideal LES model, based on the mathematics of continuous fields and filtering only, can be propagated forward into new solutions. The remaining error can be attributed to sources of uncertainty not present in the ideal LES.

2.2.1 k^{sgs} Turbulence Model

The transport equation for the LES model for the subgrid scale turbulent kinetic energy, k^{sgs}, is given by Yoshizawa [87, 88]

$$\int_V \frac{\partial \bar{\rho} k^{sgs}}{\partial t} dv + \int_{\partial V} \bar{\rho} k^{sgs} \tilde{u}_j n_j ds = \int_{\partial V} \frac{\mu_k}{\sigma_k} \frac{\partial k^{sgs}}{\partial x_j} n_j ds + \int_V \left(P_k^{sgs} - D_k^{sgs}\right) dv \tag{2.5}$$

The production and dissipation of turbulent kinetic energy, P_k^{sgs} and D_k^{sgs}, are given by

$$P_k^{sgs} = \left(2\mu_t \left(\tilde{S}_{ij} - \frac{1}{3}\tilde{S}_{kk}\delta_{ij}\right) - \frac{2}{3}\bar{\rho}k^{sgs}\delta_{ij}\right)\frac{\partial \tilde{u}_i}{\partial x_j} \tag{2.6}$$

and

$$D_k^{sgs} = C_\varepsilon \frac{(k^{sgs})^{3/2}}{\Delta} \tag{2.7}$$

After expanding the parenthesis in Eq. (2.6), the production term can be written as

$$P_k^{sgs} = 2\mu_t \left(\tilde{S}_{ij}\frac{\partial \tilde{u}_i}{\partial x_j} - \frac{1}{3}\left(\nabla \cdot \tilde{u}\right)^2\right) - \frac{2}{3}\bar{\rho}k^{sgs}\left(\nabla \cdot \tilde{u}\right) \tag{2.8}$$

For incompressible flows, with $\nabla \cdot \tilde{\mathbf{u}} = 0$, the production term reduces to

$$P_k^{sgs} = 2\mu_t \tilde{S}_{ij}\frac{\partial \tilde{u}_i}{\partial x_j} \tag{2.9}$$

The filtered strain rate tensor is written as:

$$\tilde{S}_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right) \tag{2.10}$$

and the product $\tilde{S}_{ij}\frac{\partial \tilde{u}_i}{\partial x_j}$ can be further expanded as follows

$$\frac{1}{2}\sum_i\sum_j\left(\frac{\partial \tilde{u}_i}{\partial x_j}\right)^2 + \frac{1}{2}\sum_i\sum_j\frac{\partial \tilde{u}_j}{\partial x_i}\frac{\partial \tilde{u}_i}{\partial x_j} \tag{2.11}$$

Further, the subgrid turbulent viscosity is modeled as

$$\mu_t = C_{\mu_\varepsilon}\Delta\sqrt{k^{sgs}} \tag{2.12}$$

2.3 Model Calibration

We employ a probabilistic approach, specifically a Bayesian framework, to compute posterior probabilities for the k^{sgs} model parameters, C_{μ_ε} and C_ε, introduced in Eqs. (3.2) and (3.5). In the Bayesian approach the posterior probability density for the model parameters is given by

$$p(\theta|\mathscr{D}) = L_{\mathscr{D}}(\theta)p(\theta)/p(\mathscr{D}) \tag{2.13}$$

Here, $p(\theta)$ and $p(\theta|\mathscr{D})$ are the prior and posterior probability densities, respectively, for model parameters θ. These densities represent our knowledge about these parameters before and after learning from the data $\mathscr{D}$. The likelihood function $L_{\mathscr{D}}(\theta) = p(\mathscr{D}|\theta)$ is the likelihood of the data $\mathscr{D}$ for a particular instance of θ. The denominator in Eq. (3.7), $p(\mathscr{D})$, is the "model evidence", efectively a constant normalizing factor, and not relevant for the calculations presented in this paper as it is independent of θ.

2.3.1 Calibration Data

For this study we employ the Johns Hopkins University (JHU) turbulence database [52, 35] of forced isotropic turbulence in a periodic cube. Specifically, this dataset results from a DNS using 1024^3 grid nodes overlaying a cube with sides of length $L = 2\pi$. The Taylor-scale Reynolds number, Re_λ, for this simulation is approximately 433. The simulation covers a non-dimensional time range of 2.048 with solution files saved every 10 time steps, or 2×10^{-3}. For the current study we considered the DNS data at times $t = \{0, 0.1, 0.2, \ldots, 2.0\}$.

Filtered velocity values are computed from the DNS data using a top-hat filter of width Δ in all coordinate directions. The sub-grid scale kinetic energy is then computed as $k^{sgs} = \tilde{K} - K_{res}$, where $\tilde{K} = \frac{1}{2}\widetilde{u_i u_i}$ is the kinetic energy per unit mass and $K_{res} = \frac{1}{2}\tilde{u}_i\tilde{u}_i$ is its "resolved" component, computed using the filtered values.

The boundary integrals in Eq. (3.1) are statistically zero in a periodic configuration,

$$\int_{\partial V} \bar{\rho} k^{sgs} \tilde{u}_j n_j ds = \int_{\partial V} \frac{\mu_k}{\sigma_k} \frac{\partial k^{sgs}}{\partial x_j} n_j ds \equiv 0. \tag{2.14}$$

The remaining terms in Eq. (3.1) can be written in compact form as

$$f_k(t;\Delta) = C_{\mu_\varepsilon} f_P(t;\Delta) - C_\varepsilon f_D(t;\Delta) \tag{2.15}$$

where

$$\begin{aligned} f_k(t;\Delta) &= \int_V \frac{\partial \bar{\rho} k^{sgs}}{\partial t} dv, \; f_D(t;\Delta) = \int_V \frac{k^{sgs3/2}}{\Delta} dv, \\ f_P(t;\Delta) &= \int_V \Delta \sqrt{k^{sgs}} \tilde{S}_{ij} \frac{\partial \tilde{u}_i}{\partial x_j} dv \end{aligned} \tag{2.16}$$

Spatial derivatives in the integrand for term f_P, in Eq. (2.16), were computed at each time instant using second-order centered finite differences. For f_k, time derivatives were computed using centered finite differences using the data from solution files immediately adjacent to the corresponding time stamp. In addition to the dependence on time, the terms in Eq. (2.15) also depend on the filter width Δ.

Figure 2.2 shows slices through the DNS field in the left column. These slices correspond to planes perpendicular to the coordinate directions and centered at $x = y = z = \pi/2$. The top, middle, and bottom rows show the u, v, and w velocity components, respectively. The right column in this figure shows the corresponding filtered velocity components.

Figure 2.3 shows k^{sgs} in an $x-$const slice (left column) and $y-$const slice (right column), centered at the same point as in Fig. 2.2. The dependence of k^{sgs} on the filter width is evident from this figure, as larger Δ values lead to more diffuse features for the filtered quantities.

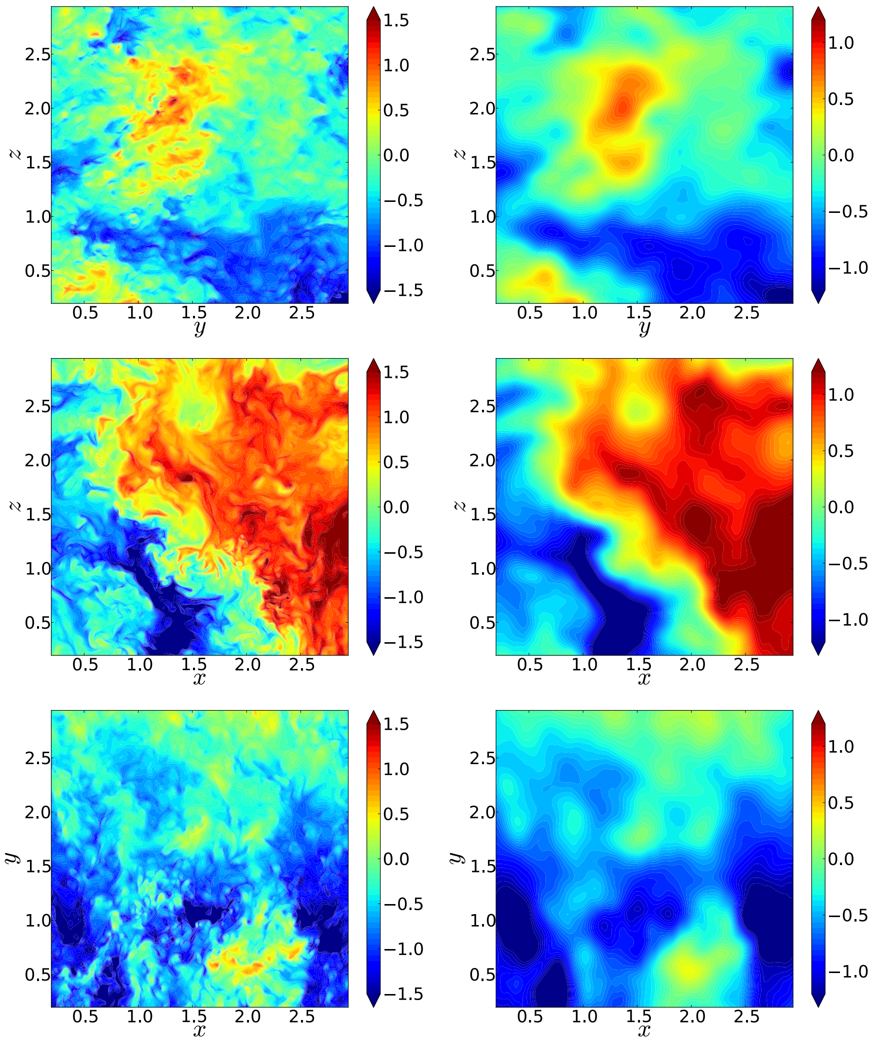


Figure 2.2: Snapshot of u, v, and w velocity components: DNS data (left column) and filtered data (right column) at $t = 0.4$ for a filter size $\Delta = L/32$.

Figure 2.4 shows the values of f_k, f_P, and f_D corresponding to the solution data considered in this study.The time derivative of the overall sub-grid scale turbulent kinetic energy, f_k varies around zero, and shows little correlation (image not shown) with the production and dissipation terms. In this configuration, the production and dissipation terms show a high degree of correlation and their magnitude depends on the filter width. This is to be expected in the forced isotropic turbulence configuration where energy is continuously injected to keep the total kinetic energy constant.

2.3.2 Likelihood Construction

In the Bayesian framework, $L_{\mathscr{D}}(\theta)$ is the likelihood of observing the data $\mathscr{D}$ for an instance of the model parameters θ at specific locations for the independent variables x. In this context $\theta = \{C_{\mu_\varepsilon}, C_\varepsilon\}$, $\mathscr{D}$ consists of the values $z = \{f_{k,i} | i = 1, \ldots, N_t\}$ while $x = \{f_P, f_D\}$. In general, the

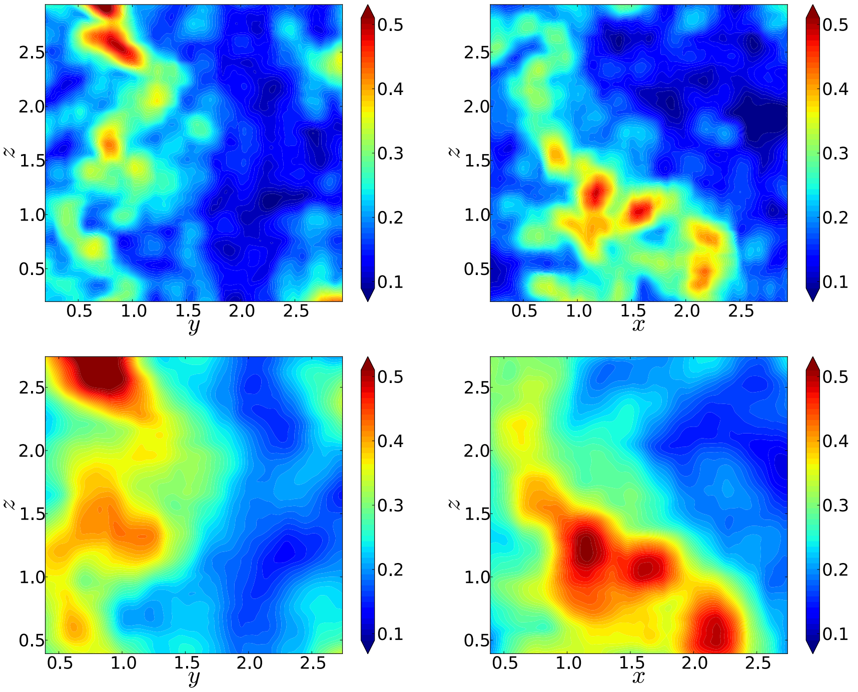


Figure 2.3: Snapshots of $\sqrt{k^{sgs}}$ field at $t = 0.4$ for filter sizes $\Delta = L/32$ (top row) and $\Delta = L/16$ (bottom row).

discrepancy between model predictions and the data can be formalized as [68]

$$z = m(x;\theta) + \varepsilon_z \tag{2.17}$$

Here, $m(x;\theta) = C_{\mu_\varepsilon} f_P - C_\varepsilon f_D$ and ε_z is the discrepancy between the data and the model, i.e. a consequence of the model only being an approximation of the true process and any imperfections in the measurement process.

In the standard approach, x is a vector of observable quantities that are known and fixed throughout the analysis. In situations, such as this work, where the independent variables x also suffer due to data collection error, one needs to consider this error when inferring the model parameters θ. The new compound discrepancy model can be formulated as a classical error-in-variables model [12]

$$\begin{aligned} z &= m(x_a;\theta) + \varepsilon_z \\ x_o &= x_a + \varepsilon_x \end{aligned} \tag{2.18}$$

Here, x_a and x_o are the actual and observed, respectively, values for the independent variable x. For this problem, we have a set of N_t observations for x_o, $\{f_{P,i}, f_{D,i} | i = 1, \ldots, N_t\}$.

We formulate a Bayesian approach to infer the values for model parameters $\theta = \{C_{\mu_\varepsilon}, C_\varepsilon\}$, the actual values for the latent variables $x_a = \{f_P^a, f_D^a\}$, as well as for other parameters for the error

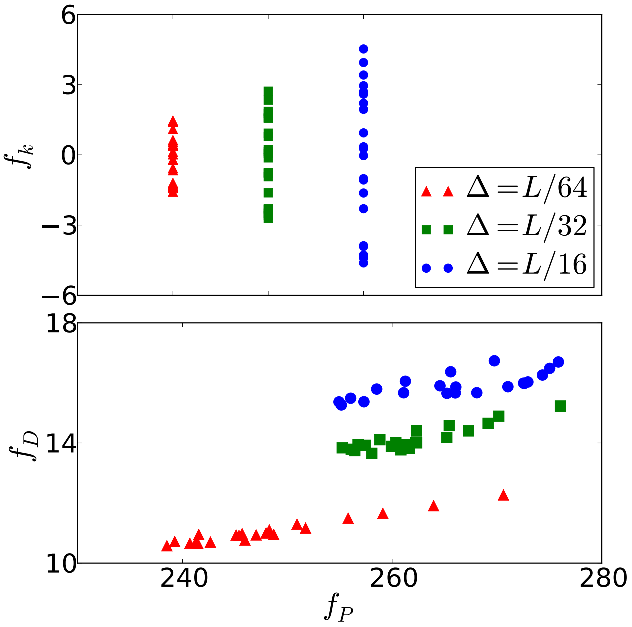


Figure 2.4: Top frame: f_k; Bottom frame: f_P and f_D (bottom frame) for several filter sizes Δ.

terms ε_z and ε_x. In the context of this paper, x_a are regarded as unknown constants. The Bayes formula can be written as

$$\begin{aligned} p(\theta,\theta_{\varepsilon,z},\theta_{\varepsilon,x},x_a|z,x_o) \quad \propto \quad & p(z|\theta,\theta_{\varepsilon,z},\theta_{\varepsilon,x},x_a,x_o)\times \\ & p(\theta,\theta_{\varepsilon,z},\theta_{\varepsilon,x},x_a|x_o) \end{aligned} \tag{2.19}$$

Here $\theta_{\varepsilon,z}$ is the vector of parameters defining ε_z and $\theta_{\varepsilon,x}$ is the vector of parameters defining ε_x, and $x_o = \{(f_{P,i}, f_{D,i})|i = 1,\ldots,N_t\}$. Given the dependencies in Eq. (2.18), the likeliihood can be written as

$$p(z|\theta,\theta_{\varepsilon,z},\theta_{\varepsilon,x},x_a,x_o) = p(z|\theta,\theta_{\varepsilon,z},x_a) \tag{2.20}$$

We further assume that $\{\theta,\theta_{\varepsilon,z}\}$ and $\{\theta_{\varepsilon,x},x_a\}$ are a priori independent and write the prior as

$$p(\theta,\theta_{\varepsilon,z},\theta_{\varepsilon,x},x_a|x_o) = p(\theta,\theta_{\varepsilon,z})p(\theta_{\varepsilon,x},x_a|x_o) \tag{2.21}$$

The second term in the right-hand side of the above expression is further expanded using Bayes theorem

$$p(\theta_{\varepsilon,x},x_a|x_o) \propto p(x_o|\theta_{\varepsilon,x},x_a)p(\theta_{\varepsilon,x},x_a) \tag{2.22}$$

The expressions in Eqs. (2.20-2.22) are then plugged back into Eq. (2.19), leading to

$$\begin{aligned} p(\theta,\theta_{\varepsilon,z},\theta_{\varepsilon,x},x_a|z,x_o) \quad \propto \quad & \underbrace{p(z|\theta,\theta_{\varepsilon,z},x_a)p(x_o|\theta_{\varepsilon,x},x_a)}_{\text{Likelihood}}\times \\ & \underbrace{p(\theta,\theta_{\varepsilon,z})p(\theta_{\varepsilon,x},x_a)}_{\text{Prior}} \end{aligned} \tag{2.23}$$

The assignment of a statistical model structure for ε_z and ε_x is not straightforward and entails a significant degree of modeling. In this study we presume a bi-variate Gaussian discrepancy with zero mean between the observed and the actual indepedent variables x,

$$\varepsilon_x \sim N(0,\Sigma_x),\ \Sigma_x = \begin{pmatrix} \sigma_P^2 & \rho_{PD}\sigma_P\sigma_D \\ \rho_{PD}\sigma_P\sigma_D & \sigma_D^2 \end{pmatrix} \tag{2.24}$$

Given this assumption, $\theta_{\varepsilon,x} = \{\rho_{PD}, \sigma_P, \sigma_D\}$.

We employ two modeling approaches to estimate ε_z. In the first approach we consider a normally distributed error model ε_z,

$$\varepsilon_x \sim N(0,\sigma_z),\ \theta_{\varepsilon,z} = \{\sigma_z\} \tag{2.25}$$

Under this assumption, $p(z|\theta,\theta_{\varepsilon,z},x_a)$ is approximated as a product of marginal densities using $f_{k,i}$ collected at several time instants

$$p(z|\theta,\theta_{\varepsilon,z},x_a) = \prod_{i=1}^{N_t} \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(f_{k,i} - m(x_a;\theta))^2}{2\sigma_z^2}\right) \tag{2.26}$$

while $p(x_o|\theta_{\varepsilon,x},x_a)$ is given by

$$p(z|\theta,\theta_{\varepsilon,z},x_a) = \frac{1}{(2\pi|\Sigma_x|)^{N_t}} \prod_{i=1}^{N_t} \exp\left(-(x_a - x_{oi})^T \Sigma_x^{-1} (x_a - x_{oi})\right) \tag{2.27}$$

Here, the dependence on the filter width was dropped to simplify the notation. One setback with this approach is that the explicit model error term is now part of the calibrated model, i.e. the evolution of k^{sgs} is now governed by a right-hand side that includes a standard normal random variable in addition to the two terms shown in Eq. (2.15).

The second approach follows Sargsyan *et al* [68] who suggest embedding the model error terms in the model components. By embedding the error model within the components themselves, the model predictions automatically contain all proposed errors without the need to specify external error terms. We begin by recasting the model constants as random variables parameterized by $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_2}$, respectively,

$$C_{\mu_\varepsilon} = C_{\mu_\varepsilon}(\boldsymbol{\alpha_1}),\ C_\varepsilon = C_\varepsilon(\boldsymbol{\alpha_2}) \tag{2.28}$$

Here, $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_2}$ are parameter vectors that now need to be estimated. This effectively re-casts the calibration problem into one of estimating densities of $\theta = (\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2})$ such that the forward-propagated PDFs of C_{μ_ε} and C_ε are consistent with select statistics of f_k. In this framework, the data model is $f_k = C_{\mu_\varepsilon}(\boldsymbol{\alpha_1}) f_P - C_\varepsilon(\boldsymbol{\alpha_2}) f_D$ where ε_z has been subsumed into the random model ansatz for C_{μ_ε} and C_ε.

For this approach we presume that C_{μ_ε} and C_ε are random variables with finite variance and we will adopt a Polynomial Chaos expansion (PCE) [19, 10] to represent these (now random) model parameters. A brief description of PCE is presented below. For an in-depth description, the reader is referred to a series of publications on this topic [80, 19, 30, 83].

Given a probability space $(\Omega, \mathscr{S}, P)$, where Ω is a sample space, $\mathfrak{S}$ is a σ-algebra on Ω, and P is a probability measure on $(\Omega, \mathfrak{S})$, we define a *germ* $\xi = \{\xi_1, \ldots, \xi_n\} : \Omega \to \Xi^n \subset \mathbb{R}^n$ where the $\{\xi_i\}_{i=1}^n$ are a set of independent random variables, and Ξ^n is the support of the density, that is the set where the probability density of the germ is non zero. The ξ_i are required to satisfy the constraint that their density is uniquely determined by their moments [17], a requirement that is satisfied by a range of standard random variables, *e.g.*, standard normals, and uniforms. It can the be shown that any random variable $X \in L^2(\Omega, \mathfrak{S}(\xi), P)$, with $\mathfrak{S}(\xi)$ being the sigma algebra generated by the chosen germ, can be written as the Polynomial Chaos expansion (PCE)

$$X = X(\xi) = \sum_{k=0}^{\infty} X_k \Psi_k(\xi). \tag{2.29}$$

Here Ψ_k are multivariate polynomials, that are products of univariate polynomials, $\Psi_k(\xi) = \Psi_{k_1}(\xi_1)\Psi_{k_2}(\xi_2)\ldots$. In a practical computational context, for a given dimensionality n, one truncates the PCE according to a desired polynomial order p. The number of terms in the resulting finite PCE

$$X(\xi) \approx \sum_{k=0}^{P-1} \alpha_k \Psi_k(\xi) \tag{2.30}$$

is given by $P = (n+p)!/n!p!$.

The univariate expansion polynomials correspond to a given choice of distribution for the ξ_i and are orthogonal with respect to the density of the random variable. In general, popular choices for (ξ, Ψ) pairs are uniform RVs with Legendre polynomials or normal RVs with Hermite polynomials.

We employ Hermite-Gauss PCEs to represent C_{μ_ε} and C_ε as

$$C_{\mu_\varepsilon} = \sum_k \alpha_{1,k} \Psi_k(\xi),\; C_\varepsilon = \sum_k \alpha_{2,k} \Psi_k(\xi). \tag{2.31}$$

Here $\xi = \{\xi_1, \ldots, \xi_n\}$ is a vector of standard normal random variables, while $\Psi_k(\xi) = \Psi_{k_1}(\xi_1) \cdots \Psi_{k_n}(\xi_n)$ are multivariate Hermite polynomials, orthogonal with respect to the density of the standard normal random variables. To illustrate the functional form representation via PCEs we employ first order expansions for both parameters

$$\begin{aligned} C_{\mu_\varepsilon} &= \alpha_{10} + \alpha_{11}\xi_1 \\ C_\varepsilon &= \alpha_{20} + \alpha_{21}\xi_1 + \alpha_{22}\xi_2 \end{aligned} \tag{2.32}$$

The first-order approximation essentially leads to multivariate normal distribution for C_{μ_ε} and C_ε for each instance of $\boldsymbol{\alpha_1}$ and $\boldsymbol{\alpha_2}$. A “triangular” form is adopted above in order to avoid rotational symmetry. Additionally, both α_{21} and α_{22} are constrained to be positive to avoid positive/negative symmetries. The likelihood, expressed as

$$L_{\mathscr{D}}(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}) = p(z|\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, x_a), \tag{2.33}$$

is the multivariate density for $z=\{f_{k,i}|i=1,\ldots,N_t\}$. Generally, this multivariate density has been shown to be degenerate [68]. Instead, we approximate this density with a product of marginal densities corresponding to each data point

$$L_{\mathscr{D}}(\alpha_{\mathbf{1}},\alpha_{\mathbf{2}})=\prod_{i=1}^{N_t}p(f_{k,i}|\alpha_{\mathbf{1}},\alpha_{\mathbf{2}},x_a), \tag{2.34}$$

Given that germs ξ_1 and ξ_2 in Eq. 2.32 are normal RV's and that the model, in Eq. 2.15, is linear, the marginal densities $p(f_{k,i}|\alpha_{\mathbf{1}},\alpha_{\mathbf{2}},x_a)$ are normal, with mean and variance given by

$$\begin{aligned}\mu_f &= \alpha_{10}f_P^a-\alpha_{11}f_D^a\\ \sigma_f^2 &= \left(\alpha_{20}f_P^a-\alpha_{21}f_D^a\right)^2+\left(\alpha_{22}f_D^a\right)^2\end{aligned} \tag{2.35}$$

In this paper we will employ acronyms AEM (additive error model) for results based on the likelihood shown in Eq. (2.26) and EEM (embedded error model) for likelihood in Eq. (2.34) based on the approach described in the latter part of this section.

2.3.3 Prior Construction

A result of the present work is that the data presented in Section 2.3.1, in particular Fig. 2.4, does not contain sufficient information to constrain both model constants C_{μ_ε} and C_ε. The production and dissipation terms are nearly proportional to each other, while f_k, showing the overall change in k^{sgs}, is small and fluctuates around 0. While it is to be expected that the balance of sub-grid kinetic energy should preferentially inform the ratio rather than the values of the two constants, it was not known *a priori* if the fluctuations in k^{sgs} would be sufficient to estimate the parameter values. It could be the case that the removal of the boundary flux terms is partially to blame, but it appears as though the time-scale of the sub-grid kinetic energy variation relative to the time-scale of the fluctuations in production and dissipation is also too small as in Fig. 2.4. With this observation, and in the absence of additional data, it is clear that one needs additional regularization to constrain the parameter estimation problem.

To this end, for the AEM approach, we choose independent Gaussian priors for the components of $\theta=\{C_{\mu_\varepsilon},C_\varepsilon\}$, $\mathscr{N}(\mu^{pr}_{C_{\mu_\varepsilon}},\sigma^{pr^2}_{C_{\mu_\varepsilon}})$ and $\mathscr{N}(\mu^{pr}_{C_\varepsilon},\sigma^{pr^2}_{C_\varepsilon})$, centered at parameter values previously recommended in the turbulence modeling literature. The first set of values considered here is the one recommended by Schumann [69] and denoted by subscript 1 below. The second set of values is recommended by Yoshizawa [87, 88] and denoted by subscript 2 below

$$\begin{aligned}\mu_1^{pr}=\{\mu^{pr}_{C_{\mu_\varepsilon}1},\mu^{pr}_{C_\varepsilon 1}\} &= \{0.0845,0.85\}\\ \mu_2^{pr}=\{\mu^{pr}_{C_{\mu_\varepsilon}2},\mu^{pr}_{C_\varepsilon 2}\} &= \{0.07,1.05\},\end{aligned} \tag{2.36}$$

For the AEM, we also explore three sets of values values for the marginal standard deviations for the prior densities, as follows

$$\begin{aligned} \sigma_1^{pr} = \{\sigma^{pr}_{C_{\mu\varepsilon}\,1}, \sigma^{pr}_{C_\varepsilon\,1}\} &= \{0.04, 0.4\} \\ \sigma_2^{pr} = \{\sigma^{pr}_{C_{\mu\varepsilon}\,2}, \sigma^{pr}_{C_\varepsilon\,2}\} &= \{0.02, 0.2\} \\ \sigma_3^{pr} = \{\sigma^{pr}_{C_{\mu\varepsilon}\,3}, \sigma^{pr}_{C_\varepsilon\,3}\} &= \{0.01, 0.1\} \end{aligned} \tag{2.37}$$

The sets of mean values provided in Eq. (2.36) are combined with the sets of marginal standard deviations, in Eq. (2.37), to construct priors for several simulations, as outlined in Table 3.1.

For EEM we construct priors for $\theta = \{\alpha_1, \alpha_2\}$ that resemble the above formulation. We first set priors for the mean values of the two constants. Given the formulation proposed in Eq. (2.32), the mean values for C_{μ_ε} and C_ε are given by

$$\mu_{C_{\mu_\varepsilon}} = \alpha_{10},\ \mu_{C_\varepsilon} = \alpha_{20} \tag{2.38}$$

We adopt independent Gaussian priors for these means, with the same values as the ones mentioned above for the classical approach, in Eqs. 2.36 and 2.37. Further we truncate and re-scale these densities so that

$$p^{pr}(\alpha_{10} \leq 0) = p^{pr}(\alpha_{20} \leq 0) = 0.$$

This ensures that the mean values for either C_{μ_ε} and C_ε remain positive. Additionally, we impose conditions on the other α coefficients to ensure only a small probability that C_{μ_ε} and C_ε are less than zero. We adopt a threshold value of 0.1% and impose

$$P(C_{\mu_\varepsilon} < 0) < 10^{-3},\ P(C_\varepsilon < 0) < 10^{-3} \tag{2.39}$$

For C_{μ_ε} this condition leads to

$$\alpha_{10} - 3|\alpha_{11}| > 0 \to p^{pr}(\alpha_{11}|\alpha_{10}) = \frac{\mathscr{H}(\alpha_{10} - 3|\alpha_{11}|)}{2\alpha_{10}/3}$$

where $\mathscr{H}$ is the Heaviside function. For C_ε, Eq. (2.39) leads to

$$p^{pr}(\alpha_{21}, \alpha_{22}|\alpha_{20}) = \frac{\mathscr{H}(\alpha_{20} - 3\sqrt{\alpha_{21}^2 + \alpha_{22}^2})}{4\pi\alpha_{20}^2/9}$$

The prior for EEM parameters is then written as

$$\begin{aligned} p^{pr}(\theta) &= p^{pr}(\alpha_{10}) \times p^{pr}(\alpha_{11}|\alpha_{10}) \times \\ &\quad\ p^{pr}(\alpha_{20}) \times p^{pr}(\alpha_{21}, \alpha_{22}|\alpha_{20}) \end{aligned} \tag{2.40}$$

The denominators in the conditional prior densities $p^{pr}(\alpha_{11}|\alpha_{10})$ and $p^{pr}(\alpha_{21}, \alpha_{22}|\alpha_{20})$ can push both α_{10} and α_{20} towards the degenerate case $\alpha_{10} = \alpha_{20} = 0$. However, this is avoided by the truncated normal densities imposed above for these two model parameters.

2.3.4 Posterior Densities of k^{sgs} Parameters

A Markov Chain Monte Carlo (MCMC) algorithm is used to sample from the posterior probability density, $p(\theta|\mathscr{D})$ in Eq. (3.7). MCMC is a class of techniques that allows sampling from a probability density by constructing a Markov Chain that has the target density as its stationary distribution [21]. In particular, we employ an adaptive Metropolis algorithm [23], which uses the covariance of the previously visited chain states to find better proposal distributions, allowing the exploration of posterior distributions in an efficient manner. We employ the Raftery-Lewis diagnostic [58] to determine when the MCMC samples converge to stationary posterior distributions. We also test the MCMC samples via the Effective Sample Size [32] (ESS) to determine if there is significant autocorrelation between chain samples.

Table 2.1: Setup for model calibration studies. μ^{pr}_{1-2} are defined in Eq. (2.36) while σ^{pr}_{1-3} are defined in Eq. (2.37).

Run	Prior mean	Prior St. Dev.
R1	μ_1^{pr}	σ_1^{pr}
R2	μ_2^{pr}	σ_1^{pr}
R3	μ_1^{pr}	σ_2^{pr}
R4	μ_2^{pr}	σ_2^{pr}
R5	μ_1^{pr}	σ_3^{pr}
R6	μ_2^{pr}	σ_3^{pr}

We start with results from the AEM approach presented in Section 2.3.2. Table 3.1 presents the prior setup for the set of trials employed in this study. For each run, the filtered DNS data employed in the calibration was generated using several choices for the filter width $\Delta = \{L/64, L/32, L/16\}$. Henceforth, results corresponding to a certain prior setup will employ a subscript corresponding to the filter size used for the data, e.g. R1_3 employs the setup for R1 with filter size $\Delta = L/16$, while R3_2 employs the setup for R3 with filter size $\Delta = L/32$. In order to maintain positivity for the standard deviation σ_z we will actually work with $\log \sigma_z$, and use an uniform prior for this parameter. This is equivalent with using a Jeffrey's prior for $1/\sigma_z$.

Figure 2.5 shows marginal posterior PDFs for R3_2. The 1D marginal densities are shown in the diagonal plots, while the contour plots show joint densities between C_{μ_ε}, C_ε, and σ, respectively. The narrow 2D joint density for the pair $(C_{\mu_\varepsilon}, C_\varepsilon)$ indicates a strong dependence between these two parameters. On the other hand, the joint posterior densities between these parameters and the standard deviation used in the likelihood construction, σ_z, suggest a weak dependence between the error term standard deviation and the model parameters. The marginal PDF results based on other filter sizes and/or other prior specifications (i.e. mean value, standard deviation) are qualitatively similar to results shown in Fig. 2.5.

Figure 2.6 shows marginal posterior densities corresponding to parameters α_{ij} that define the models for C_{μ_ε} and C_ε in the EEM aproach. These densities were computed via Kernel Density Estimate (KDE) [70, 72] using 5×10^6 MCMC samples. Parameters α_{10} and α_{20} that correspond to the means of C_{μ_ε} and C_ε exhibit a joint distribution that is very similar to the joint distribution

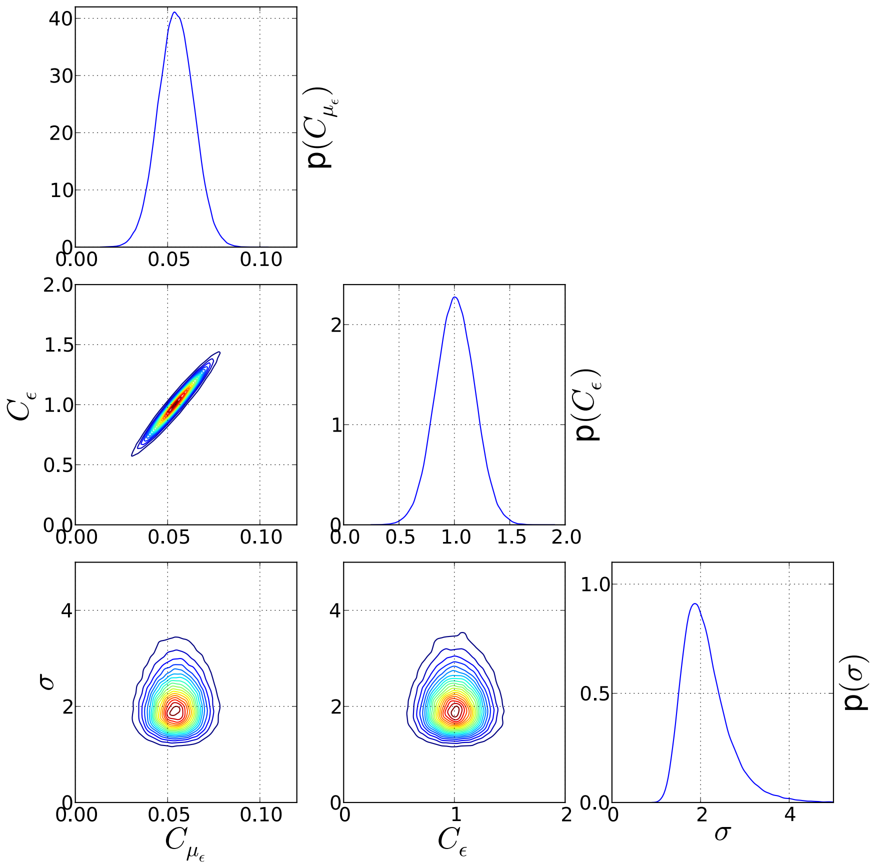


Figure 2.5: Joint posterior distribution for model parameters corresponding to $R3_2$ using AEM. For this case the filtered DNS data employed a filter size $\Delta = L/32$ and the setup for prior densities given in Table 3.1 and Eqs. (2.36) and (2.37).

for these model constants observed in Fig. 2.5 for the AEM approach. The first order coefficients, α_{11} and α_{21}, corresponding to the first random variable, ξ_1, also exhibit a strong dependence. Other 2D joint distributions show little correlation between parameters α_{ij}. It should be also noted that conditions in Eq. (2.39) are reflected in some sharp boundaries observed for several 2D joint posterior densities.

Figure 2.7 compares posterior densities for C_{μ_ε} and C_ε obtained with the AEM and EEM approaches. Each set of results is represented with five equally spaced contours. The results in this figure, corresponding to the three filter widths for R3, show consistently narrower densities for the AEM approach compared to EEM. For the former approach, the posterior densities for the model parameters only partially explain the discrepancy between the model and the data, the remaining contribution manifesting through the error terms in Eq. (2.17). For the EEM approach, the model parameters are parameterized such that the output density is consistent with the spread in the data. For this approach the discrepancy between the model, in this case the k^{sgs} sub-grid model, and the data is completely folded into the density of C_{μ_ε} and C_ε.

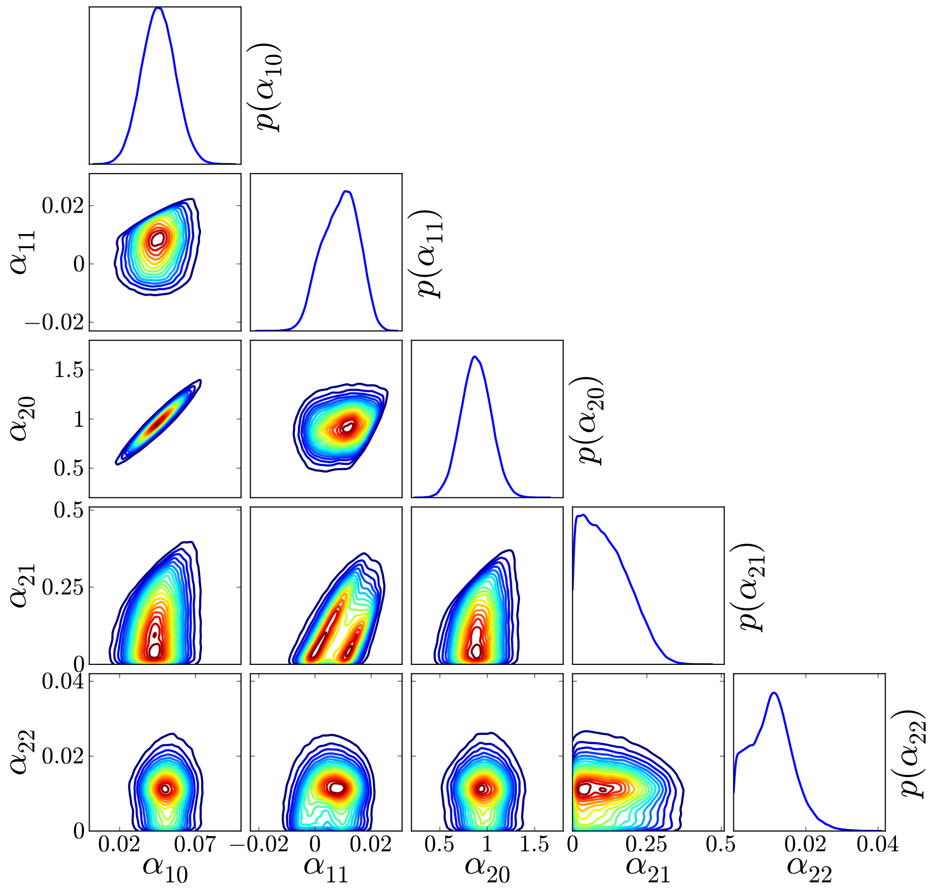


Figure 2.6: Joint posterior distribution for EEM parameters α_{ij}, corresponding to R3_2.

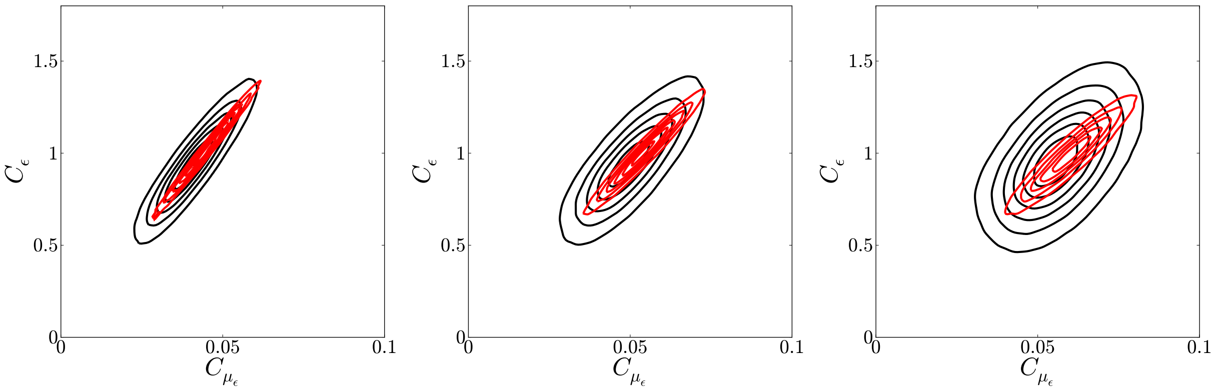
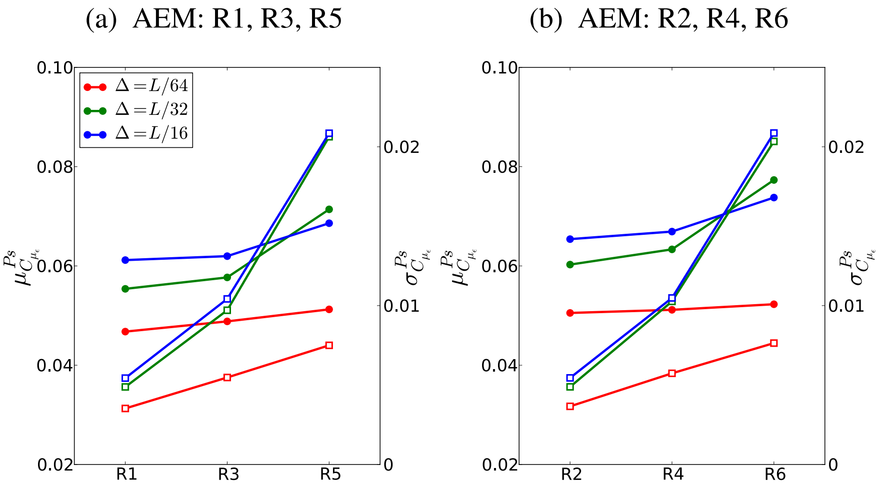
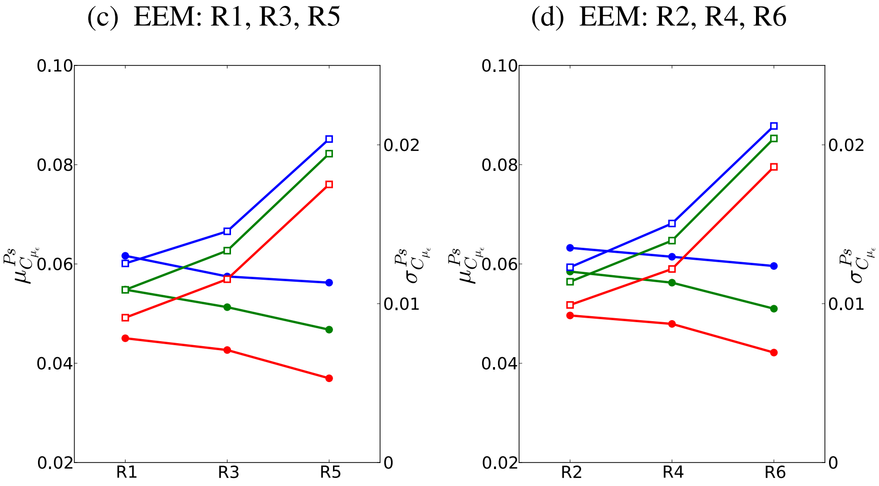

(a) R3_1 (b) R3_2 (c) R3_3

Figure 2.7: Comparison of joint posterior distribution for $(C_{\mu_\varepsilon}, C_\varepsilon)$, for AEM (red contours) and EEM (black contours). For EEM, the posterior densities for $(C_{\mu_\varepsilon}, C_\varepsilon)$ are generated by sampling the posterior densities of $(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2})$.

Figure 2.8 shows the posterior mean and standard deviations based on marginal posteriors for C_{μ_ε}. This figure show results for the three filter widths, discussed in Section 2.3.1, and prior model specifications, provided in Table 3.1. For the AEM approach, the standard deviation $\sigma^{Ps}_{C_{\mu_\varepsilon}}$ for marginal posterior density of C_{μ_ε}, in Figs. 2.8a and 2.8b, is about half compared to the one corresponding to its prior density. The corresponding results for the EEM approach, in Figs. 2.8(c,d), show larger marginal standard deviations for both model constants compared to AEM. The posterior mean $\mu^{Ps}_{C_{\mu_\varepsilon}}$ is more impacted by the filter width compared to $\sigma^{Ps}_{C_{\mu_\varepsilon}}$, while at the same time being less affected by the prior widths. Results corresponding to C_ε are qualitatively similar to the ones for C_{μ_ε}.


(a) AEM: R1, R3, R5

(b) AEM: R2, R4, R6


(c) EEM: R1, R3, R5

(d) EEM: R2, R4, R6

Figure 2.8: Posterior mean (left axis, filled circles) and standard deviation (right axis, white squares) for C_{μ_ε}.

In order to measure the degree of dependence between C_{μ_ε} and C_ε, we examine the "distance correlation" values [75] estimated based on the MCMC samples from their posterior distribution. The distance correlation is a measure of *dependence* between two random variables, and vanishes if and only if they are independent. Given random variables X and Y with finite first moments, the distance correlation $\mathscr{R}(X,Y) \in [0,1]$ is defined as

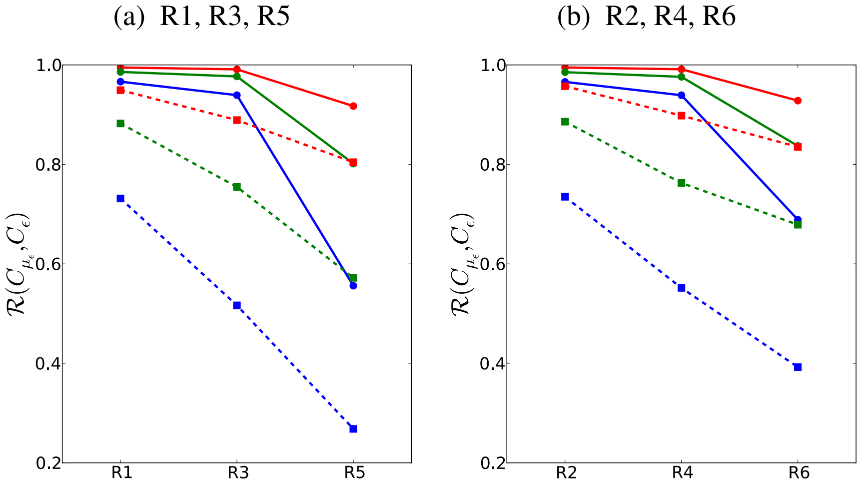
$$\mathscr{R}(X,Y) = \frac{\vartheta^2(X,Y)}{\sqrt{\vartheta^2(X)\vartheta^2(Y)}} \tag{2.41}$$

where $\vartheta^2(X,Y)$ is the “distance covariance" between X and Y and $\vartheta^2(X)$ is the “distance variance," $\vartheta^2(X) = \vartheta^2(X,X)$. The distance covariance is given by

$$\begin{aligned}\vartheta^2(X,Y) =& E(||X-X'||||Y-Y'||)) + E(||X-X'||)E(||Y-Y'||) \\ &- 2E(||X-X'||||Y-Y''||))\end{aligned}$$

where (X',Y'), (X'',Y'') are pairs of RVs with the same joint distribution as (X,Y), and $E(\cdot)$ denotes expectation.

Figure 2.9 shows the dependence of $\mathscr{R}(C_{\mu_\varepsilon}, C_\varepsilon)$ on the prior witdth for the two model parameters. For all model settings the distance correlation values between C_{μ_ε} and C_ε corresponding to the AEM approach is greater than approximately 0.6 indicating a relatively strong dependence between the two model parameters. For all filter sizes employed to filter the DNS data, these dependecies are largest when the prior is wide and decrease as the prior narrows and hence has a stronger effect on the posterior distributions. Since the priors for the two model constants are independent, this leads to weaker dependencies, i.e. smaller $\mathscr{R}(C_{\mu_\varepsilon}, C_\varepsilon)$ values. The distance correlation results corresponding to the EEM approach are consistently smaller compared to the ones for the AEM approach. This observation is consistent with the wider posterior distributions for EEM compared to AEM, shown in Fig. 2.7 for select runs.


(a) R1, R3, R5

(b) R2, R4, R6

Figure 2.9: Distance correlation dependence on the standard deviation of the marginal prior densities of C_{μ_ε} and C_ε via AEM (solid lines) and EEM (dashed lines). The color scheme corresponds to several filter sizes, and is the same as in Fig. 2.8.

This last observation is not necessarily general when comparing AEM and EEM approaches. While we do not provide counter-examples here, we believe that the degree of dependency between model parameters is case dependent and a general conclusion should not be drawn just based on the examples provided in this paper.

Posterior Predictive Distributions for Filtered DNS Data

In this section we explore the predictive skill of the k^{sgs} model given the posterior distributions for C_{μ_ε} and C_ε. We employ the Bayesian posterior predictive distribution [39] to determine how well the model covers the filtered DNS data. This test examines the probability density of the predicted data, $z^p = f_k^p$, and $x_o{}^p = \{f_P^p, f_D^p\}$, conditional on the existing data, z and x_o. Essentially it indicates how well the spread in the existing data is captured by the calibrated model. For the AEM approach the posterior predictive density $p^{pp}\left(z^p, x_o{}^p | z, x_o\right)$ is given by

$$p\left(z^p, x_o{}^p | z, x_o\right) = \int p(z^p|\theta, \theta_{\varepsilon,z}, x_a) p(x_o{}^p|x_a, \theta_{\varepsilon,x}) p(\theta, \theta_{\varepsilon,z}, x_a, \theta_{\varepsilon,x}|z, x_o) d\theta\, d\theta_{\varepsilon,z}\, dx_a\, d\theta_{\varepsilon,x} \tag{2.42}$$

Here, $p(z^p|\theta, \theta_{\varepsilon,z}, x_a) \sim N(f_P^a C_{\mu_\varepsilon} - f_D^a C_\varepsilon, \sigma_z^2)$, and $p(x_o{}^p|\theta_{\varepsilon,x}, x_a) \sim N(x_a, \Sigma_x)$. For the EEM approach the model error is embedded in the posterior densities for the model parameters. Hence there is no dependency on $\theta_{\varepsilon,z}$ in *rhs* of Eq. (2.42) for EEM.

Figure 2.10 shows the data $z = \{f_{k,i}, i = 1, \ldots, N_t\}$ with red, green, and blue symbols, respectively. The $5-95\%$ quantile range, with filled black symbold, corresponds to the marginal posterior predictive densities for f_k^p. These densities were estimated by Monte Carlo sampling, based on Eq. (2.42). The top, middle, and bottom frames, correspond to filter widths $L/64$, $L/32$, $L/16$, respectively. For both AEM and EEM, the posterior predictive results for all runs cover well the data spread observed in the fitered DNS data. The $5-95\%$ quantile range corresponding to the pushed-forward density for f_k^p is marked by open symbols for AEM. This density is obtained by neglecting the error model $\theta_{\varepsilon,z}$ in Eq. (2.42). These results show that pushed-forward densities consistently underpredict the data spread for all cases in this study.

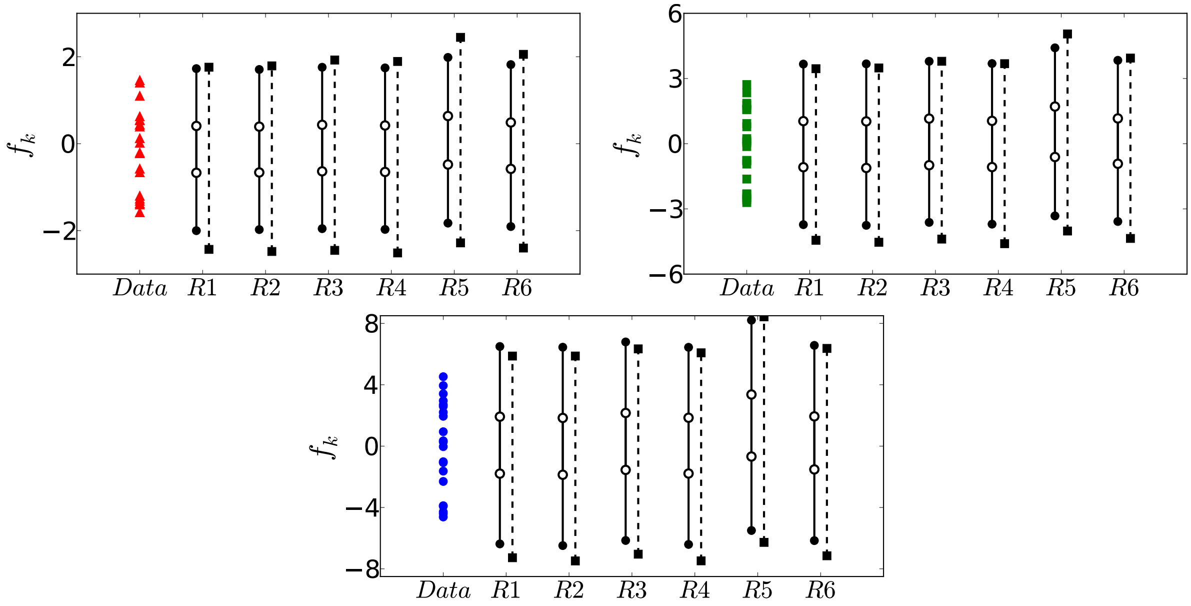


Figure 2.10: Posterior predictive $5-95\%$ quantile ranges AEM (solid lines) and EEM (dashed lines). The filtered DNS data is shown symbols using the same color scheme as in Fig. 2.4.

For the AEM approach this indicates that, in order to preserve the predictive capabilities of a model, one needs to consider both the posterior densities for model parameters as well as accounting for the error model. Since the error term is connected to a certain configuration, in this case the filtered DNS data, it is not clear how to estimate posterior predictictive distributions in configurations other than the one used for the calibration process. For the EEM approach this difficulty is circumvented by the fact that discrepancy terms are subsumed in the posterior densities for model parameters. For this approach the computation of posterior predictive densities does not require accounting for explicit error models.

We also present results for the marginal posterior predictive densities corresponding to ${x_o}^p$, to verify the approach for the error-in-variable model. Figure 2.11 shows the filtered DNS data x_o with symbols and the joint marginal posterior predictive densities, corresponding to R3, are shown with iso-contours. For clarity, we only present one iso-contour for each case, corresponding to the 5% boundary of its maximum density. These results confirm that the error model for the independent variables captures well the dependencies and spread in the independent variables f_P and f_D. The results for other runs (not shown) are nearly the same as the ones in this figure for R3 since the prior densities for $\theta = \{C_{\mu_\varepsilon}, C_\varepsilon\}$ have a negligible impact on the posterior density for x_a.

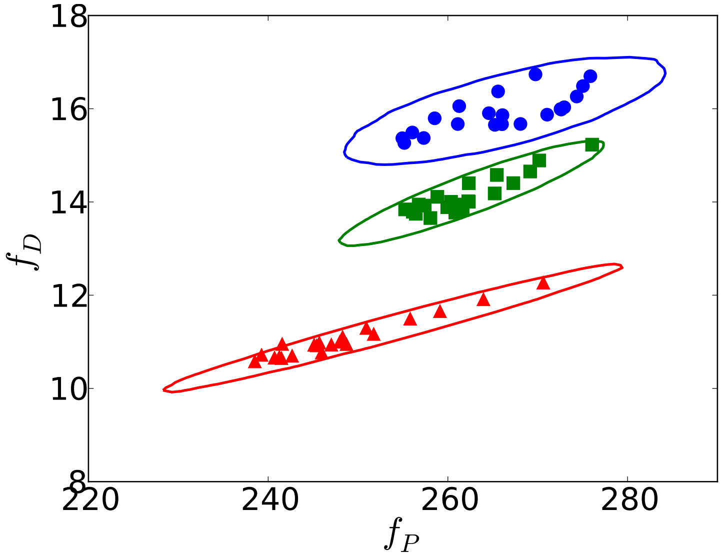


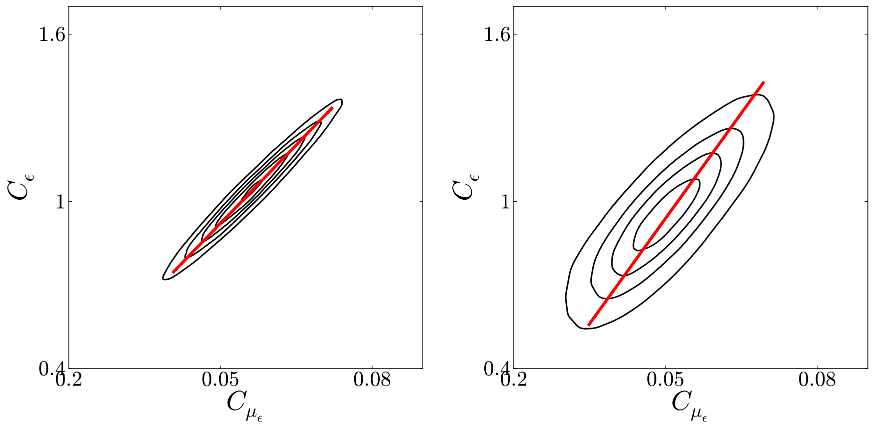
Figure 2.11: Joint posterior predictive densities for $x_o^p = (f_P^p, f_D^p)$ corresponding to $R3_1$ (red), $R3_2$ (green), and $R3_3$ (blue). The iso-contours correspond to the 5% of the maximum density for each run. The filtered DNS data x_o is shown with symbols using the same color scheme as in Fig. 2.4.

Principal Component Analysis of Joint PDFs

Given the strong dependence in the joint posterior distribution of C_{μ_ε} and C_ε for AEM, we proceed with a Principal Component Analysis [31] (PCA) of these distributions. Through the PCA approach, we aim to identify linear combinations of these parameters that explain the MCMC samples. Let $X \in \mathbb{R}^{n\times 2}$ be a matrix with n rows and two columns, first column for C_{μ_ε} and second column for C_ε. Each row represents one MCMC sample. Without loss of generality, assume that the mean values are already subtracted from X.

The principal components of X are the eigenvectors of the covariance matrix $X^T X \in \mathbb{R}^{2\times 2}$, and the corresponding non-zero eigenvalues represent the variances of the MCMC data along the principal directions. For this study there are two principal directions corresponding to the parameter space spanned by C_{μ_ε} and C_ε. For the AEM approach, the principal component analysis indicates that variances attributed to the 2^{nd} component are less that 5% of the variances along the 1^{st} componnent. This is consistent with the degree of dependence observed by visual inspection in Fig. 2.5 and by the distance correlation results in Fig. 2.9. For the EEM approach the variances along the 2^{nd} component are about $15-20\%$ of the ones for the 1^{st} component. This is consistent with the wider joint posterior distributions for C_{μ_ε} and C_ε for EEM compared to AEM.

In the next section, we will explore the dependence between C_{μ_ε} and C_ε along the 1^{st} PC for several cases. Figure 2.12 shows segments along the 1^{st} PC line for the $R3_2$ case corresponding to AEM and EEM, respectively. The endpoints of these segments correspond to the 5%-95% quantile range obtained from MCMC samples projected on the 1^{st} PC.


(a) $R3_2$ via AEM

(b) $R3_2$ via EEM

Figure 2.12: Joint posterior distribution via AEM (left frame) and EEM (right frame) for $R3_2$. The red segments show the first principal component computed from MCMC samples. The ends of each segment correspond to the 5%-95% quantile range for MCMC samples projected on the 1^{st} principal component.

2.4 Forward UQ

In this section we describe an approach for pushing forward the densities of model parameters $\theta = (C_{\mu_\varepsilon}, C_\varepsilon)$ and obtain probability densities for Quantities of Interest (QoIs) predicted by LES models. Figure 2.1 shows a schematic for the Forward UQ process in the bottom half. Below we provide a description of a non-intrusive approach, followed by a description of the LES code and the configuration setup.

The non-intrusive propagation of uncertainty from model parameters θ to QoI y, proceeds through the following procedure. Given a basis of standard random variables $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$, and a known functional form $\theta = \theta(\xi)$, we employ the representation in Eq. (2.29) to cast the QoI y derived from the output of LES model f as a PCE expansion

$$y(\theta(\xi)) \approx \sum_{k=0}^{N_t} c_k \Psi_k(\xi_1, \xi_2, \ldots, \xi_n). \tag{2.43}$$

The coefficients of this PCE are evaluated by Galerkin projection exploiting the orthogonality of basis terms

$$c_k = \langle y \Psi_k \rangle / \langle \Psi_k^2 \rangle. \tag{2.44}$$

Evaluation of the projection integrals in Eq. (2.44) requires a number of evaluations of the model f. For a small to moderate number of dimensions, numerical quadrature provides an efficient way to evaluate c_k using a number of samples much smaller compared to Monte Carlo sampling algorithms. The quadrature approach involves evaluating $y = f(\theta)$ for a particular set of parameter values $\theta^j = \theta(\xi^j)$ with $\xi^j = (\xi_1^j, \xi_2^j, \ldots, \xi_n^j)$ corresponding to the set quadrature points required by the accuracy of the method. The PCE coefficients are then computed as

$$c_k = \frac{1}{\langle \Psi_k^2 \rangle} \sum_{j=1}^{N_q} w_j f(\theta^j) \Psi_k(\xi^j) \tag{2.45}$$

where, w_j is the weight corresponding to the quadrature point ξ^j, and N_q is the number of quadrature points.

Once the PCE is constructed for QoI y, moments of the distribution of y can be estimated analytically based on the expansion coefficients c_k [19]. For example the expectation of y is given by c_0. The PCE in Eq (2.43) can also be used as a "surrogate" when the computational model is expensive and/or the probability density of the input parameters are updated. For example moments for several QoIs can be evaluated with negligible cost via PCE evaluations compared to the full model evaluations. In the following sections we describe the setup for the LES simulations, followed by additional manipulation of the input parameter space to help construct PCE for several QoIs.

2.4.1 Channel Flow Setup

In this section we describe the LES model f employed to propagate densities from model parameters to output QoIs. For this study, the model f, referenced in Fig. 2.1, is represented by the FUEGO LES software [44], developed by Sandia National Labs. It is designed to simulate turbulent reacting flow and heat transfer on massively parallel computers, and was built on the SIERRA framework [73]. In this work the low-Mach number module is used. The discretization scheme is based on the control volume finite element method [43] where the partial differential equations of mass, momentum, and energy are integrated over unstructured control volumes. We employ the values of C_{μ_ε} and C_ε, calibrated against the the canonical forced isotropic turbulence DNS data in the first part of this paper, for the forward UQ study in a parallel plane channel flow at a Reynolds number of $\mathrm{Re}_\tau = 590$. This value was chosen to match the DNS results by Moser *et.al.* [46].

The channel dimensions are $2\pi h \times 2h \times \pi h$ in the streamwise, wall-normal, and spanwise directions, respectively. The boundary conditions are periodic in both the streamwise and spanwise directions, and no slip walls are applied at both the top and bottom boundaries. There are 48 uniformly spaced grid nodes in both the streamwise and the spanwise directions. In the wall normal direction there are 103 nodes, with the first four nodes from the walls placed uniformly with a $y^+ = 1$ spacing. The grid is then stretched through a hyperbolic tangent function until the spacing at the centerline is approximately equal to the spacing in the spanwise direction. We note that our grid setup is considerably more coarse than the fully resolved DNS case for a $Re_\tau = 590$ presented by Moser *et.al.* [46] in which the number of nodes were $383 \times 257 \times 383$ in the streamwise, wall normal, and spanwise directions respectively. We employ a CFL number of about 0.5, resulting in a time step of approximately 10^{-4}. The time coordinate is normalized by the channel half-width h.

2.4.2 Rosenblatt Transformation

In order to construct the PCE presented in Eq. (2.43) we map the parameter space θ that employs dependent densities for its components to the ξ space where all coordinates have independent densities. For this study, the θ space consists of $(C_{\mu_\varepsilon}, C_\varepsilon)$, the ξ space consists of (ξ_1, ξ_2), and the map between the two can be constructed via the Rosenblatt transformation [63]. The inverse of the Rosenblatt transformation can formally be written as

$$\begin{aligned} C_{\mu_\varepsilon} &= \tilde{F}^{-1}_{C_{\mu_\varepsilon}}(\xi_1), \\ C_\varepsilon &= \tilde{F}^{-1}_{C_\varepsilon | C_{\mu_\varepsilon}}(\xi_2 | \xi_1), \end{aligned} \tag{2.46}$$

where ξ_1, ξ_2 are i.i.d. uniform random variables. Eq. (2.46) relates the regular, *mapped* domain $[0,1] \times [0,1]$ to the pair $(C_{\mu_\varepsilon}, C_\varepsilon)$ and uses the inverses of the marginal, $\tilde{F}^{-1}_{C_{\mu_\varepsilon}}$, and conditional, $\tilde{F}^{-1}_{C_\varepsilon | C_{\mu_\varepsilon}}$, cumulative distribution functions (CDF) for the random variable pair $(C_{\mu_\varepsilon}, C_\varepsilon)$.

Numerically, the Rosenblatt transformation is computed through interpolation tables constructed for the CDF functions in Eq. (2.46). To illustrate this approach we use the joint density for the uncertain parameters C_{μ_ε} and C_ε, corresponding to an average of the joint posterior densities for runs R3 and R4 via the EEM approach. This ensures that LES runs cover a region wide enough in the parameter space, leading to a PCE that can be used as a surrogate to push forward the joint distributions for C_{μ_ε} and C_ε corresponding to all runs R3 and R4.

Figure 2.13 shows the correspondence between physical parameter space $(C_{\mu_\varepsilon}, C_\varepsilon)$ and the space parameterized by uniform random variables (ξ_1, ξ_2). The markers show the location of quadrature points corresponding to a Gauss-Legendre quadrature approach. Their images in the

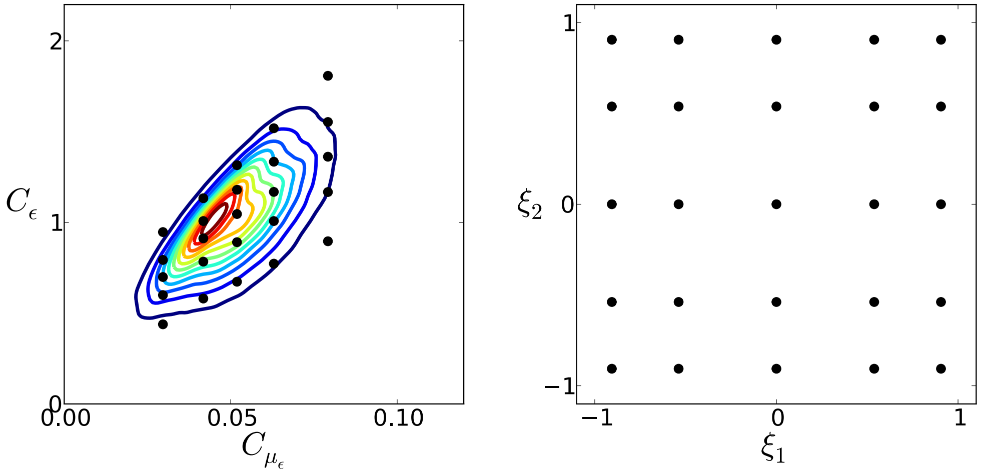


Figure 2.13: Illustration of the correspondence between samples in the parameter space and their images via the Rosenblatt map to the space parameterized by (ξ_1, ξ_2).

physical space are mostly distributed around the regions where the joint density $p(C_{\mu_\varepsilon}, C_\varepsilon)$ is significant.

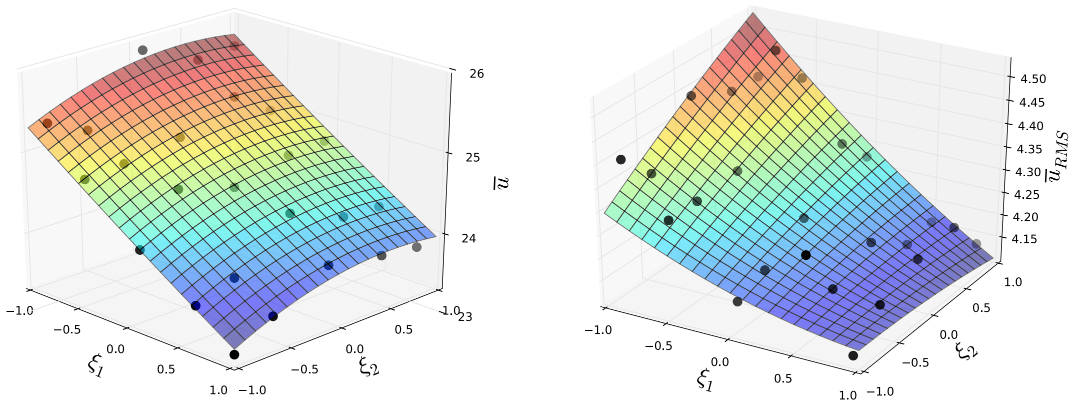
2.4.3 Polynomial Chaos Expansions for LES Quantities

We performed a set of LES simulations using the setup for the channel flow described in Section 2.4.1. Each simulation corresponds to C_{μ_ε} and C_ε values selected according to the Rosenblatt transform presented in Section 2.4.2. Using the raw LES data, we computed several QoIs. Specifically, we estimated the centerline velocity, $\overline{u}$, the maximum velocity RMS, $\overline{u}_{RMS}$, and the mass flux, $\overline{m}$. All these QoIs are averaged in time over approximately five flow-through times. Using the estimated values for these QoIs we then compute the coefficients of their corresponding PCE via the Galerkin projection approach presented in Eqs. (2.43)-(2.45). We then compared the L_2 error between 1-st, 2-nd, and 3-rd order PCEs, respectively, and the LES data, and concluded that 2-nd order polynomial approximations are sufficient to capture the QoI dependence on the model parameters.

The mean centerline velocity, in Fig. 2.14a, exhibits a quasi-linear trend in the ξ_1 direction and a weak quadratic trend in ξ_2. The mean values of the peak axial velocity RMS, in Fig. 2.14b, exhibit a quadratic dependence on the input parameters while the mean mass flux, in Fig. 2.14c, shows a shape similar to the mean centerline velocity. For this work the ξ_1 coordinate is mostly aligned with the first principal component of $(C_{\mu_\varepsilon}, C_\varepsilon)$, while ξ_2 is proportional to C_ε.

We examined the relative L_2 error between LES simulations and the PCE values for these QoIs. These errors are around 6% for $\overline{u}$ and $\overline{m}$, and less than 10% for $\overline{u}_{RMS}$. While, in principle, longer time averages can reduce the residual sample average noise in the LES values and allow lower discrepancies compared to the corresponding PCEs, we consider these errors to be reasonable for the subsequent use of the PCEs as surrogate models.

(a) Axial Velocity on the Centerline (b) Peak Axial Velocity RMS



(c) Mass Flux

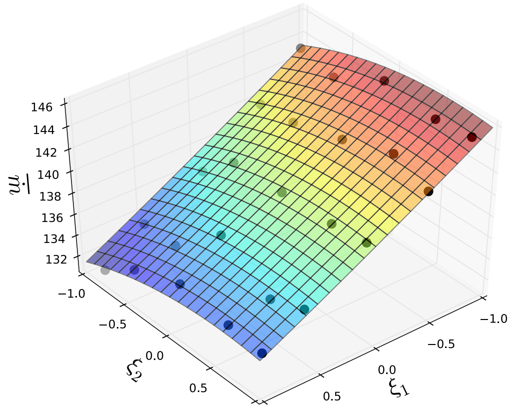


Figure 2.14: Polynomial Chaos expansions several QoI as a function of (ξ_1, ξ_2).

2.4.4 Probability Densities for Quantities of Interest

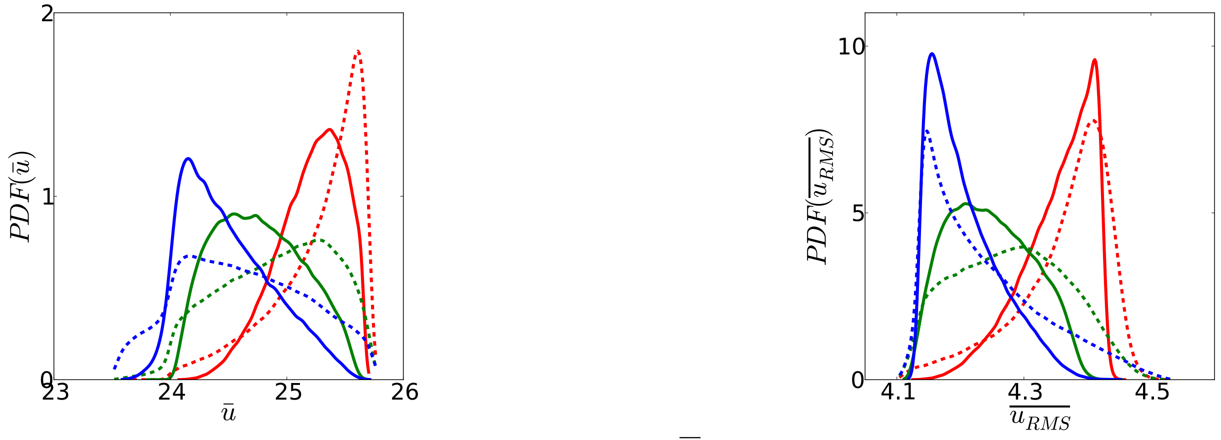
We employ the PCEs discussed in the previous section to evaluate QoIs values based on posterior densities for C_{μ_ε} and C_ε discussed in Section 2.3. The resulting PCE-approximated model evaluations are then used to compute the PDFs of the output QoIs via KDE. This procedure is repeated for all calibration studies shown in Table 3.1.

Figure 2.15a shows the densities for the mean centerline velocity $\overline{u}$ for R3 corresponding to several filter widths. The results in this figure indicate a strong effect exercised by the filter size, on the PDF of $\overline{u}$. The parameter densities corresponding to smaller filter widths, i.e. $\Delta = L/64$ (red lines), shift the center of mass to higher velocity values. Larger filter widths shift the most likely values for the centerline velocities towards smaller values which are closer to the DNS value. Earlier DNS results [46] indicate a reference value of 21.26 for the mean centerline velocity. We attribute the discrepancy between the DNS value and our LES study to other sources of error that can impact this configuration such as insuficient grid resolution, in particular near the walls.

The wider PDFs observed for $(C_{\mu_\varepsilon}, C_\varepsilon)$ with the EEM approach translate into wider PDF's for

(a) Mean Centerline Velocity, $\overline{u}$ (21.26)

(b) Mean Velocity RMS, $\overline{u_{RMS}}$ (2.74)



(c) Mean Mass Flux, $\overline{\dot{m}}$ (117.2)

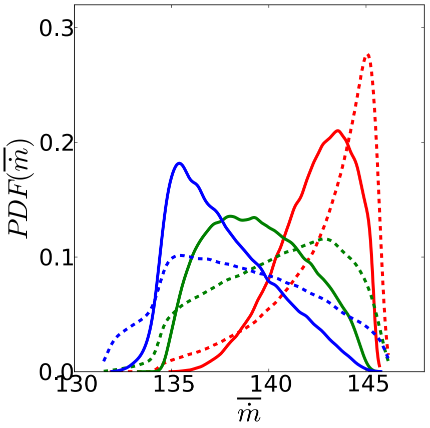


Figure 2.15: PDF's for several quantities of interest corresponding to LES results in a channel flow configuration. Red, green, and blue correspond to filter widths $\Delta = \{L/64, L/32, L/16\}$ for R3. Solid lines show pushed forward PDF's for the AEM approach while dashed show EEM results. The DNS values [46] are shown in parantheses in the corresponding captions.

the mean centerline velocity compared to AEM values. This is consistently observed for all filter widths. Nevertheless, all these runs overpredict the DNS values, suggesting that other modeling parameters, e.g. grid resolution have an substantial impact on the selected QoIs. Results corresponding to other runs are qualitatively similar to the ones observed for R3. Both the peak RMS of the axial velocity, in Fig. 2.14b, and the mass flux, in Fig. 2.14c, show output densities that are qualitatively similar to results for the centerline velocity.

Next we compare the posterior densities obtained by sampling the full joint PDFs of C_{μ_ε} and C_ε with the posterior PDFs obtained by sampling the density along the corresponding 1^{st} PCs. The results shown in Fig. 2.16 correspond to $R3_2$. For the AEM approach the the results based on the densities along the 1^{st} PC, with dashed red line, are in close agreement with the results based on full joint PDFs with solid red line. This indicates that, for this particular study, the strong dependence between the two input parameters for AEM effectively renders the output being dependent only on the linear combination of the two parameters. Unlike for AEM, the EEM results show a significant discrepancy between the full joint posterior and the values based on the 1D density along the 1^{st} PC. For this set of cases the width of the joint posterior is important to the system behavior and the reduced dimensionality parameter set does not capture the full range of results. The PCA-based results for other runs and QoIs are qualitatively similar to the ones observed in Fig. 2.16.

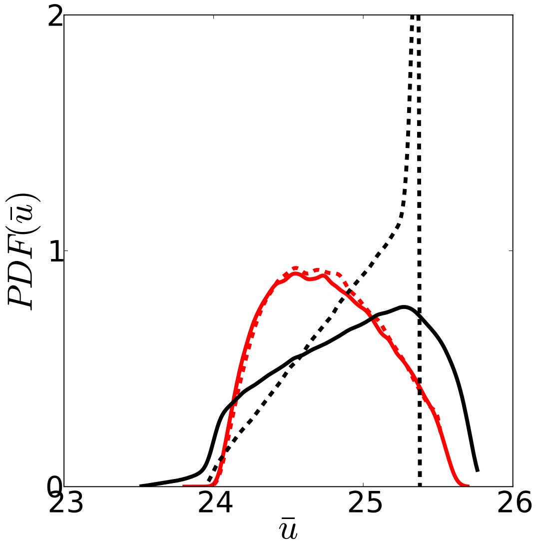


Figure 2.16: Output PDFs for mean centerline velocity based on forward propagation of the joint PDFs of C_{μ_ε} and C_ε (solid lines) and the joint density along the 1st PC (dashed lines). Results corresonding to $R3_2$, AEM results shown in red and EEM in black.

2.5 Conclusions

In this paper we present a Bayesian framework for estimating joint distributions for Large-Eddy Simulation (LES) sub-grid scale model parameters based on canonical forced isotropic turbulence Direct Numerical Simulation (DNS) data. These densities are then propagated forward through LES of channel flow to generate probability densities for several quantities of interest.

We employ DNS of forced isotropic turbulence, downloaded from the Johns Hopkins Turbulence Database, to calibrate two parameters for the sub-grid scale turbulent kinetic energy model employed in LES against fundamental configurations. We employ Gaussian priors for the model parameters, centered at values previously reported in the literature for these parameters. We then explore the effect of prior widths on the posterior densities and several summary statistics of interest. In this context we compare two approaches for treating the discrepancy between the model and the data in the Bayesian framework. In the first approach, abbreviated as AEM, the discrepancy is presumed Gaussian with zero mean, and its standard deviation is inferred together with the model parameters. In the second approach, developed recently by some of the co-authors of this paper and abbreviated as EEM, the error is embedded in the formulation of the model parameters. In both approaches we account for noise in the independent variables via a classical error-in-variables model.

We find that the filter size, employed to filter the DNS data, has a significant impact on the posterior means for the model parameters, while the effect on the marginal posterior standard deviations is weaker. We also observe that prior widths affect mostly marginal posterior standard deviations, while the priors means are less important for the posterior means. We find that the EEM approach leads to wider joint densities for the model parameters compared to the AEM approach. This also results in weaker dependencies between the two model parameters for EEM, while for AEM, these dependencies are very strong.

In the second part of the paper, the joint posterior densites for the LES model parameters are propagated forward to obtain densities for several quantities of interest (QoI). We employ a non-intrusive approach, and construct Polynomial Chaos expansions (PCEs) for these QoIs. The expansion coefficients are estimated via Galerkin projection. Specifically, the projection integrals are computed via Gauss-Legendre quadrature. Twenty-five LES channel flow simulations were run with the two model parameters set according to the corresponding quadruature points. Once the coefficients of these PCEs were available, the densities for the corresponding QoIs were evaluated through PCE sampling. We found that the first principal component corresponding to the joint posterior density leads to forward UQ results that are very close to the ones based on the full densities in the AEM approach which show narrow joint posterior densities. Unlike the AEM approach, EEM leads to wider posterior densities, accounting for model error, and the first principal component is no longer sufficient to explain the dependence on the input parameters. Despite the wider range of sub-grid model parameter values, the EEM approach still fails to predict the DNS values for the QoIs selected for this study. Preliminary studies, to be presented in a subsequent paper, indicate that grid resolution near the wall plays an important role in predicting LES values that are similar to the DNS values.

The use of informative priors was necessary to constrain the parameter space as the isotropic turbulence data was sufficient to constrain the ratio between the model parameters considered here, but not their values. It is suggested that in the future this calibration process use data from multiple complementary flows to better estimate all unknowns. Moreover, current results suggest that the filter width should also be calibrated along with other model patameters, as it is likely that the filter width will play a significant role in the resulting model form. It will also be important to match the filter width when propagating these uncertain turbulence models through a flow as this work provides data showing the impact of the filter width on important flow quantities.

Chapter 3

Calibration of LES wall models for Channel Flow

Abstract: Engineering Large Eddy Simulation (LES) is more computationally affordable and fits better into a design time frame than high-fidelity LES, which is more accurate but significantly more computationally expensive. In order for engineering LES to realize its full potential utility, efficient model calibration and error estimation techniques are required.With this approach engineering scale LES, which has a much coarser grid and faster turn-around time, can be of sufficient, quantified fidelity to be useful in the design cycle. In this study we highlight a Bayesian calibration approach for channel flow with $Re_\tau = 590$ followed by a forward uncertainty quantification study for quantities of interest (QoI). Calibration of model parameters for a subgrid scale LES turbulent kinetic energy model is done in two regions of a channel flow: the near wall region and the center of the channel. The calibration parameters are changed separately for each region for three heights of the near wall region: $y^+ \approx 16$, 32, and 48. Radial basis functions are used to construct surrogate models for the quantities of interest with respect to the model parameter values. Posterior densities for the model parameters are then estimated in a Bayesian framework employing data from Direct Numerical Simulation (DNS) of channel flow. Once we demonstrate successful calibration in the channel flow configuration, these densities are then propagated forward to obtain probability densities for several QoI. We employ non-intrusive radial basis function techniques for an efficient propagation of uncertainties from input model parameters to output QoI. We have determined that by calibrating near wall and center flow regions separately, QoI values can be more accurately predicted. We test how well this calibration works for other flows by looking at a backward facing step case at the same Re_τ as the original calibration as well as channel flows with $Re_\tau = 395$ and 950.

3.1 Introduction

For the design of mechanisms that have complex physics and chemistry, such as gas turbine engines, there is a need to balance the cost with accuracy. In Direct Numerical Simulation (DNS), the grid is fine enough that all levels of turbulence are solved for directly. This would have the accuracy needed, but is very expensive in terms of the computational time, so thus far only relatively simple geometries have been solved with DNS. Reynolds Averaged Navier-Stokes (RANS) is at the other end of the computational fluid dynamics (CFD) spectrum where only the mean flow structures are solved for and the turbulence is handled through models. While this greatly reduces computational time, it comes at the cost of accuracy since the empirical turbulence models suffer from high uncertainty in many flows of engineering relevance. Large Eddy Simulations (LES) are a compromise between these two regimes: the large turbulent scales are solved for directly, but the smaller subgrid scales are handled with turbulence models. A key question when running LES is how to set the filter width, which determines how much of the turbulence is directly calculated and how much is modeled. In academic LES, usually a smaller filter width is employed such that most all of the turbulent energy is directly resolved. However, these simulations can still take weeks to run on supercomputers, making them inconvenient to incorporate into a design process. In engineering LES, faster runtimes are achieved by using a larger filter width, such that a higher fraction of the turbulent energy is modeled using the subgrid scale model. In order to make engineering LES useful, it is critical to be able to account for the added uncertainty associated with the more active subgrid scale model. In this paper, we present a Bayesian framework to rigorously quantify that uncertainty, using fully developed channel flow as test cases.

In a previous paper [65], we used a Bayesian framework to incorporate directly filtered DNS turbulence information to estimate uncertainties in the parameters for a sub-grid scale turbulence model. This was followed by a forward UQ study on the predictive capabilities of this approach. In this previous work, we were able to successfully set up a framework. However, the predicted quantity of interest (QoI), which in this case was the mean centerline velocity of a channel flow, was significantly off from the expected DNS value. Therefore for this paper we incorporate a wall model concept where the turbulence parameters are tuned separately in a near wall region and in a bulk flow region. Unlike other wall models that impose a RANS-like wall stress and eddy viscosity [77], the wall models for this study would have the values for the subgrid scale kinetic energy, C_ε and C_{μ_ε}, in the wall region calibrated separately from the center of the channel flow.

3.2 Methods

3.2.1 Meshes and Setup

The height of the well resolved boundary layer was one of the parameters investigated for this study. To this end, three distinct meshes for a channel flow were constructed. The channel dimensions for all three remained the same and are $2\pi h \times 2h \times \pi h$ (where h is the half height of the channel) in the streamwise, wall-normal, and spanwise directions, respectively. This configuration was chosen to match simulations performed by Moser, et al. [46] so that quantities of interest could be compared. The boundary conditions are periodic in both the streamwise and spanwise directions, and no slip walls are applied at both the top and bottom boundaries. The streamwise and spanwise spacing for all three meshes was kept the same as well, namely, 49 equally spaced nodes in both of those directions. There is a total of 233k nodes, which can be compared to 37 M nodes used by Moser, et al. [46].

The wall normal node spacing was changed for each of the three cases. It was determined what spacing would fully resolve the channel flow, and then this grid was âĂIJchoppedâĂİ at three different heights from the wall. The center portion of the grid was replaced with a much coarser mesh. For the wall region, the first four nodes were placed with a spacing of $\delta y^+ \approx 1$. After that the spacing for the wall region was determined by using hyperbolic stretching over a length of h that started with a spacing that results in a y^+ value of one, and ends with close to the half of the spanwise direction spacing. When making this fine mesh, there are 96 nodes between the wall and the channel center, which is comparable to the number of nodes that would fully resolve the flow [77]. For the smallest near wall region, the first 12 intervals from this spacing were taken, corresponding with a y^+ value of approximately 16. For the near wall region with the medium height there were 18 intervals for a height of $y^+ \approx 32$, and for the largest near wall region there were 23 intervals in $y^+ \approx 48$. The spacing for the inner channel was then adjusted so that there would always be a total of 97 nodes in the wall-normal direction (across the $2h$ height of the channel) which is about half of what would be used in a fully resolved mesh. The number of nodes was kept constant to determine how to best allocate the nodes in the channel with a fixed amount of computing resources. Hyperbolic tangent stretching was also used here, starting with the spacing at the edge of the resolved near wall region and going to a spacing that is close to the same spacing as what is in the spanwise direction. While there is the same number of nodes in the wall normal direction, the concentration of where those nodes are varies for each case. A comparison of the smallest and largest wall region is shown in Figure 2.

The wall region heights correspond to different regions in the turbulent boundary layer velocity profile. The smallest, $y^+ \approx 16$, is within the viscous sublayer, the largest, $y^+ \approx 48$, is in the log layer, and the middle, $y^+ \approx 32$, is at the transition between these two in the buffer layer [81].

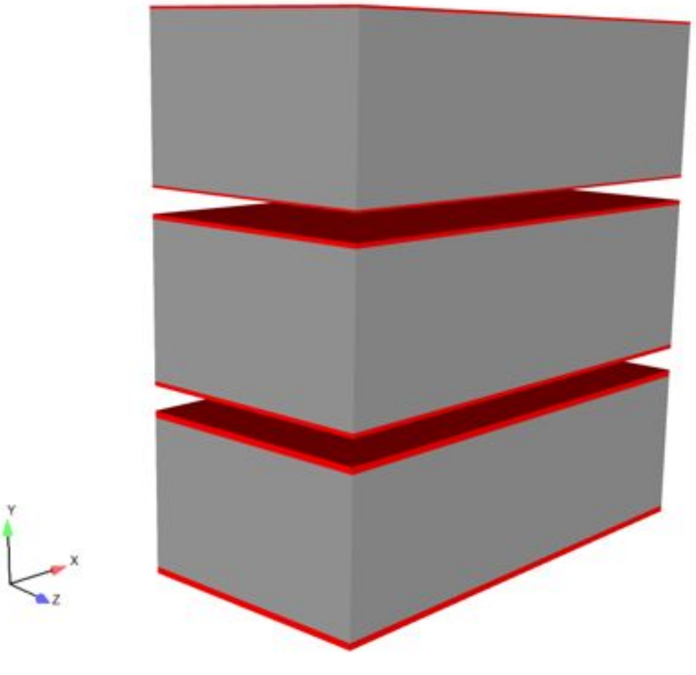


Figure 3.1: Domains with wall regions of $y^+ = 16$, 32, and 48 are shown from top to bottom. The wall region is shown in red and the center region in gray.

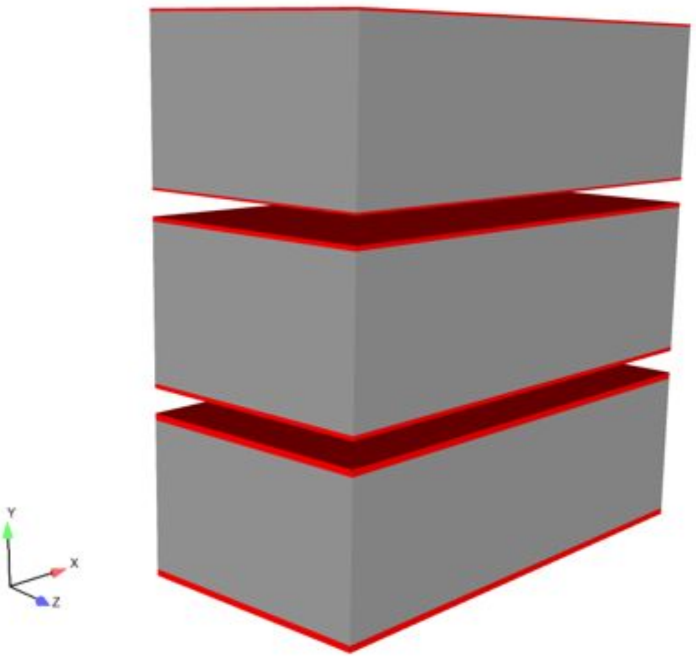


Figure 3.2: Domains with wall regions of $y^+ = 16$, 32, and 48 are shown from top to bottom. The wall region is shown in red and the center region in gray.

For the two other channel flow cases, a similar method was used to create the meshes. The spacings in the streamwise and spanwise directions were kept the same as before, but a wall region with height of $y^+ = 48$ for the specific flow velocity was created.

3.2.2 Flow Solver

Fuego is Sandia National LaboratoriesT́hermal/Fluids code designed to simulate turbulent reacting flow and heat transfer [44] on massively parallel computers built on the SIERRA framework. The code was adapted for compressible flow and combustion, and is well suited for low Mach number flows. The discretization scheme used in Fuego is based on the control volume finite element method [43], where the partial differential equations of mass, momentum, and energy are integrated over unstructured control volumes. The one equation subgrid-scale kinetic turbulent energy model (k^{sgs}) was used [7, 8], and the parameters C_ε and C_{μ_ε} for this model were calibrated using uncertainty quantification. For the calculations reported here, the first order upwind scheme was used for the convective terms. The k^{sgs} turbulence model can be written as

$$\int_V \frac{\partial \bar{\rho} k^{sgs}}{\partial t} dv + \int_{\partial V} \bar{\rho} k^{sgs} \tilde{u}_j n_j ds = \int_{\partial V} \frac{\mu_k}{\sigma_k} \frac{\partial k^{sgs}}{\partial x_j} n_j ds + \int_V \left(P_k^{sgs} - D_k^{sgs} \right) dv \tag{3.1}$$

with the dissipation term

$$D_k^{sgs} = C_\varepsilon \frac{(k^{sgs})^{3/2}}{\Delta} \tag{3.2}$$

and the production term for incompressible flows is

$$P_k^{sgs} = 2\mu_t \tilde{S}_{ij} \frac{\partial \tilde{u}_i}{\partial x_j} \tag{3.3}$$

where the filtered strain rate tensor is

$$\tilde{S}_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right) \tag{3.4}$$

and the subgrid turbulent viscosity is modeled as

$$\mu_t = C_{\mu_\varepsilon} \Delta \sqrt{k^{sgs}} \tag{3.5}$$

3.3 Results

3.3.1 Channel Flow

In a previous study, Safta et al. [65], used a Bayesian framework to incorporate filtered DNS turbulence information to estimate uncertainties in the parameters for k^{sgs} turbulence model. By varying the parameters C_ε and C_{μ_ε} found in equations 2 and 5 above for 25 channel flow cases we were able to use Polynomial Chaos Expansion (PCE) to build a response surface. When comparing the values for the mean centerline velocities of these cases with this value from a DNS simulation, it was found that the engineering LES over-predicted the centerline velocity for all values of the parameters, C_ε and C_{μ_ε}. The engineering LES values ranged from 23.3 to 26.5 which can be compared to the value of 21.26 found by Moser, Kim, and Mansour [46]. We hypothesized that a more accurate value could be reached by incorporating spatial variation into the C_ε and C_{μ_ε} parameters. In particular, we wanted to explore the potential of a wall model technique in which a different value of C_ε and C_{μ_ε} would be prescribed in the near wall region than in the bulk flow region.

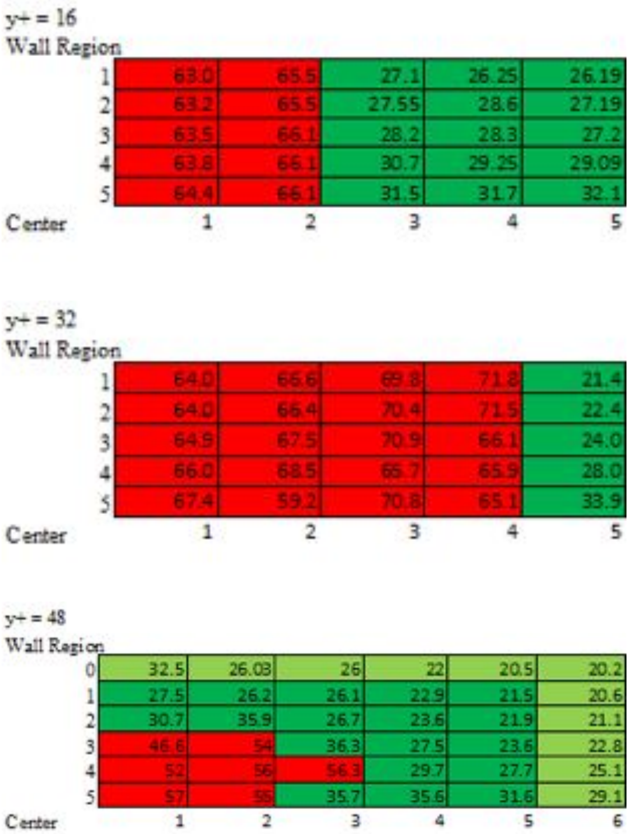
Principal Component Analysis (PCA) of C_ε and C_{μ_ε} on the filtered DNS data shows that these two parameters are highly correlated because of conservation of production and dissipation, so this was used to reduce the number of parameters calibrated in this study from 4 to 2. Because the computational cost scales exponentially with the number of parameters to be calibrated, this simplification reduced the number of simulations needed to be run by a factor of four and therefore the drastically reduced the computational resources required. Instead of varying C_ε and C_{μ_ε} separately, pairs of these values calculated from the PCA were consistently used together. Table 1 lists the values for the parameter pairs. This is an example of how DNS data/calibration can be used to make engineering calibration more tractable.

Three heights for the near wall region were investigated: $y^+ \approx 16$, 32, and 48. These heights correspond to the viscous sublayer, the buffer layer, and the log layer. For each of these heights, 25 simulations with the combinations of parameters C_ε and C_{μ_ε} 1-5 were run (see Table 3.1). It was originally intended that with this data the wall region height would also be a parameter that could be fit using PCE calibrated for the optimal height. However, not all of the runs produced turbulent flow and Figure 3.3 illustrates which became laminar and which stayed turbulent. Because so many cases failed to transition to turbulent flow for the smallest two heights, it was not possible to include the near wall region height as a calibration parameter since PCE needs a smooth function to be accurate. To move forward with the largest wall region an additional 11 cases, with value pairs zero and six from Table 1, were added to the original 25 so that the mean centerline velocity values would bracket the DNS answer.

Table 3.1: Parameters used to create response surface for the Bayesian data.

Pair	C_{μ_ε}	C_ε
0	0.1118	σ_1^{pr}
1	0.0958	σ_1^{pr}
2	0.0797	σ_2^{pr}
3	0.0563	σ_2^{pr}
4	0.0328	σ_3^{pr}
5	0.0167	σ_3^{pr}
6	0.0087	σ_3^{pr}

Nevertheless, the results from these three wall region heights are interesting from a modeling perspective. A general trend which was observed was that cases with small values for C_ε and C_{μ_ε}, and therefore less production and dissipation feedback in the k^{sgs} model (as well as generally less modeled stress), in the center region are more likely to switch to turbulent flow. For the case with the smallest wall region which place the transition in the viscous sublayer, even the turbulent cases all had mean centerline velocity values that were too high compared with the DNS data. The running averages of the mean centerline velocities for these cases are shown over the time span of the simulation in Figure 4. When the wall layer height was at the buffer layer, simulations only became turbulent for the cases where the C_ε and C_{μ_ε} values were the smallest in the center region. We speculate that a reason for this might be that the shear instabilities which generate the turbulence can be either somewhat mimicked with a model or must be resolved but are harder to capture using a combination of both. The buffer layer has a sensitive and unstable nature which makes it a poor candidate for a point at which to abruptly change modeling coefficients. However, there was one case with $y^+ = 32$ which did reach a mean centerline velocity that was very close to the DNS value. The cases with the wall region height in the log layer ($y^+ = 48$) yielded the most promising results. Only 7 of the 25 cases remained laminar, and the value for the mean centerline velocity was close to the DNS value in several cases.



y+ = 16

Wall Region

1	63.0	65.5	27.1	26.25	26.19
2	63.2	65.5	27.55	28.6	27.19
3	63.5	66.1	28.2	28.3	27.2
4	63.8	66.1	30.7	29.25	29.09
5	64.4	66.1	31.5	31.7	32.1
Center	1	2	3	4	5

y+ = 32

Wall Region

1	64.0	66.6	69.8	71.8	21.4
2	64.0	66.4	70.4	71.5	22.4
3	64.9	67.5	70.9	66.1	24.0
4	66.0	68.5	66.7	65.9	28.0
5	67.4	59.2	70.8	65.1	33.9
Center	1	2	3	4	5

y+ = 48

Wall Region

0	32.5	26.03	26	22	20.5	20.2
1	27.5	26.2	26.1	22.9	21.5	20.6
2	30.7	35.9	26.7	23.6	21.9	21.1
3	46.6	54	36.3	27.5	23.6	22.8
4	52	56	56.3	29.7	27.7	25.1
5	57	55	35.7	35.6	31.6	29.1
Center	1	2	3	4	5	6

Figure 3.3: The mean centerline x-velocities are shown for the original 75 cases plus the extra 11 cases (light green) for the wall region of height $y^+ = 48$. The rows correspond to C_ε - C_{μ_ε} pairs in the wall region and the columns represent the pairs in the center region. The values for the pairs can be found in Table 3.1. If a square is red, the simulation had laminar flow, if it is green then the flow turned turbulent.

Figure 3.4 shows the spatial mean of the velocity taken at the centerline plane. An averaging window of 40 time units was used to calculate a running time average, and the error bars represent one standard deviation for that window. The centerline velocity is the value used in the calibration. We will also see how well the calibration works on other QoI. Figure 3.5 shows the running average of the mass flux, Figure 3.6 the maximum RMS value for the flow, and Figure 3.7 the height of that maximum RMS. The mass flux follows the trends of the mean centerline velocity somewhat closely. However the maximum RMS trends are unique. All of the turbulent flow cases have a higher maximum RMS value compared to the DNS value, and the parameter set which produces the best results for the centerline velocity are not the most accurate at predicting the maximum RMS value.

Figure 3.8 compares the velocity profile for the Moser et al. DNS simulation with two cases with a wall region of $y^+ = 48$. The blue line matches the centerline velocity the best and has pair 6 from Table 3.1 for the center and pair 2 for the wall region. The green line does very well near the wall and has pair 6 for the center and pair 5 for the wall. This figure indicates that it is difficult

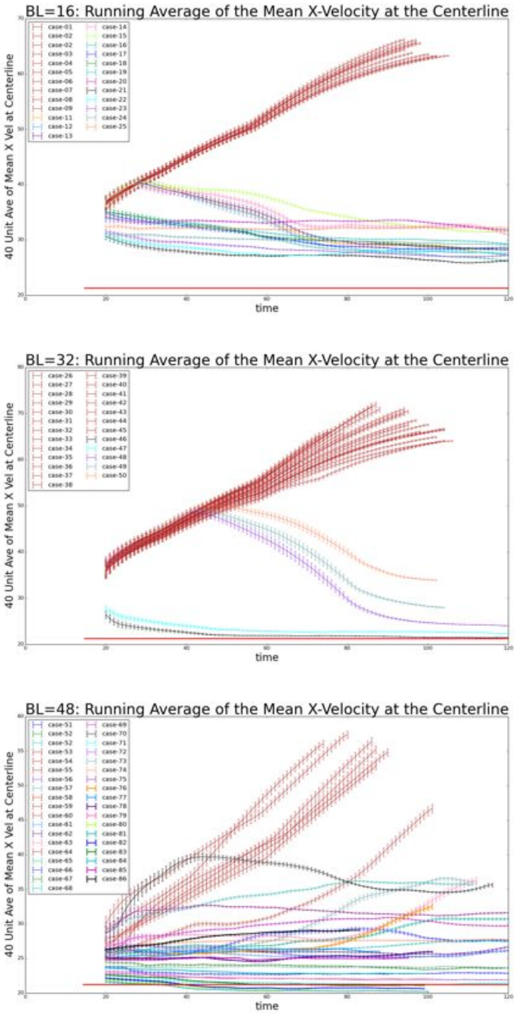


Figure 3.4: The mean x-velocity at the centerline plane was averaged in time over a window of 40 time units. The running average is plotted here along with the standard deviation in that average block. The cases that correspond to laminar flow are shown in dark red. The time averaged DNS value of 21.26 is shown as a red line.

to find a single parameter pair that would match the DNS data at both the wall and the centerline. The stress balance can be represented by the equation for the resolved convective stress, $\bar{u}v$ which is written as:

$$\bar{u}v = \frac{dP}{dx}y + \left(\frac{1}{Re} + \nu_y\right)\frac{d\bar{u}}{dy} \tag{3.6}$$

Where $1/Re(d\bar{u}/dy)$ is the viscous stress and $\nu_t(d\bar{u}/dy)$ is the modeled stress. For the case that matches the DNS centerline velocity (blue line in Figure reffig:velprof) the velocities near the wall are higher than the DNS near-wall velocities. One reason this might be is that the stresses for this case are lower than the stresses for the case that do match the DNS data near the wall (green line

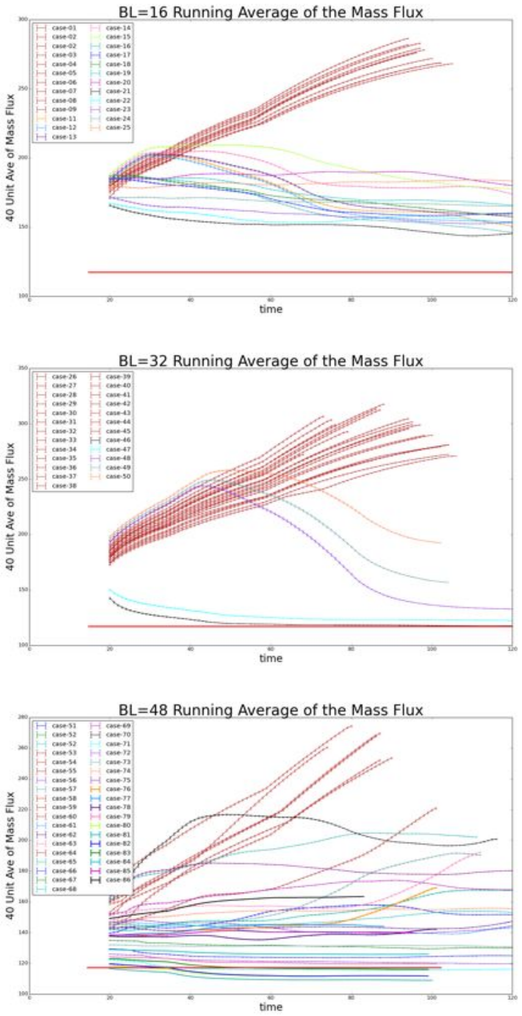


Figure 3.5: The mass flux was averaged in time over a window of 40 time units. The running average is plotted here along with the standard deviation in that average block. The cases that correspond to laminar flow are shown in dark red. The time averaged DNS value of 117 is shown as a red line.

in Figure 3.8). The stresses can be compared in Figure 3.9, which shows that overproduction of stress is mostly coming from the modeled stress. The true stress balance is not exactly recovered because we do not have access to the stress as it enters the equations in the code.

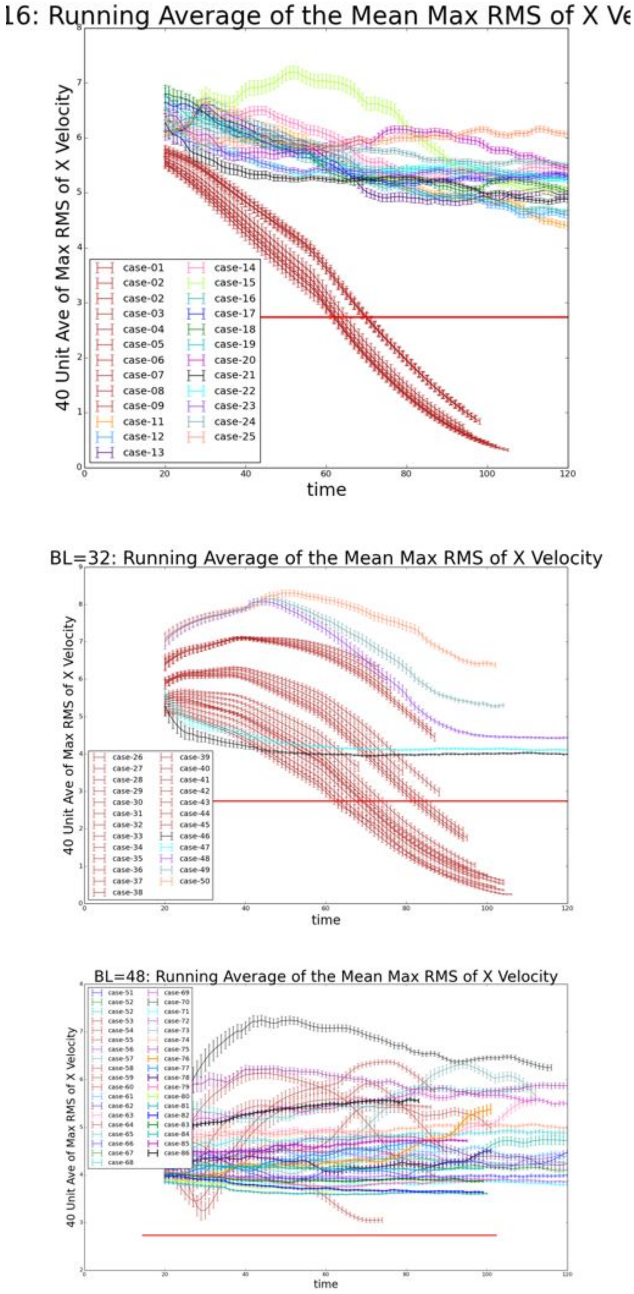


Figure 3.6: The maximum RMS value was averaged in time over a window of 40 time units. The running average is plotted here along with the standard deviation in that average block. The cases that correspond to laminar flow are shown in dark red. The time averaged DNS value of 2.74 is shown as a red line.

3.4 UQ with engineering level LES

3.4.1 Bayesian Model Calibration

We are using a Bayesian approach for the uncertainty quantification in which the posterior distribution, $P(\theta|D)$, can be expressed with Bayesformula:

$$p(\theta|\mathscr{D}) = L_{\mathscr{D}}(\theta)p(\theta)/p(\mathscr{D}) \tag{3.7}$$

Here, $p(\theta)$ and $p(\theta|\mathscr{D})$ are the prior and posterior distributions for the model parameters θ. These distributions represent knowledge about these parameters before and after learning from the dataset D, For the present study, D is based on the DNS channel flow data set [46]. The model parameters are the pairs of k^{sgs} model constants: $\theta = \{\eta_{wall}(C_\varepsilon, C_{\mu_\varepsilon})\}$, $\eta_{center}(C_\varepsilon, C_{\mu_\varepsilon})$, where η_{wall} is the set of $(C_\varepsilon, C_{\mu_\varepsilon})$ in the near wall region and η_{center} is the set in the center region. The prior distribution, $P(\theta)$, is the belief of what the parameters θ should be, and is a uniform prior in this case. The model evidence, $P(D)$, is effectively a normalizing factor and is not relevant for this study. The likelihood $P(D|\theta)$ is the probability of observing D given a particular instance of θ. We propose the following discrepancy model between the data and the LES calculations.

$$d = M(\theta) + \varepsilon = M(\theta) + N(0, \sigma) \tag{3.8}$$

Here, the model M represents Quantities of Interest (QoI) based on LES simulations corresponding to the two model parameters η_{wall} and η_{center}. The discrepancy is presumed to follow a zero-mean Gaussian distribution $\varepsilon(0, \sigma)$, and accounts for the fact that the LES computations are modeling small turbulent scales and are less accurate compared to DNS. Since data D is the result of deterministic DNS computations, in this study we neglect discrepancies due to noise in the data. With these assumptions, the likelihood is given by

$$p(\mathscr{D}|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(V_D\bar{N}S - M(\eta_{wall}, eta_{center}))^2}{2\sigma^2}\right) \tag{3.9}$$

Here, the data D consists of $V_D\bar{N}S$ which is the time and spatial averaged velocity at the centerline plane of the DNS data set and σ is standard deviation of the model error term epsilon. A Markov Chain Monte Carlo (MCMC) algorithm is used to sample from the posterior probability density, $P(D\theta)$ in Eq. (4). MCMC allows sampling from a probability density by constructing a Markov Chain that has the target density as its stationary distribution [21]. In particular, we employ an adaptive Metropolis algorithm [23], which uses the covariance of the previously visited chain states to explore posterior distributions in an efficient manner. We employ several diagnostics [58, 32] to determine when the MCMC samples converge to stationary posterior distributions. Since the MCMC approach requires a large number of samples to produce converged posterior distributions, we rely on surrogate models to replace the expensive LES computations. These surrogate models are constructed based on LES simulations corresponding to select sample values for $\theta = \eta_{wall}, \eta_{center}$, shown in Table 3.1. Originally these samples were optimally placed to construct surrogate models based on Polynomial Chaos concepts [19, 65]. As some of the model runs failed to produce valid turbulent results, resulting in an unstructured valid region, we turned our attention to Radial Basis Functions (RBFs) [6] for the construction of surrogate models for the Quantities of Interest resulted from LES computations. Specifically, we employed a Multiquadric Radial Basis Function (RBF) for the surrogate model:

$$M(\theta) \approx \sum_{i=1}^{N} \omega_i \sqrt{1+(\varepsilon||\theta-\theta_i||)^2} \tag{3.10}$$

where ε is a smoothing parameter and ω_i are weights corresponding to each RBF. Here θ_i are “valid" parameters samples, see Table 3.1. Optimal values for the smoothing parameter are determined via a cross-validation procedure by leaving out one sample at a time. The weights are computed by regression. Slices through the RBF surrogate model for the wall model cases that have a height of $y^+ = 48$ are shown in Figure 3.10 for the mean centerline velocity. RBF-based surrogate models for the mean mass flux and the mean maximum RMS of the x-velocity are also presented in Figure 3.11 and Figure 3.12. We found that, for the mean centerline velocity and the mass flux epsilon $\varepsilon = 0.001$ was an optimum value for providing a good approximation to the data. For the results shown in Figure 12, the optimum value for this parameter is approximately 0.0005.

3.4.2 Posterior Densities

Figure 3.13 shows the posterior pdf $P(\theta|D)$ for the two sets of model parameters, η_{wall} and η_{center}. These densities are constructed via Kernel Density Estimate [70] using 10^6 MCMC samples. We observed a region of Maximum A Posteriori (MAP) values shown with dark red in Figure 3.13. The values from literature for the parameter sets include $(C_{\mu_\varepsilon}, C_\varepsilon) = (0.0845, 0.85)$ for isotropic turbulence [87] and $(C_{\mu_\varepsilon}, C_\varepsilon) = (0.07, 1.05)$ for shear flow [69]. The corresponding range of values corresponding to the MAP region observed in Figure 3.13 corresponds to $(C_{\mu_\varepsilon}, C_\varepsilon)$ from (0.086, 1.65) to (0.109, 2.04) in the wall region and $(C_{\mu_\varepsilon}, C_\varepsilon)$ from (0.016, 0.35) to (0.024, 0.48) in the center region. This indicates that there is more production and dissipation in wall region and also higher stresses.

3.4.3 Forward UQ

The densities of the model parameters $\theta = \{\eta_{wall}(C_\varepsilon, C_{\mu_\varepsilon}), \eta_{center}(C_\varepsilon, C_{\mu_\varepsilon})\}$ are pushed forward to obtain the probability densities of other quantities of interest, including mass flux, maximum RMS, and the velocity profile. We first start with a posterior predictive check [39] for the mean centerline velocity. This test proceeds with the marginalization of the likelihood $P(V_c|D)$ over the posterior distribution of model parameters $P(\theta|D)$:

$$p(V_c|D) = \int p(V_c|\theta)p(\theta|D)d\theta \tag{3.11}$$

This test indicates the adequacy of the calibrated parameters and the employed noise model for the centerline velocity, V_c. The left frame in Figure 3.14 shows the posterior predictive pdf obtained based on the above expression. The experimental value, shown with the red circle, is in the vicinity of the MAP value confirming the adequacy of calibrated model for this parameter. For other QoIs we inspect push-forward posterior distributions. These are obtained via an approach similar to the one presented in Equation 3.5, with the exception of the noise model who is only available for the data employed in the calibration. The center and left frames in Figure 3.14 presents results for the mean mass flux and the maximum velocity rms. While the experimental value for the mass flux exhibits a reasonable probability density based on the pushed forward posterior for this QoI, the experimental value for the velocity rms is well outside the range of values predicted by the model This again emphasizes that care must be taken when calculation forward UQ estimates for QoI which are not closely related to the QoI on which the calibration is completed.

3.4.4 Forward Propagated PDF for Velocity Profile

A push forward UQ was also completed for the velocity profile. RBF models for the spatially averaged axial velocities were computed for grid locations across the channel cross-section. We then employed these models to push-forward the posterior density for the model parameters and obtain 1D marginal densities at each location along the profile. Summaries of these densities are shown in Figure 3.16. These results indicate a relatively narrow variation around the mean profile, with 1D marginal standard deviations between 4% near the wall and 2% near the middle section.

3.4.5 Forward Propagation for Channel Flow with Different Re_τ

In order to check the calibration of the model parameters at Reynolds numbers other than the one used for the calibration, channel flow cases at $Re_\tau = 395$ and 950 were also run. Both cases were centered around the best cases at the calibration $Re_\tau = 590$. For the lower Reynolds number, half of the flows did not stay turbulent long enough to compute statistical data on the flow, and the ones that did all had mean centerline velocities that are higher than the DNS value of 20.13 [46]. For the higher Reynolds number, all 16 flows were turbulent, but again, the mean centerline velocity was higher than the DNS value of 22.44 [29]. It would most likely be possible to find the correct combination of parameter values for this Reynolds number if there were more time fr this study.

3.4.6 Forward Propagation for Backwards Facing Step

3.5 Conclusions

Engineering LES can achieve correct quantities of interest for channel flow if the turbulence model parameters are tuned correctly. This can be achieved more easily if wall region and center region are tuned separately. We have found that the height of the wall region matters and that the best choice is to use a height that is within the log layer. Bayesian calibration and surrogate models have allowed us to explore the region of parameter space that leaded to the correct center line velocity values. The optimal parameter values for this channel flow case in the center region are lower than values found in literature, and wall region optimal values are higher. There is a sensitivity of the results to the chosen QoI. Ones that are physically similar to the output that is used for calibration are much more accurate than ones that are less similar. In this example we used the centerline velocity for calibration, and the mass flux was well predicted, while the predicted maximum RMS value was much less accurate.

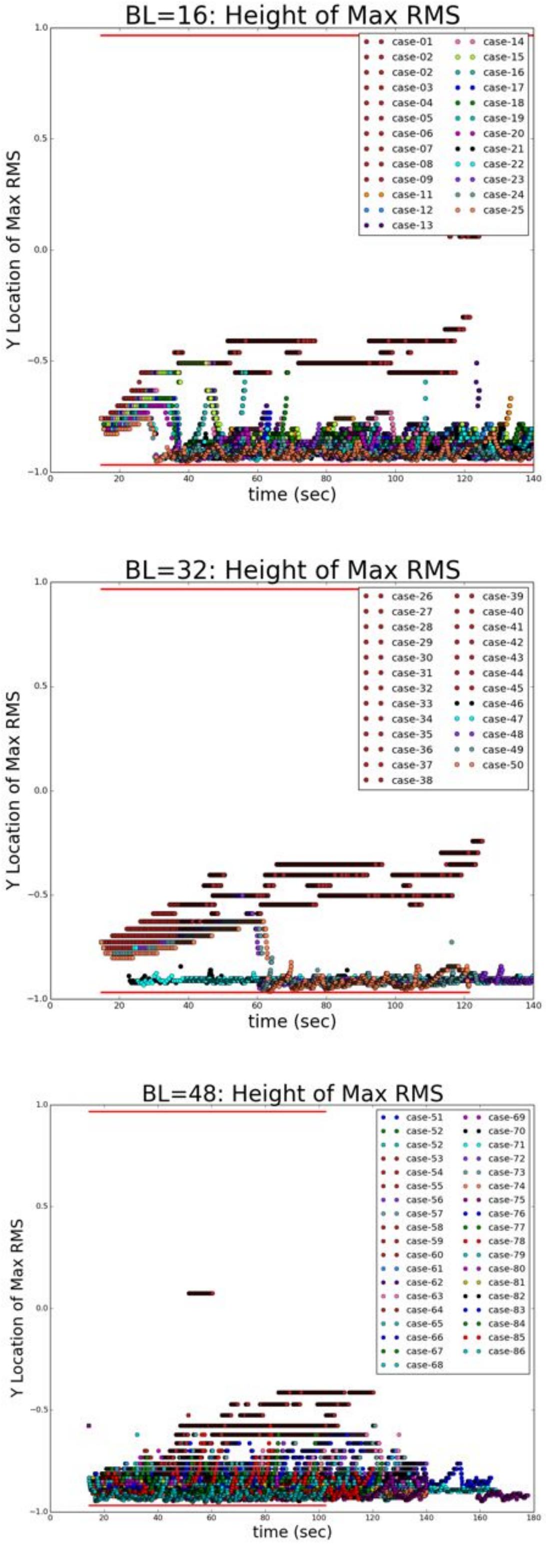


Figure 3.7: The height of the maximum RMS value is shown with respect to the DNS value (y+ = 19.39). The cases that correspond to laminar flow are shown in dark red.

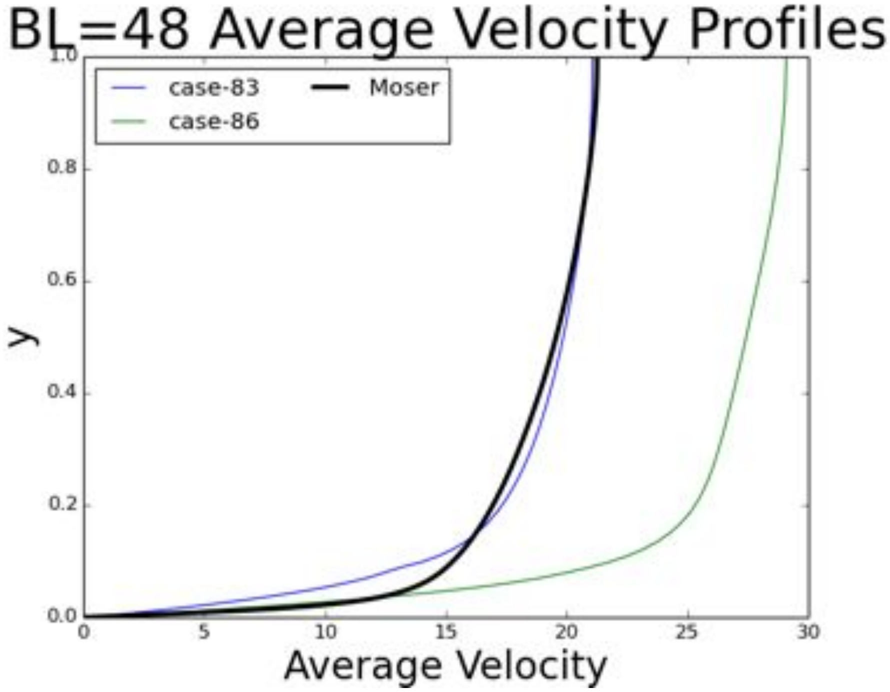


Figure 3.8: Velocity profiles for two cases compared to DNS data (black line). The blue line shows the case which has the mean centerline x-velocity closest to the DNS data. The green line is an example of a case which matches the DNS profile well in the wall region.

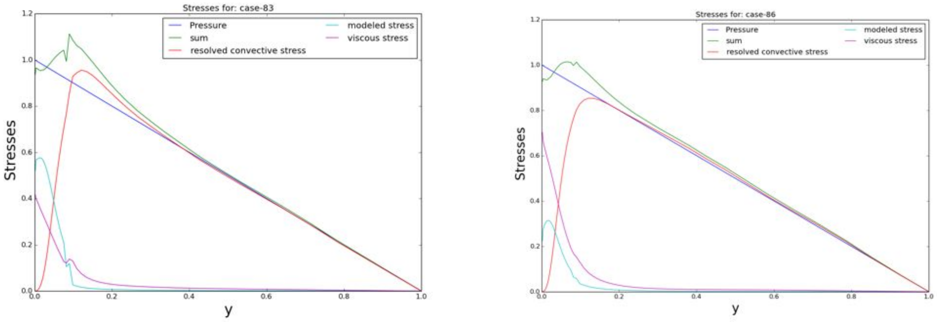


Figure 3.9: Stress balances for the two cases shown in Figure 3.8.

Figure 3.10: Comparison of Radial Basis Function surrogate, with blue lines, with data from LES runs for the mean centerline velocity, shown with green circles.

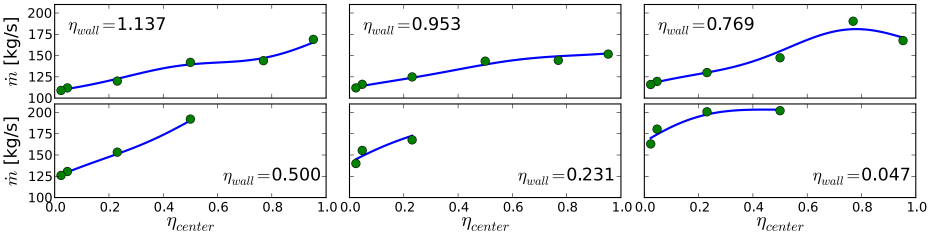


Figure 3.11: Comparison of Radial Basis Function surrogate, with blue lines, with data from LES runs for the average mass flux, shown with green circles.

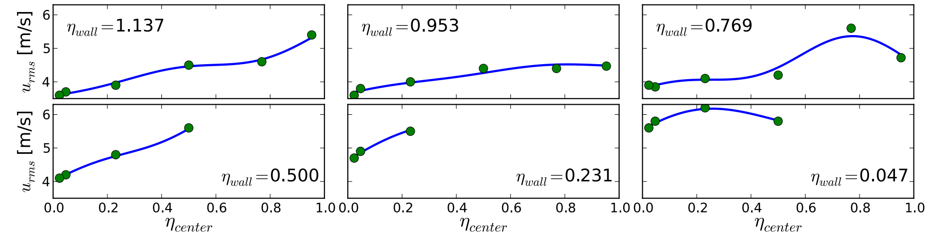


Figure 3.12: Comparison of Radial Basis Function surrogate, with blue lines, with data from LES runs for the mean RMS of the X-velocity, shown with green circles.

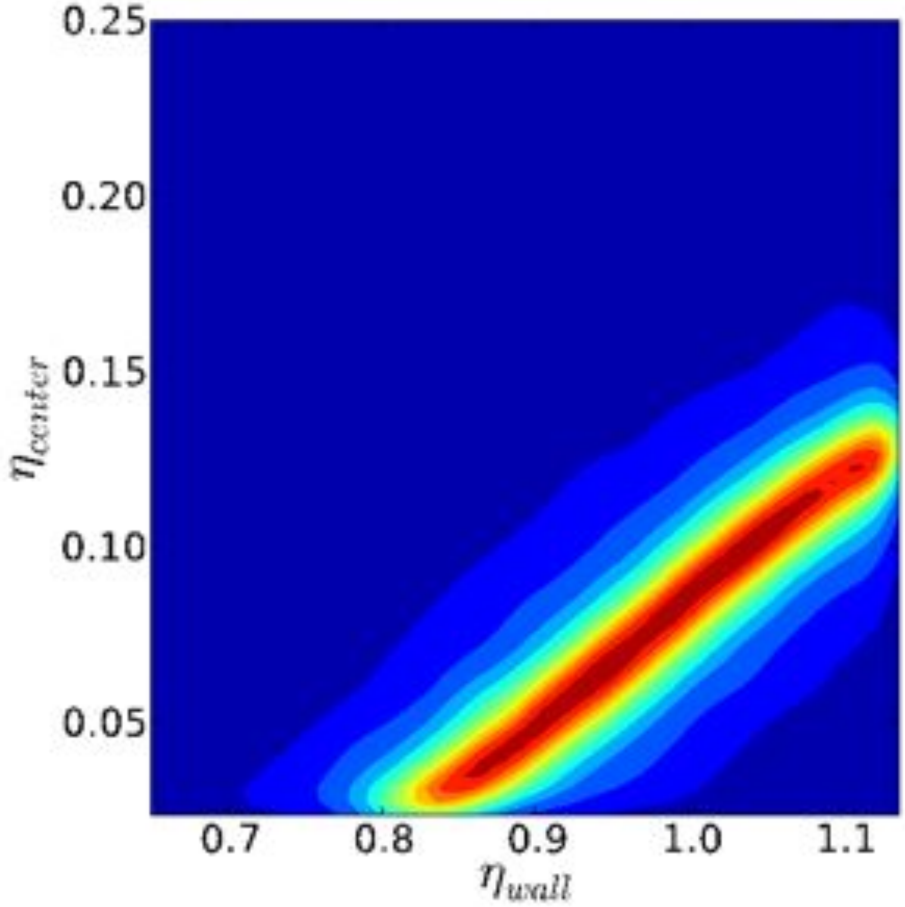


Figure 3.13: Joint PDF. Red indicates that the model produces a mean centerline velocity which is close to the DNS value while blue indicates more discrepancy.

Figure 3.14: Posterior predictive distribution for the mean centerline velocity V_c, and pushed-forward posterior distributions for the mean mass flux and maximum RMS of X-velocity. The red symbols correspond to the DNS data: 21.26, 117.2, and 2.74, respectively.

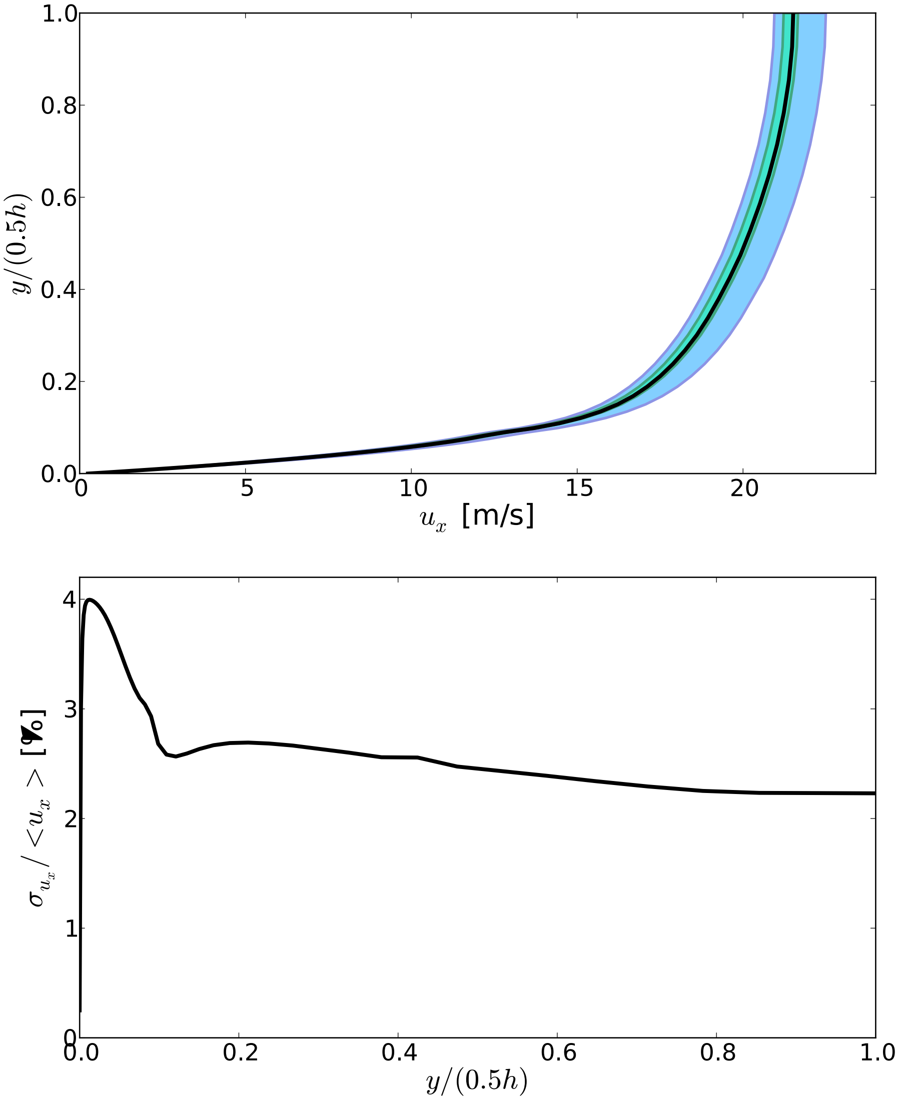
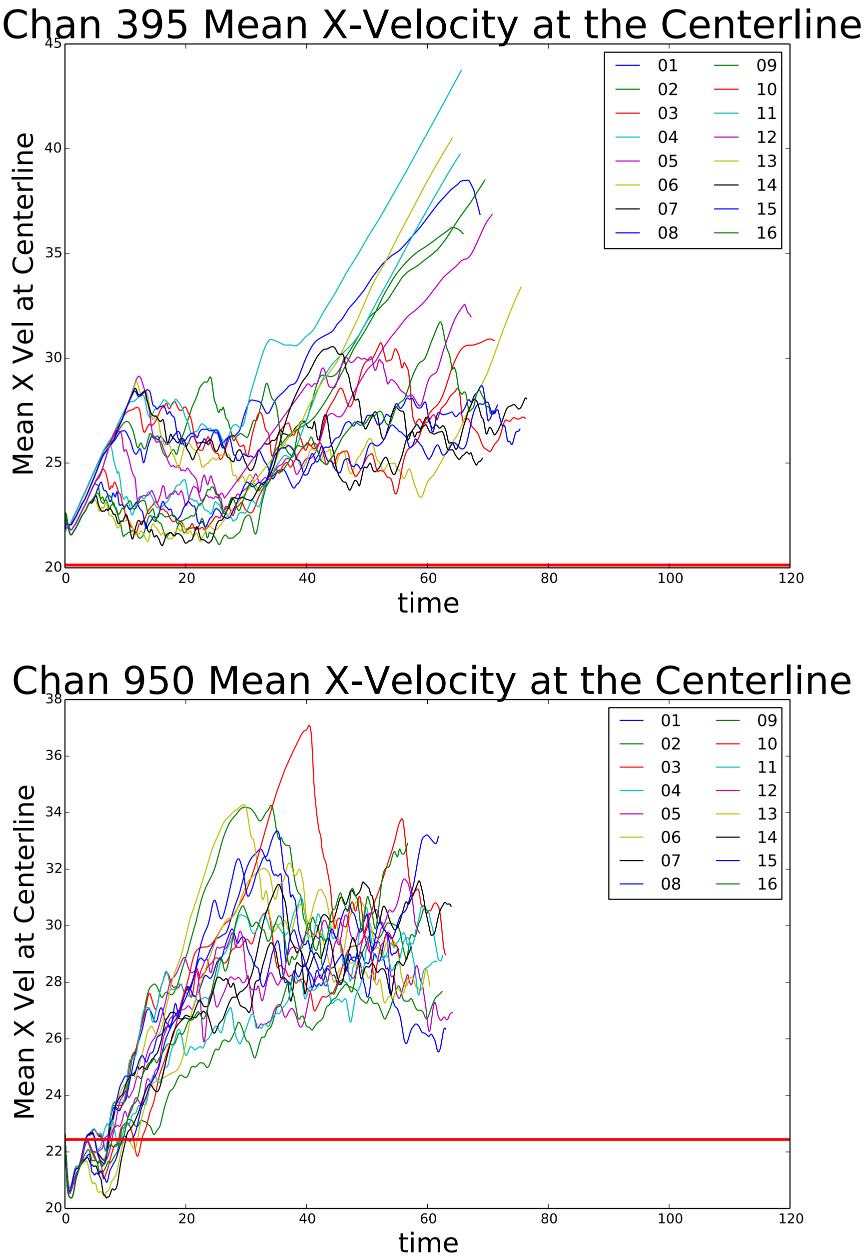


Figure 3.15: 1D Marginal pushed-forward posterior densities for the axis velocity profiles. The blue region in the left frame corresponds to 5-95% quantile range while the green corresponds to 25-75% range. The black line shows the expected u_x profile. The right plot shows the relative magnitude of the standard deviation of the 1D marginal pushed-forward posterior density.


Chan 395 Mean X-Velocity at the Centerline

Chan 950 Mean X-Velocity at the Centerline

Figure 3.16: This figure shows the mean centerline velocities for 16 cases each at $Re_\tau = 395$ and 950.

Chapter 4

Numerical Parameter Sensitivities of Open Jet Flow

4.1 Introduction

Turbulent flows are of great scientific and technological areas, ranging from astrophysical phenomena to internal combustion engines. While observation and experiment are vital tools to investigate these flows, simulation is needed in order to fully explore the various regimes in which turbulent flows occur and predict their effects for engineering applications. Unfortunately, turbulence has proven to be difficult to examine computationally. Much understanding of the fundamental physics has been learned from direct numerical simulations (DNS) of turbulent flows. These simulations use high-order numerical methods and resolve all of the energy containing scales to minimize the effect of numerical errors on the results. However, the cost of DNS exceeds the ability of many modern computers to investigate important flows [11, 2].

In lieu of resolving all the dynamically significant flow features, other approaches to computing the solutions of turbulent flows rely on resolving either none of the fluctuations, as in Reynolds Averaged Navier Stokes (RANS), or only the large scale fluctuations, as in large-eddy simulation (LES). RANS is the primary tool for engineering simulations because of its relatively low cost and well-quantified numerical errors. In particular, mesh convergence can be used to minimize the error associated with discretization, while the errors arising from the numerical methods are less significant because of the stable nature of the solution.

While RANS has these advantageous properties, its reliance on modeling all the turbulent fluctuations present in a flow can limit its predictive capabilities in some important situations such as separated flows. By resolving larger motions, LES relies less on models and has been successful at simulating scientifically interesting flows with high-fidelity. However, its utility to engineering applications has been limited because coarse, unstructured meshes and associated numerical methods introduce significant errors.

Quantifying numerical errors falls under the purview of solution verification within the Verification and Validation (V&V) process. These studies identify the key numerical parameters influencing solution fidelity and quantify the variation they induce. The most commonly investigated source of numerical errors are those associated with finite mesh resolution. For LES, numerical methods have been reported to also significantly effect the resulting solution. In this work, we consider the effect of numerical parameters on LES solutions using the software package Nalu. Nalu is an unstructured, low-Mach number solver with both RANS and LES models, and as such is an exemplar of the types of codes which will be used for engineering applications.

This chapter is organized as follows. In Section 4.2 we provide an overview of the numerical methods which will be investigated, as well as the sampling strategy and how the numerical discretization schemes are reformulated to be amenable to statistical analysis. The results of the solution verification study are presented in Section 4.3, and some conclusions are offered in Section 4.4.

4.2 Numerical Methods

In this work we use the CFD software Nalu simulating an open jet as the example for the numerical parameter study. Nalu is an unstructured, low-Mach number code which has multiple numerical methods implemented in it, making it ideal for the present study. It contains both edge and element based schemes, as well as upwinding and conservative numerical methods. For the purposes of this work, two studies will be performed and compared. The first will consider a mix of discretization schemes, while the other will examine a combination of the edge and element methods with different upwinding methods. We now present the features of the code which enable both studies.

As most of the numerical parameters in CFD codes alter how advection operators are constructed in terms of stability and accuracy, the numerical parameters used in the LHS study will span the space of various operators. Integration over a control volume face is descretized as a summation over a set of integration points, denoted by index j, as

$$\int_{CV} \rho \phi \mathbf{u} \cdot \mathbf{n} dA \approx \sum_j \rho u_j \phi_j A_j, \tag{4.1}$$

for an advected field ϕ. The choice of the method to approximate ϕ_j forms the basis of different CFD algorithms.

Nalu supports a blending of projected nodal gradients and upwinding to enable psuedo-higher order schemes using the following general approximation

$$\phi_j = \eta \phi_{upw} + (1-\eta)(\beta \phi_{cds} + (1-\beta)\phi_{4th}). \tag{4.2}$$

The parameter η can be used to select between upwind ($\eta = 1$) and central approximations ($\eta = 0$). Its value is selected using a function of the cell Peclet number, which for a given face is defined based on the adjacent nodes:

$$\mathrm{PE}_j = \frac{1}{2} \frac{(u_j^R + u_j^L)(x_j^R - x_j^L)}{\nu}, \tag{4.3}$$

where the R and L denote the face's left and right nodes, and ν is the viscosity. The upwind blending is computed using the function

$$\eta_j = \frac{\mathrm{PE}_j^2}{h + \mathrm{PE}_j^2}. \tag{4.4}$$

Denoted the hybrid upwind factor, h controls how quickly this blending occurs from $\eta = 0$ when $\mathrm{PE}_j = 0$ to $\eta = 1$ as $\mathrm{PE}_j \to \infty$.

Nalu provides additional control over the formulation of the upwind operator itself via blending between the extrapolated and interpolated state:

$$\phi_{upw} = \begin{cases} \alpha_{upw}\tilde{\phi}^L_{upw} + (1-\alpha_{upw})\phi_{cds}, & u_j > 0 \\ \alpha_{upw}\tilde{\phi}^R_{upw} + (1-\alpha_{upw})\phi_{cds}, & u_j < 0 \end{cases} \tag{4.5}$$

Extrapolated states are provided by

$$\tilde{\phi}^L_{upw} = \phi^L + d^L_j \frac{\partial \phi^L}{\partial x_j} \tag{4.6}$$

$$\tilde{\phi}^R_{upw} = \phi^R - d^R_j \frac{\partial \phi^R}{\partial x_j}, \tag{4.7}$$

with d^L_j and d^R_j are the distances between the integration point and corresponding node. For the interpolated state, ϕ_{cds}, a standard central difference scheme is used based on a linear combination of the values at the adjacent nodes similar to a finite element interpolation.

A third numerical parameter is introduced which blends the upwind and central difference operators to create a psuedo fourth order scheme:

$$\phi_{4th} = \frac{1}{2}\left(\hat{\phi}^L_{upw} + \hat{\phi}^R_{upw}\right). \tag{4.8}$$

Left and right methods are defined as

$$\hat{\phi}^L_{upw} = \alpha\tilde{\phi}^L + (1-\alpha)\phi_{cds} \tag{4.9}$$

$$\hat{\phi}^R_{upw} = \alpha\tilde{\phi}^R + (1-\alpha)\phi_{cds}. \tag{4.10}$$

Blending the upwind and central difference contributions is handled using the parameter α. Given these three numerical parameters, α_{upw}, h, and α can be combined to create a continuum of discretizations.

To determine how these parameters interact, we can combine Eqs. (4.5), (4.6), (4.7), and (4.8) into Eq. (4.2) to obtain

$$\phi_j = \eta\left[\alpha_{upw}\tilde{\phi}^U_{upw} + (1-\alpha_{upw})\phi_{cds}\right] + (1-\eta)\left[\alpha(1-\beta)\frac{\tilde{\phi}^L_{upw} + \tilde{\phi}^R_{upw}}{2} + \left(\frac{1}{2}(1-\alpha)(1-\beta) + \beta\right)\phi_{cds}\right]. \tag{4.11}$$

To simplify the notation, $\tilde{\phi}^U_{upw}$ is used to denote the upwind contribution in Eq. (4.6) or (4.7) as appropriate based on the velocity. By collecting terms, the consolidated equation for the field at the integration point is

$$\phi_j = \eta \alpha_{upw} \tilde{\phi}^U_{upw} + \alpha(1-\eta)(1-\beta) \frac{\tilde{\phi}^L_{upw} + \tilde{\phi}^R_{upw}}{2} + \left(\eta(1-\alpha_{upw}) + \frac{1-\eta}{2} (1 + \alpha\beta - \alpha + \beta) \right) \phi_{cds}, \tag{4.12}$$

with further consolidation of terms leading to

$$\begin{gathered} \phi_j = \eta \alpha_{upw} \tilde{\phi}^U_{upw} + (1-\eta)(\alpha - \alpha\beta) \frac{\tilde{\phi}^L_{upw} + \tilde{\phi}^R_{upw}}{2} + \\ \left(\eta - \eta \alpha_{upw} + \frac{1}{2} \left((1-\eta)(\alpha\beta - \alpha) + (1-\eta)(1+\beta) \right) \right) \phi_{cds}. \end{gathered} \tag{4.13}$$

From this we can see that there are three contributing parameters:

$$P_1 = \eta \alpha_{upw} \tag{4.14}$$

$$P_2 = \frac{\alpha(1-\eta)(1-\beta)}{2} \tag{4.15}$$

$$P_3 = \eta + \frac{(1-\eta)(1+\beta)}{2} \tag{4.16}$$

such that Eq. (4.13) can be expressed as

$$\phi_j = P_1 \tilde{\phi}^U_{upw} + P_2 \left(\tilde{\phi}^L_{upw} + \tilde{\phi}^R_{upw} \right) + (P_2 + P_3 - P_1) \phi_{cds}. \tag{4.17}$$

As a result, we need only vary three parameters rather than four to identify the sensitivities, which will be α, α_{upw}, and h. The presence of this structure in the numerical method allows for a robust examination of the effect of different stencils on the CFD results.

In LES the numerical method and turbulence models can interact and collectively effect the simulation results. In order to assess this effect, the turbulence model will also be varied along with the numerical parameters. The WALE model [48] will be used in this work which contains a single model constant which can be varied. The formula for the turbulent viscosity is:

$$\mu_t = \rho (C_w \Delta)^2 \frac{(S^d_{ij} S^d_{ij})^{3/2}}{(S_{ij} S_{ij})^{5/2} + (S^d_{ij} S^d_{ij})^{5/4}}. \tag{4.18}$$

The model constant, C_w is typically reported to have values between 0.325 and 0.55, and the filter width Δ is chosen as the cube root of the element's volume. The standard strain rate tensor, S_{ij} appears in the formula, as do tensors based on the velocity gradient squared:

$$S^d_{ij} = \frac{1}{2}(g^2_{ij} + g^2_{ji}), \quad g^2_{ij} = \frac{\partial u_i}{\partial x_k} \frac{\partial u_k}{\partial x_j}, \quad g^2_{ji} = \frac{\partial u_j}{\partial x_k} \frac{\partial u_k}{\partial x_i}. \tag{4.19}$$

Summation is implied over repeated indices. In this work, the model constant is varied over the range 0 to 1.

4.3 Results

4.3.1 Sampling-based Study

The first verification study was accomplished by performing a 50 evaluation Latin hypercube sample (LHS) over the three numerical parameters and turbulence model constant. Each dimension was sampled uniformly, and the overall number of samples was 60 due to the computational budget. Assessing the variability of the simulation results was done by considering the first and second moments of the velocity, as well as the pressure, at 5 heights uniformly from 20% of the domain to 100%. At each plane, the velocity was decomposed into streamwise (normal), radial, and tangential moments and averaged radially . This strategy was chosen because it enabled us to observe the most important trends in the jet's evolution.

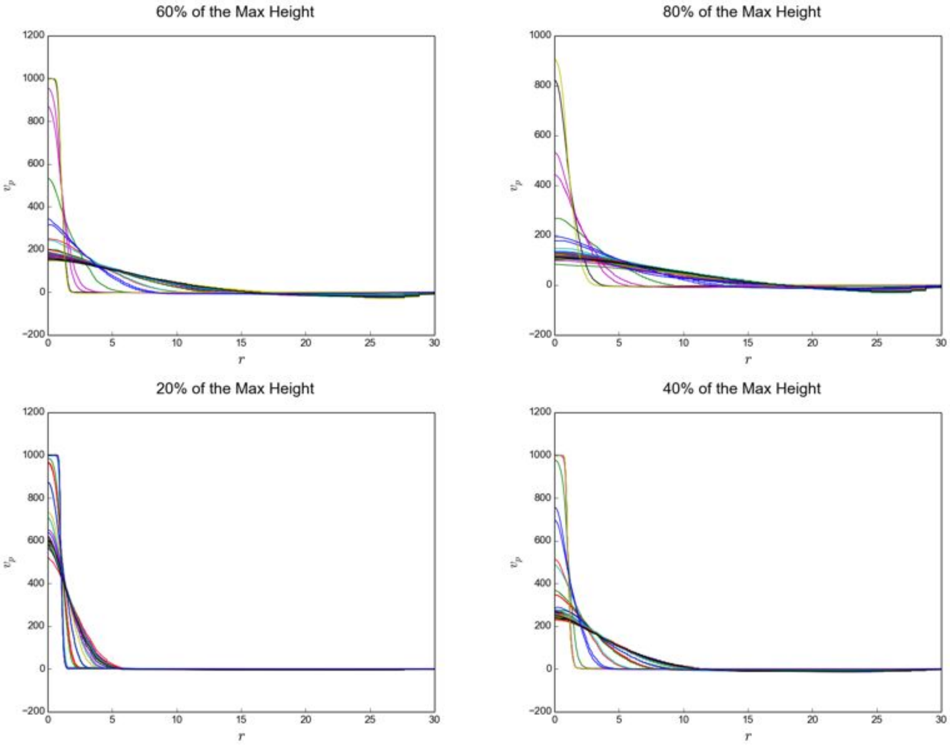


Figure 4.1: Streamwise velocity profiles at four stations away from the inlet.

Figure 4.1 shows the streamwise velocity at the 20% through 80% heights arising from each simulation. We observe that there is a cluster of profiles within the ensemble, along with a few outliers of significantly higher velocities at the centerline. To determine which parameters drive this behavior, the centerline streamwise velocities are plotted against each parameter in Figure 4.2. Correlation coefficients for each are also shown, the largest of which by a great margin is that with respect to α_{upw}. While the correlation coefficient measures the strength of the linear relationship between the parameter and the output, visual inspection reveals that while significant, the relationship is not linear. There is a rapid shift upwards in the centerline velocity starting at $\alpha_{upw} \approx 0.8$.

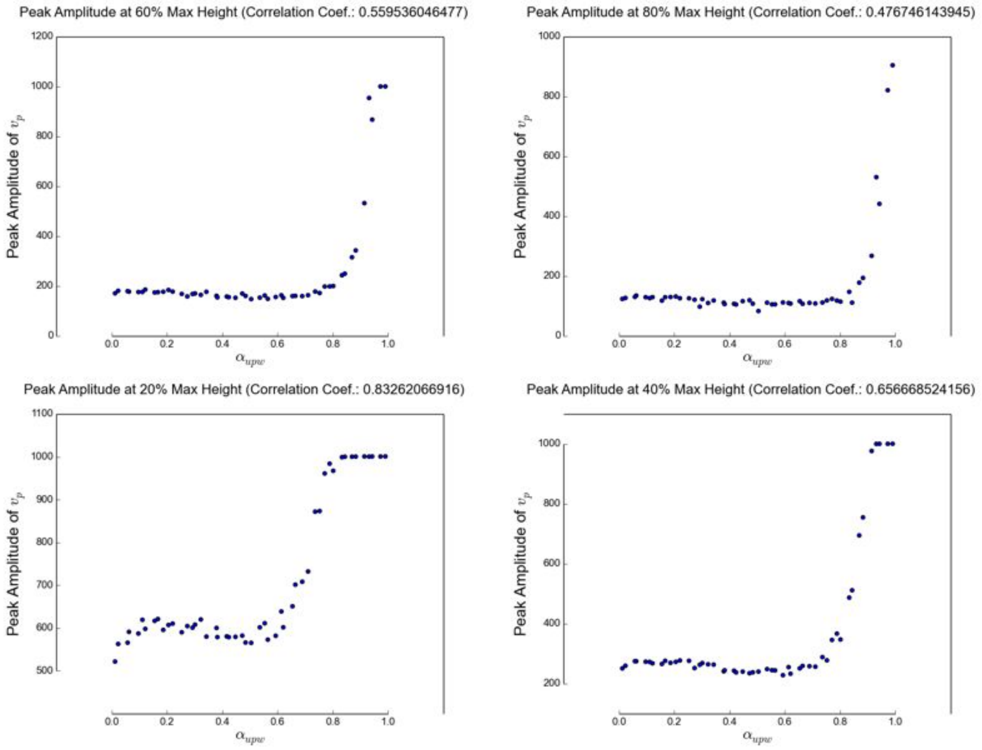


Figure 4.2: Streamwise velocity profiles at four stations away from the inlet.

While the visual data presented in the figures illustrates the phenomena, a more systematic analysis can rest upon the correlation coefficient. For each profile for all output variables, the amplitude of the first peak (or center value, if none existed) as well as its location was correlated with each input parameter. The only variable exceeding a correlation value of 0.5 for any of the quantities of interest was α_{upw}. Consistently, its effects were only present in the upper portions of its domain, ranging from 0.83 at 20% height to 0.48 at 100% height. In some cases, the correlation coefficient was quite low but the relationship was still strong but characterized by an upward then downward trend in the relationship. However, neither statistical nor visual data was obtained suggesting a relationship between any of the other parameters and the quantities of interest.

Given the results, the primary conclusion is clear: excessive damping from using upwind operators delays the onset of jet core collapse primarily through retarding the instability leading to it. Parameters affecting the accuracy of the numerical method have no statistically significant effect. Most interestingly, even the choice the turbulence model constant does not greatly impact the resulting velocity fields even though it is also dissipative in nature. A recommendation can be made based on this study that upwinding be used only to limited amounts. While well-known in the LES modeling community, these results illustrate that, at least for some cases, some upwinding can be used to either stabilize the numerics or accelerate convergence without an adverse affect on the velocity field. This result therefore enables practitioners to reduce the time to solution with only a small impact on its accuracy.

4.4 Conclusions

The primary conclusion of this study is that upwinding can have a large affect on simulation results. While previously known, these results demonstrate key quantities of interest can be incorrectly predicted by over 100%. What was discovered in this study, however, is that the variation of the overprediction is a highly non-linear function of the amount of upwinding; delineation between little to no effect and a large effect occurs over a change in the amount of upwinding of approximately 5%. Another discovery is that non of the other numerical parameters, including those which changed the order of accuracy of the method, had any statistically significant effect. More surprising was that the WALE model coefficient likewise had no statistically significant impact on the flow. We can conclude that, for this flow configuration at least, that keeping the amount of numerical dissipation below a critical threshold is the most important thing that can be done to limit numerical errors.

Chapter 5

A posteriori analysis of the WALE model for Jet-in-Crossflow

In this chapter we explore the correlations between several flow markers in LES of turbulent flow and parameters of select sub-grid models. Specifically, we study correlations between several non-dimensional parameter derived from high-resolution LES simulation and the WALE model constant C_w [48]. For this purpose we focus on the jet-in-crossflow configuration based on the experiment conducted by Su and Mungal [74]. We employ the high-resolution LES simulation data provided by Ruiz *et al* [64]. Figure 5.1 depicts the jet-in-crossflow configuration, including the configuration of the coordinate axes.

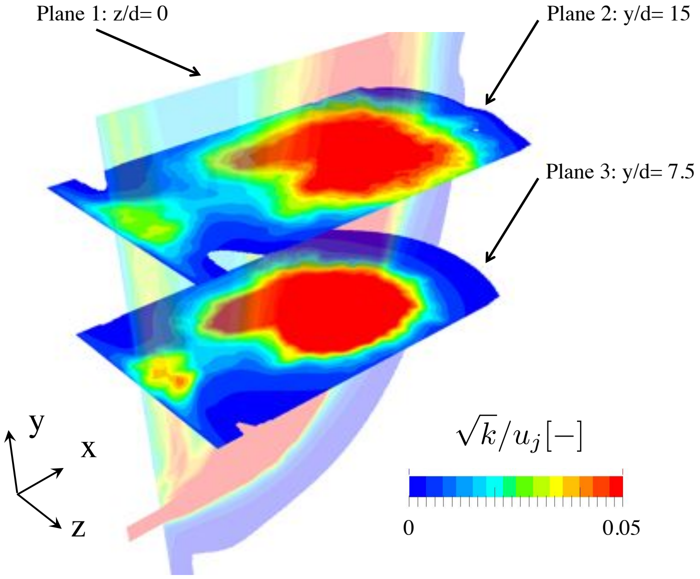


Figure 5.1: Configuration of computational domain for the jet-in-crossflow configuration.

In Section 5.1 we introduce notations used in this chapter. The derivation of the expresion for the WALE constant is presented in Section 5.2, while Section 5.3 shows the definitions for the non-dimensional parameters picked for this study and their spatial-dependent correlations with the WALE constant.

5.1 Notations

Here we define the filtering conventions for the velocity field and derived quantities. We employ a top-hat filter of width Δ to filter the high-resolution velocity field $u_i(\mathbf{y})$ provided by Ruiz *et al* [64]. Here $\mathbf{y}$ are the spatial coordinates and u_i refers to the $i-$th component of the velocity field. While the data is time dependent, we dropped the time dependency to simplify the notations below.

- Velocities

$$\begin{aligned} \tilde{u}_i &= \tilde{u}_i(\mathbf{x};\Delta) \equiv \int_\Omega u_i(\mathbf{y}) G(\mathbf{x}-\mathbf{y};\Delta)\, dv, &(5.1)\\ u_i' &= u_i - \tilde{u}_i &(5.2)\end{aligned}$$

- Kinetic energy

$$\widetilde{K} = \frac{1}{2}\widetilde{u_i u_i}, \quad K_{\text{res}} = \frac{1}{2}\tilde{u}_i\tilde{u}_i, \quad k^{sgs} = \widetilde{K} - K_{\text{res}} \tag{5.3}$$

- $\|\widetilde{u_i' u_j'}\|$ denotes the Frobenius norm

$$\|\widetilde{u_i' u_j'}\| = \sqrt{\sum_{i,j} \widetilde{u_i' u_j'}^2} \tag{5.4}$$

and

$$\widetilde{u_i' u_j'} = \widetilde{u_i' u_j'}|_{\mathbf{x};\Delta} = \int_\Omega u_i'(\mathbf{y}) u_j'(\mathbf{y})(\mathbf{y}) G(\mathbf{x}-\mathbf{y};\Delta)\, dv \tag{5.5}$$

5.2 Estimate WALE model constant c_w

Starting from the definition of the tubulent kinetic viscosity ν_t

$$\nu_t = \frac{\mu_t}{\rho} = (c_w\Delta)^2 \frac{\left(S_{ij}^d S_{ij}^d\right)^{3/2}}{\left(S_{ij}S_{ij}\right)^{5/2} + \left(S_{ij}^d S_{ij}^d\right)^{5/4}} \tag{5.6}$$

Here, S_{ij} and S_{ij}^d are given by

$$S_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right), \quad S_{ij}^d = \frac{1}{2}\left(g_{ij}^2 + g_{ji}^2\right) \tag{5.7}$$

with g_{ij}^2 given by

$$g_{ij}^2 = \frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_k}{\partial x_j} = \frac{\partial \tilde{u}_i}{\partial x_1}\frac{\partial \tilde{u}_1}{\partial x_j} + \frac{\partial \tilde{u}_i}{\partial x_2}\frac{\partial \tilde{u}_2}{\partial x_j} + \frac{\partial \tilde{u}_i}{\partial x_3}\frac{\partial \tilde{u}_3}{\partial x_j} \tag{5.8}$$

Given that ν_t can alse be defined as

$$\nu_t = \frac{(\tilde{u}_i\tilde{u}_j - \widetilde{u_i u_j})S_{ij} + \frac{2}{3}k^{sgs}\delta_{ij}S_{ij}}{2S_{kl}S_{kl}} \tag{5.9}$$

Using Eqs. (5.6) and (5.9) one can extract the WALE model constant

$$c_w = \left(\frac{(\tilde{u}_i\tilde{u}_j - \widetilde{u_i u_j})S_{ij} + \frac{2}{3}k^{sgs}\delta_{ij}S_{ij}}{2S_{kl}S_{kl}}\right) \Bigg/ \left(\Delta^2 \frac{\left(S^d_{ij}S^d_{ij}\right)^{3/2}}{\left(S_{ij}S_{ij}\right)^{5/2} + \left(S^d_{ij}S^d_{ij}\right)^{5/4}}\right) \tag{5.10}$$

5.3 Parameter of Interest

We define in Table 5.1 several non-dimensional parameters based on quantities derived from the velocity field: p_1 is based on the Frobenius norm of the filtered velocity gradient tensor; p_2 is constructed with the norm of the vorticity field based on the filtered velocity; p_3 is based on the Frobenius norm of the outer product of the filtered velocity field; p_4 is based on the magnitude of the filtered velocity vector, essentially the Re number based on the fieltered velocity; p_5 is based on the Frobenius norm of the strain rate tensor; p_6 is based on the determinant of the strain rate tensor.

Parameter	Formula	Obs.
p_1	$\Vert\frac{\partial \tilde{u}}{\partial x}\Vert\Delta^2/\nu$	$\left\Vert\frac{\partial \tilde{u}}{\partial x}\right\Vert = \sqrt{\frac{\partial \tilde{u}_i}{\partial x_j}\frac{\partial \tilde{u}_i}{\partial x_j}}$
p_2	$\Vert\omega\Vert\Delta^2/\nu$	$\Vert\omega\Vert = \sqrt{\omega_i\omega_i}$, $\omega = \nabla\times\tilde{u}$
p_3	$\sqrt{\Vert\tilde{u}\otimes\tilde{u}\Vert}\Delta/\nu$	$\Vert\tilde{u}\otimes\tilde{u}\Vert = \sqrt{\sum_{i,j}\left(\tilde{u}_i\tilde{u}_j\right)^2}$
p_4	$\sqrt{\tilde{u}\cdot\tilde{u}}\Delta/\nu$	$\tilde{u}\cdot\tilde{u} = \tilde{u}_i\tilde{u}_i$
p_5	$\Vert\mathbf{S}\Vert\Delta^2/\nu$	$\Vert\mathbf{S}\Vert = \sqrt{\sum_{i,j}{S_{ij}}^2}$
p_6	$\det(\mathbf{S})^{1/3}\Delta^2/\nu$	-

Table 5.1: Non-dimensional flow markers.

We compute the Pearson correlation coefficient between each of the parameters presented in Table 5.1 and the WALE model constant presented in Section 5.2. For example, the correlation coefficient beween p_1 and C_w is computed as

$$r_{p_1,C_w}(\mathbf{x}) = \frac{\sum_{i=1}^{N_t}\left(p_1(t_i,\mathbf{x}) - \overline{p_1}(\mathbf{x})\right)\left(C_w(t_i,\mathbf{x}) - \overline{C_w}(\mathbf{x})\right)}{\sqrt{\sum_{i=1}^{N_t}\left(p_1(t_i,\mathbf{x}) - \overline{p_1}(\mathbf{x})\right)^2}\sqrt{\sum_{i=1}^{N_t}\left(C_w(t_i,\mathbf{x}) - \overline{C_w}(\mathbf{x})\right)^2}} \tag{5.11}$$

where t_i is the time and $\mathbf{x} = \{x,y,z\}$ is the spatial coordinate. The expected values $\overline{p_1}(\mathbf{x})$ and $\overline{C_w}$ are computed with respect to time,

$$\overline{p_1}(\mathbf{x}) = \frac{1}{N_t}\sum_{i=1}^{N_t} p_1(t_i,\mathbf{x}),\ \overline{C_w}(\mathbf{x}) = \frac{1}{N_t}\sum_{i=1}^{N_t} C_w(t_i,\mathbf{x}) \tag{5.12}$$

For this study we employed approximatelt $N_t = 90$ time snapshots spanning *aaa* flow-through times.

Figures 5.2 and 5.3 show Pearson correlation coefficients between C_w and p_1, p_2, p_3, and p_5, respectively. The results corresponding to p_4 are similar to the ones bases on p_3, while the results based on p_6 are similar to the ones bases on p_5. Consequently, p_4 and p_6 are not shown for brevity. The contour plots in the center plane, $z = 0$ indicate two types of correlations. One type, corresponding to p_1 and p_5 indicate consistent correlations throught the plane after the the two flows start mixing. The correlations are stronger in the regions upstream of the turbulence breakdown point (TBP). The second type corresponds to p_2 which only exhibits relevant correlations with C_w on the edges of the mixing region between the cross-flow and the jet. Parameter p_3 exhibits little correlations with the WALE model constant.

The results in the $y = 15$ plane, in Fig. 5.3, confirm the similarities between the results based on p_1 and p_5, These parameters exhibit correlation coefficient values around $0.5 - 0.6$ throughout most of the mixing region. By this flow region, the correlations between p_2 and C_w are mostly negligible. The ocasional “spikes” are most likely due to the limited number of time steps employed for these calculations.

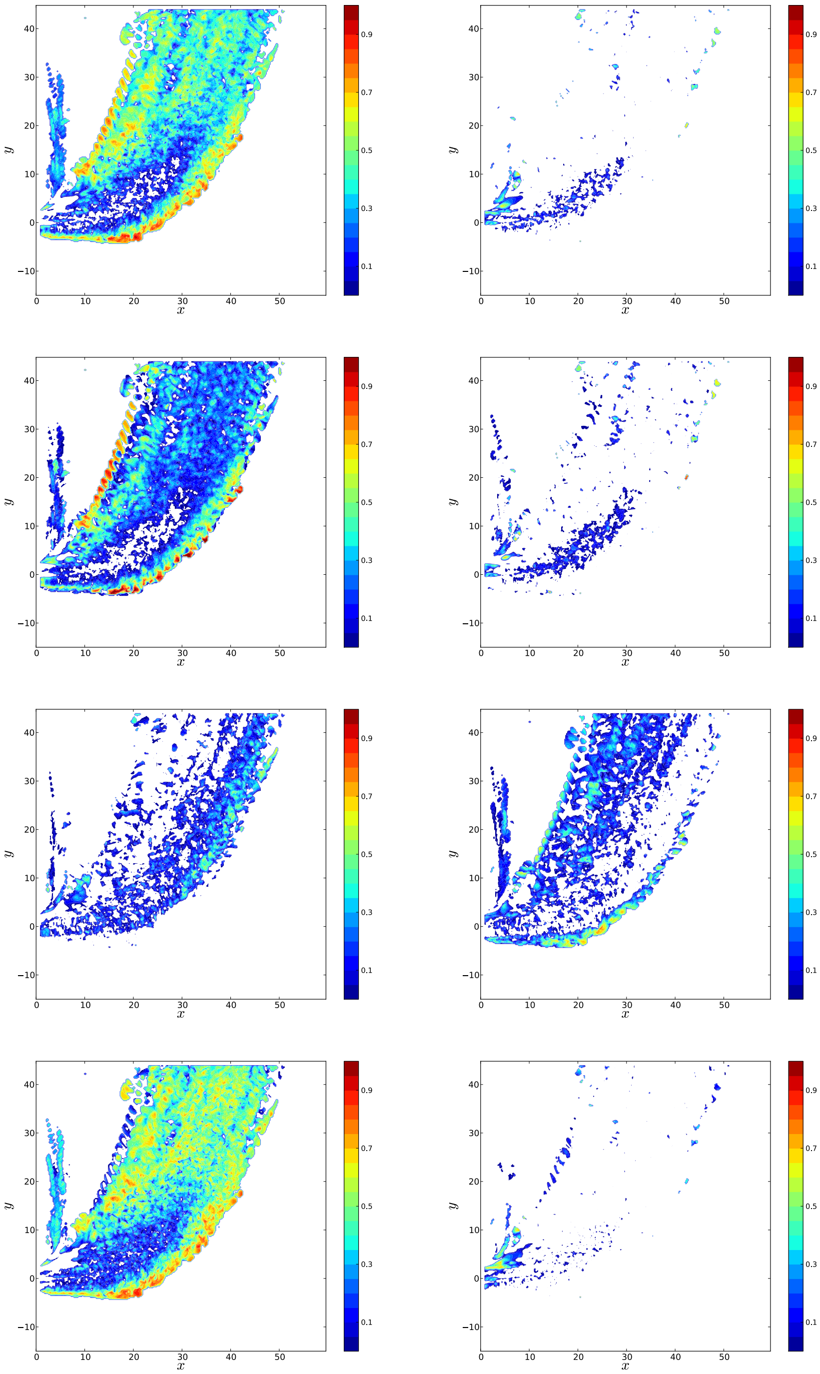


Figure 5.2: Correlation coefficient contours in the $z = 0$ plane. The left column shows contours corresponding to positive correlations, while the right column shows the absolute magnitude of the negative correlation regions. Rows correspond to correlations between C_w and p_1, p_2, p_3, and p_5, respectively.

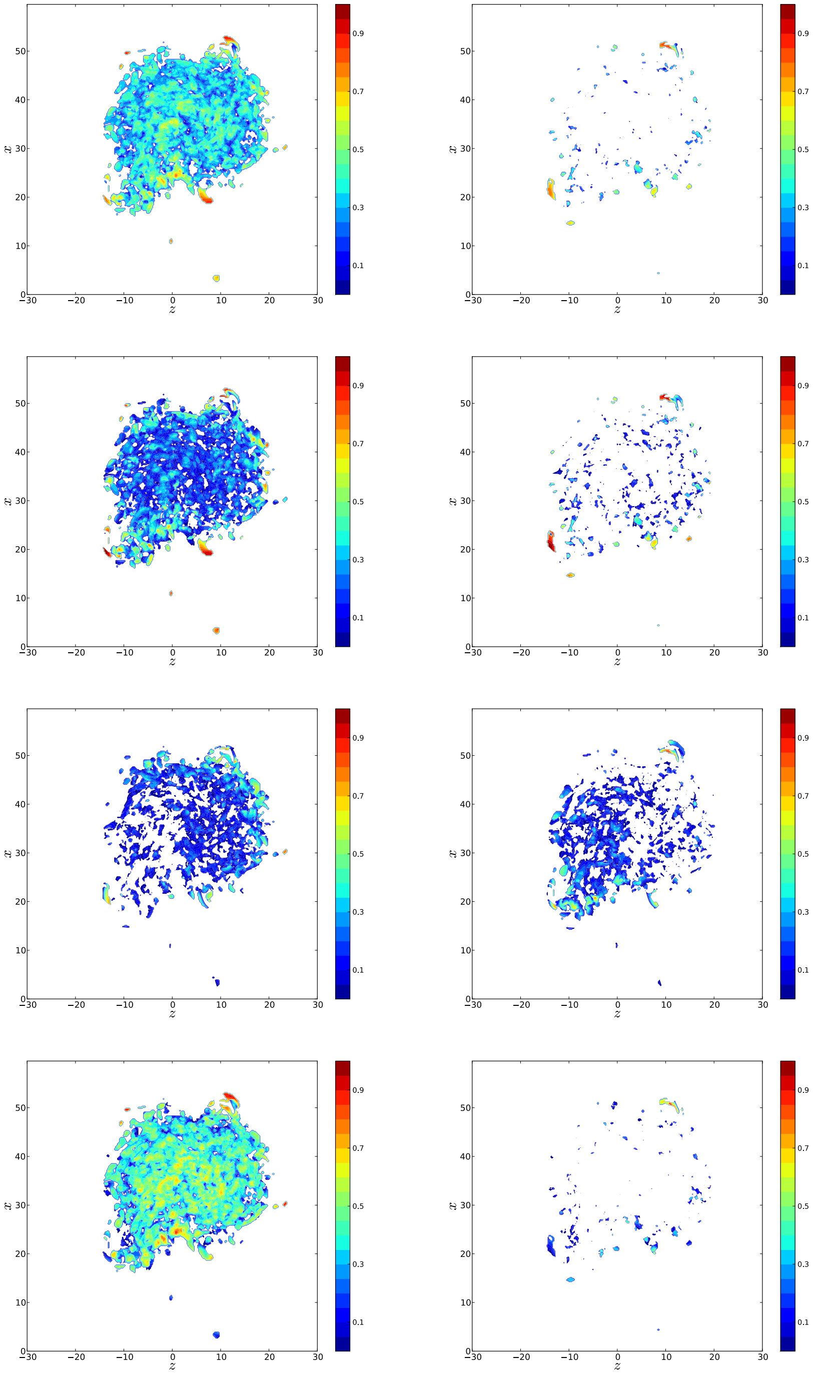


Figure 5.3: Correlation coefficient contours in the $Y = 15$ plane. The left column shows contours corresponding to positive correlations, while the right column shows the absolute magnitude of the negative correlation regions. Rows correspond to correlations between C_w and p_1, p_2, p_3, and p_5, respectively.

Chapter 6

Sensitivity of Combustion Models to Physical Variation

6.1 Introduction

The existence of a critical dissipation rate, above which a nonpremixed flame is extinguished, has been known for decades. Recent advances in modeling have allowed the simulation of turbulent nonpremixed flames that include local extinction as a consequence of the stochastic variation in mixing rates. In this chapter we present the critical dissipation impulse magnitude that will lead to extinction even if the mean dissipation rate is well below the criteria for a steady flame. This critical impulse magnitude depends on the time-integrated excess dissipation rate, stoichiometric factors and the form of the S-curve describing the steady-state flame. This criteria is evaluated in a diverse set of flames including n-heptane, diluted n-heptane and $CO/H_2/N_2$ mixtures. A published paper presents additional results, includes further discussion of the time scales over which this extinction criterion is applicable, and further discusses the results in the context of prior literature [25].

A criteria for which turbulent flames are extinguished is key to predicting such diverse phenomena as fire suppression and flame stabilization. While the theoretical basis for extinction of an individual laminar nonpremixed flamelet (here used broadly to indicate a stoichiometric fluid element where chemical reactions take place due to molecular mixing as opposed to the turbulent flame which is a collection of flamelets) has been known for decades [41], the details of extinction in turbulent flames have been coming to light more recently as stochastic simulations of localized extinction and reignition are conducted [27, 57, 84]. The present report is a step toward quantifying the frequency of turbulent flamelet extinction without the need to carry out a stochastic simulation. To accomplish this, a criteria for extinction due to an unsteady variation in the mixing rate first proposed in [27] is evaluated.

For an individual nonpremixed flamelet, extinction may occur if the rate of heat loss exceeds the rate of chemical heat release. This rate of heat loss can be characterized by the scalar dissipation rate, χ, defined in terms of the gradient of the mixture fraction, Z, as $\chi = 2D|\nabla Z|^2$. All dissipation rates in this report are evaluated at the stoichiometric point. The response of a flamelet to a steady dissipation rate can be characterized using the well-known S-curve showing the maximum temperature as a function of dissipation rate. The upper and middle branches are shown in Fig. 6.1 for boundary conditions considered in this work. The upper branch represents stably burn-

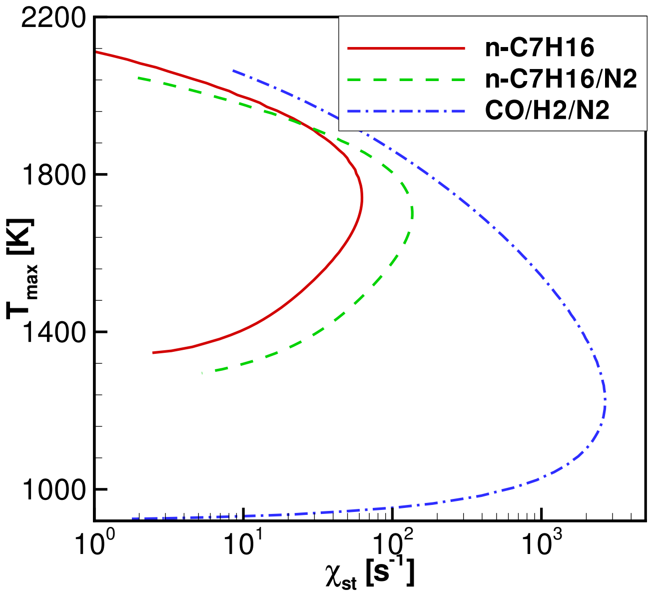


Figure 6.1: The maximum flame temperature as a function of χ for steady laminar flamelets with three fuel mixtures (left). Possible trajectories for unsteady flamelets are shown relative to the steady state by dotted arrows at right.

ing flames while the middle branch is unstable; flames will tend to move away from the middle branch toward the upper or lower branches. Note that the S-curve abscissa in this report is χ rather than the Damköhler number, which is inversely proportional to χ; using χ along the abscissa allows easier interpretations of changes in the mixing rate, but results in a reversed S-curve. Above a critical dissipation rate, heat losses exceed heat-release rates and a steady flame cannot exist; this point, χ_q, is associated with extinction for steady flames.

One problem that arises in the application of the S-curve describing steady flamelets is that the instantaneous value of χ varies with time in turbulent flows about whatever effective mean value is determined by the large-scale mixing. The distribution of χ can be approximated as log normal, and for moderate to large Reynolds numbers χ may vary over several orders of magnitude. In direct numerical simulations, the local dissipation rate has been shown to vary over time scales comparable to the Kolmogorov time scale [86], so that large fluctuations in χ may occur over short periods of time. The present report addresses how the magnitude and time scale of rapid fluctuations in χ determine whether or not extinction will occur.

A number of studies addressing the response of individual flames to unsteady mixing time scales have identified the fact that flames are more resistant to extinction for high frequency or short period mixing rate fluctuations. Ghoniem et al. [20] showed that by increasing the frequency one can go to higher amplitude fluctuations before the flame is extinguished. Mauss et al. [40] showed that, for a fixed peak dissipation rate, exceeding χ_q for a longer duration required a greater subsequent reduction in the dissipation rate to prevent extinction. These two observations can be coupled together in an extinction criteria first proposed in [27] and described in the subsequent section.

6.2 Method and Approach to Analyzing Results

For ease in interpretation of the results and to simplify the form of the equations, a simplified version of the flamelet equations written in the mixture fraction coordinate is employed. While detailed chemical kinetics and thermo chemistry (i.e. enthalpies) are retained, all transport properties are assumed identical so that Lewis numbers are unity, and the species diffusion is presumed to follow Fick's Law so that the effect of gradients in the molecular weight on diffusion are ignored. Simplified species and energy equations are

$$\frac{dY_i}{dt} = \frac{\chi}{2}\frac{d^2Y_i}{dZ^2} + \frac{\omega_i}{\rho} \tag{6.1}$$

$$\frac{dT}{dt} = \frac{\chi}{2}\frac{d^2T}{dZ^2} - \frac{\chi}{2}\sum_i\left(1-\frac{c_{p,i}}{c_p}\right)\frac{dY_i}{dZ}\frac{dT}{dZ} - \sum_i\frac{\omega_i h_i}{\rho c_p}. \tag{6.2}$$

Here Y_i, ω_i, $c_{p,i}$ and h_i are the mass fraction, production rate, specific heat and enthalpy, respectively, of species i; ρ and c_p are the mixture density and specific heats. The scalar dissipation rate, χ is allowed to vary in the mixture fraction coordinate with the standard counterflow configuration form, $\chi(Z) = \chi_0\exp(-2[\mathrm{erfc}^{-1}(2Z)]^2)$. These equations are evolved forward in time from a steady-state initial condition using the FlameMaster code [56]. During the temporal evolution the value of χ_0 is varied so that the stoichiometric dissipation rate follows either a square-shaped, triangle-shaped or sinusoidal impulse with an initial magnitude χ_1 and a peak value of χ_2 as indicated in Fig. 6.2. In order to determine χ_q, to define the shape of the S-curves in Fig. 6.1 and to generate initial conditions for the transient flames, a series of steady flamelets are also computed using continuation methods.

Computations are conducted using varied fuel mixtures to provide a wide range of parameter space. Fuel mixtures employed are pure n-heptane, n-heptane diluted with nitrogen (50% by mass) and a $CO/H_2/N_2$ mixture (55.4%/3%/41.6% by mass). The chemical-kinetic mechanisms employed are available from the author and are similar to those described in [37, 24]. The present results are insensitive to the mechanism.

Unsteady extinction results can be explained in terms of the S-curve [40]. Variations in χ tend to move the state away from the S-curve as depicted with dotted arrows in Fig. 6.3. An increase in the dissipation rate above χ_q will result in the flame temperature dropping with time (segments 1 and 2 in Fig. 6.3). If the duration for which $\chi > \chi_q$ is sufficiently short, the subsequent reduction in the dissipation rate leads to a flame state in between the upper and middle branches of the S-curve and the flame, given sufficient time, returns to the steady state value on the upper branch (segments 3 and 4 in Fig. 6.3). Conversely, if the duration for which $\chi > \chi_q$ is sufficiently long, a reduction in the dissipation rate will lead to a flame state below the middle branch, in which case the flame temperature will continue to drop (segments 5 and 6 in Fig. 6.3). Though it is beyond the scope of the present work, we note in passing that the steady-state middle branch only represents an approximate dividing line because, particularly with detailed chemistry, the transient flame state differs from that at steady-state.

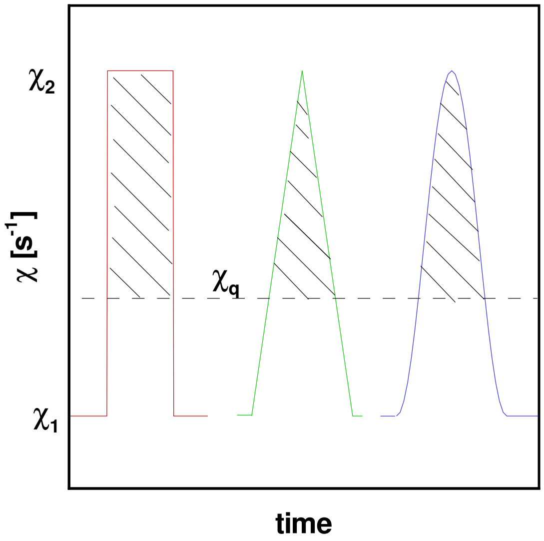


Figure 6.2: Temporal evolution of χ and the region contained in the integral of Ξ for square-shaped, triangle-shaped and sinusoidal impulse forms.

In Ref. [27] a criteria was proposed for extinction when fluctuations in the dissipation rate exceed χ_q for a period in the course of turbulent flame evolution. This criteria is based on estimates for the heat-release rate and heat-dissipation rate. The maximum heat-release rate can be approximated by the steady mixing rate at the quenching condition, equating the second and fourth terms in the energy equation in Eq. 6.2

$$\left(\sum_i \frac{\omega_i h_i}{\rho c_p}\right)_{max} \approx -\frac{\chi_q}{2}\frac{d^2T}{dZ^2}. \tag{6.3}$$

It is presumed that while χ_q is exceeded, but while the temperature is still above the middle branch, heat release may continue at approximately this rate. The diffusive term can be estimated as $-\chi T Z_{st}^{-2}(1-Z_{st})^{-2}$ so that, for $\chi > \chi_q$ and temperatures above the middle branch, the rate of change in the thermochemical state, represented by temperature, can be estimated (neglecting the enthalpy flux for simplicity) as

$$\frac{dT}{dt} = \frac{\chi}{2}\frac{d^2T}{dZ^2} - \left(\sum_i \frac{\omega_i h_i}{\rho c_p}\right)_{max} \approx -\frac{(\chi_q-\chi)}{2}\frac{d^2T}{dZ^2} \approx \frac{(\chi_q-\chi)T}{Z_{st}^2(1-Z_{st})^2}. \tag{6.4}$$

This creates a plausible time scale for extinction. The time scale, the inverse of the right-hand side of Eq. 6.4, is the exponential decay constant since the solution of Eq. 6.4 is of the form

$$\frac{T_2-T_0}{T_1-T_0} = \exp(-A\Xi) \tag{6.5}$$

where

$$\Xi = \frac{\int_{\chi>\chi_q}(\chi-\chi_q)dt}{Z_{st}^2(1-Z_{st})^2}. \tag{6.6}$$

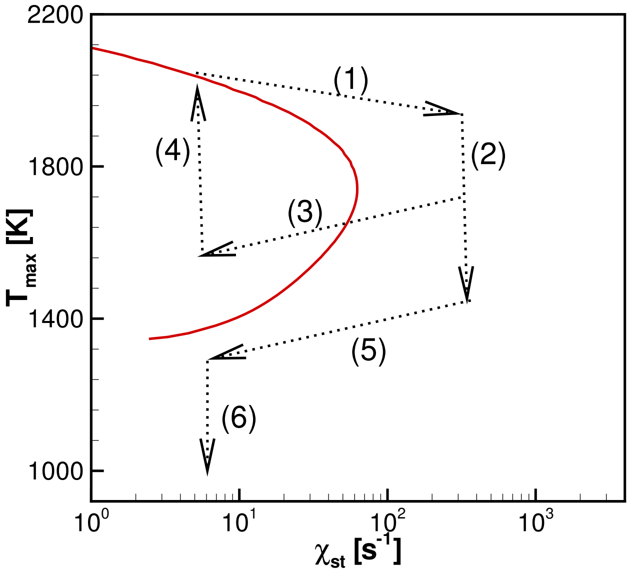


Figure 6.3: Possible trajectories for unsteady flamelets shown relative to the steady state by dotted arrows.

Here A is a proportionality constant since Eq. 6.4 is only an approximate equality. A better approximation would involve a prediction of the reaction zone thickness as can be obtained through asymptotic means, for example [41, 28]. Here T_0 is the temperatures of the reactants, T_1 the steady flame temperature before the scalar dissipation rate rises and T_2 the predicted unsteady flame temperature after χ is again less than χ_q. In the numerator of Ξ the integral $\int_{\chi>\chi_q}(\chi_q-\chi)dt$ appears. This is the integral over time of the excess dissipation above χ_q as indicated by the hatched region in Fig 6.2. If the value of T_2 predicted by Eq. 6.6 when χ returns to its original value, χ_1, is below the middle branch, then we expect the flame to be extinguished.

6.3 Results and Discussion

In order to evaluate Eq. 6.6 as an unsteady extinction criteria, unsteady simulations were performed as described in the previous section. Using a bisection search algorithm, the critical value of Ξ that delineates the division between extinguished and non-extinguished flames, Ξ_q, is identified as a function of the initial dissipation rate, χ_1, and the maximum dissipation rate, χ_2, for the three profiles indicated in Fig. 6.2 by varying the duration of the pulse. As an example of the evolution of the thermochemical state, Fig. 6.4 shows the temperature and dissipation rate as a function of time along with the temperature-dissipation phase plot for three pairs of simulations with durations characterized by Ξ somewhat above (dashed lines) and below Ξ_q (solid lines). It is noted that a square temporal profile for χ results in virtually no change in temperature between χ_1 and χ_2 and the solution moves well away from the stable S-curve solution (the duration of the transition is 1% of the total impulse period). For a triangle (or sinusoidal) temporal profile, however, the magnitude and period of the pulse determine whether or not the temperature follows the S-curve. For χ_2/χ_q

not very large, the duration of the impulse for Ξ near Ξ_q is long and the gradual changes in χ allow the transient state to largely follow the S-curve (i.e. the middle panes of Fig. 6.4). For larger χ_2/χ_q (lower panes of Fig. 6.4) the solution increasingly moves away from the S-curve.

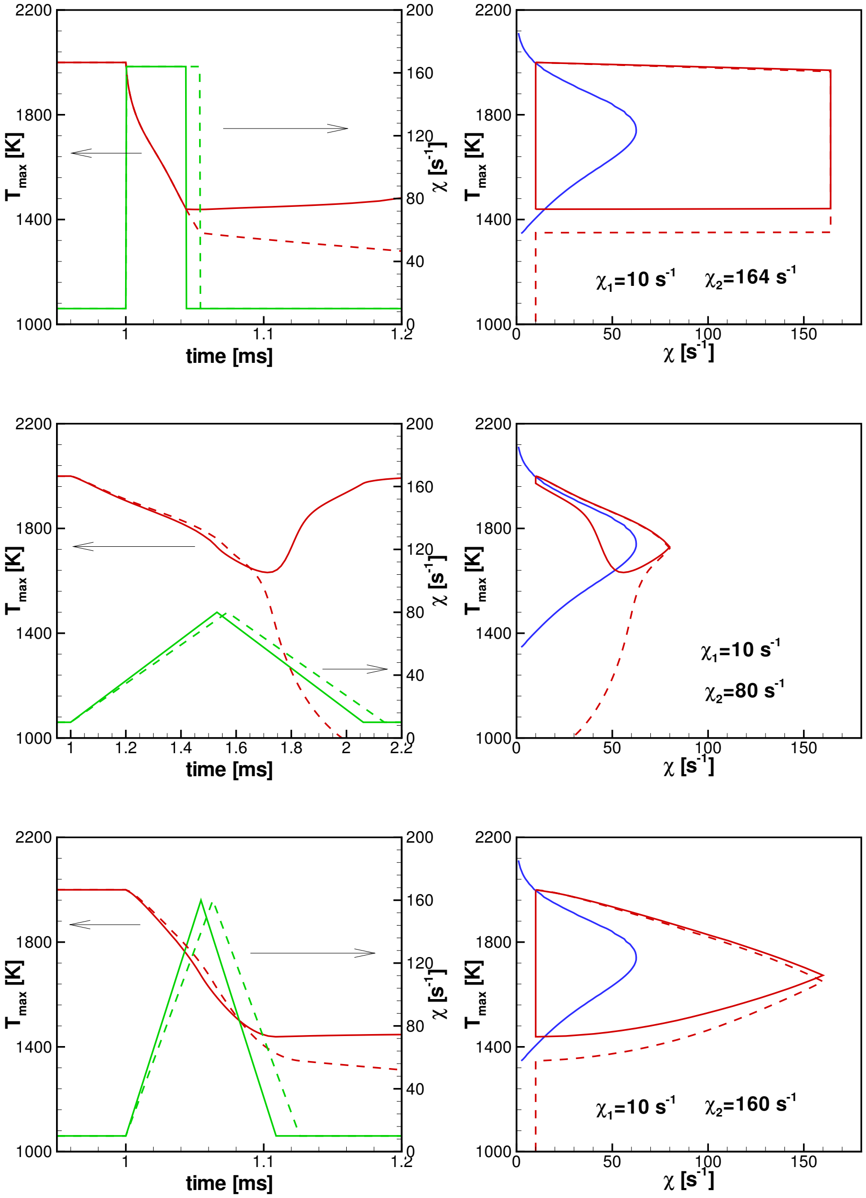


Figure 6.4: The temporal evolution of χ and the temperature (left) and the temperature-dissipation phase plot (right) for one square and two triangle-shaped dissipation impulses.

The value of Ξ_q is determined to within 1% for a wide range of $\chi_1/\chi_q < 1$, $\chi_2/\chi_q > 1$ and for the three fuel mixtures indicated. These are plotted in Fig. 6.5, and a number of observation can be made. For smaller χ_1/χ_q where the required reduction in the temperature to cross the middle branch of the S-curve is greater, the magnitude of Ξ_q is generally greater as predicted by Eq. 6.6.

For sinusoidal and triangle-shaped impulses, as χ_2/χ_q approaches unity the magnitude of of Ξ_q approaches the value for a flame near extinction ($\chi_1/\chi_q \rightarrow 1$). This occurs because the flame has time to adapt to the near-extinction state as indicated in the middle panes of Fig. 6.4. Because the change in χ is rapid for square impulses, this is not observed in those curves. To quantify the meaning of rapid versus gradual changes in the dissipation rate, one must compare dT/dt from Eq. 6.4 with $(dT/d\chi)(d\chi/dt)$. If the latter is small relative to the former, then the flame state will follow the S-curve, until such a condition is violated (as when $dT/d\chi \rightarrow -\infty$ at χ_q). As a preliminary estimate of the significance of rapid fluctuations, we note that Yeung and coworkers [86, 85] looked at the dynamics of flame surface straining using direct numerical simulations. They observed that the integral time scales for variation in strain rates are only somewhat larger than the Kolmogorov time scale; in Ref. [86] they indicate this is three times the Kolmogorov scale. Therefore, we expect rapid fluctuations in the dissipation rate are typical in large Reynolds number turbulence.

For rapid changes in χ, as occurs for large χ_2/χ_q and for square profiles, the value of Ξ_q becomes independent of the shape of the impulse and depends on the initial conditions. In this case, we identify the magnitude of T_1 and T_2 appearing in Eq. 6.6 by reading from the appropriate S-curve in Fig. 6.1 at the location χ_1 and normalize Ξ_q by $\ln(T_2 - T_0)/(T_1 - T_0)$ giving the constant of proportionality, A, in Eq. 6.6. This is plotted in Fig. 6.6. It is observed in Fig. 6.6 that for large χ_2/χ_q the curves come together (for small χ_2/χ_q the change in χ is gradual so that T_1 should be taken near extinction resulting in the convergence evident in Fig. 6.5) indicating that the relationship in Eq. 6.6 is a suitable predictor of unsteady extinction criteria for large fluctuations in the dissipation rate.

There is still some systematic variation in the value of A indicated in Fig. 6.6. In particular, as $\chi_1/\chi_q \rightarrow 1$, the appropriate magnitude of A is larger. While space limitations prevent a graphical demonstration, this occurs because the reaction rates are initially higher in flames with initially high dissipation rates. The reaction rates are observed to *transiently* exceed the reaction rates at χ_q leading to greater resistance to extinction than predicted in Eq. 6.6. This is a purely transient phenomena occurring as the temperature remains above the steady-state temperature at χ_q (so that the kinetic rates are still fast) and the high dissipation rates mix reactants more rapidly. This phenomena will require further investigation.

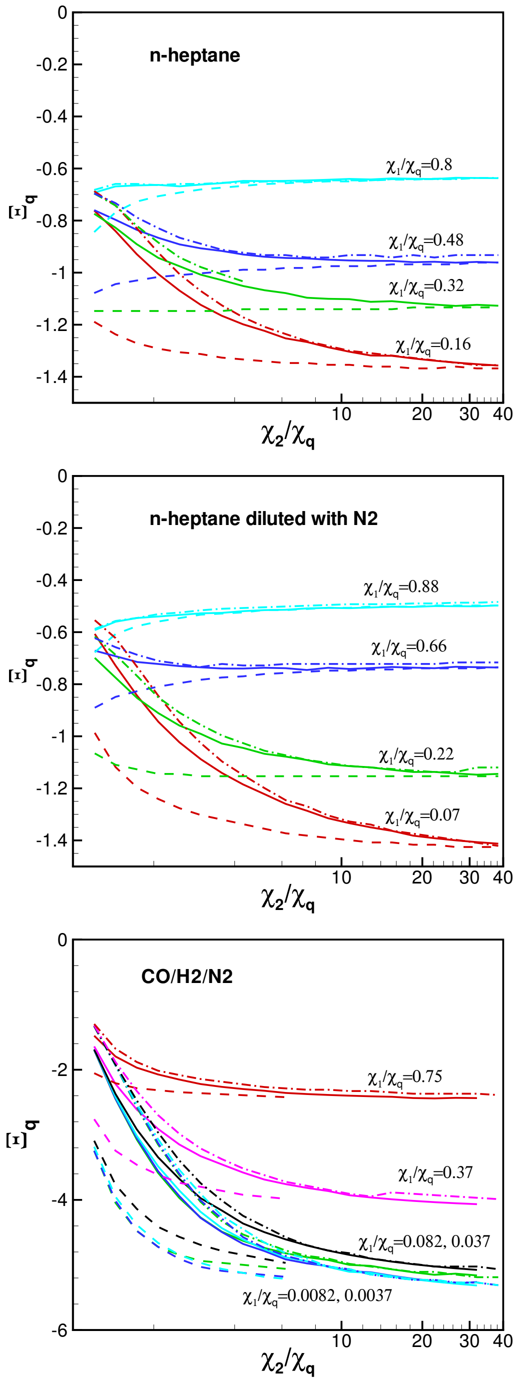


Figure 6.5: The critical value of Ξ corresponding to extinction for various dissipation impulses. Solid lines are sinusoidal, dashed lines are square-shaped and dash-dot lines are triangle-shaped profiles.

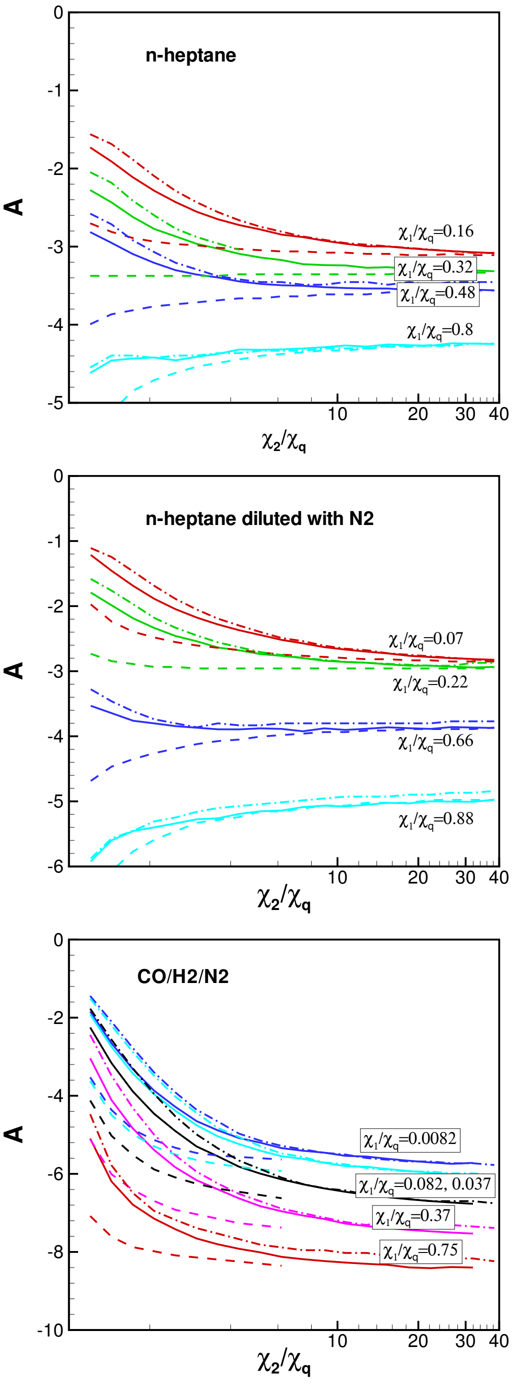


Figure 6.6: Ξ_q normalized by $\ln(T_2 - T_0)/(T_1 - T_0)$ giving the constant of proportionality, A in Eq. 6.6. Solid lines are sinusoidal, dashed lines are square-shaped and dash-dot lines are triangle-shaped profiles.

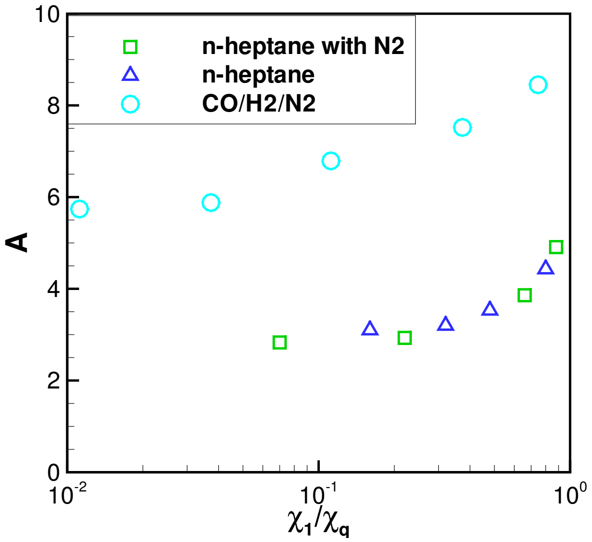


Figure 6.7: The value of A for large χ_2/χ_q.

To clarify the variation of A in the limit of large χ_2/χ_q, the approximate limiting values are plotted in Fig. 6.7. There, the variation in A as $\chi_1/\chi_q \rightarrow 1$ described in the previous paragraph is clear. Furthermore, for n-heptane with and without nitrogen dilution, the curves directly coincide indicating that the effect of dilution is captured. Conversely, the predicted value of A for the CO/H_2/N_2 flames is observed to be substantially greater. As indicated in the discussion of Eq. 6.6, the approximations, particular that for the heat loss, would best use a true measure of the reaction zone thickness. Since the reaction zone thickness for hydrocarbon flames is generally less than that for CO/H_2 flames, the greater magnitude of A for the CO/H_2/N_2 flames is expected until a better measure of the heat losses is provided. The better measure of the heat losses would be $-\chi T/[Z_{st}(1-Z_{st})\varepsilon]$ where ε is the reaction zone thickness in mixture fraction space, obtainable from asymptotic analysis [41, 28].

6.4 Conclusion

A criteria indicating the magnitude of a scalar-dissipation impulse leading to extinction has been investigated. The quantity Ξ, involving the time-integrated excess dissipation over χ_q and appropriate stoichiometric factors, has been shown to provide a sufficient description of the heat losses during a brief period where the dissipation reaches values much greater than χ_q so that, when coupled with knowledge of the S-curve temperature-dissipation relationship, extinction can be reasonably well predicted. When the rate of change in dissipation is fast relative to the rate of response of the temperature to changes in χ, this criteria is independent of the profile of the dissipation impulse, but does depend on the temperatures at the steady-state condition along the S-curve prior to the impulse. When the dissipation rate changes gradually, the criteria is independent of the initial steady-state conditions, but depends on the state along the S-curve near extinction or wherever the rate of change in the dissipation rate becomes large. For initial and or maximum dissipation rates near χ_q, this criteria based on Ξ slightly over predicts heat losses, and would thus predict extinction too soon if not corrected. A further study with greater detail is available separately [25].

Chapter 7

Development of Reacting Flow Capabilities in Nalu

7.1 Low Mach Equation Set

The low Mach number equations are a subset of the fully compressible equations of motion (momentum, continuity and energy), admitting large variations in gas density while remaining acoustically incompressible. The low Mach number equations are preferred over the full compressible equations for low speed flow problems as the accoustics are of little consequence to the overall simulation accuracy. The technique avoids the need to resolve fast-moving acoustic signals. Derivations of the low Mach number equations can be found in found in Rehm and Baum, [60], or Paolucci, [51]. A condensed version of the derivation can also be found in [14].

The equations are derived from the compressible equations using a perturbation expansion in terms of the lower limit of the Mach number squared; hence the name. The asymptotic expansion leads to a splitting of pressure into a spatially constant thermodynamic pressure and a locally varying dynamic pressure. The dynamic pressure is decoupled from the thermodynamic state and cannot propagate acoustic waves. The thermodynamic pressure is used in the equation of state and to determine thermophysical properties. The thermodynamic pressure can vary in time and can be calculated using a global energy balance.

7.1.1 Conservation of Mass

The continuity equation is always solved in the variable density form.

$$\int \frac{\partial \bar{\rho}}{\partial t} dV + \int \bar{\rho} \tilde{u}_i n_i dS = 0 \tag{7.1}$$

Since Nalu uses equal-order interpolation (variables are collocated) stabilization is required. The stabilization choice will be developed in the pressure stabilization section.

7.1.2 Conservation of Momentum

The integral form of the Favre-filtered momentum equations used for turbulent transport are

$$\begin{aligned}\int \frac{\partial \bar{\rho} \tilde{u}_i}{\partial t} \mathrm{d}V + \int \bar{\rho} \tilde{u}_i \tilde{u}_j n_j \mathrm{d}S + \int \bar{P} n_i \mathrm{d}S &= \int \bar{\tau}_{ij} n_j \mathrm{d}S + \int \tau_{u_i u_j} n_j \mathrm{d}S \\ &+ \int (\bar{\rho} - \rho_\circ) g_i \mathrm{d}V, \end{aligned} \tag{7.2}$$

where the turbulent stress $\tau_{u_i u_j}$ is defined as

$$\tau_{u_i u_j} \equiv -\bar{\rho}\left(\widetilde{u_i u_j} - \tilde{u}_i \tilde{u}_j\right). \tag{7.3}$$

In a low Mach flow, as described in the low Mach theory section, the above pressure, $\bar{P}$ is the purturbation about the thermodynamic pressure, P^{th}. In a low speed compressible flow, i.e., flows confinded to a closed domain with energy or mass addition in which the continuity equation has been modifed to accomodate accoustics, this pressure is interpreted at the thermodynamic pressure itself.

For LES, $\tau_{u_i u_j}$ in Equation 7.2 represents the subgrid stress tensor. The deviatoric part of the subgrid stress tensor is defined as

$$\begin{aligned}\tau^D_{u_i u_j} &\equiv \tau_{u_i u_j} - \frac{1}{3} \tau_{u_k u_k} \delta_{ij} \\ &= \tau_{u_i u_j} + \frac{2}{3} \bar{\rho} q^2 \delta_{ij}, \end{aligned} \tag{7.4}$$

where the subgrid turbulent kinetic energy is defined as $q^2 \equiv \frac{1}{2}(\widetilde{u_k u_k} - \tilde{u}_k \tilde{u}_k)$. The deviatoric part of the subgrid stress tensor is then modeled similar to RANS closures as,

$$\tau^D_{u_i u_j} = 2\mu_t \left(\tilde{S}_{ij} - \frac{1}{3} \tilde{S}_{kk} \delta_{ij} \right). \tag{7.5}$$

Substituting this into Equation 7.4 yields the modeled form of the full subgrid stress tensor

$$\tau_{u_i u_j} = 2\mu_t \left(\tilde{S}_{ij} - \frac{1}{3} \tilde{S}_{kk} \delta_{ij} \right) - \frac{2}{3} \bar{\rho} q^2 \delta_{ij}. \tag{7.6}$$

For low Mach-number flows, a vast majority of the turbulent kinetic energy is contained at resolved scales. For this reason, the subgrid turbulent kinetic energy q^2 will not be directly treated and will instead be included in the pressure as an additional normal stress. The Favre-filtered momentum equations then become

$$\begin{aligned}&\int \frac{\partial \bar{\rho} \tilde{u}_i}{\partial t} \mathrm{d}V + \int \bar{\rho} \tilde{u}_i \tilde{u}_j n_j \mathrm{d}S + \int \left(\bar{P} + \frac{2}{3} \bar{\rho} q^2 \right) n_i \mathrm{d}S = \\ &\quad \int 2(\mu + \mu_t) \left(\tilde{S}_{ij} - \frac{1}{3} \tilde{S}_{kk} \delta_{ij} \right) n_j \mathrm{d}S + \int (\bar{\rho} - \rho_\circ) g_i \mathrm{d}V, \end{aligned} \tag{7.7}$$

where LES closure models for the subgrid turbulent eddy viscosity μ_t are either the constant coefficient Smagorinsky, WALE or the constant coefficient k_{sgs} model (see the turbulence section).

7.1.3 Filtered Mixture Fraction

The optional quantity used to identify the chemical state is the mixture fraction, Z. While there are many different definitions of the mixture fraction that have subtle variations that attempt to capture effects like differential diffusion, they can all be interpreted as a local mass fraction of the chemical elements that originated in the fuel stream. The mixture fraction is generally considered to be a conserved scalar that varies between zero in the secondary stream and unity in the primary stream and is transported in laminar flow by the equation,

$$\frac{\partial \rho Z}{\partial t} + \frac{\partial \rho u_i Z}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\rho D \frac{\partial Z}{\partial x_i}\right), \tag{7.8}$$

where D is an effective molecular mass diffusivity. For some applications, the mixture fraction is not conserved and a source term will apply. Many of these applications involve multiphase flow not addressed here and the mixture-fraction source is associated with the interphase reactions. When species diffusion coefficients are not equal, an effective source term also appears in Eq. 7.8, but turbulent fluctuations act to reduce the effect of this, and it is ignored in the filtered equations [26].

Applying either temporal Favre filtering for RANS-based treatments or spatial Favre filtering for LES-based treatments yields

$$\int \frac{\partial \bar{\rho} \tilde{Z}}{\partial t} \mathrm{d}V + \int \bar{\rho} \tilde{u}_j \tilde{Z} n_j \mathrm{d}S = -\int \tau_{Zu_j} n_j \mathrm{d}S + \int \bar{\rho} D \frac{\partial \tilde{Z}}{\partial x_j} n_j \mathrm{d}S, \tag{7.9}$$

where sub-filter correlations have been neglected in the molecular diffusive flux vector and the turbulent diffusive flux vector is defined as

$$\tau_{Zu_j} \equiv \bar{\rho}\left(\widetilde{Zu_j} - \tilde{Z}\tilde{u}_j\right). \tag{7.10}$$

This sub-filter correlation is modeled in both RANS and LES closures with the gradient transport approximation

$$\tau_{Zu_j} \approx -\bar{\rho} D_t \frac{\partial Z}{\partial x_j}, \tag{7.11}$$

where D_t is the turbulent mass diffusivity, modeled as $\bar{\rho} D_t = \mu_t / \mathrm{Sc}_t$ where μ_t is the modeled turbulent viscosity from momentum transport and Sc_t is the turbulent Schmidt number. The molecular mass diffusivity is then expressed similarly as $\bar{\rho} D = \mu / \mathrm{Sc}$ so that the final modeled form of the filtered mixture fraction transport equation is

$$\frac{\partial \bar{\rho} \tilde{Z}}{\partial t} + \frac{\partial \bar{\rho} \tilde{u}_i \tilde{Z}}{\partial x_i} = \frac{\partial}{\partial x_i}\left[\left(\frac{\mu}{\mathrm{Sc}} + \frac{\mu_t}{\mathrm{Sc}_t}\right)\frac{\partial \tilde{Z}}{\partial x_i}\right]. \tag{7.12}$$

In integral form the mixture fraction transport equation is

$$\int \frac{\partial \bar{\rho} \tilde{Z}}{\partial t} dV + \int \bar{\rho} \tilde{u}_i \tilde{Z} n_i \, dS = \int \left(\frac{\mu}{\mathrm{Sc}} + \frac{\mu_t}{\mathrm{Sc}_t}\right) \frac{\partial \tilde{Z}}{\partial x_i} n_i \, dS. \tag{7.13}$$

7.1.4 Conservation of Energy

The integral form of the Favre-filtered static enthalpy energy equation used for turbulent transport is

$$\begin{aligned}\int \frac{\partial \bar{\rho}\tilde{h}}{\partial t}\mathrm{d}V + \int \bar{\rho}\tilde{h}\tilde{u}_j n_j \mathrm{d}S &= -\int \bar{q}_j n_j \mathrm{d}S - \int \tau_{hu_j} n_j \mathrm{d}S - \int \frac{\partial \bar{q}_i^r}{\partial x_i}\mathrm{d}V \\ &\quad + \int \left(\frac{\partial \bar{P}}{\partial t} + \tilde{u}_j \frac{\partial \bar{P}}{\partial x_j}\right)\mathrm{d}V + \int \overline{\tau_{ij}\frac{\partial u_i}{\partial x_j}}\mathrm{d}V. \end{aligned} \tag{7.14}$$

The above equation is derived by starting with the total internal energy equation, subtracting the mechanical energy equation and enforcing the variable density continuity equation. Note that the above equation includes possible source terms due to thermal radiatitive transport, viscouss dissipation, and pressure work.

The simple Fickian diffusion velocity approximation is assumed, so that the mean diffusive heat flux vector $\bar{q}_j$ is

$$\bar{q}_j = -\overline{\left[\frac{\mu}{\mathrm{Pr}}\frac{\partial h}{\partial x_j} - \frac{\mu}{\mathrm{Pr}}\sum_{k=1}^{K} h_k \frac{\partial Y_k}{\partial x_j}\right]} - \overline{\frac{\mu}{\mathrm{Sc}}\sum_{k=1}^{K} h_k \frac{\partial Y_k}{\partial x_j}}. \tag{7.15}$$

If Sc = Pr, i.e., unity Lewis number (Le = 1), then the diffusive heat flux vector simplifies to $\bar{q}_j = -\frac{\mu}{\mathrm{Pr}}\frac{\partial \tilde{h}}{\partial x_j}$. The viscous dissipation term is closed by

$$\begin{aligned}\overline{\tau_{ij}\frac{\partial u_i}{\partial x_j}} &= \left((\mu + \mu_t)\left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right) - \frac{2}{3}\left(\bar{\rho}\tilde{k} + \mu_t \frac{\partial \tilde{u}_k}{\partial x_k}\right)\delta_{ij}\right)\frac{\partial \tilde{u}_i}{\partial x_j} \\ &= \left[2\mu\tilde{S}_{ij} + 2\mu_t\left(\tilde{S}_{ij} - \frac{1}{3}\tilde{S}_{kk}\delta_{ij}\right) - \frac{2}{3}\bar{\rho}\tilde{k}\delta_{ij}\right]\frac{\partial \tilde{u}_i}{\partial x_j}. \end{aligned} \tag{7.16}$$

The turbulent diffusive flux vector τ_{hu_j} in Equation 7.14 is defined as

$$\tau_{hu_j} \equiv \bar{\rho}\left(\widetilde{hu_j} - \tilde{h}\tilde{u}_j\right). \tag{7.17}$$

For RANS simulations, τ_{hu_j} represents the turbulent energy diffusive flux vector and is simplified to the form $\tau_{hu_j} = \overline{\rho h'' u_j''}$ by substitution of the Favre decomposition of each variable. It is then modeled by

$$\tau_{hu_j} = \overline{\rho h'' u_j''} = -\frac{\mu_t}{\mathrm{Pr}_t}\frac{\partial \tilde{h}}{\partial x_j}, \tag{7.18}$$

where Pr_t is the turbulent Prandtl number and μ_t is the modeled turbulent eddy viscosity from momentum closure. For LES, τ_{hu_j} represents the subgrid turbulent energy diffusive flux vector, and is modeled in the same way as

$$\tau_{hu_j} = -\frac{\mu_t}{\mathrm{Pr}_t}\frac{\partial \tilde{h}}{\partial x_j}, \tag{7.19}$$

where Pr_t is the subgrid turbulent Prandtl number and μ_t is the modeled subgrid turbulent eddy viscosity from momentum closure.

The resulting filtered and modeled turbulent energy equation for both RANS and LES is given by,

$$\begin{aligned}\int \frac{\partial \bar{\rho}\tilde{h}}{\partial t}\mathrm{d}V + \int \bar{\rho}\tilde{h}\tilde{u}_j n_j \mathrm{d}S &= \int \left(\frac{\mu}{\mathrm{Pr}} + \frac{\mu_t}{\mathrm{Pr}_t}\right)\frac{\partial \tilde{h}}{\partial x_j} n_j \mathrm{d}S - \int \frac{\partial \bar{q}_i^r}{\partial x_i}\mathrm{d}V \\ &+ \int \left(\frac{\partial \bar{P}}{\partial t} + \tilde{u}_j \frac{\partial \bar{P}}{\partial x_j}\right)\mathrm{d}V + \int \overline{\tau_{ij}\frac{\partial u_j}{\partial x_j}}\mathrm{d}V. \end{aligned} \tag{7.20}$$

The turbulent Prandtl number must have the same value as the turbulent Schmidt number for species transport to maintain unity Lewis number.

Review of Prandtl, Schmidt and Unity Lewis Number

For situations where a single diffusion coefficient is applicable (e.g., a binary gas system) the Lewis number is defined as:

$$\mathrm{Le} = \frac{\mathrm{Sc}}{\mathrm{Pr}} = \frac{\alpha}{D}. \tag{7.21}$$

If the diffusion rates of energy and mass are equal,

$$\mathrm{Sc} = \mathrm{Pr} \text{ and } \mathrm{Le} = 1. \tag{7.22}$$

For completeness, the thermal diffusivity, Prandtl and Schmidt number are defined by,

$$\alpha = \frac{\kappa}{\rho c_p}, \tag{7.23}$$

$$\mathrm{Pr} = \frac{c_p \mu}{\kappa} = \frac{\mu}{\rho \alpha}, \tag{7.24}$$

and

$$\mathrm{Sc} = \frac{\mu}{\rho D}, \tag{7.25}$$

where c_p is the specific heat, κ, is the thermal conductivity and α is the thermal diffusivity.

7.1.5 Multi-physics

The code base allows for multi-physcis coupling including congugate heat transfer, i.e., coupling of fluid mechanics to a thermal heat conduction physics and coupling to a participating media radiation physics. More detail can be found in [14]

7.2 Discretization

Nalu supports two discretizations: control volume finite element and (CVFEM) edge-based vertex centered (EBVC). Each are finite volume forumations and each solve for the primitives are are each considered vertex-based schemes. For large eddy simulation, a low dissipation skew symmetric scheme has been developed (see subsequent section).

For generalized unstructured meshes that have poor quality, CVFEM has been shown to excell in accuracy and robustness. This is mostly due to the inhearant accuracy limitation for the non-orthogonal correction terms that appear in the diffusion term and pressure stabilization for the EBVC scheme. For generalized unstructured meshes of decent quality, either scheme is ideal. Finally, for highly structured meshes with substantail aspect ratios, the edge-based scheme is ideal.

In general, the edge-based scheme is at least two times faster per iteration than the element-based scheme. For some classes of flows, it can be up to four times faster. However, due to the lagged coupling between the projected nodal gradient equation and the dofs, on meshes with high non-orthogonality, nonlinear residual convergence can be delayed.

The CVFEM scheme lends itself to promotion to higher order methods (greater than second order) while the edge-based scheme is limited to spatially second order. More detail on the supported formulations with detailed description of the comparison of CVFEM low order, EBVC low order and third order CVFEM can be found in [13].

Finally, time integration is managed thorugh an implicit Backward Euler (first order accurate) and a BDF2 second order temporal accurate scheme.

7.3 Low Dissipation Operators

In general, advection for both the edge and element-based scheme is identical with standard exception of the location of the integration points. The full advection term is simply written as,

$$ADV_{\phi} = \int \rho u_j \phi_{ip} A_j = \sum \dot{m} \phi_{ip}, \tag{7.26}$$

where ϕ is u_i, Z, h, etc.

The evaluation of ϕ_{ip} defines the advection stabilization choice. In general, the advection choice is a cell Peclet blending between higher order upwind (ϕ_{upw}) and a generalized unstabilized central (Galerkin) operator, ϕ_{gcds},

$$\phi_{ip} = \eta \phi_{upw} + (1 - \eta) \phi_{gcds}. \tag{7.27}$$

In the above equation, η is a cell Peclet blending, The generalized central operator can take on a pure second order or psuedo fourth order form (see below). In general, a hybrid upwind factor, γ, can be used to ensure that no stabilization is added ($\eta = 0$) or that full upwind stabilization is included (as will be shown, even with limiter functions). The hybrid upwind factor allows one to modify the functional blending function; values of unity result in the normal blending function response; values of zero yield a pure central operator, i.e., blending function of zero; values $>>$ unity result in a blending function value of unity, i.e., pure upwind. The constant A is implemented as above with a value of 5. This value can not be changed via the input file.

The cell-Peclet number is computed for each sub-face in the element from the two adjacent left (L) and right (R) nodes,

$$\mathrm{Pe} = \gamma \frac{\frac{1}{2}\left(u_{R,i} + u_{L,i}\right)\left(x_{R,i} - x_{L,i}\right)}{\nu}. \tag{7.28}$$

A dot-product is implied by repeated indices. The cell Peclet blending function is, therefore, the following:

$$\eta = \frac{\mathrm{Pe}^2}{5 + \mathrm{Pe}^2}. \tag{7.29}$$

The upwind operator, ϕ_{upw} is computed based on a blending of the extrapolated state (using the projected nodal gradient) and the linear interpolated state. Second or third order upwind is provided based on the value of α_{upw} blending

$$\begin{aligned} \phi_{upw} = {} & \alpha_{upw} \tilde{\phi}^{L}_{upw} + (1 - \alpha_{upw}) \phi_{cds}; \dot{m} > 0, \\ & \alpha_{upw} \tilde{\phi}^{R}_{upw} + (1 - \alpha_{upw}) \phi_{cds}; \dot{m} < 0. \end{aligned} \tag{7.30}$$

The extrapolated value based on the upwinded left (ϕ^L) or right (ϕ^R) state,

$$\begin{aligned}
\tilde{\phi}^L_{upw} &= \phi^L + d^L_j \frac{\partial \phi^L}{\partial x_j}, \\
\tilde{\phi}^R_{upw} &= \phi^R - d^R_j \frac{\partial \phi^R}{\partial x_j}.
\end{aligned} \tag{7.31}$$

The distance vectors are defined based on the distances between the L/R points and the integration point (for both edge or element-based),

$$\begin{aligned}
d^L_j &= x^{ip}_j - x^L_j, \\
d^R_j &= x^R_j - x^{ip}_j.
\end{aligned} \tag{7.32}$$

In the case of all transported quantities, a Van Leer limiter of the extrapolated value is supported and can be activated withing the input file (using the solution options “limiter” specification).

Second order central is simply written as a linear combination of the nodal values,

$$\phi_{cds} = \sum N^{ip}_k \phi_k. \tag{7.33}$$

where N^{ip}_k is either evaluated at the subcontrol surface or edge midpoint. In the case of the edge-based scheme, the edge midpoint evaluation provides for a skew symmetric form of the operator.

The generalized central difference operator is provided by blending with the extrapolated values and second order Galerkin,

$$\phi_{gcds} = \frac{1}{2}\left(\hat{m}\phi^L_{upw} + \hat{m}\phi^R_{upw}\right), \tag{7.34}$$

where,

$$\begin{aligned}
\hat{m}\phi^L_{upw} &= \alpha \tilde{\phi}^L_{upw} + (1-\alpha)\,\phi_{cds}, \\
\hat{m}\phi^R_{upw} &= \alpha \tilde{\phi}^R_{upw} + (1-\alpha)\,\phi_{cds}.
\end{aligned} \tag{7.35}$$

The value of α provides the type of psuedo fourth order stencil and is specified in the user input file.

The above set of advection operators can be used to define an idealized one dimensional stencil denoted by ($i-2$, $i-1$, i, $i+1$, $i+2$), where i represents the particular row for the given transported variable. Below, in Table 7.1 the stencil can be noted for each value of α and α_{upw}.

It is noted that by varying these numerical parameters, both high quality, low dissipation operators suitable for LES usage or limited, monotonic operators suitable for RANS modeling can be accomodated.

Table 7.1: Edge-based stencils for various combinations of α and α_{upw}

$i-2$	$i-1$	i	$i+1$	$i+2$	α	α_{upw}
0	$-\frac{1}{2}$	0	$+\frac{1}{2}$	0	0	n/a
$+\frac{1}{8}$	$-\frac{6}{8}$	0	$+\frac{6}{8}$	$-\frac{1}{8}$	$\frac{1}{2}$	n/a
$+\frac{1}{12}$	$-\frac{8}{12}$	0	$+\frac{8}{12}$	$-\frac{1}{12}$	$\frac{2}{3}$	n/a
$+\frac{1}{4}$	$-\frac{5}{4}$	$+\frac{3}{4}$	$+\frac{1}{4}$	0	$\dot{m} > 0$	1
0	$-\frac{1}{4}$	$-\frac{3}{4}$	$+\frac{5}{4}$	$-\frac{1}{4}$	$\dot{m} < 0$	1
$+\frac{1}{6}$	$-\frac{6}{6}$	$+\frac{3}{6}$	$+\frac{2}{6}$	0	$\dot{m} > 0$	$\frac{1}{2}$
0	$-\frac{2}{6}$	$-\frac{3}{6}$	$+\frac{6}{6}$	$-\frac{1}{6}$	$\dot{m} < 0$	$\frac{1}{2}$

7.4 Reacting Flow

Reacting flow in the code base is provided through three avenues including infinitely-fast Burke-Schumann chemistry, Strained, Steady Laminar-Flamelet model (SSLF) and finite-rate chemistry. Two of the available approaches for reacting flows, Burke-Schumann chemistry and SSLF, rely on the separation of time scales between fluid mixing and chemical reactions and area appropriate when chemical mixing is the limiting process. The third approach with finite-rate chemistry is appropriate when chemistry is the limiting process as in premixed systems. Although implemented, this project did not fully exercise the reacting flow capability.

7.4.1 Infinitely Fast Burke-Schumann Chemistry

In gas-phase systems Pr and Sc are of order unity for most species. If combinations of species can be taken that have similar transport properties and also represent conserved quantities, like element conservation, significant simplifications arise. If the chemical reactions occuring can be considered to be inifinitely fast, then greater simplifications arise. This section addresses the limit where the transport properties are sufficiently similar and the chemistry is represented by infinitely fast and irreversible reactions. In this limit, the mass fractions can all be expressed as linear functions of the mixture fraction between the two source streams except for a slope change where reactions occur. The next section addresses the limit of similar transport properties and either equilibrium (fast and reversible) chemistry or chemistry limited by a mixing time scale.

For the purposes of demonstration we make the assumption that the fuel is a hydrocarbon for which the global reaction is

$$\mathrm{C_mH_n} + \nu_{\mathrm{O_2}}\mathrm{O_2} \rightarrow \mathrm{mCO_2} + (\mathrm{n}/2)\mathrm{H_2O} \tag{7.36}$$

where $\nu_{\mathrm{O_2}} = m + n/4$ is the stoichiometric coefficient for the oxidizer relative to the fuel. Linear combinations of the form

$$\beta = \frac{Y_{\mathrm{O_2}}}{\nu_{\mathrm{O_2}} W_{\mathrm{O_2}}} - \frac{Y_{\mathrm{C_mH_n}}}{W_{\mathrm{C_mH_n}}} \qquad \text{or} \qquad \beta_{\mathrm{CO_2}} = \frac{Y_{\mathrm{CO_2}}}{m W_{\mathrm{CO_2}}} - \frac{Y_{\mathrm{C_mH_n}}}{W_{\mathrm{C_mH_n}}} \tag{7.37}$$

are conserved scalars. The mixture fraction is generally defined to be that conserved scalar that is normalized to unity in the fuel stream and zero in the oxidizer stream (denoted by subscripts 1 and 0, respectively) so that

$$Z = \frac{\beta - \beta_0}{\beta_1 - \beta_0} \tag{7.38}$$

is the standard definition of the mixture fraction for any linear combination β that defines a conserved scalar. The flame is located at a fixed value of the mixture fraction, the stoichiometric mixture fraction, Z_{st}, where $\beta = 0$. From Eq. 7.38 using β from Eq. 7.37 one obtains

$$Z_{st} = \left[1 - \frac{\beta_1}{\beta_0}\right]^{-1} = \left[1 + \frac{\nu_{\mathrm{O_2}} Y_{\mathrm{C_mH_n},1} W_{\mathrm{O_2}}}{Y_{\mathrm{O_2},0} W_{\mathrm{C_mH_n}}}\right]^{-1}. \tag{7.39}$$

In the limit of irreversible, infinitely-fast, single-step kinetics, fuel and oxidizer do not coexist and the reaction is a Dirac delta function at Z_{st}, the profiles are linear except for this point. These relations are different for rich and lean mixtures. For $Z \leq Z_{st}$ one obtains

$$\begin{aligned} Y_{\mathrm{C_mH_n}} &= 0 \\ Y_{\mathrm{O_2}} &= Y_{\mathrm{O_2},0}(Z_{st} - Z) \end{aligned} \tag{7.40}$$

and for $Z \geq Z_{st}$ one obtains

$$\begin{aligned} Y_{\mathrm{C_mH_n}} &= Y_{\mathrm{C_mH_n},1}(Z - Z_{st})/(1 - Z_{st}) \\ Y_{\mathrm{O_2}} &= 0. \end{aligned} \tag{7.41}$$

The mass fractions of $\mathrm{CO_2}$ and $\mathrm{H_2O}$ can then be expressed as

$$\begin{aligned} Y_{\mathrm{CO_2}} &= m(W_{\mathrm{CO_2}}/W_{\mathrm{C_nH_m}})\,(Y_{\mathrm{C_mH_n},1}Z - Y_{\mathrm{C_mH_n}}) \\ Y_{\mathrm{H_2O}} &= (n/2)(W_{\mathrm{H_2O}}/W_{\mathrm{C_nH_m}})\,(Y_{\mathrm{C_mH_n},1}Z - Y_{\mathrm{C_mH_n}}) \end{aligned} \tag{7.42}$$

Initially and in adiabatic systems the enthalpy will also be linear in the mixture fraction

$$h = h_0 + (h_1 - h_0)Z \tag{7.43}$$

although participating media radiation and wall heat transfer alter this relation. The relationship between the temperature and enthalpy is

$$h = \sum_i Y_i(Z)h_i(T) \tag{7.44}$$

where h is the enthalpy as evolved by the fluids solve as per Eq. 7.14 (not necessarily adiabatic) and

$$h_i(T) = h_i^0 + \int_{T^0}^{T} c_{p,i}\,dT \tag{7.45}$$

splits the enthalpy into a reference chemical enthalpy and a sensible enthalpy where the superscript 0 indicates properties evaluated at a reference temperature, T^0. The fluid solve evolves the enthalpy, but since it is typically easier to express boundary conditions in terms of the temperature, the initial adiabatic enthalpy values can be determined from direct evaluation of Eq. 7.44. The temperature is obtained by nonlinear solution of this same equation given the enthalpy field from the fluids solve. Relations such as these translate the unknown space from $(\vec{Y},h)$ into a need for simply (Z,h). Similar relations will be described later for more sophisticated treatments.

The mixture molecular mass can be expressed in terms of the species molar masses, W_i, as

$$\bar{W} = \left(\sum_i \frac{Y_i}{W_i} \right)^{-1} \tag{7.46}$$

and this is used to obtain the density

$$\rho = \frac{p\bar{W}}{RT} \tag{7.47}$$

where p is the pressure and R is the ideal gas constant. There is a nonlinearity in the expression for the mixture molecular mass that can be significant if there are strong molar mass differences.

7.4.2 Tabulated State Variables

For equilibrium or strained laminar flamelet models (SLFM) characterized by a mixing rate, Nalu takes advantage of tabulated state variables where the tabulation can be a function of the mixture fraction, the enthalpy and the scalar dissipation rate, which provides the mixing time scale. These state relations are generated externally and stored in HDF5 tables that Nalu reads. Table generation can be performed using the Sierra/Thermal-Fluids code tabular_props. Inputs for the table with the SLFM come from solution of the flamelet equations described in the following subsection; the output is filtered in the tabulation process to provide the quantities evolved in LES as described afterward.

The Flamelet Equations

The species conservation equations can be transformed in several ways [82, 34, 53] to obtain a set of equations for which the independent spatial variables are replaced by an independent mixture fraction variable. The resulting transformed equation is

$$\frac{d\rho Y_i}{dt} - \frac{\rho\chi}{2}\frac{d^2 Y_i}{dZ^2} = W_i \sum_j \nu_{i,j}\omega_j(\vec{Y},T). \tag{7.48}$$

where $\chi = 2D_Z|\nabla Z|^2$ is the scalar dissipation rate. Similarly, an energy equation can be written

$$\frac{d\rho c_p T}{dt} - \frac{\rho c_p \chi}{2}\frac{d^2 T}{dZ^2} - \left[\frac{dc_p}{dZ} + \sum_i (c_p - c_{p,i}) \frac{dY_i}{dZ} \right] \frac{dT}{dZ} = -\sum_i W_i h_i \sum_j \nu_{i,j}\omega_j(\vec{Y},T). \tag{7.49}$$

The primary advantage of this form of the equations is that the stiff reacting flow problem is reduced to a lower dimensional space. The connection to the fluid mixing is through the scalar dissipation rate, $\chi = 2D|\nabla Z|$, and the fluid solve evolves only those variables that evolve over the time scales of the flow, typically the filtered mixture fraction and enthalpy.

Tables are generated by solving Eqs. 7.48 and 7.49 for a range of boundary conditions and χ, and the resulting solutions are converted into multidimensional HDF5 table format using tabular_props.

The scalar dissipation rate varies across the mixture fraction range, and flamelets are tabulated using a reference value, χ_0, that is usually the stoichiometric dissipation rate, and a functional form in the mixture fraction coordinate, $F_\chi(Z)$. A functional form often employed is

$$F_\chi(Z) = \exp\left(-2[\mathrm{erfc}^{-1}(2Z)]^2\right) \tag{7.50}$$

that corresponds to the (constant density) counterflow configuration. Given χ_0, the scalar dissipation profile in the Z coordinate is

$$\chi(Z) = \chi_0 \frac{F_\chi(Z)}{F_\chi(Z_0)}. \tag{7.51}$$

Filtered variables

In turbulent simulations, a Favre filtered form of the governing equations is evolved. A Favre-filtered quantity is represented by $\tilde{\phi} \equiv \overline{\rho\phi}/\bar{\rho}$, and Favre-filtered variables in the flamelet library must be calculated. This is performed by convoluting the variables with a presumed PDF for the mixture fraction. Such a convolution can be costly and therefore filtered values will be precomputed and stored in another flamelet library. Filtering the variables in the flamelet library takes the form

$$\tilde{\phi}(\tilde{Z}, \widetilde{Z''^2}, \chi_0, ...) = \frac{\int_0^1 \rho(Z, \chi_0, ...)\phi_T(Z, \chi_0, ...)\; P_Z(Z; \tilde{Z}, \widetilde{Z''^2})\, \mathrm{d}Z}{\int_0^1 \rho(Z, \chi_0, ...)\; P_Z(Z; \tilde{Z}, \widetilde{Z''^2})\, \mathrm{d}Z}. \tag{7.52}$$

Here the elipsis refers to the list of any other independent variables and it is understood that these can be expressed as function of the filtered variables evolved by the fluid region as $\chi_0 = \chi_{st}(\tilde{Z}, \widetilde{Z''^2}, \tilde{\chi}, ...)$.

The independent variables for such a library would include $\tilde{Z}$ and $\widetilde{Z''^2}$, which would replace Z in the original library. The remainder of the independent variables would remain the same (i.e. χ_0). The latter greatly facilitates the maintenance of the table as an orthogonal hypercube since parameterizing in terms of moments of, for example, χ could result in an oddly shaped, non-orthogonal table. Instead $\tilde{\chi}$ is computed as an output of the filtered table. Since the fluid region evolves (or knows how to find) $\tilde{\chi}$, there is a reverse translation to get variables like χ_{st} in which the library is parameterized described below.

Accessing Flamelet Libraries

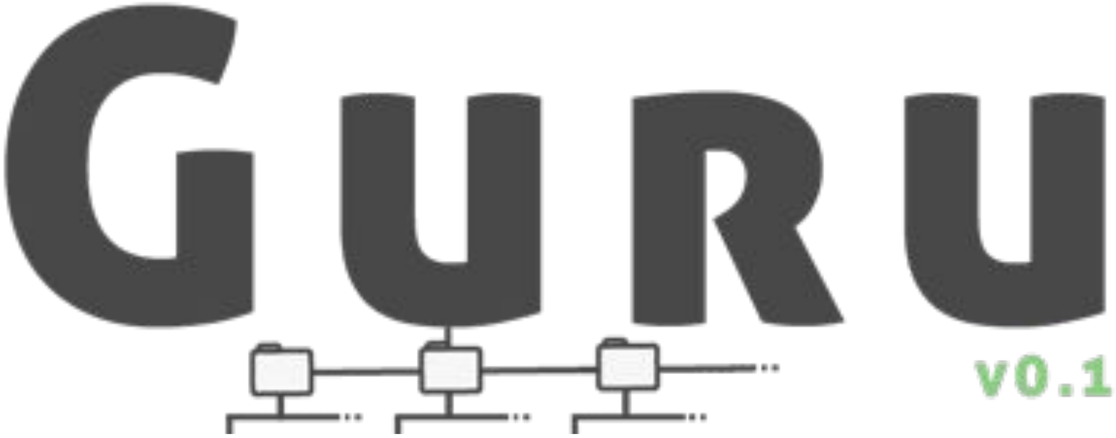
As described above, the flamelet library is generated with independent variables being filtered quantities. The independent variables $\tilde{Z}$ and $\widetilde{Z''^2}$ correspond to table coordinates in a straightforward manner. The table coordinate associated with the variance is scaled by $(1-\tilde{Z})\tilde{Z}$ that gives the upper bound for the variance to provide a structured range of values for that table dimension. In order to generate orthogonal and structured tables, some independent variables like χ are tabulated using reference values, χ_0. The mapping from the filtered $\tilde{\chi}$ to χ_0 is done by taking advantage of the known function, $F_\chi(Z)$. The filtered value of this function is also tabulated and the reference dissipation rate is obtained from

$$\chi_0(\tilde{Z},\widetilde{Z''^2},\tilde{\chi}) = \tilde{\chi}\frac{F_\chi(Z_0)}{\widetilde{F_\chi}(\tilde{Z},\widetilde{Z''^2})}. \tag{7.53}$$

This χ_0 is the resulting independent table coordinate in the dissipation direction.

Chapter 8

Guru: A UQ study management framework



8.1 Introduction

Guru provides tools for constructing, maintaining, and analyzing uncertainty quanitfication (UQ) studies, aimed to support common UQ workflows, reduce analyst effort, and enable more complex studies. It is written purely in Python and requires only a stable build of Python 2.6 or 2.7 for installation; there are no mandatory third-party dependencies.

In its current form, *Guru*'s public interface is the **guru** command-line program. The plan is to launch the next minor version of the framework with a public application program interface (API) in Python. *Guru* is written purely in Python and adopts the standard packaging and distribution model for a third-party Python package, including a setup script, *setup.py*, that handles building and installing.

8.2 Installation

After aquiring a guru-v**X**.**X**.**X**.tar.gz archive, we prepare for installation with the following commands (we are using a Bash shell hereforth):

```
$ tar zxvf guru-v0.1.tar.gz  # unpack the source package
$ cd guru-v0.1               # move into the source pacakge
```

Basic Installation

Building and installing *Guru* can be done in one command:

```
$ python setup.py install  # install guru!
```

Installing Locally

However, if we do not have permission to build or install *Guru* as a global package (this is often the case with an account on a shared computing resources) or prefer not to, we can utilize any one of the alternative methods that can be found in Python's official, web-accessible documentation for installing Python modules. Our preferred alternative is to install with the --*user* option.

```
$ python setup.py install --user
```

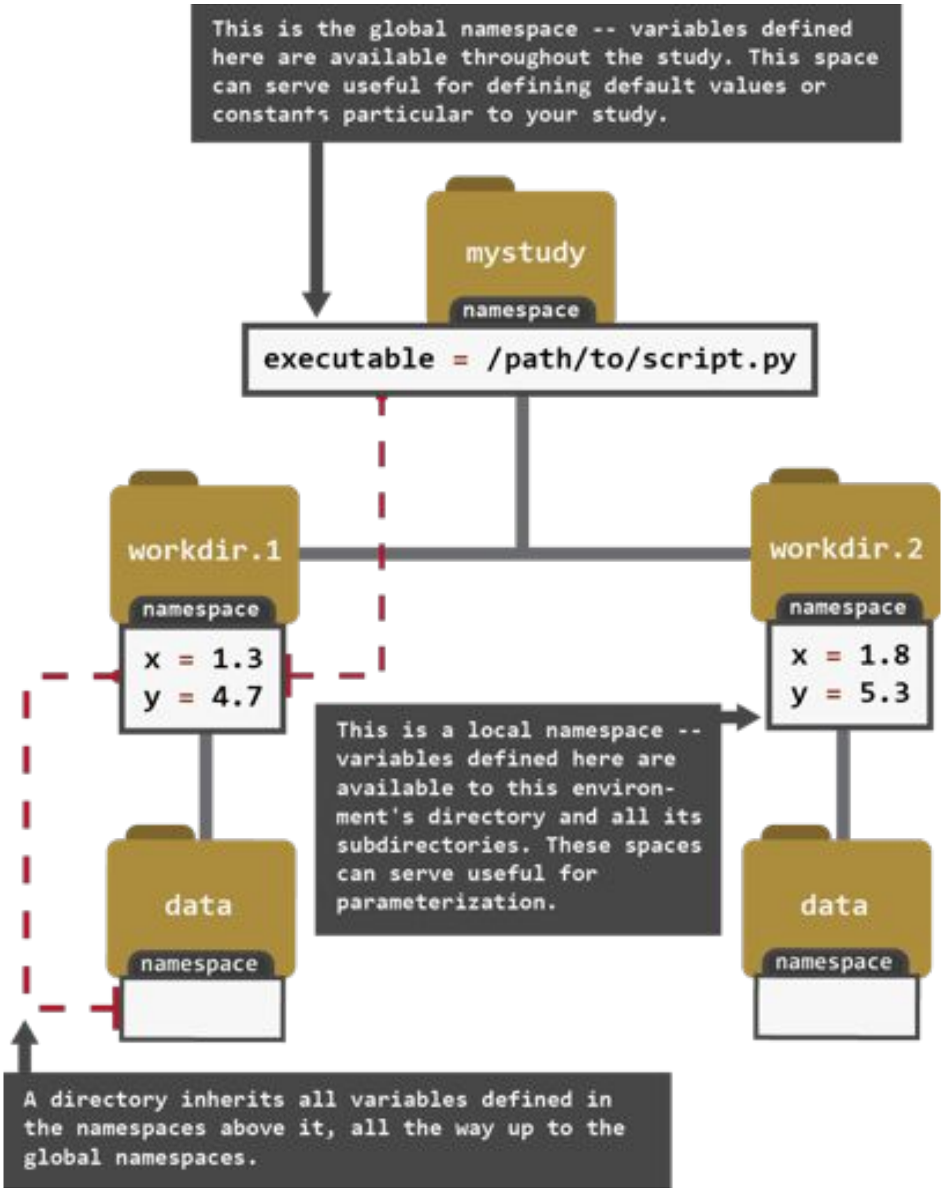
When installed locally, the command-line program is installed into a subdirectory of *site.USER_BASE* (see interactive python session below) and the framework to *site.USER_SITE*. To ensure access to these resources, we append the former directory to the **PATH** environmental variable and the latter to **PYTHONPATH**, within our startup script (for a Bash shell, use either .bash_profile or .bashrc).

```
$ python
>> import site
>> # Append the below to PATH
>> print(site.USER_BASE + '/bin')
/path/to/local/bin
>> # Append the below to PYTHONPATH
>> print(site.USER_SITE)
/path/to/local/lib/python/site-packages
```

8.3 The Environment

Upon initialization with a study, Guru creates a new subdirectory named *.guru* directly underneath the study's top-level directory (referred to as **study root** hereafter). The study root and all its subdirectories, regardless whether created before or after initialization, are provided a **namespace**, a container for mapping names to values (i.e. variables). Each namespace is dynamic – we can (re-)declare and delete variables as the study progresses – and persistant – the state of a study's environment is saved within the study's .guru directory as a pickled Python dictionary. Furthermore, a hierarchal relationship exists between a directory's namespace and its parent directory. Working upwards and terminating at the study root, Guru searches this hierarchy to resolve name-to-value lookups, returning the most immediate value mapped to the target name. The study root's namespace is refered to as the **global namespace** since all names mapped here could resolve anywhere within the study.

Figure 8.1: Example of Guru environment



For example, take the directory tree described in figure 8.1. In this case, the study root is *mystudy*, and we've assigned the value '/path/to/script.py' to the name 'executable' in the global namespace. Also, the names 'x' and 'y' are each mapped to different values in *workdir*.1 and *workdir*.2's local namespaces, forming some sense of a parametric study. Refering back to 'executable', A name-to-value loopkup for this name within any subdirectory of the study (e.g. the *data* directory underneath *workdir*.1) would always yield '/path/to/script.py' defined at the global namespace, the most immediate ancestor with a mapping to that key, making 'executable' a de facto **study constant**. The value of the names 'x' and 'y' depend on which immediate subdirectory of the study root – workdir.1 or workdir.2 – we are under.

Referencing the Environment

Guru utilizes the environment to perform preprocessing tasks.

8.4 The Command-Line Tool

In the usage examples below, we construct and perform a few management tasks for the study defined in section 8.3. The only setup needed is to create the *mystudy* directory, which, for this example, will be underneath our home directory (∼).

```
$ # current path: ~
$ mkdir mystudy && cd mystudy
$ # current path: ~/mystudy
```

8.4.1 guru init

We initilize Guru with the subcommand **guru init**. This creates a *.guru* directory inside the current directory.

```
$ # current path: ~/mystudy
$ guru init
Initialized Guru repository in ~/mystudy/.guru.
```

8.4.2 guru setpars

We now want to create the *workdir*.1 and *workdir*.2 directories and appropriately set the 'x' and 'y' values inside each directory's namespace as shown in Figure 8.1. **guru setpars** can help with both tasks. Below is the general usage of this command:

```
$ guru setpars [OPTIONS] inputfile pattern [pattern ...]
```

The Input File

The input file is where we define the study's parameter space. This file should contain a table of (in general) m rows by n columns – each column corresponds to a parameter, and each row, consisting of n values (one for each parameter), holds the values of the parameters for one directory.

The first line may contain the parameter names. Alternatively, the parameter names can be passed through the command-line with the $-n/--name$ option. One of these two methods for naming parameters must be used (we recommend the former as it also serves as a method of documentaion).

For our example, we'll name our input file *pars.in* and place the text below inside it – lines that start with '#' are considered comments and are ignored.

```
# file: ~/mystudy/pars.in
x    y
1.3  4.7
1.8  5.3
```

Pattern

The second argument is the pattern – or patterns – for naming the directories associated to each row.

Chapter 9

Conclusion

This report has described several activities undertaken to develop UQ-based calibration methods for LES models and make these methods more accessible for complex flows via open source software. When taken as a whole, several important conclusions can be drawn. The first set of conclusions come in the form of new capabilities which have been developed and tested in the course of this work. Primary amongst them is that engineering LES models can be calibrated using high-fidelity simulation data. Importantly, the LES calibration process is more difficult than many other calibration processes because the models must encompass the range of turbulent fluctuations and the quantities of interest are often non-smooth. This report describes methods to overcome these challenges. Complementary to the conclusion that LES can be calibrated is the conclusion that forward propagation of uncertainty through LES can also be performed. As such, the impact of uncertainties at the modeling level on engineering quantities of interest can be determined. In terms of innovations, the final conclusion is that mesh and numerical method uncertainty are often more significant than model uncertainty, and new ways to quantify them have begun to be developed.

In addition to the positive conclusions mentioned above, there is also a significant negative conclusion that has been reached: model form error for LES must be captured for simulations to be predictive. One example is that most LES models are strictly dissipative in nature, but as meshes are coarsened, it is known that the amount of backscatter of turbulent energy increases. It is hypothesized that models which cannot represent backscatter will have high errors in predicting certainty quantities, for example the location and intensity of the near-wall peak in streamwise fluctuations. Further, because mesh changes must be accounted for in the LES models, existing approaches to solution verification using mesh refinement are inappropriate. Methods to determine innovative approaches to model formulation would be needed to have any chance at overcoming these challenges.

The final set of conclusions from this work come in the arena of open source software development. In order to successfully collaboration with external organizations, it is necessary for Sandia to have codes and support packages which are amenable to research and can be fully examined by outside partners. To faciliate these multi-institutional efforts, this work has significantly added to the capabilities of the Nalu low-Mach number flow solver. It as been shown to be accurate on unstructured meshes and scale well on large numbers of processors. New capabilities have also been added for modeling reacting flows of interest to industry. This work has also demonstrated the ability to manage the large, coordinated sampling process required for UQ. Studies can be created and executed much more easily than previously possible, as well as monitored and altered while running to ensure the needed data is obtained.

As an overall conclusion, many steps have been taken towards realizing the goal of engineering LES with quantified cost versus accuracy for flows of technological interest. Calibration and uncertainty propagation methods suitable for engineering LES have been developed and demonstrated in a variety of canonical flows. Open source software has been created to support collaborations between Sandia and industry on problems of mutual interest. Finally, the next research steps have been identified that are necessary to obtaining the overall goal to improve engineering when turbulent flows are present.

References

[1] R.J. Adrian. Stochastic estimation of the structure of turbulent fields. Technical Report TAM Report No. 800, UILU-ENG-95-6022, Theoretical and Applied Mechanics, University of Illinois, 1995.

[2] J.S. Baggett, J. Jimenez, and A.G. Kravchenko. Resolution requirements in large-eddy simulation of shear flows. In *Center for Turbulence Research Annual Research Briefs*, 1997.

[3] A. Bhattacharya, A. Das, and R.D. Moser. A filtered-wall formulation for large-eddy simulation of wall-bounded turbulence. *Physics of Fluids*, 20:115104, 2008.

[4] J Borggaard and T. Iliescu. Approximate deconvolution boundary conditions for large eddy simulation. *Applied Mathematics Letters*, 19(9):735–740, 2006.

[5] S. Bose, P. Moin, and D. You. Grid-independent large-eddy simulation using explicit filtering. *Physics of Fluids*, 22:105103, 2010.

[6] M.D. Buhmann. *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, Cambridge, 2003.

[7] S.H. Cheung, T.A. Oliver, E.E. Prudencio, S. Prudhomme, and R.D. Moser. Bayesian uncertainty analysis with applications to turbulence modeling. *Reliability Engineering and System Safety*, 96:1137–1149, 2011.

[8] M. Christie, V. Demyanov, and D. Erbas. Uncertainty quantification for porous media flows. *Journal of Computational Physics*, 217:143–158, 2006.

[9] P.M. Congedo, C. Duprat, G. Balarac, and C. Corre. Numerical prediction of turbulent flows using Reynolds-averaged Navier-Stokes and Large-Eddy Simulations with uncertain inflow conditions. *International Journal for Numerical Methods in Fluids*, 72(3):341–358, 2012.

[10] B.J. Debusschere, H.N. Najm, P.P. Pébay, O.M. Knio, R.G. Ghanem, and O.P. Le Maître. Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM Journal on Scientific Computing*, 26(2):698–719, 2004.

[11] S. Deck, N. Renard, R. Laraufie, and P. Sagaut. Zonal detached eddy simulation (zdes) of a spatially developing flat plate turbulent boundary layer over the reynolds number range 3 150 $<= re_\theta <=$ 14 000. *Physics of Fluids*, 26(2):pp. 025116, 2014.

[12] Petros Dellaportas and David A. Stephens. Bayesian Analysis of Errors-in-Variables Regression Models. *Biometrics*, 51(3):pp. 1085–1095, 1995.

[13] S. Domino. A comparison between low-order and higher-order low-mach discretization approaches. Technical report, Proceedings of the 2014 Center for Turbulence Research Summer Program, July 2014.

[14] S. Domino. Nalu theory manual. Technical Report SAND2015-3107W, Sandia National Laboratories, Albuquerque, NM, May 2015.

[15] K. Duraisamy and P. Chadrashekar. Goal-oriented uncertainty propagation using stochastic adjoints. *Computers and Fluids*, 66:10–20, 2012.

[16] M. Emory, J. Larsson, and G. Iaccarino. Modeling of structural uncertainties in Reynolds-averaged navier-stokes closures. *Physics of Fluids*, 25:110822, 2013.

[17] O.G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On The Convergence of Generalized Polynomial Chaos Expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:317–339, 2012.

[18] Andrew Fox, Mathew Williams, Andrew D. Richardson, David Cameron, Jeffrey H. Gove, Tristan Quaife, Daniel Ricciuto, Markus Reichstein, Enrico Tomelleri, Cathy M. Trudinger, and Mark T. Van Wijk. The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. *Agricultural and Forest Meteorology*, 149(10):1597–1615, 2009.

[19] R.G. Ghanem and P.D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer Verlag, New York, 1991.

[20] A. F. Ghoniem, M. C. Soteriou, O. M. Knio, and B. Cetegen. Effect of steady and periodic strain on unsteady flamelet combustion. *pci*, 24:223–230, 1992.

[21] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, 1996.

[22] J. Gullbrand and F.K. Chow. The effect of numerical errors and turbulence models in large-eddy simulations of channel flow, with and without explicit filtering. *Journal of Fluid Mechanics*, 495:323–341, 2003.

[23] H. Haario, E. Saksman, and J. Tamminen. An adaptive Metropolis algorithm. *Bernoulli*, 7:223–242, 2001.

[24] J. C. Hewson. *Pollutant Emissions from Nonpremixed Hydrocarbon Flames*. PhD thesis, Univ. of California, San Diego, La Jolla, CA, 1997.

[25] J. C. Hewson. An extinction criterion for nonpremixed flames subject to brief periods of high dissipation rates. *cf*, 160(5):887–897, 2013.

[26] J. C. Hewson and A. R. Kerstein. Stochastic simulation of transport and chemical kinetics in turbulent $CO/H_2/N_2$ flames. *ctm*, 5:669–697, 2001.

[27] J. C. Hewson and A. R. Kerstein. Local extinction and reignition in nonpremixed turbulent $CO/H_2/N_2$ jet flames. *cst*, 174(5-6):35–66, 2002.

[28] J. C. Hewson and F. A. Williams. Rate-ratio asymptotic analysis of methane-air diffusion flame structure for predicting production of oxides of nitrogen. *cf*, 117:441–476, 1999.

[29] Sergio Hoyas and Javier Jimenez. Reynolds number effects on the reynolds-stress budgets in turbulent channels. *Physics of Fluids*, 20, 2008.

[30] S. Janson. *Gaussian Hilbert Spaces*. Camb. Univ. Press, Cambridge, UK, 1997.

[31] I. T. Jolliffe. *Principal Component Analysis*. Springer, 2002.

[32] R.E. Kass, B.P. Carlin, A. Gelman, and R.M. Neal. Markov Chain Monte Carlo in Practice: A Roundtable Discussion. *The American Statistician*, 52(2):93–100, 1998.

[33] C.M. Kaul and V. Raman. A posteriori analysis of numerical errors in subfilter scalar variance modeling for large eddy simulation. *Physics of Fluids*, 23:035102, 2011.

[34] A. Y. Klimenko and R. W. Bilger. Conditional moment closure for turbulent combustion. *pecs*, 25:595–687, 1999.

[35] Y. Li, E. Perlman, M. Wan, Y. Yang, C. Meneveau, R. Burns, S. Chen, A. Szalay, and G. Eyink. A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence. *Journal of Turbulence*, 9(31):1–29, 2008.

[36] J. Ling, F. Coletti, S. Yapa, and J. Eaton. Experimentally informed optimization of turbulent diffusivity for a discrete hole film cooling geometry. *International Journal of Heat and Fluid Flow*, 44:348–357, 2013.

[37] S. Liu, J. C. Hewson, J. H. Chen, and H. Pitsch. Effects of strain rate on high-pressure nonpremixed n-heptane autoignition in counterflow. *cf*, 137:320–339, 2004.

[38] D. Lucor, J. Meyers, and P. Sagaut. Sensitivity analysis of large-eddy simulations to subgrid-scale-model parametric uncertainty using polynomial chaos. *Journal of Fluid Mechanics*, 585:255–279, 2007.

[39] S.M. Lynch and B. Western. Bayesian posterior predictive checks for complex models. *Sociological Methods and Research*, 32(3):301–335, 2004.

[40] F. Mauss, D. Keller, and N. Peters. A lagrangian simulation of flamelet extinction and reignition in turbulent jet diffusion flames. *pci*, 23:693–698, 1990.

[41] A. Liñán. The asymptotic structure of counterflow diffusion flames for large activation energies. *Acta Astronautica*, 1(7-8):1007–1039, July 1974.

[42] J. Meyers, B.J. Geurts, and M. Baelmans. Database analysis of errors in large-eddy simulation. *Physics of Fluids*, 15(9):2740–2755, 2003.

[43] W. J. Minkowycz, E. M. Sparrow, G. E. Schneider, and R. H. Pletcher. *Handbook of Numerical Heat Transfer*. Wiley-Interscience, 1988.

[44] C. D. Moen, G. H. Evans, S. P. Domino, and S. P. Burns. A multi-mechanics approach to computational heat transfer. In *ASME 2002 International Mechanical Engineering Congress and Exposition*, volume 6, pages 25–32. ASME, 2002.

[45] R.D. Moser, N.P. Malaya, H. Chang, P.S. Zandonade, Vedula P., A. Bhattacharya, and A. Haselbacher. Theoretically based optimal large-eddy simulation. *Physics of Fluids*, 21:105104, 2009.

[46] Robert D. Moser, John Kim, and Nagi N. Mansour. Direct numerical simulation of turbulent channel flow up to Re=590. *Physics of Fluids*, 11(4):943–945, 1999.

[47] H.N. Najm. Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics. *Annual Review of Fluid Mechanics*, 41(1):35–52, 2009.

[48] F. Nicoud and F. Ducros. Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow, Turbulence and Combustion*, 62(3):183–200, 1999.

[49] T.A. Oliver and R.D. Moser. Bayesian uncertainty quantification applied to RANS turbulence models. *Journal of Physics: Conference Series*, 318:042032, 2011.

[50] T.A. Oliver and R.D. Moser. Accounting for uncertainty in the analysis of overlap layer mean velocity models. *Physics of Fluids*, 24:075108, 2012.

[51] S. Paolucci. On the filtering of sound waves from the navier-stokes equations. Technical Report SAND-828257, Sandia National Laboratories, Livermore, CA, December 1982.

[52] E. Perlman, R. Burns, Yi Li, and C. Meneveau. Data exploration of turbulence simulations using a database cluster. In *Supercomputing, 2007. SC '07. Proceedings of the 2007 ACM/IEEE Conference on*, pages 1–11, Nov 2007.

[53] N. Peters. *Turbulent Combustion*. Cambridge University Press, Cambridge, United Kingdom, 2000.

[54] P. Pettersson, G. Iaccarino, and J. Nordstróm. Numerical analysis of the Burgers' equation in the presence of uncertainty. *Journal of Computational Physics*, 228:8394–8412, 2009.

[55] P. Pettersson, J. Nordstróm, and G. Iaccarino. Boundary procedures for the time-dependent Burgers' equation under uncertainty. *Acta Mathematica Scientia*, 30B(2):539–550, 2010.

[56] H. Pitsch. *FlameMaster, V3.1*. Stanford University, 2004.

[57] H. Pitsch, C. M. Cha, and S. Fedotov. Flamelet modeling of non-premixed turbulent combustion with local extinction and re-ignition. *ctm*, 7:317–332, 2003.

[58] A.E. Raftery and S. Lewis. How Many Iterations in the Gibbs Sampler? In *Bayesian Statistics 4*, pages 763–773. Oxford University Press, 1992.

[59] M.T. Reagan, H.N. Najm, R.G. Ghanem, and O.M. Knio. Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection. *Combustion and Flame*, 132:545–555, 2003.

[60] R. G. Rehm and H. R. Baum. The equations of motion for thermally driven buoyant flows. *Journal of Research of the National Bureau of Standards*, 83:279, 1978.

[61] Daniel M. Ricciuto, Anthony W. King, D. Dragoni, and Wilfred M. Post. Parameter and prediction uncertainty in an optimized terrestrial carbon cycle model: Effects of constraining variables and data record length. *Journal of Geophysical Research: Biogeosciences*, 116(G1), 2011.

[62] F. Rizzi, M. Salloum, Y. Marzouk, R. Xu, M. Falk, T. Weihs, G. Fritz, and O. Knio. Bayesian inference of atomic diffusivity in a binary Ni/Al system based on molecular dynamics. *Multiscale Modeling and Simulation*, 10(2):550–584, 2011.

[63] M. Rosenblatt. Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics*, 23(3):470 – 472, 1952.

[64] A. M. Ruiz, G. Lacaze, and J. C. Oefelein. Flow topologies and turbulence scales in a jet-in-cross-flow. *Physics of Fluids*, 27(4):–, 2015.

[65] C. Safta, M. Blaylock, J. Templeton, S. P. Domino, K. Sargsayn, and H. N. Najm. Parameter uncertainty in les of channel flow. *Inernational Journal for Numerical Methods in Fluids*, In Review., 2015.

[66] P. Sagaut. *Large Eddy Simulation for Incompressible Flows*. Springer, 2005.

[67] M. Salloum and J.A. Templeton. Inference and uncertainty propagation of atomistically informed continuum constitutive laws, part 2: Generalized continuum models based on guassian processes. *International Journal for Uncertainty Quantification*, 4(2):171–184, 2014.

[68] K. Sargsyan, H. Najm, and R. Ghanem. On the Statistical Calibration of Physical Models. *International Journal for Chemical Kinetics*, 47(4):246–276, 2015.

[69] U. Schumann. Subgrid length-scales for large-eddy simulation of stratified turbulence. *Theoretical and Computational Fluid Dynamics*, 2(5-6):279–290, 1991.

[70] D.W. Scott. *Multivariate Density Estimation. Theory, Practice and Visualization*. Wiley, New York, 1992.

[71] M. Shoeybi and J.A. Templeton. Three-dimensional wall filtering formulation for large-eddy simulation. In *Center for Turbulence Research Annual Research Briefs*, 2006.

[72] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.

[73] James R. Stewart and H. Carter Edwards. The SIERRA Framework for Developing Advanced Parallel Mechanics Applications. In L. T. Biegler, M. Heinkenschloss, O. Ghattas, and B. van Bloemen Waanders, editors, *Large-Scale PDE-Constrained Optimization*, volume 30 of *Lecture Notes in Computational Science and Engineering*, pages 301–315. Springer Berlin Heidelberg, 2003.

[74] L. K. Su and M. G. Mungal. Simultaneous measurements of scalar and velocity field evolution in turbulent crossflowing jets. *Journal of Fluid Mechanics*, 513:1–45, 8 2004.

[75] G.J. Székely, M.L. Rizzo, and N.K. Bakirov. Measuring and Testing Dependence by Correlation of Distances. *Annals of Statistics*, 35:2769–2794, 2007.

[76] J.A. Templeton and M. Shoeybi. Towards wall-normal filtering for large-eddy simulation. *Multiscale Modeling and Simulation*, 5(2):420–444, 2006.

[77] Jeremy A. Templeton, Gorazd Medic, and Georgi Kalitzin. An eddy-viscosity based near-wall treatment for coarse grid large-eddy simulation. *Physics of Fluids*, 17(10):105101, 2005.

[78] S. Völker, R.D. Moser, and P. Venugopal. Optimal large eddy simulation of turbulent channel flow based on direct numerical simulation statistical data. *Physics of Fluids*, 14(10):3675–3691, 2002.

[79] Q. Wang, K. Duraisamy, J.J. Alonso, and G. Iaccarino. Risk Assessment of Scramjet Unstart Using Adjoint-Based Sampling Methods. *AIAA Journal*, 50(3):581–592, 2012.

[80] N. Wiener. The Homogeneous Chaos. *Am. J. Math.*, 60:897–936, 1938.

[81] David C. Wilcox. *Turbulence Modeling for CFD*. DCW Industries, Inc., third edition edition, 2010.

[82] F. A. Williams. Crocco variables for diffusion flames. In C. Casci, editor, *Recent Advances in Aerospace Sciences*, page 415. Plenum, New York, 1980.

[83] D. Xiu and G.E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002.

[84] J. Xu and S. B. Pope. PDF calculations of turbulent nonpremixed flames with local extinction. *cf*, 123:281–307, 2000.

[85] P. K. Yeung, J. G. Brasseur, and Q. Wang. Dynamics of large-small scale couplings in coherently forced turbulence: concurrent physical and fourier-space views. *jfm*, 283:43–95, 1995.

[86] P. K. Yeung, S. S. Girimaji, and S. B. Pope. Straining and scalar dissipation on material surfaces in turbulence: Implications for flamelets. *cf*, 79:340–365, 1990.

[87] A. Yoshizawa. Statistical theory for compressible turbulent shear flows, with the application to subgrid modeling. *Physics of Fluids (1958-1988)*, 29(7):2152–2164, 1986.

[88] A. Yoshizawa. Bridging between eddy-viscosity-type and second-order turbulence models through a two-scale turbulence theory. *Phys. Rev. E*, 48:273–281, Jul 1993.

DISTRIBUTION:

1	MS 0899	Technical Library, 8944 (electronic)
1	MS 0123	D. Chavez, LDRD Office, 1011

Sandia National Laboratories