AIAA JOURNAL
Vol. 60, No. 1, January 2022

# Assessment of Regularized Ensemble Kalman Method for Inversion of Turbulence Quantity Fields

Xin-Lei Zhang*
*State Key Laboratory of Nonlinear Mechanics, Chinese Academy of Sciences, 100190 Beijing, People's Republic of China*
Heng Xiao†
*Virginia Tech, Blacksburg, Virginia 24060*
and
Guowei He‡
*State Key Laboratory of Nonlinear Mechanics, Chinese Academy of Sciences, 100190 Beijing, People's Republic of China*

https://doi.org/10.2514/1.J060976

**This paper introduces an ensemble-based field inversion framework to augment the turbulence models by incorporating prior physical knowledge. Different types of prior knowledge such as smoothness, prior values, and sparsity are enforced to improve the inference of the eddy viscosity and laminar–turbulent intermittency. This work first assesses the method on the problems of inferring eddy viscosity in the Reynolds-averaged Navier–Stokes equation from the velocity observation data for separated flows over periodic hills. Further, the method is used to infer the intermittency field in the transport equation of turbulent kinetic energy from measurements of the friction coefficient for transitional flows over a plate. The results demonstrate the performance of the regularized ensemble method by enforcing prior knowledge into the inference. The method serves as a useful inverse modeling tool to augment the turbulence model from observation data.**

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

## Nomenclature

| | | |
|---|---|---|
| $C_f$ | = | friction coefficient |
| $D$ | = | dimension of observation space |
| $\mathcal{G}$ | = | regularization function |
| $H$ | = | height of hill crest or half channel |
| $\mathsf{H}$ | = | tangent linear operator; $\mathsf{H} \in \mathbb{R}^{D\times N}$ |
| $J$ | = | cost function |
| $\mathcal{K}$ | = | kernel function |
| $k$ | = | turbulent kinetic energy |
| $M$ | = | size of ensemble |
| $N$ | = | dimension of parameter space |
| $\mathcal{N}$ | = | Gaussian process |
| $\mathcal{P}$ | = | turbulent kinetic energy production term |
| $\mathsf{P}$ | = | covariance of model error; $\mathsf{P} \in \mathbb{R}^{N\times N}$ |
| $\mathsf{R}$ | = | covariance of observation errors; $\mathsf{R} \in \mathbb{R}^{D\times D}$ |
| $S$ | = | sigmoid function |
| $T$ | = | Chebyshev basis |
| $U$ | = | mean velocity |
| $u_\tau$ | = | friction velocity |
| $\mathsf{W}$ | = | weight for regularization term |
| $w$ | = | mode coefficients |
| $\mathsf{x}$ | = | state vector; $\mathsf{x} \in \mathbb{R}^{N}$ |
| $x$ | = | spatial coordinate |
| $x_{\text{tr}}$ | = | location of transition point |
| $\mathsf{y}$ | = | observation vector; $\mathsf{y} \in \mathbb{R}^{D}$ |
| $y_{\text{crit}}$ | = | critical curve for the interface between laminar flow and turbulent flow |
| $\partial$ | = | finite difference operator |
| $\beta^*, \sigma_k$ | = | model parameter in turbulent kinetic energy transport equation |
| $\gamma$ | = | laminar–turbulent intermittency |
| $\delta$ | = | regularization correction |
| $\epsilon$ | = | dissipation rate |
| $\lambda$ | = | regularization parameter |
| $\nu$ | = | molecular viscosity |
| $\nu_t$ | = | eddy viscosity |
| $\sigma_1, \sigma_2$ | = | parameters in sparse representation of intermittency field |
| $\tau$ | = | Reynolds stress |
| $\phi$ | = | basis of Karhunen–Loève expansion |
| $\omega$ | = | specific dissipation rate |

*Subscripts*

| | | |
|---|---|---|
| $i$ | = | index of spatial dimension |
| $j$ | = | index of sample |
| $n$ | = | index of mode |

*Superscripts*

| | | |
|---|---|---|
| $a$ | = | analysis |
| $f$ | = | forecast (propagated by using dynamic model) |
| $l$ | = | index of spatial location |
| $\mathsf{T}$ | = | transpose |
| 0 | = | prior |
| $\prime$ | = | gradient |
| $-$ | = | mean |

Received 7 June 2021; revision received 24 October 2021; accepted for publication 24 October 2021; published online 24 November 2021. Copyright © 2021 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved. All requests for copying and permission to reprint should be submitted to CCC at www.copyright.com; employ the eISSN 1533-385X to initiate your request. See also AIAA Rights and Permissions www.aiaa.org/randp.

*Postdoctoral Research Associate, Institute of Mechanics; also School of Engineering Sciences, University of Chinese Academy of Sciences, 100049 Beijing, People's Republic of China; zhangxinlei@imech.ac.cn.

†Associate Professor, Kevin T. Crofton Department of Aerospace and Ocean Engineering; hengxiao@vt.edu.

‡Professor, Institute of Mechanics; also School of Engineering Sciences, University of Chinese Academy of Sciences, 100049 Beijing, People's Republic of China; hgw@lnm.imech.ac.cn.

## I. Introduction

THE Reynolds-averaged Navier–Stokes (RANS) method is one of the most widely used approaches for estimation of turbulent mean flows, in which the Reynolds stress requires modeling based on the mean flow quantities. The Reynolds stress-related models have been developed over decades, but there is still no universal model that makes flow predictions accurately and robustly. Field inversion has

emerged as a promising method to assist the RANS modeling by inferring modeled quantities, e.g., eddy viscosity [1], from available observation data. However, the field inversion problem is usually ill-posed. That is, the optimal solutions are not unique, and different fields can lead to good agreement with observation data. To this end, enforcing additional regularization based on prior knowledge is usually required to alleviate the ill-posedness of the field inversion problem and thus improve the accuracy of the inferred field.

Various types of prior knowledge are desired for the modeled quantities in turbulence models. For example, Wang et al. [2] embedded some specific types of prior knowledge (e.g., realizability of Reynolds stresses, variance of model error in different regions) in their problem formulation and demonstrated that these types of prior knowledge could improve inversion results and reduce the amount of required data. However, it is more desirable to incorporate more general prior knowledge such as smoothness, prior values from existing models, and sparsity on a chosen basis.

Both the adjoint-based [3] and ensemble-based [4] method can be used to solve the inverse problem and meanwhile enforce the general prior knowledge. The adjoint method is able to enforce prior knowledge through adding the regularization term in the objective function. Various regularization terms have been formulated to ensure the desired property of inferred fields [1,5] and applied for field inversion in different flow configurations [6,7]. However, most computational fluid dynamics solvers have no readily available adjoint capabilities and need intrusive modifications to develop the adjoint solver. On the other hand, ensemble methods use an ensemble of simulations to estimate the model sensitivity and thus do not require extra efforts for the development of adjoint solver. For this reason, the ensemble-based methods have been increasingly used to solve inverse problems in turbulence modeling [8–11], flow state estimation [12–14], and fluid mechanics in general applications [15].

Traditional ensemble methods are incapable of using regularization to enforce general prior knowledge, e.g., smoothness, as in adjoint-based methods. The common approach to enforce the desired properties with the ensemble methods is through the model error covariance, which can indicate the spatial properties of the inferred solution. Specifically, smoothness can be achieved by generating smooth samples based on the Gaussian process where the Gaussian kernel is chosen as the model error covariance. These prior realizations can enforce the smoothness, because the inferred field is within the subspace spanned by these smooth samples [16]. As to constraints from prior values, the model error covariance can be prescribed with specific variance. That is, one can assign small values to the prior variance, indicating the high degree of confidence on the prior values, to avoid large deviation from prior. Nevertheless, enforcing these properties through specific covariance is not sufficient on some occasions. For example, in the case where we need to ensure the nonnegativity of the inferred field, one commonly takes the logarithm on the inferred quantity. As a result, smoothing the prior samples can only ensure the smoothness in the logarithmic space, which still allows large gradients in the linear space. On the other hand, the prior information about model error covariance is often not available and cannot be enforced during the inference. Moreover, the inference can deviate significantly from the prior value due to the repeated use of the observation data in the steady-state scenario [17]. To this end, it is desired to devise regularization schemes for enforcing general prior knowledge in the ensemble method. In this context, a regularized ensemble Kalman method [18] was proposed to enforce general regularization for ensemble-based inversion.

The regularized ensemble Kalman method inherits the advantage of ensemble methods in that it does not require extra efforts in developing an adjoint solver. Moreover, it provides a derivative-free update scheme capable of enforcing general regularization, with only minor modifications to the conventional ensemble Kalman method. It has been demonstrated that this method is capable of enforcing straightforward constraints on simple inference problems [18]. However, enforcement of more general prior knowledge such as smoothness and prior values has not been explored and warrants further investigations. Such constraints are more common in practical applications.

In this work we demonstrate applications of the regularized ensemble Kalman method on practical turbulence modeling problems. Specifically, our contributions include 1) the derivation and implementation of such common constraints as smoothness, prior values, and sparsity in the general framework proposed in [18], and 2) demonstration of the merits in incorporating such prior knowledge in inferring turbulence quantity fields such as eddy viscosity and intermittency. Moreover, this work highlights the need of regularization in the inference for turbulence quantity fields and demonstrates the capability of ensemble-based approach in enforcing such general regularization.

The rest of the paper is structured as follows. In Sec. II, the ensemble-based field inversion framework is introduced along the problem of inferring turbulence quantity fields. In Sec. III, the method is assessed on two canonical cases of separated flows and transitional flows, demonstrating the merits of the proposed framework. Finally, the paper is concluded in Sec. IV.

## II. Ensemble-Based Inverse Modeling Methodology

In this section, we describe the inverse modeling methodology based on the regularized ensemble Kalman method. Moreover, the turbulence quantity to be inferred is described as well as the prior knowledge on the turbulence field.

### A. Formulation of Regularized Ensemble Kalman Method

We employ a regularized ensemble Kalman method [18] to enforce regularization into the inference process. The ensemble method uses the Monte Carlo technique to draw an ensemble of samples for the inferred quantities and estimates the model error covariance with these samples. Further, the optimal solution is searched in the space spanned by these samples through minimizing a cost function. Specifically, for a generic constraint $\mathcal{G}[\mathsf{x}]$ on the inferred state $\mathsf{x}$, the cost function with such regularization can be written as

$$J = \|\mathsf{x}_j^a - \mathsf{x}_j^f\|_{\mathsf{P}^{-1}}^2 + \|\mathsf{y}_j - \mathsf{H}\mathsf{x}_j^a\|_{\mathsf{R}^{-1}}^2 + \lambda\|\mathcal{G}[\mathsf{x}_j^a]\|_{\mathsf{W}^{-1}}^2 \tag{1}$$

In this equation, superscripts "$a$" and "$f$" denote "analysis" and "forecast" following the convention of nomenclature in data assimilation, $\mathsf{x}$ is the quantities to be inferred, $\mathsf{H}$ is the model operator that maps the inferred variable $\mathsf{x}$ to the observation space, $\mathsf{P}$ is the model error covariance, $\mathsf{R}$ is the observation error covariance, $\lambda$ is the regularization parameter to control the tradeoff between the penalty term and the data fit, and $\mathsf{W}$ is the weight for the penalty function. The first term on the right-hand side of Eq. (1) is the departure of the analysis from the forecast state $\mathsf{x}$, the second term represents the data discrepancy, and the third term is the regularization. The sum of the three terms, which form the cost function $J$, is to be minimized.

By setting the gradient of the cost function to be zero, the analysis scheme of the regularized ensemble Kalman method can be formulated as [18]

$$\tilde{\mathsf{x}}_j^f = \mathsf{x}_j^f + \delta, \quad \text{with } \delta = -\lambda\mathsf{P}\mathcal{G}'[\mathsf{x}_j^f]\mathsf{W}^{-1}\mathcal{G}[\mathsf{x}_j^f] \tag{2a}$$

$$\text{and } \mathsf{x}_j^a = \tilde{\mathsf{x}}_j^f + \mathsf{P}\mathsf{H}^{\mathsf{T}}(\mathsf{H}\mathsf{P}\mathsf{H}^{\mathsf{T}} + \mathsf{R})^{-1}(\mathsf{y}_j - \mathsf{H}\tilde{\mathsf{x}}_j^f) \tag{2b}$$

where $\delta$ is the regularization correction term, and $\mathcal{G}'[\mathsf{x}]$ is the derivative of penalty function with respect to state $\mathsf{x}$. The scheme above uses a similar Kalman update as in the traditional ensemble Kalman method but adds an extra regularization step in Eq. (2a). The method inherits the advantages of the ensemble Kalman method, which is applicable to nonlinear problems [19,20].

### B. Formulation of Prior Knowledge for Turbulence Quantity Fields

Different prior knowledge can be used to improve the inference of the turbulence quantities. Here we formulate the regularization in the ensemble-based framework to enforce three representative constraints on the inferred fields: 1) smoothness, 2) prior values, and 3) sparsity.

The regularization term for smoothness can be formulated as the functional derivatives or the total variation in a discrete manner. The

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

functional derivative measures the change rate of value in adjacent grids, and the total variation measures the difference among the values in the adjacent mesh cells. Both the functional derivative and total variation are able to measure the smoothness of a field and have been adopted in different contexts [1,21]. In this work, we use the total variation as the regularization to favor that the values in the adjacent grids differ only moderately. The regularization term can be written as

$$\mathcal{G}[\mathsf{x}] = \partial\mathsf{x}, \quad \text{and } \mathcal{G}'[\mathsf{x}] = \partial \tag{3}$$

where $\partial$ is the discrete difference operator. Many latent fields, such as eddy viscosity, laminar–turbulent intermittency, and thermal diffusivity, are assumed to be positive. In such cases, the logarithm is usually imposed on the turbulence fields to ensure the nonnegativity, and thus the logarithmic quantity is regarded as the inferred quantity. The smoothness property is often desired for the inferred field in the linear space instead of the logarithmic space. Accordingly, the regularization term need to be formulated as

$$\mathcal{G}[\mathsf{x}] = \partial e^{\mathsf{x}}, \quad \text{and } \mathcal{G}'[\mathsf{x}] = \partial \mathrm{diag}(e^{\mathsf{x}}) \tag{4}$$

Note that in this work we use the first-order operator to enforce the smoothness, but the method can be extended to use high-order operators to measure the smoothness. The weight for the constraint $\mathsf{W}$ can be constructed simply as an identity matrix. For field inference problems, Karhunen–Loève (KL) expansion is often used to reduce the dimensionality of the problem, and only the mode coefficients are inferred [8]. In such a representation, the present framework can also be used to impose smoothness constraints. This is derived and presented in Appendix A.

The constraint on the prior value is essentially implemented in the conventional ensemble Kalman method as in the first term of the Eq. (1). However, in the stationary scenario, the iterative ensemble Kalman method updates the state with artificial dynamics. This results in the prior distribution shifting away from the initial guess [17]. To this end, an additional regularization term can be formulated in the regularized ensemble Kalman method as

$$\mathcal{G}[\mathsf{x}] = \mathsf{x} - \mathsf{x}^0, \quad \text{and } \mathcal{G}'[\mathsf{x}] = \mathsf{I} \tag{5}$$

to minimize the departure from the initial estimation $\mathsf{x}^0$.

The sparsity is often desired in reduced-order representations for the inferred field with a given basis such as the KL modes. In this case, the inferred quantities are the mode coefficients, i.e., $\mathsf{x} = \boldsymbol{w}$. It is usually desirable to represent the field with fewer modes. In other words, the number of nonzero mode coefficients $w$ is expected to be as small as possible. To this end, the regularization for sparsity constraint can be formulated as

$$\mathcal{G}[\mathsf{x}] = \mathsf{x}, \quad \text{and } \mathcal{G}'[\mathsf{x}] = \mathsf{I} \tag{6}$$

Lasso regularization based on L1 norm is often used to yield sparsity [22]. Here, to facilitate analytical derivations of the derivative $\mathcal{G}'[\mathsf{x}]$, we use shrinkage method based on L2 norm [23] along with thresholding method [22] to achieve sparsity. Specifically, we set $\mathsf{x} = 0$ for $\mathsf{x} < \sqrt{2\lambda\|\mathsf{W}\|}$.

Other constraints, such as symmetry and physical realizability, are also able to be employed in the framework. For instance, the symmetric constraints can be enforced by penalizing the difference at the symmetric locations. The physical realizability can be considered by formulating the inequality constraints to bound the constructed Reynolds stress within the Lumley triangle. The inequality constraint can be achieved by formulating the regularization term as a specific bounding function. The reader is referred to Ref. [18] for details about enforcing inequality constraints with the regularized method. Besides, other physical constraints such as the divergence-free condition for incompressible flows can be also enforced by penalizing the divergence of the reconstructed velocity field.

### C. Turbulence Model Uncertainties

The RANS method is not able to predict the flows accurately in the presence of separation and adverse pressure gradient, which is mainly due to the model-form uncertainties in Reynolds stress [24]. The Reynolds stress indicates the nonlinear effects of small-scale turbulence and usually requires modeling based on the mean flow quantities. Various models, such as the linear eddy viscosity model [25], the nonlinear eddy viscosity model [26], the explicit algebraic Reynolds stress model [27], and the Reynolds stress transport model [28], have been developed over years. All these models and the data-driven counterparts (e.g., [8,29–32]) based thereon remain an important topic in the turbulence modeling community today. Because of the ease of implementation, the linear eddy viscosity model is still the most widely used model, particularly in industrial applications. Hence, in this work, we focus on the uncertainties of the linear eddy viscosity model. The extensively used linear eddy viscosity models, such as Spalart–Allmaras model [33], $k-\varepsilon$ [25], $k-\omega$ [34], and $k-\omega$ SST model [35], estimate the turbulent eddy viscosity through introducing additional assumptions such as the mixing length hypothesis and equilibrium turbulence [36]. These modeling assumptions often lead to the poor predictive performance of the RANS method. For this reason, turbulence model uncertainties have been proposed and introduced in different manners, such as in the momentum equation [37], in the Reynolds stress [8,38,39], in the eddy viscosity [1], and in the model transport equations for turbulence quantities [5]. Two types of the RANS model uncertainties are investigated in this work: 1) uncertainties in the modeled eddy viscosity, and 2) uncertainties in the equation governing the turbulence quantities, e.g., the turbulent kinetic energy (TKE). That is,

$$\begin{aligned} \nu_t &= \nu_t^{\mathrm{RANS}} \exp[f_{\nu_t}], \\ \frac{\partial k}{\partial t} + \frac{\partial (U_j k)}{\partial x_j} &= \gamma \mathcal{P} - \beta^* k\omega + \frac{\partial}{\partial x_j}\left((\nu + \sigma_k \nu_t)\frac{\partial k}{\partial x_j}\right) \end{aligned} \tag{7}$$

where $f_{\nu_t}$ is a multiplicative correction to the RANS-modeled eddy viscosity, and $\gamma$ is the laminar–turbulent intermittency, which can be obtained from transition models [40]. In the TKE transport equation, the term $\gamma\mathcal{P}$ represents the turbulence production, and the term $\beta^* k\omega$ is the turbulence dissipation, where the parameter $\beta^*$ is chosen as 0.09 in this work. The introduced uncertainty $f_{\nu_t}$ is flexible to correct the magnitude of the Reynolds stresses. However, it has been observed that nonphysical correction with strong gradient is often obtained due to the ill-posedness [6], and hence smoothness regularization is required to penalize these strong gradients during the inference. On the other hand, the intermittency $\gamma$ in the TKE transport equation can provide the smooth $\nu_t$ field, due to the advection–diffusion effects in the transport equation [6]. In this work, we will consider the inference of the two turbulence quantify fields to assess the ensemble-based inverse modeling framework.

### D. Practical Implementation

The framework solves inverse problems in a statistical formulation. First, it requires generating an ensemble of realizations from the Gaussian process. We draw the prior samples $\mathsf{x}$ based on the KL decomposition to ensure the smoothness of the prior samples and then estimate model error covariance $\mathsf{P}$ with these samples. The specific inference procedure is as follows. Given the prior field $\mathsf{x}^0$, kernel function $\mathcal{K}$ of a Gaussian process, and the observation error covariance $\mathsf{R}$, the procedure is as follows:

1) *Sampling:* Construct the prior ensemble from a Gaussian process as $\mathsf{x} = \{\mathsf{x}_j\}_{j=1}^M \sim \mathcal{N}(\mathsf{x}^0, \mathcal{K})$.

2) *Propagation:* Propagate each sample to the observation space through the forward model as $\hat{\mathsf{y}}_j = \mathsf{H}\mathsf{x}_j$.

3) *Regularization:* Construct the constraint term $\mathcal{G}[\mathsf{x}]$ as well as the gradient of the constraint term $\mathcal{G}'[\mathsf{x}]$ for the desired properties, and then obtain the regularized correction field $\bar{\mathsf{x}}^f$ based on Eq. (2a).

(4) *Kalman update:* Update the regularized field $\tilde{\mathsf{x}}$ based on Eq. (2b). Return to step 2 until the convergence is reached.

The convergence is considered achieved when a) the data discrepancy error is smaller than the observation error and b) the penalty term is smaller than the constraint error in the objective function in Eq. (1). The regularization parameter $\lambda$ in Eq. (2a) is set to ramp up gradually to its full value in about 10 steps to ensure the robustness of the inference. See [18] for further details. The regularized ensemble Kalman method is implemented in the open-source code Data Assimilation and Field Inversion (DAFI) [41].

## III. Assessment for Performance of Ensemble-Based Inverse Modeling Method

Two canonical test cases are used to assess the ensemble-based method for the inference of turbulence quantity fields: 1) separated flow over periodic hills, and 2) the transitional flow over flat plates. We apply the ensemble-based inverse modeling to improve the mean flow estimation by inferring the eddy viscosity for the periodic hill case and the laminar–turbulent intermittency for the flat plate case. Moreover, the method is validated on a simpler case of inferring eddy viscosity from mean velocities in the plane channel flow. This is presented in Appendix B.

### A. Separated Flows over Periodic Hills

The turbulent flow over periodic hills [42] is one of the canonical test cases for validation of numerical simulations. Here, we apply this configuration to demonstrate the ability of the proposed method to infer the eddy viscosity. The Reynolds number based on the height of crest and bulk velocity is $Re = 5600$. The unstructured mesh is generated with 50 cells in the streamwise direction and 30 cells in the normal-to-wall direction as previous works [8,43]. The periodic boundary condition is imposed on the inlet and outlet, and the bottom and top boundaries apply the no-slip wall condition. The $k-\varepsilon$ model is used as the baseline, and the time-averaged results from the direct numerical simulation (DNS) are used as the truth. The DNS data are obtained from Ref. [44], and the DNS mesh has $n_x \times n_y = 768 \times 385$ cells per plane. The velocity along the profiles at $x_1/H = 0, 1, \ldots, 8$ is used as observation data. In this case, we prescribe as prior knowledge that the inferred field is smooth and also does not deviate significantly from the prior value obtained with the $k-\varepsilon$ model. Hence, we employ the regularization of the smoothness and the prior value simultaneously. Note that penalizing the deviation from the prior value can also enforce the smoothness because the prior is smooth in this case.

The comparison of the inference results and propagated velocity without and with regularization is shown in Fig. 1. The eddy viscosity of DNS is computed by projecting the deviatoric part of DNS Reynolds stress $\tau$ onto the mean strain rate $\boldsymbol{S}$ [30], i.e.,

$$\nu_t^{\mathrm{DNS}} = \frac{\mathrm{dev}(\boldsymbol{\tau}) : \boldsymbol{S}}{\boldsymbol{S} : \boldsymbol{S}} \tag{8}$$

The obtained eddy viscosity $\nu_t^{\mathrm{DNS}}$ is an optimal solution in the *a priori* sense (i.e., achieving the least error norm $\|\mathrm{dev}(\boldsymbol{\tau}) - \nu_t \boldsymbol{S}\|$ in reconstructed Reynolds stresses). Because of the misalignment between the Reynolds stress and the strain rate, there exist negative values of eddy viscosity, and hence they are clipped to zero. The peak of the eddy viscosity in some regions such as the downhill is caused by the near-zero strain rate (i.e., singularity) in these areas. Both ensemble-based methods without and with regularization can provide similar velocity fields in good agreements with the DNS data. The reconstructed velocity with the regularized method is not as close to DNS as the unregularized case. This is likely because the regularized method has an objective function that aims to achieve the compromise between the data misfit and regularization. As a result, the regularization term is enforced at the expense of data fitting accuracy. On the other hand, the inferred eddy viscosity from the standard ensemble Kalman method without regularization leads to a very rough field, which has a large departure from the prior value in most areas. Particularly, at $x_1/H = 0$, the magnitude of the inferred eddy viscosity without regularization is many times larger than that provided by the $k-\varepsilon$ model. In contrast, the regularized ensemble Kalman method can infer a relatively smooth field of eddy viscosity by enforcing the regularization on the smoothness and prior values. The inferred eddy viscosity field is significantly improved in the magnitude and the smoothness compared with the inference results without regularization. Admittedly, the eddy viscosity field is still different from the DNS in this case, but noticeable improvement is achieved in the magnitude and smoothness compared with the unregularized method. It is emphasized that the main aim of the present work is to demonstrate the superiority of the regularized method over the unregularized method in inferring the turbulence quantities such as the eddy viscosity. To further improve the inference, one needs to incorporate additional observation data, explore better representation of inferred fields [30], and leverage additional physical constraints, e.g., from the boundary condition [45] and the characteristics of the governing equations [46]. Additionally, we emphasize that the eddy viscosity is not a physical concept in a strict sense, particularly for regions where Boussinesq hypothesis is invalid, i.e., where the Reynolds stress and strain rate are not aligned. For this reason, the comparison in the eddy viscosity between the inferred field and the DNS should be interpreted with caution. Here the inferred field can be considered as the optimal eddy viscosity in the *a posteriori* sense (i.e., it allows us to achieve the best agreement

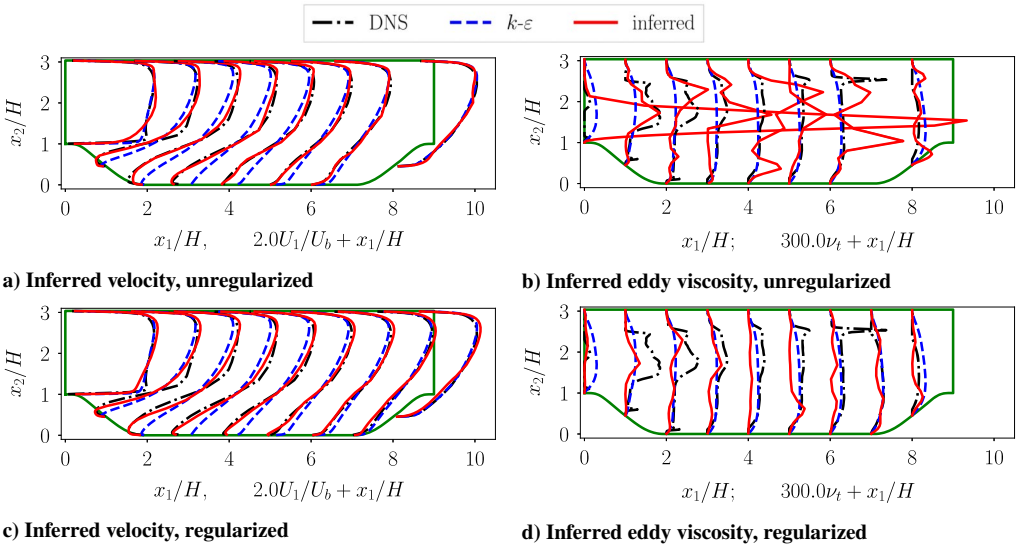



**Fig. 1 The inferred eddy viscosity and the propagated streamwise velocity field by use of the ensemble methods without and with regularization for periodic hill case.**

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

with the velocity data), which is fundamentally different from the DNS eddy viscosity computed by tensor projection in Eq. (8).

The contour plots of reconstructed velocity are presented in Fig. 2 to show the structure of the velocity field in comparison with the $k-\varepsilon$ model and the DNS results. The streamlines show that the baseline with $k-\varepsilon$ underestimates separation bubble sizes significantly compared with DNS data. The reconstruction results without and with regularization both provide better estimates in the separation bubble size compared with the baseline. The structure of the reconstructed separation bubble is noticeably improved compared with the prediction by $k-\varepsilon$ model but still deviates from the DNS. That is consistent with the slight deviation in the velocity field as shown in Fig. 1.

The contour plots of the eddy viscosity are provided in Fig. 3 to show the structure of the inferred fields. It is clear that the regularized EnKF method can significantly reduce the magnitude of the inferred eddy viscosity and smooth the entire field compared with the unregularized method. In this case, the inferred eddy viscosity has no apparent patterns. In the separation region, the local eddy viscosity is around zero, which is smaller than the $k-\varepsilon$ modeled values. The inferred eddy viscosity field embeds underlying information for the turbulence modelers to construct an improved predictive closure model for the RANS methods. The regression technique such as sparse regression and backpropagation is able to be used for training the algebraic [47] or neural-network-based [48,49] models based on the inferred fields.

### B. Transitional Flows over a Flat Plate

In the second case, we demonstrate inference of the laminar–turbulent intermittency in the equation governing turbulence kinetic energy. The test case is the turbulent flow over a flat plate, which is often used for investigations of the by-pass transition problem [50]. The inflow turbulence intensity is 0.033. The inlet bulk velocity $U_b$ is 5.4 m/s. The kinetic viscosity $\nu$ is $1.5\times10^{-5}$ m$^2$/s. The turbulent viscosity is set to $\nu_t = 12\nu$. The computational mesh has $10^4$ cells. The inlet is imposed with a uniform velocity $U_b$, and the outlet has a zero-gradient condition for velocity. The top boundary is set as the freestream, and the plate is a solid wall with the no-slip condition.

The intermittency $\gamma$ field has the value of one except in the laminar and transition region. Inferring such a field is highly ill-posed, because the intermittency away from the wall does not affect the flow quantities near the wall where the observation, e.g., friction coefficient, is often available. To this end, a sparse representation is needed for this field. Based on this prior knowledge, we use the Chebyshev basis and the sigmoid function to achieve this goal. First, the Chebyshev basis is used to represent the interface between the laminar region ($\gamma = 0$) and turbulent region ($\gamma = 1$), which is a line $y(x_1) = \sum_{n=1}^{N} w_n \hat{T}_n(x_1)$ (truncated to $N = 16$ modes). Across the interface, a sigmoid function is used to represent variation of $\gamma$ from 0 to 1 in the wall-normal direction. Further, another sigmoid function $f_{\text{tr}}(x_1)$ is used to allow the transition point $x_{\text{tr}}$ in the streamwise direction. The length scales of the streamwise and wall-normal transitions are denoted as $\sigma_1$ and $\sigma_2$, respectively. Other basis functions, such as the wavelet basis and Fourier basis, also can be used to represent the interface. Specifically, the wavelet basis is often used for the image with sharp edges, whereas the Fourier basis is widely used for periodic signals. As for the intermittency field, the turbulent/nonturbulent interface is smooth and also not subjected to the periodic conditions. To this end, we choose the Chebyshev basis to construct the interface. The sigmoid function is one of the most widely used activation functions, and hence we apply it to represent the transition from 0 to 1. It is noted that other activation functions, such as the hyperbolic tangent function, can be also used for this purpose. The representation is usually application specific and defined in this case based on the prior knowledge. That is, the intermittency field is demarcated by a turbulent/nonturbulent interface, and a ramp-up region exists around the interface. In a scenario where such prior knowledge about the inferred field is lacking, how to represent the field in an inference problem is a challenge and remains an open question. However, note that this work specifically targets scenarios where such vague prior knowledge is available but difficult to represent with traditional methods. The field representation as illustrated in Fig. 4 and the details are provided in Appendix C. With this scheme, the intermittency field $\gamma(x_1, x_2)$ is parameterized with parameter set $\{w_n\}_{n=1}^{16}$, $x_{\text{tr}}$, $\sigma_1$, and $\sigma_2$. The prior parameters are given as $w_1 = 0.5$, $w_n = 0$ for $n = 2, \ldots, 16$, $x_{\text{tr}} = 0.9$, and $\sigma_1 = \sigma_2 = 0.9$.

The experimental measurements [51] on the friction coefficients are used as the observation data. We use 16 measurement points that are evenly distributed along the plate to infer the intermittency field. Based on our numerical tests, the inversion results are sensitive to the measurements at the transition region where the friction coefficient exhibits noticeable differences between the data and the prediction from the baseline model. The $k-\omega$ SST model with $\gamma - Re_\theta$ transition model [40] is used as a reference to evaluate our inference, because this model has been validated to simulate the by-pass transitional flow accurately. In this case, we enforce the sparsity constraint on the mode coefficients such that fewer modes are used to avoid complex field structures. Additionally, we penalize the large departure from the initially guessed location of the transition point as well as the length scales $\sigma_1$ and $\sigma_2$. Therefore, in this case the sparsity and prior are enforced simultaneously to improve inference results.

We present the results with the unregularized and regularized ensemble method in Fig. 5. The ensemble methods with and without regularization can both reconstruct the friction coefficient and

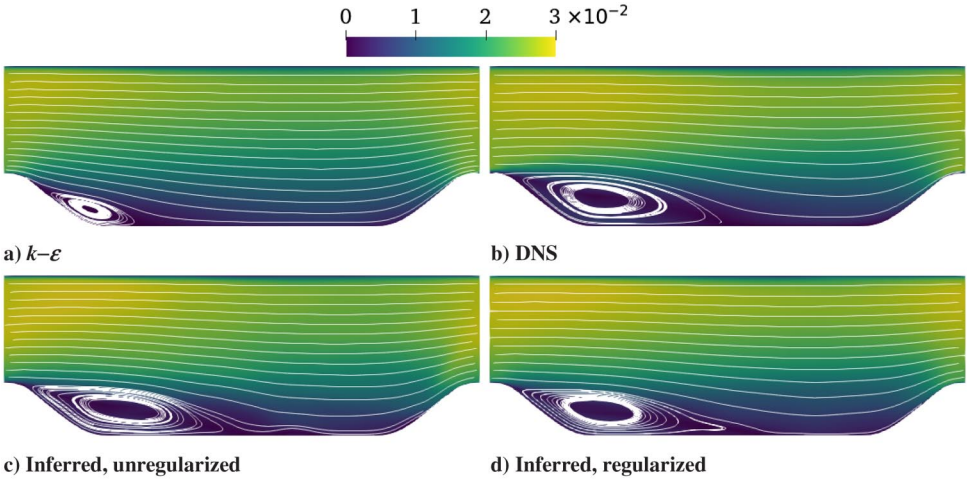



**Fig. 2 Contour plots of reconstructed velocity with unregularized and regularized ensemble methods in comparison to the prediction with $k-\varepsilon$ model and the DNS for the periodic hill case.**

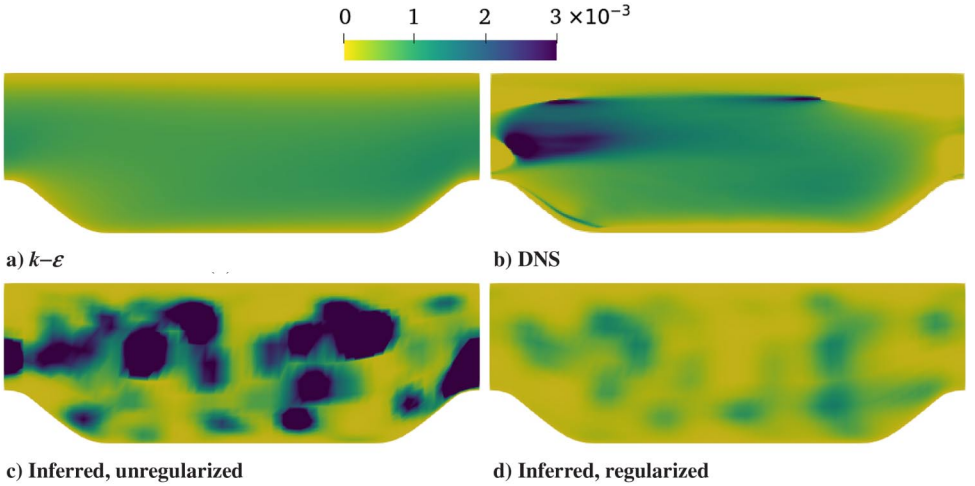



**Fig. 3 Contour plots of inferred eddy viscosity with unregularized and regularized ensemble methods in comparison to eddy viscosity from $k-\varepsilon$ model for periodic hill case.**

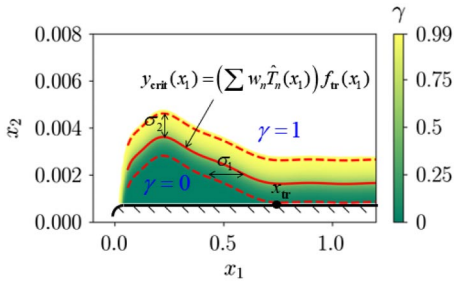



**Fig. 4 Parameterization of the intermittency field.**

velocity in good agreements with the reference. The similar reconstruction between the regularized and unregularized methods is not surprising, because the good agreement in the velocity and friction coefficient is enforced by the observation data. Concretely, the initial intermittency field leads to the large discrepancy in the friction coefficient. By incorporating the experimental data of friction coefficients, the reconstructed results are improved by comparison with the prediction from the transition model as shown in Figs. 5a and 5c. Moreover, around $x_1 = 0.6$ the inference leads to better estimation on the friction coefficient compared with the transition model in this case. Figures 5b and 5d show a good agreement between the inferred velocity and that from the transition model. At $x_1 = 0.6$, the gradient of the velocity is slightly less than that from the transition model, leading to the relatively small friction coefficient. Also, we present the reconstructed TKE and dissipation rate with the unregularized and regularized methods in Fig. 6. Here the dissipation rate is computed as $\epsilon = \beta^* k\omega$. It can be seen that the regularized method is able to provide better estimation in the TKE and dissipation rate, which is closer to the prediction of the Langtry–Menter model compared with the unregularized method.

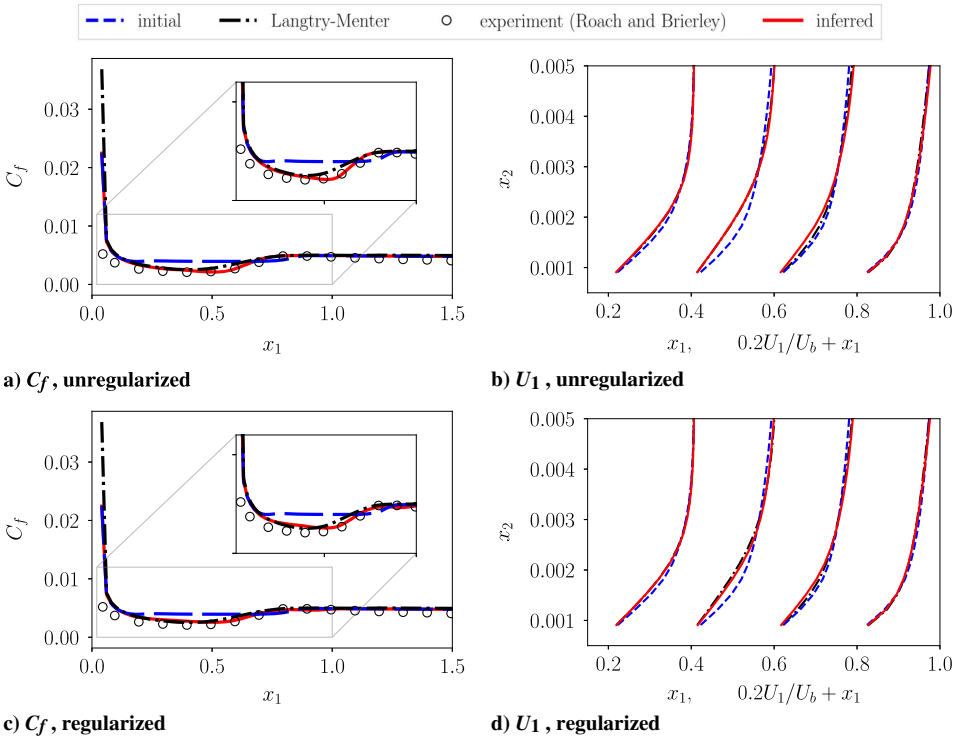
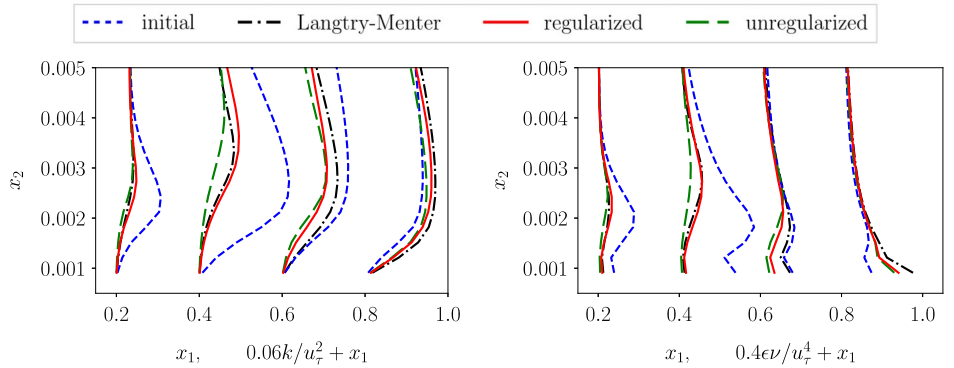


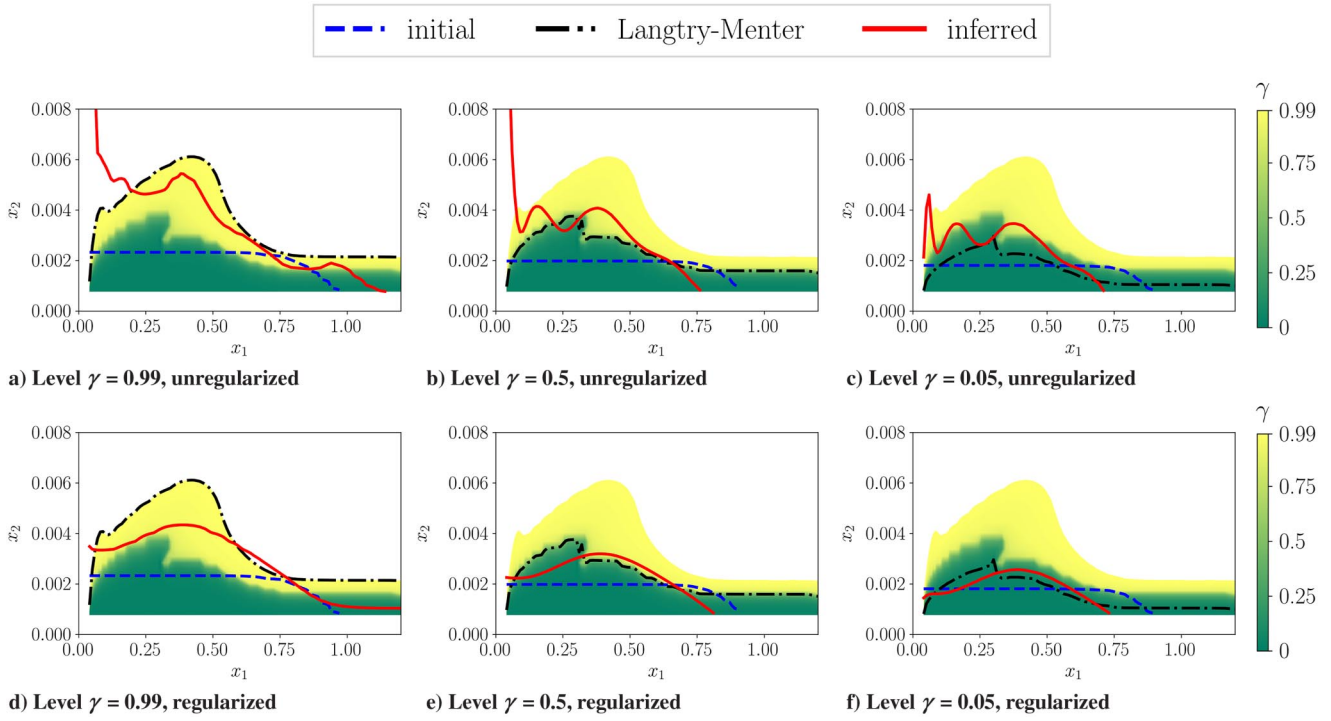

**Fig. 5 Reconstructed results of the friction coefficient and velocity with the regularized ensemble method in comparison among the initial prediction, the Langtry–Menter model [40], and the experimental data of Roach and Brierley [51] for the T3A case.**

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976




**a) Turbulent kinetic energy, $k$** **b) Dissipation rate, $\varepsilon$**

**Fig. 6 Reconstructed results of turbulent kinetic energy and dissipation rate with the unregularized and regularized ensemble method in comparison with the initial prediction and the Langtry–Menter model [40] for the T3A case.**




**a) Level $\gamma = 0.99$, unregularized** **b) Level $\gamma = 0.5$, unregularized** **c) Level $\gamma = 0.05$, unregularized**

**d) Level $\gamma = 0.99$, regularized** **e) Level $\gamma = 0.5$, regularized** **f) Level $\gamma = 0.05$, regularized**

**Fig. 7 Inferred intermittency field $\gamma$, showing a comparison of inferred $\gamma$ field with the initial estimation and the Langtry–Menter model [35]. The background color contour is the $\gamma$ field from the Langtry–Menter model.**

The inferred intermittency fields with and without regularization are shown in Fig. 7. It can be seen that the intermittency field without regularization has strong gradients, particularly at the range from $x_1 = 0$ to 0.25. In contrast, the inferred intermittency field with the regularization is much smoother than that without the regularization, because the sparsity regularization drives the coefficients of high-order modes to be zero. However, the results in the eddy viscosity without the regularization can also be smooth, as presented in Fig. 8. That is because the advection–diffusion function about the TKE is still able to provide the relatively smooth eddy viscosity [6]. Also, the high intermittency value at the front of the plate has no significant effects on the eddy viscosity fields. That is because the production of TKE diminishes in the region near the leading edge of the plate, whereas, on the other hand, the intermittency influences the mean field only through scaling the production term, i.e., $\gamma\mathcal{P}$ in the TKE transport equation (7). Consequently, the intermittency cannot be inferred from the friction coefficient measurements (or any other mean field quantities). In other words, even if the inferred intermittency can admit arbitrary values in this region (e.g., with very large magnitude and/or steep gradients), the mean field does not feel such anomaly due to the small production term.

## IV. Conclusions

This work demonstrates the application of an ensemble-based inverse modeling method for the field inversion of turbulence quantities by incorporating prior knowledge. The formulations are derived to implement general prior knowledge such as smoothness, prior value, and sparsity in the ensemble methods. Two different turbulence quantities, including the eddy viscosity and the intermittency field in the equation governing the TKE, are inferred from the observation of velocity and friction coefficient. The separated flow over periodic hills and transitional flow over a flat plate are used to demonstrate the merits of the ensemble-based inference framework. The results highlight the necessity of incorporating the prior knowledge into the inference of turbulence quantities and demonstrate the capability of the ensemble-based inference framework in implementing general regularization.

This work mainly demonstrates the merits of the regularized ensemble Kalman method in 2D scenarios. To further demonstrate the practicality of the method, future works will be conducted to apply the present method in 3D flow configurations. Moreover, the proposed method can be extended to infer uncertain terms in other

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

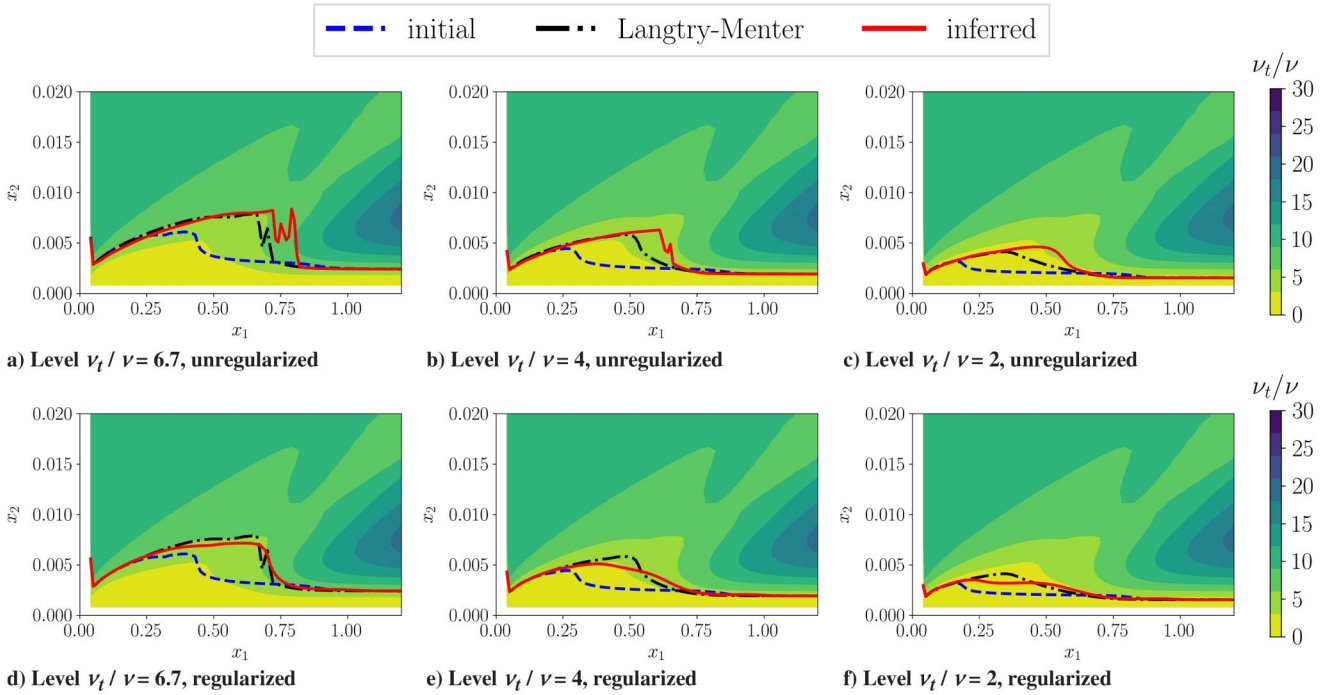



**Fig. 8 Inferred eddy viscosity field $\nu_t$ (normalized by molecular viscosity $\nu$), showing a comparison of inferred $\nu_t$ field with the initial estimation and the Langtry–Menter model [35]. The background color contour is the $\nu_t/\nu$ field from Langtry–Menter model.**

RANS models, such as the pressure–mean-strain correlation term in the Reynolds stress model, and other applications such as the flow state estimation problem. Finally, it is worthwhile to further investigate enforcing different physical constraints with this method, such as symmetry, physical realizability, and divergence-free condition for incompressible flows, which can further improve the inference.

## Appendix A: Imposing Smoothness Constraints for States Represented with KL Expansion

In some scenarios, we apply the KL expansion and take the KL coefficients $w$ as the inferred parameters to reduce the dimension of field inversion problems. That is, the inferred quantity, e.g., the eddy viscosity $\nu_t$, is written as

$$\log \nu_t = \log \nu_t^0 + \sum_{n=1}^{N} w_n \phi_n$$

where $\nu_t^0$ is the baseline of eddy viscosity from the RANS model. In such cases, the total variation method cannot provide the derivative of the penalty term to the inferred KL coefficients $w$. However, the functional derivative can be formulated to enforce smoothness constraints for the ensemble method. The cost function can be written as

$$J = \|w_j^a - w_j^f\|_{\mathsf{P}}^2 + \|\mathsf{H}w_j^a - \mathsf{y}_j\|_{\mathsf{R}}^2 + \lambda \sum_{i=1}^{3} \left\| \frac{\partial \nu_t}{\partial x_i} \right\|_{\mathsf{W}}^2 \tag{A1}$$

where $j$ is index of samples and $i$ is the index of spatial dimension. Note that

$$\log \nu_t = \log \nu_t^0 + \sum_{n=1}^{N} w_n \phi_n[x] = \log \nu_t^0 + \boldsymbol{\Phi w} \tag{A2a}$$

$$\log\left[\frac{\nu_t}{\nu_t^0}\right] = \boldsymbol{\Phi w} \tag{A2b}$$

$$\nu_t = \nu_t^0 \exp[\boldsymbol{\Phi w}] \tag{A2c}$$

where $n$ is the index of mode, $\boldsymbol{\Phi} = [\phi_1(x), \phi_2(x), \ldots, \phi_N(x)]$, $\phi(x) \in \mathbb{R}^{P\times 1}$, $\boldsymbol{\Phi} \in \mathbb{R}^{P\times N}$, $\boldsymbol{w} \in \mathbb{R}^{N\times 1}$, and $P$ is the dimension of physical space. For each direction $i$, at the specific location $x^l$ the smoothness penalty can be written as

$$\begin{aligned}\frac{\partial \nu_t(x^l)}{\partial x^l} &= \frac{\partial \nu_t^0 \exp\left[\sum_{n=1}^{N} w_n \phi_n\right]}{\partial x^l} \\ &= \frac{\partial \nu_t^0}{\partial x^l} \exp\left[\sum_{n=1}^{N} w_n \phi_n\right] + \nu_t^0 \exp\left[\sum_{n=1}^{N} w_n \phi_n\right] \sum_{n=1}^{N} w_n \frac{\partial \phi_n}{\partial x^l} \\ &= \frac{\nu_t}{\nu_t^0}\frac{\partial \nu_t^0}{\partial x^l} + \nu_t \frac{\partial}{\partial x^l}(\log[\nu_t/\nu_t^0])\end{aligned} \tag{A3}$$

The penalty function can be written as

$$\mathcal{G}_i[w_j] = \left[\frac{\partial \nu_t}{\partial x^1}, \cdots, \frac{\partial \nu_t}{\partial x^l}, \cdots, \frac{\partial \nu_t}{\partial x^P}\right] \tag{A4}$$

The gradient of the penalty function on the mode coefficient $w$ can be written as

$$\mathcal{G}_i'[w_j] = \begin{bmatrix} \frac{\partial^2 \nu_t}{\partial x^1 \partial w_1} & \cdots & \frac{\partial^2 \nu_t}{\partial x^1 \partial w_k} & \cdots & \frac{\partial^2 \nu_t}{\partial x^1 \partial w_N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \nu_t}{\partial x^l \partial w_1} & \cdots & \frac{\partial^2 \nu_t}{\partial x^l \partial w_k} & \cdots & \frac{\partial^2 \nu_t}{\partial x^l \partial w_N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \nu_t}{\partial x^P \partial w_1} & \cdots & \frac{\partial^2 \nu_t}{\partial x^P \partial w_k} & \cdots & \frac{\partial^2 \nu_t}{\partial x^P \partial w_N} \end{bmatrix} \tag{A5}$$

where

$$\frac{\partial^2 \nu_t(x^l)}{\partial x^l \partial w_k} = \frac{\partial}{\partial w_k}\left(\frac{\partial \nu_t}{\partial x^l}\right) \tag{A6}$$

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

$$= \frac{\partial}{\partial w_k}\left(\frac{\partial \nu_t^0}{\partial x^l}\exp\left[\sum_{n=1}^{N} w_n\phi_n\right] + \nu_t^0 \exp\left[\sum_{n=1}^{N} w_n\phi_n\right]\sum_{n=1}^{N} w_n \frac{\partial \phi_n}{\partial x^l}\right) \tag{A7}$$

$$= \frac{\partial \nu_t^0}{\partial x^l}\exp\left[\sum_{n=1}^{N} w_n\phi_n\right]\phi_k + \nu_t^0 \exp\left[\sum_{n=1}^{N} w_n\phi_n\right]\phi_k \sum_{n=1}^{N} w_n \frac{\partial \phi_n}{\partial x^l} + \nu_t^0 \exp\left[\sum_{n=1}^{N} w_n\phi_n\right]\frac{\partial \phi_k}{\partial x^l} \tag{A8}$$

$$= \frac{\partial \nu_t^0}{\partial x^l}\frac{\nu_t}{\nu_t^0}\phi_k + \nu_t\phi_k \frac{\partial \log\left[\frac{\nu_t}{\nu_t^0}\right]}{\partial x^l} + \nu_t \frac{\partial \phi_k}{\partial x^l} \tag{A9}$$

Equation (A9) is the gradient of the regularization term for smoothness when using the KL coefficient as inferred parameters.

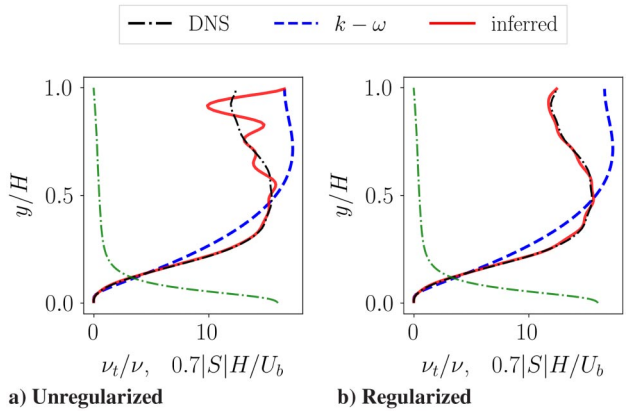



**Fig. 9 Inferred eddy viscosity by using the ensemble method with and without regularization for the channel case. The green/light dash-dotted line denotes the norm of mean rate-of-strain $S = \partial U_1/\partial x_2$.**

## Appendix B: Validation of Ensemble-Based Field Inversion in Channel Flow

We validate the capability of the ensemble method to enforce the smoothness properties in inferring eddy viscosity in the channel case. The Reynolds number based on friction velocity and half channel height is 180. The inferred parameter $x$ is the eddy viscosity field $\nu_t$, which can be formulated as $\nu_t = \{\nu_t^{(1)}, \ldots, \nu_t^{(l)}, \ldots, \nu_t^{(90)}\}$. Accordingly, the forward difference operator can be constructed as

$$\partial = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \\ & & & 0 \end{bmatrix} \tag{B1}$$

The ensemble-based field inversion is used to infer the eddy viscosity from the DNS data of the time-averaged velocity. The DNS data are from Ref. [52], which are widely used for numerical validation and analysis. The initial field $\boldsymbol{x}^0$ is from the $k-\omega$ model. The inference results without and with smoothness regularization are shown in Fig. B1. The propagated velocity without and with regularization both have good agreements with DNS, and hence the plots are omitted. It is clear that the unregularized method infer the eddy viscosity with strong gradients in the center region of the channel, where the velocity is insensitive to the eddy viscosity due to the small value on the norm of mean rate-of-strain $|S|$, as it is the shear stress $\tau_{12} = \nu_t(\partial U_1/\partial x_2)$ that appears in the RANS momentum equation. As such, enforcing the smoothness regularization provides better inference in the eddy viscosity.

## Appendix C: Parameterization of the Intermittency Field

The intermittency field is parameterized by using a sigmod function as follows:

$$\gamma(x_1, x_2) = S\left[\frac{x_2 - y_{\text{crit}}[x_1]}{\sigma_2}\right]$$

where $\sigma_2$ is the length scale of transition in the vertical direction, and the location of transition $y_{\text{crit}}(x_1)$ (i.e., the contour line $\gamma = 0.5$) is

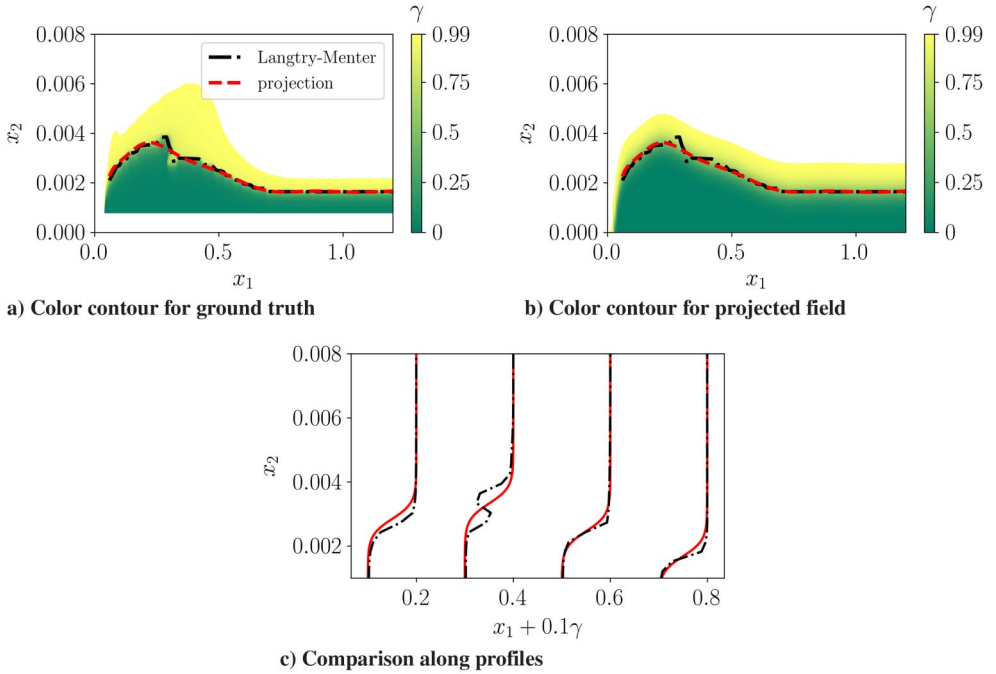



**Fig. 10 The comparison between the projection and the result from the Langtry–Menter transition model. The projected curve $y_{\text{crit}}$ with the Chebyshev basis based on the intermittency at level $\gamma = 0.5$ from Langtry–Menter transition model.**

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

parameterized by using a shifted Chebyshev basis for dimension reduction purposes:

$$y_{\text{crit}}(x_1) = \left(\sum_{n=1}^{N} w_n \hat{T}_n(x_1)\right) f_{\text{tr}}[x_1], \quad \text{with } f_{\text{tr}}[x_1] = S\left[-\frac{x_1 - x_{\text{tr}}}{\sigma_1}\right]$$

where $x_{\text{tr}}$ is the location of laminar–turbulent transition on the plate, and $\sigma_1$ is the length scale of transition. The standard Chebyshev polynomial of the first kind is an orthogonal basis on the domain $[-1, 1]$ defined based on the following recursive relation [53]:

$$T_0[x] = 1, \quad T_1[x] = x, \quad \ldots, \quad T_{n+1}[x] = 2xT_n[x] - T_{n-1}[x] \quad \text{(C1)}$$

For example, the first few basis functions are $T_2[x] = 2x^2 - 1$, $T_3[x] = 4x^3 - 3x$, $T_4[x] = 8x^4 - 8x^2 + 1$, and $T_5[x] = 16x^5 - 20x^3 + 5x$. When $x_1 \in [x_s, x_e]$, the following transformation should be performed to obtain polynomials $\hat{T}_n[x]$ that are orthogonal on the new domain:

$$\hat{T}_n[x] = T_n[a + bx]$$

where the argument to the basis function $T_n$ is $a + bx$ with coefficients $a = -1 - 2x_s/(x_e - x_s)$ and $b = 2/(x_e - x_s)$. Therefore, the intermittency field is parameterized by the parameters $\{w_n\}_{n=1}^{N}$, the location of transition $x_{\text{tr}}$, and the length scales $\sigma_1$ and $\sigma_2$ over which $\gamma$ varies from 0 to 1. We project the contour line of intermittency at level $\gamma = 0.5$ from the Langtry–Menter transition model on the Chebyshev basis, and the projected curve represents that of the reference data very well, as is shown in Fig. C1a. Further, the best fit of the $\gamma(x_1, x_2)$ on the parameterization above is shown in Fig. C1b. These results suggest that the proposed parameterization is suitable for the representation of underlying intermittency field. The plots along the profiles are provided in Fig. C1, clearly showing that the projected profiles have a good alignment with the results from Langtry–Menter model.

## Acknowledgments

Xin-Lei Zhang and Guowei He are supported by the National Natural Science Foundation of China (NSFC) Basic Science Center Program for "Multiscale Problems in Nonlinear Mechanics" (No. 11988102), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB22040104), and the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (QYZDJ-SSWSYS002). Xin-Lei Zhang acknowledges support from the China Postdoctoral Science Foundation (No. 2021M690154). Heng Xiao did not receive external funding for this research.

## References

[1] Dow, E., and Wang, Q., "Quantification of Structural Uncertainties in the $k - \omega$ Turbulence Model," AIAA Paper 2011-1762, 2011. https://doi.org/10.2514/6.2011-1762

[2] Wang, J.-X., Wu, J.-L., and Xiao, H., "Incorporating Prior Knowledge for Quantifying and Reducing Model-Form Uncertainty in RANS Simulations," *International Journal for Uncertainty Quantification*, Vol. 6, No. 2, 2016, pp. 109–126. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2016015984

[3] Giles, M. B., Duta, M. C., Muller, J.-D., and Pierce, N. A., "Algorithm Developments for Discrete Adjoint Methods," *AIAA Journal*, Vol. 41, No. 2, 2003, pp. 198–205. https://doi.org/10.2514/2.1961

[4] Evensen, G., "Analysis of Iterative Ensemble Smoothers for Solving Inverse Problems," *Computational Geosciences*, Vol. 22, March 2018, pp. 885–908. https://doi.org/10.1007/s10596-018-9731-y

[5] Singh, A. P., and Duraisamy, K., "Using Field Inversion to Quantify Functional Errors in Turbulence Closures," *Physics of Fluids*, Vol. 28, No. 4, 2016, Paper 045110. https://doi.org/10.1063/1.4947045

[6] Franceschini, L., Sipp, D., and Marquet, O., "Mean-Flow Data Assimilation Based on Minimal Correction of Turbulence Models: Application to Turbulent High Reynolds Number Backward-Facing Step," *Physical Review Fluids*, Vol. 5, No. 9, 2020, Paper 094603. https://doi.org/10.1103/PhysRevFluids.5.094603

[7] Belligoli, Z., Dwight, R. P., and Eitelberg, G., "Reconstruction of Turbulent Flows at High Reynolds Numbers Using Data Assimilation Techniques," *AIAA Journal*, Vol. 59, No. 3, 2021, pp. 855–867. https://doi.org/10.2514/1.J059474

[8] Xiao, H., Wu, J.-L., Wang, J.-X., Sun, R., and Roy, C., "Quantifying and Reducing Model-Form Uncertainties in Reynolds-Averaged Navier–Stokes Simulations: A Data-Driven, Physics-Informed Bayesian Approach," *Journal of Computational Physics*, Vol. 324, Nov. 2016, pp. 115–136. https://doi.org/10.1016/j.jcp.2016.07.038

[9] Mons, V., Chassaing, J.-C., Gomez, T., and Sagaut, P., "Reconstruction of Unsteady Viscous Flows Using Data Assimilation Schemes," *Journal of Computational Physics*, Vol. 316, July 2016, pp. 255–280. https://doi.org/10.1016/j.jcp.2016.04.022

[10] Zhang, X., Gomez, T., and Coutier-Delgosha, O., "Bayesian Optimisation of RANS Simulation with Ensemble-Based Variational Method in Convergent-Divergent Channel," *Journal of Turbulence*, Vol. 20, No. 3, 2019, pp. 214–239. https://doi.org/10.1080/14685248.2019.1622016

[11] Yang, M., and Xiao, Z., "Improving the k—Ar Transition Model by the Field Inversion and Machine Learning Framework," *Physics of Fluids*, Vol. 32, No. 6, 2020, Paper 064101. https://doi.org/10.1063/5.0011655

[12] Darakananda, D., da Silva, A. F. C., Colonius, T., and Eldredge, J. D., "Data-Assimilated Low-Order Vortex Modeling of Separated Flows," *Physical Review Fluids*, Vol. 3, No. 12, 2018, Paper 124701. https://doi.org/10.1103/PhysRevFluids.3.124701

[13] da Silva, A. F., and Colonius, T., "Ensemble-Based State Estimator for Aerodynamic Flows," *AIAA Journal*, Vol. 56, No. 7, 2018, pp. 2568–2578. https://doi.org/10.2514/1.J056743

[14] Le Provost, M., and Eldredge, J. D., "Ensemble Kalman Filter for Vortex Models of Disturbed Aerodynamic Flows," *Physical Review Fluids*, Vol. 6, No. 5, 2021, Paper 050506. https://doi.org/10.1103/PhysRevFluids.6.050506

[15] Pawar, S., and San, O., "Data Assimilation Empowered Neural Network Parametrizations for Subgrid Processes in Geophysical Flows," *Physical Review Fluids*, Vol. 6, No. 5, 2021, Paper 050501. https://doi.org/10.1103/PhysRevFluids.6.050501

[16] Iglesias, M. A., Law, K. J., and Stuart, A. M., "Ensemble Kalman Methods for Inverse Problems," *Inverse Problems*, Vol. 29, No. 4, 2013, Paper 045001. https://doi.org/10.1088/0266-5611/29/4/045001

[17] Zhang, X.-L., Xiao, H., He, G., and Wang, S., "Assimilation of Disparate Data for Enhanced Reconstruction of Turbulent Mean Flows," *Computers & Fluids*, Vol. 224, June 2021, Paper 104962. https://doi.org/10.1016/j.compfluid.2021.104962

[18] Zhang, X.-L., Michelén-Ströfer, C., and Xiao, H., "Regularized Ensemble Kalman Methods for Inverse Problems," *Journal of Computational Physics*, Vol. 416, Sept. 2020, Paper 109517. https://doi.org/10.1016/j.jcp.2020.109517

[19] Evensen, G., "Advanced Data Assimilation for Strongly Nonlinear Dynamics," *Monthly Weather Review*, Vol. 125, No. 6, 1997, pp. 1342–1354. https://doi.org/10.1175/1520-0493(1997)125<1342:ADAFSN>2.0.CO;2

[20] Evensen, G., and Van Leeuwen, P. J., "An Ensemble Kalman Smoother for Nonlinear Dynamics," *Monthly Weather Review*, Vol. 128, No. 6, 2000, pp. 1852–1867. https://doi.org/10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2

[21] Scales, J. A., Docherty, P., and Gersztenkorn, A., "Regularisation of Nonlinear Inverse Problems: Imaging the Near-Surface Weathering Layer," *Inverse Problems*, Vol. 6, No. 1, 1990, p. 115. https://doi.org/10.1088/0266-5611/6/1/011

[22] Hastie, T., Tibshirani, R., and Wainwright, M., *Statistical Learning with Sparsity: The Lasso and Generalizations*, Chapman and Hall/CRC, New York, 2019, Chap. 4.

[23] Bishop, C. M., *Pattern Recognition and Machine Learning*, Springer, New York, 2006, pp. 144–145.

[24] Xiao, H., and Cinnella, P., "Quantification of Model Uncertainty in RANS Simulations: A Review," *Progress in Aerospace Sciences*, Vol. 108, July 2019, pp. 1–31. https://doi.org/10.1016/j.paerosci.2018.10.001

[25] Launder, B. E., and Sharma, B., "Application of the Energy-Dissipation Model of Turbulence to the Calculation of Flow near a Spinning Disc," *Letters in Heat and Mass Transfer*, Vol. 1, No. 2, 1974, pp. 131–137. https://doi.org/10.1016/0094-4548(74)90150-7

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976

[26] Craft, T., Launder, B., and Suga, K., "Prediction of Turbulent Transitional Phenomena with a Nonlinear Eddy-Viscosity Model," *International Journal of Heat and Fluid Flow*, Vol. 18, No. 1, 1997, pp. 15–28. https://doi.org/10.1016/S0142-727X(96)00145-2

[27] Wallin, S., and Johansson, A. V., "An Explicit Algebraic Reynolds Stress Model for Incompressible and Compressible Turbulent Flows," *Journal of Fluid Mechanics*, Vol. 403, Jan. 2000, pp. 89–132. https://doi.org/10.1017/S0022112099007004

[28] Launder, B. E., Reece, G. J., and Rodi, W., "Progress in the Development of a Reynolds-Stress Turbulence Closure," *Journal of Fluid Mechanics*, Vol. 68, No. 3, 1975, pp. 537–566. https://doi.org/10.1017/S0022112075001814

[29] Ling, J., Kurzawski, A., and Templeton, J., "Reynolds Averaged Turbulence Modelling Using Deep Neural Networks with Embedded Invariance," *Journal of Fluid Mechanics*, Vol. 807, Nov. 2016, pp. 155–166. https://doi.org/10.1017/jfm.2016.615

[30] Wu, J.-L., Sun, R., Laizet, S., and Xiao, H., "Representation of Stress Tensor Perturbations with Application in Machine-Learning-Assisted Turbulence Modeling," *Computer Methods in Applied Mechanics and Engineering*, Vol. 346, April 2019, pp. 707–726. https://doi.org/10.1016/j.cma.2018.09.010

[31] Wang, J.-X., Huang, J., Duan, L., and Xiao, H., "Prediction of Reynolds Stresses in High-Mach-Number Turbulent Boundary Layers Using Physics-Informed Machine Learning," *Theoretical and Computational Fluid Dynamics*, Vol. 33, Feb. 2019, pp. 1–19. https://doi.org/10.1007/s00162-018-0480-2

[32] Ströfer Michelén, C. A., and Xiao, H., "End-to-End Differentiable Learning of Turbulence Models from Indirect Observations," *Theoretical and Applied Mechanics Letters*, Vol. 11, No. 4, 2021, Paper 100280. https://doi.org/10.1016/j.taml.2021.100280

[33] Spalart, P. R., and Allmaras, S. R., "A One-Equation Turbulence Model for Aerodynamic Flows," AIAA Paper 1992-439, 1992. https://doi.org/10.2514/6.1992-439

[34] Wilcox, D. C., *Turbulence Modeling for CFD*, 3rd ed., DCW Industries, La Cañada, CA, 2006, Chap. 4.

[35] Menter, F. R., "Two-Equation Eddy-Viscosity Turbulence Models for Engineering Applications," *AIAA Journal*, Vol. 32, No. 8, 1994, pp. 1598–1605. https://doi.org/10.2514/3.12149

[36] Pope, S. B., *Turbulent Flows*, Cambridge Univ. Press, Cambridge, England, U.K., 2000, pp. 366–450.

[37] Foures, D. P., Dovetta, N., Sipp, D., and Schmid, P. J., "A Data-Assimilation Method for Reynolds-Averaged Navier-Stokes-Driven Mean Flow Reconstruction," *Journal of Fluid Mechanics*, Vol. 759, Nov. 2014, p. 404. https://doi.org/10.1017/jfm.2014.566

[38] Emory, M., Larsson, J., and Iaccarino, G., "Modeling of Structural Uncertainties in Reynolds-Averaged Navier-Stokes Closures," *Physics of Fluids*, Vol. 25, No. 11, 2013, Paper 110822. https://doi.org/10.1063/1.4824659

[39] Hao, Z., and Gorlé, C., "Quantifying Turbulence Model Uncertainty in Reynolds-Averaged Navier–Stokes Simulations of a Pin-Fin Array. Part 1: Flow Field," *Computers & Fluids*, Vol. 209, Sept. 2020, Paper 104641. https://doi.org/10.1016/j.compfluid.2020.104641

[40] Langtry, R. B., and Menter, F. R., "Correlation-Based Transition Modeling for Unstructured Parallelized Computational Fluid Dynamics Codes," *AIAA Journal*, Vol. 47, No. 12, 2009, pp. 2894–2906. https://doi.org/10.2514/1.42362

[41] Michelén Ströfer, C. A., Zhang, X.-L., and Xiao, H., "DAFI: An Open-Source Framework for Ensemble-Based Data Assimilation and Field Inversion," *Communications in Computational Physics*, Vol. 29, No. 5, 2021, pp. 1583–1622. https://doi.org/10.4208/cicp.OA-2020-0178

[42] Fröhlich, J., Mellen, C. P., Rodi, W., Temmerman, L., and Leschziner, M. A., "Highly Resolved Large-Eddy Simulation of Separated Flow in a Channel with Streamwise Periodic Constrictions," *Journal of Fluid Mechanics*, Vol. 526, March 2005, pp. 19–66. https://doi.org/10.1017/S0022112004002812

[43] Zhang, X.-L., Xiao, H., Gomez, T., and Coutier-Delgosha, O., "Evaluation of Ensemble Methods for Quantifying Uncertainties in Steady-State CFD Applications with Small Ensemble Sizes," *Computers & Fluids*, Vol. 203, May 2020, Paper 104530. https://doi.org/10.1016/j.compfluid.2020.104530

[44] Xiao, H., Wu, J.-L., Laizet, S., and Duan, L., "Flows over Periodic Hills of Parameterized Geometries: A Dataset for Data-Driven Turbulence Modeling from Direct Simulations," *Computers & Fluids*, Vol. 200, March 2020, Paper 104431. https://doi.org/10.1016/j.compfluid.2020.104431

[45] Ströfer Michelén, C. A., Zhang, X.-L., Xiao, H., and Coutier-Delgosha, O., "Enforcing Boundary Conditions on Physical Fields in Bayesian Inversion," *Computer Methods in Applied Mechanics and Engineering*, Vol. 367, Aug. 2020, Paper 113097. https://doi.org/10.1016/j.cma.2020.113097

[46] Wu, J.-L., Michelén Ströfer, C. A., and Xiao, H., "Physics-Informed Covariance Kernel for Model-Form Uncertainty Quantification with Application to Turbulent Flows," *Computers & Fluids*, Vol. 193, Oct. 2019, Paper 104292. https://doi.org/10.1016/j.compfluid.2019.104292

[47] Schmelzer, M., Dwight, R. P., and Cinnella, P., "Discovery of Algebraic Reynolds-Stress Models Using Sparse Symbolic Regression," *Flow, Turbulence and Combustion*, Vol. 104, March 2020, pp. 579–603. https://doi.org/10.1007/s10494-019-00089-x

[48] Zhou, Z., He, G., Wang, S., and Jin, G., "Subgrid-Scale Model for Large-Eddy Simulation of Isotropic Turbulent Flows Using an Artificial Neural Network," *Computers & Fluids*, Vol. 195, Dec. 2019, Paper 104319. https://doi.org/10.1016/j.compfluid.2019.104319

[49] Zhou, Z., He, G., and Yang, X., "Wall Model Based on Neural Networks for LES of Turbulent Flows over Periodic Hills," *Physical Review Fluids*, Vol. 6, No. 5, 2021, Paper 054610. https://doi.org/10.1103/PhysRevFluids.6.054610

[50] Duraisamy, K., Zhang, Z. J., and Singh, A. P., "New Approaches in Turbulence and Transition Modeling Using Data-Driven Techniques," AIAA Paper 2015-1284, 2015. https://doi.org/10.2514/6.2015-1284

[51] Roach, P., and Brierley, D., "The Influence of a Turbulent Free-Stream on Zero Pressure Gradient Transitional Boundary Layer Development: Part 1. Test Cases T3A and T3B," *ERCOFTAC Workshop*, Cambridge Univ. Press, Cambridge, England, U.K., 1990, pp. 319–347.

[52] Kim, J., Moin, P., and Moser, R., "Turbulence Statistics in Fully Developed Channel Flow at Low Reynolds Number," *Journal of Fluid Mechanics*, Vol. 177, April 1987, pp. 133–166. https://doi.org/10.1017/S0022112087000892

[53] Boyd, J. P., *Chebyshev and Fourier Spectral Methods*, Dover, New York, 2001, p. 497.

K. Taira
*Associate Editor*

Downloaded by ZHEJIANG UNIVERSITY on January 13, 2022 | http://arc.aiaa.org | DOI: 10.2514/1.J060976