



Informing Turbulence Closures With Computational and Experimental Data

Karthik Duraisamy and Anand Pratap Singh

*Department of Aerospace Engineering,
University of Michigan, Ann Arbor, MI 48109.*

Data from high fidelity computational simulations and experimental measurements is used to quantify errors in the functional form of turbulence closure models. A Bayesian formalism is adopted to infer a full-field of discrepancy that provides otherwise inaccessible modeling information. The utility of the approach is demonstrated by applying it to a number of problems including channel flow, shock-boundary layer interactions, and flows with curvature and separation. In all these cases, the posterior model correlates well with the data. Furthermore, it is shown that even if limited data (such as surface pressures) is used, the accuracy of the inferred solution is improved over the entire computational domain. By directly addressing the connection between physical data and model discrepancies, the field inversion approach enhances the value of computational and experimental data for model improvement. Based on the results, it could be argued that some of the criticism of eddy viscosity models that is usually attributed to the isotropic stress-strain relationship is partially a result of poor functional forms within the model rather than the limiting nature of the Boussinesq assumption itself. Models should rather be assessed by evaluating the type and variability of corrections that will be required in a diverse set of problems. The information from the inversion procedure can be used by the modeler as a guiding tool to design more accurate model forms, or serve as input to machine learning algorithms to directly replace deficient modeling terms.

I. Introduction

Near-wall turbulence modeling in both a RANS and hybrid RANS-LES context is recognized to be the pacing item in applied computational fluid dynamics.¹ Despite their widespread use, near-wall turbulence models lack accuracy in a number of flows, especially those involving pressure gradients, separation, curvature, etc. While there are several encouraging recent efforts exploring advanced second,² third³ and even fourth⁴ moment closures for complex flows, the predictive capability of turbulence models has not improved significantly over the past two decades.

As part of a NASA-funded initiative, the authors are engaged in a broad program of turbulence modeling⁵⁻⁹ with a focus on the innovative use of data to improve closure models. Since RANS models involve empiricism, it can be well-argued that these models have always been data-driven. At the same time, tremendous advances in computing power and measurement resolution have provided datasets containing complete information in problems of reduced scope or partial information in regimes of interest. Despite the availability of such high quality DNS, LES and PIV data, the improvement in closure models has not been significant.

The key impediment to improving closure models from higher fidelity data is the issue of model discrepancy. The role of model discrepancy can be illustrated by the example of a typical turbulence closure model. Since turbulence models are formulated to provide first and second moments, secondary variables such as the dissipation rate typically assume different values in closure models compared to their values in the real-world or in higher-fidelity simulations. Thus, approaches that attempt to extract *a priori* information from DNS may not be very useful in a predictive setting, since the **extracted information will be inconsistent with the information processed by the model**. Further, even if individual terms in the closure equations are modeled accurately, small discrepancies can lead to poor results. For instance, Poroseva and Murman¹⁰ observed that even if some terms in a second moment closure are extracted directly from DNS, the overall

prediction was not satisfactory. It is thus the **balance** between every term in the model equation that is important. To obtain useful data-driven models, one should pursue methods that produce **information** within the context of the model.

At this stage, it would be convenient to categorize modeling errors into those resulting from a) the model structure (such as the choice of an algebraic Reynolds stress model or $k - \epsilon$ model), b) the chosen functional forms within the model structure, and c) parameters in the chosen functional form. Much of the data-driven approaches that have been applied in turbulence modeling have involved the calibration of a set of model parameters. Yarlanki et al.,¹¹ Kato and Obayashi,¹² and Edeling et al.¹³ apply modern data-assimilation techniques to infer model parameters. These techniques provide insight into the variability of model outputs, but are limited in scope since the functional form of the model is frozen, thus restricting the scope and generality of the calibration. Other attempts have approached the problem from the viewpoint of *characterizing* structural errors. Dow and Wang¹⁴ infer the error in the coefficient of the eddy viscosity relationship by assuming it to be a Gaussian random field. Cheung et al.¹⁵ employ Bayesian model averaging to calibrate turbulence models. Emory et al.¹⁶ inject realizable perturbations to the Reynolds stresses computed by eddy viscosity models to obtain uncertainty bounds on the predictions. Ling and Templeton¹⁷ employ machine learning techniques directly on high and low-fidelity data to identify regions in which high turbulence modeling uncertainty may be present.

In this work, a more comprehensive approach towards extracting modeling information from data is pursued, with the goal of ultimately improving our qualitative and quantitative understanding of model errors. In contrast to inferring calibration parameters, the functional form of the model discrepancy is targeted. This information can be used by the modeler as a guiding tool to design parametric model forms, or serve as input to machine learning algorithms to supplement models.^{5,7,9}

The rest of the paper is organized as follows: Section II introduces the inversion framework and methodology for functional error and uncertainty quantification and sections III – VI present the application of the inversion framework to four different turbulence modeling scenarios. The summary and future perspectives are provided in section VII.

II. Methodology

The proposed method can be adapted for near-wall modeling in LES, or for improving closures in general. Our focus in this paper is on a subset of RANS-based turbulence models. The examples are intended to demonstrate how useful modeling information can be derived from data. A key challenge is the accurate formulation of source terms in the transport equations as they have to be modified to represent important phenomena such as streamline curvature, relaminarization, etc. It is pertinent to recognize that modifying the source terms does indeed change the balance of the entire model. Further, because of the underlying transport equation, a change in the source term affects the solution domain in a non-local sense.

Without a loss of generality, the applications considered in this work involve steady flows, and thus the inferred functional forms are of a spatial nature. A spatially-varying function $\beta(\mathbf{x})$ is introduced as a multiplier of a source term $S(U(\mathbf{x}))$ ^a, typically the production term, the destruction term, or the balance between the two terms. The multiplication by $\beta(\mathbf{x})$ is equivalent to adding a new discrepancy term $\delta(\mathbf{x}) = (\beta(\mathbf{x}) - 1)S(U(\mathbf{x}))$ in the transport equation. The multiplication to the source rather than a stand-alone correction function, $\delta(\mathbf{x})$ presents several advantages to the inversion process as $\beta(\mathbf{x})$ is, by definition, non-dimensional and has a simple prior value of unity. This structure is also favorable to the conditioning of the inverse problem. Note that the inference of $\beta(\mathbf{x})$ is not physically restrictive. As an example, if $\beta(\mathbf{x})$ is a multiplier of the production term and the production is zero, the turbulence model is not relevant in the region. Alternately, functions $\beta_p(\mathbf{x})$ and $\beta_d(\mathbf{x})$ can be used to multiply production and destruction, respectively.

Once $\beta(\mathbf{x})$ is introduced into the solution, the next step involves inference. Given data from a higher fidelity simulation or experiment, $\beta(\mathbf{x})$ can be estimated by solving an inverse problem as described in the following section.

Solution of individual inverse problems can be considered to be a step towards the larger goal of improving predictive capabilities using a data-driven framework.⁷ The inferred function $\beta(\mathbf{x})$ can be mapped to $\beta(\boldsymbol{\eta})$, where $\boldsymbol{\eta}$ may include quantities such as mean velocity gradients, turbulent time scales, etc. By considering a sufficiently diverse set of inverse problems, the form of $\beta(\boldsymbol{\eta})$ can be generalized, for instance, using a machine

^a $U(\mathbf{x})$ are the Reynolds averaged flow variables

learning algorithm. The inverse solution can, by itself, be used by a modeler to understand and correct for deficiencies in existing models.

II.A. Bayesian Inversion

Inverse problems can be used to extract causal relationships from data. While inverse problem can be formulated in a deterministic setting, we are also interested in characterizing the impact of uncertainties in the data, model inadequacies and existing knowledge, on the inferred outputs. These statistical measures can be used to quantify the confidence on the inferred model, and by extension, can be further propagated in predictive modeling.

The Bayesian approach¹⁸ provides a formalism to quantify the posterior probability distribution $q(\boldsymbol{\beta}|\mathbf{d})$, given a prior distribution ($p(\boldsymbol{\beta})$), data vector, \mathbf{d} , and a likelihood function, $h(\mathbf{d}|\boldsymbol{\beta})$. The Bayes theorem states that,

$$q(\boldsymbol{\beta}|d) = \frac{h(\mathbf{d}|\boldsymbol{\beta})p(\boldsymbol{\beta})}{c}, \quad (1)$$

where $c = \int h(\mathbf{d}|\boldsymbol{\beta})p(\boldsymbol{\beta})d\boldsymbol{\beta}$. In the current form, the solution to equation 1 is intractable using a sampling based method because of the infinite dimensional nature of $\boldsymbol{\beta}$. The problem is made finite-dimensional by re-defining it such that value of $\boldsymbol{\beta}$ is inferred at every discrete location in the computational domain of the RANS grid^b. If $\boldsymbol{\beta}$ and $h(\mathbf{d}|\boldsymbol{\beta})$ are assumed to be Gaussian, the Maximum a posteriori (MAP) estimate¹⁹ can be computed by solving the following deterministic minimization problem:

$$\boldsymbol{\beta}_{MAP} = \arg \min_{\boldsymbol{\beta}} \mathfrak{J}(\boldsymbol{\beta}) = \arg \min_{\boldsymbol{\beta}} \frac{1}{2} [\mathbf{F}^T \mathbf{C}_{obs}^{-1} \mathbf{F} + (\boldsymbol{\beta} - \boldsymbol{\beta}_{prior})^T \mathbf{C}_{prior}^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}_{prior})], \quad (2)$$

where \mathbf{F} is a vector with i^{th} element,

$$f_i = d_{i,RANS} - d_{i,benchmark},$$

d_i is the i^{th} data point. d can be any flow dependent quantity such as pressure, skin-friction, velocity etc. \mathbf{C}_{obs} and \mathbf{C}_{prior} are the observational and the prior covariance matrices. $\boldsymbol{\beta}_{prior}$ is the prior mean of the parameters (and also corresponds to the *base* model).

Turbulence models constitute complex non-linear systems, and thus the posterior distribution cannot be expected to be perfectly Gaussian. To promote tractability in the already high-dimensional setting, the posterior covariance is *approximated* by linearizing the estimate about the MAP point. This leads to a definition of the covariance in terms of the inverse of the Hessian of the objective function, $\mathfrak{J}(\boldsymbol{\beta})$,

$$\mathbf{C}_{posterior} = \left[\frac{d^2 \mathfrak{J}(\boldsymbol{\beta})}{d\boldsymbol{\beta}d\boldsymbol{\beta}} \right]^{-1} \Bigg|_{\boldsymbol{\beta}_{MAP}}. \quad (3)$$

The posterior distribution $\boldsymbol{\beta}$ can be obtained by first performing a Cholesky decomposition of the form^c

$$\mathbf{R}^T \mathbf{R} = \mathbf{C}_{posterior}, \quad (4)$$

and generating realizations using

$$\boldsymbol{\beta} = \boldsymbol{\beta}_{MAP} + \mathbf{R}^T \mathbf{s}, \quad (5)$$

where \mathbf{s} is a vector of the same size as $\boldsymbol{\beta}$ and contains normally distributed random numbers with zero mean and unit standard deviation. The realizations of the posterior $\boldsymbol{\beta}$ are used to construct the posterior realizations of various flow quantities by solving the RANS equations for each realization of $\boldsymbol{\beta}$. The prior realizations are constructed in the similar manner using \mathbf{C}_{prior} .

It is well-recognized that the posterior distribution is sensitive to the assumptions on the prior and likelihood and the structure of the observational covariance matrix, \mathbf{C}_{obs} . Some of the implications of these

^bNote that is still a very high-dimensional problem.

^c \mathbf{R} is an upper triangular matrix

assumptions are discussed in a companion work by Parish & Duraisamy.⁵ In the scope of the present work, however, a Gaussian assumption is pursued as it is basic and results in reduced complexity.

As mentioned previously, the posterior distribution is approximated by the inverse of the Hessian (Eq.3). In the examples presented in this work, computations of the exact Hessian are extremely expensive and ill-conditioned as the dimensions scale linearly with the number of the unknowns (in this case, the number of mesh points). Therefore, a Gauss-Newton approximation is employed, the details of which is presented in Ref. 5.

II.B. Flow and Adjoint Solver

The unstructured compressible finite volume solver JOE²⁰⁻²² is used to solve the RANS equations. A cell-centered, second order accurate spatial discretization is employed, with a backward Euler implicit scheme to march towards steady state. The inversion procedure requires gradients with respect to every grid point, thus necessitating a discrete adjoint²³ procedure. The partial derivatives required in the adjoint procedure are computed efficiently using automatic differentiation via the ADOL-C framework.²⁴

The following sections demonstrate the application of the inference in a number of turbulence modeling situations.

III. Turbulent Channel Flow using $k-\omega$

The first case to be considered is a fully developed planar turbulent channel flow. This canonical problem has been heavily utilized to calibrate model parameters throughout the history of turbulence modeling, although most models do not account for the variability due to Reynolds number. The question of choosing the right combination of length and time scales is still an outstanding one in modeling. Rich Direct Numerical Simulation (DNS) datasets²⁵⁻²⁷ exist for such flows. These datasets are, in our opinion, under-utilized for modeling purposes. In the current work, the functional form of discrepancies in the Wilcox $k-\omega$ model²⁸ is extracted using DNS data for a wide range of Reynolds numbers.

α^*	σ^*	γ	α	σ
0.09	0.6	13/25	0.09	0.5

Table 1: Wilcox $k-\omega$ model constants.

III.A. Bayesian Inversion Setup

The Reynolds-averaged momentum equation for a fully developed, incompressible and steady channel flow is given by,

$$\frac{\partial}{\partial y} \left[\mu \frac{\partial u}{\partial y} - \rho \overline{u'v'} \right] - \frac{\partial p}{\partial x} = 0, \quad (6)$$

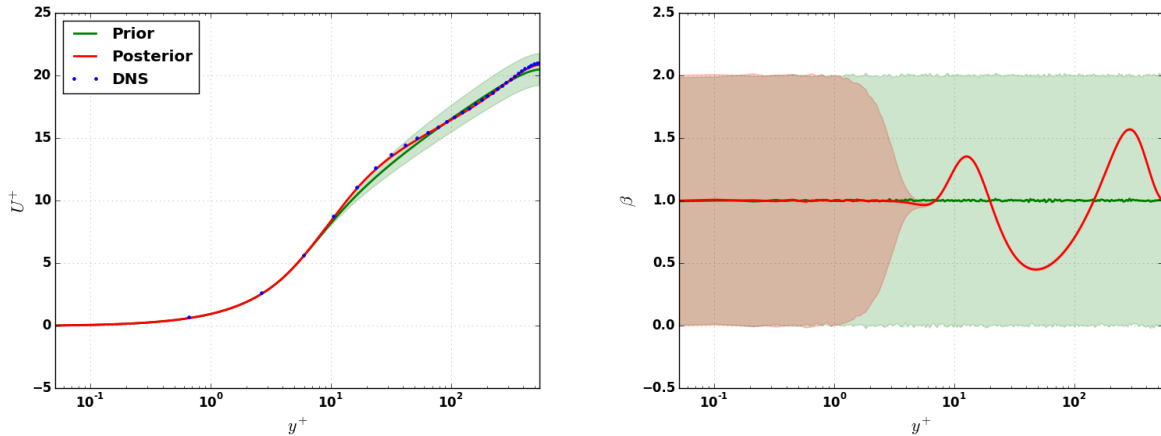
where, u represents the mean streamwise velocity and the Reynolds stress $-\overline{u'v'}$ is modelled as $\nu_t \frac{\partial u}{\partial y}$ in eddy viscosity models. In the Wilcox $k-\omega$ model, $\nu_t = \frac{k}{\omega}$, where k and ω satisfy the following transport equations,

$$\nu_t \left(\frac{\partial U}{\partial y} \right)^2 - \alpha^* k \omega + \frac{\partial}{\partial y} \left[\left(\nu + \sigma^* \frac{k}{\omega} \right) \frac{\partial k}{\partial y} \right] = 0 \quad (7)$$

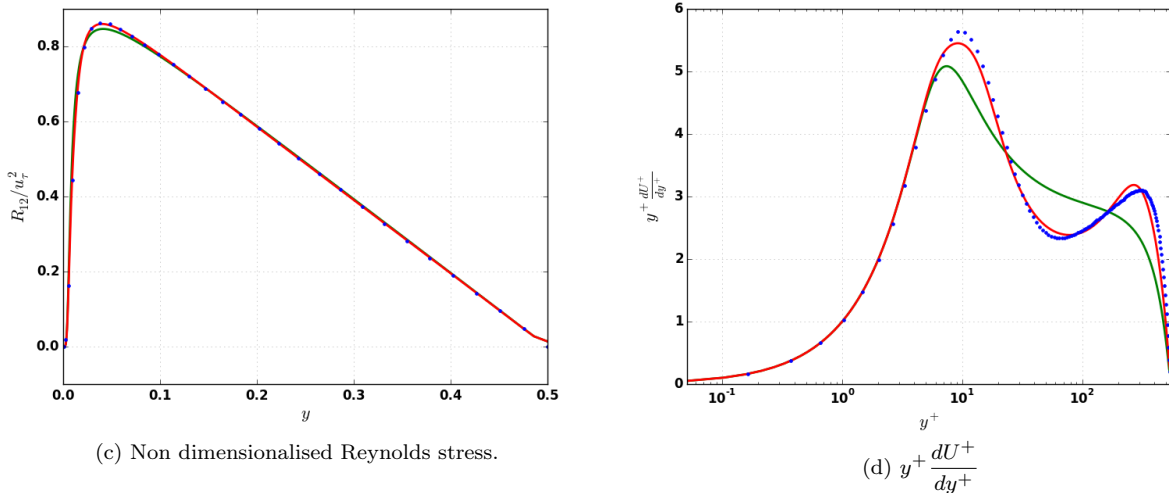
$$\gamma \left(\frac{\partial U}{\partial y} \right)^2 - \alpha \omega^2 + \frac{\partial}{\partial y} \left[\left(\nu + \sigma \frac{k}{\omega} \right) \frac{\partial \omega}{\partial y} \right] = 0. \quad (8)$$

Table 1 contains the values of the model constants. In this work, a correction function $\beta(\mathbf{y})$ is introduced as a multiplier to the production term in the ω equation (equation 8). The system is solved on a geometrically

graded mesh consisting of 201 points with the first grid point placed well into the viscous sublayer at $y^+ \approx 0.05$. Inversion is performed for set of five different Reynolds numbers, $Re_\tau \in [180, 550, 950, 2000, 4200]$. The ensemble averaged velocity –at every grid point– from the DNS solution²⁶ is used as the benchmark data. A diagonal observational covariance with standard deviation, σ_{obs} , of 10^{-8} is used. The very low value represents a high degree of confidence on the DNS data. A diagonal prior covariance with $\sigma_{prior} = 0.5$ is used, which corresponds to a nearly uninformative prior. The MAP solution is obtained as described in the previous section. The posterior and prior distributions are constructed using 10000 realizations.



(a) Non dimensionalized velocity, U^+ . Mean is shown with the solid line; shaded region represents the 95% confidence interval. (b) Correction function, β . Mean is shown with the solid line; shaded region represents the 95% confidence interval.



(c) Non dimensionalised Reynolds stress.

(d) $y^+ \frac{dU^+}{dy^+}$

Figure 1: The prior, the posterior and the DNS solution for $Re_\tau = 550$ with $k - \omega$ model.

III.B. Results

Figure 1 shows the prior, posterior and DNS values for flow properties at $Re_\tau = 550$. The posterior velocity profile is almost identical to the DNS profile with the standard deviation collapsing to a very low value. The accuracy of the posterior can be further verified from Figure 1d, which shows the variation of $y^+ du^+/dy^+$, and further highlights the differences. As turbulent production is negligible for $y^+ < 1$, the posterior solution is not affected by $\beta(y)$ and hence $\beta(y)$ remains unchanged from its prior value. Correspondingly, there is also no reduction in uncertainty in the viscous sublayer for $y^+ < 2$. Outside of the sublayer, $\beta(y)$ changes to match the specified benchmark velocity profile and a drastic reduction in the posterior standard deviation is noticeable.

Figure 2 shows the inferred function, β , for all Reynolds numbers. In every case, β in the overlap and the log-layer scales with wall units. In the outer region, β scales with wall height. This is a physically reasonable variation and represents the systematic nature of modeling deficiency. The large variation in β around $y^+ = 25$ shows the inability of the underlying model to provide the correct time scale near the interface of the overlap region and the log-layer. It is particularly revealing that the nature of the required correction is almost universal (with second-order viscous corrections) for $y^+ < 50$, a fact that can be used to develop improved models.

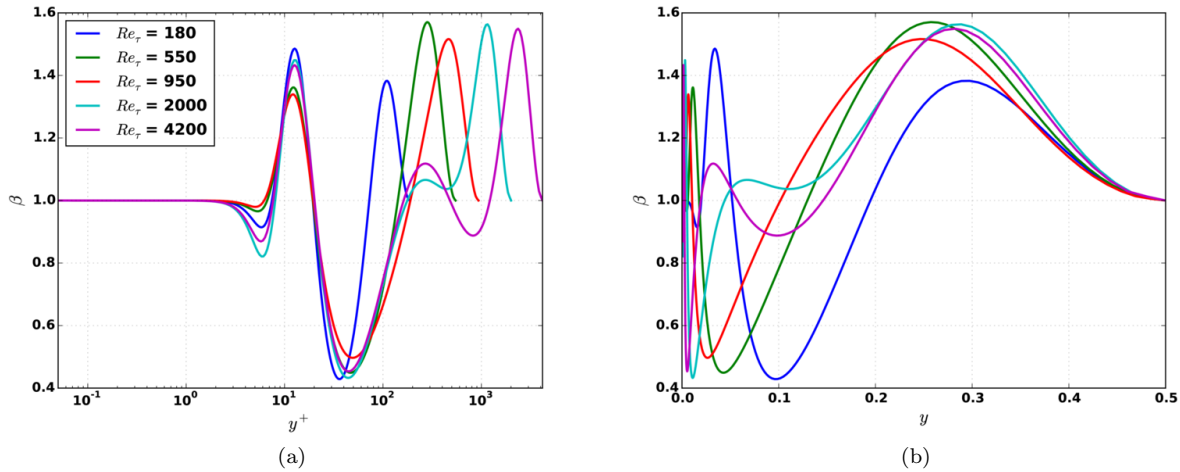


Figure 2: The inferred correction function, β_{MAP} , for $Re_\tau \in [180, 550, 950, 2000, 4200]$ with $k - \omega$ model.

IV. Turbulent Channel Flow using *Stress* – ω

The next two cases are extensions of the previous case, employ the Wilcox *Stress* – ω^{28} second-moment closure model. Reynolds stress-based models have the potential to better represent the turbulence evolution and are inherently advantageous owing to the exact representation of production and convective processes and their ability to represent anisotropic effects. However, the inexact pressure strain correlation and the assumption of local isotropy while calculating turbulent dissipation can result in large errors. The associated numerical stiffness also presents robustness issues. In second moment closure models, the scale equation is often blamed for the poor performance of the model. In the following, we infer the correction function for the *stress* – ω model. The model is adapted in two different ways: (1) by introducing a source term in the ω transport equation, and, (2) by introducing source terms in the Reynolds stress equations. For the inference, the following objective function is used,

$$\begin{aligned} \mathfrak{J}(\beta) = & \frac{100w_u^2}{\sigma_{obs}^2} \mathbf{F}_u^T \mathbf{F}_u + \frac{w_{R11}^2}{\sigma_{obs}^2} \mathbf{F}_{R11}^T \mathbf{F}_{R11} + \frac{w_{R12}^2}{\sigma_{obs}^2} \mathbf{F}_{R12}^T \mathbf{F}_{R12} + \frac{w_{R22}^2}{\sigma_{obs}^2} \mathbf{F}_{R22}^T \mathbf{F}_{R22} + \frac{w_{R33}^2}{\sigma_{obs}^2} \mathbf{F}_{R33}^T \mathbf{F}_{R33} \\ & + \frac{(\beta - \beta_{prior})^T (\beta - \beta_{prior})}{\sigma_{prior}^2}, \end{aligned}$$

where w 's are the weights taken to be the reciprocal of the averaged value. \mathbf{F} 's are the error vectors for the 5 flow quantities. A higher weight velocity term is required to infer accurate velocity, which is our primary objective, while minimizing the error in the Reynolds stresses. The numerical setup and the variances are identical to those in the previous case. The results of the inversion are presented in the following sections. The statistical distribution could not be obtained because of the numerical stiffness of the governing equations, hence only the maximum a posteriori estimate is presented.

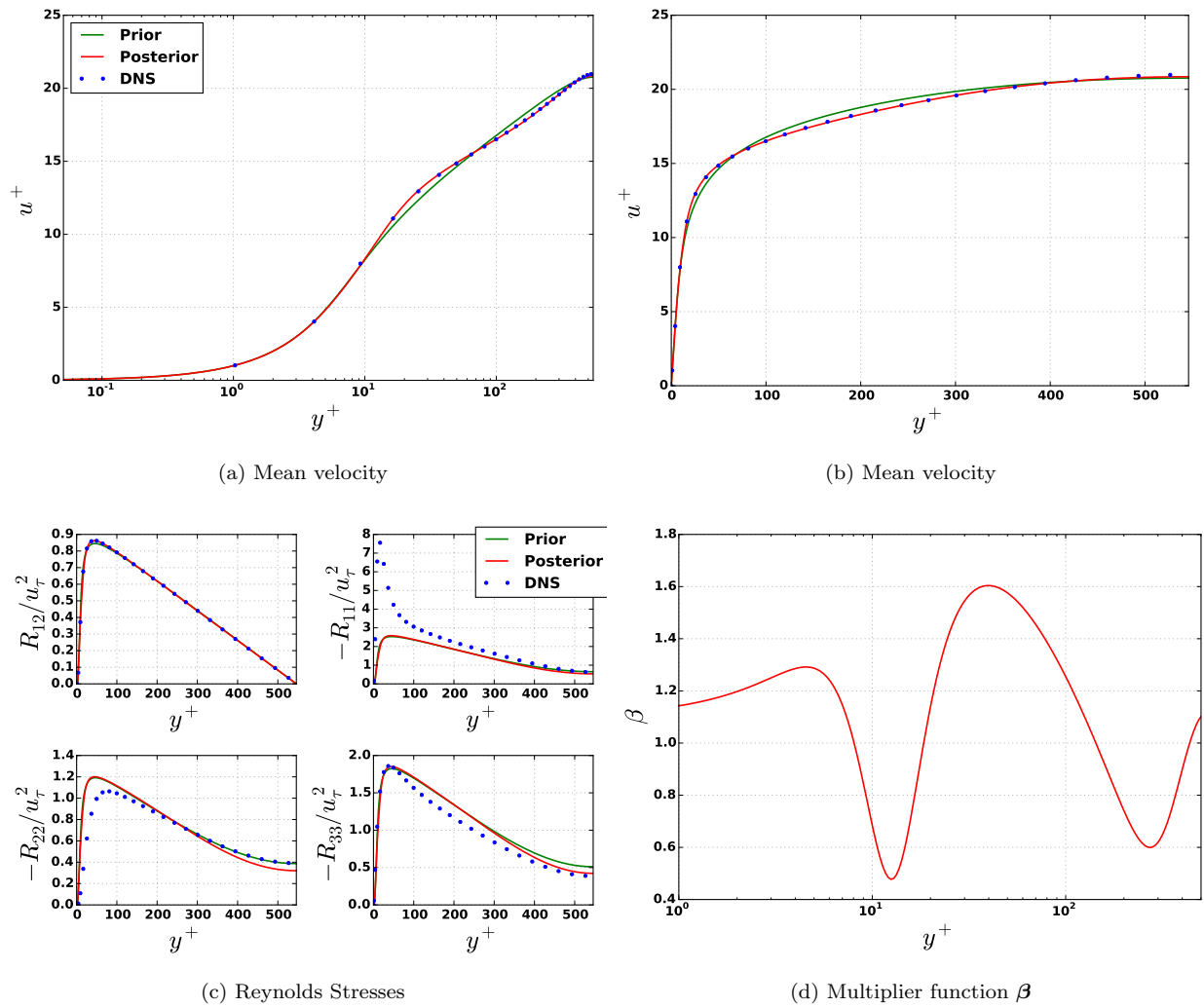


Figure 3: The prior, posterior and the DNS solution for $Re_\tau = 550$ with *Stress* - ω model and modified ω transport equation.

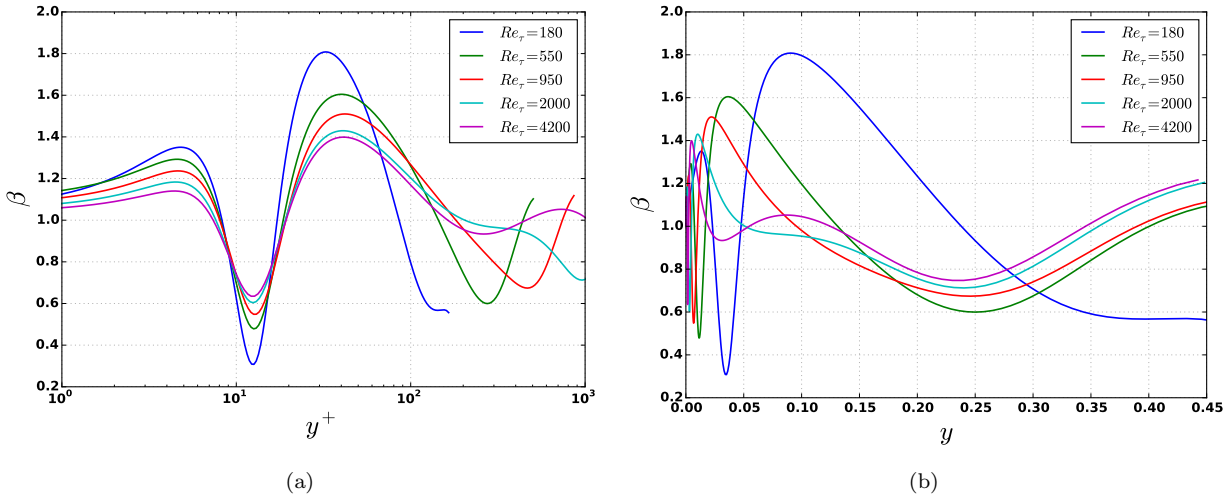


Figure 4: The inferred function $\beta(y)$ for $Re_\tau \in [180, 550, 950, 2000, 4200]$ with *Stress*– ω model and modified ω transport equation.

IV.A. Modifying the ω transport equation

A source term is introduced by multiplying a correction function, $\beta(y)$, to the dissipation term of the transport equation for ω , in a manner similar to that in the previous section. Figure 3 shows the results of the inversion process. Clearly, the inversion is able to infer a function $\beta(y)$ which results in a better estimate of the velocity. However, the normal Reynolds stresses do not show significant improvements, while some even degrade. This points to the inherent deficiencies in the transport equation for the Reynolds stresses which arise from the inexact pressure strain term and the assumption of the isotropy of the dissipation term. The form of the correction function is similar to that in the previous case with a significant disparity near y^+ of 12 and y^+ of 120, which implicates a prevailing discrepancy in the turbulence models in those regions. Figure 4 shows the inferred function for a set of Reynolds numbers, which delineate a trend in the correction for $y^+ < 50$. The inability of the current case to predict correct Reynolds stresses is addressed in the next case by adding additional source terms in the Reynolds stress equations.

IV.B. Modifying the Reynolds stress transport equations

In this case, the dissipation term in the normal stress (R_{ii}) transport equation is multiplied by a function $\beta_{ii}(y)$. In the R_{12} equation, a function β_{12} is introduced as a multiplier of the pressure strain and the turbulent transport term. It should be noted that by multiplying by β we are introducing an additional source term. Figure 5 shows the results of the inversion process. All the five posterior quantities used in the objective function are in close agreement with the DNS solution. Again, significant discrepancies are observed in all the transport equations near y^+ of 12. The anisotropic nature of the dissipation is reflected in the trends of the function β_{ii} . These trends can be utilized to correct for anisotropy and to build a better pressure strain correlation term. Figure 6 highlights that the discordance in the correction function for a set of Reynolds numbers in the near wall region follows the previously observed trend. Also, the anisotropy in the spanwise direction can be observed to be sensitive to the Reynolds number.

V. Turbulent Flow With Curvature

Turbulent flows over curved surfaces are subjected to streamwise pressure gradients and additional mean strain. Curvature acts to stabilize (or destabilize) the flow via the diverging (or impinging) inviscid core.^{29–31} The flow field is complicated by the presence of complex small-scale vortical structures.³² Linear eddy viscosity closures have to be explicitly sensitized to capture curvature effects.^{33–35} In this section, the inversion procedure is applied to boundary layer flows subject to (a) convex and (b) concave curvature. A comparison of the posterior solution with the analytically-sensitized rotational correction (SARC) model of

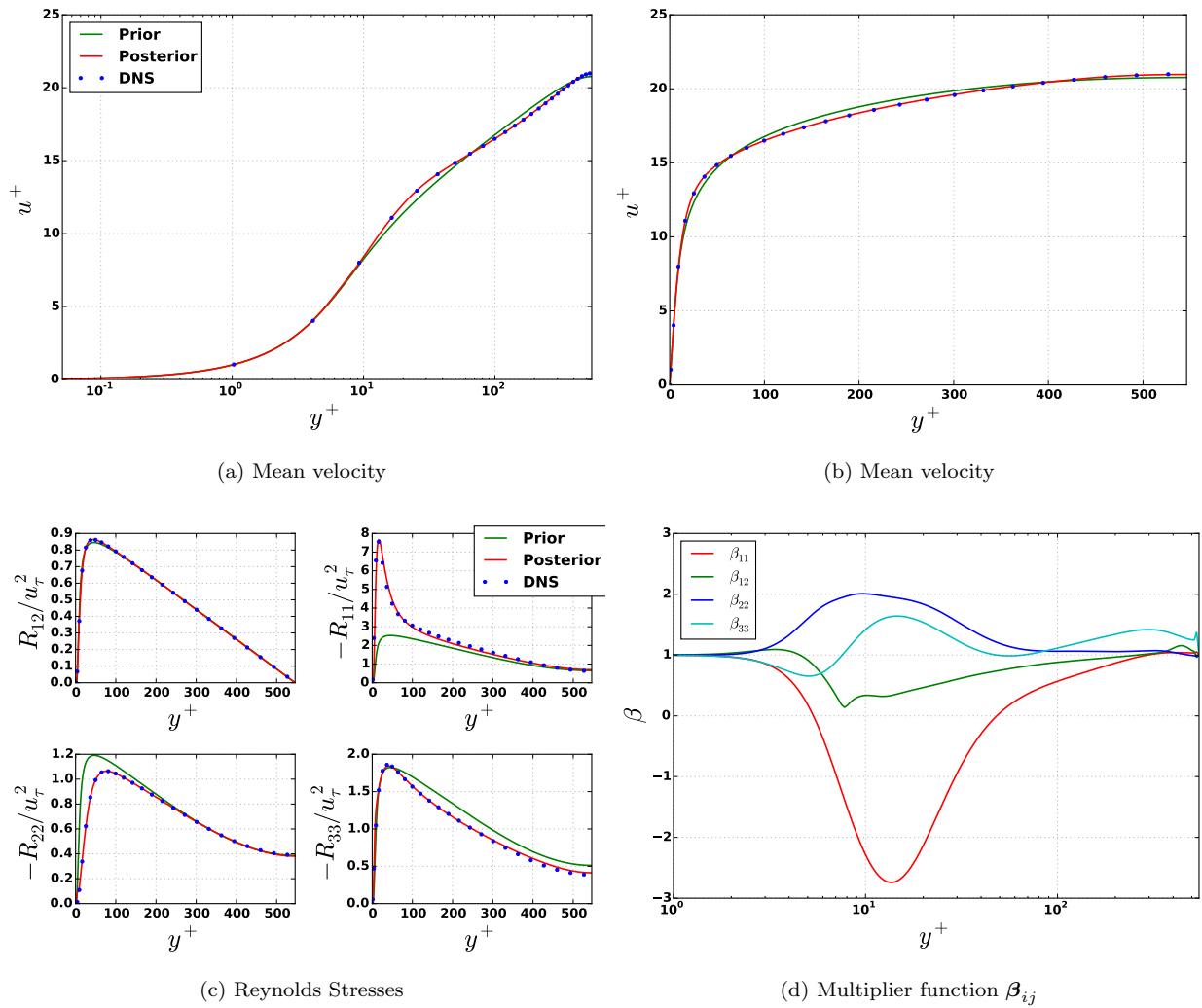


Figure 5: The prior, posterior and the DNS solution for $Re_\tau = 550$ with *Stress* - ω model and modified R_{ij} transport equations.

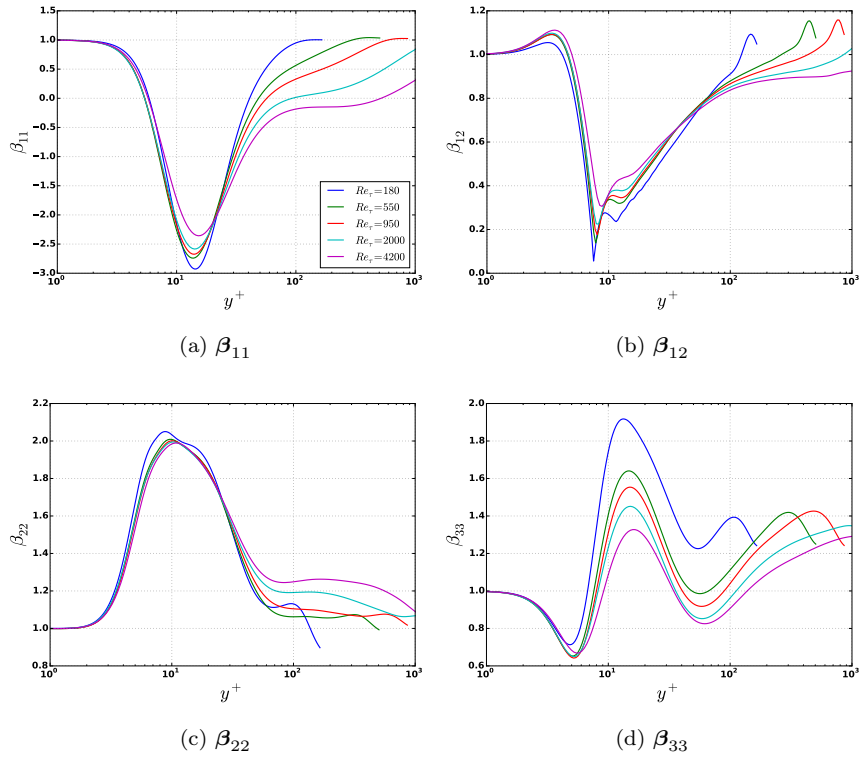


Figure 6: The inferred function for $Re_\tau \in [180, 550, 950, 2000, 4200]$ with *Stress* – ω model and modified R_{ij} transport equations.

Spalart³⁶ is also presented.

V.A. Flow Conditions

The benchmark data is obtained from a Large Eddy Simulation.^{37,38} In the LES, an inlet boundary layer with Re_θ of 2000 is generated using the recycling-rescaling procedure.³⁷ The inlet boundary layer thickness, δ_0 , is 0.04 grid units and the inlet size is $4\delta_0$. The recycling and outflow sections were of length $5\delta_0$ and $7\delta_0$, respectively. The LES solution is performed over 13 flow-through times and statistics are computed from the last 9 flow-through times. The grid inlet in the RANS is extended to minimize the effects of the inlet profile, such that $Re_\theta = 1500$ at the new grid entrance and it ensures $Re_\theta = 2000$ near the original grid entrance. The extended mesh consists of 473 points in the streamwise direction and 111 points in the wall-normal direction. The upper wall is assigned an inviscid slip boundary condition while the lower wall is given an adiabatic no-slip boundary condition. The inlet velocity, density, and eddy viscosity profiles are obtained from a precursor zero pressure gradient flat plate simulation. The pressure at the inlet is extrapolated from interior. At the outlet, pressure is interpolated from the exterior and the rest of the quantities are extrapolated.

V.B. Bayesian Inversion Setup

A correction function $\beta(x, y)$ is multiplied to the production term of the Spalart–Allmaras (SA) model in the form

$$\frac{D\tilde{\nu}_t}{Dt} = \beta(x, y)P(\tilde{\nu}_t, \mathbf{U}) - D(\tilde{\nu}_t, \mathbf{U}) + T(\tilde{\nu}_t, \mathbf{U}). \quad (9)$$

The coefficient of friction (C_f) at the lower wall, from the LES solution, is used as the benchmark data for inversion. The variation in pressure for this case is small, therefore comparing pressure from our compressible RANS to the incompressible LES is prone to error. Hence, pressure-related quantities are not used for

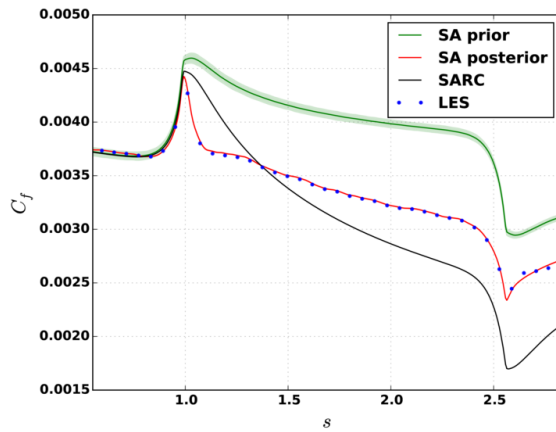


Figure 7: The prior, posterior, SARC and LES coefficient of friction (C_f) at the lower wall for the convex channel with SA model. Shaded region represents the 95% confidence interval.

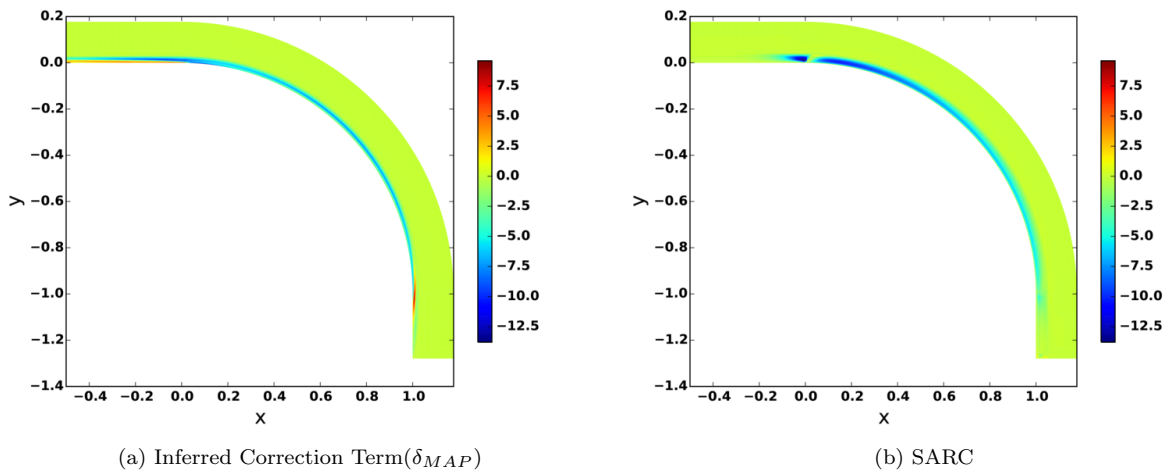


Figure 8: Contour plot showing the inferred correction term (δ_{MAP}) and the SARC correction term for the convex channel.

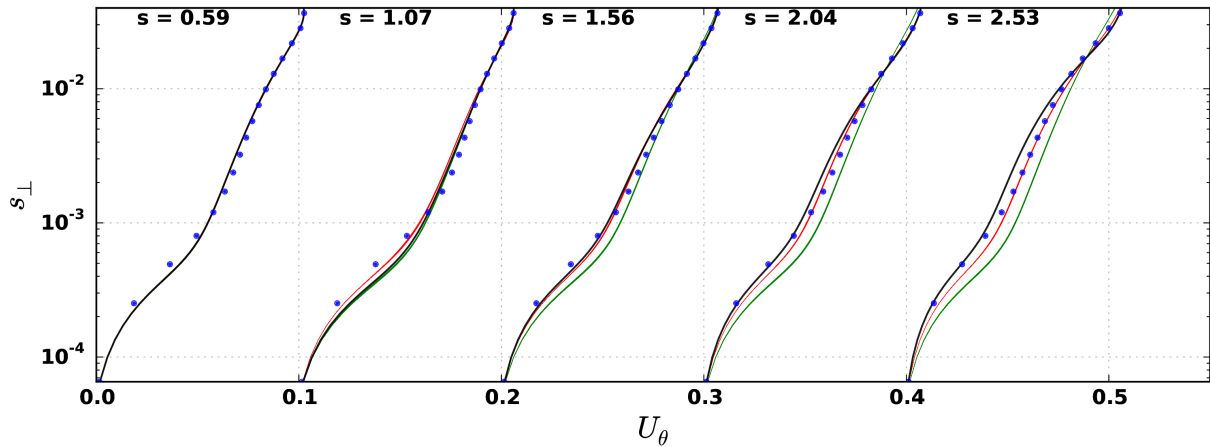


Figure 9: The prior, posterior, SARC and LES stream-wise velocity at various locations for the convex channel with SA model. s_{\perp} refers to the perpendicular distance from the lower wall. Refer Fig. 7 for the legend.

benchmarking. A diagonal prior covariance with a standard deviation, $\sigma_{prior} = 1$ is used. A diagonal observation covariance with standard deviation, σ_{obs} , of 10^{-6} is used. This corresponds to a standard deviation of around 0.03% on LES C_f , and represents a high degree of confidence in the LES. The posterior covariance and realizations are computed as detailed in section II.A. The flow solver is evaluated for 4000 realizations of β to obtain both the prior and the posterior realizations.

V.C. Results

Figure 7 shows the prior and the posterior C_f with 95% confidence intervals alongside LES and SARC results for the convex channel. The LES data points lie within the bounds or close to the posterior C_f , while the SARC model is not very accurate for this case. The confidence level in the posterior solution is seen to be very high. Figure 8 shows the inferred MAP correction term, δ_{MAP} , and the analytically-defined correction term from the corresponding SARC model. Qualitatively, the correction term is similar for both models, but the magnitudes are locally different. The trend in the correction is consistent with the expectation that the convex curvature reduces turbulence intensity. The quantitative nature of the results should provide valuable information to the modeler.

Figure 9 shows the variation of the streamwise velocity with respect to the distance from the wall at various streamwise locations. The posterior velocity is seen to correlate well with the LES solution compared to the prior and the SARC prediction. The results suggest that the SARC model requires improvements in the log layer.

Figure 10 shows the skin-friction predictions for a channel with concave curvature. The posterior C_f is close to the LES solution except in a small region downstream. It is to be noted that linear eddy viscosity-based models may be deficient in their ability to perform better. The correction term in Figure 11 is consistent with the increase in turbulent production in a concave boundary layer.

Figure 12 shows the streamwise velocity profiles. Clearly, the posterior velocity prediction is closer to the LES velocity profile compared to the prior. The SARC model appears to perform reasonably well in this case.

As observed in both the concave and convex cases, while the posterior C_f matches the LES extremely well, the velocity profiles show some degree of variability. This is because C_f is a local quantity and a small local change in the velocity profile may be enough to match C_f . Nevertheless, the improvement in the posterior velocity profiles adds more credibility to the inversion procedure.

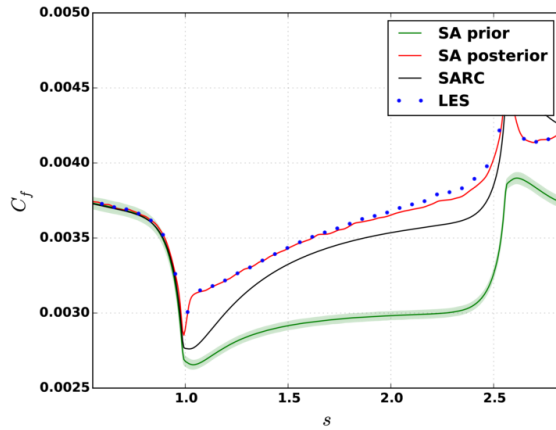


Figure 10: The prior, posterior, SARC and LES coefficient of friction (C_f) at the lower wall for the concave channel with SA model. Shaded region represents the 95% confidence interval.

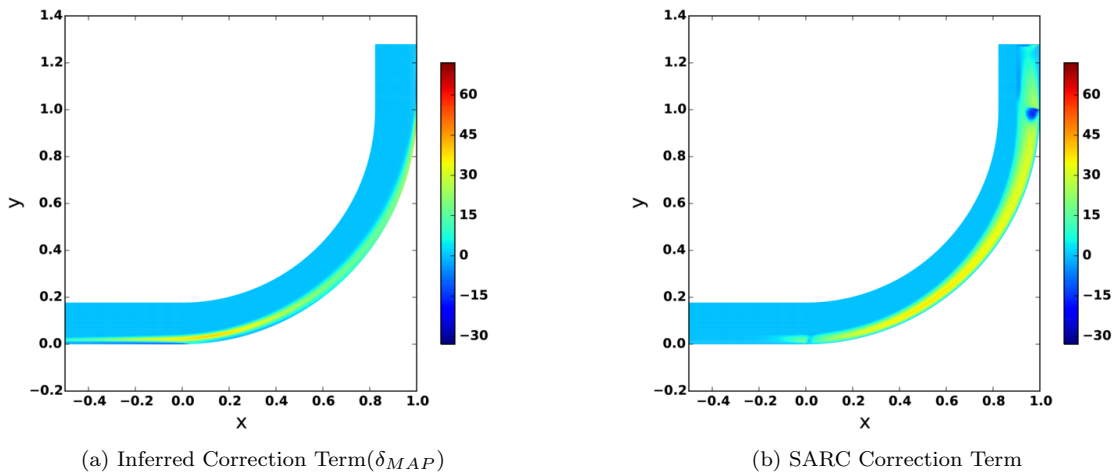


Figure 11: Contour plot showing the inferred correction term (δ_{MAP}) and the SARC correction term for the concave channel.

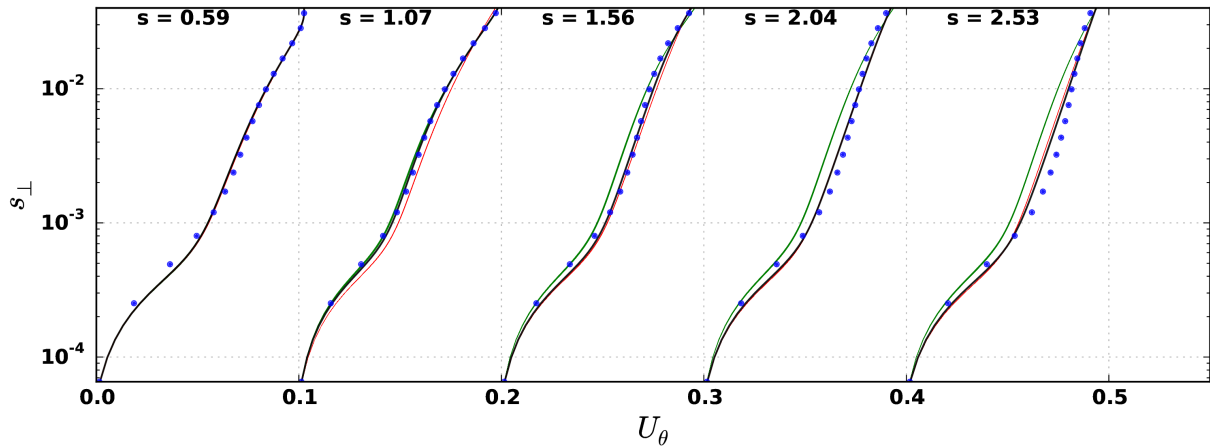


Figure 12: The prior, posterior, SARC and LES wall parallel velocity (U_θ) at various stream wise locations for the concave channel with SA model. s_\perp refers to the perpendicular distance from the lower wall. Refer Fig. 10 for the legend

VI. Oblique Shock Turbulent Boundary Layer Interaction (OSTBLI)

The next test problem considers the interaction of an oblique shock with a spatially evolving turbulent boundary layer. Shock boundary layer interactions have been studied extensively^{39–41} over the years. Knight & Degrez⁴² present an extensive survey for this case and found RANS models to be inaccurate, except for very weak interactions. This behavior is attributed to the inability of RANS models to respond to a sharp pressure rise and the use of a single turbulent scale in an inherently multi-scale environment. We use the Bayesian formulation to infer an improved time-scale estimate for use within a $k - \omega$ model solution.

The flow conditions attempt to emulate the benchmark setup in Morgan et. al,⁴³ which employs a well-resolved Large Eddy Simulation. The incoming boundary layer has a free-stream Mach number of 2.28, an inlet momentum thickness Reynolds number of 1500, and a thickness of 0.87 grid units. The undisturbed shock is aligned at an angle of 8° to the freestream. The inlet boundary layer is generated using a precursor flat plate simulation. Rankine-Hugoniot conditions are used to obtain the shock jump conditions, which are then forced at the inlet plane.

Similar to the previous case, a correction function $\beta(\mathbf{x}, \mathbf{y})$ is multiplied to the production term of the ω equation. The coefficient of pressure (C_p) at the lower wall, from the LES solution,⁴³ is used as the benchmark data for inversion. A diagonal prior covariance with a standard deviation, $\sigma_{prior} = 1$ is used. This corresponds to a nearly uninformative prior. The selection of the prior can indeed be guided by additional insights gained during the inversion process. A diagonal observation covariance with standard deviation, σ_{obs} , of 10^{-6} is used. The posterior covariance and realizations are computed as detailed in section II.A. The flow solver is evaluated for 1000 realizations of β to obtain both the prior and the posterior realizations.

VI.A. Results

Fig. 13 compares the pressure and friction coefficients using the prior, the posterior model, and the LES. The posterior C_p tracks the LES consistently. As evident in the C_f plot, location and length of the separation bubble is more accurately predicted with the posterior model. In contrast with the previous case, the posterior variance has increased for this problem. The Gaussian assumption of the probability distributions is less accurate for this case because of the increased non-linearity. The inadequacy of the Gaussian assumption was verified from the skewed probability distribution of the likelihood function.

Figs. 14a,b show the streamwise-velocity, density profiles along various streamwise locations for the prior and the posterior model. Both of these quantities show strikingly good agreement with data, given the data used for inversion was just the surface pressure, showcasing the ability of the inverse approach to predict the complete flow field using very limited information. These quantities are compared in contour form in Figs. 15- 18. Figure 14c,d compare the Reynolds shear stress and turbulent kinetic energy (TKE). The fact

that the velocity field, pressure, density, TKE and Reynolds stress correlate well with the LES suggests that the improvements are physically realizable and offers the promise of model improvement.

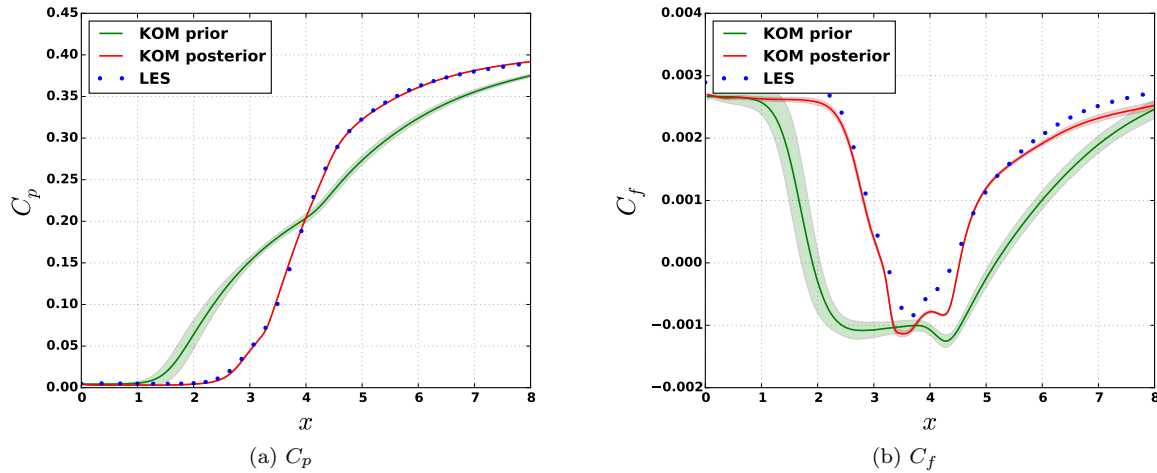


Figure 13: Pressure and friction coefficient for the LES, prior and the posterior $k - \omega$ model. Shaded region represents the 95% confidence interval.

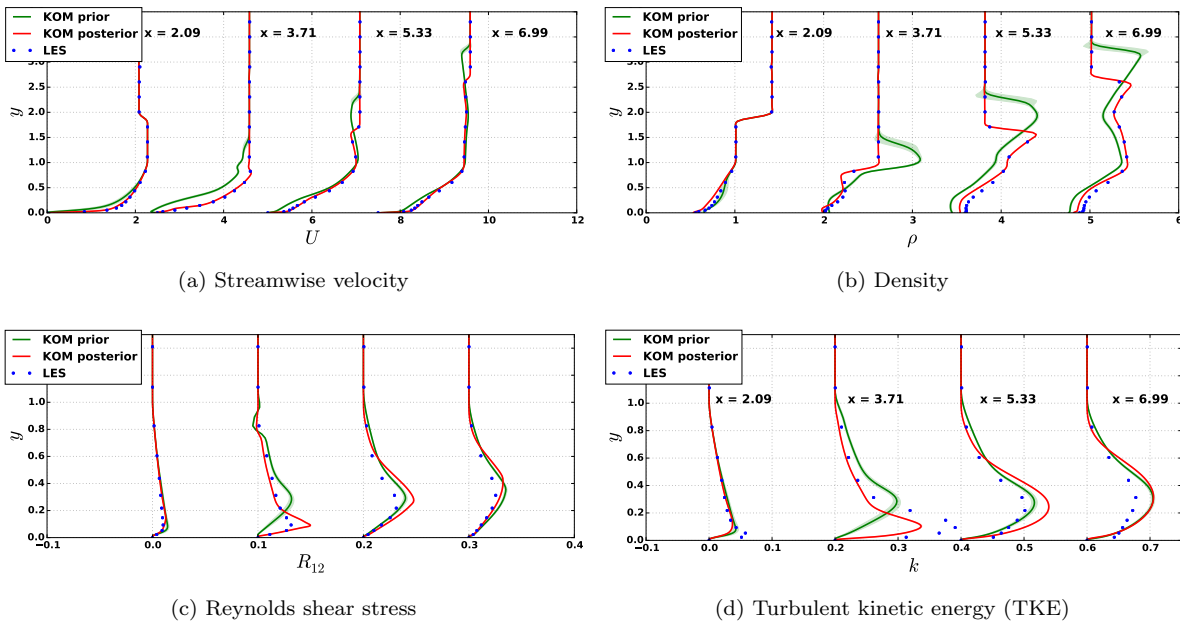


Figure 14: Flow quantities at various streamwise locations with $k - \omega$ model. Shaded region represents the 95% confidence interval.

VII. High Angle of Attack Flow over S809 Airfoil

The next case involves the S809 airfoil, which is extensively used in the blades of horizontal axis wind turbines (HAWTs). At a Reynolds number of 2 million, Somers⁴⁴ observed a laminar separation bubble near mid-chord for angles of attack α between 0 and 5 degrees, which leads to transition to turbulence. A further increase in the angle of attack moves the transition point towards the leading edge. At $\alpha = 9.22^\circ$, the transition point is close to the leading edge, and turbulent separation occurs near the trailing edge. Beyond

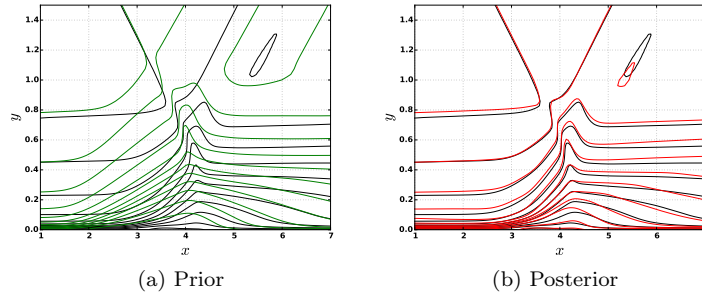


Figure 15: X-velocity contours near the interaction region with $k - \omega$ model. The contours are superimposed with the LES velocity (black) for comparison. The full flow field is well predicted with the inferred MAP $k - \omega$ model.

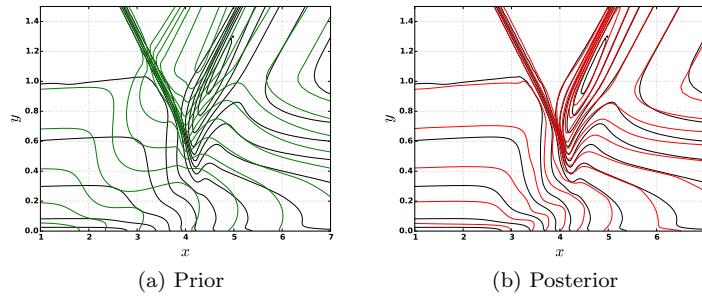


Figure 16: Density contours near the interaction region with $k - \omega$ model. The contours are superimposed with the LES density (black) for comparison.

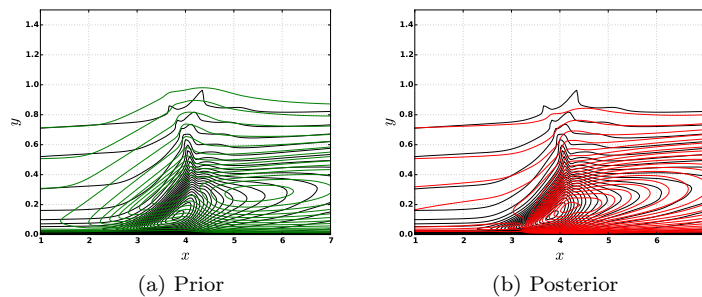


Figure 17: Turbulent kinetic energy (TKE) contours near the interaction region with $k - \omega$ model. The contours are superimposed with the LES TKE (black) for comparison.

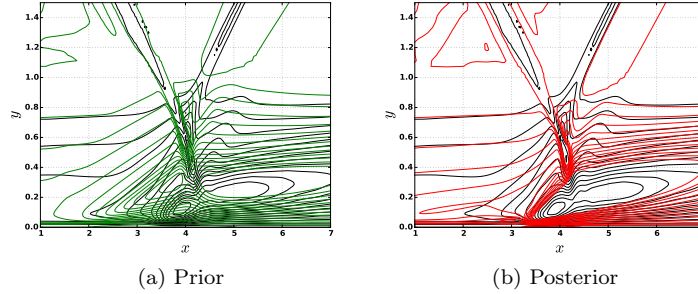


Figure 18: Reynolds shear stress contours near the interaction region with $k - \omega$ model. The contours are superimposed with the LES Reynolds shear stress (black) for comparison.

this point, the flow separates near mid chord. At the design Reynolds number, RANS based models are fairly accurate for flows at low angles of attack ($\leq 9^\circ$). Beyond this regime, the presence of a large region of turbulent separation with an adverse pressure gradient renders RANS inaccurate. Previous work⁴⁵⁻⁴⁷ indicates consistent failure of the RANS models beyond the point of separation. In this work, Bayesian inversion is applied to quantify discrepancies in the one equation Spalart-Allmaras (SA) model by utilizing the experimentally acquired pressure data.

VII.A. Flow Conditions

A Mach number of 0.2 and a Reynolds number of 2 million is used to match the experimental Reynolds number by Somers.⁴⁴ A C-grid with 291 points in the wrap-around direction and 111 points in the wall perpendicular direction is used. This corresponds to 200 grid points on the airfoil surface. The far field boundaries are at least 35 chord lengths away from the airfoil surface. Characteristic based freestream boundary conditions are used for the flow variables at the far-field and the eddy viscosity is set to the fully turbulent value, $\nu_{t,\infty}/\nu_\infty = 3$.

VII.B. Bayesian Inversion Setup

As in the case of curved channel flow, a multiplicative correction function $\beta(\mathbf{x}, \mathbf{y})$ is introduced into the production term. Experimental coefficient of pressure (C_p) is used as the benchmark for inversion. The experimental data is extracted from the report by Somers.⁴⁴ The data extraction is prone to error, especially near the leading edge, where the pressure changes rapidly. The effect of this error is minimized by utilizing *only the suction-side pressure taps, starting from 0.05c to 0.8c*, for the inversion. A diagonal prior with a standard deviation, $\sigma_{prior} = 1.0$ and a diagonal observational covariance with standard deviation, $\sigma_{obs} = 10^{-2}$ are used. This corresponds to a maximum error of 0.2% in the C_p . Note that σ_{obs} should ideally be derived from the experimental data. However the experimental error is unavailable for this case, and the posterior results were confirmed to be insensitive to small changes in the prior covariance. The optimization process is identical to the previous cases. The prior and the posterior model are constructed using 1000 realizations of β .

	$\alpha = 9.22^\circ$	$\alpha = 12.2^\circ$	$\alpha = 14.24^\circ$
Experiment	1.010	1.007	1.083
Prior SA (% error in mean)	$1.166 \pm 0.005(15\%)$	$1.411 \pm 0.017(40\%)$	$1.453 \pm 0.039(34\%)$
Posterior SA (% error in mean)	$1.051 \pm 0.005(4\%)$	$1.028 \pm 0.004(2\%)$	$1.140 \pm 0.004(5\%)$

Table 2: Comparison of prior, posterior and experimental lift coefficients for $\alpha = 9.22^\circ, 12.2^\circ, 14.24^\circ$ with SA model.

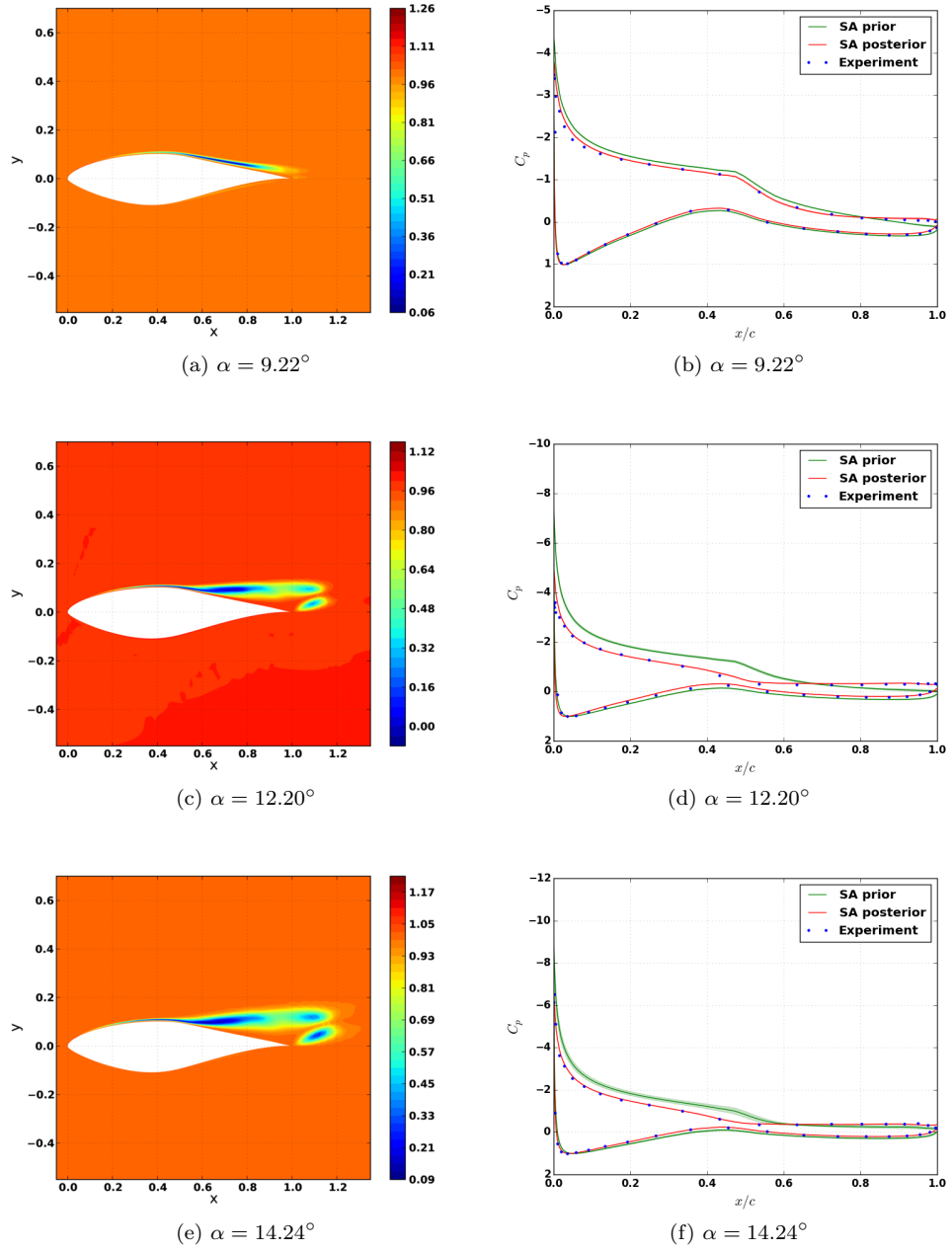


Figure 19: The inferred correction multiplier, $\beta(\mathbf{x}, \mathbf{y})_{MAP}$ (left) and prior/posterior pressure coefficient (right) with SA model. Shaded region represents the 95% confidence interval. Error bars for the experimental results are not provided in the original document.⁴⁴

VII.C. Results

Figure 19 shows the inferred field β_{MAP} for all the cases. Turbulent production is reduced on the suction side. The tendency of the inferred field is to significantly reduce the production of eddy viscosity starting in an extremely thin region close to the leading edge and continuing into the edge of the boundary layer and into the separated shear layer. This reduction is consistent with the transitioning and separating boundary layer and may be used to construct a transition model and a model for adverse pressure gradients. Figure 19 shows that the experimental C_p is well-matched by the inversion, and the posterior variance in C_p is reduced significantly. Though the results are not shown here, it was confirmed that solely by using the pressure data at the **three most upstream locations** on the suction side, the posterior pressure distribution could be improved to the level in Figure 19 over the complete airfoil surface.

Figure 20 compares the near-wall streamlines for the base SA model and the MAP, showing the large differences in the prior and the posterior separation bubble. For $\alpha = 12.20^\circ$ and 14.24° , the posterior model predicts the separation at $0.50c \pm 0.01c$ (which represents a 68% confidence interval) and $0.48c$, respectively, compared to the prior predictions of $0.73c \pm 0.04c$ and $0.56c \pm 0.02c$ respectively. The predictions are consistent with the experiments, where the separation is observed near the mid chord for these angles. For $\alpha = 9.22^\circ$, turbulent separation is observed in the experiment but not in the prior SA model. The posterior, however, predicts the separation at $0.71c \pm 0.01c$. Table 2 compares the prior and the posterior lift coefficient (C_l) with the experiment. For $\alpha = 14.24^\circ$, the error in the C_l is reduced from 34% to 5% and the trend is preserved for other angles of attack. The inferred field β_{MAP} is thus a good starting point for model improvements.

VIII. Summary and Perspectives

A data-driven approach was presented with the objective of quantifying functional errors and uncertainties in turbulence closure models. Bayesian inversion techniques were adapted to extract field variables (source terms) that are consistent with their use in closure equations. The problem is made finite-dimensional by redefining it such that the distribution of the unknown functions is sought at every point in the computational domain of the RANS grid. With the additional assumption of a Gaussian distribution, the inversion is reduced to the solution of a deterministic optimization problem to compute the maximum a posteriori estimate and sampling to obtain the posterior variabilities.

A number of problems are studied, including one-dimensional channel flows, boundary layers with curvature, airfoil flows, and oblique shock-turbulent boundary layer interactions. In all of these cases, given a few data points (typically wall-pressure or skin-friction), complete realizations of the closure model are obtained. For the one-dimensional channel flow, the posterior model suggests a universal time-scale correction in the near wall region. For flows with curvature, the entire flowfield is improved by matching the skin-friction at the wall and the corrections to turbulence production were compared to existing analytical methods. The entire flow field is well-predicted in the oblique shock-turbulent boundary layer interaction problem by matching the surface pressure. The point of separation and lift is more accurately predicted for the S809 airfoil by matching the experimental surface pressure at a few points on the suction side.

Having extracted pertinent data, it remains to use it to improve models. As a first step, a larger, and more diverse set of problems will have to be considered. In this process, care has to be devoted to the choice of the objective function. For instance, the coefficient of pressure is found to be a more *global* benchmark when compared to the friction and leads to a model that is closer to reality. Friction, a highly localized quantity, may lead to a degree of over-fitting by adjusting the model in a small near-wall region in many problems. *The inferred functions can be used to assess the structural form of the model and make improvements.* As an alternate step, machine learning algorithms^{5,6,17} may also be directly used to reconstruct model terms.

Based on the posterior results in this paper, it could also be argued that some of the criticism of eddy viscosity models that is usually attributed to the isotropic stress-strain relationship is partially a result of poor functional forms within the model rather than the limiting nature of the Boussinesq assumption itself. Models should be assessed by evaluating the type and variability of corrections that will be required in a diverse set of problems. Another indicator of the robustness of a model structure and its improvement via field inversion is to inspect the posterior distributions of variables other than those that were used in the inversion.

We note that additional work may be required to assess the role of numerical errors^d, sensitivity to error

^dIn this work, all the RANS results (prior to inversion) were grid-converged

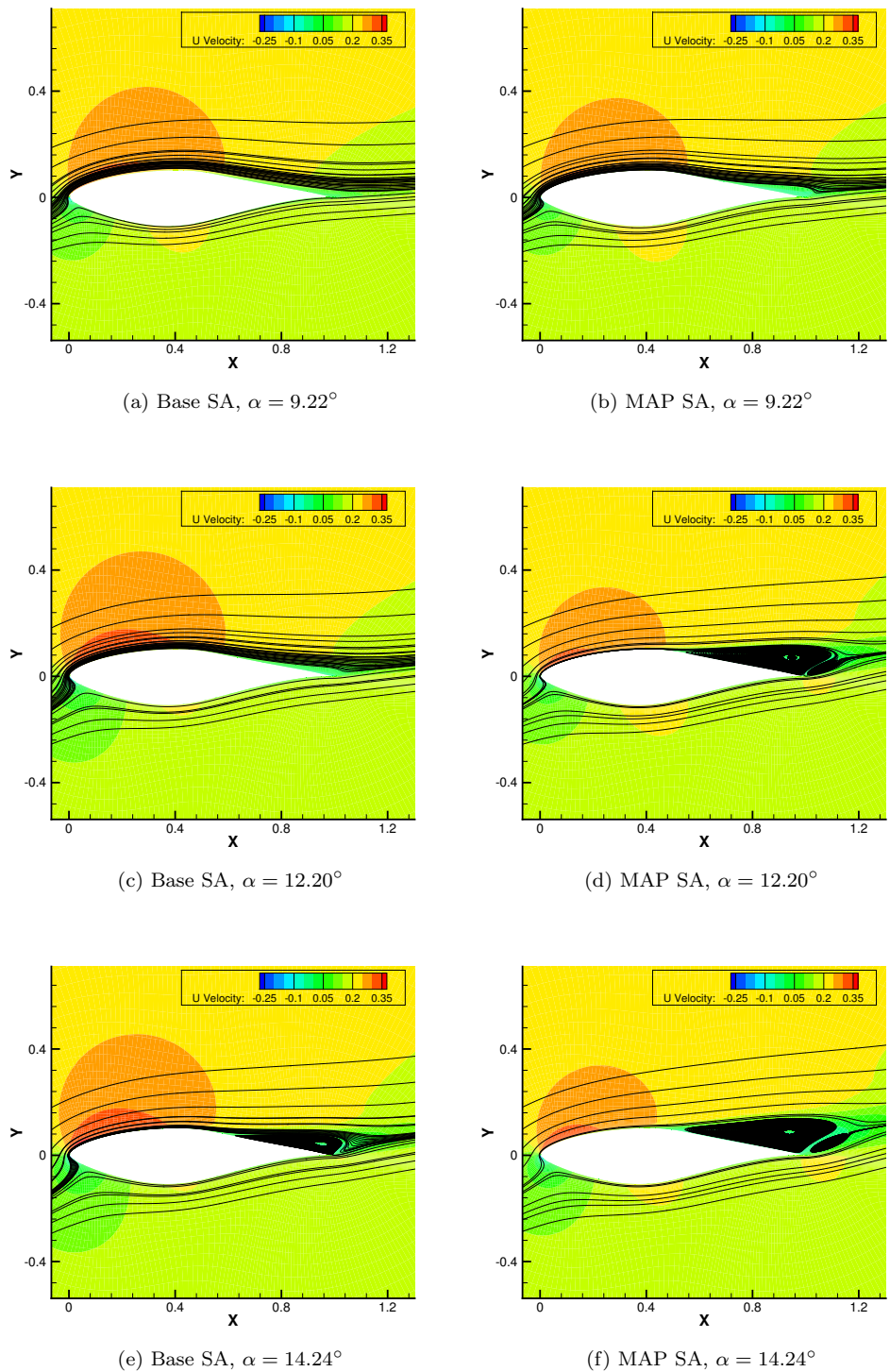


Figure 20: X-velocity contour along with the streamlines near the surface of the airfoil with SA model.

in the data, impact of the Gaussian assumption, etc. Efficiencies can be gained by decomposing the inferred functions in terms of a priori bases and estimating coefficients of basis functions.

Acknowledgments

This work was supported by the NASA Aeronautics Research Institute (NARI) under the Leading Edge Aeronautics Research for NASA (LEARN) fund, with Dr. Gary Coleman as the technical monitor.

References

- ¹Slotnick, J., Khodadoust, A., Alonso, J., Darmofal, D., Gropp, W., Lurie, E., and Mavriplis, D., “CFD vision 2030 study: a path to revolutionary computational aerosciences,” *NASA technical report, NASA/CR-2014-218178, NF1676L-18332*, 2014.
- ²Gerolymos, G. A., Sauret, E., and Vallet, I., “Contribution to single-point closure Reynolds-stress modelling of inhomogeneous flow,” *Theoretical and Computational Fluid Dynamics*, Vol. 17, No. 5-6, 2004, pp. 407–431.
- ³Olsen, M. E., “Prediction of Separation with a Third-Order-Moment Model,” *AIAA SciTech*, Jan 2015.
- ⁴Poroseva, S. and Murman, S. M., “Velocity/Pressure-Gradient Correlations in a FORANS Approach to Turbulence Modeling,” *AIAA Aviation*, Jun 2014.
- ⁵Parish, E. and Duraisamy, K., “A Paradigm for Data-Driven Modeling using Functional Inversion and Machine Learning,” *Accepted at Journal of Computational Physics*, 2015.
- ⁶Tracey, B. D., Duraisamy, K., and Alonso, J. J., “A Machine Learning Strategy to Assist Turbulence Model Development,” *AIAA SciTech*, Jan 2015.
- ⁷Duraisamy, K., Zhang, Z. J., and Singh, A. P., “New Approaches in Turbulence and Transition Modeling Using Data-driven Techniques,” *AIAA SciTech*, Jan 2015.
- ⁸Parish, E. and Duraisamy, K., “Quantification of Turbulence Modeling Uncertainties Using Full Field Inversion,” *AIAA Aviation*, Jun 2015, 0.
- ⁹Zhang, Z. J. and Duraisamy, K., “Machine Learning Methods for Data-Driven Turbulence Modeling,” *AIAA Aviation*, Jun 2015.
- ¹⁰Poroseva, S. V. and Murman, S. M., “Reynolds-Stress simulations of a fully-developed channel flow using a new velocity/pressure-gradient model,” *International Symposium on Turbulence and Shear Flow Phenomena (TSFP-9)*, 2015.
- ¹¹Yarlanki, S., Rajendran, B., and Hamann, H., “Estimation of turbulence closure coefficients for data centers using machine learning algorithms,” *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012 13th IEEE Intersociety Conference on*, IEEE, 2012, pp. 38–42.
- ¹²Kato, H. and Obayashi, S., “Data Assimilation for Turbulent Flows,” *AIAA SciTech*, Jan 2014.
- ¹³Edeling, W., Cinnella, P., Dwight, R. P., and Bijl, H., “Bayesian estimates of parameter variability in the $k-\epsilon$ turbulence model,” *Journal of Computational Physics*, Vol. 258, 2014, pp. 73–94.
- ¹⁴Dow, E. and Wang, Q., “Quantification of Structural Uncertainties in the $k-\omega$ Turbulence Model,” *Structures, Structural Dynamics, and Materials and Co-located Conferences*, Apr 2011.
- ¹⁵Cheung, S. H., Oliver, T. A., Prudencio, E. E., Prudhomme, S., and Moser, R. D., “Bayesian uncertainty analysis with applications to turbulence modeling,” *Reliability Engineering & System Safety*, Vol. 96, No. 9, 2011, pp. 1137–1149.
- ¹⁶Emory, M., Larsson, J., and Iaccarino, G., “Modeling of structural uncertainties in Reynolds-averaged Navier-Stokes closures,” *Physics of Fluids*, Vol. 25, No. 11, 2013.
- ¹⁷Ling, J. and Templeton, J., “Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty,” *Physics of Fluids (1994-present)*, Vol. 27, No. 8, 2015, pp. 085103.
- ¹⁸Stuart, A. M., “Inverse problems: a Bayesian perspective,” *Acta Numerica*, Vol. 19, 2010, pp. 451–559.
- ¹⁹Aster, R. C., Borchers, B., and Thurber, C. H., *Parameter estimation and inverse problems*, Academic Press, 2013.
- ²⁰Duraisamy, K. and Alonso, J., “Adjoint based techniques for uncertainty quantification in turbulent flows with combustion,” *42nd AIAA Fluid Dynamics Conference and Exhibit*, 2012, p. 2711.
- ²¹Pecnik, R., Terrapon, V. E., Ham, F., Iaccarino, G., and Pitsch, H., “Reynolds-averaged Navier-Stokes simulations of the HyShot II scramjet,” *AIAA journal*, Vol. 50, No. 8, 2012, pp. 1717–1732.
- ²²Wang, Q., Duraisamy, K., Alonso, J. J., and Iaccarino, G., “Risk assessment of scramjet unstart using adjoint-based sampling methods,” *AIAA journal*, Vol. 50, No. 3, 2012, pp. 581–592.
- ²³Giles, M. B., Duta, Mihai C., J.-D. M., and Pierce, N. A., “Algorithm developments for discrete adjoint methods,” *AIAA journal*, Vol. 41, No. 2, 2003, pp. 198–205.
- ²⁴Walther, A. and Griewank, A., “Getting started with ADOL-C,” *Combinatorial Scientific Computing*, 2012, pp. 181–202.
- ²⁵Hoyas, S. and Jiménez, J., “Scaling of the velocity fluctuations in turbulent channels up to $Re_\tau = 2003$,” *Physics of Fluids (1994-present)*, Vol. 18, No. 1, 2006, pp. 011702.
- ²⁶Lee, M. and Moser, R. D., “Direct numerical simulation of turbulent channel flow up to $Re_\tau = 5200$,” *arXiv preprint arXiv:1410.7809*, 2014.
- ²⁷del Alamo, J. C. and Jimenez, J., “Direct numerical simulation of the very large anisotropic scales in a turbulent channel,” *arXiv preprint arXiv:1309.2322*, 2013.
- ²⁸Wilcox, D. C., *Turbulence modeling for CFD*, Vol. 3, D C W Industries, 2006.
- ²⁹Bradshaw, P., “The analogy between streamline curvature and buoyancy in turbulent shear flow,” *Journal of Fluid Mechanics*, Vol. 36, No. 01, 1969, pp. 177–191.

- ³⁰Gillis, J. and Johnston, J., "Turbulent boundary-layer flow and structure on a convex wall and its redevelopment on a flat wall," *Journal of Fluid Mechanics*, Vol. 135, 1983, pp. 123–153.
- ³¹Patel, V. and Sotiropoulos, F., "Longitudinal curvature effects in turbulent boundary layers," *Progress in Aerospace Sciences*, Vol. 33, No. 1, 1997, pp. 1–70.
- ³²Saric, W. S., "Gortler vortices," *Annual Review of Fluid Mechanics*, Vol. 26, No. 1, 1994, pp. 379–409.
- ³³Spalart, P. and Shur, M., "On the sensitization of turbulence models to rotation and curvature," *Aerospace Science and Technology*, Vol. 1, No. 5, 1997, pp. 297–302.
- ³⁴Hellsten, A., "Some improvements in Menter's $k-\omega$ SST turbulence model," *Fluid Dynamics and Co-located Conferences*, Jun 1998.
- ³⁵Duraisamy, K. and Iaccarino, G., "Curvature correction and application of the v2-f turbulence model to tip vortex flows," *Center for Turbulence Research Annual Research Briefs*, 2005, pp. 157–168.
- ³⁶Shur, M. L., Strelets, M. K., Travin, A. K., and Spalart, P. R., "Turbulence modeling in rotating and curved channels: Assessing the Spalart-Shur correction," *AIAA journal*, Vol. 38, No. 5, 2000, pp. 784–792.
- ³⁷Arolla, S. K. and Durbin, P. A., "LES of spatially developing turbulent boundary layer over a concave surface," *Journal of Turbulence*, Vol. 16, No. 1, 2015, pp. 81–99.
- ³⁸Durbin, P., Personal communication, Iowa State University.
- ³⁹Ferri, A., "Experimental Results with airfoils tested in the High-speed Tunnel at Guidonia," *National Advisory Committee for Aeronautics*, 1940.
- ⁴⁰Dolling, D. S. and Murphy, M. T., "Unsteadiness of the Separation Shock Wave Structure in a Supersonic Compression Ramp Flowfield," *AIAA Journal*, Vol. 21, No. 12, Dec 1983, pp. 1628–1634.
- ⁴¹Dolling, D. S., "Fifty Years of Shock-Wave/Boundary-Layer Interaction Research: What Next?" *AIAA Journal*, Vol. 39, No. 8, Aug 2001, pp. 1517–1531.
- ⁴²Knight, D. D. and Degrez, G., "Shock wave boundary layer interactions in high Mach number flows: A critical survey of current CFD prediction capabilities," *Hypersonic experimental and computational capability, improvement and validation*, 1998.
- ⁴³Morgan, B., Duraisamy, K., Nguyen, N., Kawai, S., and Lele, S., "Flow physics and RANS modelling of oblique shock/turbulent boundary layer interaction," *Journal of Fluid Mechanics*, Vol. 729, 2013, pp. 231–284.
- ⁴⁴Somers, D., "Design and Experimental Results for the S809 Airfoil," *NREL/SR-440-6918*, 1997.
- ⁴⁵Chang, Y., Yang, S., and Arici, O., "Flow field computation of the NREL S809 airfoil using various turbulence models," Tech. rep., American Society of Mechanical Engineers, New York, NY (United States), 1996.
- ⁴⁶Yang, S., Chang, Y., and Arici, O., "Incompressible Navier-Stokes computation of the NREL airfoils using a symmetric total variational diminishing scheme," *Journal of solar energy engineering*, Vol. 116, No. 4, 1994, pp. 174–182.
- ⁴⁷Wolfe, W. and Ochs, S., "CFD calculations of S809 aerodynamic characteristics," *Aerospace Sciences Meetings*, Jan 1997.